Tatu Haataja

PEROXISOMAL MULTIFUNCTIONAL ENZYME TYPE 2 (MFE-2)

THE CATALYTIC DOMAINS WORK AS INDEPENDENT UNITS
TATU HAATAJA

PEROXISOMAL MULTIFUNCTIONAL ENZYME TYPE 2 (MFE-2)
The catalytic domains work as independent units

Academic dissertation to be presented with the assent of the Faculty of Science of the University of Oulu for public defence in Kuusamonsali (Auditorium YB210), Linnanmaa, on 25 November 2011, at 12 noon

UNIVERSITY OF OULU, OULU 2011
Haataja, Tatu, Peroxisomal multifunctional enzyme type 2 (MFE-2). The catalytic domains work as independent units
University of Oulu, Faculty of Science, Department of Biochemistry, P.O. Box 3000, FI-90014
University of Oulu, Finland; University of Oulu, Biocenter Oulu, P.O. Box 5000, FI-90014
University of Oulu, Finland
Acta Univ. Oul. A 582, 2011
Oulu, Finland

Abstract

Lipids are necessary for living organisms and have various roles as energy sources, hormone precursors and as components of membrane structures. Fatty acid \(\beta\)-oxidation is a pathway of energy metabolism, in which the fatty acyl-CoAs are degraded in several steps. Peroxisomal \(\beta\)-oxidation systems are found in all eukaryotes studied thus far, but the existence of a mitochondrial system is established in mammals only.

Multifunctional enzyme type 2 (MFE-2) has been characterized from various species and is responsible for catalyzing the second and third steps in the R-specific peroxisomal \(\beta\)-oxidation pathway. MFE-2 accepts a wide range of substrates and displays great variation in domain organization and overall molecular mass. The crystal structures of individual domains of MFE-2 from several species have been determined previously.

In this study, the structural knowledge of MFE-2 is further extended from the domain level to the assembly of the full-length enzyme. The crystal structure of Drosophila melanogaster MFE-2 (DmMFE-2) was solved at 2.15 Å resolution. The enzyme is a homodimer with 3R-hydroxyacyl-CoA dehydrogenase and 2E-enoyl-CoA hydratase 2 activities residing on each of the polypeptide subunits. Kinetic data combined with information from the structure, suggest that the catalytic domains of DmMFE-2 work as separate entities. It also appears that the enzyme does not assemble into dimers in vitro when the catalytic subunits are introduced in a solution as stand-alone proteins. The data were confirmed by two different methods, static light scattering and small-angle X-ray scattering.

However, the primary use of SAXS was not to monitor the formation of dimers in solution, but instead it was used for structure determination of the human MFE-2. During the process, the structural information from DmMFE-2 was used as a scaffold for the human homolog. After collecting a wide range of SAXS data and numerous calculations, a plausible low resolution solution structure of human MFE-2 was obtained. The model reveals the overall assembly of the enzyme and the locations of the C-terminal SCP-2L domains (an unspecific lipid carrier), thus enabling further hypotheses regarding the possible role of the SCP-2L domain in the enzymatic reaction.

Keywords: multifunctional enzyme type 2, peroxisomes, small-angle X-ray scattering, X-ray crystallography, \(\beta\)-oxidation
Haataja, Tatu, Peroksisomaalinen monitoiminen entsyymi tyyppi 2 (MFE-2). Katalyyttiset domeenit toimivat itsenäisinä yksiköinä

Oulun yliopisto, Luonnontieteellinen tiedekunta, Biokemian laitos, PL 3000, 90014 Oulun yliopisto; Oulun yliopisto, Biocenter Oulu, PL 5000, 90014 Oulun yliopisto

*Acta Univ. Oul. A 582, 2011*

Oulu

**Tiivistelmä**

Lipidit eli rasva-aineet ovat välttämättömiä eliöille ja niillä on lukemattomia rooleja mm. energianlähteinä, kalvojen rakenteina ja hormonien esiasteina. Rasvahappojen hajotettu monivaliehissellä metabolireitillä, jota kutsutaan β-oksidaatiossa, peroksisomaalinen rasvojen hajotusreitti on löydetty kaikista tähän asti tutkituista mitotumaisista, mutta mitokondrioissa tapahtuva rasvojen β-oksidaatio on löydetty vain nisäkkäiltä.

Peroksisomaalinen monitoiminen entsyymi tyyppi 2 (MFE-2) katalysoi toisen ja kolmannen reaktion R-spekifisellä rasvahappojen hajotusreitillä ja se on karakterisoitu useilta eri lajeilta. MFE-2 muodostaa lajista riippuen riippuen hyvin erilaisia ja molekyylimassaltaan erikokoisia alayksikköhyökkäsmiä, jotka pystyvät katalysoimaan erityyppisten substraattien hapetuksen. Eri lajien MFE-2:n yksittäisten alayksiköiden kiderakenteet ovat olleet tunnettuja jo vuosia.


**Asiasanat:** monitoiminen entsyymi tyyppi 2, peroksisomit, röntgenkristallografia, röntgenpikkulakrosaattana, β-oksidaatio
Acknowledgements

This work was carried out at the Department of Biochemistry and Biocenter Oulu, University of Oulu during years 2005–2011.

I want to express my deepest gratitude to my supervisors, Docent Tuomo Glumoff and Professor Kalervo Hiltunen for giving me the opportunity to do my Ph.D. work in the group and for the excellent guidance and expertise throughout the project. I want to thank Dr. Kristian Koski for all the valuable expertise and support with protein crystallography, but also for being such an excellent and patient mentor. I would also like to thank all the professors and group leaders for creating an excellent research environment at the Department of Biochemistry.

I wish to thank Docent Tommi Kajander and Dr. Werner Schmitz for careful reviewing of the manuscript of the thesis and Dr. Deborah Kaska for the revision of the language. I am indebted to all my collaborators and co-authors for their crucial contributions. I am especially grateful to Dr. Antti Vasala and BioSilta, Dr. Dmitri Svergun and Dr. Clémént Blanchet from EMBL, Hamburg for fruitful collaboration during my project. I wish to address special thanks to Dr. Päivi Pirilä for skillful synthesis of the substrates, Dr. Ulrich Bermann and Eeva-Liisa Stefanius for their expertise in mass spectrometry and Ville Ratas for running the crystallization facility smoothly. In addition, I want to thank all the structural biologists in Oulu for creating and maintaining such a great working facility and also for all the help and support during these years.

From my colleagues, I would like to address special thanks to M.Sc. Maija Mehtälä not only for great company in the laboratory and office, but particularly for the fruitful discussions and collaboration over these years. I want to thank also Dr. Mari Ylianttila and Dr. Antti Haapalainen for all the practical assistance especially in the beginning of my project. Then I would like to thank Alex, Anne, Antti, Fumi, François, Juha, Katri, Laura, Leena, Maike, Mari, Marjut, Miia, Samuli, Silke, Tonja, Vasily and all the other present and former members of the TG- and KH-groups for the enjoyable atmosphere and nice discussions. I would like to extend my thanks to the staff of our Department: Ari-Pekka and Miki for setting up and maintaining the computer systems, Anneli, Jaakko, Jari H, Jari K, Kyösti, Maila, Maire, Pentti, Petra, Pia, Pirjo and Tuula for running everything smoothly. I am grateful to all the staff members for the enjoyable working atmosphere at the Department.

I thank my friends for all the support and for reminding me about the other important aspects of life.
My warmest thanks belong to my mother Sirpa and her life-companion Anse, to my father Seppo and my brother Jani for all their love and support. I am also grateful to my parents-in-law, MaijaRiitta and Timo for all the encouragement, support and the memorable days spent in the Kainuu area. I wish to thank Saana and Joni for enjoyable discussions and great company. Finally, I would like to thank the most important people of my life, Salla, Ilari and Ida, for their love, support, patience and all the moments we have shared.

The work was financially supported by the Academy of Finland, Biocenter Oulu and Sigrid Jusélius Foundation.

Oulu, October 2011

Tatu Haataja
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>C-</td>
<td>carboxy (terminal)-</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>C4</td>
<td>2E-butenoyl-CoA</td>
</tr>
<tr>
<td>C6</td>
<td>2E-hexenoyl-CoA</td>
</tr>
<tr>
<td>C10</td>
<td>2E-decenoyl-CoA</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DmMFE-2</td>
<td><em>Drosophila melanogaster</em> multifunctional enzyme type 2</td>
</tr>
<tr>
<td>DH</td>
<td>3R-hydroxyacyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>FADH₂</td>
<td>flavine adenine dinucleotide, reduced form</td>
</tr>
<tr>
<td>HsMFE-2</td>
<td><em>Homo sapiens</em> multifunctional enzyme type 2</td>
</tr>
<tr>
<td>H2</td>
<td>2E-Enoyl-CoA hydratase 2</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>MFE-1</td>
<td>peroxisomal multifunctional enzyme type 1</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>N-</td>
<td>amino (terminal)-</td>
</tr>
<tr>
<td>NAD⁺</td>
<td>nicotinamide adenine dinucleotide, oxidized form</td>
</tr>
<tr>
<td>NADH</td>
<td>nicotinamide adenine dinucleotide, reduced form</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>nickel-nitrilotriacetic acid</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>protein data bank</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PTS</td>
<td>peroxisomal targeting signal</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>SAXS</td>
<td>small-angle X-ray scattering</td>
</tr>
<tr>
<td>ScMFE-2</td>
<td><em>Saccharomyces cerevisiae</em> multifunctional enzyme type 2</td>
</tr>
<tr>
<td>SCP-2L</td>
<td>sterol carrier protein 2-like</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SDR</td>
<td>short-chain dehydrogenase/reductase</td>
</tr>
<tr>
<td>SLS</td>
<td>static light scattering</td>
</tr>
<tr>
<td>TFE</td>
<td>mitochondrial trifunctional enzyme</td>
</tr>
<tr>
<td>YPD</td>
<td>yeast extract peptone dextrose</td>
</tr>
<tr>
<td>Å</td>
<td>ångström, $10^{-10}$ m</td>
</tr>
</tbody>
</table>
List of original articles

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


* Equal contribution
## Contents

Abstract ................................................. 7

Tiivistelmä .................................................. 7

Acknowledgements 7

Abbreviations 9

List of original articles 11

1 Introduction ............................................. 15

2 Review of the literature ............................... 17

2.1 Peroxisomes ........................................... 17

2.1.1 Peroxisome biogenesis and proliferation .............. 18

2.1.2 Protein import into peroxisomes ......................... 18

2.1.3 Trafficking of metabolites across the peroxisomal membrane ................................................. 19

2.1.4 The function and interactive role of peroxisomes .......... 21

2.2 Fatty acid β-oxidation ................................... 22

2.2.1 Peroxisomal β-oxidation of fatty acids .................. 23

2.3 Multifunctional enzymes and enzyme complexes in β-oxidation ............. 25

2.3.1 Peroxisomal multifunctional enzyme type 1 ................ 26

2.3.2 Peroxisomal multifunctional enzyme type 2 ............... 28

2.3.3 Mitochondrial trifunctional enzyme (TFE) ................. 34

2.3.4 Bacterial β-oxidation complex ............................ 35

2.4 Substrate channelling in multifunctional enzymes and enzyme complexes ............................................. 37

2.4.1 Tunnel-forming enzymes .................................... 37

2.4.2 Eccentric ways of substrate channeling .................. 40

3 Aims of the study ....................................... 43

4 Materials and methods ............................... 45

4.1 Plasmids, strains and culture media ........................ 45

4.2 Production of recombinant proteins in *E.coli* (I-III) ........... 45

4.3 Protein purification (I-III) .................................... 46

4.4 Enzyme activity assays (II) ................................... 47

4.5 Crystallization of *Dm*MFE-2 (II) ............................. 47

4.6 Data collection and structure determination of *Dm*MFE-2 (II) ...... 48

4.7 Rescue experiments with *Dm*MFE-2 in a yeast mutant devoid of endogenous MFE-2 (II) ................................................. 48

4.8 Static light scattering (SLS) experiments (II) ............... 49
4.9 SAXS data collection, data refinement and construction of ab initio models of DmMFE-2 and HsMFE-2 (III) ......................... 49

5 Results 51
5.1 Purification of DmMFE-2 and the separate enzymatic domains .......... 51
5.2 Crystallization of DmMFE-2 ................................................................. 52
5.3 The crystal structure of DmMFE-2 .................................................. 52
  5.3.1 Structure determination of DmMFE-2 ........................................... 53
  5.3.2 The overall structure of DmMFE-2 monomer and dimer .......... 54
  5.3.3 3R-Hydroxyacyl-CoA dehydrogenase domain of DmMFE-2 ........ 56
  5.3.4 2E-Enoyl-CoA hydratase 2 domain of DmMFE-2 .................. 57
5.4 Kinetic properties of DmMFE-2 in vitro ........................................ 59
5.5 Rescue experiments in vivo using DmMFE-2 in a yeast mutant devoid of endogenous MFE-2 ........................................... 60
5.6 Investigation of heteromeric complex formation of separate domains of DmMFE-2 in vitro by static light scattering .......... 61
5.7 SAXS data collection and the low resolution models of fruit fly and human MFE-2 ................................................................. 62

6 Discussion 67
  6.1 The structure of DmMFE-2 reveals the assembly of the full-length enzyme ......................................................... 67
  6.2 The localization and kinetic studies of DmMFE-2 ......................... 69
  6.3 The current model of HsMFE-2 ...................................................... 70

7 Conclusions 73
References 75
Original articles 93
1 Introduction

Lipids are a structurally diverse group of hydrophobic molecules that are vital for living organisms where they serve as building blocks for membrane structures, energy sources for cells and precursors for hormones or intracellular messengers. Some lipids are called essential, since the molecules cannot be made via the major synthetic or degradation pathways in the body. β-Oxidation is a pathway of energy metabolism in which fatty acids are degraded by enzymes either in mitochondria or peroxisomes. In the latter, the second and third steps are catalyzed by two structurally non-related multifunctional enzymes, MFE-1 and MFE-2 that display opposite chiral specificity. NADH and FADH\textsubscript{2}, the intermediates of β-oxidation, are subsequently oxidized in the electron transport chain of mitochondria yielding a high-energy compound ATP, or in peroxisomes FADH\textsubscript{2} which donates electrons to oxygen to form H\textsubscript{2}O\textsubscript{2}.

Peroxisomal MFE-2 has been characterized from various species, including human, rat, fish, fruit fly and yeast. The enzymatic activities of MFE-2 from the N-terminal end include 3R-hydroxyacyl-CoA dehydrogenase and 2E-enoyl-CoA hydratase, respectively. In some species the C-terminus of the enzyme is extended by a non-specific sterol carrier protein 2-like domain of unknown biological function. The domain count and organization of various MFE-2s also display great variation depending on the species. The crystal structures of the separate domains of MFE-2 are known from previous studies. Mutations in the enzymatic domains of MFE-2 can cause dysfunction of the protein and consequently MFE-2 deficiency. Protein crystallography and structural enzymology have enabled the studies on structure-function relationships at atomic resolution that have increased our understanding of the mechanisms leading to the development of disease in the case of dysfunctional proteins in mammals.

In this study, the structural knowledge of MFE-2 is further expanded from the domain level to the domain assembly level by solving the crystal structure of fruit fly MFE-2 and the low resolution model of human MFE-2. The crystal structure of fruit fly MFE-2 reveals the interface between the two enzymatic domains. Quite surprisingly, the center of the biologically active dimer is a hole that appears to be too narrow to allow the bulky fatty acyl-CoA substrate to pass through. The combined results from kinetics experiments and structure analysis suggest that the dehydrogenase and hydratase domains work as separate entities and the efficiency of the enzyme catalysis does not benefit from substrate channeling.
2 Review of the literature

2.1 Peroxisomes

The family of microbodies is a collection of organelles that include peroxisomes, glyoxysomes (Graham 2008), glycosomes (Opperdoes & Szikora 2006) and Woronin bodies (Liu et al. 2008). Microbodies were discovered in 1954 by Rhodin and over a decade later these particles were identified as organelles with oxidase and catalase activities and renamed peroxisomes (De Duve & Baudhuin 1966). The highest number of peroxisomes (comprising 2% of the total protein in the organ) can be found in the liver (Leighton et al. 1968). The diameter of the spherical or oval peroxisome varies between 0.1 to 1 µm. For instance, peroxisomes isolated from the kidney or liver are larger (0.3–0.9 µm) than those isolated from the bowels (0.1–0.3 µm) (Islinger et al. 2010). One explanation for the differences in size could be the distinct roles of these tissues in the metabolism of lipids and xenobiotics. A single lipid bilayer membrane separates the proteinaceous peroxisomal matrix from the cytosol (van den Bosch et al. 1992). Unlike mitochondria, peroxisomes do not contain DNA and thus are solely dependent on the protein synthesis that takes place on free polyribosomes in the cytosol (Goldman & Blobel 1978, Lazarow & Fujiki 1985).

Quite remarkably, the number, morphology and enzyme content of peroxisomes vary significantly between different species and tissues (Islinger et al. 2010). Good examples of such phenomena can be found in rodents that display significant levels of peroxisome proliferation when fed polyunsaturated fatty acids (Flatmark et al. 1988). In addition, higher levels of β-oxidation activity can be monitored in hibernating rats compared to control animals housed in a common laboratory environment (Pollera et al. 1983). The ability of peroxisomes to adjust to changes of nutritional and physiological state, e.g. starvation and diabetes or other external signals, e.g. ROS (reactive oxygen species) or ethanol is fascinating (Issemann & Green 1990, Heinemann & Just 1992, Schrader et al. 1994, Erdmann & Blobel 1995, Titorenko et al. 2000, Islinger et al. 2010). As described in several reviews, the fact that peroxisomes are indispensable to human cells became evident after the discovery of a wide spectrum of lethal or severe diseases caused by problems in peroxisome biogenesis (Steinberg et al. 2006) or mutations in proteins targeted into peroxisomes (Wanders et al. 2010).
2.1.1 Peroxisome biogenesis and proliferation

The number of peroxisomes in a cell is dependent on multiple processes that include de novo biogenesis (Matsuzono et al. 1999, South & Gould 1999), degradation by autophagy (Yokota 1993) and division (Lazarow & Fujiki 1985). The major mechanism that regulates all these processes is unknown, but two classes of proteins, namely dynamin-related proteins and Pex11-type peroxisome proliferators are known to be involved (Thoms & Erdmann 2005). However, it has been demonstrated that lipids trigger peroxisome proliferation, which in mammals is mediated by the peroxisome proliferator-activated receptor α (PPARα) (Kliewer et al. 1992, Gurvitz & Rottensteiner 2006). The proliferation process in which the number and/or size of peroxisomes rapidly increases, can also be induced by drugs (Issemann & Green 1990).

Peroxisome biogenesis has been studied extensively. The process is mediated by a group of proteins called peroxins that are encoded by PEX genes (Distel et al. 1996). To date 32 peroxins have been identified that are required not only for biogenesis, but also for maintenance of functional peroxisomes (Heiland & Erdmann 2005, Kiel et al. 2006). The basic process of peroxisomal biogenesis consists of three main steps; membrane formation with the insertion of peroxisomal membrane proteins (PMPs), matrix protein import and the proliferation of peroxisomes (Eckert & Erdmann 2003, Fidaleo 2010). Pex3p, Pex19p and in some organisms also Pex16p are involved in membrane biogenesis (Höhfeld et al. 1991, Götte et al. 1998, South & Gould 1999). In de novo biogenesis, the membrane elements are derived from the endoplasmic reticulum (ER) (Hoepfner et al. 2005). The matrix proteins and PMPs, in turn, are imported using different partners; Pex5p and Pex7p recognize targets destined for the matrix (Brocard et al. 1994, Marzioch et al. 1994) while the majority of the PMPs are assisted by Pex19p. However, some PMPs are targeted into peroxisomes via the ER and are independent of Pex19p (Tabak et al. 2008).

2.1.2 Protein import into peroxisomes

The peroxisomal protein import mechanism has been proposed to have four main steps. These include the recognition of fully folded proteins by receptors in the cytosol, transport and subsequent receptor-cargo complex docking at the peroxisomal membrane, translocation of the complex towards the luminal side of the organelle, followed by cargo release and receptor recycling back into the
cytosol (Platta & Erdmann 2007, Rucktachsel et al. 2011). The precise sequence of events and the exact mechanisms involved are not well understood and are currently somewhat controversial (Girzalsky et al. 2010, Rucktachsel et al. 2011).

Proteins destined to peroxisomes are imported by diverse mechanisms. Although two peroxisomal targeting signals (PTS) exist, the carboxy terminal PTS1 sequence, (S/A/C)-(K/R/H)-(L/M), is the most predominant regardless of the species (Gould et al. 1989). Even gold particles could be targeted into peroxisomes when decorated with SKL (Walton et al. 1995). Later the PTS1 signal was redefined as a dodecapeptide (Brocard & Hartig 2006). The use of the PTS2 pathway is organism dependent and exploits the amino terminal targeting sequence, (R/K)-(L/V/I)-X_{5}-(H/Q)-(L/A/F) that is often cleaved off after the import (Osumi et al. 1991, Swinkels et al. 1991, Lazarow 2006). The PTS2 targeting pathway is absent in Caenorhabditis elegans (Motley et al. 2000). The targeting signals are recognized by different receptors: Pex5p interacts with PTS1 labeled proteins while Pex7p interacts with PTS2 decorated targets (Brocard et al. 1994, Marzioch et al. 1994). Some peroxisomal membrane proteins are labeled with an internal mPTS that contains a membrane anchor sequence and a Pex19p binding motif (Rottensteiner et al. 2004). Proteins lacking PTS can also enter peroxisomes as stowaways after oligomerizing with PTS containing proteins. The process is usually referred to as the “piggyback” mechanism (McNew & Goodman 1994). If the import of matrix proteins becomes blocked while integral PMPs continue to be imported normally, peroxisomal membrane “ghosts” are formed (Santos et al. 1988).

2.1.3 Trafficking of metabolites across the peroxisomal membrane

Translocation of various solutes across the peroxisomal membrane occurs through the action of non-selective channels that allow the passage of molecules smaller than 400 Da, and also of highly selective transporters for larger and bulkier subjects (Antonenkov et al. 2010). Quite recently the existence of a peroxisomal channel was reported in mouse, called Pxmp2, which works as a general pore for diffusion (Rokka et al. 2009). Similar kinds of pore forming proteins were also discovered in yeast peroxisomes allowing the traffic of small metabolites between the lumen and cytosol (Grunau et al. 2009). The transportation of fatty acids into the peroxisomes is more complex and differs from the mitochondrial system that utilizes carnitine esters and carnitine palmitoyl transferase I, carnitine acylcarnitine translocase and carnitine palmitoyl transferase II in order to transfer
acyl-goups from the cytosol into mitochondria for the mitochondrial β-oxidation (Bartlett & Eaton 2004, Wanders 2004). This is an unlikely scenario in peroxisomes, and instead, it has been proposed that fatty acids enter peroxisomes in a non-esterified form or as CoA esters (Visser et al. 2007b). The former import mechanism requires the assistance of acyl-CoA synthetases (Knoll et al. 1994). The latter molecules require the action of so-called half-transporters in order to enter peroxisomes. In mammals these include ABCD1/ALDP (Contreras et al. 1994, Mosser et al. 1994), ABCD2/ALDRP (Lombard-Platet et al. 1996) and ABCD3/PMP70 (Kamijo et al. 1990), which are members of the ATP-binding-cassette (ABC) transporter family (van Roermund et al. 2011). There are two half-ABC transporters in S. cerevisiae, namely Pxa1p and Pxa2p that are involved in the translocation of long-chain fatty acyl-CoAs into peroxisomes (Hettema et al. 1996). Complementation studies of the human homologs of ALDP and ALDRP with the yeast half-transporter deletion strains have demonstrated that indeed the human proteins participate in fatty-acyl-CoA import and even possess different substrate specificities. ALDP seems to be more active towards the straight chain fatty acyl esters, such as C24:0-CoA and C26:0-CoA, while ALDRP favors the unsaturated lipids, like C22:6-CoA and C24:0-CoA (van Roermund et al. 2011).

Reconstitution of peroxisomal membrane proteins in liposomes has been used to demonstrate the existence of a bile acid transport activity in peroxisomes (Visser et al. 2007a). However, the protein responsible for the activity is yet to be identified. The transportation of adenine nucleotides across the peroxisomal membrane is mediated by specific transporters that belong to the mitochondrial carrier family (MCF). These kinds of ATP transporters have been identified from human and yeasts (Sakai et al. 1996, Wylin et al. 1998, Palmieri et al. 2001, van Roermund et al. 2001). The protein characterized from S. cerevisiae, Ant1p (YPR128cp), appears to be functionally involved in peroxisome proliferation and in fatty acid β-oxidation of medium-chain fatty acids (van Roermund et al. 2001). In addition, it was determined that the protein has a high affinity towards the adenine nucleotides and works primarily as a transporter of cytosolic ATP into peroxisomes but also as an exchanger of AMP generated inside peroxisomes (Palmieri et al. 2001, van Roermund et al. 2001).

It has been suggested that there are transporters for smaller compounds as well, such as (pyro)phosphate or 2-ketoglutarate and isocitrate (Visser et al. 2005, Visser et al. 2006), but the nature of these transporters remains to be identified. Furthermore, certain ions can pass through non-selective channels, but specific
ion transporters are also known. In fruit fly, a Ca$^{2+}$/Mn$^{2+}$-ATPase of peroxisomal origin was found to alter calcium levels and therefore was suggested to take part in calcium transport (Southall et al. 2006).

2.1.4 The function and interactive role of peroxisomes

Knowledge of the function of this organelle has increased remarkably during the past two decades and now peroxisomes are clearly recognized as multifunctional. Peroxisomes house numerous enzymes that are required for β-oxidation and α-oxidation of fatty acids, cholesterol biosynthesis, metabolism of bile acids, biosynthesis of fatty acids, synthesis of plasmalogens, glyoxylate metabolism, degradation of purines and response to oxidative stress (Fidaleo 2010, Islinger et al. 2010).

The various oxidative reactions require a substantial amount of oxygen. Studies with rat suggest that 20% of the total oxygen consumption in liver is due to peroxisomes (De Duve & Baudhuin 1966). Reactive oxygen species (ROS) are formed as by-products of flavin-dependent oxidative reactions, which relates peroxisomes to ageing (Titorenko & Terlecky 2011). ROS-metabolizing enzymes, such as catalase (De Duve & Baudhuin 1966), superoxide dismutase (Keller et al. 1991) and epoxide hydrolase (Pahan et al. 1996) are synthesized in cells for scavenging ROS that include hydrogen peroxide (H$_2$O$_2$), superoxide anion (O$_2$-) and epoxides.

Some functions of peroxisomes are shared with other cell compartments. For example Fis1 has been suggested to have a role in fission both in mitochondria and peroxisomes (Koch et al. 2005). In addition, β-oxidation of long-chain fatty acids (Amery et al. 2000) is also performed by these two organelles as well as ketogenesis (Ashmarina et al. 1999). ROS-metabolizing enzymes including catalase and superoxide dismutase are known to be active in mitochondria as well (Radi et al. 1991, Bodyl & Mackiewicz 2007). Not surprisingly, mitochondrial abnormalities are detected if the peroxisomal metabolic machinery is impeded (Baumgart et al. 2001, Dirnx et al. 2005).

The interactive role with other organelles requires an efficient translocation of the metabolites through the peroxisomal membrane. It has been established that non-selective channels as well as highly selective transporters exist that enable the traffic of various metabolites through the organelle (Rokka et al. 2009, Antonenkov et al. 2010, van Roermund et al. 2011). Moreover, the discovery of vesicular traffic between mitochondria and peroxisomes that is mediated by a
novel mitochondrial outer-membrane protein (MAPL) has revealed an alternative mechanism for cargo transport between these organelles (Neuspiel et al. 2008).

2.2 Fatty acid β-oxidation

The degradation of fatty acids in the β-oxidation spiral is one example of metabolic compartmentalization. The spiral consists of four steps: dehydrogenation (FAD-dependent), hydration, dehydrogenation (NAD-dependent) and thiolytic cleavage. The reaction pathway is partially shared with peroxisomes and mitochondria in mammals, while most fungi and plants solely oxidize the fatty acid esters in peroxisomes (Poirier et al. 2006). Despite the identical degradation mechanism, the enzymes taking part in the oxidative reactions are different in these organelles (Osumi & Hashimoto 1984). In mammals, β-oxidation goes to completion in mitochondria, while in peroxisomes it is terminated at the medium-chain intermediates stage due to the substrate specificity of palmitoyl-CoA oxidase in the first step of the spiral (Osumi et al. 1980). Mitochondrial β-oxidation is capable of handling straight-chain substrates up to long-chain, whereas the peroxisomal system deals mostly with the very long-chain and branched-chain fatty acid esters (Mannaerts & van Veldhoven 1996).

The final outcome of β-oxidation differs in peroxisomes and mitochondria, at least in terms of ATP. Less energy is produced in peroxisomal β-oxidation simply because the organelle lacks the acetyl-CoA utilizing citric acid cycle and the respiratory chain that would directly use the electrons from NADH and FADH₂ for subsequent reactions (Schrader & Yoon 2007). In contrast to mitochondria, the electrons from the first dehydrogenation reaction in peroxisomes are donated straight to molecular oxygen, which results in H₂O₂ production and loss of energy as heat (Cooper & Beevers 1969). Finally, the β-oxidation products from peroxisomes need to be shuttled into mitochondria in which the final oxidation to carbon dioxide and water occurs. In mammals, the straight chain fatty acids become degraded up to short acyl-CoA chains (of 6-8 carbons) and acetyl-CoA. The former product can be further processed by a group of enzymes called ACOTS (Acyl-CoA thioesterases/hydrolases) in subsequent reactions to yield free fatty acids and free CoASH. Alternatively the acyl-CoA can serve as a substrate for carnitine octanoyl transferase (Farrell et al. 1984) that results in a carnitine-dependent translocation of the acyl-chains into mitochondria (Poirier et al. 2006). Acetyl-CoA is processed in a similar manner by carnitine acetyl transferase
(Miyazawa et al. 1983). In yeast, acetyl-CoA formed in the β-oxidation spiral can have two final destinations. It is either used as a substrate for citrate synthase that catalyzes the condensation of acetyl-CoA with oxaloacetate to form citrate or, similar to mammals, acetyl-CoA can serve as a substrate for carnitine acetyl transferase (Poirier et al. 2006).

In a wider perspective, peroxisomes and mitochondria play key roles not only in energy production from various lipids via oxidative reactions, but also in creating balanced lipid homeostasis in cells. In addition, these organelles contribute to thermogenesis at least in the brown adipose tissue (Nedergaard et al. 1980, Schrader & Yoon 2007).

2.2.1 Peroxisomal β-oxidation of fatty acids

Fatty acid β-oxidation that does not take place in mitochondria was first characterized in glyoxysomes from castor bean endosperm (Cooper & Beevers 1969). A similar system that could be enhanced by hypolipidemic drug treatment was found in rat liver peroxisomes (Lazarow & De Duve 1976). Peroxisomal β-oxidation machinery accepts a wide range of substrates, including very long-chain fatty acids (Singh et al. 1984), dicarboxylic acids (Cerdan et al. 1988, Leighton et al. 1989), di- and trihydroxycholestanolic acids (Pedersen & Gustafsson 1980, Kase et al. 1986), eicosanoids (Diczfalusy & Alexson 1988, Schepers et al. 1988, Shirley & Murphy 1990), pristanic acid (Singh et al. 1990) and some xenobiotics (Gatt et al. 1988, Suzuki et al. 1990). It has also been demonstrated that polyunsaturated fatty acids are oxidized more rapidly in peroxisomes (Hiltunen et al. 1986). A schematic presentation of β-oxidation in peroxisomes is presented in figure 1.

The fatty acids entering the peroxisomal β-oxidation spiral are activated to their CoA esters. This requires the action of acyl-CoA synthetases of which one is active towards the very long-chain fatty acids and is localized in peroxisomes and ER (Singh et al. 1987, Wanders et al. 1987, Singh & Poulos 1988). Besides peroxisomes, another synthetase utilizing the long-chain substrates is also found in mitochondria and ER (Shindo & Hashimoto 1978, Mannaerts et al. 1982). Most of the peroxisomal β-oxidation substrates, excluding the very long-chain fatty acids, become activated in the ER (Vamecq et al. 1985, Casteels et al. 1988, Prydz et al. 1988, Schepers et al. 1989). Since peroxisomal β-oxidation mostly utilizes substrates that are not degraded by the mitochondrial system, impairment
of peroxisomal fatty acid oxidation causes toxic effects due to the accumulation of the hydrophobic substrates in the body (Fidaleo 2010).

Fig. 1. Peroxisomal β-oxidation pathway comprising of four steps. The electrons from the first dehydrogenation reaction are transferred to molecular oxygen for H₂O₂ production. Multifunctional enzymes, namely multifunctional enzymes type 1 or 2 (MFE-1 or MFE-2) catalyze the second and third reactions (hydration and subsequent dehydrogenation) of the spiral. In the fourth step, the substrate becomes shortened by two carbon atoms after thiolytic cleavage. Finally, the chain-shortened acyl-CoA enters another cycle of β-oxidation or is exported outside the peroxisome.
Some of the substrates that undergo β-oxidation have structures that are not initially recognized by the enzymes catalyzing the oxidative reactions in the spiral. Alternative mechanisms are used to modify the structures. For instance the 3-methyl group of phytic acid blocks β-oxidation and therefore α-oxidation is required to transform the substrate into a 2-methyl fatty acid (Avigan et al. 1966, Tsai et al. 1969). It seems that in higher eukaryotes the process is solely carried out in peroxisomes (Croes et al. 1996). In addition, the utilization of substrates with trans-bonds at odd-numbered positions or cis-bonds in general, requires the assistance from an additional set of enzymes. A group of auxiliary enzymes has been demonstrated to be involved in the process (Poirier et al. 2006). At least 2,4-dienoyl-CoA reductase, Δ¹,Δ²-enoyl-CoA isomerase and Δ³,Δ⁵,Δ²,4-dienoyl-CoA isomerase are known to modify their substrates so that they can be completely degraded in the β-oxidation spiral (Kunau et al. 1995, Novikov et al. 1999, Hiltunen et al. 2003).

The short-term regulation of fatty acid oxidation in peroxisomes is not well understood. β-Oxidation in mitochondria is largely determined at the level of free fatty acids, the availability of carnitine and the inhibition of carnitine palmitoyltransferase I by malonyl-CoA (Hashimoto 1999). Both of the systems are induced by cold or alternatively with hypolipidemic drugs or physiological ligands through the peroxisome proliferator-activated receptor alpha (PPARα), which is regarded as long-term regulation (Nedergaard et al. 1980, Aoyama et al. 1998, Vamecq & Latruffe 1999). It has been demonstrated that this receptor subsequently activates genes that encode enzymes involved in β-oxidation, such as the multifunctional enzyme type 1, acyl-CoA oxidase and 3-ketoacyl-CoA thiolase (Reddy et al. 1980, Reddy et al. 1986, Nicolas-Frances et al. 2000).

2.3 Multifunctional enzymes and enzyme complexes in β-oxidation

Fatty acid β-oxidation from yeast to humans requires numerous enzymes. A few of these are multifunctional enzymes that contain more than one activity in a single polypeptide. In a similar fashion, several proteins with individual activities can form a large enzyme complex to carry out a task, a system that is more often encountered in bacteria. The most obvious significance or advantage of these units would be the channeling of substrates between active sites or the isolation of hydrophobic substrates or intermediates from aqueous surroundings to enable faster enzyme catalysis.
2.3.1 Peroxisomal multifunctional enzyme type 1

The crystal structure of full-length 79 kDa rat peroxisomal multifunctional enzyme type 1 (MFE-1), operating in the S-specific fatty acid oxidation pathway, was recently solved at 2.8 Å resolution (Kasaragod et al. 2010). The history of the enzyme dates back to 1980s, when it was first characterized from rat liver (Osumi & Hashimoto 1979, Furuta et al. 1980, Osumi & Hashimoto 1980). MFE-1 consists of Δ^3, Δ^2–enoyl-CoA isomerase, 2E-enoyl-CoA hydratase 1 and 3S-hydroxyacyl-CoA dehydrogenase activities (Palosaari & Hiltunen 1990). The five-domain protein has two active sites of which the isomerase and hydratase activities reside in the same N-terminal domain (domain A) and the dehydrogenase active site is situated in the C-terminal part (domains C-E), while leaving the linker helix (domain B) between the catalytic units (Kasaragod et al. 2010). The recent structure of MFE-1 suggests the existence of substrate channeling between the active sites situated 40 Å from each other (Kasaragod et al. 2010). The structural finding is consistent with the earlier published kinetic data (Yang et al. 1986, Palosaari & Hiltunen 1990). The crystal structure of full-length rat MFE-1 is presented in figure 2.

The exact physiological role of MFE-1 is still a mystery, but there are a few suggestions based on the fact that the enzyme has broad substrate specificity (Palosaari & Hiltunen 1990, Xu & Cuebas 1996, Kurosawa et al. 2001). A study done using human fibroblasts from patients suffering from β-oxidation disorders suggest that MFE-1 has a role in long-chain dicarboxylic acid β-oxidation together with MFE-2 and sterol carrier protein X (Ferdinandusse et al. 2004). A second study performed on α-methylacyl-CoA racemase (Amacr) deficient mice showed that MFE-1 could take part in bile acid synthesis allowing the mutant mouse to survive (Savolainen et al. 2004). Even though the S-specific fatty acid β-oxidation pathway catalyzed by MFE-1 has properties that overlap with those of the R-specific pathway of MFE-2, the functional similarity does not correlate at the structural level (Bhaumik et al. 2005). Patients with MFE-1 deficiency have not been identified and most interestingly, MFE-1 deficient mice are viable and are able to produce offspring. Most likely in mice the oxidation of fatty acids in this pathway is diverted to the R-specific route (Qi et al. 1999).
Fig. 2. The crystal structure of rat multifunctional enzyme type 1 (MFE-1). The 79 kDa enzyme (PDB ID 2X58) has been shaded to match the linear representation of the domain organization in MFE-1. The N-terminal part carrying the hydratase 1 and Δ^2,Δ^3-enoxy-CoA isomerise activities (domain A) has been shaded in white, while the linker helix (domain B) and the domains residing with the 3S-hydroxyacyl-CoA dehydrogenase activity (domains C-E) has been shaded in black and gray, respectively. The structure was solved in the presence of ligands, coenzyme A (CoA) and adenosine diphosphate (ADP). The figure has been modified from (Kasaragod et al. 2010).
2.3.2 Peroxisomal multifunctional enzyme type 2

Peroxisomal multifunctional enzyme type 2 (MFE-2) was first characterized from fungi (Hiltunen et al. 1992) and later from mammals almost simultaneously by several groups (Adamski et al. 1995, Corton et al. 1996, Dieuaide-Noubhani et al. 1997, Jiang et al. 1997, Qin et al. 1997a). MFE-2 is responsible for fatty acid β-oxidation in the R-specific pathway of 3R-hydroxy intermediates (Hiltunen et al. 1992, Adamski et al. 1995). The enzyme accepts a wide range of substrates, including very long-chain and 2-methyl branched fatty acids and bile acid intermediates (Reddy & Hashimoto 2001). The domain organization of MFE-2 as well as the molecular mass of the polypeptides varies considerably between different species. The human enzyme consists of three domains, two of which have enzymatic activities (Adamski et al. 1995). The N-terminus is occupied by the 3R-hydroxyacyl-CoA dehydrogenase domain (DH), while the 2E-enoyl-CoA hydratase 2 domain (H2) is situated in the middle of the full-length enzyme. The C-terminus consists of a 14 kDa sterol carrier protein 2-like domain (SCP-2L) that is capable of binding a number of different lipids (Haapalainen et al. 2001). The physiological role of the SCP-2L domain is unknown. The domain organization in various MFE-2s is presented in figure 3.

Fig. 3. Domain organization of MFE-2 displaying great variation between the species. The diamond pattern represents the 3R-hydroxyacyl-CoA dehydrogenase domain with a molecular mass of 35 kDa. The open rectangles represent the 2E-enoyl-CoA hydratase 2 domains with a molecular mass of 33 kDa. The 14 kDa SCP-2L domain is represented using a striped pattern and is missing from the enzyme in the fruit fly and in yeasts (Saccharomyces cerevisiae and Candida tropicalis). The worm (Caenorhabditis elegans) expresses the hydratase 2 domain as an individual polypeptide, while the yeast enzyme has duplicated the dehydrogenase domain to A and B, which have different substrate length specificities (Qin et al. 1999).
3R-hydroxyacyl-CoA dehydrogenase domain

More than 47000 enzymes have been classified as belonging to the short-chain dehydrogenase/reductase superfamily (Kallberg et al. 2010). The members of the superfamily share a few distinct features that are conserved through the phylogenetic tree from bacteria to plants and mammals. Despite the size and diversity of this protein family, all of the members have the “Rossmann fold” that consists of a core of parallel β-sheets lined up with a group of parallel α-helices following a topological order of βαβ. The fold is further labeled with the characteristic G-X-X-X-G-X-G phosphate binding consensus sequence (Rossmann et al. 1974, Wierenga et al. 1985). The catalytic site is also conserved; the Y-X-X-X-K –motif has proven to be crucial for the catalytic reaction (Borrás et al. 1989), and later a conserved serine has been suggested to be important for the catalysis (Cols et al. 1997).

The reaction mechanism is most likely the same in all of the SDR-members and requires the binding of the coenzyme first in the apo form of the enzyme. The residues in the catalytic triad each form hydrogen bonds to one of the three water molecules located in the vicinity the side chains. Two of the hydrogen bonds are broken when the binding of the NAD⁺ molecule takes place. In the second phase, the proton is extracted from the tyrosine residue of the catalytic triad (S-Y-K) and transferred to the NAD⁺. During the transfer the remaining water molecule forms an intermediate alcohol that finally donates the hydride ion to the coenzyme. In the third phase, the products, a ketone and NADH, leave the active site (Benach et al. 1999). One could expect significant changes in the protein conformation upon binding of the bulky coenzyme, but some evidence suggests this is not the case (Grimm et al. 2000).

The 3R-hydroxyacyl-CoA dehydrogenase domain of MFE-2, responsible for the third step of the β-oxidation cycle, possesses the “Rossmann fold” nucleotide-binding core typical for the SDR (short-chain dehydrogenase/reductase) superfamily (Jörnvall et al. 1995) and an extended C-terminal subdomain of 60 residues that is unique for MFE-2s (Haapalainen et al. 2003). The dehydrogenase domain of MFE-2 is biologically active when assembled into dimers. The dimerization is mediated via the C-terminal extension, which allows the two monomers to align on top of each other in the C-terminal part. Additional strength for the dimerization is provided by the network of mainly hydrophobic interactions between the N-terminal α-helices in the nucleotide-binding subdomain (Haapalainen et al. 2003). The mode of dimerization of the
dehydrogenase domains in yeast MFE-2s is slightly different due to the duplication of the dehydrogenase domain. The yeast dehydrogenase forms a structural dimer between the A- and B-domains within the same polypeptide (Ylianttila et al. 2006). It is not clear, how the substrates of varying chain length are guided to the correct active sites that have different substrate chain length specificities (Qin et al. 1999). Malfunction of the catalytic unit in humans causes one type of MFE-2 deficiency (van Grunsven et al. 1998). The dehydrogenase domain of rat MFE-2 is presented in figure 4.

**Fig. 4.** Diagram of the 3R-hydroxyacyl-CoA dehydrogenase dimer of rat MFE-2 (PDB ID 1GZ6). The molecular mass of one monomer is approximately 34.5 kDa. White and dark gray colors represent each of the monomers. The structure contains nicotinamide adenine dinucleotide (NAD⁺) as a natural ligand of the dehydrogenase in both of the monomers.

**2E-Enoyl-CoA hydratase 2 domain**

The second step of β-oxidation cycle in the R-specific route is catalyzed by the 2E-enoyl-CoA hydratase 2 domain that belongs to the hot-dog fold superfamily (Leesong et al. 1996). A hot-dog fold is formed when a long central α-helix (the sausage) is intertwined around a bundle of highly curved β-sheets forming the “bun”. Structure database searches produced approximately 160 hits with this fold. Despite the similarities at structural level, the enzymes with the hot-dog fold have significant differences in their substrate specificity, amino acid sequence and
oligomerization (Pidugu et al. 2009). The range of substrates in the superfamily include for example saturated and unsaturated acyl-CoAs of various lengths, acetyl-CoA, phenylacetyl-CoA, 4-hydroxybenzoyl-CoA and 2,5-dihydroxybenzoyl-CoA (Dillon & Bateman 2004). In contrast to other MFE-2s, *C. elegans* expresses the hydratase 2 domain as a separate polypeptide apart from the dehydrogenase and SCP-2L domains that form a fusion protein. In the hydratase reaction, the hydroxyl group is added to the R-position in the 2E-enoyl-CoA substrate. Regardless of the species, the domain is able to catalyze the hydration reaction from short chain (C4) to very long chain (>C20) fatty acyl-CoAs (Malila et al. 1993, Jiang et al. 1996, Qin et al. 1997b, van Grunsven et al. 1999b). In humans dysfunction of the catalytic unit leads to accumulation of very long-chain substrates in the body (van Grunsven et al. 1999b). It is noteworthy that yeasts do not have mitochondrial fatty acid β-oxidation, but instead all of the available substrates are processed in the peroxisomes (Poirier et al. 2006).

The broad substrate specificity can be explained at the structural level. When comparing the monomer of eukaryotic 2E-enoyl-CoA hydratase of MFE-2 (Koski et al. 2004) and the dimeric bacterial R-hydratase from *A. caviae* (Hisano et al. 2003), it appears that there is only one catalytic site in the eukaryotic hydratase. At the expense of the second active site, the additional space in the reaction center of the eukaryotic hydratase enables the utilization of substrates up to very long-chain fatty acyl-CoAs (Koski et al. 2004). The crystal structure of the liganded hydratase 2 dimer from yeast *C. tropicalis* is presented in figure 5.
Fig. 5. Dimeric 2E-enoyl-CoA hydratase 2 from yeast MFE-2 (PDB ID 1PN4). The enzyme is liganded with 3R-hydroxydecanoyl-CoA revealing the spacious substrate binding area. The interactions between the two monomers are mainly hydrophobic. There are two hot-dog folds in a 32 kDa monomer, which are indicated in dark gray and white respectively.

Another characteristic feature in the hydratase 2 domain is the conserved hydratase 2 fingerprint Y-L-R-(G)-S-G-D-X-N-P-L-H-X-D-P-X-X-A (Qin et al. 2000). Complementation analysis in yeast *S. cerevisiae* revealed that the human enzyme contains at least two amino acids that are crucial for the activity, D510 and E366 (Qin et al. 2000). However, the liganded structure of the yeast counterpart revealed that residues important for catalysis in addition to the D510 are most likely N513, H515 and G533 (Koski et al. 2004). Due to extensive conservation, these amino acids are likely to be important for the enzyme catalysis from yeast to humans.

**Sterol carrier protein 2-like domain**

The sterol carrier protein 2-like (SPC-2L) domain is present in mammalian MFE-2s as the C-terminal part of the large polypeptide, but is expressed as a separate unit apart from the fruit fly catalytic polypeptide chain and fungal homologs with one known exception, the arbuscular mycorrhizal fungus, *Glomus mosseae*. The biological function of this small domain is unknown, but it is not crucial for the oxidation of the fatty acyl-CoAs. Despite having no role in the actual enzyme catalysis, the domain is capable of binding various kinds of natural lipids (Dyer et
al. 2003, Dyer et al. 2009, Singarapu et al. 2010). The crystal structure of human SCP-2L was solved in complex with Triton X-100, which is a lipid-like molecule commercially available (Haapalainen et al. 2001). The structure displays a fold with a large hydrophobic binding pocket that has also been characterized from human, rabbit and mosquito homologs (Choinowski et al. 2000, Garcia et al. 2000, Dyer et al. 2003).

The mechanism by which SCP-2L becomes targeted into peroxisomes is still under discussion. A ligand assisted targeting mechanism has been proposed in which the binding of a lipid molecule is prerequisite for the PTS1 targeting signal to be accessible for Pex5 (Lensink et al. 2002). However, in the case of the mosquito protein, AedesSCP-2, the data indicate that the protein secondary structure is not altered when comparing the apo and holo forms of the structure. Instead, it was suggested that ligand binding to AedesSCP-2 significantly reduces the ability of the protein to move. Therefore it was proposed that different SCP-2s have preformed binding pockets that allow the binding of fatty acids in alternative configurations (Singarapu et al. 2010). The structure of human SCP-2L is presented in figure 6.

Fig. 6. The crystal structure of human SCP-2L at 1.75 Å resolution liganded with Triton X-100 (PDB ID 1IKT). The lipid binding pocket is large enough to accommodate long-chain fatty acids, such as palmitoyl-CoA. The PTS1-targeting signal is exposed in the structure at the end of the long α-helix that is situated above the Triton X-100 molecule. The protein has been shaded according to the secondary structures. Dark gray represents α-helices, the β-sheets are pale gray and the remaining loops are black. Triton X-100 is shown in plain gray.
Multifunctional enzyme type 2 deficiency

The first patients with an inborn autosomal recessive MFE-2 deficiency were described before the actual enzyme was even characterized (Aubourg et al. 1986, Goldfischer et al. 1986, Schram et al. 1987). The most common symptoms in MFE-2 deficient patients include neonatal hypotonia, various dysmorphic features of the head, hepatomegaly, enlarged peroxisomes and neuronal seizures (Huyghe et al. 2006). The symptoms are caused by the accumulation of the MFE-2 substrates in the body (Schram et al. 1987). These clinical states were originally called pseudo-Zellweger syndrome and neonatal adrenoleukodystrophy. However, a more recent study helped to reclassify the disorders as MFE-2 deficiency (Ferdinandusse et al. 2002). The first recognized MFE-2 deficiency was described more than a decade ago (Suzuki et al. 1997) and after reinvestigation of older cases with proposed MFE-1 deficiency, more patients with defects in MFE-2 were discovered (van Grunsven et al. 1998, van Grunsven et al. 1999a).

This single enzyme peroxisomal disorder can be further classified into three subgroups that include (i) patients lacking MFE-2, (ii) patients with mutations in the hydratase 2 domain and (iii) patients with mutations in the dehydrogenase domain (Huyghe et al. 2006). Further investigations demonstrated correlations between the genotypes and phenotypes that were based on structural data available from the catalytic domains (Ferdinandusse et al. 2006). The mutations affect mainly the coenzyme binding area, the two catalytic sites and the extensive dimerization interfaces. MFE-2 deficiency is not treatable and depending on the severity of the mutations, the patients live usually less than ten years (Ferdinandusse et al. 2006b).

2.3.3 Mitochondrial trifunctional enzyme (TFE)

The mitochondrial inner membrane associated trifunctional enzyme (TFE) was almost simultaneously discovered by three different groups (Carpenter et al. 1992, Uchida et al. 1992, Luo et al. 1993). The protein is involved in the mitochondrial β-oxidation spiral and accepts up to long-chain substrates. In humans the monomeric enzyme consists of two different subunits. The α-subunit with a molecular mass of 71 kDa has enzymatic domains displaying enoyl-CoA hydratase (ECH) and 3-hydroxyacyl-CoA dehydrogenase (HACD) activities. The 47 kDa β-subunit has 3-ketoacyl-CoA thiolase (KACT) activity. It has been suggested that the human enzyme is a α2β2 heterotetramer (Carpenter et al. 1992).
In contrast, the enzyme purified from rat liver displays an α₄β₄ assembly (Uchida et al. 1992). The three-dimensional crystal structure of TFE is currently not known, but efforts have been made by homology modeling (Ishikawa et al. 2004) and electron microscopy (Fould et al. 2010) to predict the structure.

Mutations in genes encoding the catalytic subunits of TFE, like mutations affecting many enzymes involved in fatty acid metabolism, result in severe defects. The first case of TFE deficiency was reported already the same year the enzyme was discovered (Wanders et al. 1992). Further investigations revealed that the α-subunit is more prone to null mutations and the β-subunit to missense mutations. Three phenotypes are associated with the deficiency of which the most severe cardiomyopathic form is lethal (Moczulski et al. 2009).

2.3.4 Bacterial β-oxidation complex

The bacterial β-oxidation enzyme complex is known from Escherichia coli (Binstock et al. 1977) and Pseudomonas fragi (Imamura et al. 1990). The multienzyme complex from E. coli is called FOM and the P. fragi counterpart was designated HDT. The similarity of these bacterial enzymes to the mammalian TFE at the amino acid level is relatively high (Ishikawa et al. 2004), but in addition to the hydratase, dehydrogenase and thiolase activities, the bacterial enzyme has also 3-hydroxyacyl-CoA epimerase and 3,2-enoyl-CoA isomerase activities (Pramanik et al. 1979). The function of the bacterial β-oxidation system is highly similar to the eukaryotic systems, although the substrates are not identical. The substrate specificity in the bacterial β-oxidation complexes is undoubtedly broad (Kunau et al. 1995). Like in TFE, the hydratase and dehydrogenase activities reside in subunit α, and the thiolase activity is found in subunit β.

Interestingly, HDT from P. fragi was not only crystallized (Ishikawa et al. 1997), but later the large structure was also solved at high resolution (Ishikawa et al. 2004). The structure allowed the detailed characterization of the assembly of the biologically active enzyme, and also revealed a novel substrate channeling mode in which the complex utilizes a special substrate-anchored mechanism to avoid the diffusion of intermediates into the surrounding solvent between the reactions. The crystal structure of the bacterial fatty acid β-oxidation complex is presented in figure 7.
Fig. 7. The crystal structure of the bacterial fatty acid \( \beta \)-oxidation complex (HDT) from \textit{P. fragi} at 2.5 Å resolution (PDB ID 1WDK). The liganded 242.3 kDa heterotetramer consists of 2 \( \alpha \)-subunits and 2 \( \beta \)-subunits forming an \( \alpha_2\beta_2 \) assembly. Pale gray and dark gray shading represent one \( \alpha\beta \)-subunit each. 2-enoyl-CoA hydratase (ECH) and 3S-hydroxyacyl-CoA dehydrogenase (HACD) are present in the larger polypeptide of 73 kDa and the 3-ketoacyl-CoA thiolase (KACT) in the smaller 42 kDa polypeptide (Ishikawa \textit{et al.} 1997).

As expected, the multistep reactions in such a large enzyme require substantial coordination. The bulky nature of the large fatty acyl-CoA substrates induces more complexity in the reaction mechanism. When considering the actual reaction sequence in \( \beta \)-oxidation, the overall architecture of the enzyme complex is not obvious since the hydratase is located between the dehydrogenase and thiolase activities (Ishikawa \textit{et al.} 2004). The ECH and HACD subunits share the same binding pocket for the 3’-phosphate ADP moiety, but in this sense KACT is autonomous. The reaction starts by providing 2E-enoyl-CoA to the ECH subunit that hydrates the substrate to 3S-hydroxyacyl-CoA. Subsequently the acyl-group of the substrate swings into the HACD active site that generates the 3-ketoacyl-CoA. The substrate is anchored in the shared binding pocket via the adenine moiety that in practice functions as a pivoting center during these two reactions. The 3-ketoacyl-CoA becomes detached from the first binding pocket after
substantial conformational change prior to the thiolase reaction that weakens the affinity of the 3’-phosphate moiety for the first binding pocket. The overall rearrangement of the structure thus enables the correct positioning of the last intermediate in the thiolase subunit. It is crucial to keep the acyl-group attached to the thiolase subunit during the relocation of the intermediate. The reaction end products are released from the complex after the thiolytic cleavage (Ishikawa et al. 2004).

2.4 Substrate channelling in multifunctional enzymes and enzyme complexes

An event, in which the intermediates become translocated non-covalently between the catalytic sites without releasing the molecules into the surrounding solvent during enzyme catalysis, is called substrate channeling. This kind of substrate channeling can occur in multifunctional enzymes or multienzyme complexes (Huang et al. 2001, Weeks et al. 2006). Quite often the term substrate tunneling is used if the intermediate is transferred to the subsequent reaction site within the same enzyme (Weeks et al. 2006). This can even involve a real physical tunnel inside the enzyme if the substrate is small enough, like carbon monoxide or ammonia (Raushel et al. 2003).

Almost any enzymatic reaction involving an effective channeling mechanism can provide significant benefits in the overall reaction outcome. First, the reaction times are reduced, because of the direct transfer of the intermediates to the following reaction center. Secondly, labile substrates can be protected from degradation, which prevents the loss of intermediate products. In addition, any toxic intermediates can easily be isolated from the cell environment and competing pathways excluded from interfering. In tightly coupled reactions, it is also possible to prevent the accumulation of certain intermediates during a specific step (Huang et al. 2001, Milani et al. 2003). Although the channeled intermediates are often similar, the tunnels characterized thus far are structurally distinct (Weeks et al. 2006).

2.4.1 Tunnel-forming enzymes

The first substrate tunnel was verified in *Salmonella typhimurium* tryptophan synthase using X-ray crystallography at 2.5 Å resolution (Hyde et al. 1988). This bifunctional enzyme, which catalyzes the last step in L-tryptophan biosynthesis,
consists of two separate polypeptides that assemble into a 143 kDa $\alpha_2\beta_2$-heterotetramer in bacteria. The 29 kDa tryptophan synthase subunit $\alpha$ is responsible for cleaving the indole and glyceraldehyde-3-phosphate intermediates from the substrate indole-3-glycerol phosphate in the first half of the reaction. In the second phase, the intermediate indole is condensed with L-serine in the 43 kDa $\beta$-subunit (Crawford & Yanofsky 1958, Hyde et al. 1988). The tunnel observed in tryptophan synthase is approximately 25 Å long and connects the two catalytic sites of the enzyme allowing the indole to pass through (Hyde et al. 1988). The data from various kinetic experiments, like quench-flow and stopped-flow, demonstrate that the indole indeed is channeled through the enzyme (Matchett 1974, Leja et al. 1995). The structure of tryptophan synthase is presented in figure 8A.

A substantially longer substrate tunnel that connects three active sites was discovered in carbamoyl phosphate synthetase (CPS). The protein is formed from two subunits, $\alpha$ and $\beta$, that have molecular masses of 42 kDa and 118 kDa, respectively. The active enzyme converts into an $\alpha_4\beta_4$-heterooctamer and the distance between the first and last reaction centers is 96 Å within an $\alpha\beta$-heterodimer (Thoden et al. 1997). The smaller subunit hydrolyzes the glutamine yielding ammonia and glutamate, while the larger subunit is responsible for assembling the end-product carbamoyl phosphate in sequential reactions that involve bicarbonate and carbamate phosphorylation in tandem with generation of the ATP necessary for the reaction (Raushel et al. 1978, Thoden et al. 1997). The substrate channeling mechanism that requires synchronization of the active sites has been validated by rapid quench investigations (Miles & Raushel 2000). The substrate channel present in CPS is shown in figure 8B.

Glutamate synthase is an iron-sulfur protein that possesses a tunnel of 31 Å in length for intermediate channeling (Binda et al. 2000). Two molecules of L-glutamate are formed when ammonia extracted from glutamine is used for sequential reductive transfer to 2-oxoglutarate (Tempest et al. 1970). The *Azospirillum brasilense* 150 kDa $\alpha$-subunit and 50 kDa $\beta$-subunit of glutamate synthase form $\alpha\beta$-heterodimers. The larger subunit provides the tools for the glutamine hydrolysis, the iron-sulfur cluster as well as the binding sites for 2-oxoglutarate and flavin mononucleotide. NADPH is oxidized in the smaller subunit that is instantly followed by the reduction of FAD (Vanoni et al. 1996).
Fig. 8. Examples of substrate tunnels within enzymes. Pale gray represents the α-subunits and dark gray the β-subunits of the enzymes. Arrows with black and pale gray frames represent the approximate locations of the substrate tunnels that pass from one subunit to another during subsequent reactions. (A) Tryptophan synthase monomer from *S. typhimurium* (PDB ID 2TRS) displaying a 29 Å tunnel between the α-subunit and the β-subunit. (B) *E. coli* carbamoyl phosphate synthetase monomer has a tunnel that is 96 Å in length and connects the three catalytic sites (PDB ID 1JDB). (C) Acetyl-CoA synthase/carbon monoxide dehydrogenase dimer from *Moorella thermoacetica* (PDB ID 1OAO) displays a passage that is 138 Å long and is formed by four of the iron-sulfur clusters located in the two β-subunits. The diagrams were adapted from reviews focused on substrate channeling (Huang et al. 2001, Weeks et al. 2006).
Ammonia, which is small and toxic, seems to be a rather common intermediate for channeling. Ammonia passage has been characterized from several enzymes including asparagine synthetase (Larsen et al. 1999), glucosamine 6-phosphate synthase (Teplyakov et al. 2001) imidazole glycerol phosphate synthase (Douangamath et al. 2002), cytidine triphosphate synthetase (Endrizzi et al. 2004) and a tRNA-dependent amidotransferase (Schmitt et al. 2005). The 310 kDa acetyl-CoA producing acetyl-CoA synthase/carbon monoxide dehydrogenase from *M. thermoacetica* (figure 8C), in turn, tunnels toxic carbon monoxide through a hydrophobic cavity of 138 Å (Doukov et al. 2002, Darnault et al. 2003). Tunneling of a toxic compound also occurs in the bacterial 4-hydroxy-2-ketovalerate aldolase/acetylating acetaldehyde dehydrogenase (DmpF) that is involved in the last steps of the degradation of toxic aromatic intermediates. In the enzyme, acetaldehyde is passed through a 29 Å tunnel to protect the cell until it is converted to acetyl-CoA (Manjasetty et al. 2003).

2.4.2 Eccentric ways of substrate channeling

A substrate tunnel is a very sophisticated way to prevent the diffusion of intermediates into the surrounding solvent, but not all the characterized enzymes use tunnels for channeling the molecules between the catalytic sites. Reaction molecules, for example, can be swung between the active sites that are close to each other or alternatively translocated via the surface of the protein or by novel mechanisms yet to be proposed and discovered. The previously mentioned bacterial β-oxidation complex is a good example of a mechanism that is considered substrate channeling because the substrate pivots between active sites, but does not involve a tunnel (Ishikawa et al. 2004).

An ‘electrostatic highway’ was discovered in *Leishmania major* bifunctional dihydrofolate reductase-thymidylate synthase (Knighton et al. 1994, Stroud 1994) and the data were supported by kinetic experiments using bacterial enzymes (Meek et al. 1985, Trujillo et al. 1996, Liang & Anderson 1998). The specific channeling mechanism was suggested based on the highly unusual positive surface charge distribution provided by basic amino acid residues of arginine, histidine and lysine. The folate binding sites on the two domains are separated by approximately 40 Å (Knighton et al. 1994). A similar kind of channeling mechanism could be considered in the case of MFE-1 as well, although the positively charged surface area is not so extended (Kasaragod et al. 2010). In addition, an electrostatic channeling mechanism in formiminotransferase-
cyclodeaminase has been proposed based on kinetic evidence and a partial structure of the enzyme (Paquin et al. 1985, Kohls et al. 2000).

Structural and functional characterization of the separated units of the 2-oxo acid dehydrogenase complex (that is structurally and functionally related to the pyruvate dehydrogenase complex and branched-chain oxo acid dehydrogenase multienzyme complex) allowed the identification of a novel channeling mechanism that is referred to as covalent substrate channeling. The complex consists of dozens of copies of three major parts that include thiamine diphosphate-dependent 2-oxo acid dehydrogenase, E1 (Aevaarsson et al. 1999, Frank et al. 2007), dihydrolipoyl acyltransferase, E2 (Berg et al. 1996, Hübner et al. 1998) and dihydrolipoamide dehydrogenase, E3 (Toyoda et al. 1998) and has a total molecular mass up to 10 MDa (Perham 1991). A reaction intermediate that is covalently bound to the enzyme is translocated to the following catalytic site by channeling (Huang et al. 2001). The mechanism involves at least one hydrophobic funnel-shaped channel that leads to the reaction center of E1 (Aevaarsson et al. 1999).
3 Aims of the study

The purpose of the work was to:

1. Improve the expression and purification of a full-length MFE-2;
2. Solve the high resolution structure of full-length DmMFE-2;
3. Apply structural data and enzyme kinetics for comparison of the substrate transfer mechanism in DmMFE-2 to the known cases among multifunctional enzymes.
4 Materials and methods

The biggest obstacle in the structural project is often to obtain a sufficient amount of full-length recombinant protein for crystallization experiments. In the current project, the problem was approached by selecting more than one species for the initial expression and purification screens. The most promising candidate, *Dm*MFE-2, was selected for further processing based on the preliminary tests. The detailed description of the materials and methods is available in the original articles referred to by their Roman numerals (I-III).

4.1 Plasmids, strains and culture media

The construction of the recombinant expression plasmids for yeast and bacteria, the expression hosts used and the preparation of different culture media and plates are described in detail in the original articles. The plasmid for expressing the human hydratase 2 (HsH2) domain of MFE-2 was kindly provided by Dr. M.K. Koski. The enzyme was expressed and purified as previously described (Koski et al. 2005).

4.2 Production of recombinant proteins in *E.coli* (I-III)

The overexpression plasmids of *Dm*MFE-2, *Dm*DH, *Dm*H2, *Hs*DH and *Hs*SCP-2L were transformed into the *E. coli* BL21(DE3) pLysS strain and selected on fresh Luria Bertani (LB) plates containing glucose (2 mg/ml), carbenicillin (50 µg/ml) and chloramphenicol (34 µg/ml). The human overexpression plasmid was transformed into Lemo21(DE3) cells. A liquid pre-culture from cells carrying the *Dm*MFE-2 or *Hs*DH or *Hs*SCP-2L expression plasmid was prepared in LB-medium supplemented with glucose and antibiotics at the concentrations mentioned above. After culturing the cells to saturation at 30 ºC, 1 liter expression cultures were started in M9ZB-medium using 20 ml of the pre-culture. The cells were incubated in a shaker at 37ºC and the production of recombinant protein was induced using 0.4 mM to 0.6 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) after the OD₆₀₀ reached 0.8. The induction occurred at 37 ºC for 5 hours after which the cells were harvested, washed with PBS, re-suspended in Ni-NTA lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 1.0 M urea) and flash frozen using liquid nitrogen prior to storage at -70 ºC. For further improvement of the expression of recombinant *Dm*MFE-2, the enzyme was
produced in commercial medium, EnBase Flo, according to the instructions provided by the manufacturer (BioSilta). In addition, the final yield of purified DmMFE-2, as well as the expression in E. coli from cell lysates on Coomassie blue-stained SDS-polyacrylamide gels, was compared after expressing the cells on different media (LB and M9ZB).

The expression of the full-length HsMFE-2 and the individual domains of DmMFE-2 (DmDH and DmH2), were also done using EnBase Flo. The high-yield recombinant protein expression was carried out in 0.5 liter volume according to the guidelines provided by the kit. IPTG (0.4 mM to 1 mM) was added to the culture medium for recombinant protein induction at 30 °C. The cells were harvested after 24 hours and subsequently treated like the cells from the full-length DmMFE-2 production, except that the bacterial pellet containing DmDH was resuspended in Profinity eXact lysis buffer (100 mM sodium phosphate pH 7.2, 300 mM sodium acetate).

4.3 Protein purification (I-III)

The purification of DmMFE-2, DmDH, DmH2, HsMFE-2, HsDH and HsSCP-2L was initiated by thawing the bacterial cell suspension at 30 °C. The lysis of the cells was completed by adding lysozyme (500 µg/ml), DNase I (50 µg/ml) and RNase A (50 µg/ml). After 30 min of incubation at 30 °C, the lysate was cleared by centrifugation (30 000 g, 45 min, 4 °C). The soluble fraction of DmMFE-2 was applied at 1 ml/min to a Ni-NTA column (Qiagen) that was equilibrated in Ni-NTA lysis buffer and the weakly bound proteins were flushed using Ni-NTA wash buffer (50 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazol). The target protein was released by Ni-NTA elution buffer (50 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazol) and subsequently diluted with 10 volumes of distilled water. The diluted DmMFE-2 was applied to a 14 ml Red Sepharose CL-6B affinity column at 0.5 ml/min that was equilibrated in 50 mM potassium phosphate buffer pH 7.0. A linear KCl gradient (0-2.0 M in 140 ml) was used to elute the bound protein from the column. Fractions containing DmMFE-2 were pooled and concentrated in the presence of 10% glycerol. The concentrated protein was finally applied to a size exclusion column, Superdex 200 10/300 GL (GE Healthcare) that was equilibrated in 50 mM sodium phosphate buffer pH 7.5, 200 mM NaCl, 1 mM EDTA, 5% glycerol, 1 mM NaN₃. The same buffer was also used for the elution of the protein sample. The major peaks were fractionated, and DmMFE-2 containing fractions were pooled and concentrated to 2.9 mg/ml.
The purification of \textit{Dm}H2, \textit{Hs}DH and \textit{Hs}SCP-2L followed the same protocol as the purification of \textit{Dm}MFE-2, except that the Red Sepharose CL-6B purification step was omitted and an imidazol gradient was used for the elution of \textit{Hs}DH. A single 500 ml culture expressing \textit{Dm}H2 using the EnBase Flo medium resulted in 113 mg of purified protein. The \textit{Dm}DH however, was initially purified using a commercial Profinity eXact affinity column that was equilibrated in Profinity eXact lysis buffer. After binding the target protein in the column, the column was flushed with wash buffer (100 mM sodium phosphate pH 7.2, 150 mM sodium acetate). The tag-free \textit{Dm}DH was released from the column by elution buffer (100 mM sodium phosphate pH 7.2, 200 mM NaF). From this step on the purification procedure followed that of \textit{Dm}H2. A 500 ml culture expressing \textit{Dm}DH in EnBase Flo yielded 5 mg of purified \textit{Drosophila} dehydrogenase. The yield from the rest of the samples was typically less than five milligrams from a one liter culture.

The full-length \textit{Dm}MFE-2 was also expressed and purified as tag-free following the protocol developed for \textit{Dm}DH utilizing the pPAL7-vector (Bio-Rad). This was done in order to see if the tag-free protein would produce better crystals for the structure determination.

4.4 Enzyme activity assays (II)

The method described by (Hiltunen et al. 1989) was applied in the determination of enzyme activities of \textit{Dm}MFE-2 towards the selected substrates. 2E-butenoyl-CoA, 2E-hexenoyl-CoA and 2E-decenoyl-CoA were synthesized as described (Qin et al. 1999). The measurements were carried out as double determinations using quartz cuvettes (0.5 ml) at 22 °C. GraFit 5 (Erithacus Software) was used for calculating \(K_M\), \(V_{max}\) and \(k_{cat}\) values.

4.5 Crystallization of \textit{Dm}MFE-2 (II)

The initial crystallization screening of \textit{Dm}MFE-2 was executed by using the sparse matrix screening method (Jancarik & Kim 1991). A Tecan Freedom Evo liquid handling robot was used for setting vapour-diffusion crystallization plates. Equal volumes of protein solution and mother liquor (1 µl) were used for setting up the drops at 22 °C. The best tetragonal crystals were grown in 100 mM Tris pH 8.0, 1.0 M NaCl, 20% (w/v) PEG 5000 MME, 5 mM NAD\(^+\). The crystal dimensions were 0.23 mm x 0.21 mm x 0.05 mm. Prior to data collection, the
crystal was transferred into a cryo-protecting mother liquor containing 20% (v/v) glycerol in addition to the original mother liquor, incubated for one minute and flash frozen in a suitable loop at 100 K using liquid nitrogen.

4.6 Data collection and structure determination of \( \text{DmMFE-2 (II)} \)

Synchrotron radiation at the ID29 beamline, ESRF, Grenoble was used for collecting a dataset (90° of diffraction data in 180 frames). XDS (Kabsch 2010) was used for processing the images and setting the high resolution limit to 2.15 Å. The structure was initially solved by molecular replacement using Phaser (McCoy et al. 2007). The existing crystal structures of rat dehydrogenase (1GZ6) and yeast hydratase 2 domains of MFE-2 (1PN4) were used as models in the rotation and translation search. After solving the structure, the amino acids sequences of the homology models were treated with CHAINSaw (Stein 2008) to match the \( \text{DmMFE-2} \) sequence. The missing parts of the structure were built manually in Coot (Emsley et al. 2010) and were refined using restrained refinement cycles by Refmac (Murshudov et al. 1997, Winn et al. 2001) or PHENIX (Adams et al. 2010). From the beginning of the structure refinement, TLS (Traslation-Liberation-Screw-rotation displacement) were introduced in the restrained refinement cycles both in Refmac and PHENIX by dividing the structure in four TLS groups. Group A comprised of residues 8-251, group B residues 252-312, group C residues 313-450 and group D residues 459-592. The obtained structure coordinates of \( \text{DmMFE-2} \) were finally validated using MolProbity (Davis et al. 2007) to estimate the geometry and steric of the structure.

4.7 Rescue experiments with \( \text{DmMFE-2} \) in a yeast mutant devoid of endogenous MFE-2 (II)

For the yeast complementation studies, the pYE352::\( \text{DmMFE-2} \) expression plasmid was transformed (Chen et al. 1992) into the \( \text{S. cerevisae} \) MFE-2 deletion strain BY4741 \( \Delta \text{fox2} \) (YKR009c, Euroscarf, accession code Y0508). In addition, the \( \Delta \text{fox2} \) strain was also transformed with pYE352::\( \text{CTA1} \) (Filppula et al. 1995) and pYE352::\( \text{ScMFE-2} \) (Qin et al. 1999). The yeasts were prepared on thin YP-plates supplemented with 0.14% oleic acid and 0.2% dextrose. BY4741 wt and BY4741 \( \Delta \text{fox2} + \) pYE352::\( \text{ScMFE-2} \) were used as the positive controls in the complementation experiment, while BY4741 \( \Delta \text{fox2} \) and BY4741 \( \Delta \text{fox2} + \) pYE352::\( \text{CTA1} \) served as the negative controls. The plates were incubated for one
week at 30 °C to force the yeast to utilize oleic acid as the sole carbon source. Clearing zones appeared around the positive samples after incubating the plates an additional week at 4 °C.

4.8 Static light scattering (SLS) experiments (II)

To investigate possible complex formation of the separate enzymatic domains in vitro, purified protein samples were analyzed using the multi-angle light scattering device miniDAWN TREOS (Wyatt Technology Corp.). Sample concentrations were 7.4 mg/ml (DmMFE-2), 5.7 mg/ml (DmDH), 6.1 mg/ml (DmH2) and 5.7 mg/ml (DmDH + DmH2 in a stoichiometric ratio). The device was connected to an ÄKTA purifier (GE Healthcare) and a Superdex 200 10/300 GL (GE Healthcare) size exclusion column equilibrated in 50 mM sodium phosphate pH 7.5, 200 mM NaCl, 1 mM EDTA, 5% glycerol, 1 mM NaN₃. 50 µl of protein was used for one injection and during the separation at 17 °C the system flow was set to 250 µl/min. The data were analyzed using Astra V software (Wyatt Technology Corp.). The same approach was used when monitoring the monodispersity of the protein samples prepared for SAXS data collection.

4.9 SAXS data collection, data refinement and construction of ab initio models of DmMFE-2 and HsMFE-2 (III)

Small-angle X-ray scattering data were collected at X33 beamline in DESY, Hamburg using the 1M PILATUS detector (Roessle et al. 2007) at a fixed 1.500 Å wavelength. Exposures of 30 seconds were used for measuring the scattering patterns from the monodisperse samples of HsSCP-2L and DmMFE-2 and the single domain fragments of the Drosophila protein at 15 °C. The corresponding exposure time for the rest of the HsMFE-2 samples was 15 seconds. The sample concentrations were 11.2 mg/ml and 7.6 mg/ml for DmMFE-2, 6.4 mg/ml and 3.2 mg/ml for DmDH, 18.9 mg/ml and 9.5 mg/ml for DmH2, and 6.4 mg/ml and 3.2 mg/ml for DmDH + DmH2 in a stoichiometric ratio. Several measurements were made with sample concentrations of 5.5 mg/ml and 4.0 mg/ml for HsMFE-2, 5.0 mg/ml and 1.9 mg/ml for HsDH, 5.2 mg/ml and 2.8 mg/ml for HsH2, and 8.4 mg/ml and 4.2 mg/ml for HsSCP-2L. In addition, reference data were also collected from BSA (MM = 66 kDa). A measurement of the buffer was done before and after all of the protein samples. The signal from the buffer was
subsequently subtracted from the scattering of protein prior to further data reduction. In addition to the use of automatic data processing at the X33 beamline (Petoukhov et al. 2007), further data processing were done using PRIMUS (Konarev et al. 2003) software. Guinier approximation was applied to evaluate the radius of gyration ($R_g$) and forward scattering ($I_0$) of the samples. $D_{\text{max}}$ and the interatomic distance distribution function $p(r)$ were calculated using the software GNOM (Svergun 1992). The molecular mass of a sample was evaluated by comparing the forward scatter of the reference sample (BSA) to each of the measured samples. The low resolution data from SAXS were compared to existing crystal structures using CRYSOL (Svergun et al. 1995). Finally, the ab initio models of each protein were reconstructed by the software DAMMIN (Svergun 1999) or MONSA (Svergun 1999) and finally averaged by DAMAVER (Volkov & Svergun 2003). Rigid body modeling was done using SASREF and BUNCH (Petoukhov & Svergun 2005).
5 Results

5.1 Purification of DmMFE-2 and the separate enzymatic domains

The full-length DmMFE-2 and the single enzymatic domains, dehydrogenase and hydratase, were purified to high homogeneity using affinity and size exclusion chromatography. The purification was initially monitored by SDS-polyacrylamide gel electrophoresis and Coomassie staining. Further information about the sample homogeneity was obtained from static light scattering analysis. Comparison of the final yields of crystallization grade DmMFE-2 expressed using different media revealed a significant difference. The use of EnBase Flo (BioSilta) medium resulted in a 13-fold increase of pure protein from a liter culture when compared to LB medium. The same phenomenon was also seen when the hydratase 2 domain of DmMFE-2 was expressed. This result proved to be crucial for the continuation of the work and obviated the need to explore the use of expression systems other than E. coli. The comparison of the expression of DmMFE-2 in two different media is presented in figure 9.

Fig. 9. The comparison of expression of recombinant DmMFE-2 in E. coli BL21 (DE3) using EnBase Flo (EB) and Luria Bertani (LB) media. Total protein and soluble fractions are compared after 24 hours of induction at 30 °C. The results demonstrate that the use of EnBase increases the yield of soluble fraction and total protein. The effect is approximately the same when using a mineral medium, such as M9ZB, instead of LB. The picture has been modified from (Krause et al. 2010) and the original publisher of the figure is BioMed Central.
5.2 Crystallization of DmMFE-2

The preliminary crystallization screening was done using double His-tagged DmMFE-2 in sitting drop vapor diffusion experiments. The initial screens produced crude crystals under several conditions in just a few hours. In addition, the number of nucleations seemed to be too high in most the drops resulting in large number of tiny crystals. The most promising hits were transformed into optimization grid screens using the hanging drop method by preparing fresh drops from the protein solution and the selected precipitants. The aim was to decrease the number of nucleations by gradually reducing the concentration of the precipitants. Originally this experiment appeared to fail, because most of the drops precipitated immediately after setting up. Nevertheless, the drops were left for incubation at room temperature. Six months after setting up the screens, some of the drops had cleared, and rather large and smooth tetragonal shaped crystals had appeared in the hanging drops (60 mM MOPS pH 7.5, 1.2 M (NH₄)₂SO₄, 2% PEG 400) at 21 °C. The crystals gave diffraction up to 3.0 Å when using a rotating anode X-ray source for the initial tests. The original expression and purification protocols of recombinant DmMFE-2 were optimized and the initial crystallization screens were repeated. The original protocol was lacking the second affinity chromatography step (Red Sepharose CL-6B) that significantly improved the purity grade of the final protein preparation. The quality improvement of the protein sample resulted in a higher number of hits in the preliminary crystallization screens, both with and without His-tag in the target protein. The best diffraction power was obtained from crystals grown in 0.1 M Tris pH 8.0, 1.0 M NaCl, 20% PEG 5000 MME at 21 °C. The crystal packing appeared to be similar regardless of the crystallization mother liquor and whether the protein was affinity tagged or not. Table 1 includes the conditions for obtaining the best DmMFE-2 crystals.

5.3 The crystal structure of DmMFE-2

The crystal structure of DmMFE-2 is first discussed at a general level, starting from the structure determination of the enzyme, and finally the results are viewed at the domain level.
5.3.1 Structure determination of DmMFE-2

The crystal structure of full-length DmMFE-2 was solved by molecular replacement at 2.15 Å resolution. The best solution in the rotation and translation search was obtained when using the space group P4_{3}2_{1}2. The asymmetric unit was occupied by one DmMFE-2 monomer. Interpretable electron density covered 80% of the total sequence, but some unclear regions were not constructed in the final structure. The missing regions were 1-6, 44-55, 88-89 and 200-204 in the 3R-hydroxyacyl-CoA dehydrogenase domain and 373-387, 451-457 and 593-598 in the 2E-enoyl-CoA hydratase 2 domain. Efforts to improve the data quality (including the resolution and better coverage of the electron density map) by producing better crystals by co-crystallizing and soaking the enzyme with natural ligands failed. For example electron density for NAD^{+} was not detected in the dehydrogenase domain. The final structural refinement statistics are summarized in table 1. The atomic coordinates and final structure factors were deposited in the RCSB Protein Data Bank under the ID 3OML.

Table 1. Crystallization, data collection and preliminary data processing statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DmMFE-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein concentration (mg/ml)</td>
<td>2.9</td>
</tr>
<tr>
<td>Crystallization temperature (°C)</td>
<td>21</td>
</tr>
<tr>
<td>Precipitant solution</td>
<td>0.1 M Tris pH 8.0, 1.0 M NaCl, 20% PEG 5000 MME</td>
</tr>
<tr>
<td>Ligand</td>
<td>5 mM NAD^{+}</td>
</tr>
<tr>
<td>Final resolution range (Å)</td>
<td>28.5–2.15</td>
</tr>
<tr>
<td>Crystal mosaicity (°)</td>
<td>0.61</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9310</td>
</tr>
<tr>
<td>Rotation range (°)</td>
<td>90</td>
</tr>
<tr>
<td>Data collection temperature (K)</td>
<td>100</td>
</tr>
<tr>
<td>Space group</td>
<td>P4_{3}2_{1}2</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a (Å)</td>
<td>114.48</td>
</tr>
<tr>
<td>b (Å)</td>
<td>114.48</td>
</tr>
<tr>
<td>c (Å)</td>
<td>89.11</td>
</tr>
<tr>
<td>Number of molecules in the asymmetric unit</td>
<td>1</td>
</tr>
<tr>
<td>Matthews coefficient (Å^{3}/Da)</td>
<td>2.25</td>
</tr>
</tbody>
</table>
5.3.2 The overall structure of DmMFE-2 monomer and dimer

The structure of MFE-2 is now extended from domain level to domain assembly level after solving the structure of the DmMFE-2. For the first time, the organization of the catalytic domains into the biologically active enzyme is described. In addition, the between-domain loop region and the factors contributing to the inter-domain interactions are revealed. The monomer, consisting of 3R-hydroxyacyl-CoA dehydrogenase in the N-terminus and 2E-enoyl-CoA hydratase 2 in the C-terminus, is roughly L-shaped when viewed from the front (figure 10) and hour-glass shaped in the side view with the between-domain loop region representing the structure of the hour-glass. The similarity of the current domain structures is high when compared to the previously published domain structures of MFE-2.

The biologically active DmMFE-2 dimer can be described in two ways. First, the crosswise alignment of the two polypeptides creates a complex interaction network between the two monomers, which form a homodimer at the 2-fold symmetry axis. Secondly, the enzyme could also be considered “double dimeric”, since both identical catalytic domains form their own dimers that are connected via short polypeptide linkers in the center of the molecule. The overall silhouette of the DmMFE-2 dimer is rectangular when viewed from the front and hourglass shaped from the side view, like the monomer. Unexpectedly, the between-domain space is not filled, but instead there is a small opening (⌀ = 9Å) that in theory could provide access for substrates between the domains (figure 10). Interestingly, a small number of positively charged amino acids residues, including H256, K285, K287, K305 and K372 from both of the polypeptides surround this opening. These residues could provide a positive surface potential patch to interact with the substrates that are covalently linked to a negatively charged carrier molecule, CoA.

Several hydrogen bonds, aromatic and hydrophobic interactions that are already known from the previously published catalytic domain structures (Haapalainen et al. 2003, Koski et al. 2004, Koski et al. 2005), dominate the interactions between the two DmMFE-2 monomers. A few salt bridges between the two polypeptides provide additional strength to the biologically active dimer. The bonds in the dehydrogenase domain between E215 and R261’, E215’ and R261, E241 and R254’, E241’ and R254 belong to this class. The hydratase 2 domain includes the ionic bonds between R492 and E347’ and between R492’ and E347. The two DmMFE-2 monomers could also be further cross-linked via
hydrogen bonds between S335 and D486 and between S335’ and D486 since the side chains of these residues are located only 2.6 Å apart.

Fig. 10. The DmMFE-2 dimer as ribbon and surface presentations. (A) Front view of the dimer. (B) Side view of the dimer after rotating the molecule 90° clockwise. (C) DmMFE-2 dimers as a surface representation. The black and white shades indicate the two DmMFE-2 monomers forming a dimer. The large ellipsoids emphasize the location of the dehydrogenase (DH) and hydratase 2 (H2) domains forming two dimers of their own.
5.3.3 3R-Hydroxyacyl-CoA dehydrogenase domain of DmMFE-2

Amino acid residues 1-307 cover the 3R-hydroxyacyl-CoA dehydrogenase domain (DmDH) of DmMFE-2. The first 244 residues display the Rossmann fold (Jörnvall et al. 1995) with an α/β doubly-wound structure and the last 63 residues in the C-terminal part of the domain constitute the extension specific to MFE-2 proteins. The overall structural topology of DmDH is highly similar to the previously published rat and yeast structures (PDB IDs 1GZ6, 2ET6), which is not surprising when the high homology between the members of the MFE-2 family is considered. The high structural homology between the enzymes is presented in figure 11. The conserved phosphate binding consensus sequence GXXXGXXG is ordered in the structure, but the non-conserved loop area that contributes to the binding of NAD⁺ (between β2 and α2) is completely absent. When compared to the other dehydrogenase domain structures (PDB IDs 1GZ6, 2ET6), it is clear that the polypeptide after the helix α2 does not form a clear secondary structure. In the current structure, the NAD⁺ region is open for cofactor binding, but the actual binding may be inhibited by a loop between sheets β12 and β13 from a symmetry related molecule. The nearest side chain of the symmetry molecule (Q474) is only 3.62 Å apart from the supposed location of the coenzyme binding site in the DmMFE-2 dehydrogenase domain.
Fig. 11. Structural comparison of 3R-hydroxyacyl-CoA dehydrogenase domains of MFE-2. The *Drosophila* dehydrogenase dimer has been superimposed on the homologous structures of rat and yeast *C. tropicalis* dehydrogenases of MFE-2 (PDB IDs 1GZ6, 2ET6). *DmDH* is shown in black, the rat dehydrogenase in white and the yeast dehydrogenase in gray. The overall structural topology is the same and the largest differences can be seen in the loop regions.

The C-terminal extension, a feature that is not common in the SDR superfamily (Jörnvall *et al.* 1995), is well-ordered in the current structure as well as the catalytic triad of the dehydrogenase, S154, Y167 and K171. These features were also seen in the rat dehydrogenase of MFE-2 (PDB ID 1GZ6). One interesting detail in the rat dehydrogenase dimer is the intersection of two tryptophan residues at position 246 in the longest β-sheet (corresponding to β7 in the current structure). This residue is conserved in human, rat, fish and *C. tropicalis* dehydrogenases, but during evolution this residue has been replaced by methionine in *DmMFE-2*. It is worth mentioning that a point mutation of W249G, among several others, results in MFE-2 deficiency in humans (Ferdinandusse *et al.* 2006).

**5.3.4 2E-Enoyl-CoA hydratase 2 domain of *DmMFE-2***

The 2E-enoyl-CoA hydratase 2 domain (*DmH2*) is formed from amino acids 314-598 in *DmMFE-2*. The domain consists of seven α-helices (α10-α16) and ten β-
sheets (β8-β17) that display strong electron density in the Fo-Fc difference map. The domain can be divided into two subdomains that are connected via a linker between sheets β12 and β13. In the cores of the subdomains a characteristic fold for the R-hydratase family, the hot-dog fold, is formed. The N-terminal hot-dog fold includes helix α12 that is covered by β-sheets β8-β12 and the C-terminal counterpart consists of helix α16 and sheets β13-β17.

The similarity of the hydratase 2 domain of DmH2 to the previously published domain structures of yeast and human MFE-2s (PDB IDs 1PN4, 1PN2 and 1S9C) is also high. However, one interesting detail appears when comparing the structures of liganded yeast hydratase 2 (PDB ID 1PN4) and Aeromonas caviae R-hydratase (PDB ID 1IQ6) with the current structure. The previous study with the yeast enzyme demonstrated the ability of the hydratase 2 to accept long chain substrates because the lack of the central α-helix (α12) creates extra space in the core of the domain (Koski et al. 2004). The corresponding α-helix in the A. caviae structure is fully folded, which allows the chemical catalysis of short chain substrates only. This helix (α12) is of intermediate length in the DmMFE-2 structure, but yet the enzyme has not lost the ability to utilize substrates of various chain lengths. When the ligand (3R-hydroxydecanoyl-CoA) from the yeast hydratase 2 structure is superimposed on the Drosophila hydratase 2, it can be seen that the helix α12 is in close contact with the fatty acid tail of the ligand. Despite this fact, enzyme kinetics indicate that the C10 substrate is preferred by DmMFE-2 over the short chain (C4-C6) substrates. Our unpublished results demonstrate that DmMFE-2 is capable of degrading substrates at least to chain length C16. The structural comparison of R-hydratases is presented in figure 12.
Fig. 12. Structural comparison of R-hydratases from various species. (A) Monomeric 31.5 kDa 2E-enoyl-CoA hydratase 2 from *C. tropicalis* MFE-2 (PBD ID 1PN4) liganded with 3R-hydroxydecanoyl-CoA. (B) 30.8 kDa monomeric 2E-enoyl-CoA hydratase 2 of *Dm*MFE-2. Helix α-12 is labeled in the figure and the position of the front loop is indicated by the black arrow. (C) Homodimeric R-hydratase from the bacterium *A. caviae* (PBD ID 1IQ6) with two 14 kDa subunits and two active sites. The secondary structures are shaded as follows, α-helices are black, β-sheets are gray and the remaining loops are white. The interior of the bacterial R-hydratase is completely filled, while the eukaryotic hydratases from yeast and fruit fly have a variable amount of space to accept a wider range of substrates. Both of the eukaryotic hydratases possess only one active site.

5.4 Kinetic properties of *Dm*MFE-2 in vitro

The kinetic properties of the full-length enzyme were compared with those of the enzymatic domains expressed as separate polypeptides using three different substrates from C4 to C10. According to the $K_M$ and $k_{cat}$ values, *Dm*MFE-2 seems to prefer C10 over the short chain substrates. The most significant observation is that the enzyme catalysis occurs just as efficiently regardless of whether the dehydrogenase and hydratase activities are part of the same polypeptide or function as stand-alone proteins. The results from the kinetic assays have been summarized in table 2.
Table 2. Comparison of kinetic properties of DmMFE-2 and the separate enzymatic domains. The separate monofunctional polypeptides were used in a 1:1 ratio.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>C4</th>
<th>C6</th>
<th>C10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full-length DmMFE-2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{\text{max}}$ ($\mu$mol min$^{-1}$ mg$^{-1}$)</td>
<td>1.07 ± 0.3</td>
<td>15.0 ± 1.5</td>
<td>31.4 ± 0.5</td>
</tr>
<tr>
<td>$K_M$ ($\mu$M)</td>
<td>85.3 ± 38.5</td>
<td>66.7 ± 11.8</td>
<td>1.14 ± 0.06</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>0.4</td>
<td>210</td>
<td>1100</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$</td>
<td>0.005</td>
<td>3.2</td>
<td>970</td>
</tr>
<tr>
<td><strong>DmDH + DmH2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{\text{max}}$ ($\mu$mol min$^{-1}$ mg$^{-1}$)</td>
<td>0.36 ± 0.03</td>
<td>9.6 ± 2.6</td>
<td>56.0 ± 2.1</td>
</tr>
<tr>
<td>$K_M$ ($\mu$M)</td>
<td>83.5 ± 15.1</td>
<td>70.7 ± 33.0</td>
<td>1.2 ± 0.15</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>7.5</td>
<td>200</td>
<td>1170</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$</td>
<td>0.09</td>
<td>2.8</td>
<td>980</td>
</tr>
</tbody>
</table>

5.5 Rescue experiments in vivo using DmMFE-2 in a yeast mutant devoid of endogenous MFE-2

The ability to use oleic acid as a sole carbon source was tested with selected yeast strains on thin YP-plates supplemented with 0.125% oleic acid and 0.2% D+ glucose. The formation of a clear halo around yeast colonies was monitored. As expected, the two negative controls were not able to use oleic acid for growth. However, clear halos were formed around the yeast colonies of the positive controls during the experiment. The growth of the MFE-2 deletion strain $\Delta$fox2 transformed with pYE352::DmMFE-2 was then compared to the negative and positive controls. The results, presented in figure 13, show that the Drosophila MFE-2 complements the yeast MFE-2 deletion strain well.
Fig. 13. Activity of \textit{DmMFE-2} \textit{in vivo}: the enzyme is able to rescue the growth of the yeast MFE-2 deletion strain, $\Delta$fox2. 1. $\Delta$fox2 (halo-), 2. $\Delta$fox2 + pYE352::CTA1 (halo-), 3. wt (halo+), 4. $\Delta$fox2 + pYE352::ScFOX2 (halo+), 5. $\Delta$fox2 + pYE352::\textit{DmMFE-2} (halo+). Samples 1. and 2. do not display the formation of a halo due to the lack of a functional fox2 gene. Samples 3.-5. are able to use oleic acid as a carbon source, which is indicated by the clearing of the plates around the yeast colonies.

5.6 Investigation of heteromeric complex formation of separate domains of \textit{DmMFE-2} \textit{in vitro} by static light scattering

Previous structural studies have shown that MFE-2 forms dimers, (Haapalainen \textit{et al.} 2003, Koski \textit{et al.} 2004). Since the domains retain their enzymatic activities when expressed and purified as separate polypeptides, static light scattering was performed to investigate whether the separately purified enzymatic domains would form heteromeric complexes \textit{in vitro}. In the experiment, four different samples were processed identically to obtain easily comparable data. The static light scattering of the full-length protein was compared to the signals of individual domains alone and also to a mixed solution of the two catalytic domains that had been purified as separate polypeptides. The results clearly indicate that the dehydrogenase and hydratase subunits of \textit{DmMFE-2} do not assemble together in the tested conditions as summarized in table 3. Small-angle
X-ray scattering experiments confirm the result from SLS (unpublished data). However, the data do not exclude the existence of weak interaction forces between the catalytic domains.

Table 3. The results from static light scattering experiments. The catalytic domains of DmMFE-2 purified as separate polypeptides do not form heteromeric complexes in vitro as indicated by the absolute molecular masses. The monodispersity of the samples is very high. The numbers in parenthesis indicate the error in the measurement automatically defined by the Astra software (Wyatt Technology Corp.).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Polydispersity (Mw/Mn)</th>
<th>Molecular mass (g/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmMFE-2</td>
<td>1.000 (0.1%)</td>
<td>$1.22 \times 10^5$ (0.1%)</td>
</tr>
<tr>
<td>DmIDH</td>
<td>1.000 (0.1%)</td>
<td>$6.16 \times 10^4$ (0.2%)</td>
</tr>
<tr>
<td>DmH2</td>
<td>1.000 (0.6%)</td>
<td>$6.93 \times 10^4$ (0.4%)</td>
</tr>
<tr>
<td>DmIDH + DmH2</td>
<td>1.000 (0.3%)</td>
<td>$6.46 \times 10^4$ (0.2%)</td>
</tr>
</tbody>
</table>

5.7 SAXS data collection and the low resolution models of fruit fly and human MFE-2

Small-angle X-ray scattering data were collected from both human and fruit fly purified recombinant protein samples of high homogeneity at X33 beamline in DESY, Hamburg. Data were obtained from various preparations of DmMFE-2 and HsMFE-2, including the full-length protein, separate catalytic domains as individuals and the catalytic domains in the same solution, but as separate polypeptides. The SAXS data collection with averaged statistics from automatic data processing (Petoukhov et al. 2007) is summarized in table 4.
Table 4. Small-angle X-ray scattering data collection and data refinement from recombinant protein samples of MFE-2. The data allow direct comparison between the samples of radius of gyration ($R_g$), $I_0$, maximum size of the particle ($D_{\text{max}}$) and particle volume. The molecular mass is derived from the direct comparison of $I_0$ values between the BSA (66 kDa) reference solution and the actual samples of MFE-2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>MM (kDa)</th>
<th>$R_g$ (nm)</th>
<th>$I_0$</th>
<th>$D_{\text{max}}$ (nm)</th>
<th>Volume (nm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmMFE-2</td>
<td>80 ± 20</td>
<td>3.6 ± 0.1</td>
<td>90.5</td>
<td>12 ± 1</td>
<td>200</td>
</tr>
<tr>
<td>DmDH</td>
<td>40 ± 10</td>
<td>2.7 ± 0.1</td>
<td>47.3</td>
<td>9 ± 1</td>
<td>110</td>
</tr>
<tr>
<td>DmH2</td>
<td>40 ± 10</td>
<td>2.9 ± 0.1</td>
<td>48.3</td>
<td>9 ± 1</td>
<td>110</td>
</tr>
<tr>
<td>DmDH + DmH2</td>
<td>40 ± 10</td>
<td>2.8 ± 0.1</td>
<td>48.8</td>
<td>10 ± 1</td>
<td>100</td>
</tr>
<tr>
<td>HsMFE-2</td>
<td>100 ± 20</td>
<td>4.5 ± 0.1</td>
<td>129.0</td>
<td>15 ± 1</td>
<td>330</td>
</tr>
<tr>
<td>HsDH</td>
<td>60 ± 15</td>
<td>3.0 ± 0.1</td>
<td>76.5</td>
<td>10 ± 1</td>
<td>140</td>
</tr>
<tr>
<td>HsH2</td>
<td>40 ± 10</td>
<td>2.9 ± 0.1</td>
<td>48.3</td>
<td>10 ± 1</td>
<td>120</td>
</tr>
<tr>
<td>HsSCP-2L</td>
<td>10 ± 5</td>
<td>1.9 ± 0.1</td>
<td>14.5</td>
<td>6 ± 1</td>
<td>90</td>
</tr>
<tr>
<td>BSA</td>
<td>66</td>
<td>3.0 ± 0.1</td>
<td>75.1</td>
<td>10 ± 1</td>
<td>100</td>
</tr>
</tbody>
</table>

According to the data, HsMFE-2 is clearly the largest particle of all the measured specimens as one would expect. The catalytic domains of human and fruit fly MFE-2 display small differences in molecular masses, but also in other measured parameters. The $I_0$-value for HsDH deviates significantly from the expected value, although this could be explained by minor sample aggregation. Otherwise, the collected data are of high quality and consistent when comparing the samples’ molecular masses and the extracted particle values. The data were compared to the existing crystal structures of MFE-2’s and were used for calculating ab initio bead models, as well as for rigid body modeling. The most interesting piece of data in the table is that the fruit fly catalytic domains do not form a stable complex in solution. Neither the measured volume, nor the $D_{\text{max}}$ of the protein complex is comparable with that determined for the full-length enzyme. This agrees well with the results obtained from SLS experiments that clearly indicated that this complex does not form.

The DmMFE-2 SAXS data were directly compared to the theoretical scattering calculated from the monomeric and dimeric forms of the crystal structure. The data do not fit together with the monomeric form of DmMFE-2 that in practice is the content of the asymmetric unit in the crystal structure. However, when comparing the data with the biological assembly, the gained fit is perfect indicating that indeed the protein is dimeric in buffered solutions. The fit between
the small-angle X-ray scattering data from *Dm*MFE-2 and the dimeric crystal structure of the protein is presented in figure 14.

![Figure 14](image)

**Fig. 14. Comparison of measured and theoretical small-angle scattering of full-length *Dm*MFE-2.** The intensity of the trace measures light scattering by the suspended enzyme particles, while the thin light gray line is calculated from the dimeric crystal structure (PDB ID 3OML) using CRYSOL (Svergun et al. 1995). The fit is good and validates the existing structure.

The data were further utilized to create *ab initio* dummy atom models of *Dm*MFE-2. Several models calculated in the P2-symmetry mode using DAMMIN (Svergun 1999) produced the most reasonable models when evaluated by eye. The dimeric nature of the enzyme enabled to introduce the P2-symmetry (local two-fold axis) for the calculations at early stage. The models obtained were averaged to make the final dummy atom model. The model was superimposed using SUPCOMB (Kozin & Svergun 2001) and the figures were drawn with PyMol (DeLano Scientific LLC). The resolution of the method does not allow the location of the catalytic domains to be distinguished, but nevertheless provides a model that is relatively easy to fit with the crystal structure. Again, the high quality of the data is reflected by the fact that the hole situated between the catalytic domains in the crystal structure is also visible in the dummy atom model. The low resolution model of *Dm*MFE-2 and the comparison with the crystal structure is presented in figure 15.
Fig. 15. *Ab initio* model of *DmMFE-2*. (A) The crystal structure of the enzyme (PDB ID 3OML) is presented in dark gray and (B) the dummy atom model as white large beads. (C) Superimposed view of the high and low resolution models. The resolution of the *ab initio* model is limited and therefore it is not possible to distinguish the exact location of the catalytic domains. The hole in the middle of the bead model is larger than that seen in the crystal structure. There are also other regions in the SAXS-model that are not filled by the parts of the high resolution structure. This can be explained by the mobility of the protein parts in solution and only the average positions of the domains can be seen.

The SAXS data from *HsMFE-2* did not allow a direct approach to building the low resolution model of the enzyme. Instead, a multiphase approach was applied in which the data from *DmMFE-2*, *HsDH* and *HsH2* provided additional support for the calculations from the *HsMFE-2* data. This method was applied to the *ab initio* dummy atom model refinement. In rigid body modeling, the theoretical scattering curves of *DmMFE-2* and *HsSCP-2L* were fitted against the scattering data from *HsMFE-2*. As a result, the physical location of the SCP-2L domain of human MFE-2 was identified and models of the full-length enzyme were constructed. The models were obtained independently using different approaches with the same data and thus are represented either as a dummy atom model or a rigid body model, in which the crystal structures have been fitted to the shape of the molecule studied. Despite the differences in the approaches, the two models fit well together when superimposed. The locations of the SCP-2L domains in *HsMFE-2* are very similar to the previously suggested model (Koski *et al.* 2005) in which these domains are situated in the close proximity to the hydratase domains. In contrast to the *DmMFE-2*, the model is lacking the central hole and in general is more extended. This is concordant with the processed data presented in
The current low resolution model of human MFE-2 is presented in figure 16.

Fig. 16. Current model of *Hs*MFE-2 constructed using small-angle X-ray scattering data. (A) Surface presentation of the rigid body model of *Hs*MFE-2. The rings indicate the positions of the three domains in the dimeric enzyme and in addition, the regions linking the hydratase 2 domains to the SCP-2L domains have been included. (B) Dummy atom model of *Hs*MFE-2 presenting the overall shape of the molecule. (C) The rigid body model of *Hs*MFE-2 (dark gray) superimposed on the *ab initio* dummy atom model (white) of *Hs*MFE-2. The fit between the two models built independently is good.
6  Discussion

Expression, purification and crystallization of recombinant DmMFE-2 enabled the structure of full-length MFE-2 to be solved and explored for the first time. The structure revealed the assembly and organization into a ‘double-dimeric’ multifunctional enzyme. In addition, the data obtained from the kinetic experiments together with the structural information question the possible existence of substrate channeling in DmMFE-2. Moreover, the low resolution data from small-angle X-ray scattering of DmMFE-2 provided ab initio models of exceptional quality that are in good agreement with the current crystal structure. These data were subsequently used as a scaffold for building the first model of HsMFE-2, the first SCP-2L-containing MFE-2 structure.

6.1 The structure of DmMFE-2 reveals the assembly of the full-length enzyme

The DmMFE-2 structure solved at 2.15 Å resolution revealed a ‘double-dimeric’ biological assembly that could be understood as a dimer of two dehydrogenase and hydratase 2 dimers or a dimer of the two dehydrogenase/hydratase monomers of the same polypeptide. Since the dimer is formed ‘crosswise’, it is not clear if the chemical reactions are catalyzed by the catalytic domains in the same polypeptide or whether the intermediates also translocate ‘crosswise’. The linker region between the two catalytic domains is short, limiting the domain movements and this is likely one of the factors that initially enabled the packing of the protein into a crystalline form. A second factor favoring the crystallization was probably the lack of an SCP-2L domain. It is clear that the general domain structures of MFE-2s are well-conserved at the domain level, even if there are differences in the amino acid composition. This is not surprising, because the same MFE-2 gene can still be functional when introduced into a different host, such as yeast (Qin et al. 2000).

The absence of the few regions in the overall structure can be partially explained by the lack of ligands as part of the structure. At least the N-terminal parts in the dehydrogenase domain would be more ordered in the presence of the natural cofactor, NAD⁺ as seen in the structure of the rat dehydrogenase domain of MFE-2 (Haapalainen et al. 2003). In addition, acyl-CoAs would stabilize the flexible loops in the structure, but most likely the preferred 3R-hydroxy substrates would not easily be captured in the active sites if the activity of the enzyme was
not reduced. The problem of the missing parts was addressed by co-crystallizing the protein in the presence of NAD\(^+\) and by removing the affinity tags of the recombinant enzyme. In theory, any tag could interrupt ligand access to the reaction sites of target proteins. In addition, the crystallization conditions were changed in an effort to produce a crystal that would allow binding of the ligand. Despite these changes, the same crystal packing, the absence of ligands as well as certain parts of the polypeptide chain were observed in the resulting structures. Most probably substantial modifications of the polypeptide are required to enable the molecules to pack into a different crystal form.

Examination of the hydratase 2 domain of DmMFE-2 and comparison with the other members of the enzyme family reveals small differences between the fruit fly, yeast and bacterial counterparts. The homodimeric R-hydratase from A. caviae (PDB ID 1IQ6) clearly has a more restricted cavity for shorter chain fatty-acyl-CoA binding (Hisano et al. 2003), whereas the yeast hydratase domain of MFE-2 (PDB ID 1PN4) can accept substrates of long chain and beyond. The yeast enzyme lacks the second active site and utilizes this additional space for processing a broader spectrum of various substrates (Koski et al. 2004). In the current structure the particular helix \(\alpha_{12}\), involved in the substrate binding, is of intermediate length and appears to be more restrictive than the corresponding short \(\alpha\)-helix in the yeast homolog. This does not directly correlate with substrate specificity \textit{in vivo}, because DmMFE-2 can complement the yeast MFE-2 deletion strain when using oleic acid as the sole carbon source.

The main forces holding the two monomers together are hydrophobic interactions between the two dimers formed by the catalytic domains. These interactions were already characterized when the single domain structures of MFE-2 were solved (Haapalainen et al. 2003, Koski et al. 2004, Koski et al. 2005). Within a polypeptide, only a few stabilizing non-covalent interactions are found that keep the dehydrogenase and hydratase 2 domains together. An even more interesting feature in the DmMFE-2 structure is the physical hole situated between the two dimers of the catalytic domains. The existence of the hole was also validated by the small-angle X-ray scattering data. The role of the hole is currently unknown and definitely a liganded structure would be required in order to draw any further conclusions. In addition, the surface of the DmMFE-2 dimer contains a significant number of positively charged residues (H/K/R) that potentially could provide an interface for the channeling of acyl-CoA substrates from the hydratase domain to the dehydrogenase. However, when comparing this positively charged surface to other known examples of electrostatic channeling,
like the dihydrofolate reductase-thymidylate synthase (Knighton et al. 1994), it is unlikely that this surface area is large enough to facilitate efficient substrate translocation. A surface presentation of the DmMFE-2 dimer is presented in figure 17.

![Surface presentation of the DmMFE-2 dimer](image)

**Fig. 17.** Surface presentation of the DmMFE-2 dimer. The two monomers are shown in white and dark gray and the positively charged residues have been highlighted in black. The structure has been superimposed with 3R-hydroxydecanoyl-CoA molecules that have been found experimentally (Koski et al. 2004) or in silico to interact with the active site residues of both enzymatic domains (Ylianttila et al. 2006).

### 6.2 The localization and kinetic studies of DmMFE-2

A complementation study with yeast S. cerevisiae devoid of endogenous MFE-2 was done in order to investigate the ability of DmMFE-2 to rescue the yeast growth on oleate media. The experiment would indirectly show if the protein is targeted to peroxisomes in yeast. When comparing the growth of the selected yeast cells, the results clearly demonstrate that indeed DmMFE-2 can rescue the growth of the yeast deletion strain and therefore is targeted into peroxisomes.
In order to investigate the existence of the substrate channeling mechanism, a series of kinetic data were collected using short and medium chain 2E-enoyl-CoAs and the full-length DmMFE-2 as well as the catalytic domains expressed and purified as separate polypeptides. The hypothesis was that substrate channeling in MFE-2 enhances the efficiency of the catalytic reaction. Data comparison clearly indicated that the postulated existence of an electrostatic channeling mechanism in DmMFE-2 is unlikely (or if a channeling mechanism exists, it does not contribute to catalytic efficiency), because the reactions proceeded equally well in both cases regardless of the fatty acyl-CoA chain length. Clearly, this is different from MFE-1 that operates in the same pathway, but in practice these two enzymes are not structurally related (Kasaragod et al. 2010). The results from SLS, which demonstrate that DmMFE-2 does not assemble into a complex when the two catalytic units are introduced in the same solution as separate polypeptides, also does not lend support to the existence of a channeling mechanism. In other words: the interaction between the two different catalytic domains is looser than the interaction of the catalytic units dimerized on their own. It remains to be determined whether a substrate molecule is required to bring the two domains together.

6.3 The current model of HsMFE-2

The existing crystal structures and good quality low resolution small-angle X-ray scattering data allowed the location of the SCP-2L domains in HsMFE-2 to be distinguished, thus revealing the probable structure of the full length enzyme. The SCP-2L domains are situated logically in the vicinity of the hydratase 2 domains (as in the sequence) and the overall shape of HsMFE-2 resembles the shape of the DmMFE-2 dimer. The obtained low resolution model of HsMFE-2 is similar to the previously predicted model that was based on single domain structures of MFE-2 (Koski et al. 2005). The ab initio model of DmMFE-2 is in good agreement with the existing crystal structure (PDB ID 3OML) and even the hole in the biological assembly can be seen in the model. The quality of the raw data was very good and the light scattering of the sample fits perfectly to the theoretical scattering of the DmMFE-2 dimer. Very often good SAXS data equates to the ability of the protein to crystallize readily. This was true in the case of the fruit fly MFE-2. However, when the models were examined further, the hole between the dehydrogenase and hydratase domains seems to be absent in HsMFE-2. This detail could be related to the difference in the linkers that are of
different lengths thus allowing greater movements not only for the SCP-2L domains, but also for the long linker between the catalytic domains, making the possible hole virtually undetectable. In contrast, the lack of NAD$^+$ in the data collection could also explain the extended bumps in the ab initio dummy atom model of HsMFE-2. Since the loops that participate in the binding of the cofactor are not fixed due to the absence of NAD$^+$, it is possible that these loops can be seen in the mobile solution model, but not in the rigid body model, because it is restricted by the static crystal structures of the domains.

The comparison of rudimentary scattering data from human and fruit fly MFE-2s, display clear consistency in terms of the theoretical molecular mass, maximum size of the particle and the overall shape of the molecules. This is a good indication of the success in data collection and especially in sample preparation. The ab initio models of the full length MFE-2s show areas that are not filled by the volumes of the crystal structures. These areas are situated in regions that are known to be flexible and require the presence of cofactors and substrates for stabilization. Crystal structures present a static picture of a molecule, but SAXS data represent a dynamic image and it is likely that additional features in the SAXS model compared to the crystal structure reveal flexible areas.

The low resolution model of HsMFE-2 is sufficient to allow consideration of the possible role of the SCP-2L domain in MFE-2. Since the SCP-2L domain is missing from the fruit fly and yeast enzymes, it is clearly not necessary for the actual enzymatic reactions to occur. The versatile ability of SCP-2L to bind almost any kind of lipids and lipid like molecules does not explain its functional significance in MFE-2. It is not likely that this domain would exist only to serve as a carrier of the short PTS1 signal either, and possible lipid-binding as a requirement for the actual targeting event, is a controversial issue (Lensink et al. 2002, Stanley et al. 2006). The linker regions in human MFE-2 between the catalytic domains and between the hydratase 2 and SCP-2L domains are rather extended providing flexibility for the enzyme. However, the wide-ranging dimerization contacts between the dehydrogenase and hydratase domains do not allow large between-domain movements, but between the hydratase and SCP-2L domain this limitation does not exist. Therefore, the extended connecting loop in HsMFE-2 could possibly bind to or act in concert with the SCP-2L domain when the substrate is introduced to the enzymatic domains. In the first stage, the SCP-2L domain could bind to any of the poorly soluble substrates from the surrounding solvent and feed the substrates to the hydratase domain. After the hydration, the intermediate could be translocated into the second active site for
oxidation with a currently unknown mechanism that could benefit from the assistance of SCP-2L. The role of the domain in the second enzymatic reaction could be simply to provide assistance with the intermediate transfer, especially when that molecule is bulky and very hydrophobic. It is neither clear whether the linker regions are extended enough to allow such movements, nor how exactly the substrates become translocated between the two reaction centers. In any case, this mobility provides a plausible explanation for the non-crystallizability of any SCP-2-containing MFE-2 construct. Naturally, proving the between-domain movements would require liganded crystal structures and additional SAXS data from the different stages of the catalytic reaction.
7 Conclusions

The structure of the full-length DmMFE-2 was solved. This goal was achieved when the structure of DmMFE-2 was verified at 2.15 Å resolution by molecular replacement. Previously, only isolated domain structures of MFE-2 were known. The full-length structure solved in this work provides novel information about the previously unknown assembly of the full-length enzyme and the nature of the linker region between the dehydrogenase and hydratase catalytic domains. DmMFE-2 is able to complement a yeast deletion strain of MFE-2 and therefore can be considered as a peroxisomal protein. Surprisingly, DmMFE-2 does not exploit substrate channeling for efficient catalysis as demonstrated by the comparative kinetic measurements between the full-length enzyme and catalytic units purified as separate polypeptides. The data are supported by the fact that the dehydrogenase and hydratase domains do not form heterogenous complexes in vitro when sharing the same aqueous environment and monitored with SLS or SAXS. The small-angle X-ray scattering data agrees well with the crystal structure proving that the enzyme forms dimers in solution, and also validates the existence of the small hole between the catalytic domains.

A low resolution model of HsMFE-2 was constructed. The collected SAXS data verified the locations of the SCP-2L domains. In addition, an ab initio dummy atom model was constructed based on the data. There are similarities in the shapes between the human and fruit fly enzymes, but also logical differences due to the different domain composition. The current model allows a plausible hypothesis that SCP-2L could have a role in the enzyme catalysis, despite the fact that the enzymatic domains are also active without SCP-2L. A crystal structure would be needed to address this question further, but requires large amounts of purified enzyme for crystallization screening. It would be interesting to determine how the linker region is organized between the dehydrogenase and hydratase domains in HsMFE-2 and whether the linker between the hydratase and SCP-2L domains is long enough to allow movements of the SCP-2L domain between the catalytic domains.

The fact that the catalytic domains seem not to form heteromeric complexes in vitro when purified as stand-alone proteins, does not exclude the existence of weak interaction forces. In the simplest case, the strength of interaction between the different domains of MFE-2 could be studied using surface plasmon resonance technique. The method, however, does not allow distinguishing the locations of the interaction surfaces, and it is rather difficult to control in which
orientations the studied molecules become immobilized on the used chip. To study the dynamics of the SCP-2L domain in relation to the hydratase and dehydrogenase domains is far more tricky. One approach could be to apply nuclear magnetic resonance (NMR), but due to the limitations in the molecular mass of the studied molecules, large macromolecules or complexes are not often measurable. Besides traditional protein crystallography and SAXS, the distance of two different domains in a multifunctional protein or protein complex can be estimated by fluorescence or electron spin resonance (ESR) techniques. When simplified, the latter technique requires the use of labelling the studied macromolecules with unpaired electron spins (spin-label). The changes of the distance between the labelled particles in the magnetic field can be monitored in the change of the ESR spectra. The use of fluorescence techniques in practice can be difficult, because more invasive polypeptide modifications have to be applied in order to get reliable signal.
References


Original articles


Reprinted with permission from (I) PubMed Central and (II) Portland Press Ltd.

Original publications are not included in the electronic version of the dissertation.
566. Helikala, Ritva (2011) The many facets of an inter-organisational information system project as perceived by the actors

567. Niittyvuopio, Anne (2011) Adaptation to northern conditions at flowering time genes in Arabidopsis lyrata and Arabidopsis thaliana


570. Kaarinen, Salla (2011) Space use and habitat selection of the wolf (Canis lupus) in human-altered environment in Finland


574. Petsalo, Aleksanteri (2011) Development of LC/MS techniques for plant and drug metabolism studies

575. Leppäät, Mirva (2011) Successional changes in vegetation and carbon dynamics during boreal mire development


577. Alahuhta, Janne (2011) Patterns of aquatic macrophytes in the boreal region: implications for spatial scale issues and ecological assessment

578. Moody, Gregory (2011) A multi-theoretical perspective on IS security behaviors

579. Karjalainen, Mari (2011) Improving employees’ information systems (IS) security behavior: Toward a meta-theory of IS security training and a new framework for understanding employees’ IS security behavior


581. Nykänen, Kari (2011) Tietoturvakohtausen vaikuttavuuden arviointi yksilön ja organisaation tietoturvaväkeystäytymisestä

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Tatu Haataja

PEROXISOMAL MULTIFUNCTIONAL ENZYME TYPE 2 (MFE-2)

THE CATALYTIC DOMAINS WORK AS INDEPENDENT UNITS