Jukka Taskinen

PROTEIN CRYSTALLOGRAPHIC STUDIES OF COA-DEPENDENT PROTEINS: NEW INSIGHT INTO THE BINDING MODE AND EXCHANGE MECHANISM OF ACYL-COA
Jukka Taskinen

Protein Crystallographic Studies of Coa-Dependent Proteins: New Insight into the Binding Mode and Exchange Mechanism of Acyl-Coa

Academic Dissertation to be presented with the assent of the Faculty of Science, University of Oulu, for public discussion in Raahensali (Auditorium L10), Linnanmaa, on April 28th, 2006, at 12 noon

Oulun Yliopisto, Oulu 2006
Taskinen, Jukka, Protein crystallographic studies of CoA-dependent proteins: new insight into the binding mode and exchange mechanism of acyl-CoA

Faculty of Science, Department of Biochemistry, University of Oulu, P.O.Box 3000, FI-90014 University of Oulu, Finland

Oulu, Finland

Abstract

Multifunctional enzyme type 1 (MFE-1) is a monomeric member of the hydratase/isomerase superfamily (H/I) involved in the β-oxidation of fatty acids. MFE-1 has 2-enoyl-CoA hydratase-1, Δ^3-Δ^2-enoyl-CoA isomerase, and several other enoyl-CoA isomerase activities at the N-terminus. The C-terminus has (3S)-hydroxyacyl-CoA dehydrogenase activity. MFE-1 can also convert certain hydroxylated C_{27} bile acid synthesis intermediates.

In these studies, a domain assignment of MFE-1 by sequence alignment with the H/I family (domains A and B in MFE-1) and mitochondrial monofunctional 3-hydroxyacyl-CoA dehydrogenases (HAD, domains C, D and E) was proposed. This was further improved with the structural information obtained from the crystal structure of the construct containing domains B, C, D and E (MFE1-DH). The structure of MFE1-DH resembles the bilobal structure of the α-subunit of the bacterial fatty acid metabolising complex and the mammalian HAD enzyme. The N-terminal linker helix of MFE1-DH (domain B) corresponds to helix-10 of the hydratase/isomerase enzymes having residues important for substrate contacts. Domain C adopts the classical Rossmann fold and forms the first lobe of the MFE1-DH structure. The C-terminal domains D and E form the second lobe and have local symmetry between each other. This local symmetry corresponds to the D domain-mediated dimerisation of the HAD dimer. The domain deletion studies showed that the presence of domains D and E, but not domain C, was essential to obtain a functional hydratase 1 enzyme; this can be understood from stabilising contacts from domain E to the linker helix, as seen in the MFE1-DH structure.

The structure of human ACBP from liver was determined with and without a physiological ligand. This structure adopts the classical four-helix bundle of the ACBP family. The ligand binding mode seen in the presence of myristoyl-CoA shows that one ligand molecule is bound jointly by the two protein molecules of the asymmetric unit such that the fatty acid tail is bound by one protein molecule, and the 3'-phosphate AMP moiety of the CoA is bound by the other protein molecule, essentially as in known complexed ACBP structures in the monomeric binding mode. The observed ligand binding mode suggests a new model for the ACBP-mediated ligand transfer observed in biochemical in vitro studies.

Keywords: β-oxidation, 3-hydroxyacyl-CoA dehydrogenase, acyl-CoA binding protein, multifunctional enzyme, X-ray crystallography
Taskinen, Jukka, Proteiinikristallografisia tutkimuksia CoA-riippuvaisista proteiineista: uusi näkemys asyyli-CoA:n sitoutumiseen ja välitysmekanismiin
Luonnontieteellinen tiedekunta, Biokemian laitos, Oulun yliopisto, PL 3000, 90014 Oulun yliopisto
Oulu

Tiivistelmä
Tyypin 1 monitoiminen entsyymi (MFE-1) on hydrataasi/isomeraasiperheen (H/I) jäsen ja se osallistuu rasvahappojen β-oksidaatioon. MFE-1:n N-päädyssä on 2-enoyyli-CoA-hydrataasi 1- ja Δ³,Δ⁴-enoyyli-CoA-isomeraasiksi aktiivisuus sekä useita muita enoyyli-CoA-isomeraasi-aktiivuuksia. C-päädyssä on β-hydroksiasyyli-CoA-dehydrogenaasiaktiivisuus. MFE-1 voi myös katalysoida tiettyjen hydroksyloituksen koossa 27-sappihappo-reaktiota

proteins are molecules made of atoms, not letters

Joël Janin
Acknowledgements

This work was carried out at the Department of Biochemistry, University of Oulu, during the years 1999–2006.

I wish to express my gratitude to my supervisors Professor Rik Wierenga and Professor Kalervo Hiltunen. Their knowledge and expertise in structural biology and fatty acid metabolism has guided this work and enabled its completion. They have also provided the necessary resources that have made it possible to carry out this study. Professor Jens Knudsen's generous donation of ACBP initiated a series of events that allowed me to learn most of what I know about protein crystallography today. I also wish to thank Dr. Daan van Aalten for giving me an opportunity to work and learn with the ACBP project. Professors and principal investigators at the Department of Biochemistry are acknowledged for providing the facilities and supportive atmosphere for research. I also wish to thank all colleagues and staff members at the department for help and discussions.

I am very much grateful to Professor Stefan Alexson and Professor Mark Johnson for their interest on this thesis. Their encouraging comments and critical notes helped me to further improve the manuscript. Dr. Peter Cura's efforts in revisioning the language are acknowledged. I also want to express my appreciation to all of my co-authors whose contribution to this work has been important.

The members of the RW and KH groups are acknowledged for their collective efforts in unravelling yet more wonders of the lipid metabolism, and also on helping with the daily struggle in the lab. Especially the presence of Aare, Antti, Ilkka, Juha, Kalle, Kristian, Seppo, Tiila and Tomi has made the days pass quickly. The adventures at the synchrotrons around Europe with many of you have been unforgettable.

Help from the system managers Seppo Kilpeläinen and Ari-Pekka Kvist deserve special thanks. More recent efforts of Jyrki Hänninen and Miki Kallio on the same field are also much appreciated.

Finally, I am deeply grateful to my love Kati for her presence and support. My parents Pekka and Mirja, sister Kirsti and her husband Peter deserve warm thanks for their support and encouragement that have been essential to get this work completed.

This study was partially financed by the Academy of Finland, Oulu University Research Foundation and Sigrid Jusélius Foundation.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACBP</td>
<td>acyl-CoA binding protein</td>
</tr>
<tr>
<td>ACS</td>
<td>acyl-CoA synthetase</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bACBP</td>
<td>bovine acyl-CoA binding protein</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>ECH</td>
<td>2-enoyl-CoA hydratase</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>FA</td>
<td>fatty acid</td>
</tr>
<tr>
<td>FABP</td>
<td>fatty acid binding protein</td>
</tr>
<tr>
<td>FOM</td>
<td>bacterial fatty acid metabolising complex</td>
</tr>
<tr>
<td>H/I</td>
<td>hydratase-1/isomerase superfamily</td>
</tr>
<tr>
<td>HAD</td>
<td>(3S)-hydroxyacyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>hACBP</td>
<td>human acyl-CoA binding protein</td>
</tr>
<tr>
<td>hmECI</td>
<td>human mitochondrial 2-enoyl-CoA isomerase</td>
</tr>
<tr>
<td>hmHAD</td>
<td>human mitochondrial (3S)-hydroxyacyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>MFE</td>
<td>multifunctional enzyme</td>
</tr>
<tr>
<td>MFE1-DH</td>
<td>(3S)-hydroxyacyl-CoA dehydrogenase part of the rat peroxisomal multifunctional enzyme type 1</td>
</tr>
<tr>
<td>MTP</td>
<td>mitochondrial trifunctional protein</td>
</tr>
<tr>
<td>NAD⁺</td>
<td>nicotinamide adenine dinucleotide (oxidised form)</td>
</tr>
<tr>
<td>PfACBP</td>
<td><em>Plasmodium falciparum</em> acyl-CoA binding protein</td>
</tr>
<tr>
<td>rmHYD</td>
<td>rat mitochondrial enoyl-CoA hydratase</td>
</tr>
<tr>
<td>SAD</td>
<td>single wavelength anomalous diffraction</td>
</tr>
<tr>
<td>SCP</td>
<td>sterol carrier protein</td>
</tr>
<tr>
<td>VLCFA</td>
<td>very-long-chain fatty acid</td>
</tr>
<tr>
<td>ypECI</td>
<td>yeast peroxisomal enoyl-CoA isomerase</td>
</tr>
</tbody>
</table>
List of original articles

The present thesis is based on the following articles, which are referred to in the text by Roman numerals:


## Contents

Abstract  
Tiivistelmä  
Acknowledgements  
Abbreviations  
List of original articles  
Contents  

1 Introduction ................................................................................................................... 17  
2 Review of the literature ................................................................................................. 19  
2.1 Physical and chemical properties of lipids ............................................................. 19  
2.1.1 Simple fatty acids ............................................................................................ 19  
2.1.2 Acyl-coenzyme A ............................................................................................ 21  
2.1.3 Acyl carrier protein ......................................................................................... 22  
2.2 Fatty acid activation and transport.......................................................................... 24  
2.2.1 Intracellular trafficking of fatty acyl ................................................................ 25  
2.2.1.1 Fatty acid binding proteins ....................................................................... 26  
2.2.1.2 Sterol carrier protein ................................................................................. 27  
2.2.1.3 Acyl-CoA binding proteins ....................................................................... 28  
2.2.2 Protein-acyl-CoA interactions ......................................................................... 30  
2.3 Fatty acid degradation ............................................................................................ 31  
2.3.1 Multifunctional β-oxidation enzymes in mammals ......................................... 32  
2.3.1.1 Multifunctional enzyme type 1 ................................................................. 32  
2.3.1.2 Mitochondrial trifunctional enzyme ........................................................... 34  
2.3.2 Substrate channelling in the bacterial fatty acid metabolising complex ........ 35  
3 Aims of the present study ............................................................................................. 38  
4 Materials and methods ................................................................................................. 39  
4.1 Bacterial strains and plasmids ................................................................................ 39  
4.1.1 Construction of MFE-1 domain variants ......................................................... 39  
4.2 Protein expression and purification ........................................................................ 40  
4.2.1 Production of MFE-1 domain variants ............................................................ 40  
4.2.2 Purification of the (3S)-hydroxyacyl-CoA dehydrogenase for crystallisation........ 40
1 Introduction

Fatty acids have fundamental roles in living organisms. They provide not only energy, but they are also building blocks that are used to make a variety of biological compounds that are required to support cellular processes. Fatty acids are substrates for the generation of membranes, and they have a role in many signalling pathways. They are also precursors of hormones and bile acids. The essential role of fatty acids in the cell is emphasised by the identification of several genetic diseases related to defects in proteins participating in fatty acid metabolism.

The free fatty acid concentration in the cell is kept low even though the availability of fatty acids to the cell varies greatly depending on the nutritional state. Dietary fats raise the chylomicron concentration rapidly and ultimately activate regulatory pathways resulting in adjustments in gene expression levels of genes involved in fatty acid synthesis and degradation. The fatty acid fluxes of these two apparently opposite pathways are kept separate, for example by using different cofactors (like coenzyme A, CoA; and acyl carrier protein, ACP) to prepare activated fatty acids for each metabolic pathway, and by performing the reactions in compartmentalised cellular locations. Fatty acid activated transcription factors in the nucleus play a fundamental role in regulating the energy balance by sensing changes in the fatty acid fluxes.

Because of the essential functional role of the diverse fatty acids and because of their physico-chemical characteristics, understanding the mechanisms that are used to ensure the uptake, transport and processing of these compounds is of great importance. Lipids are chiral, complex molecules that differ in their physical and chemical properties in aqueous solutions. Even though the major routes of lipid metabolism have been characterised in detail, recent studies have revealed many alternative routes for metabolic processes. Furthermore, the structural features of the enzymes processing the various classes of lipids are still much uncharacterised and only few enzymes exist from which structures are available with each reaction intermediate.

All characterised β-oxidation systems use multifunctional enzymes (MFEs). Peroxisomal MFEs catalyse the second and third reaction of the β-oxidation cycle. MFEs can also form multienzyme complexes with other enzymes in the pathway, providing in some cases a full set of enzymes to complete the fatty acid degradation cycle. There is evidence that while being processed in the enzyme complex, the substrate molecule is not
fully detached from the enzyme when it moves from one active site to the next, but it moves through a “substrate channelling” mechanism.

The majority of cellular fatty acids are bound by proteins. Some of these proteins are merely buffering molecules that keep the free fatty acid content low, whereas others are able to carry a specific fatty acid derivative to sites in which the specific metabolite is needed. Binding between the protein and its ligand fatty acid derivative can be covalent or non-covalent. Non-covalent binding is often mediated by interactions between the protein’s binding pocket residues and the fatty acid derivative and especially with its functional moieties. Unspecific interactions also exist, especially between simple lipids and the proteins binding them.

The aim of the present study was to increase understanding of the mechanism of the acyl-CoA-protein interactions. The study concerns protein crystallographic studies of two proteins, rat peroxisomal MFE-1 and human cytosolic acyl-CoA binding protein (ACBP) from liver. Altogether three crystal structures will be described: unliganded MFE-1, unliganded ACBP, and liganded ACBP.
2 Review of the literature

The cellular lipidome comprises over a thousand different lipid molecules (van Meer 2005). Lipids are processed in a complex network of metabolic pathways, in which they are transported, activated and modified. There exist several receptors and transcription factors that sense the extracellular and intracellular concentration of lipids and adapt the cell to prevailing environmental and physiological conditions by regulating the expression levels and catalytic activity of the proteins involved in lipid metabolism. The present review discusses the physico-chemical properties characterising a particular class of lipids, fatty acids (FA), and their derivatives. In addition, certain key proteins that regulate cellular lipid distribution and participate in lipid metabolism are presented, for which detailed structural data are available.

2.1 Physical and chemical properties of lipids

Lipids comprise a diverse class of biomolecules and recently, a new classification system was designed for lipids (Fahy et al. 2005). In this system, the lipids are defined as “hydrophobic or amphipathic small molecules that may originate entirely or in part by carbanion-based condensations of thioesters and/or by carbocation-based condensation of isoprene units”. Lipids are categorised into eight classes, of which FAs constitute a class of one of the most fundamental categories of biological lipids. In biological systems, FAs exist as oxygen esters in a storage form, whereas more reactive thioesters are used in metabolic pathways.

2.1.1 Simple fatty acids

Fatty acids are characterised by repeating methylene units that give rise to their hydrophobic character. There are several subclasses in the FA category, of which the most reduced members are the straight-chain saturated FAs that possess a terminal carboxylic acid group and a hydrophobic acyl chain. The acyl chain of the FA may have various
modifications, including one or more double bonds or even triple bonds. The acyl chain may also contain heteroatoms like oxygen, halogen, nitrogen, and sulfur atoms. Additionally, simple methyl groups may branch off the acyl chain, or branching can be further extended resulting in cyclic compounds that are further varied by the inclusion of functionalities such as alcohols, epoxides, and nitriles. Also included in the FA subcategory are important thioesters such as with coenzyme A, acyl carrier protein and esters with carnitine and adenylates, all of which are involved in metabolic pathways.

For convenience of discussion in the context of the metabolism, straight chain fatty acids are divided into short (less than 8 carbon atoms), medium (8 to 12 carbon atoms), long (14 to 18 carbon atoms) and very-long (more than 20 carbon atoms) chain fatty acids based on the chain length. The stereochemical aspects of fatty acids are crucial for efficient processing in cells. For example, the position of double bonds either in even (e.g., between C2-C3) or odd (e.g., between C3-C4) numbered carbon atoms as well as the cis/trans isomerism of this bond determines whether it can be directly converted by the core β-oxidation enzymes or if auxiliary enzymes are required.

Solubility of the free FAs and fatty acyl-CoAs in water is poor (Table 1), and it decreases as the acyl chain grows in length. Roughly, the critical micelle concentration (CMC) is halved by the addition of one methylene unit to a straight-chain hydrophobic group attached to a single terminal hydrophilic group, like coenzyme A or a carboxylic acid. The ionic strength of the medium and the presence of Mg²⁺ ions also decrease the solubility of acyl-CoAs. (Constantinides & Steim 1985.) Due to poor solubility, FAs and their CoA thioesters readily form micelles and aggregates in aqueous solutions even at very low concentrations, and little data is available on the solubility of long-chain fatty acids in physiological aqueous solutions. Similarly, the micelle formation of, for example, nutritious lipids must be decreased in the gastrointestinal tract by bile acids that increase their solubility and allow the digestion and uptake of lipid compounds. In blood, the majority of the free fatty acids are bound to serum albumin and minor amounts are carried by plasma lipoproteins.

Table 1. Critical micelle concentration of some physiologically important fatty acyl-CoAs and their free fatty acids.

<table>
<thead>
<tr>
<th>Fatty acyl-CoA</th>
<th>No. of carbon atoms in the acyl chain</th>
<th>CMC (µM, ±5%)¹</th>
<th>Solubility of the free fatty acid (µM)³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lauroyl-CoA</td>
<td>12</td>
<td>1.130</td>
<td>CMC 1 580², solubility &gt; 500 µM³</td>
</tr>
<tr>
<td>Myristoyl-CoA</td>
<td>14</td>
<td>2.10</td>
<td>CMC 223², solubility 20–30 µM³</td>
</tr>
<tr>
<td>Pentadecanoyl-CoA</td>
<td>15</td>
<td>114</td>
<td>NA</td>
</tr>
<tr>
<td>Palmitoyl-CoA</td>
<td>16</td>
<td>42, 77⁴</td>
<td>&lt; µM³</td>
</tr>
<tr>
<td>Heptadecanoyl-CoA</td>
<td>17</td>
<td>23</td>
<td>NA</td>
</tr>
<tr>
<td>Stearoyl-CoA</td>
<td>18</td>
<td>7.5</td>
<td>solubility &lt; µM³</td>
</tr>
<tr>
<td>Oleoyl-CoA</td>
<td>18:1</td>
<td>32</td>
<td>solubility &lt; µM³</td>
</tr>
</tbody>
</table>

¹ (Smith & Powell 1986); ² HEPES buffer at pH 7.3 (Kimura et al. 1998); ³ Phosphate buffer at pH 7.4 (Vorum et al. 1992); ⁴ Phosphate buffer at pH 7.4 (Constantinides & Steim 1985); NA, data not available.
2.1.2 Acyl-coenzyme A

The CoA molecule was first isolated from porcine liver as a heat stable cofactor by Lippmann in 1946 (Lippmann & Kaplan 1946) and the covalent structure was reported in 1953 (Baddiley et al. 1953). Coenzyme A is one of the most widely used cofactors in Nature and acyl-CoA derivatives are abundant in several primary and secondary metabolic pathways and regulatory processes (Kragelund et al. 1999a). Additionally, acetyl-CoA containing two carbon atoms in the acyl chain (Fig. 1) is a key metabolite in the processing of various biological compounds such as amino acids, carbohydrates and fats.

<table>
<thead>
<tr>
<th>acetyl coenzyme A</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3'-phosphate ADP</td>
<td>pantetheine</td>
</tr>
</tbody>
</table>

Fig. 1. Chemical structure of the acetyl-CoA molecule. The pantetheine unit has one chiral carbon atom (*) and two peptide bonds (P). The charges at the phosphate oxygen at neutral pH are indicated (-). The acetyl carbons are labelled and properties of the C1 and C2 are indicated. (Adopted from Mishra & Drueckhammer 2000.)

The CoA moiety can be divided into three modules, as depicted in Fig. 1. The first β-mercaptoethylamine module possesses the terminal thiol functionality that serves as a linkage point for the hydrophobic acyl chain. The β-mercaptoethylamine moiety is connected via an amide link to a pantothenic acid moiety. The relatively rigid pantothenic acid limits the dynamic properties of the whole molecule which contributes favourably to the entropic cost associated with, for example, binding on separate active sites in multienzyme complexes. (D'Ordine et al. 1995.) It also has one chiral centre (marked with an asterisk in Fig. 1; it has the substituents -H and -OH in addition to the large 3'-phosphate ADP and β-mercaptoethylamine moieties) in the pantetheine moiety and includes the dimethyl groups, also called “geminal dimethyls”. It has also been suggested
that the methyl substituents play an important role in determining the dynamics and conformational preferences of the CoA moiety, but they probably do not have a direct role in the reactions of the FA acyl group during metabolism (D’Ordine et al. 1995). Solution studies suggest that some interactions between protons of the pantetheine group and adenine and 3’-phosphate group occur (D’Ordine et al. 1995, Wu et al. 1998). The pantetheine unit is joined to 3’-phosphate adenosyl diphosphate via a 5’-pyrophosphate bridge, which includes mobile phosphoanhydrides. The 3’ and 5’-phosphates are charged at a neutral pH and they are often involved in hydrogen bonding or salt bridge contacts with the charged residues of binding pockets in proteins. The adenine ring is another important determinant in ligand binding and it often stacks with aromatic residues in protein structures (Engel & Wierenga 1996).

While the structure of the CoA molecule is fairly complex (Fig. 1), functionally it is a simple molecule. The presence of the CoA moiety activates the FA molecule for cellular metabolism and the enzymatic reactions of the CoA concern predominantly the thiol group, whereas reactions of the CoA thioesters involve the thioester group and/or the acyl moiety. An exception is found in nudix hydrolases, the recently found mammalian peroxisomal member, which has acyl-CoA diphosphatase activity towards long and very-long-chain acyl-CoAs producing acyl-4’-phosphopantetheine and 3’-phosphate AMP (Ofman et al. 2006). Generally, in acyl-CoA processing the rest of the CoA molecule serves as a recognition element for binding to enzyme active sites. Acylation of the thiol group results in an acyl derivative that is activated in two ways: the thioester carbonyl can react as an electrophile towards an attack by a nucleophilic cosubstrate, and the thioesters α-carbon can react as a nucleophile after deprotonation. (Mishra & Drueckhammer 2000.) The thioester sulfur atom makes the bond linking the acyl chain to CoA more reactive compared to an oxygen ester bond. The presence of a sulfur atom instead of oxygen causes the carbonyl atom to be more reactive to a nucleophilic attack and the α-carbon atom to be more easily deprotonated. In the transition state of both of these reactions the thioester oxygen develops a negative charge, which is often stabilised by an oxyanion hole in the protein structure. (Kursula et al. 2002, Bhaumik et al. 2005.)

2.1.3 Acyl carrier protein

Acyl carrier protein has a central role in type II fatty acid synthesis, which is the pathway utilising soluble monomeric enzymes in prokaryotes and plants. By contrast, in type I fatty acid synthesis the ACP is covalently bound in the enzyme complex (White et al. 2005). ACP is an acidic, 9 kDa protein and all ACPs, with the exception of the cleavable presequences that distinguish prokaryotic ACPs from the mitochondrial ones, are structurally similar and share a high level of sequence similarity with the ACP from E. coli (Rawlings & Cronan 1992). Recent studies have shown that in mitochondria, a soluble ACP and a respiratory complex I-associated ACP coexist. The latter is essential for the stability of the complex (Schulte 2001).

All of the intermediates required along the type II fatty acid synthesis pathway are carried by ACP. The acyl intermediates involved in the fatty acid biosynthesis are bound to ACP through a covalent thioester linkage with the terminal sulfhydryl of the 4’-
phosphopantetheine group (Hancock et al. 1972, Zhang et al. 2001). The prosthetic sulfhydryl is the only thiol group in Escherichia coli ACP, and it is attached to the protein via a phosphodiester bond from Ser36 of the fully conserved Asp-Ser-Leu motif at the end of helix 2 (Holak et al. 1988, Kim et al. 1990). The substrate is shuttled between the various enzymes that do not interact directly with each other, but require substrate transfer mediated by ACP (Honeyman & Fawcett 2000). In addition to fatty acid synthesis, ACP also has an important role in the synthesis of other compounds, like bacterial glycerolipids and polyamino acid antibiotics. (Lipmann 1980, Rock & Cronan 1996.)

Detailed structural information is available for E. coli ACP (Holak et al. 1988). Since the publication of the first structure, other studies have characterised the structure of the ACP from several other species and with different ligands, demonstrating a well conserved fold. The protein is composed of four α-helices with a C-terminal structured loop and a loop linking helices 1 and 2 (Fig. 2). The most prominent differences between ACPs of different origin are located at the loop regions, and they may be due to different experimental methods (Xu et al. 2001, White et al. 2005). It is currently thought that the interaction between ACP and the modifying enzyme occurs via specific interactions with helix 2 (Zhang et al. 2003).

The binding site for the prosthetic group on the surface of ACP has been mapped to an exposed hydrophobic pocket around Ser36 (Fig. 2A). The binding site of the attached acyl intermediate in butyryl-ACP has been assigned using NMR chemical shift perturbation studies, and the butyryl group shows additional interactions with the end of helix 3 (side of the pocket) and the loop connecting helices 2 and 3 (bottom of the pocket). The interactions identified in the NMR studies are consistent with the idea that ACPs interact with the bound acyl intermediates by using the hydrophobic cavity that resides in a cleft flanked by helices 2, 3 and 4 as is also seen in Fig. 2B. (Roujeinikova et al. 2002.) In this X-ray structure, the butyryl chain was not visible and it was assigned using molecular modelling resulting in a binding site different from the one identified in NMR experiments. It has been proposed (White et al. 2005), that the acyl chain and the prosthetic group go through a rapid two-state exchange (one, in which the prosthetic group is bound to the ACP surface; and the other, in which it is solvent exposed) in solution resulting in a change in the two-state equilibrium depending on the chain length of the acyl group. As the acyl chain length increases, hydrophobic interactions with the protein become more favourable than those which expose the acyl chain to the solvent. This hypothesis is further supported by the ability of the acyl chain to stabilise the protein against pH-induced denaturation (Rock et al. 1981).
2.2 Fatty acid activation and transport

Before the energy contained in fatty acids can be obtained in β-oxidation and subsequently in the Krebs cycle, the fatty acids must be activated to their corresponding CoA derivatives. Acyl-CoA synthetases (ACSs) catalyse this first reaction of fatty acid metabolism by ligating fatty acid and CoA in a two-step reaction requiring energy from ATP (Bar-Tana et al. 1973). There are distinct synthetases for the long-chain (ACSL) and very-long-chain fatty acids (ACSVL). The mechanism of the branched chain fatty acid and bile acid intermediate activation is not fully understood, but there are reports indicating that branched chain FAs can be activated by both ACSL (Vanhove et al. 1991) and ACSVL (Steinberg et al. 1999). Activation of the bulky bile acid intermediates has been reported to occur only in liver microsomes and it requires cholesteryl-CoA synthetase and trihydroxyxypocprostanoyl-CoA synthetase (Van Veldhoven et al. 1996; for a review see Watkins 1997).

In recent studies, the mitochondrial ACSs have been reported to be the ones exclusively activating short and medium-chain fatty acids (Fujino et al. 2001) even though some discrepancy exists about the ability of the peroxisomal membrane ACSL to

Fig. 2. The X-ray structure of *E. coli* butyryl-ACP (PDB ID 1L0I, (Roujeinikova et al. 2002)) with the butyryl-pantetheine group modelled in the structure. A: Ribbon diagram showing the organisation of the four α-helices. The conserved Ser36 is shown as a thick ball and stick model pointing out from helix 2, and the prosthetic butyryl-pantetheine group attached to it is shown as a thinner grey ball-and-stick model. B: Surface representation of the protein (rotated ca. -10° around y axis with respect to A) shows the overall shape of the protein with the butyryl acyl chain buried in a hydrophobic pocket lined by residues in helices 2, 3 and 4. The figure was prepared with Ribbons (Carson 1991).
activate medium-chain FAs (Reddy & Hashimoto 2001). Long-chain FAs are activated on
the cytosolic face of the outer mitochondrial membrane (Aas 1971, Hesler et al. 1990)
and they are actively transported into the mitochondrial matrix by a carnitine-dependent
route. During the transport, the long-chain acyl molecule is first linked to the carnitine by
a carnitine acyltransferase I, which is located on the outer mitochondrial membrane, and
it is subsequently exchanged across the inner mitochondrial membrane by the carnitine
acylcarnitine translocase for a carnitine molecule in the matrix (Pande 1975, Eaton et al.
Vaz & Wanders 2002). Once inside the mitochondrial matrix, the acyl moiety is
detached from the carnitine and religated to CoA by carnitine acyltransferase II on the
matrix side of the inner mitochondrial membrane (Kerner & Hoppel 2000). Unlike long-
chain fatty acids, medium-chain and short-chain fatty acids are able to cross the
mitochondrial membranes freely and they are activated in the matrix by short-chain and
medium-chain ACSs (Eaton et al. 1996).

Peroxisomal membranes contain at least two distinct acyl-CoA synthetases. One exists
on the cytosolic side of the membrane and it activates long-chain fatty acids for
subsequent transport into peroxisomes (Mannaerts et al. 1982, Lazo et al. 1990). The
other characterised synthetase activates very-long-chain FAs (Hashimoto 1996).
Activation of 2-methyl-branched fatty acids is not well understood, but it has been
suggested that the long-chain acyl-CoA synthetase residing on the peroxisomal
membrane is responsible for this activity (Vanhove et al. 1991, Wanders et al. 1992).
Transport of FAs across the peroxisomal membrane is still subject to speculation.
Experiments with purified peroxisomes have shown that the membrane is permeable to
< 1 kDa molecules, which comprise small solutes but not large molecules such as acyl-
CoA (Kerner & Hoppel 2000, Antonenkov et al. 2004).

The peroxisomal ACSL enzyme is identical at the protein level with the ones found in
mitochondria and ER (Suzuki et al. 1990). Very-long-chain FAs, on the other hand, are
activated inside peroxisomes by ACSL, which is associated with the peroxisomal
membrane (Lazo et al. 1990, Singh et al. 1992). This enzyme can also be found on ER
but not in mitochondria (Suzuki et al. 1990), which may explain why very-long-chain
FAs are β-oxidised exclusively in peroxisomes and not in mitochondria. Based on their
biochemical and immunochemical properties, as well as on primary structure comparison,
ACSL and ACSVL are different enzymes. Human peroxisomes also have an ACSVL,
which is able to activate fatty acids with both 2-methyl branched and very-long-chain
fatty acids (Mihalik et al. 2002) in addition to ACSL activity.

2.2.1 Intracellular trafficking of fatty acyl

As FAs are poorly soluble in cellular aqueous environments and they are readily
embedded in cellular membrane systems, specific trafficking mechanisms are required for
their delivery to various cellular compartments. Several proteins are dedicated to
transporting cellular lipids into different sites in the cell. Here, the properties of the fatty
acid binding protein (FABP), the sterol carrier protein type 2 (SCP-2) and ACBP are
discussed, as structural information is available for these transport proteins.
2.2.1.1 Fatty acid binding proteins

The large family of the fatty acid binding proteins (also called intracellular lipid-binding proteins) has an important role in controlling lipid trafficking in different tissues. Fatty acid binding proteins comprise a family of intracellular 14–16 kDa proteins that bind various amphiphilic molecules, like fatty acids, bile acids and retinoids (Ockner et al. 1972, Veerkamp & Maatman 1995). The FABP family has predecessors from 1 billion years ago and it is one of the few large protein families, from which dozens of representative X-ray and NMR structures are available in apo and holo forms (Fig. 3, for a review, see e.g., Lücke et al. 2003).

Because of its ability to bind a variety of essential lipids, several roles have been suggested for FABP. However, knock-out mice without FABP are still viable and these studies have failed to determine any critical role for FABPs. FABPs have been suggested to be both passive “buffering molecules” helping FAs and other lipids remain soluble, and active chaperones facilitating the transport of the lipids to specific targets, like membranes and other proteins (Glatz & van der Vusse 1996, Bernlohr et al. 1997). FABPs are actively expressed in tissues with high lipid metabolism, like skeletal muscle, liver, intestine and adipose tissues. The level of FABPs in cells may reach 1–3% in adipocytes and even 8% in other tissues (Coe & Bernlohr 1997, Schaap et al. 1998).

It has been proposed that the lipid sorting mechanism of various FABPs, which generates specificity for certain type of lipids, is linked to properties such as the electrostatic surface potential, which would result in the ligand approaching the binding pocket of FABP at a different rate (Simpson et al. 1999). Thus, the affinity for different ligands would be controlled by changes in the surface residue pattern that either enhance or suppress protein-ligand contacts. Another suggested mechanism is based on the altered folding stability resulting from ligand binding, which would cause a change in the protein stability, and consequently in the cellular half life (Bernlohr et al. 1999).

Unlike other lipid binding proteins from which high resolution structures are available, such as human serum albumin, maize non-specific lipid transfer protein and peroxisomal proliferator activated receptor δ (PPARδ) that have a helical fold (Hamilton et al. 2001, Hamilton 2002, Reese & Banaszak 2004), the structure of FABP consists of ten antiparallel β-strands that wrap around a central cavity as a β-barrel fold, also called “β-clam” (Fig. 3). The barrel is capped from the top by two short α-helices, and the bottom is closed with hydrophobic residues from the surrounding β-strands. Differences between the liganded and the unliganded lipid binding proteins involve mainly the region surrounding the beta C β-D β turn and α-helix II, which is often referred to as the “portal region” defining the most common entry point for ligand binding. The lipid binding site is in the interior of the barrel and it is lined by both hydrophobic and charged residues. The bound lipid occupies only about 40% of the binding cavity and there are several ordered solvent molecules, which are presumed to account for both fold stability and lipid binding. (LaLonde et al. 1994, Ory & Banaszak 1999.)

FABPs are classified into four subfamilies that separate these proteins according to amino acid sequence similarities and ligand binding properties. The subfamily I binds retinoids, i.e., vitamin A derivatives. In these proteins, the functional head group of the ligand is always buried in the binding cavity and it interacts non-covalently with its
carboxyl group with conserved charged residues. (Noy 2000.) The members of subfamily II bind relatively bulky lipids in addition to fatty acids. The lipid binding mode is different from that of type I, as L-FABPs of this subfamily are able to bind two long-chain fatty acids at the same time: one FA is bound in a bent conformation at the bottom of the binding cavity and the other low affinity FA interacts with its hydrophilic end with the mid section of first FA, and it extends its polar carboxyl group outside the binding pocket. (Sacchettini et al. 1990.) In subfamily III there is only one member, intestinal-type FABP, which binds a single long-chain FA in a slightly curved conformation in a similar fashion as seen in subtype I. However, FA is buried deeper in the pocket and its ω-end interacts with a conserved Phe55, which has been suggested to function as a portal lid for the binding site. (Sacchettini et al. 1988.) The last subfamily is type IV, which contains the most diverse set of FABPs. The endogenous ligand is not known for this subtype, but structures obtained in the presence of FAs have shown that long-chain molecules are bound inside the barrel fold in a U-shaped conformation while very-long-chain polyunsaturated ligands adopt a spiral conformation. (Cowan et al. 1993.)

Fig. 3. Soluble “small” lipid binding proteins. The structures of type I fatty acid binding protein with a trans-retinol ligand (PDB ID 1CRB, (Cowan et al. 1993)), sterol carrier protein type 2-like domain with Triton X-100 (PDB ID 1IKT, (Haapalainen et al. 2001)) and acyl-CoA binding protein with palmitoyl-CoA (PDB ID 1ACA, (Kragelund et al. 1993)) are shown. The figure was prepared with Ribbons (Carson 1991).

2.2.1.2 Sterol carrier protein

The high asymmetry in the distribution of cholesterol between the aqueous phase and membranes as well as its extremely poor solubility requires that there are specific mechanisms that transport cholesterol and prevent it from diffusing to the outer membrane leaflet. While it is not currently known whether soluble cholesterol carriers have a role in target-specific intracellular cholesterol trafficking, some candidates have been characterised that are able to promote intermembrane change of sterols in vitro. (Zurzolo et al. 1994.)

The best studied sterol carrier is SCP-2, whose role in peroxisomes has been suggested to be similar to that of ACBP in the cytosol. More specifically, it may protect the various
CoA esters from unspecific hydrolysis by the peroxisomal thioesterases or it may be a part of the peroxisomal antioxidiant mechanism hiding the potential double bonds of the fatty acyl-CoA from superoxide peroxidation. The structure of the apo SCP-2 was initially determined by NMR spectroscopy (Szyperski et al. 1993) and a more complete structure of the apo form followed (Garcia et al. 2000). The crystal structure of apo rabbit (Choinowski et al. 2000) SCP-2 appeared to be very similar to the NMR structures. From these studies the binding site could not be confirmed, but a putative hydrophobic tunnel lined with non-polar side chains was suggested as a candidate site for accommodating hydrophobic lipid molecules. In addition to independent entities, SCP-2s are also present as domains in certain larger proteins, like MFE-2 and SCP-2/SCP-x. (Seedorf et al. 1998.) The structure of the peroxisomal MFE-2 SCP-2-like domain with a bound Triton X-100 molecule is known (Haapalainen et al. 2001). This structure is composed of an α/β-fold (Fig. 3) with the ligand binding pocket on a β-sheet platform lined by two helices. Interestingly, the peroxisomal targeting signal was ordered and solvent exposed in this liganded form, which is not the case in unliganded SCP-2. (Choinowski et al. 2000, Haapalainen et al. 2001.) Hydrophobic tunnel-like cavities like those seen in SCP-2 structures are also present in non-specific lipid transfer proteins in plants, and these have been identified as lipid binding sites (Hoh et al. 2005). Although these proteins are even smaller than SCP-2 and apparently share no sequence similarity, the characteristics of the tunnels in both proteins, namely width, shape, hydrophobicity and length are very similar, suggesting a common functional role.

2.2.1.3 Acyl-CoA binding proteins

Long-chain fatty acyl-CoAs are bound in a 1:1 stoichiometry by specific intracellular carrier proteins called acyl-CoA-binding proteins. These 10 kDa proteins are conserved from yeast to mammal. In addition to independent proteins, ACBP domains are also found in several multidomain proteins, like membrane associated proteins with an N-terminal ACBP (Chye et al. 1999), multifunctional β-oxidation enzymes having both the ACBP and the Δ^3-Δ^2-enoyl-CoA isomerase (Geisbrecht et al. 1999), and proteins with an ACBP domain and an ankyrin repeat (Chye et al. 2000). Currently, it is not known whether these ACBP domains are able to bind long-chain acyl-CoA. However, bovine, yeast and Arabidopsis thaliana ACBPs are able to bind both saturated and unsaturated acyl-CoA, but not fatty acids, acyl carnitines or cholesterol. (Rasmussen et al. 1990, Rosendal et al. 1993, Knudsen et al. 1994, Faergeman et al. 1996.)

NMR structures have been obtained for bovine ACBP (bACBP) with palmitoyl-CoA (Kragelund et al. 1993). X-ray structures have also been determined for apo bovine ACBP and holo Plasmodium falciparum ACBP (PfACBP) (van Aalten et al. 2001). The polypeptide chain folds into a conserved four-helix bundle fold (Fig. 4), which is different from the topology of ACP. The NMR and the X-ray studies have shown that the overall fold is practically identical for the bovine apo and holo forms except for a slight but significant compacting of the A4 helix upon ligand binding.
The protein-ligand interactions can be grouped for the three regions of acyl-CoA such that each unit of acyl-CoA (the adenine ring, the 3'-phosphate, and the acyl chain) interacts with different parts of the protein. The adenine ring and the 3'-phosphate binding sites are structurally most conserved between bovine and PfACBP and at the sequence level also in other ACBPs. (Burton et al. 2005.) The 3'-phosphate group, which contributes 40% of the binding energy, hydrogen bonds to residues Tyr28, Lys32 and Lys54. The aromatic ring of Tyr31 stacks with the adenine ring of CoA forming a tight binding pocket for the adenine ring together with Tyr73 and Phe5. (Kragelund et al. 1993, Kragelund et al. 1999a.)

![NMR-structure of bovine ACBP complexed with palmitoyl-CoA (PDB ID 1ACA (Kragelund et al. 1993)). The helices are labelled as A1, A2, A3 and A4. Residues involved in ligand binding are shown in dark grey as thin sticks. Palmitoyl-CoA is shown as thick sticks, and the acyl chain and the CoA moieties are coloured in white and dark grey, respectively. The thioester sulphur and the ω-end of the ligand are indicated with an arrow. The topology of this four-helix bundle is different from ACP shown in Fig. 2. The figure was prepared with Ribbons (Carson 1991).](image)

The ω-end of the acyl chain of the ligand is buried in a hydrophobic groove between helices A2 and A3. In bovine ACBP, the ω-end actually resides in a tunnel formed by Lys50 from helix A3 folding over towards helix A2 Asp21 and CoA 5'-phosphate. The ligand binding groove continues towards the loop connecting helices A1 and A2, and the
acyl chain is protected from the bulk solvent by the CoA end, which folds over the acyl chain in a spiral-like conformation. The ω-end of the acyl chain makes several non-polar contacts with the cleft between helices A2 and A3. Specifically, this involves residues Met24, Leu25 and Ala53 of the bovine protein. (Kragelund et al. 1993.) The binding specificities between bACBP and PfACBP are different and the bovine protein was shown to have a peak at C₁₈ whereas the PfACBP peak was C₁₄ (van Aalten et al. 2001). The difference in the binding specificity can be explained by a structural comparison of the two proteins, which shows small but significant differences in the architecture of the binding pocket. The loop connecting helices A1 and A2 is two residues longer in PfACBP. Substitutions Ala53Lys, Lys50Ile and Asp21Asn, in turn, change the ω-end of the binding groove such that shorter acyl chains bind more easily to PfACBP, which correlates well with the fact that C₁₄-CoA is used in large amounts to construct the GPI-anchored protein coat around the parasite. In this process PfACBP donates C₁₄-CoA directly to the fatty acid remodelling machinery (Milne & Ferguson 2000). Even though the mammalian ACBP function in vivo is currently not known, for P. falciparum the gene coding ACBP is essential (Milne et al. 2001).

In vitro studies have shown that the acyl-CoA-ACBP complex is the preferred substrate of e.g. carnitine-palmitoyl transferase I (Abo-Hashema et al. 2001) and microsomal acyl-CoA cholesterol acyltransferase (Chao et al. 2003). Moreover, ACBP is able to extract acyl-CoA from artificial liposomes, transport the obtained acyl-CoA to another membrane, and donate it to mitochondrial β-oxidation (Rasmussen et al. 1994). The mechanism by which ACBP transfers acyl-CoA directly with other proteins without release into the bulk solvent is not known in detail. However, based on atomic force microscopy studies with synthetic phospholipid bilayer membrane containing C₁₂ fatty acyl-CoA and ACBP (Cohen Simonsen et al. 2003), a mechanism is proposed in which fatty acyl-CoA initially has its fatty acyl chain buried in the outer membrane leaf and the CoA end is exposed to the solvent phase. Binding of the prominent CoA end to the ACBP binding pocket for CoA moieties begins the desorption leaving the acyl chain binding groove initially empty. Subsequently, the acyl chain detaches from the membrane and binds in the hydrophobic groove remaining protected from bulk solvent.

### 2.2.2 Protein-acyl-CoA interactions

The polarity of the acyl-CoA molecule requires that their binding pockets in proteins contain areas with both hydrophobic and hydrophilic character. Furthermore, the hydrophobic acyl chain of the acyl-CoA is poorly soluble in the aqueous phase, which efficiently drives it to localise into hydrophobic environments, for example a hydrophobic binding pocket on a protein surface, or a membrane. In the case of myristoyl-CoA, the acyl chain creates a hydrophobic cylinder with a length of about 14 Å. The forces contributing to the overall free energy of binding (ΔG) of this molecule can be separated to two groups: van der Waals interactions between hydrocarbon surfaces (ΔH) and the loss of water from the active site associated with binding (ΔS). (Watanabe et al. 1993.)
From the available structures of protein-acyl-CoA complexes it is seen that the CoA molecule can adopt a wide range of conformations. Despite the conformational freedom, certain preferred rotamers found in NMR studies of soluble CoA seem to be observed often also in protein-acyl-CoA complexes (D’Ordine et al. 1995). In a study by Engel and Wierenga (1996), fourteen structures of varying topology and containing the acyl-CoA ligand were analysed. The CoA molecule was found to adopt various conformations ranging from a sharply bent conformation in 2-enoyl-CoA hydratase (ECH) to an extended one in the case of biosynthetic thiolase (Engel & Wierenga 1996, Kim & Battaile 2002). This study did not reveal any conserved features for the acyl-CoA-protein interactions except that the 3’-phosphate was always pointing into the bulk solvent and the adenine ring was pointing into the core of the protein. The sequences involved in CoA contacts did not show any conservation (Engel & Wierenga 1996). The 3’-phosphate of the CoA is in some cases coordinated by conserved residues. For example, in the ACBP family the binding energy decreases significantly if substrate analogues missing this group are used in binding studies (Faergeman et al. 1996).

2.3 Fatty acid degradation

FAs are processed to smaller metabolites in cells by three major oxidative pathways: $\alpha$-oxidation, $\omega$-oxidation and $\beta$-oxidation. $\alpha$-Oxidation is required when a methyl group at C3 position prevents the conversion by 3-hydroxyacyl-CoA dehydrogenase to the 3-keto form in $\beta$-oxidation (Vanhove et al. 1991, Van Veldhoven et al. 1991). After the 4-step conversion, $\alpha$-oxidation eventually returns a branched chain FA with an asymmetric 2-methyl-substituted carbon atom. $\omega$-Oxidation is a cytochrome P450-dependent pathway that is used to process the $\omega$ or $\omega$-1 carbon of monocarboxylic FAs (Kundu et al. 1993).

$\beta$-Oxidation, the major degradative pathway of fatty acids, is compartmentalised in mammals into peroxisomes and mitochondria. In one cycle of either mitochondrial or peroxosomal $\beta$-oxidation, the fatty acid is shortened by two carbon atoms. $\beta$-Oxidation cycles are analogous in both subcellular compartments and they consists of four enzymatic steps of dehydrogenation/oxidation, hydration, NAD$^+$-dependent dehydrogenation and thiolytic cleavage (for a review, see e.g. Hiltunen & Qin 2000). Despite the same reactions, distinct peroxisomal and mitochondrial enzymes have evolved for each step (Eaton et al. 1996). The function of $\beta$-oxidation in peroxisomes and mitochondria differs: mitochondrial $\beta$-oxidation is directly linked to the electron transfer chain and oxidative phosphorylation providing a source of energy for cells via the citric acid cycle, while in peroxisomal $\beta$-oxidation the electrons released in the oxidation are donated directly to molecular oxygen and thus no energy is harvested in this process.

In mammals, the mitochondrial $\beta$-oxidation enzymes are responsible for the degradation of the major portion of short, medium and long-chain fatty acids (Eaton et al. 1996). In general, the peroxisomal $\beta$-oxidation enzymes recognise more complex substrates, including (poly)unsaturated fatty acids that are not accepted or only slowly metabolised by the mitochondrial enzymes. After peroxisomal $\beta$-oxidation, the chain-shortened products are targeted to mitochondrial oxidation. (Reddy & Mannaerts 1994, Eaton et al. 1996, Mannaerts et al. 2000.)
2.3.1 Multifunctional \(\beta\)-oxidation enzymes in mammals

The mitochondria, peroxisomes and glyoxisomes of some eukaryotes contain multienzyme complexes for \(\beta\)-oxidation, but their detailed architectures are different from the multienzyme complexes of bacteria. It remains unclear why these enzyme activities are sometimes unified in the form of a multienzyme complex, while the same activities are also separately carried out by the individual monofunctional enzymes. However, it has been suggested that more efficient and better controlled processing of the metabolites is achieved in multienzyme complexes in which the active centres are close to each other. In mammalian peroxisomes, two multifunctional enzymes, type 1 (MFE-1) and type 2 (MFE-2), coexist. These two enzymes are the only enzymes that catalyse the second and third steps of the \(\beta\)-oxidation spiral in the peroxisome. Despite their ability to catalyse similar reactions, they do not share significant sequence similarity and they have separate evolutionary origins. In the following section, the structural properties of MFE-1 and related enzymes are reviewed.

2.3.1.1 Multifunctional enzyme type 1

Peroxisomal MFE-1 (Osumi & Hashimoto 1979, Furuta et al. 1980) is a soluble monomeric protein having a homologue in mitochondria as the \(\alpha\)-chain of the trifunctional enzyme (TFE) (Uchida et al. 1992), which is a membrane bound dimer of the fatty acid \(\beta\)-oxidation multienzyme (FOM) \(\alpha_2\beta_2\) complex resulting eventually in an \(\alpha_4\beta_4\) assembly. In bacteria, this complex occurs as a soluble \(\alpha_2\beta_2\) monomer. Unlike other known \(\beta\)-oxidation multifunctional enzymes that occur as a complex consisting of several polypeptide chains, peroxisomal MFE-1 is monomeric.

Reactions catalysed by MFE-1 are shown in Fig. 5. In addition to 2-enoyl-CoA hydratase-1 activity (Furuta et al. 1980), the N-terminal part of mammalian MFE-1 also catalyses several isomerisation reactions in \textit{in vitro} studies, including isomerisation of \textit{cis}-3 into \textit{trans}-2-enoyl-CoA, \textit{trans}-3-enoyl-CoA into \textit{trans}-2-enoyl-CoA and \textit{trans}-2,\textit{cis}-5-dienoyl-CoA into \textit{trans}-3,\textit{cis}-5-dienoyl-CoA. The latter is catalysed with the highest rate and has also been proposed to be the \textit{in vivo} function of peroxisomal MFE-1 (Zhang et al. 2002). In addition to CoA substrates with a linear acyl chain, MFE-1 also accepts hydroxylated C\textsubscript{27} bile acid synthesis intermediates and dicarboxylic acids (Reddy & Hashimoto 2001, Ferdinandusse et al. 2005). The catalytic residues comprising the oxyanion hole for the hydratase and isomerase reactions in the mitochondrial enzyme (Palosaari et al. 1991, Engel et al. 1996) are also fully conserved in rat MFE-1 (residues Glu103, Glu123, Ala61 and Gly100).

The dehydrogenase part of MFE-1 has affinity for linear (3S)-hydroxy substrates with a chain length between C\textsubscript{4} and C\textsubscript{16} with highest activity occurring with the C\textsubscript{10} substrate (Osumi & Hashimoto 1980). The dehydrogenase activity for bile acid intermediates is even more stereospecific and is functional only towards the (24S)-hydroxy,(25R)-methyl intermediate (Kurosawa et al. 2001).
Fig. 5. Reactions catalysed by peroxisomal MFE-1. The vertical arrows describe the classical β-oxidation reactions and the horizontal inbound arrows outside the frame mark the reactions of the auxiliary pathway importing substrates into the classical β-oxidation. The outbound horizontal arrow indicates the route to auxiliary dienoyl-CoA processing enzymes that ultimately produce trans-3-enoyl-CoA. MFE-1 catalyses the isomerisation, hydration and dehydrogenation reactions shown inside the frame. (Osumi & Hashimoto 1979, Furuta et al. 1980, Palosaari et al. 1991, Ferdinandusse et al. 2005.)

The expression of MFE-1, like the expression of the other peroxisomal enzymes of the “classical” β-oxidation pathway, is up-regulated by peroxisome proliferators and other biological ligands that are recognised by the peroxisome proliferator-activated receptor α (PPARα) (Reddy & Hashimoto 2001). Despite its ability to respond to changes in metabolic environment and being one of the most abundant proteins in the peroxisome, the biological role of MFE-1 is still open. The physiological significance of the isomerase activity of peroxisomal MFE-1 has been verified by expressing heterologous rat MFE-1 in ecclΔ yeast strains, in which the gene for the Δ3-Δ2-enoyl-CoA isomerase is removed. The strain regains its ability to grow on unsaturated fatty acids in the presence of heterologous expression of rat MFE-1, demonstrating the in vivo isomerase activity (Gurvitz et al. 1998). MFE-1(-/-) mice have been described and found to be clinically symptomless, which suggests that functional MFE-1 is not crucial for the peroxisomal β-oxidation for the tested diet (Qi et al. 1999). However, a double knock-out of MFE-1(-/-) and MFE-2(-/-) in mouse (Ferdinandusse et al. 2005) results in a more severe phenotype.
than seen in MFE-2(-/-) mice (Baes et al. 2000), which suggests a functional link between these two proteins that could be related to the peroxisomal bile acid metabolism (Ferdinandusse et al. 2005).

The metabolism of the (24S)-hydroxy-cholesterol derivatives via MFE-1 and an α-methylacyl-CoA racemase has been suggested as an alternative route for the conversion of cholesterol into bile acids in the case of MFE-2 deficiencies (Baes et al. 2000, Cuebas et al. 2002, Savolainen et al. 2004). MFE-1 hydratase is able to convert Δ24-THCS-CoA to the (24R,25S)-diastereomer of 24-OH-THCA-CoA, but this substrate is not accepted directly by MFE-1 dehydrogenase (Xu & Cuebas 1996, Dieuaide-Noubhani et al. 1997, Qin et al. 1997, Cuebas et al. 2002). However, the peroxisomal 2-methylacyl-CoA racemase is able to convert the 24R,25S substrate into the 24S,25R form, which is accepted by the MFE-1 dehydrogenase providing 24-keto-THCA-CoA, which is finally converted to cholyl-CoA by SCPx/thiolase (Cuebas et al. 2002). The combined action of MFE-1 and 2-methylacyl-CoA racemase in peroxisomal bile acid synthesis is supported by the observation that MFE-1 is up-regulated in mice having either MFE-2 or racemase knocked out (Baes et al. 2000, Savolainen et al. 2004).

### 2.3.1.2 Mitochondrial trifunctional enzyme

The mitochondrial matrix contains a set of soluble enzymes capable of catalysing the reactions of the β-oxidation pathway. Additionally, mitochondria also have a membrane-bound multifunctional enzyme complex (MTP) possessing 2-enoyl-CoA hydratase 1, and (3S)-hydroxyacyl-CoA dehydrogenase (α subunit) and 3-ketoacyl-CoA thiolase (β subunit) activities (Carpenter et al. 1992, Uchida et al. 1992). This enzyme complex is associated on the matrix side of the inner mitochondrial membrane and it assembles into an α₄β₄ complex which is a dimer of two α₂β₂ complexes. The enzyme expresses the highest catalytic activity towards acyl chain derivatives that contain more than 12 carbon atoms; it has activity for substrates with C₁₀ to C₁₆ acyl chains with the highest level for the C₁₂ substrate. (Uchida et al. 1992.)

The α-subunit has 65% sequence identity with peroxisomal MFE-1 (Kamijo et al. 1993). It has been shown that the dissociated α-subunit retains its hydratase and dehydrogenase activities, which suggests close resemblance of the subunit architecture between the α-subunit and MFE-1. For the β-subunit thiolase activity to be functional, a full enzyme complex is required (Ishikawa et al. 1997), although observations to the contrary have been made (Kamijo et al. 1994). The significant sequence identity of 35% of the Pseudomonas fragi α₄β₄ complex with the human MTP allowed creating a homology model of MTP. The model was reported to have reasonable dimer interface interactions between the subunits and revealed a concentration of positively charged residues in the centre of the complex, potentially localising the complex near negatively charged lipid bilayers or attracting substrate molecules with partially negative charges at their CoA end. (Ishikawa et al. 2004.) The MTP activity involves substrate channelling because of the absence of detectable β-oxidation intermediates (Yao & Schulz 1996), although cases inconsistent with this have also been reported (Eaton & Bartlett 1999).
2.3.2 Substrate channelling in the bacterial fatty acid metabolising complex

Many enzymes assemble into multifunctional enzyme complexes in which sequential enzyme reactions are carried out by transferring the reaction intermediates between different subunits without dissociation into the bulk solvent. This type of substrate channelling mechanism appears to be employed in various cellular processes, including glycolysis, citrate and urea cycle, and β-oxidation (Heath & Rose 1985, Watford 1991, Agius et al. 2002, Ishikawa et al. 2004). The structural features of these complexes remain uncharacterised for many systems, but, for example, the bacterial ~9 MDa pyruvate dehydrogenase multienzyme complex with covalently bound intermediates has been investigated in detail by electron microscopy (Milne et al. 2002) and more recently by X-ray crystallography (Frank et al. 2005). The substrate channelling mechanism in this massive complex is mediated by a lipoyl domain shuttling in the interdomain space. Also enzymes involved in fatty acid metabolism, such as fatty acid synthesis and β-oxidation, tend to be assembled such that efficient kinetics is achieved.

The bacterial fatty acid oxidation multienzyme complex (FOM) is composed of two α-subunits and two β-subunits. (Yang et al. 1991a, Yang et al. 1991b.) In E. coli and in P. fragi, the α-subunit is a 79 kDa multifunctional enzyme that has 2-enoyl-CoA hydratase-1 (ECH), (3S)-hydroxyacyl-CoA dehydrogenase (HAD) and Δ3-Δ2-enoyl-CoA isomerase activities (Pawar & Schulz 1981, DiRusso 1990, Sato et al. 1992). The amino acid sequence comparisons suggest a common ancestor with, for example, mammalian peroxisomal MFE-1 and mitochondrial MTP (DiRusso 1990). The β-subunit is a 41 kDa 3-ketoacyl-CoA thiolase (Pawar & Schulz 1981). All subunits of the bacterial FOM have 30–40% sequence identity with the corresponding monofunctional enzymes from various sources (DiRusso 1990, Ishikawa et al. 2004).

Recently, the crystal structure of the αβ2 β-oxidation multienzyme complex from P. fragi was published suggesting the first comprehensive model for the substrate channelling mechanism in β-oxidation (Ishikawa et al. 2004). The P. fragi FOM subunits are assembled into a ring-like structure where the catalytic centres surround a region open to bulk solvent (Fig. 6D). The sequence of β-oxidation reactions, however, does not follow the architecture of the subunits in this complex, in which the ECH is located between HAD and thiolase. The protruding helix-turn-helix motifs between the two HAD α-subunits form a four-helix bundle by polar and hydrophobic interactions and bind the α-subunits together. These HAD α-subdomains do not interact directly with the thiolase β-subunits. In the human mitochondrial monofunctional HAD dimer, these motifs are also present (Fig. 6B), but they do not participate in the dimer formation, which is accomplished in this enzyme by the C-terminal domains (Barycki et al. 1999). The ECH domains have no direct contacts between each other, but they both interact through the end of a helix and a loop with the exposed surface loops of the thiolase β-subunits.
Fig. 6. Architecture of the *P. fragi* FOM complex. A: Human mitochondrial enoyl-CoA isomerase with an acetyl-CoA (PDB ID 1SG4 (Partanen et al. 2004)). B: human mitochondrial 3-hydroxyacyl-CoA dehydrogenase (PDB ID 1F0Y (Barycki et al. 2000)) with acetyl-CoA and NAD⁺. C: α-Subunit of the fatty acid metabolising complex with acetyl-CoA and NAD⁺ (PDB ID 1WDM, (Ishikawa et al. 2004)). The 3'-phosphate ADP moiety is bound to the linker helix, the orientation of the domains is as in A and B. D: α₂β₂ architecture of FOM; NAD⁺ bound on the α-subunit HAD surface is shown in black sticks, acetyl-CoA on the thiolase surface is in light grey. Compared to panel C, the molecule was rotated ca. 180° clockwise. E: stereo pair showing the α-subunit of FOM in roughly the same orientation as in panel C, the ECH domain is shown in pale grey, the linker helix in dark grey, the HAD middle region in grey and the HAD C-terminal region in darkest grey. The figure was prepared with Ribbons (Carson 1991) and PyMOL (DeLano 2002).
The $\alpha_2\beta_2$ structure suggests that as the three sequential $\beta$-oxidation reactions proceed in the FOM, the substrate molecules are anchored during the first two reactions by their 3'-phosphate ADP moiety. The binding pocket of the adenine moiety is formed by Lys142 and Tyr249 of the ECH core, Leu290 of the linker helix (Figure 6), and Leu666 of HAD. This pocket is located such that the substrate is accessible to the active sites in both ECH and HAD, and the hydrophobic character of this region is conserved among the multifunctional enzymes (Ishikawa et al. 2004). It is suggested that this pocket serves as a pivot point for substrate transfer between the ECH and HAD active sites. The linker $\alpha$-helix participating in the ECH/HAD substrate binding is typical for the $\beta$-oxidation multifunctional enzymes and it connects ECH and HAD both in $\alpha$FOM and in MFE-1. It is also present in the monofunctional ECHs (Figure 6), such as human mitochondrial ECH (referred to as helix-10), in which it provides residue Leu255 for hydrophobic interaction with the adenine ring, and Lys256 and Lys259 that interact with the phosphate groups of the CoA (Partanen et al. 2004).

The NAD$^+$-binding domain in HAD is virtually identical with the monofunctional HAD as are the active site residues (Barycki et al. 1999, Barycki et al. 2000, Barycki et al. 2001). On the other hand, the CoA binding site is distinctly different between these two HAD enzymes. In human monofunctional HAD, the residues of the “inner” helix in the helix-turn-helix of the dimer form the hydrophobic platform for the adenine ring and polar contacts for the 3'-phosphate, whereas in bacterial FOM the dual-mode binding pocket is close to the linker helix (Figure 6). (Barycki et al. 2000, Ishikawa et al. 2004.) For the thiolase reaction to proceed, the substrate has to be pivoted to the more distant $\beta$-subunit in the multienzyme complex.
3 Aims of the present study

Degradation of cellular fatty acids requires efficient processing of reaction intermediates. The various enzymes catalysing fatty acid β-oxidation reactions occur as monofunctional and multifunctional enzymes that often assemble as complexes composed of several individual enzymes. The substrates are recognised in these complexes with high specificity with respect to the length and topology of the acyl chain. Coenzyme A and other activating groups, like ACP, also play an important role in facilitating fatty acid metabolism by activating the fatty acid substrate. For certain proteins, it has been shown that the preferred substrate is not the activated fatty acyl derivative alone, but the fatty acyl derivative carried by a transporter protein, like ACBP, and that this combination is more readily accepted. The substrate transfer between the transporter and target enzyme, or within a multifunctional enzyme, requires controlled events in binding and release of the substrate. In this study, two proteins that bind acyl-CoA derivatives and participate in their processing and trafficking were studied.

The following specific aims were laid out for the present study:

1. To predict the domain structure of rat peroxisomal MFE-1
2. To crystallise the dehydrogenase part of MFE-1 for structure determination
3. To determine the structure of MFE-1 by the X-ray diffraction method
4. To determine the structure of human ACBP
4 Materials and methods

In the following sections, the materials and methods used in the study are described. A more detailed description of the methods is provided in the original articles (papers I-IV).

4.1 Bacterial strains and plasmids

In the construction of expression plasmids, pUC18 subcloning vector was used together with the SureClone ligation kit (Amersham Pharmacia Biotech AB, Uppsala, Sweden) to generate plasmids containing the desired cDNA. For overexpression in *E. coli* cells, these cDNAs were digested out and ligated into pET vectors (Novagen, Madison, USA). For overexpression in *Pichia pastoris* yeast, pAO815 vector (Invitrogen, Groningen, The Netherlands) was used instead. A more detailed description of the cloning procedure and site directed mutagenesis of the MFE-1 variants is described in the original article (I).

4.1.1 Construction of the MFE-1 domain variants

For the studies described in paper I, various MFE-1 variants, in which specific domains were deleted, were constructed. These variants were made using the overlap extension PCR method (Ho et al. 1989) as described in the original article (I). The C-terminal truncations of the MFE-1 were generated using a QuickChange mutagenesis kit, and by PCR using Pfu DNA polymerase (Stratagene, Promega, Madison, USA) as described in the original articles (I and II).
4.2 Protein expression and purification

4.2.1 Production of the MFE-1 domain variants

The recombinant proteins (I, II, III) were produced in *E. coli* by overexpression in the pET expression vector (Novagen, Madison, WI). The BL21(DE3) pLysS bacterial cell line was used for expression. From an overnight culture, a 1:100 dilution was made to M9ZB medium supplemented with 34 µg/ml chloramphenicol and 50 mg/ml carbenicillin and grown at 37°C until the OD$_{600}$ reached 0.4–1.0. Expression was initiated by adding IPTG to a final concentration of 0.4 mM. Incubation was continued at 30°C or 35°C for three hours, and the cells were collected by centrifugation, washed and stored at -70°C.

4.2.2 Purification of the (3S)-hydroxyacyl-CoA dehydrogenase for crystallisation

The MFE-1 (3S)-hydroxyacyl-CoA dehydrogenase (MFE1-DH) protein purification was started from thawed cells (approximately 5.0 g) that were suspended in 50 ml of cold lysis buffer (50 mM HEPES pH 7.6, 1 mM EDTA, 0.5 mM PMSF, 0.5 mM benzamidine-HCl, 0.5 mM DTT, 100 mg/ml lysozyme). The suspension was incubated for 15 min at room temperature and the cells were disrupted by sonication. The cell debris was removed by centrifugation and the soluble fraction was applied onto a DEAE-Sephacel column for chromatography, followed by chromatography with Resource S (Amersham Pharmacia Biotech, Uppsala, Sweden), POROS S (PerSeptive Biosystems, Framingham, USA) and SuperDex 75 HR (Amersham Pharmacia Biotech, Uppsala, Sweden) in the final buffer containing 50 mM Tricine pH 7.6, 200 mM NaCl, and 1.5 mM NaN$_3$.

4.3 Enzyme assays

The 2-enoyl-CoA hydratase-1 (hydratase-1) and Δ$^3$-Δ$^2$-enoyl-CoA isomerase (isomerase) reactions were measured by following the oxidation of the (3S)-hydroxyacyl-CoA esters by an auxiliary (3S)-hydroxyacyl-CoA dehydrogenase in the presence of NAD$^+$ and MgCl$_2$ and the production of the stable Mg$^{2+}$ complex, which has an absorption maximum at 303 nm. *Trans*-2 and *trans*-3-enoyl-CoAs were used as substrates and Na-pyruvate and L-lactic dehydrogenase were added to regenerate NAD$^+$. All reactions were carried out at 22°C, pH 9.0. For combined activity, (3S)-hydroxyacyl-CoA dehydrogenase was not included in the reaction. MFE-1DH activity was measured as described (Palosaari et al. 1991) using (3S)-hydroxydecanoyl-CoA as a substrate.
4.4 Crystallisation

All crystallisation experiments were carried out using the hanging drop vapour diffusion method. Initial screening was done using sparse matrix screens from Hampton Research Inc. (Aliso Viejo, USA) and as described by Zeelen (Zeelen et al. 1992). In all experiments, an equal amount of protein and well solution was used.

4.4.1 Crystallisation of apo MFE1-DH

Before crystallisation experiments, the protein solution was concentrated to a concentration of 5 mg/ml and filtered through a 0.4 µm syringe filter. Crystallisation screens were set up to equilibrate at both +4 °C and +22°C. After 14 days, crystals were observed under conditions consisting of 30% poly(ethylene glycol) (PEG) 4000; 0.1 M sodium citrate, and 0.2 M ammonium acetate at pH 5.6. These conditions were further optimised and the protein concentration was increased to 10 mg/ml. Optimal crystallisation was obtained from a well solution of 26–27% PEG 4000; 0.2 M sodium citrate, 20 mM ammonium acetate, pH 5.6.

4.4.2 Crystallisation of apo hACBP

Lyophilised protein was dissolved to a concentration of 20 mg/ml in 50 mM MES, pH 7.0; 50 mM NaCl. Before setting up crystallisation experiments, Pb(NO₃)₂ was added to a concentration of 5 mM. Hampton Research Crystal Screens I and II were used to find initial crystallisation conditions. One microlitre of protein-Pb(NO₃)₂ solution was mixed with one microlitre of well solution. Single microcrystals were observed under light microscopy after 3 h incubation at +4 °C in 35% PEG 550 MME; 100 mM MES, pH 6.5; 10 mM ZnSO₄ and 3% 1-propanol. To facilitate the growth of larger crystals, similar drops were set up but without 1-propanol, which was found out to be essential for fast nucleation. In these drops without 1-propanol, heavy precipitate was observed after 10 h and single microcrystals from the previous drops were transferred approximately 3 h later to drops without 1-propanol. After additional equilibration for 10 h at +4°C, these microseeded crystals were flash frozen for data collection without additional cryoprotection.

4.4.3 Crystallisation of myristoyl-CoA hACBP

The lyophilised protein was dissolved to a 15 mg/ml solution in 50 mM MES, pH 7.0; 50 mM NaCl. Myristoyl-CoA (Sigma), dissolved in water, was added to a concentration of 1.8 mM to obtain ligand:protein molar excess. The screening for crystallisation conditions was performed as for the apo protein above. Single crystals were grown from 24% PEG MME 550; 100 mM MES, pH 6.5; 10 mM ZnSO₄.
4.5 X-ray diffraction data collection, processing and reduction

4.5.1 MFE1-DH

A single crystal was taken from the drop and directly frozen without adding any cryoprotectant in a 100 K nitrogen stream using a Hampton Research Cryoloop. Data were collected using a Marresearch 165 CCD detector at beamline ID711 at the MAXLab synchrotron, Lund, Sweden, to a 1.9 Å resolution. Indexing and data reduction were carried out using XDS (Kabsch 1993).

4.5.2 Apo hACBP

The data collection was carried out to 1.6 Å resolution at beam line X11 at the EMBL outstation at the DESY synchrotron, Hamburg, Germany. 0.5 degree oscillations were exposed on the MAR CCD detector using a wavelength of 0.91 Å. Indexing and scaling of the images with the HKL suite (Otwinowski & Minor 1997) suggested that the space group was $P_{6_5}2_2$ or $P_{6_1}2_2$ with the unit cell parameters $a = b = 43.89$ Å, $c = 177.57$ Å. A Matthews coefficient of $V_m = 2.6 \text{ Å}^3$/Dalton for one monomer per asymmetric unit suggested one monomer per asymmetric unit.

4.5.3 Myristoyl-CoA hACBP

The 1.4 Å dataset was collected from a crystal cryoprotected by soaking for 60 seconds in the mother liquor with an increased (35%) PEG MME 550 concentration before flash freezing. Data collection was carried out at beam line ID14-4 at the European Synchrotron Radiation Facility, Grenoble, France. 0.5 degree oscillations were exposed on the ADSC Q4 CCD detector. Data processing was performed with the program XDS and this suggested the space group $I2_3$ or $I2_13$ with the unit cell parameters $a = b = c = 118.49$ Å. A $V_m = 3.61 \text{ Å}^3$/Dalton for two monomers per asymmetric unit suggests two monomers per asymmetric unit.

4.6 Determination of the structure factor phases

4.6.1 Molecular replacement phasing of MFE1-DH

Several heavy metal atom derivative datasets were collected but none of these could be used for initial phasing. However, eventually the molecular replacement method was used successfully. The search model consisted of the dehydrogenase domains (residues 311–681) of the bacterial fatty acid metabolising complex (PDB ID 1WDK, Ishikawa et al.)
including domains C and D of MFE-1. The sidechains other than proline or glycine of this fragment were mutated to alanine to reduce model bias and molecular replacement was done with the program Phaser (McCoy et al. 2005) in both of the possible space groups. The solution obtained in $P4_12_12$ showed no overlapping symmetry mates and had a Z score of 15.07 indicating a plausible orientation of the protein molecule in the crystal lattice.

Before initiating manual model building, the phases were further improved by solvent flattening with DM (Cowtan & Zhang 1999).

### 4.6.2 SAD phasing of apo hACBP

The heavy metal atoms were located using the CNS package (Brunger et al. 1998) in space group $P6_522$. One Pb$^{2+}$ site and one Zn$^{2+}$ site were found. Density modification was performed using the “solvent flipping” method as implemented in CNS. The phases were combined with the observed amplitudes of the structure factors to calculate the electron density maps.

### 4.6.3 SAD phasing of myristoyl-CoA hACBP

A SAD approach similar to that for the apo protein was followed with the liganded crystals. Eight Zn$^{2+}$ ions were located using the program CNS. After density modification with CNS, the initial electron density map was used in ARP/wARP (Perrakis et al. 1999) for building the first molecular model.

### 4.7 Model building, refinement and validation of the structures

Manual model building was done using primarily $2f_o-f_e$ maps and $f_o-f_e$ difference maps. $R$ and $R_{free}$ (Brunger 1993) were followed together with deviations from ideal bond angles and lengths to maintain structural validity during the refinement process. For automatic model building programs ARP/wARP (Perrakis et al. 1999), Maid (Levitt 2001) and RESOLVE (Terwilliger 2000) were used. O (Jones et al. 1991) and Xfit (McRee 1999) were used for manual model building and visualisation of the electron density maps during structure refinement.

### 4.7.1 Model building and refinement of MFE1-DH

Solvent flattened phases were used to calculate initial electron density maps. These maps were used to manually add side chains to the polyalanine backbone with the program Xfit. The backbone was deleted in regions with poor electron density. Substantially continuous electron density was observed for the residues in domain D and in the buried
portions of domain E. The improved model was refined with the program Refmac5 (Murshudov et al. 1997). Subsequently, the model was further improved by doing manual and automatic model building with Resolve and Maid followed by Refmac5 refinement.

After the majority of the polyalanine backbone was completed with fitted side chains, TLS refinement as implemented in Refmac5 (Winn et al. 2001) was used with domains C, D and E as TLS-groups. The three C-terminal residues corresponding to peroxisomal targeting signal (PTS-1) showed poor electron density and were omitted from the final model. At the end of the refinement the R-factor was 17.0% (R$_{free}$ = 22.1 %) at 1.9 Å resolution.

The final coordinates were deposited in the Protein Data Bank (Berman et al. 2000) under the PDB ID 1ZCI.

### 4.7.2 Model building and refinement of apo hACBP

The density modified phases and the experimental amplitudes were used in the ARP/wARP program for automated model building. The resulting model was refined with the CNS package and manual model building was done with the program O to add alternate side chain conformations and ordered water molecules. Individual B-factors were refined for every atom. The final R-factor was 19.3% (R$_{free}$ = 23.0 %) at 1.56 Å resolution.

### 4.7.3 Model building and refinement of myristoyl-CoA hACBP

The density modified phases and experimental amplitudes were used in ARP/wARP for automatic model building. This model was further refined using CNS. O was used to manually adjust the side chain conformations and to add residues missing from the model. Further refinement was done with Refmac5. After including all sidechains, TLS refinement was carried out using each monomer as an individual TLS group.

Three 3’-phosphate AMP moieties of the ligand molecules had clear continuous density in the F$_o$-F$_c$ maps and these ligand fragments were added in the model (molecules C1, D1 and P1; C1 is bound to monomer B, molecules D1 and P1 are complexed to monomer A).

At this stage the F$_o$-F$_c$ difference maps indicated the presence of additional ligand fragments. Continuous density was observed starting from the 5’-phosphate group of the C1 molecule and extending from molecule B to molecule A. A stretch of positive density corresponding to the carbon atoms of an acyl chain was also observed in the acyl chain binding pocket of monomer B. These moieties were added in subsequent model building/refinement cycles and the data range was slowly extended to 1.4 Å, and anisotropic B-factors were refined for the protein atoms. Finally, alternative conformations for some side chains were added when appropriate and more ordered water molecules with a satisfactory hydrogen bonding network were added. At the end of the refinement the R-factor was 17.0% (R$_{free}$ = 19.3%) at 1.4 Å resolution.
4.7.3.1 Confirmation of the heavy metal positions in complexed hACBP

In order to confirm the location of heavy atoms, two additional datasets were collected from a single crystal at appropriate wavelengths to measure the anomalous signals of Zn ($\lambda = 1.2818$ Å), and S and P atoms ($\lambda = 1.7453$ Å). The 1.4 Å structure of the myristoyl-CoA complex (without ligands) was refined against these two datasets and anomalous difference maps were calculated to visualise the location of anomalous difference peaks in the asymmetric unit.

4.7.4 Structure validation and analysis

All structures were analysed and evaluated by using the programs WHATIF (Vriend 1990), PROCHECK (Laskowski et al. 1993), DSSP (Kabsch & Sander 1983), O (Jones et al. 1991) and PyMOL (DeLano 2002).

4.8 Sequence and structure comparison

Amino acid sequences were compared with the programs ClustalW (Thompson et al. 1994) and GeneDoc (Nicholas et al. 1997). Protein structures were superimposed with lsqkab (Kabsch 1976) and visualised with O, PyMOL, Xfit and Ribbons.
5 Results

5.1 Domain structure of rat peroxisomal MFE-1 (I)

Rat peroxisomal MFE-1 has 2-enoyl-CoA hydratase 1, several enoyl-CoA isomerase and (3S)-hydroxyacyl-CoA dehydrogenase activities. Unlike other \( \beta \)-oxidation multifunctional enzymes, MFE-1 is monomeric. To characterise the domain architecture of MFE-1, the polypeptide chain was divided into functional regions denoted as A, B, C, D, and E, based on sequence similarity with the monofunctional hydratases/isomerases and HADs (Fig. 7). The region A (residues 1–190) belongs to the hydratase/isomerase family. The region B (residues 191–280) does not show significant sequence similarity with the H1/I family but it was eventually found to include a helix-10 known to be important for the hydratase activity. The residues 281–467 (region C) comprise a region showing similarity to the NAD\(^+\)-binding domain of HAD, and residues 479–583 (region D) have sequence similarity with the C-terminal dimerisation domain characteristic of the monofunctional HADs. The ultimate C-terminus of MFE-1 (region E, residues 584–722) shows 30% sequence similarity with region D of the MFE-1 (residues 479–526) and 48% with the HAD dimerisation domain. The last three C-terminal residues are the type I peroxisomal targeting signal (PTS-1) SKL (Gould et al. 1988).

The H1/I family members usually assemble as trimers, or dimers of trimers. The stability of MFE-1 was studied by generating MFE-1 deletion variants and following the hydratase 1 activity. The variants were engineered based on the sequence alignment and structural data available from other family members. The variants were expressed in recombinant expression vectors in \( E. coli \) cells and the H1 activity was followed. The variant with a deleted C region (NAD\(^+\)-binding domain of the HAD) showed an increased hydratase-1 activity (Fig. 7C and Fig. 7D). When the deleted region was extended by 28 residues to include the N terminus of region D, or when the whole D region was deleted (Fig. 7E), no hydratase 1 activity was observed. The importance of the integrity of regions D and E was further supported by the observation that when using a variant in which the final 17 carboxy terminal residues were deleted, no HAD activity was detected.
Fig. 7. Domain structure of MFE-1 and the effects of domain deletions on the hydratase 1 activity. The rectangles in the panel of each MFE-1 variant show the domains of MFE-1 as determined in paper I. The structure of the crystallised variant (MFE1-DH, missing region A and most of B, paper III) is shown as cylinders coloured according to domain identity and superimposed on the backbone of the bacterial fatty acid metabolising complex α-subunit (Ishikawa et al. 2004). The catalytic residues of H1 are indicated with arrows. B–D: deletion variants missing domain C but showing hydratase 1 activity. The part of the MFE1-DH structure containing the corresponding residues is also shown to illustrate the interactions available for the H/I part. Constructs in panel C and D showed higher H1 activity than the wild type enzyme. E: inactive variant missing region E. F: inactive variant missing regions B and C.
The functional deletion variant was further studied by modifying the N and C termini of the construct resulting in the variant HisMFE(288–479)Δ, which retained the hydratase-1 activity and could be expressed in *E. coli*. This variant was also expressed together with the wild type MFE-1 (MFE-1wt) in *P. pastoris* and the hydratase 1 was assayed. In the HisMFE(288–479)Δ variant, the hydratase 1 activity of the purified protein was approximately 2-fold higher compared to the MFE-1wt with the two tested substrates, *trans*-2-decenoyl-CoA and *trans*-2-hexenoyl-CoA, whereas the isomerase activity assayed with *trans*-3-hexenoyl-CoA was reduced by about 80%. The *Km* values for both *trans*-3-hexenoyl-CoA and *trans*-2-hexenoyl-CoA were higher in HisMFE(288–479)Δ than in the MFE-1wt, but no significant change was observed when the *trans*-2-decenoyl-CoA substrate was used.

The secondary structure content of MFE-1wt and HisMFE(288–479)Δ was analysed with CD spectrum analysis in the far UV region (195–250 nm). The comparison of the spectra from HisMFE(288–479)Δ and MFE-1wt proteins did not reveal significant differences in the composition of the secondary structure elements. However, the stability of deletion variant HisMFE(288–479)Δ was less than that of MFE-1wt as shown by a drop of 9°C in the melting temperature.

### 5.2 Crystallisation of the rat peroxisomal MFE-1 dehydrogenase (II)

The stable deletion variant MFE(1–259)Δ (MFE1-DH, Fig. 7A) was found to have retained the MFE-1 dehydrogenase activity, and it was selected for crystallisation experiments. This protein comprises amino acids 260–722 of MFE-1 and also corresponds to the identified physiological degradation product of MFE-1 (Dieuaide-Noubhani *et al.* 1996). This variant was monomeric in gel filtration analysis and showed a specific dehydrogenase activity of 2.2 µmol/min/mg when assayed with (3S)-hydroxydecanoyl-CoA substrate. This value is comparable to that reported for the wild type MFE-1 (Palosaari & Hiltunen 1990).

Single crystals were obtained from conditions consisting of 27% PEG 4000; 0.2 M sodium citrate/20 mM ammonium acetate, pH 5.6; at room temperature. A full data set could be collected from a single crystal frozen in a 100 K nitrogen stream using a rotating anode home source to 2.45 Å resolution. Oscillations of 0.5° per frame were collected on a Mar345 image plate detector. The scaling of the data suggests point group *P*422 and the systematic absences along the *h00* and *00l* axes suggest the space group *P*41212 or *P*43212 with the unit-cell parameters *a* = *b* = 125.9 Å, *c* = 60.2 Å.

### 5.3 Crystal structure of the peroxisomal MFE-1 dehydrogenase (III)

The 1.9 Å structure of the C-terminal dehydrogenase variant MFE1-DH of rat peroxisomal MFE-1 was solved by molecular replacement using a polyalanine model for residues 311–681 of the bacterial fatty acid metabolising complex α-subunit (Ishikawa *et al.* 2004) as a search model. All residues of the MFE1-DH construct could be included and refined in the model, except for the final three C-terminal residues providing the
PTS-1 signal peptide SKL. Three regions of the backbone $C_{\alpha}$ atoms showed relatively high B-factors for regions near residues 355, 540 and 610. The first region provides the protruding helix-loop region near helix CH2 in the C domain. The second region is a protruding loop in the D domain, and the third region is the loop connecting domain D to domain E. There is only one outlier in the Ramachandran plot: the conformation of the residue Phe478 ($\phi = 46$, $\psi = -126$) is similar to the corresponding residue in the HAD structure.

The structure begins with an $\alpha$-helix extending from the C-terminal lobe to the N-terminal lobe of MFE1-DH (region B in paper I). Based on sequence and structure analysis, this helix corresponds to helix-10 of hmECI. Unlike in hmECI, in the apo MFE1-DH structure this helix is kinked. The dehydrogenase region (domain C) fold is similar to the HAD structure (Barycki et al. 1999). However, there is an important difference concerning the protruding helix-loop-helix motif that partly opens in the MFE-1DH structure. In $\alpha$FOM this helix-loop-helix motif is also present dimerising the two $\alpha$-subunits. Domain D of the MFE-1 corresponds to the dimerisation domain of the HAD enzyme. The D domain (approximately 130 residues) is connected to the E domain (approximately 110 residues) via a short loop (residues 607–614). Domains D and E share a sequence identity of 18% and they are related by a pseudo two-fold symmetry which relates the six corresponding helices of the two domains. The high B-factor loop of domain D is missing in domain E.

Based on the sequence and structural alignment, the catalytic site residues (His431 and Glu443; Asn481 and Ser410) between domains C and D in HAD and MFE-1DH structures are conserved. In the HAD structure, Glu443 modulates the catalytic properties of the catalytic histidine. The Asn-Ser pair interacts with the keto group of the substrate. There is also a water-mediated interaction from Asn481 of domain D to catalytic Glu443 in domain C.

5.4 The crystal structure of the human acyl-CoA binding protein (IV)

5.4.1 Structure quality

Crystal structures were determined for apo hACBP (at 1.6 Å resolution) and liganded hACBP (at 1.4 Å resolution) using the single wavelength anomalous diffraction (SAD) method. The apo structure contains one ACBP polypeptide chain (from residues Ser2 to Ile87) in the asymmetric unit with a Pb$^{2+}$ and a Zn$^{2+}$ ion. The apo structure was refined to $R = 19.3\%$ and $R_{\text{free}} = 23.0\%$. In the liganded form, two polypeptide chains are found per asymmetric unit comprising residues Ser2 to Ile87 for molecule A and residues Gln3 to Ile87 for molecule B. In the asymmetric unit there are four ligand molecules, 7 Zn$^{2+}$ ions and one sulfate ion. The complexed structure was refined to final $R = 17.0\%$ and $R_{\text{free}} = 19.3\%$. None of the three structures have high B-factor loops and there are no outliers in the Ramachandran plot.
5.4.2 The crystal structure of apo hACBP

Apo hACBP was crystallised in the presence of Pb\(^{2+}\) and Zn\(^{2+}\) ions. The structure resembles very much the same four-helix fold that is seen in bACBP (Kragelund et al. 1993) and myristoyl-CoA-PfACBP (van Aalten et al. 2001) crystal structures. When compared with the bovine crystal structures without a ligand (PDB IDs 1HB6 and 1HBK, (van Aalten et al. 2001)), the most significant difference is in the backbone of the loop connecting helices A2 and A3, which is more compactly associated with the core of the protein in the bovine structures. The single Pb\(^{2+}\) ion is bound in an interface between two crystallographic symmetry mates on the “back side” of the protein relative to the ligand binding site.

5.4.3 The crystal structure of myristoyl-CoA hACBP

The crystals of the hACBP in complexed form were obtained by co-crystallisation with ZnCl\(_2\) and myristoyl-CoA. The two molecules (molecules A and B) in the asymmetric unit are related by translational symmetry and both molecules are liganded. The folds of the two polypeptide chains are very similar with each other and also with the apo structure. Molecule A is located near a crystallographic three-fold axis, which generates a trimer of the three A molecules, which together with the three B molecules in the asymmetric unit results in a hexameric assembly. The three A molecules at the centre interact via a cluster of seven Zn\(^{2+}\) ions and ligand fragments. Molecules A and B interact both directly and via two Zn\(^{2+}\) ions (ions Zn4 and Zn5) and a myristoyl-CoA ligand linker.

The peaks in the anomalous difference density maps (9.7–28.7 σ from the 1.2818 Å dataset) correspond to the positions of the eight Zn\(^{2+}\) ions used in the initial phasing. Two of the unique crystallographically related Zn\(^{2+}\) ions together with the special Zn\(^{2+}\) ion on the three-fold axis create the central seven-Zn\(^{2+}\) cluster. This Zn\(^{2+}\) cluster is stabilised by a sulfate ion in its centre and by nine phosphate groups from the six interacting 3’-phosphate-AMP moieties (molecules D1 and P1 and their three crystallographically related copies). Each of the Zn\(^{2+}\) ions in the cluster is tetrahedrally coordinated to phosphate oxygen atoms of the ligand. The distances observed in this complex between the zinc and oxygen atoms agree with the published data (Harding 2004).

In the hexameric complex of myristoyl-CoA-hACBP, only one complete myristoyl-CoA molecule was seen (C1) between molecules A and B (Fig. 8). This molecule is bound by its CoA end to molecule B and by its fatty acyl chain part to molecule A. To confirm this dimeric mode of binding, the \(\lambda = 1.7453\) Å data set was used to calculate anomalous difference maps as described above. These maps showed peaks at positions corresponding to Zn, S and P atoms in the model. The highest anomalous peak (8.0 σ) was observed at the centre of the Zn\(^{2+}\) cluster, supporting the presence of the modelled sulfate ion. Another remarkable peak (4.0 σ) in this map was located at the position corresponding to the thioester sulphur of the full length myristoyl-CoA molecule (C1), confirming the extended mode of binding of this myristoyl-CoA molecule across monomers A and B.
Fig. 8. The structure of the myristoyl-CoA-hACBP complex shown as a stereo pair. Chain A is shown in dark grey and chain B in light grey ribbon. The ligand molecules with their ω-ends are indicated with arrows. The residues involved in ligand and dimer contacts are labelled with the corresponding residue number. The two hydrogen bonding interactions between chain A and chain B are highlighted with dotted lines (between Asn60 (A) and Thr36 (B) as well as between Glu61 (A) and Tyr32 (B)). The CoA end of the ligand C1 is bound to the polypeptide B binding pocket and its acyl chain resides in the ligand binding groove of polypeptide A. The 3'-phosphate AMP moiety (bound to molecule B) has additional hydrogen bonds from molecule A residues Arg44 and Lys53 to the phosphate moieties, and from Asp57 to the ribose moiety. The figure was prepared with PyMOL (DeLano 2002).

The hACBP dimer is shown in Fig. 8. The CoA part of ligand C1 has its adenine ring bound as it is in known structures of bovine and P. falciparum ACBP. The ring stacks on residue Tyr32 and is hydrogen bonded to Arg14 and Tyr74. The hydrophobic surface of the adenine ring is covered by ligand fragment C2. The 3’-phosphate of the C1 ligand hydrogen bonds to the conserved Tyr29, Tyr32 and Lys55 residues. In addition to these interactions, which are present in all known ACBP structures, in the myristoyl-CoA-hACBP structure there are additional hydrogen bonds from monomer A to the CoA part of ligand C1: from Arg44 to 3’-phosphate, from Lys53 to 5’-phosphate, and from Asp57 to the ribose hydroxyl group.
6 Discussion

6.1 Domain organisation in rat peroxisomal MFE-1 (I)

The monomeric nature of peroxisomal MFE-1 is a unique feature in the hydratase/isomerase superfamily. Other members of this family are either homotrimers or homohexamers having the subunits assembled as trimeric disc-like structures that have strong interactions with the adjacent subunits. Additionally, these multimeric proteins are not functional as individual monomers. (Engel et al. 1996, Modis et al. 1998, Mursula et al. 2001.)

To study the role of the different parts of MFE-1, the protein was dissected into five regions (A, B, C, D, and E, as shown in Figure 7) by structure based sequence alignment with related proteins. The sequence alignment showed that domain A (residues 1–190) exhibits similarity with the hydratase/isomerase superfamily. In MFE-1, a sequence of 89 residues separates domain A from the region similar to the NAD⁺-binding domain of HAD (domain C). This domain, domain B, is not related by sequence to the trimerisation domain of other multimeric hydratase/isomerase proteins, while its C-terminus aligns well with the so-called helix-10 of the hydratase/isomerase superfamily. Deletion experiments with MFE-1 showed that a variant consisting of domains A and B (domains C, D and E deleted) and resembling the mitochondrial, monofunctional hydratases, did not have any hydratase-1 activity. However, the variant containing domains A, B, D, and E was fully active (Figure 7). Stability assays using CD spectroscopy indicated that this variant was not as stable as the wild type MFE-1, but the composition of secondary structure elements remained unchanged, which suggests a similar fold.

The differences in the kinetic parameters suggest that the deletion of domain C induces changes in the catalytic site resulting in a preference for hydratase-1 activity over isomerase activity. The two active site residues of the MFE-1 hydratase-1 involve residues Glu103 and Glu123 (Glu144 and Glu164 in ECH). The proposed active site residue for the isomerase active centre is Glu123. Site-directed mutagenesis experiments with MFE-1 confirmed the essential role of these residues for the activity of hydratase-1 and isomerase reactions as hydratase-1 and isomerase activities of Glu103Ala and Glu123Ala variants were reduced 2000-fold and 100-fold, respectively.
From the structural similarity of domain A with the enzymes of the hydratase/isomerase superfamily, it appears that domain A must be accompanied by interactions from the other domains to yield a functional enzyme. Intramolecular sequence similarity between domains D and E, and the fact that domain D is the dimerisation domain of the HAD enzyme suggest that domain D and E of MFE-1 form a compact structure with each other and that E and D subdomains have a similar fold. A similar C-terminal extension showing sequence similarity to the (3S)-hydroxyacyl-CoA dehydrogenase dimerisation domain is also present in the mammalian mitochondrial trifunctional enzyme (Uchida et al. 1992) and in the E. coli fatty acid metabolising multienzyme complex (DiRusso 1990, Ishikawa et al. 2004). Furthermore, deletion studies showed that the D/E-domains are required for full hydratase activity.

6.2 Crystallisation of the dehydrogenase part of rat MFE-1 (II)

The C-terminal part of MFE-1 (MFE1-DH) has 35% sequence identity with the monofunctional mitochondrial HAD enzymes in mitochondria (Barycki et al. 2000). The expressed variant is active as a monomer whereas the other monofunctional HADs are dimers (Noyes & Bradshaw 1973, Osumi & Hashimoto 1980, Barycki et al. 1999). To obtain a stable crystallisable construct, it was essential to reduce the mobility of the N-terminus by discarding the His6 tag initially used and to include all the residues at the C-terminus. The N-terminus was also trimmed to match the N-terminus of the reported degradation product of MFE-1 (Dieuaide-Noubhani et al. 1996) which was shown to yield a variant with similar enzymatic activity as reported for the wild type enzyme and which could be used to obtain protein crystals suitable for structure determination. The catalytic activity of rat MFE-1 is lower than that of rat HAD, but both of these enzymes express broad substrate chain length specificity and accept fatty acyl groups from C4 to C16. (Osumi & Hashimoto 1980.) Additionally, bile acid synthesis intermediates and dicarboxylic acids are acceptable substrates of MFE-1 (Osumi & Hashimoto 1980, Reddy & Hashimoto 2001, Ferdinandusse et al. 2004). For the dehydrogenase activity, the (24S-hydroxy, 25R-methyl)-bile acid intermediate is the only valid substrate (Kurosawa et al. 2001).

6.3 Crystal structure of the rat MFE-1 dehydrogenase (III)

The crystal structure of the dehydrogenase part of MFE-1 allowed making a very informative structure-based sequence alignment with αFOM and human mitochondrial 2-enoyl-CoA isomerase (hmECI), yeast peroxisomal enoyl-CoA isomerase (ypECI) and rat mitochondrial enoyl-CoA hydratase (rmHYD) of the hydratase/isomerase superfamily. While the crystallised MFE-1 variant does not contain all of the MFE-1 residues, the structure nevertheless provides insight into the structure of the full length MFE-1 as will be briefly discussed here.

The hydratase/isomerase fold is characterised by a parallel 4-stranded β-sheet (sheet B) facing the solvent and providing the contact surface for the CoA moiety (Holden et al. 2001).
Highly conserved sequence patterns in the loops following the β-strands of this sheet have been correlated with functional properties related to CoA binding and catalysis. These conserved loop regions are also present in the MFE-1 sequence. The findings from this alignment suggest that the hydratase/isomerase domain of MFE-1 is more similar with hmECI (21% sequence identity) and rmHYD (30% sequence identity) than with ypECI (13% sequence identity). In addition, the catalytic residues of hmECI and rmHYD are conserved in MFE-1 (Glu103 and Glu123). Glu103 and Glu123 are key residues in the conserved sequence patterns of loop-3 and loop-4. In the hmECI structure, the CoA moiety is bound to β-sheet B such that the adenine ring interacts with loop-1 (Partanen et al. 2004) and the thioester oxygen of the fatty acyl-CoA is hydrogen bonded to the peptide NH-groups in the loops following β-strand B2 and β-strand B3. These NH-groups are provided by the second residue in the conserved GxD tripeptide (loop-2, in MFE-1 this is sequence Gly60–Ala61–Asp62) and the middle residue of a conserved glycine rich region after β-strand B3 (loop-3, in MFE-1 this is Gly99–Gly100–Gly101) (Partanen et al. 2004). The presence of these sequence signatures of loop-1, loop-2 and loop-3, as well as the conserved catalytic residues of loop-3 and loop-4 suggests that in the MFE-1 hydratase/isomerase CoA binds in a similar way to that generally seen in the hydratase/isomerase superfamily.

The sequence alignment also reveals a large insertion between helices 8 and 10 in domain A of MFE-1 and αFOM with respect to the classical hydratase/isomerase fold. In hmECI, helix-10 is preceded by helices helix-7, helix-8, and helix-9 such that these helices form a linear array of helices around the core of the subunit (Fig. 6A). This feature is also present in the hydratase/isomerase domain of the αFOM enzyme, except that two additional helices have been inserted between helix-8 and helix-9, referred to as helices 9a and 9b, whereas helix-9c of αFOM corresponds to helix-9 of the hydratase/isomerase fold (Fig. 6C, Fig. 9). Helix-9c then continues into the αFOM linker helix, which corresponds to the linker-1 helix of MFE1-DH and to helix-10 of enoyl-CoA isomerase.

The structure of crystallised MFE1-DH begins with an N-terminal region that is folded as an α-helix extending from near the E domain of the C-terminal lobe to the Rossmann fold domain. The structure-based sequence analysis of MFE-1 with αFOM, hmHAD and pig mitochondrial HAD showed that the linker-1 helix in MFE-1 corresponds to helix-10 of the hydratase/isomerase domain (Fig. 9). Conserved sequence patterns of helix-10 important for interaction with the substrate include residues that are always basic, and which interact with the phosphate moieties of the CoA, as discussed in section 2.3.2 and also shown in Fig. 6. (Mursula et al. 2004.) Consequently, it can be suggested that residues Lys275 and Lys279 of the linker-1 helix of MFE-1 will interact with the substrate in the active site of the MFE-1 hydratase/isomerase domain (domain A). This predicted mode of binding of the CoA moiety, involving interaction with the side chains of the linker-1 helix as well as with the residues of sheet B of the hydratase/isomerase A domain, would require movement of the hydratase/isomerase A domain towards the linker-1 helix and the remaining C-terminal dehydrogenase lobes (Fig. 9). Interestingly, helix-10 of hmECH (corresponding to the MFE-1 linker helix) is known to become disordered and move further away from the core of the protein when the ligand is not bound in the active site interacting with the linker helix residues (Partanen et al. 2004). This observation, together with the common sequence patterns of the linker helices
between hmEC1 and MFE-1 further support the proposed ligand binding mode in the MFE-1 H/I active site and the proposed concerted movement of the subdomains that occur when the active site becomes occupied by the substrate. The structure analysis also confirms that the linker-1 helix is stabilised by interactions from the D and E domains, which explains the previous observation that full activity of the MFE-1 variant was obtained with a variant including residues 1–287 and the D and E domain residues (paper I, Fig. 7).

Fig. 9. Stereo picture showing the comparison of the MFE1-DH structure (dark grey) with human mitochondrial hydratase 1 with a ligand (medium grey, PDB ID 1SG4, (Partanen et al. 2004)) and bacterial fatty acid metabolising complex α-subunit (light grey, PDB ID 1WDK, (Ishikawa et al. 2004)). The features discussed in the text are indicated with labels and arrows, N and C termini of MFE1-DH are labelled. The arrow below linker helix-1 illustrates the two-fold symmetry axis relating domains D and E. The figure was prepared with PyMOL (DeLano 2002).

The sequence identity between MFE1-DH and human HAD is 33%. The NAD+-binding domain (domain C) is fully present in the MFE1-DH structure and it has adopted the classical Rossmann fold (Wierenga et al. 1986). In both structures, helix CH2 is much longer than usually seen among the Rossmann fold structures and the parallel 6-stranded β-sheet (β-strands CB1 to CB6) is extended by two antiparallel β-strands (CB7 and CB8) at the C-terminus of this subdomain. Crystallographic binding studies with the HAD enzyme have shown that the longer α-helix CH2 provides some of the binding contacts for the fatty acid CoA substrate. In HAD, the extended helix is connected via a second
helix back to the core of the protein while in the MFE1-DH structure this region of the polypeptide chain is folded as an extended high B-factor loop suggesting increased mobility. HAD is known to adopt an open, unliganded conformation in the absence of the 3-hydroxyacyl-CoA substrate and a closed, liganded conformation when both the cofactor (NAD\(^+\)) and the product (3-acetoacetyl-CoA) are bound. Comparison with HAD revealed that the lobes of the crystallised MFE1-DH resemble the unliganded HAD structure. In the liganded HAD structure, the C-terminal lobe (corresponding to domain D in MFE1-DH) rotates towards the NAD\(^+\) binding site of the N-terminal lobe as compared to the apo structure. The hinge region has been identified between the lobes at the N-terminus of the first helix of the D domain in both HAD and MFE-1DH. (Barycki et al. 2000.)

MFE-1 is active as a monomer but the HAD enzyme is a tight dimer with the substrate molecule bound partly in the dimer interface. The D domain of MFE1-DH corresponds to the HAD dimerisation domain. This dimerisation domain is transformed by the HAD dimer twofold axis into the corresponding domain of the second subunit of the HAD dimer. In MFE1-DH, a local two-fold axis (Fig. 9), corresponding to the HAD dimer twofold axis, relates domain D to domain E, and therefore domain E prevents the MFE-1 monomer from dimerising like the HAD dimer. Domains D and E have six corresponding \(\alpha\)-helices, of which the first three \(\alpha\)-helices form the interface between the domains. Domain D (approximately 130 residues) is connected to domain E (approximately 110 residues) by a short loop region of residues 607–614. Domain D is larger than domain E because the high B-factor loop of domain D (after helix DH4) is missing in domain E and because in domain D, the region between helix DH5 and helix DH6 is longer. This region adopts a three-stranded antiparallel \(\beta\)-meander fold which is located at the site of the C-terminus of the HAD enzyme in the overlaid structures.

The \(\alpha\) chain of the bacterial fatty acid metabolising multienzyme complex (\(\alpha\)FOM) is the best structurally characterised MFE-1 orthologue. The linker helix of the \(\alpha\)FOM structure corresponds to the N-terminal helix of MFE1-DH. In MFE1-DH this helix (19 residues long vs. 31 residues in \(\alpha\)FOM) is kinked whereas in \(\alpha\)FOM it is evenly curved. The MFE1-DH helix-1 starts from the N-terminus of the construct (residue 261) and continues till residue 279. This linker-1 helix is stabilised by hydrogen bonding and van der Waals contacts with loop residues 657–659. This protruding loop, between the helices EH1 and EH2, is also present in the \(\alpha\)FOM structure, but not in the D domain.

The sequence identity between the MFE1-DH and the \(\alpha\)FOM dehydrogenase part is 28% and the structural comparison shows a good alignment for both the N-terminal lobe (domain C) as well as for the C-terminal lobe (domains D and E) of the dehydrogenase part. The position of the protruding CH2 Rossmann fold helix with respect to the bulk of the protein is different for MFE1-DH and \(\alpha\)FOM. An interesting structural difference is located after the protruding CH2 helix, as in \(\alpha\)FOM this helix is connected back to the bulk of the protein via another helix, whereas in MFE1-DH the returning helix is replaced with a loop. In \(\alpha\)FOM this helical pair is the dimerisation motif of the \(\alpha_2\)-dimer of the bacterial \(\alpha_2\beta_2\) tetramer and in the HAD structure the helix makes contacts with the CoA moiety of the substrate. At the moment the mode of binding of the CoA substrate to MFE-1 is not known.

An important structural difference between MFE1-DH and \(\alpha\)FOM involves the high B-factor loop after helix-4 (helix DH4) of the D domain of MFE1-DH (residues 537–
545), which is absent in αFOM and HAD. Helix DH4 in MFE1-DH is pointing away from the bulk of the protein, different from what is seen in αFOM and HAD. For αFOM it is proposed that a narrow hydrophobic tunnel between helix DH4 and helix DH1 is the binding cavity for the extended linear alkyl chain. The substrates of MFE-1 also include 3-hydroxyacyl chains with a bulky steroid moiety and for binding such bulky groups, much more space would be required. The movement of helix DH4 and the wider hydrophobic pocket correlate with the substrate specificity differences of MFE1-DH with αFOM and HAD.

Another functional difference between domain E of MFE1-DH and αFOM is that in MFE1-DH, the longer C-terminal tail presents the PTS-1 signal peptide for interaction with the peroxisomal import machinery. The other functions of domain E, such as stabilising the conformation of the linker-1 helix, and possible involvement in binding the substrate CoA moiety are apparently conserved among MFE-1, αFOM and HAD. The latter functionality of domain E can be deduced from the mode of binding of CoA to HAD, where the CoA binding pocket is completed by the corresponding dimerisation domain of the second monomer.

6.4 Crystal structure of human ACBP (IV)

The structure of human ACBP adopts the typical four-helix bundle of the ACBP family. The apo structure fold is essentially identical with the four helix bundle seen in the bACBP and myristoyl-CoA-PfACBP crystal structures. The major findings of the present structural studies on human ACBP will be briefly discussed here.

Compared with the bovine crystal structures without a ligand, (PDB IDs 1HB6 and 1HBK (van Aalten et al. 2001)), the most significant difference concerns the loop between helices A2 and A3. In both apo bACBP structures this loop is more compactly associated with the core of the protein. In hACBP the loop is more extended, which is also seen in the case of the PfACBP structure. Apo hACBP was crystallised in the presence of heavy metal ions Pb²⁺ and Zn²⁺. The Pb²⁺ ion is bound at an interface between two crystallographic symmetry mates. The terminal carboxyl group of residue Asp69 coordinates Pb²⁺ together with one ordered water molecule and a chloride ion. Glu61 completes hydrogen bonds on the other side of the Pb²⁺ ion with another water molecule from the neighbouring asymmetric unit. Human ACBP has been reported to bind lead in vivo in the kidneys of humans as a result of chronic low level exposure to lead (Smith et al. 1998). In the apo hACBP structure, the acyl-CoA binding pocket contains only solvent H₂O molecules and the Pb²⁺ binding occurs on the “back side” of the protein without apparent conformational changes in the residues involved in the ligand binding.

The liganded structure of hACBP contains two liganded protein molecules (molecules A and B) in the asymmetric unit, related by translational symmetry. The polypeptide fold in both of the molecules is very similar compared to each other and to the apo structure, contrary to the previous observations reporting relatively large ligand binding induced conformational change in the tertiary structure of liver ACBP (Frolov & Schroeder 1998). Molecule A is located near a crystallographic three-fold axis, which generates a trimer of three A molecules. Together with the three B molecules in the asymmetric unit the
polypeptides assemble as a hexamer. The three A molecules at the centre of this hexamer do not have direct contacts with each other. Instead, a cluster of seven Zn$^{2+}$ ions and ligand fragments stabilise the trimeric assembly of the dimers in each asymmetric unit. Molecules A and B interact both directly and via Zn$^{2+}$ ions (Zn4 and Zn5) and the myristoyl-CoA ligand linker C1.

Compared with the bovine NMR structure with bound palmitoyl-CoA, the most significant differences with the myristoyl-CoA complex of hACBP concern the loop between helices A2 and A3. In bACBP, this loop together with helix A3 is closer to the core of the protein forming a more compact structure, whereas in the hACBP-myristoyl-CoA complex these structural elements are further away from the ligand binding site. Comparison with the liganded structure of PfACBP also reveals that the backbone trace is nearly identical with that of hACBP-myristoyl-CoA complex. In PfACBP there is an insertion of two residues in the C-terminus of helix A1 introducing changes that extend to the following loop and change the binding specificity of PfACBP towards shorter chain acyl-CoAs by blocking the acyl chain binding groove from longer hydrocarbon chains (van Aalten et al. 2001).

Binding of the acyl-CoA molecule to ACBP comprises essential interactions with two parts of the ligand: the 3'-phosphate AMP moiety of CoA and the ω-end of the acyl chain. In the monomeric CoA binding mode seen in bACBP and PfACBP, the adenine ring is buried deep in the core of the protein and stacked between Tyr32 and the acyl chain of the ligand (also illustrated in Fig. 4). The extended pantetheine unit has only few contacts with the protein and it turns back towards the acyl binding groove starting from the loop between helices A1 and A2. Finally, the acyl chain continues towards the N-termini of helices A2 and A3 in the hydrophobic groove. In the dimeric binding mode of the hACBP-myristoyl-CoA complex, the hydrophobic interactions and hydrogen bonds observed in the monomeric binding mode are also present, and the 3'-phosphate AMP moiety of ligand C1 is essentially bound as in the monomeric binding mode. However, additional interactions from the other A molecule affecting the ligand binding in molecule B are seen, concerning the 3'-phosphate AMP moiety. The 3'-phosphate interactions of the CoA account for 40% of the total binding energy (Faergeman et al. 1996) and they involve hydrogen bonding contacts with residues Lys33, Lys53 and Tyr74 of the bACBP (Kragelund et al. 1999b). In the hACBP-myristoyl-CoA structure, there are additional hydrogen bonds from monomer A (near the end of the hydrophobic acyl chain binding groove between helices A2 and A3) to the CoA part of ligand C1: from Arg44 to 3'-phosphate, from Lys53 to 5'-phosphate, and from Asp57 to ribose oxygen. In the liver form of ACBP, these residues are well conserved with the single exception of Lys53 to Arg. This conserved set of charged residues seems to provide an alternative binding surface for the CoA and may keep the CoA end of the ligand exposed so that the monomer B can bind it.

In molecule B of the myristoyl-CoA-hACBP complex, the other hydrophobic counterpart of the adenine ring of ligand C1 comes from another ligand molecule, C2. In apo hACBP, residue Arg14 points away from the adenine ring binding pocket and leaves the pocket open. It moves inwards to form a hydrogen bond with the ligand adenine ring in both monomeric and dimeric modes of the complexed structures. Additional changes involving residues Lys19 and Met25 are induced when a ligand is bound in the dimeric mode. These movements consequently evoke a constriction in the acyl binding groove in
the presence of the $\omega$-end of the fatty acyl chain (C1 bound to molecule A), while in the presence of the pantetheine part (C2 bound to molecule B) the groove is wider. In the myristoyl-CoA-hACBP complex dimeric binding mode, the myristoyl-CoA (molecule C1) is in reverse orientation with respect to molecule A and the reported monomeric binding mode. Ligand C1 is bound such that its $\omega$-end is between residues Met25 and Arg14, whereas in the monomeric mode this portion of the groove is occupied by the mid section of the acyl chain instead (and as is also seen for molecule B and ligand C2).

The previously published structures of ACBP with ligands show that the binding occurs in 1:1 (protein:ligand) stoichiometry. In the present work, however, a dimeric binding mode involving two protein molecules per one fully bound ligand was also observed. The ability of ACBP to donate its ligand directly to receiving proteins makes it intriguing to propose a model for the mechanism of the ligand-linked dimer of hACBP based on the dimeric binding mode. In one plausible model, the myristoyl-CoA initially binds with its 3'-phosphate AMP part as seen in molecule A and with its fatty acyl moiety pointing into solution. Subsequently, another fatty acyl binds in the fatty acyl binding pocket as seen in molecule A, and its CoA head group binds to molecule B; this mode is preferred because of the additional contacts from molecule A to the 3'-phosphate AMP moiety, resulting in even stronger binding forces than in the monomeric binding mode. Currently, there is no further structural data from ACBP supporting the observed dimeric binding mode but biochemical analyses have shown that the acyl-CoA-ACBP complex is the preferred substrate of various lipid processing enzymes (discussed in section 2.2.1.3) and direct lipid transfer obviously requires an intermediate state, in which the ligand is bound both by ACBP and the receiving enzyme. Additionally, atomic force microscopy studies with synthetic phospholipid bilayer membrane containing C$_{12}$ fatty acid CoA and ACBP have shown that ACBP is able to bind and desorb membrane-bound acyl-CoA (Cohen Simonsen et al. 2003). In this model, the fatty acyl-CoA has its fatty acyl chain initially buried in the outer membrane leaf and the CoA end is exposed to the bulk solvent. Binding of the exposed CoA end to the ACBP begins the desorption leaving the acyl chain binding groove of ACBP empty. The binding interactions between ACBP and the 3'-phosphate were shown to be strong enough to keep the ACBP molecule associated on the membrane by being anchored to the CoA moiety. (Cohen Simonsen et al. 2003.) Our ACBP structure with myristoyl-CoA correlates well with this proposed desorption mechanism showing that the flexibility of the acyl-CoA molecule allows different ligand binding modes in ACBP.
7 Conclusions

In this study two intracellular proteins involved in fatty acid processing were studied using protein crystallographic methods. The aim was to increase understanding of the acyl-CoA-protein interactions in rat peroxisomal MFE-1 and human liver ACBP.

Initially five functional domains (domain A, B, C, D and E) were assigned to MFE-1 based on sequence analysis and enzyme activity measurements. From the structural studies it became clear that domain A and domain B together form the 2-enoyl-CoA hydratase 1$\Delta^3$-$\Delta^2$-enoyl-CoA isomerase part of MFE-1. The isomerase part is directly linked to the Rossmann fold (domain C) of the bilobal (3$S$)-hydroxyacyl-CoA dehydrogenase part, and domains D and E complete the dehydrogenase part. Domains D and E were both shown to have sequence similarity with the dimerisation domain of the mitochondrial HAD enzyme. Interactions from domains D and E were essential for the hydratase activity of domain A/B, but the presence of domain C was not required. The structure of MFE1-DH confirms the domain duplication of domain D into E observed initially in the sequence analysis studies, and the structural analysis reveals a pseudo 2-fold symmetry between the $\alpha$-helical elements of these two subdomains. Furthermore, the protruding loop of domain E was shown to interact with the connecting helix of domain B, supporting the role of domain E as a stabilising counterpart for the hydratase/isomerase of domain A/B. The monomeric nature of MFE-1, unlike other $\beta$-oxidation multifunctional enzymes, was shown to be caused by absence of dimerisation motifs seen in closely related proteins, like the $\alpha$-subunit of the bacterial FOM complex (helix-turn-loop instead of helix-turn-helix in the Rossmann fold) and human mitochondrial HAD (duplication of domain D resulting in a D-E pair instead of dimerisation via a D-D pair). Additionally, a large hydrophobic groove was located in the putative acyl chain binding site of the protein supporting the observations that MFE-1 may also be involved in processing bile acid intermediates with a bulky acyl moiety.

The apo and liganded forms of human liver ACBP were crystallised in the presence of heavy metal cations and their structures were determined by X-ray crystallographic methods. The apo structure was shown to adopt the classical four-helix bundle fold typical for the ACBP family. The liganded structure was determined in the presence of Zn$^{2+}$ ions and myristoyl-CoA. The asymmetric unit of the liganded crystals contains two ACBP molecules and one complete myristoyl-CoA molecule, plus additional ligand
fragments. Each of the polypeptides has a 3’-phosphate-AMP moiety and a fatty acyl moiety in the binding pocket, but these ligands originate from distinct myristoyl-CoA molecules. The present study shows that the ACBP molecule is able to bind two different myristoyl-CoA molecules at the same time. Furthermore, in the novel dimeric binding mode the 3’-phosphate-AMP part of the completely ordered ligand binds on one ACBP molecule as in the monomeric binding mode, but the fatty acid moiety of the same ligand binds in the corresponding pocket of an adjacent molecule in reverse orientation compared to the one reported previously.

For some enzymes, the ACBP-fatty acyl-CoA complex is preferred as a substrate over free fatty acyl-CoA. For these enzymes the fatty acyl-CoA molecule has to be transferred from the ACBP binding pocket to the active site in the receiving enzyme. The novel binding mode of complexed hACBP suggests a model for transferring a protein-bound fatty acid CoA molecule. The dimeric binding mode of the present study showed that hACBP is able to bind either the CoA end, or the acyl chain end of the ligand molecule, leaving the other end accessible for other interactions and mimicking the ligand transfer complex formed by liganded ACBP and the acyl-CoA receiving enzyme.
References


Original articles


Permission to reprint the original articles has been obtained from the following copyright owners:


II International Union of Crystallography (http://journals.iucr.org)

III Elsevier.

Original articles are not included in the electronic version of the dissertation.

439. Tero, Niina (2005) Genetic structure at different spatial scales in metapopulations of Silene tatarica

440. Säkkinen, Hannele (2005) Variation in the blood chemical constituents of reindeer. Significance of season, nutrition and other extrinsic and intrinsic factors


443. Jäkäläniemi, Anne (2005) Adaptation, population viability and colonization-extinction dynamics of Silene tatarica in riparian habitats


447. Harjumaa, Lasse (2005) Improving the software inspection process with patterns


452. Autio, Jyrki (2006) Environmental factors controlling the position of the actual timberline and treeline on the fells of Finnish Lapland

Jukka Taskinen

PROTEIN CRYSTALLOGRAPHIC STUDIES OF COA-DEPENDENT PROTEINS: NEW INSIGHT INTO THE BINDING MODE AND EXCHANGE MECHANISM OF ACYL-COA