



LUND UNIVERSITY

Structural characterization of proteins to investigate their roles in diseases: Focus on MID & LTA4H

Hasan, Mahmudul

2014

[Link to publication](#)

Citation for published version (APA):

Hasan, M. (2014). *Structural characterization of proteins to investigate their roles in diseases: Focus on MID & LTA4H*. [Doctoral Thesis (monograph), Biochemistry and Structural Biology]. Department of Chemistry, Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Structural characterization of proteins to investigate their roles in diseases:

Focus on MID & LTA4H

Mahmudul Hasan



LUND
UNIVERSITY

Structural characterization of proteins to investigate their roles in diseases: Focus on MID & LTA4h

Copyright © Mahmudul Hasan

Faculty of Science
Department of Biochemistry and Structural Biology
Center for Molecular Protein Science
PO Box 124
22100 Lund
Sweden

ISBN 978-91-7422-355-2

All previously published articles are reprinted with permission from the publisher.

Printed in Sweden by Tryckeriet E-Huset, Lund University
Lund 2014

*To my family & my late
grandfather*

Abstract

Protein molecules are responsible for many biological functions in cells. In order to fulfill their various biological roles, these chain-like molecules must fold into precise three-dimensional shapes. The knowledge of accurate molecular structures is a prerequisite for rational drug design and for structure based functional studies. Getting structural information of proteins can be a very difficult task, especially when it comes to high resolution.

Moraxella catarrhalis is widely recognized human-restricted gram-negative bacterium for which it has become clear that it is a true pathogen of both the upper and lower respiratory tract. After *Haemophilus influenzae* and *Streptococcus pneumoniae*, it is the third most common cause of otitis media in children. The bacterium can directly stimulate B-cells without any recognition of T-cells and it can therefore be classified as a T-cell independent antigen. The mitogenic activity of *Moraxella catarrhalis* is performed by a 2139 residue long outer membrane protein MID. An IgD binding domain (MID⁹⁶²⁻¹²⁰⁰) has been described and the colonization to human respiratory tract cells is mediated by a 150-residue adhesin domain (MID⁷⁶⁴⁻⁹¹³). Small-angle x-ray (SAXS) studies on the IgD binding domain showed that this domain has an elongated 3-fold organization and that there is the presence of unordered/flexible structures. Circular dichroism (CD) data and prediction of secondary structure for both of the domains indicated the presence of large amounts of (~33%) β -sheet and ~10% α -helix content. Native datasets for MID⁹⁶²⁻¹²⁰⁰ to 2.3 Å resolution and for MID⁷⁶⁴⁻⁹¹³ to 2.7 Å resolution are collected and processed.

Vertebrate leukotriene A₄ hydrolases are zinc metalloenzymes with an epoxide hydrolase and aminopeptidase activity belonging to the M1 family of aminopeptidases. The human enzyme produces LTB₄, a powerful mediator of inflammation and is implicated in a wide variety of rheumatoid diseases. The yeast homolog *scLTA4H* contains only a rudimentary epoxide hydrolase activity and was shown to undergo a large conformational change upon binding of the inhibitor bestatin. In SAXS studies *XiLTA4H* shows a more compact form upon bestatin binding, but *humLTA4H* did not. It was confirmed that the LTA4H from *Xenopus* is a dimer and that it seems to contract in size upon bestatin binding. In contrast the human enzyme does not show any major difference in SAXS patterns upon inhibitor binding and it seems therefore that this enzyme does not display larger conformational changes.

Keywords: *Moraxella* IgD binding protein, adhesin, leukotriene A₄ hydrolase, small angle x-ray scattering, crystallography, circular dichroism

LIST OF PAPERS

This thesis is based upon the following papers, which are referred to in the text by their Roman numerals

Paper I

Mahmudul Hasan, Kristian Riesbeck & Marjolein Thunnissen. SAXS measurements and Preliminary crystallographic investigations show an elongated three-fold organization of MID⁹⁶²⁻¹²⁰⁰.

Manuscript to be submitted.

Paper II

Mahmudul Hasan, Vivek Srinivas & Marjolein Thunnissen. Crystallization, circular dichroism and SAXS studies on the adhesin domain MID⁷⁶⁴⁻⁹¹³ from *Moraxella catarrhalis*.

Manuscript.

Paper III

Helgstrand C, **Hasan M**, Uysal H, Haeggström JZ & Thunnissen MMGM. A Leukotriene A4 Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding.

J. Mol. Biol. (2011) 406, 120–134.

Paper IV

Stsiapanava A, Tholander F, Kumar R, Qureshi A, Niegowski D, **Hasan M**, Thunnissen M, Haeggström J. Z. and Rinaldo-Matthis A. Product Formation Controlled By Substrate Dynamics In Leukotriene A4 Hydrolase.

Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics (2014), 1844, 439–446.

Paper V

Mahmudul Hasan, Agnes Rinaldo-Matthis, Jesper Z. Haeggström & Marjolein Thunnissen. SAXS studies on leukotriene A4 hydrolases reveal conformational differences upon inhibitor binding.

Manuscript.

Additional papers from the author not included in the thesis

Paper I

Different methods of HIV vaccination, efficacy and their delivery system: A review.

Mahmudul Hasan & Mahbub-E-Sobhani. (2003).

Journal of Medical Sciences. Vol 3 (1): 1-23.

Paper II

Transmission, biochemical manifestation and CD4+ cell count of HIV: A review.

M.M. Ahasan, M.M. Billah, **M.M. Hasan**, K.M.D. Islam & J.A. Shilpi. (2004).

Pakistan Journal of Biological Sciences 7 (2): 292-300.

Paper III

Creation of primate-specific fusion genes through segmental duplications.

Daniel Svenback, **Mahmudul Hasan** & Patrik Medstrand.

Manuscript.

Author's contribution to the papers

Paper I

MH has done the cloning of **MID**⁹⁶²⁻¹²⁰⁰ and developed the purification protocol. **MH** has optimized the stability, started crystallization trials and optimized the crystallization conditions. **MH** collected the crystal diffraction data at ESRF, France and processed the data. **MH** has collected SAXS data, done initial processing and built the *ab initio* (DAMMIN) model. **MH** purified the protein, collected and processed CD data and wrote the initial manuscript.

Paper II

MH has developed the purification protocol and optimized the stability of **MID**⁷⁶⁴⁻⁹¹³. **MH** started crystallization trials and optimized the crystallization conditions. **MH** collected the crystal diffraction data at the MAX IV laboratory, Sweden and processed part of the data. **MH** has collected SAXS data, done initial processing and further analyzed the data. **MH** purified the protein, collected and processed CD data and wrote the initial manuscript.

Paper III

MH has expressed, optimized and purified the *S. cerevisiae* **LTA4H** and then crystallized/optimized the inhibitor bound form. **MH** participated until data collection at MAX IV laboratory, Lund, Sweden and took part in writing the manuscript.

Paper IV

MH has further purified the *Xenopus* **LTA4H** by SEC and collected data on inhibitor bound form at BM29, ESRF, France. **MH** did the initial processing of the data and build the *ab initio* model (GASBOR). **MH** took part in writing the manuscript.

Paper V

MH has purified *scLTA4H*, partially purified *humLTA4H* & *xLTA4H* and optimized the protocol for all three proteins for SAXS data collection. **MH** has collected, initially processed and analyzed the SAXS data. **MH** wrote the initial manuscript.

Table of Contents

List of papers	VII
Authors contribution to the papers	IX
Abbreviations	XII
1. Introduction	1
1.1. <i>Moraxella</i> Immunoglobulin-D (IgD) binding protein (MID)	1
1.1.1. MID is a trimeric autotransporter	1
1.1.2. MID binds to IgD	3
1.1.3. MID is also an adhesin	4
1.1.4. MID contains Hep-Hag repeats	5
1.1.5. MID contains a membrane anchoring domain and fibrils.....	8
1.1.6. MID ⁹⁶²⁻¹²⁰⁰ purification	11
1.2. Leukotriene A4 hydrolases (LTA4H)	12
1.2.1. Properties of LTA4H	13
1.2.2. Structure of Human LTA4H (<i>hum</i> LTA4H).....	15
1.2.3. <i>Saccharomyces cerevisiae</i> LTA4H (<i>sc</i> LTA4H)	16
1.2.3.1. <i>Sc</i> LTA4H purification	18
1.2.4. <i>Xenopus laevis</i> LTA4H (<i>xl</i> LTA4H)	18
2. Crystallography	20
2.1. Growing protein crystals	20
2.2. Symmetry	22
2.3. Data collection	23
2.4. Diffraction Resolution	25
2.5. Mosaicity and mechanical stress	25
2.6. Data quality	26
2.7. B Factor	27
2.8. The Phase Problem	28
2.8.1. Molecular Replacement	28
2.8.2. Isomorphous replacement	29
2.8.3. Multiple/Single Anomalous Dispersion	30

3. Small angle x-ray scattering	33
3.1. Sample preparation and measurement	34
3.2. Data processing and analysis	35
3.2.1. Guinier approximation	36
3.2.2. Pair distance distribution functions	37
3.2.3. Flexibility assessment by Kratky plot	38
3.2.4. Shape estimation	38
3.2.4.1. DAMMIN and GASBOR	39
3.2.5. Comparison of SAXS data with a crystal structure	40
4. Circular dichroism spectroscopy	41
4.1 Synchrotron CD Spectroscopy	43
4.2 Data Processing	43
4.2.1. Delta Epsilon	43
4.2.2. Mean Residue Ellipticity (MRE)	44
4.3. Data analysis and analysis software	44
5. Summary of specific results obtained from projects	45
6. Conclusion and future prospects	50
7. Acknowledgements	52
8. References	53

ABBREVIATIONS

5-LO	5- Lipo-Oxygenase
AA	Arachidonic acid
CD	Circular Dichroism
GST	Glutathione S-transferase
Ig	Immunoglobulin
LTA ₄	<i>5S-trans-5,6-oxido-7,9-trans-11,14-cis</i> eicosatetraenoic acid
LTA ₄ H	Leukotriene A ₄ hydrolase
LTB ₄	<i>5S,12R-dihydroxy-6,14-cis-8,10-trans</i> -eicosatetraenoic acid
MAD	Multiple anomalous dispersion
MID	<i>Moraxella</i> IgD binding protein
NMR	Nuclear Magnetic resonance
OCA	Oligomeric coiled-coil adhesin
OMP	Outer membrane protein
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein Data Bank
SAD	Single anomalous dispersion
SAXS	Small Angle X-Ray Scattering
SC-57461A	(3-[Methyl[3-[4-(phenylmethyl)phenoxy]propyl]amino]propanoic Acid HCl)
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SRCO	Synchrotron radiation circular dichroism
TAA	Trimeric autotransporter adhesin
TEV	Tobacco etch virus
TPSS	Two-partner secretion system

1.INTRODUCTION

1.1. *MORAXELLA* IMMUNOGLOBULIN-D (IgD) BINDING PROTEIN (MID)

Moraxella catarrhalis is a gram-negative diplococcus bacterium. It was discovered about a century ago and for a long time thought as being a rather harmless human commensal bacterium. However, over the last couple of decades it has emerged as a true human pathogen, mostly associated with respiratory tract infections. After *Haemophilus influenzae* and *Streptococcus pneumoniae* it is the third most common cause of otitis media in children and an important cause of upper and lower respiratory tract infections. It is found to be the second most common cause of chronic obstructive pulmonary disease after *H. influenzae* [1]. In the USA, *M. catarrhalis* is the 3rd most common cause of acute otitis media and sinusitis in children. In adults and the elderly it causes lower respiratory tract infections. Almost all *M. catarrhalis* strains contain genes expressing β -lactamases [2] and more than 95% of all clinical isolates are penicillin resistant [3,4] and this is a sobering statistic and a cause of worry. An overview of *M. catarrhalis* and its pathology can be found in Murphy & Parameswaran, 2009 [5].

1.1.1. MID is a trimeric autotransporter

One of the key virulence factors of *M. catarrhalis* is a 200 kDa protein named as *Moraxella* IgD binding protein (MID). It belongs to the trimeric autotransporter family [6] and is considered a superantigen [7]. Autotransporter proteins possess a range of virulence properties such as adherence, aggregation, invasion and biofilm formation. Bacteria produce sticky fibers, termed adhesins, which mediate bacterial attachment to biotic or abiotic surfaces (e.g. host cells or catheters) and are essential for colonization of their ecological niche. Gram-negative bacteria express a large repertoire of adhesins, including fimbrial and autotransporter (AT) adhesins.

Trimeric autotransporter adhesins (TAAs) form one of the many families of bacterial surface proteins and are representatives of a group of non-fimbrial adhesins, which are found in proteobacteria. In this group of proteins there are many prominent pathogenic determinants including *M. catarrhalis* UspA1 and A2, *Yersinia* YadA, *Haemophilus*

influenzae Hia and Hsf, and more. Typical autotransporters have three different segments in their chain. The first part (i) is a N-terminal signal peptide, the second an N-terminal passenger domain that confers the protein's biological function and contains in general stalk regions, binding or adhesin domains, and (iii) a C-terminal translocator domain responsible for pore formation, which is required for translocation of the protein across the outer membrane. Autotransporter proteins are defined by the ability to drive their own secretion across the bacterial outer membrane. The C-terminus translocator domain forms a pore in the outer membrane and the N-terminus parts are pulled through this pore either by threading (N-terminus first) or a hairpin transport (C-terminal part of the passenger domain first). After the translocation the protein is folded into its functional entities (for schematic overview, see figure 1). Different autotransporter systems have been described and they are classified into monomeric and trimeric autotransporters, and two-partner secretion systems (TPSS) [8].

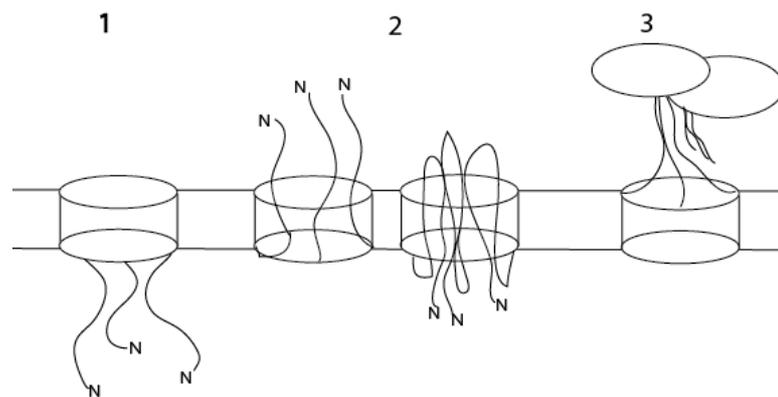


Figure 1: Schematic presentation of the translocation of MID over the membrane. Step1: Pore formation by the translocator domain. Step 2: Threading (N-terminus first, left) or hairpin transport (C-terminal part first, right) of the N-terminal part. Step 3: folding of functional domains.

The subcellular localization of the recombinant protein showed that MID was equally located in the cytoplasmic and the membrane fractions, but not in the periplasmic space.

1.1.2. MID binds to IgD

M. catarrhalis can directly stimulate B-cells without any recognition of T-cells and it is therefore classified as a T-cell independent antigen or superantigen. In contrast with gram-positive bacteria, only very few gram-negative bacteria have the capacity to bind Ig in a non-immune fashion. Since the soluble and membrane bound forms of IgD have many similarities it is likely that the mitogenic signal is carried through the membrane bound form of IgD on the surface of the B-cell. Binding is proposed to be to the CH1 domain of IgD. The OMP that has been identified [9] to be responsible for the IgD binding of *M. catarrhalis* was the Moraxella IgD-binding protein MID [10], also known as HAG [11].

From MID-deficient Moraxella mutants it could be seen that MID is the only outer membrane protein responsible for IgD binding in *M. catarrhalis*. This protein binds specifically to both soluble as well as membrane bound IgD on B-cells. However MID does not bind to other immunoglobulins such as IgG, IgM, IgA, or IgE myeloma proteins. Apart from IgD binding, MID is also an important adhesin capable of agglutinating human erythrocytes and it binds to type II alveolar epithelial cells. It is also responsible for the adherence of *M. catarrhalis* to air-liquid interface cultures of normal human bronchial epithelium exhibiting mucociliary activity [12]. MID does not share any homology with other Ig binding proteins including protein D from *H. influenzae*.

MID is thus a multifunctional protein and apart from the binding to IgD it has also adherence properties to different cell types. Even though other adhesion proteins are present in *M. catarrhalis*, deletion of MID (Hag) caused an overall decrease of adherence [13]. At present, two major functional domains: MID⁹⁶²⁻¹²⁰⁰ (binds to IgD) and MID⁷⁶⁴⁻⁹¹³ (binds to human lung epithelial cell) have been described, and several other regions have been pointed out as significant for adherence.

The capacity for IgD binding of MID has been located in an IgD binding domain of 239 amino acids (MID⁹⁶²⁻¹²⁰⁰) [14]. The truncated MID⁹⁶²⁻¹²⁰⁰ is efficiently attracted to a standard IgD serum and purified myeloma IgD but not to IgG, IgM, or IgA myeloma sera. Early ultracentrifugation and gel electrophoresis experiments suggested that

MID⁹⁶²⁻¹²⁰⁰ is a tetramer and that this multimeric form binds to IgD 20-fold more efficiently than the monomeric form. The multimeric state of the protein would be in contrast with all other members of the trimeric autotransporter family and later studies showed that while MID is clearly an oligomer, it was unclear whether there are three or four copies in the assembly [6]. MID⁹⁶²⁻¹²⁰⁰ itself can activate human B-cells independent from T-cells, while in contrast a mutated form of MID lacking the IgD binding domain is not able to activate B-cells [15]. Results obtained by introducing five amino acids randomly into MID⁹⁶²⁻¹²⁰⁰ using transposons suggested that α -helical structures might be important for the IgD binding. As compared with immunoglobulin (Ig)-binding regions derived from other Ig-binding proteins, MID⁹⁶²⁻¹²⁰⁰ is relatively large (239 amino acids). For example, the Ig-binding domains of the streptococcal protein A and protein L from *Peptostreptococcus magnus* are 58 and 72–76 amino-acid residues, respectively [16].

1.1.3. MID is also an adhesin

The second functional domain MID⁷⁶⁴⁻⁹¹³ is capable of binding to human alveolar epithelial cells and causes agglutination of human erythrocytes [13]. It was shown that antibodies generated against the MID⁷⁶⁴⁻⁹¹³ portion of the total protein could effectively inhibit adherence to the human alveolar epithelial cells, both in in vitro as well as in mouse model studies [17]. Furthermore, since the homology of this segment is very high between the different strains of *Moraxella* it makes this area an attractive target for vaccine-development.

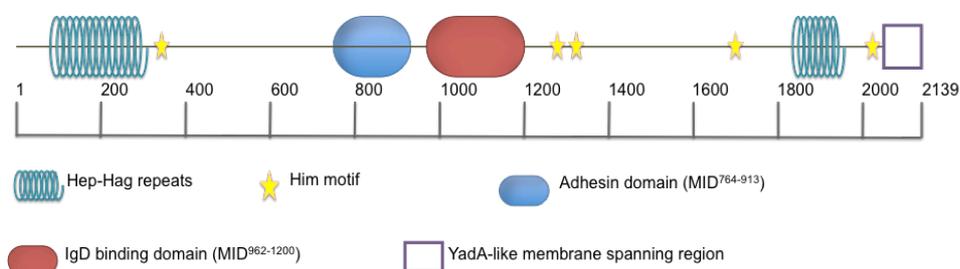


Figure 2: Whole sequence of MID¹⁻²¹³⁹ with two major domain adhesive MID⁷⁶⁴⁻⁹¹³ (colored in blue) and IgD-binding domain MID⁹⁶²⁻¹²⁰⁰ (colored in red) and repeat regions are marked.

Other functionalities of MID are located in residues 385-705 of MID as these are important for adherence to middle ear and NCIH292 lung cells while the area between residues 706 and 1194 was found to be important for the adherence to collagen [12].

1.1.4. MID contains Hep-Hag repeats

For most members of the trimeric autotransporter family, the binding domains itself are located at the N-terminal end of the sequence. These members are in general between 45 and 80 kDa big. MID belongs to a subclass within the family that is larger and as mentioned above it has a size of around 200 kDa. In the case of MID the two functional domains are located in the middle of the sequence while so-called Hep-Hag repeats that are characteristic for a β -roll structure can be found in both the N-terminal and C-terminal regions (see figure 2 and figure 3).

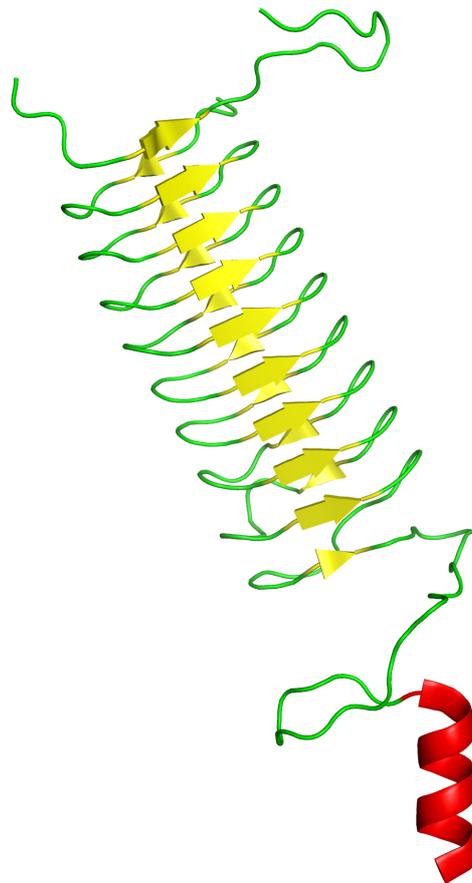


Figure 3: Hep/Hag repeats as seen in the structure of the collagen-binding domain of YadA, pdb code: 1p9h [18].

The Hep/Hag motif is a seven-residue repeat that makes up the majority of the sequence of a family of bacterial hemagglutinins and invasins. As many as ten copies of the Hep/Hag motif can be present in these proteins. A family of immunodominant antigens identified in *Burkholderia mallei* (*Pseudomonas mallei*) and *Burkholderia pseudomallei* (*Pseudomonas pseudomallei*) is known as Hep_Hag autotransporter (BuHA) proteins, and they have been found to share protein domain architectures with hemagglutinins and invasins [19].

Several short hemmagglutinin motifs (HIM) can also be located in the sequence. In many other trimeric autotransporter family members it is these Hep-Hag repeats and HIM-motifs that are important for the adhesive properties, however in the case of MID, the adhesive domains are not located in these Hep-Hag regions. Thus MID differs in its structure-function relationships from those of other trimeric autotransporters.

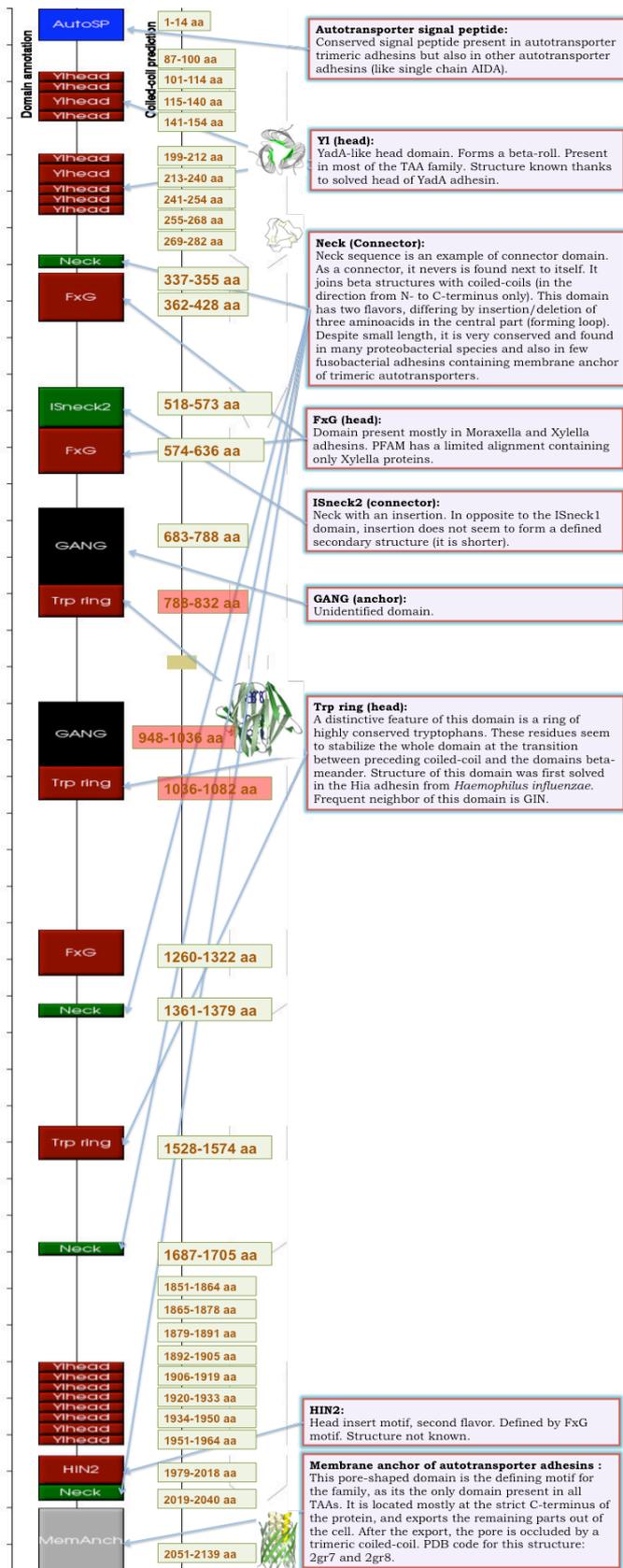


Figure 4: Predicted possible domain in MID¹⁻²¹³⁹ protein generated by dATAA server [20, 21]. The location two domains presented in paper I & II are colored in red.

1.1.5. MID contains a membrane anchoring domain and fibrils

In gram-negative bacteria, the cell envelope is composed of two hydrophobic membranes separated by a hydrophilic space and can be named as inner and outer membranes [22]. Like many other trimeric autotransporters [23, 24], MID forms an extended lollipop type structure and becomes anchored on the outer membrane [6] by a β -sheet porin like anchoring domain. The structure of this membrane-spanning domain has been solved for the Hia transporter and is shown in figure 5.

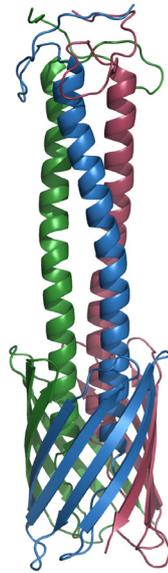


Figure 5: Porin-like membrane spanning domain and part of the stalk region from the Hia autotransporter [25]. Figure made with Pymol and coordinates from PDB file 3emo.

The structure of the fiber is not homogenous – it is a mixture of globular domains and coiled coils. On a sequence level, they have lots of internal repeats, heavily biased residue composition. Their domain composition and architecture varies by protein. The only conserved part in all TAAs is the porin-like autotransport domain. From electron microscopy work it has been seen that MID is a fibrous protein with a globular head like the other trimeric transporters, however by using immuno electron microscopy it could be observed that MID is folding back upon itself (figure 6) [6]. The protein is membrane anchored in the C-terminal region and has a unique double folded stalk protruding approximately 100 nm in length. MID is about double the size of YadaA and its stalk is longer and more fibril-like.

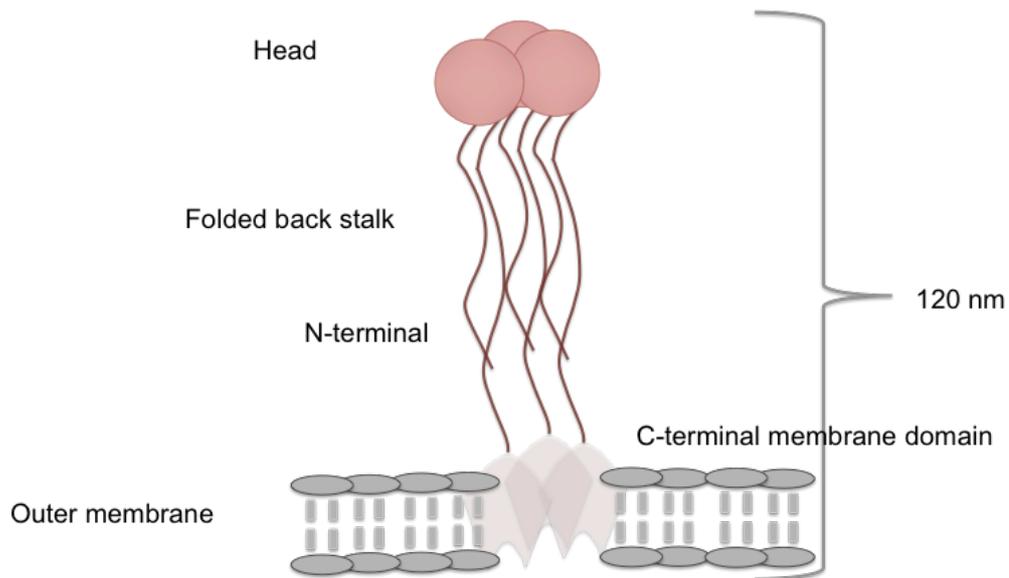


Figure 6: Fold prediction of MID¹⁻²¹³⁹ protein depicted as cartoon representation based on electron microscopy experiments [6].

The lollipop head forms the bulkier part of the structure and is responsible for the functional properties of the proteins such as binding to proteins of the extra-cellular matrix or autoagglutination and for MID it contains the functional domains for IgD binding and adhesion. It is resided at the very tip of the protein, distal to the cell membrane. Several structures of head domains have been solved by X-ray crystallography e.g. those from YadA, Hia BadA and BpaA. All of these show a trimeric arrangement but each of these head domains has a different arrangement and can include different folds (see figure 7 for an overview). Common motifs found in the different head domains include Trp-ring [6], GIN, and neck motifs while Ylhead, FGG, HANS and HIN2 motifs are also observed (for review see Łyskowski et al., 2011) [26]. Threading fold recognition has shown that several of these motifs can also be found for the full MID protein (figure 4).

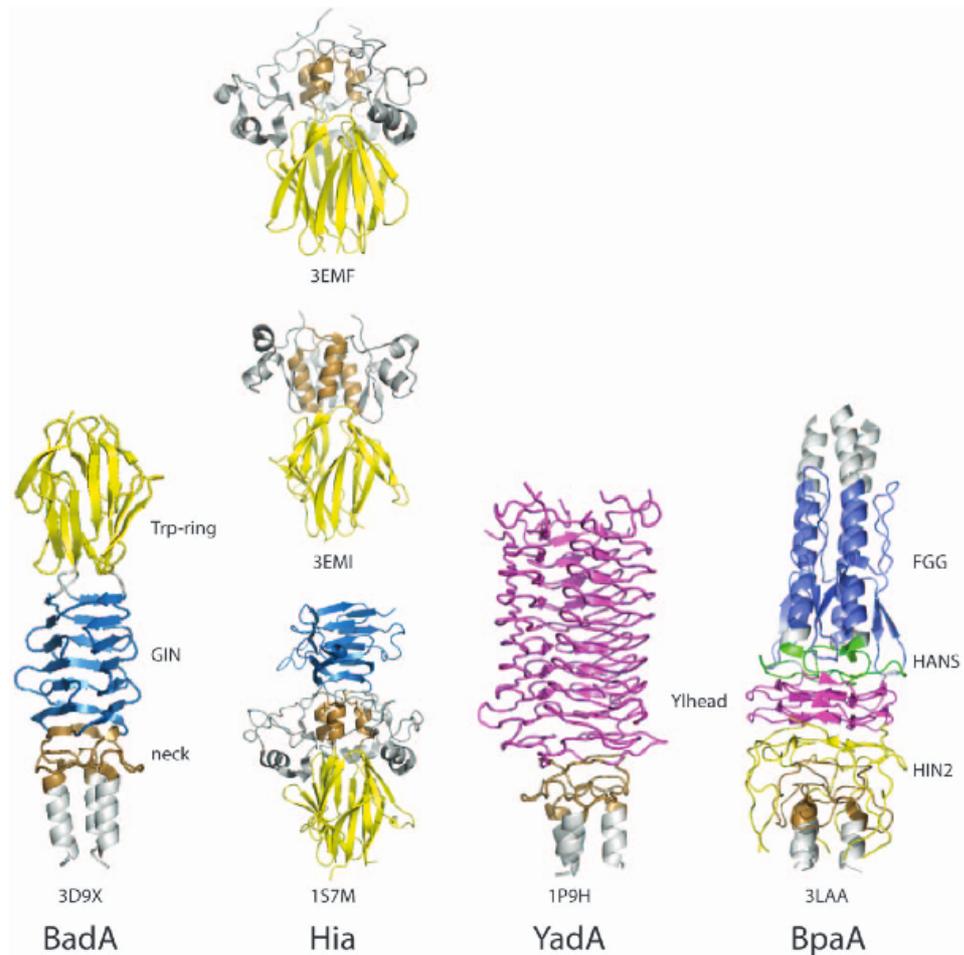


Figure 7. Different structures of head domains from trimeric autotransporters. Figure adapted from Edwards *et al.*, 2010. Trimeric structures are shown for the TAA head domains of BadA from *Bartonella henselae* (PDB ID 3d9x), HiaBD2 (PDB ID 3emf), KG1-W3 (PDB ID 3emi), and HiaBD1 (PDB ID 1s7m), from *Haemophilus influenzae*, YadA from *Yersinia enterocolitica* (PDB ID 1p9h), and BpaA from *B. pseudomallei* (PDB ID 3laa) [27].

A major global issue is the rise of antibiotic-resistance in bacterial pathogens or ‘super-bugs’. In response, new drug targets need to be identified and new vaccines need to be developed. Fundamental aspects of bacterial cell biology need to be understood if we are to take creative approaches to defeat these ‘new’ bacterial pathogens. The MID protein could be a possible good candidate for these kind of developments as there is a high homology between the genes from different clinical strains (sequence comparison of MIDs from five strains show that the overall identity ranged between 65.3% and 85.0%, while some regions showing an identity up to 97%) [28] and furthermore experiments with antibodies generated against the different domains have been

successful. Structural studies of MID would further enhance the knowledge on this atypical trimeric autotransporter.

1.1.6. MID⁹⁶²⁻¹²⁰⁰ purification

MID⁹⁶²⁻¹²⁰⁰ was cloned with a cleavable GST and His6-tag. Detailed composition of buffer used is mentioned in paper I. In figure 8 size exclusion chromatography is not shown. This method helps to produce cleaner protein. Similar principle is also used in *Saccharomyces cerevisiae* LTA4H purification (paper III & V).

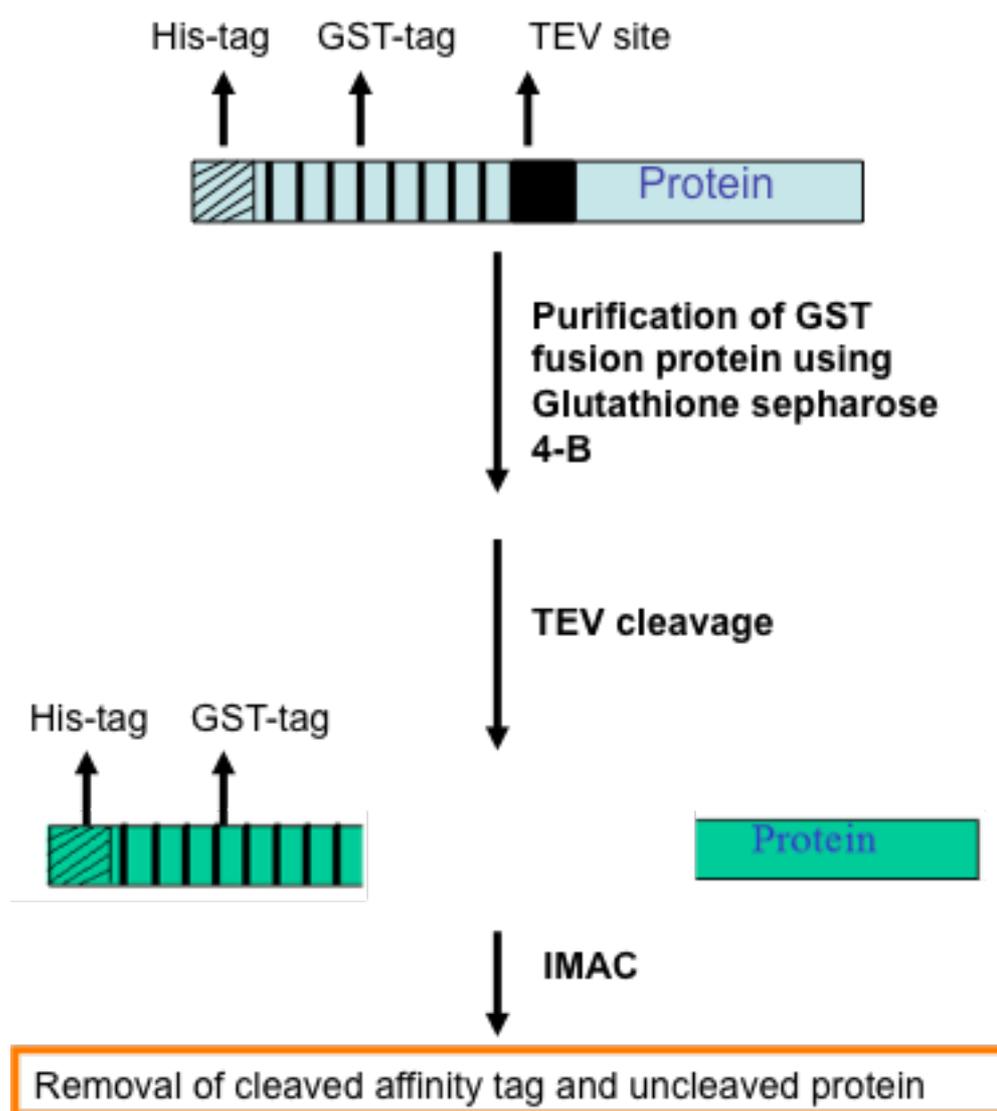


Figure 8: Schematic representation of steps involved in MID⁹⁶²⁻¹²⁰⁰ purification.

1.2. LEUKOTRIENE A₄ HYDROLASES (LTA₄H)

Inflammation is a pathophysiological mechanism by which an organism clears and rebuilds a site of infection or trauma and triggers the synthesis of an array of molecules among which the class of molecules known as leukotrienes (LTs) [29]. Leukotrienes are an important family of eicosanoid lipid mediators derived from the metabolism of arachidonic acid (AA) [30]. The biosynthetic pathway starts with a receptor-mediated influx of calcium ions that cause translocation to the cell membrane of a phospholipase enzyme, cytosolic phospholipase A₂, which selectively cleaves arachidonic acid from the perinuclear cell membranes.

Being triggered by increased calcium level, arachidonic acid is converted sequentially to 5-hydroperoxyeicosatetraenoic acid and then to leukotriene A₄ (5,6-oxido-7,9-trans-11,14-cis-eicosatetraenoic acid) by a catalytic complex consisting of 5-lipoxygenase (5-LO) and the 5-LO activating protein (FLAP). FLAP plays a critical role for leukotriene synthesis by binding to arachidonic acid to present it to 5-LO. The human zinc metalloenzyme leukotriene A₄ (*hum*LTA₄) hydrolase (LTA₄H) catalyzes the hydrolysis of the unstable epoxide LTA₄ (5*S*-*trans*-5,6-oxido-7,9-*trans*-11,14-*cis*-eicosatetraenoic acid) into LTB₄ (5*S*,12*R*-dihydroxy-6,14-*cis*-8,10-*trans*-eicosatetraenoic acid). Leukotriene A₄ is also unstable and may get hydrolyzed unspecifically in solution. It can be transformed specifically through the action of the enzyme leukotriene A₄ epoxide hydrolase in polymorphonuclear leukocytes into leukotriene B₄, which is involved in eosinophil and neutrophil chemotaxis. LTB₄ is a powerful mediator of inflammation and present in many autoimmune diseases, such as arthritis, psoriasis, irritable bowel syndrome and more.

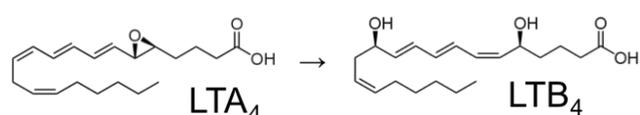


Figure 9: Conversion of LTA₄ into LTB₄ by LTA₄H.

Alternatively, in the presence of *LTC*₄ synthase, a glutathione is adducted to the C6 position of leukotriene A₄ in eosinophils, mast cells, and alveolar macrophages to yield the molecule known as leukotriene C₄ (LTC₄). Once exported from the cytosol to the

extracellular microenvironment, the glutamic acid moiety of LTC₄ is cleaved by glutamyltranspeptidase to form the active entity, leukotriene D₄. Cleavage of the glycine moiety from leukotriene D₄ by a variety of dipeptidases results in the formation of leukotriene E₄. LTC₄, leukotriene D₄ and leukotriene E₄ are all known as the cysteinyl leukotrienes since each one contains a cysteine. Together, these three molecules comprise the biological mixture that is formally known as the slow-reacting substances of anaphylaxis. All of these reactions take place in mast cells, eosinophils and alveolar macrophages, all of which have been implicated as critical effector cells in the pathobiology of asthma. Airway epithelial cells and pulmonary vascular endothelial cells may also produce leukotrienes via transcellular metabolism [31,32].

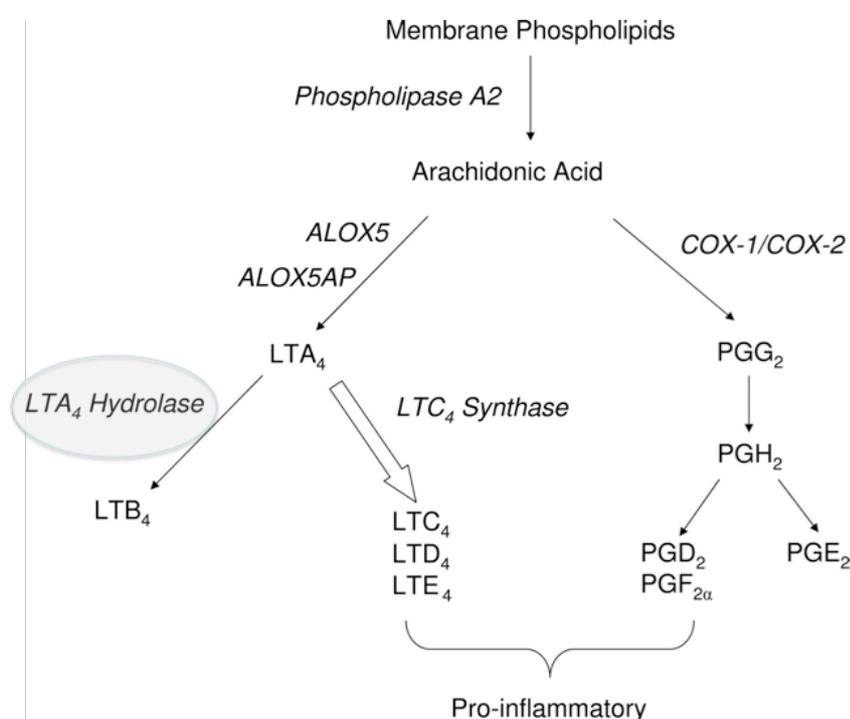


Figure 10: Pathways of LT biosynthesis. LTA₄ hydrolase is circled in grey [32].

1.2.1. Properties of LTA₄H

LTA₄ hydrolase is a 69 kDa enzyme with a Zinc ion in the active site with no apparent homology to other epoxide hydrolases. In addition to its epoxide hydrolase activity, the enzyme also possesses an intrinsic aminopeptidase activity towards a variety of substrates such as certain arginyl di- and tripeptides as well as *p*-nitroanilide derivatives

of Ala and Arg which are hydrolysed with high efficiencies [33]. While the physiological role of the epoxide hydrolase activity is quite well studied, the aminopeptidase role is less clear but recent studies point in a role of binding the chemotactic tripeptide Pro-Gly-Pro [34].

The catalytic activity of LTA4H has been extensively studied through a combination of site directed mutagenesis, enzymology and X-ray diffraction [35]. The two reaction mechanisms (aminopeptidase and epoxide hydrolase) do not share a common mechanism, but there is an overlap of their active sites. Important residues for either of these mechanisms have been identified and studied and for an overview see figure 11 adapted from Haeggström et al., 2007.

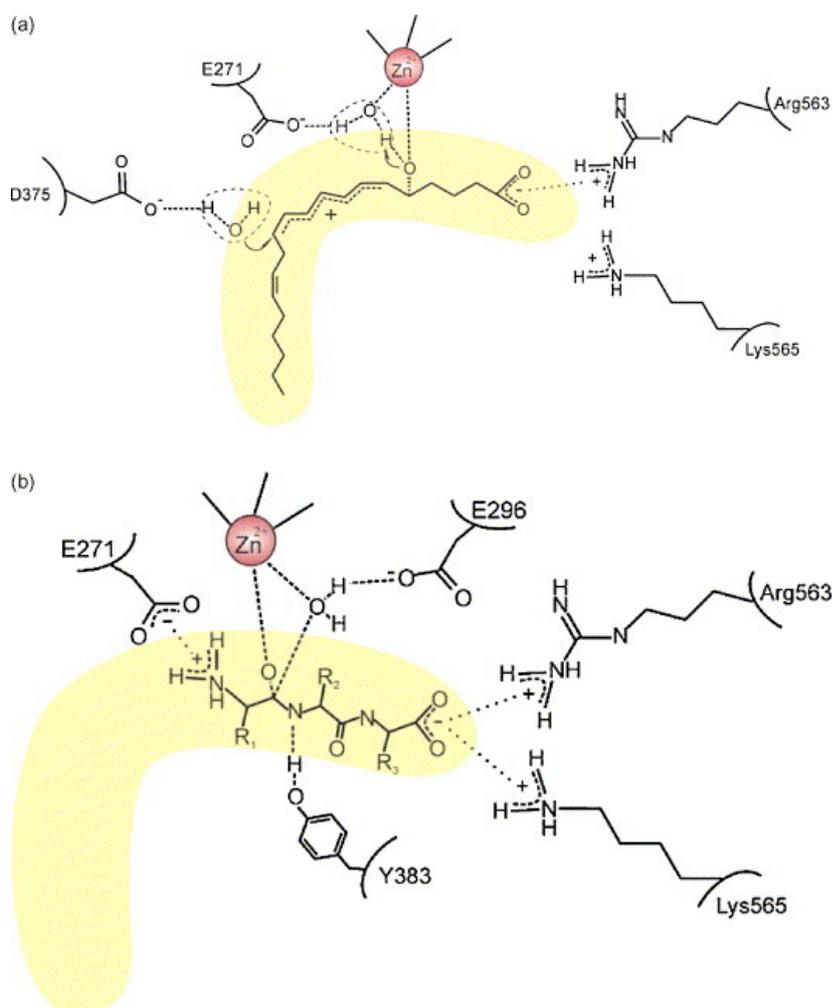


Figure 11: The Catalytic mechanisms of LTA4H. (a) The epoxide hydrolase activity and (b) the aminopeptidase activity. The substrate binding pocket is shown in yellow.

Homologues to LTA4H can be found in many organisms, from bacteria to mammals. While the enzymes in the higher vertebrates all have the ability to hydrolyse epoxides as well as tripeptides, the enzymes in the lower organisms, despite high sequence identity, only possess the aminohydrolase activity [36].

1.2.2. Structure of Human LTA4H (*hum*LTA4H)

The crystal structure of human LTA4H (*hum*LTA4H) was solved at 1.95 Å resolution in complex with its competitive inhibitor Bestatin [37]. The protein is folded into three domains: an N-terminal domain, a catalytic domain and a C-terminal domain organized in a superhelical fashion.

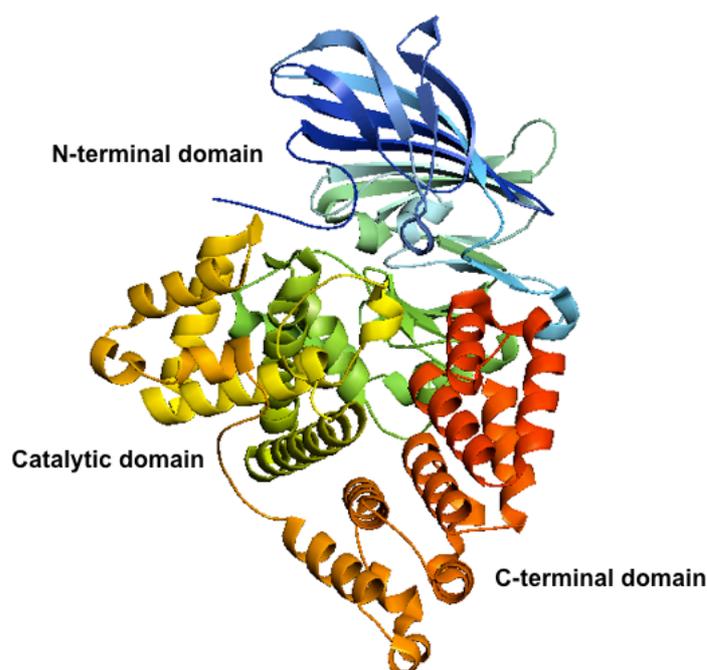


Figure 12: Crystal structure of human LTA4H generated by Pymol from PDB file 1hs6.

The three domains are packed together forming a molecule with approximate dimensions of 85 Å x 65 Å x 50 Å. Through this packing arrangement, a deep cavity is formed in the interface between the three domains. The cavity wall is mostly surrounded by residues of the catalytic and C-terminal domains and only minor contribution comes from the N-terminal domain. The cavity consists of a hydrophilic part separated by zinc binding site. At the zinc site itself the cleft extends deeper into the enzymes resulting

into a narrow and bent tunnel that penetrates into the catalytic domain. The bestatin molecule is located partly in the hydrophilic cleft and its phenyl ring moiety partially occupies the narrow channel. It is directly bound to the catalytic zinc ion. It is thought that LTA₄ binds into the full narrow binding pocket.

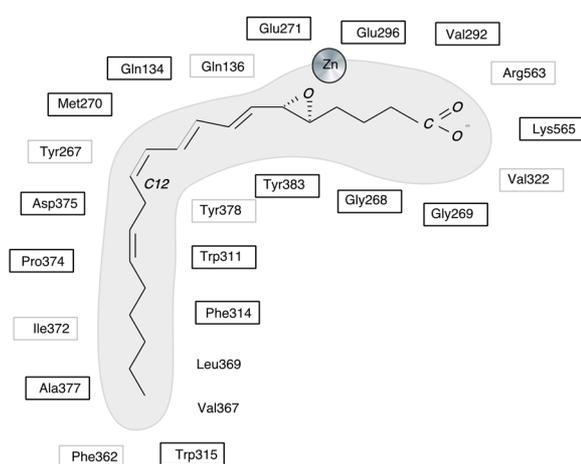


Figure 13: Schematic presentation of LTA₄ binding into the pocket of *hum*LTA4H

1.2.3. *Saccharomyces cerevisiae* LTA4H (*sc*LTA4H)

The related LTA4H enzyme from *S. cerevisiae* displays only a rudimentary epoxide hydrolase activity and should be considered mainly a leucine aminopeptidase. From sequence comparisons it is not clear why the epoxide hydrolase activity is lacking while all identified essential residues for the catalytic mechanism are present in the protein. From initial homology modeling it has been proposed that the size and nature of the epoxy-substrate binding site is an essential determinant for the presence of the epoxide hydrolase activity but no clear picture has emerged [38].

From *S. cerevisiae* both the apo and inhibitor bestatin bound crystal structures were solved (see paper III in this thesis). Like the human protein, *sc*LTA4H is folded into three different domains (an all- β N-terminal domain, residues 40–252; a mixed α - β catalytic domain, residues 253–506; and an all helical C-terminal domain, residues 507–671) (figure 14). In comparison with the human enzyme, *sc*LTA4H has an extension at the N terminus, some small deletions (3–4 residues) in the N-terminal domain, and two larger insertions (6 and 12 residues, respectively) located in loop areas

between the helices in the C-terminal domain. The catalytic domain is the most conserved (56% identity, 76% similarity at the amino acid level).

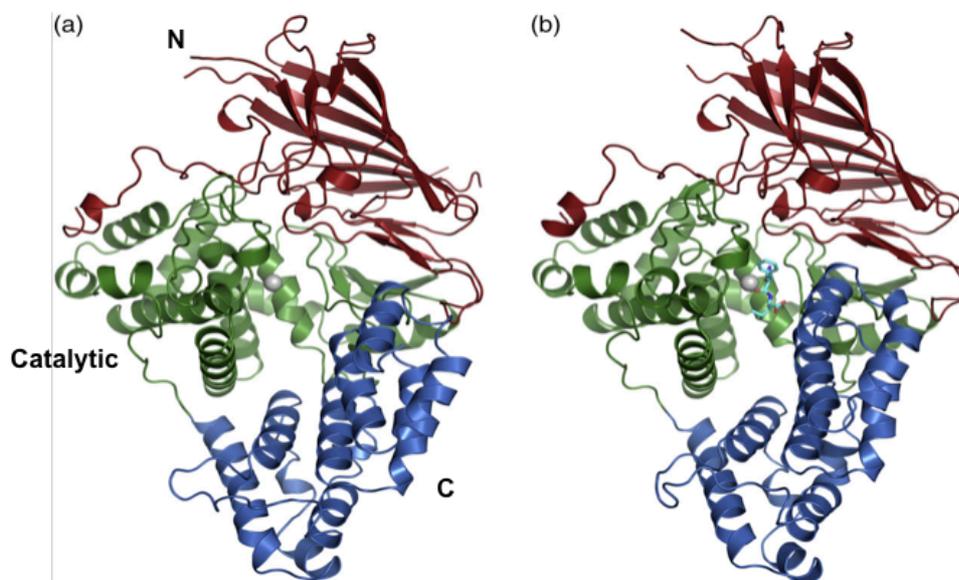


Figure 14: Cartoon presentation of *scLTA4H*. (a) Structure of the enzyme alone; (b) structure of the enzyme in complex with the bestatin inhibitor [39].

Compared to the human protein, in the apo-form of the *scLTA4H*, the C-terminal domain was rotated away from the interface between the three domains and the cavity inside the protein was increased in size. The structure of the complex of *scLTA4H* with the inhibitor bestatin shows that upon binding of the inhibitor a rearrangement of the structure occurs and the active site closes (paper III). This points to an induced fit mechanism where the full active site is only formed upon binding of the substrate. The lack of leukotriene A₄ hydrolase activity within the yeast homologue seems to be a result of an obstruction of the protected cavity which is therefore no longer present, while other parts of the binding pocket have become more solvent accessible [39].

1.2.3.1. *Sc*LTA4H purification

*Sc*LTA4H is purified as mentioned previously [39]. In the purification procedure the protein is expressed with an extra protein (GST) which is used as fusion tag. The tag is then digested off using a TEV protease [40]. An schematic flowchart is available at figure 8.

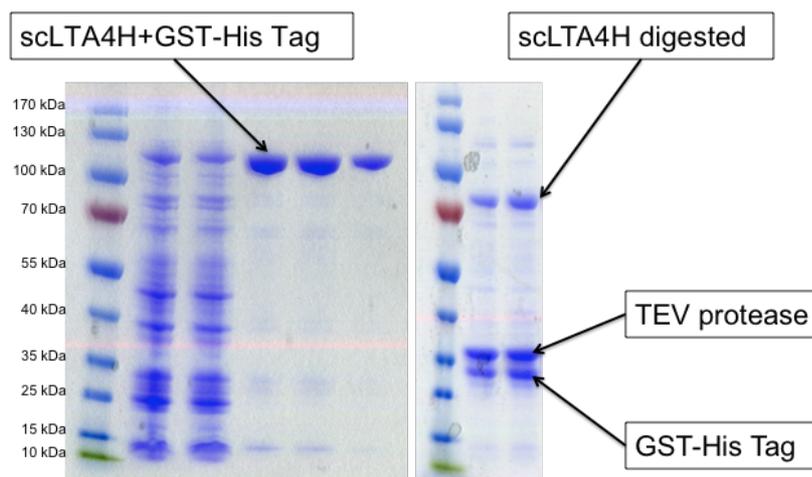


Figure 15: SDS-PAGE gel of *sc*LTA4H during purification steps. From left, the third column is crude protein, fourth is GST tagged *sc*LTA4H. On right image the digested protein mixture is shown. The first column in both image contains the molecular weight marker.

1.2.4. *Xenopus laevis* LTA4H (*x*LTA4H)

*X*LTA4H enzyme displays both the aminopeptidase and the leukotriene hydrolase activities, although without the precise product specificity. The structure of *x*LTA4H was recently solved (see paper IV of this thesis). *X*LTA4H forms a dimer in the asymmetric unit, which is unlike *hum*LTA4H or *sc*LTA4H. The main difference in the active sites of the human and the amphibian enzymes is the tyrosine to a phenylalanine exchange at position 375, closely situated to the catalytic residue Y380 and near the site where the LTA₄ triene moiety is thought to bind. Y383 (Y380 in *x*LTA4H) is important for LTB₄ formation. In the human enzyme the two tyrosines are hydrogen bonded to each other. Y378 has been suggested to have an important role to stabilize Y383 for efficient catalysis in the human enzyme. The *x*LTA4H with a Phe in position 375 is not able to form a hydrogen bond to Y380 and has a less stabilizing effect on the

positioning of Y380 which as a result means a decreased product specificity for this enzyme [41].

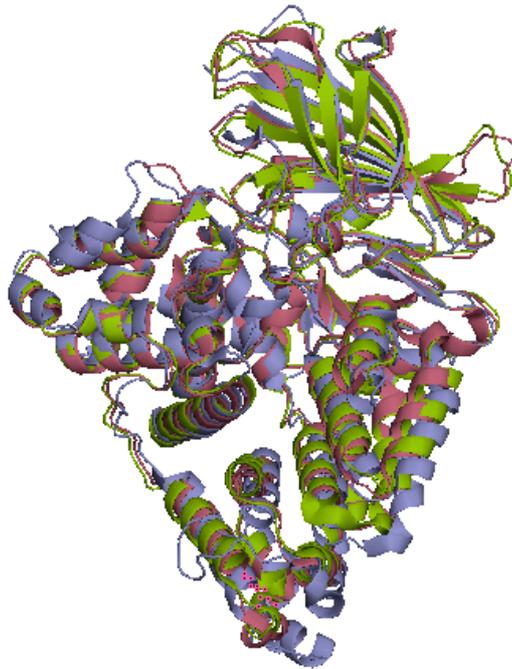


Figure 16: Superimposed cartoon structure of bestatin bound *hum*LTA4H (colored red, PDB code 1hs6), *sc*LTA4H (colored blue, PDB code 2xq0) and *x*/LTA4H (colored green). Monomeric form of *x*/LTA4H was generated by deleting one chain from PDB file 4gaa.

2.1. CRYSTALLOGRAPHY

The knowledge of accurate molecular structures at high resolution is a prerequisite for rational drug design and for structure based functional studies. Crystallography can reliably provide the answer to many structure related questions, from global folds to atomic details of bonding. X-ray diffraction studies have made outstanding contributions to structural molecular biology and many other life science fields.

2.1. Growing protein crystals

Crystallization remains the main bottleneck in determining a structure by X-ray diffraction methods. Unfortunately, the exact reasons why a certain protein crystallizes or not, are still not known. Purity and homogeneity of the protein sample are important criteria, but many other variables are present as well. To obtain crystals for protein 3D structure determination, a wide array of crystallization conditions is usually screened. These conditions have to be explored in a rational fashion. A break-through in the way that protein crystallizations are performed was the development of the sparse-matrix method [42]. In this method, instead of using brute-force methods of screening all conditions, only those conditions that previously had produced crystals for several other proteins were used. As this set of conditions is very much varied, and the pH or protein & precipitant concentrations are not finely sampled, often only initial hit are identified that needs to be optimized. Different sparse-matrix screens have been developed since and are commercially available.

Thus, starting with different sets of initial screening conditions, the search for a suitable condition can be narrowed by further developing conditions centered around a promising condition that provides some kind of crystalline form. Formation of protein crystals gets initiated by nucleation of protein molecules in a crystallization drop [43]. Subsequently a growing step occurs that should result in the formation of a single protein crystal.

Failure to crystallize a protein may be due to in part lack of insufficient intermolecular site-specific interactions needed for stable lattice formation [44]. To aid crystallization sometime a stabilizing molecule or additive can be used [45]. Additive screens have

Seeding is a technique to induce artificial nucleation or accelerate crystal growth in a crystallization experiment. Often the exact conditions for good nucleation and perfect growth are not the same, which can lead to too many very small crystals, imperfect growth or the incorporation of faults or no growth at all. Seeding, where the conditions can be tailored for each of these steps is a way to separate nucleation and growth and obtain well diffracting crystals [47-50].

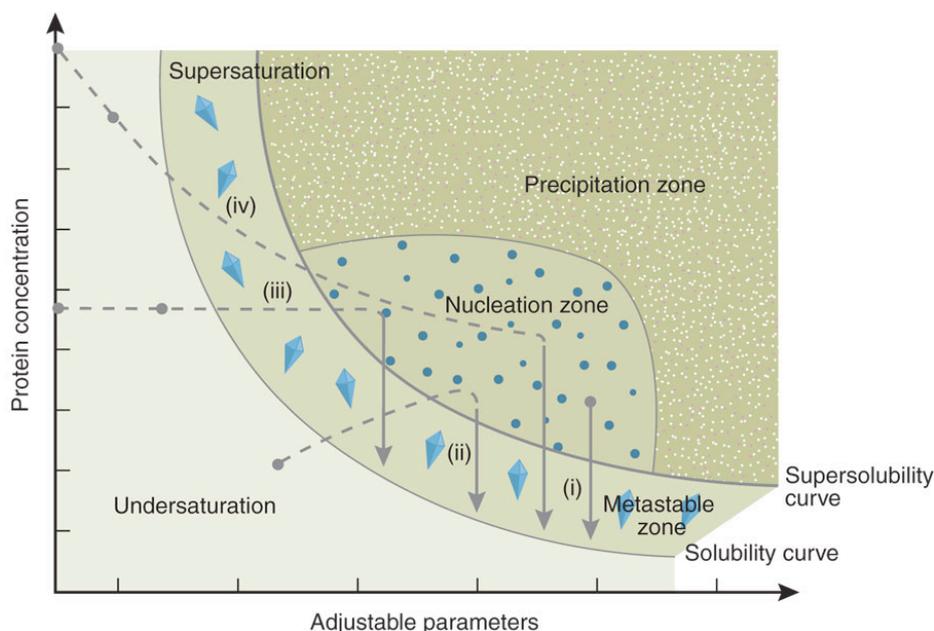


Figure 18: Phase diagram for a crystallization experiment, showing the path that in a successful batch experiment takes. The horizontal and vertical axes in this case are undetermined functions of the precipitant and macromolecule concentration, respectively [51].

2.2. Symmetry

An asymmetric unit is the smallest unit of volume that contains all of the structural information and that by application of the symmetry operations can reproduce the unit cell. There are 7 lattice systems, which combined with centering gives rise to 14 Bravais lattice types (shape). Rotations, reflections and inversions give 32 point groups of crystal that can be combined the Bravais types. The result is that there are 230 space groups. However as protein molecules contain asymmetry, symmetry operations that contain reflections or inversions are not possible and thus for protein crystallography only 65 space groups need to be considered.

2.3. Data collection

In order to solve structure using X-ray diffraction, datasets with diffraction data needs to be collected. For that purpose, crystals are exposed to X-rays, either on a home source or at a synchrotron beamline. As crystals are prone to radiation damage, especially at the modern undulator synchrotron beamlines, data is collected at cryogenic temperatures. Crystals are flash-cooled in the presence of a suitable cryoprotectant and exposed to X-rays at 100K. The data frames are inspected and crystal symmetry, unit cell parameters, crystal orientation and resolution limit are determined. For the calculations of electron density maps it is important to both have completeness of the data (over 90%) as well as a reliable quality (see figure 19).

Many parameters come into play for data collection. Shortly during data collection, the crystal is rotated with small increments and data is collected for each increment. An adequate D° is chosen, over which the crystal is rotated during the collection of a frame of data. This D° is typically for CCD detector between 0.5 and 1° . If the data frames contain many spots, e.g. in the case of a very long cell axis, the oscillation might be needed to be further reduced. The exposure time for each frame is determined, it needs to be such that reflections are reliable with a good signal to noise, but not too many overloads, due to detector limitations, occur. If too many overloads occur while the outer shells are still weak, several passes of dataset collection might have to happen with different exposure times.

Data frames are carefully inspected to see where the resolution limit is and the detector is moved such that the detector to sample distance matches this limit (see figure 29 for relationship between detector to sample distance and resolution). The amount of data that needs to be collected depends on the symmetry of the crystal. Lower symmetry space groups (e.g. P1) require that data over a larger total rotation are collected (up to 180° for native data sets), while for higher symmetry crystal forms this can be reduced considerably if the starting point for data collection is chosen right. There are programs such as EDNA [52] and BEST [53] that give guidance for the strategy to apply during data collection regarding the many parameters that need to be supplied.

$$\rho(x,y,z) = \frac{1}{V} \sum_{hkl} F_{hkl} 2\pi i (\exp(hx + ky + lz))$$

Completeness \nearrow hkl \nwarrow Quality

$$I_{hkl} \approx |F_{hkl}|^2$$

Figure 19: Fourier Transformation, relating electron density to diffraction data.

Ideally in a crystal, all molecules should be well ordered and the macromolecules should have the same conformation to determine the 3 dimensional high-resolution structure of a protein using X-ray diffraction. In a crystal the large number of repeating units or molecules amplifies the signal and the overall diffraction intensity is proportional to the number of unit cells in the crystal present. From an ordered crystal strong diffraction results can be obtained from constructive interference of photons. Irregularity in the orientations or translations of the molecules limits the order and usefulness of a crystal. Crystal suffering from disorder destroys the periodicity and this leads to streaky, weak, fuzzy diffraction patterns [54] (see also figure 22). Annealing, dehydration, soaking and cross-linking are suggested crystal manipulation techniques to improve the diffraction quality [55].

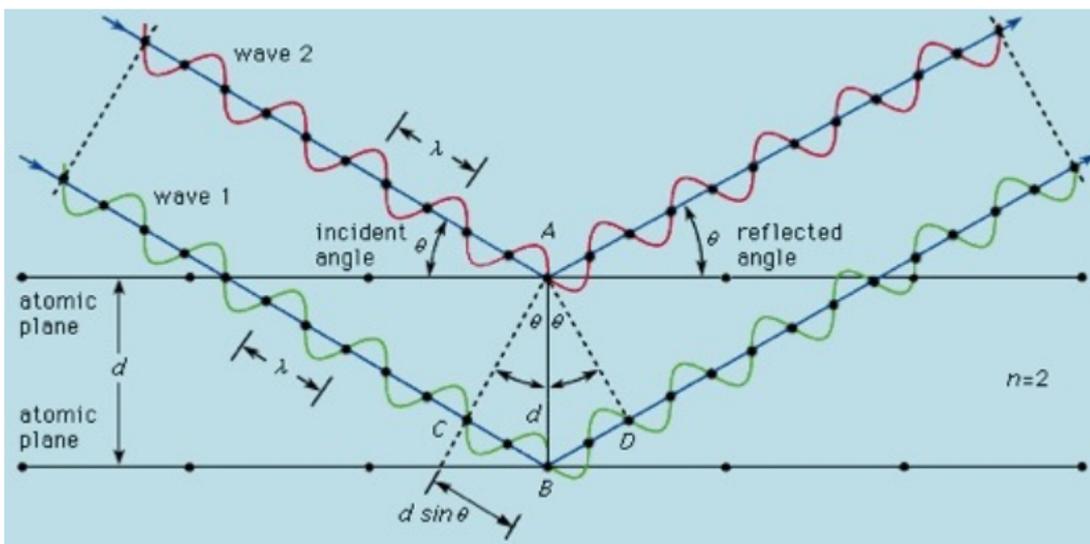


Figure 20: Bragg's law. The beams reflected by a set of parallel planes amplify each other when the path difference between reflections from successive planes is equal to 1λ , 2λ , etc.

2.4. Diffraction Resolution

The maximum scattering angle $(2\theta)_{\max}$ at which diffraction peak intensities can be reliably measured determines the diffraction resolution $d_{\min} = \lambda/2\text{Sin}[(2\theta)_{\max}/2]$ and this limits the spatial resolution of the electron density map that can be derived from the diffraction data [56].

2.5. Mosaicity and mechanical stress

Mosaicity arises from the presence of small individual domains in the crystal which are misaligned to each other. As these small domains will diffract with slightly different settings, reflection spots will elongate and with increasing mosaicity overlap and no longer deconvolute. Mosaicity can be a problem to obtain reliable diffraction. Mechanical stress within the crystal can also lead to a deterioration of the diffraction pattern.

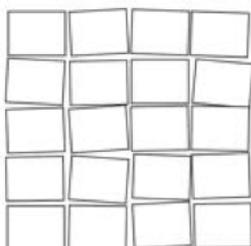


Figure 21: Simplified presentation of the presence of mosaic blocks in a crystal.

The mosaic width of a reflection is defined as the range of angles $\Delta\theta$ over which a crystal will continue to diffract strongly at a fixed, well-defined 2θ when the crystal is rotated about an axis θ perpendicular to the plane defined by the incident and diffracted X-ray beams. The measured mosaic width $\Delta\theta$ is a convolution of the intrinsic width η of the crystal and an instrumental resolution $(\Delta\theta)_{\text{IR}}$ [56].

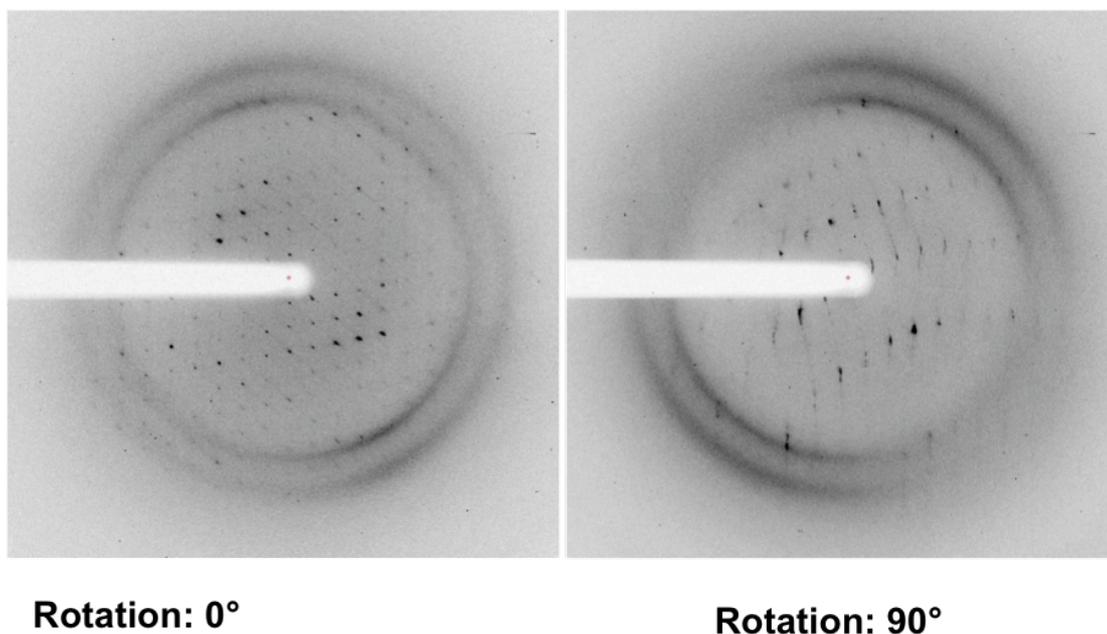


Figure 22: Diffraction image from crystal of MID⁷⁶⁴⁻⁹¹³ depicting presence of rotational disorder on the right side image (see paper II).

2.6. Data quality

Diffraction experiments yields to a series of 3 dimensional grids of pixels containing intensities of the individual reflections. The quality of diffraction data are indicated by parameters such as resolution, completeness, I/σ (or signal-to-noise ratio), and R_{merge} , overall and in the highest resolution shell. If a proper diffraction data collection experiment is not performed or if rapid decay of diffraction occurs due to radiation damage, some reflections might not be measured at all and this leads to an inability to reach to a sufficient completeness (above 90%) of the data. Each value of the electron density map is correctly calculated by Fourier transform only with the contribution of all reflections (see figure 19). Thus the lack of completeness will negatively influence the quality and interpretability of maps computed from incomplete data sets.

Multiple observations of each reflection are necessary to statistically assess the quality of each reflection. They are needed to identify and reject potential outliers which might have occurred due to possible instrumental problems. For a good quality dataset the number of such rejections should be minimal. Multiple observations and observations from symmetry related reflections are used to measure the statistical error of a dataset by the R_{merge} . As these multiple and symmetry related reflections should have the same

values, a comparison of these reflections will indicate the level of error within the dataset. A very good set of diffraction data should be characterized by an R_{merge} value <4–7 although with well optimized experimental systems it can be even lower.

In principle, high multiplicity (or redundancy) of measurements is desirable, as it improves the quality of the resulting merged data set, with respect to both the intensities and their estimated uncertainties. This effect may be aggravated by radiation damage, initiated in protein crystals by ionizing radiation, especially at the very intense synchrotron beamlines. The meaningfulness of measured intensities can be gauged by the average signal-to-noise ratio, $\langle I/\sigma(I) \rangle$. This measure is not always absolutely valid because it is not trivial to accurately estimate the uncertainties of the measurements $[\sigma(I)]$. Usually the diffraction limit is defined at a resolution where the $\langle I/\sigma(I) \rangle$ value decreases to 2.0 although more recent approaches have been more inclusive to even lower values [57,58].

2.7. B Factor

The B or “temperature” factor is used to characterize the fall-off of the diffracted intensity I with the scattering angle, according to $I \propto \exp[-2B \sin^2 \theta / \lambda^2]$. Unlike the diffraction resolution, the overall B factor obtained from a Wilson analysis is determined primarily by the properties of the crystal, and can be more reliably used to compare crystals measured in different laboratories. For a typical protein crystal, the B factors range from 5–100 Å² compared with <1 Å² for small molecule crystals, and these correspond (in a simple Debye–Waller analysis) to r.m.s. atomic displacements on the order of 1% of the molecular diameter [56].

$$\exp \left[-B \frac{\sin^2 \theta}{\lambda^2} \right],$$

B is called the thermal or temperature parameter; θ is the Bragg angle.

2.8. The Phase Problem

Each reflection of the diffraction pattern or structure factor corresponds to a wave consisting of an amplitude and a phase. An X-ray detector can only record intensities but not the phases of the electromagnetic waves during data collection. As a result, from the diffraction pattern, we can only obtain the intensities $I(hkl)$ of the reflections (hkl) and the phase $\alpha(hkl)$ cannot be measured. The X-ray experiment measures thus only the amplitudes where reflection intensities are taken as the square of the Fourier amplitudes. This phenomenon in protein crystallography is known as the phase problem. We need to know the amplitudes of the diffracted x-rays as well as their relative phases to compute the Fourier transform and to obtain the vital information required for the determination of the electron density distribution in the crystal. The Phase problem in protein crystallography can be solved by molecular replacement, isomorphous replacement and anomalous scattering.

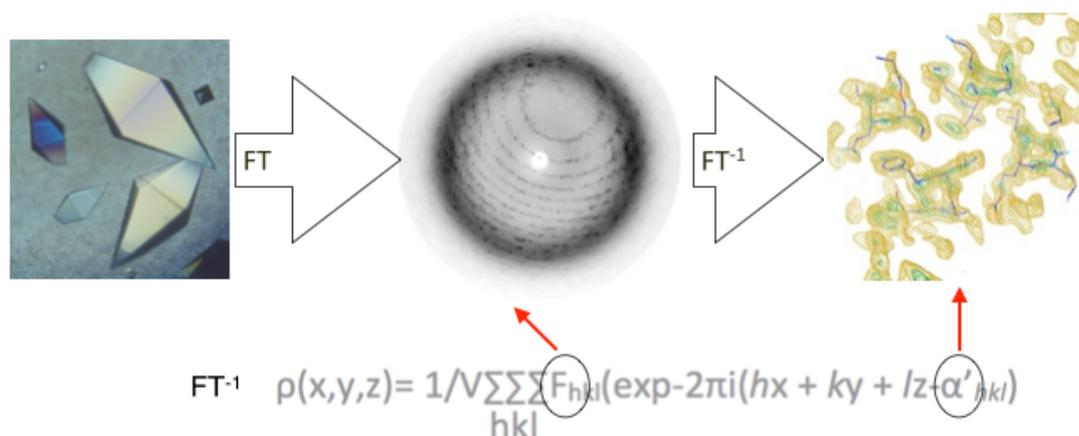


Figure 23: The phase problem in crystallography. As we can only measure I_{hkl} , ($I_{hkl} \approx F_{hkl}^2$) and not the $\alpha(hkl)$, the Fourier transformation required for the conversion from diffraction to electron density contains a problem.

2.8.1. Molecular Replacement

If a homologous protein structure with a sequence of sufficient identity (preferably at least 40%) is known then this can be used to approximate the initial phases of the structure factors. For that purpose a Patterson map is calculated from the experimental

intensities and compared to a theoretical Patterson map of the homologous structure coordinates. In Patterson maps all phases are set to 0, which results in a density where the peaks correspond to distance vectors between individual atoms. Patterson maps can be rotated and then translated relative to each other and the position of the molecules can be found and initial phases can be determined.

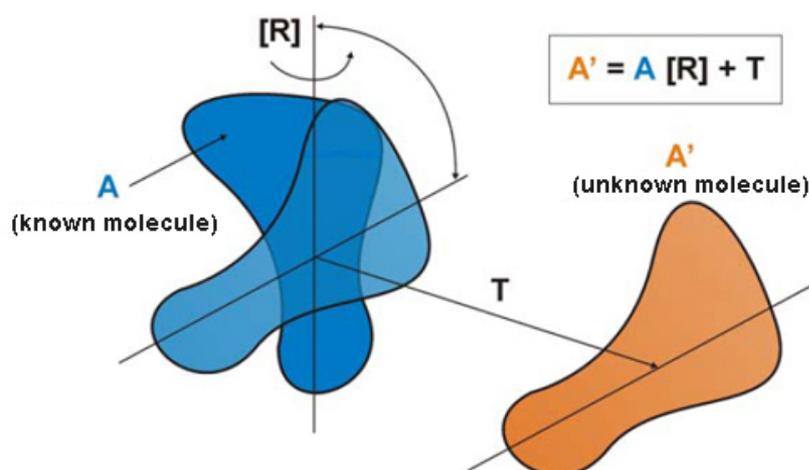


Figure 24: Simplified scheme of the molecular replacement (MR) method. The molecule with known structure (A) is rotated through the [R] operation and shifted through T to bring it over the position of the unknown molecule (A').

2.8.2. Isomorphous replacement

Isomorphous replacement is based on the idea that introduction of a small molecule into a protein or nucleic acid crystal does not or hardly alter the structure of the macromolecule. This method can be used by introducing one or more “heavy” atoms to the protein and this method was awarded the Nobel price to Max Perutz in 1962. The introduction of the heavy atoms (e.g. atoms such as Hg, containing many more electrons than C, O and N) changes the intensities of the reflections in the diffraction pattern, while the spacing of the diffraction pattern itself should remain the same (i.e. no changes of cell dimensions should occur and crystals should remain isomorphous). From these small differences the positions of the heavy atoms in the unit cell can be determined. This information can be used as a clue to determine the phase angles of the protein reflections using either direct or Patterson methods. If more than one type of heavy atom is used the technique is called multiple isomorphous replacement (MIR)

[59]. Multiple datasets need to be collected within this method, at least one for the native crystal (without addition of the heavy atoms) and subsequent ones for the heavy atom derivatives.

2.8.3. Multiple/Single Anomalous Dispersion

Anomalous scattering or dispersion of heavy atoms can also be used to obtain the information about the positions of the heavy atoms using only one single crystal. If the wavelength of the X-ray corresponds to a transition between different electron shells of the atom, which occurs at an absorption edge, there will be a modification of the phase of the diffracted wave. This so-called anomalous scattering results in changes in the diffraction pattern such that it can be distinguished from the rest (e.g. unequal amplitudes of the Friedel's pairs). Each reflection will be modified by an anomalous component that contains a real Df' and an imaginary Df'' component. For a normal diffraction experiment, Friedel's law is valid, which states that the intensities of the reflection $(h; k; l)$ and its inverse $(-h; -k; -l)$ are equal and that the phases of the underlying structure factor have opposite signs, $\phi(h; k; l) = -\phi(-h; -k; -l)$. For heavy atoms, the wavelength of X-rays lies in a region where this is no longer true under all circumstances due to the anomalous behavior (see figure 25).

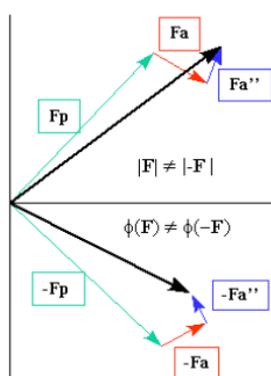


Figure 25: Breakdown of Friedel's Law due to the presence of an anomalous scatterer.

This wavelength is different for every type of atom and normally has to be determined before data collection by a fluorescence scan (scattering of X-rays at right angle to the incident beam). The difference in intensities can be exploited by a Harker construction similar to isomorphous replacement, but with j_{FT} and j_{FP} replaced with $j_F(h; k; l)$

and $jF(-h; -k; -l)j$. With this SAD (single-wavelength anomalous dispersion) approach, the two-fold ambiguity for the phases remains. Modern phase modification approaches as supplied by solvent flattening or symmetry averaging techniques can solve this ambiguity and a structure can be solved.

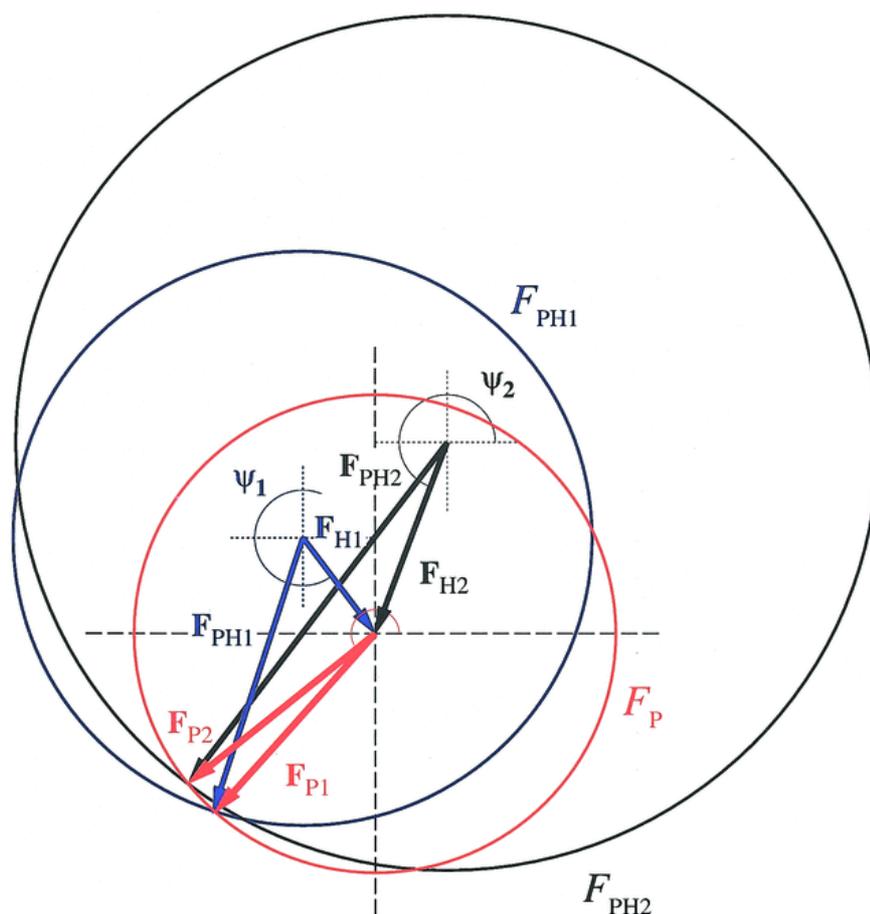


Figure 26: Argand diagram of the determination of unknown phases from a protein crystal structure using multiple isomorphous derivatives. protein structure factors F_P (red) are vectors that radiate from the origin of coordinates, with amplitude F_P and phase α ; heavy-atom structure factors F_{H_i} (blue for $i = 1$, black for $i = 2$) are represented by vectors of amplitudes F_{H_i} and phases ψ_i , that start at $(-F_{H_i}\cos\psi_i, -F_{H_i}\sin\psi_i)$ and end at the origin of coordinates; derivative structure factors are represented by vectors beginning at the origin of their respective heavy-atom vectors and finishing at the end of the protein structure factor vectors, such that $F_{PH_i} = F_{H_i} + F_P$. Possible solutions for the protein complex structure factor F_{P_i} are given by the intersections between the circle of radius F_P centered at the origin and the derivative circles of radii F_{PH_i} centered at their respective F_{H_i} origins. In absence of any error, all circles should intersect at only one point, which would give the correct solution for F_P . In practice, every derivative circle

may intersect with the native circle in a different position or may not even intersect at all [60].

As the contributions of f' and f'' to the diffraction change over the absorption edge of an anomalous atom, very small changes in the diffraction pattern will occur when data is collected at different wavelengths spanning the edge. Therefore careful data collection over several wavelengths, i.e. one dataset at the maximum of f'' (Peak), one at the minimum of f' (Inflection) and one at a remote wavelength can give a set of data for which the phase problem can be solved unambiguously. This is called the Multiple Anomalous Dispersion (MAD) method. Each of the wavelengths can be treated as a different heavy atom derivative and the position of the anomalous scatterer can be determined.

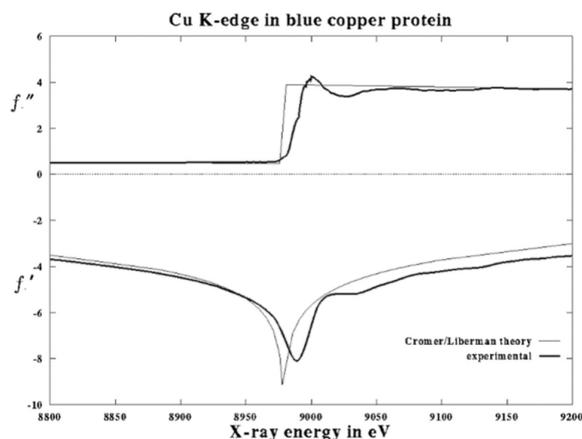


Figure 27: Behavior of f' and f'' during an absorption edge scan.

For the MAD method, if the protein contains Methionine residues, these can be substituted by a Selenium containing methionine analogue, Se-Met (see figure 28). Natural metal containing cofactors such as Iron clusters can be exploited or an anomalous scatterer can be introduced similar as in the MIR method.



Figure 28: Methionine residue (a) and its selenium analogue Seleno-Methionine (b).

3. SMALL ANGLE X-RAY SCATTERING (SAXS)

Small angle X-ray scattering (SAXS) is a widely used technique today and can be used to characterize protein structure and their complex with ligand or inhibitor at a low resolution. In this technique radiation is elastically scattered by a sample, which leads to an interference effect and results in a scattering pattern. This scattering pattern can be further analyzed to get information about size and shape of a protein molecule. The signal strength in SAXS is obtained by squaring the volume of the particle and as a result small particles are hardly visible in the presence of big particles. On the other hand, SAXS is very sensitive to polydispersity or formation or growth of large particles and care needs to be taken to avoid the presence of such particles [61,62].

As the main information yielded from SAXS experiments will come from shape reconstruction, the method can be used to rule out differences in structural information obtained from different techniques or can be used as cross validation tool [63]. It has the advantage that it is a method where the proteins are studied in solution and possible artifacts due to crystal contacts are not present. It can also be used to see changes that might occur upon ligand or inhibitor binding to protein molecules, complex formation for bigger macromolecules assemblies or flexibility of macromolecules in solution [64-67], or even to track any abnormalities in a protein [68,69].

The resolution criteria in SAXS are the same as those in microscopy. The closer the lens to the object (the larger the aperture or the scattering angle), the smaller is the detail that can be resolved. The farther away the object is from the lens (the smaller the aperture or the scattering angle), the bigger is the largest object that can be brought into the picture [70]. One should however always keep in mind that the SAS models are low-resolution ones. Typically, *ab initio* bead models utilize the scattering data up to about 1.5–2 nm resolution (momentum transfer to about $s = 0.3\text{--}0.4 \text{ nm}^{-1}$), but also for rigid body modeling this range is most sensitive to changes in the quaternary structure.

Incorporation of higher resolution data (to 0.5-1 nm resolution) is useful for domain structure analysis and the methods employing dummy residues, which also provide somewhat more detailed but still low resolution models. Wide-angle X-ray scattering patterns up to $s = 2\text{--}3 \text{ nm}^{-1}$ (0.2–0.3 nm resolution) are sensitive to the internal structure and can be utilized to probe, for example, ligand-induced conformational changes in

proteins with known high-resolution structure and can be used to study smaller changes [71]. WAXS studies can be used to study conformational changes using time resolved approaches [72].

3.1. Sample preparation and measurement

SAXS, apart from radiation damage as observed at modern SAXS beamline at synchrotrons, normally requires only minimum effort of sample preparation. Care needs to be taken however that the samples are on a homogeneous state and that aggregation is not apparent. Often samples are checked just before a SAXS experiment for the presence of such particles by e.g. dynamic light scattering, or the sample is injected right after a size exclusion gel filtration purification.

For a normal weight protein the SAXS measurements are done at least at three different concentrations in a range between 1-10 mg/ml. In a SAXS experiment the scattering intensity $I(q)$ changes upon variation of the scattering angle 2θ as a function of the scattering vector q as follows.

$$q = 4\pi\sin\theta/\lambda$$

Where λ is the wavelength of the radiation.

According to Bragg's law, the corresponding distance in real space of the scattering vector q is given by $d = 2\pi/q$. The obtained scattering curve $I(q)$ contains information in the reciprocal space on the structure of the object in solution [73].

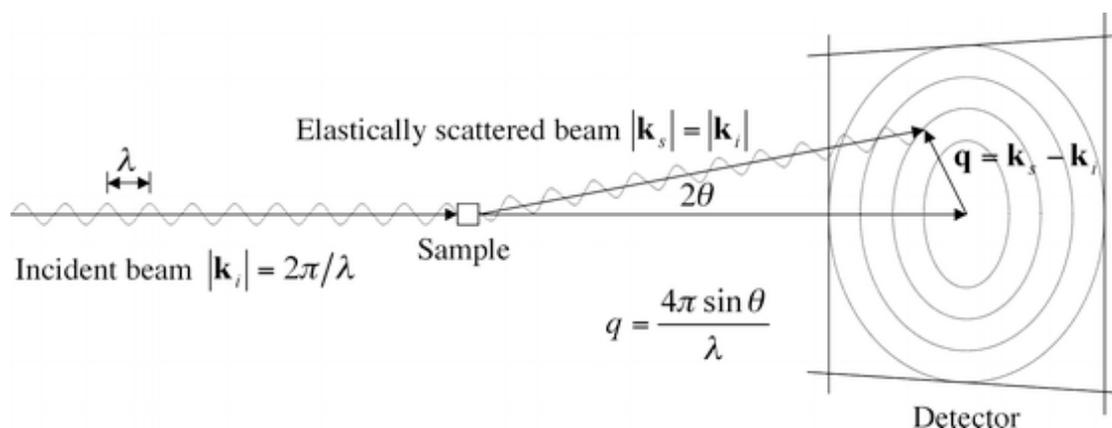


Figure 29: Schematic beamline geometry for a regular SAXS experiment [74].

3.2. Data processing and analysis

The whole sample is illuminated during SAXS data collection and as a result the average values of the structure parameters are obtained. By indirect Fourier transform methods, the average is taken over all objects and all orientations. From a SAXS experiment it is possible to measure molecular weight [75] and provide more accurate information when other methods are unable (e.g. spectroscopy when there is a lack of enough aromatic residues) [76,77].

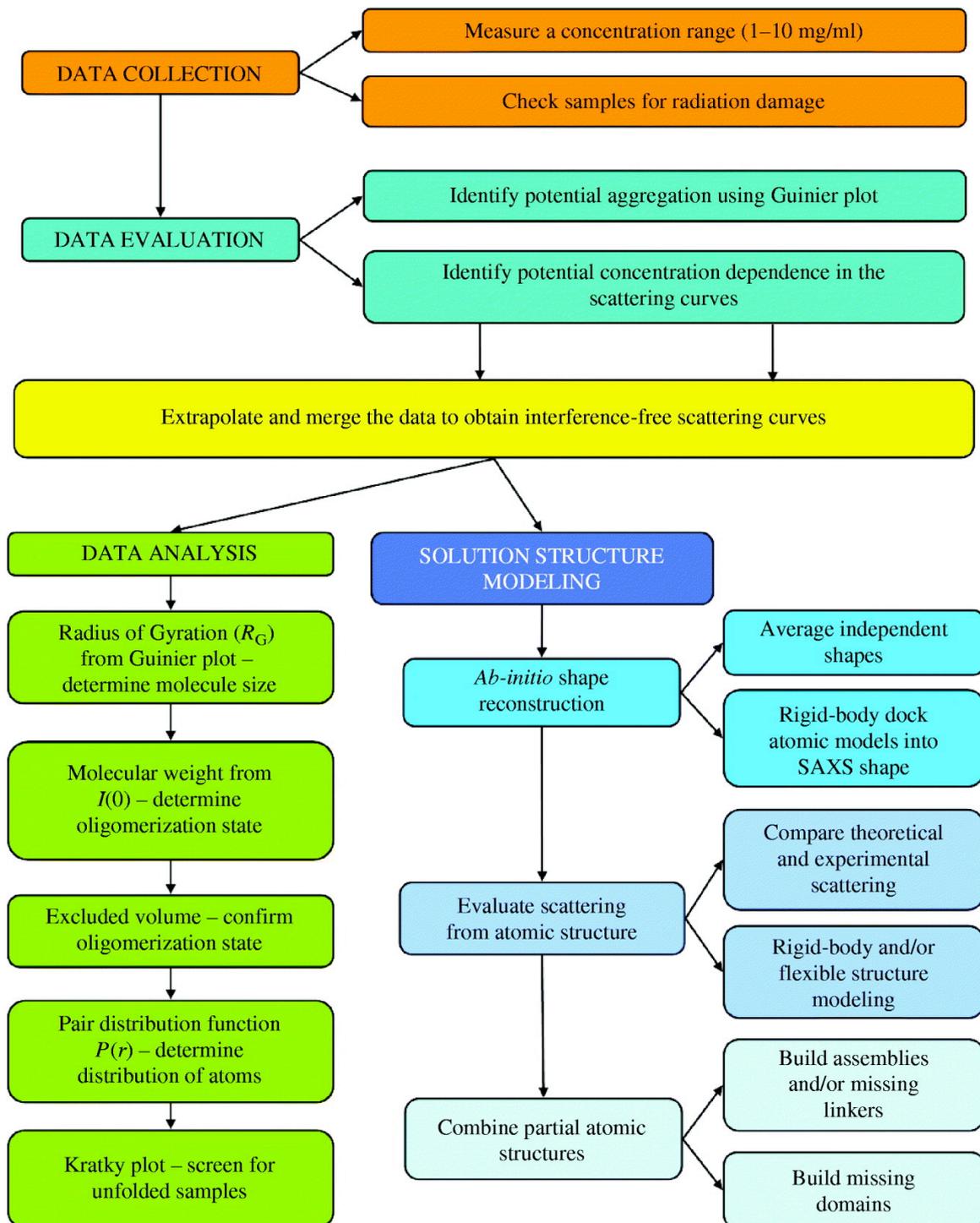


Figure 30: Steps involved in SAXS experiments [78].

3.2.1. Guinier approximation

Generally the Guinier region ($q \cdot R_g < 1.3$) is the lowest observed resolution region of a SAXS profile. This region, which can be approximated by a linear relationship, is used

to estimate for a globular protein its radius of gyration [79]. For ideal monodisperse systems, the Guinier plot $\{\ln I(q) \text{ vs. } q^2\}$ should be linear and the slope yields the radius of gyration R_g (figure 3a in paper I). Since for monodisperse samples the Guinier plot will have linearity, it can be considered as a test of monodispersity of the sample and as a test of the rigorousness of the purification protocol [61]. This parameter can also be used to determine whether the protein is well defined or if any irregularities are present in the protein molecule [79,80].

3.2.2. Pair distance distribution functions

The scattering data obtained from a SAXS experiment can be Fourier transformed to obtain information about the distances between electrons within the scattering particles [78, 81]. This is called a pair-distance distribution function $P(r)$ and upon solving it will provide the radius of gyration R_g , zero intensity $I(0)$ and the maximum length of the particle D_{max} . The shape of the pair-distance distribution function $P(r)$ itself contains information on the shape of the protein under investigation (Figure 31).

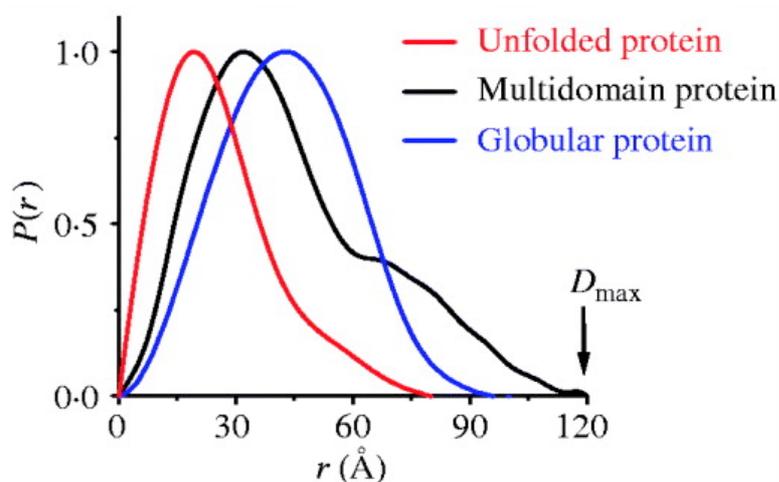


Figure 31: Globular macromolecules have a $P(r)$ function with a single peak, while elongated macromolecules have a longer tail at large r and can have multiple peaks. The maximum length in the particle, D_{max} , is the position where the $P(r)$ function returns to zero at large values of r . [78].

3.2.3. Flexibility assessment by Kratky plot

If the scattering data is plotted as $q^2I(q)$ vs. q then the plot is called a Kratky plot and this provides information on the folded nature of the scattering protein [78]. A well-folded protein gives a parabola shaped Kratky plot, while an unfolded protein either just continues to rise or reaches at maximum and follows a linear path (e.g. red curve in figure 32).

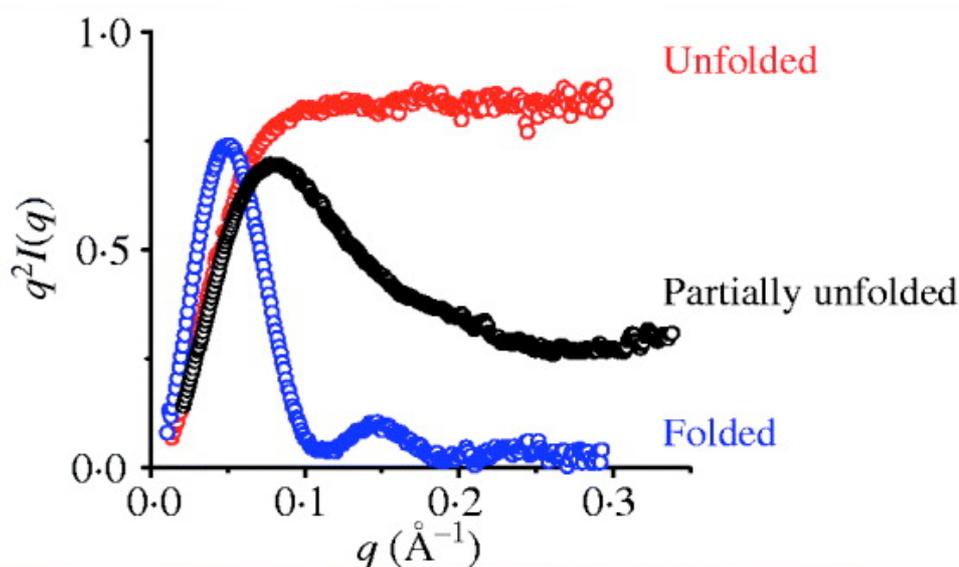


Figure 32: Kratky plots for properly folded, partially unfolded and unfolded proteins [78].

The Porod volume (hydrated volume, V_p) is calculated without normalization of the intensity and the molecular mass (MM) can be assessed from the Porod volume (V_p) provided that the partial specific volume and hydration of the particle are known [82]. The Porod law describes that the slope of the plot $q^4I(q)$ vs. q or q^4 represents the interface and fractal dimension of the scattering objects. At high q , a slope of -2 is a signature of Gaussian chain in a dilute solution, whereas a slope of -1 points to rigid rods [83].

3.2.4. Shape estimation

In a bio-SAXS experiment probably the most used tool is *ab initio* modeling. Generally, the algorithm uses a simulated annealing procedure and produces various shapes for the

scattering particle and compares SAXS profile of the model in order to optimize the similarity [84,85]. It is suggested that such shapes obtained should be viewed with caution as multiple models could describe the data and it is important to remember that resolution is very low, even with a well-behaved homogenous sample. Since SAXS data creates the average state of the sample in solution, contaminants will contribute negatively and will have an influence on the calculated shape of the experimental protein. On the other hand if there are multiple oligomerization states of the protein in solution, an average shape of all these states will be built. But if it is known how much of each of the components is present the information can be deconvoluted. The *ab initio* model producing algorithms assumes there is no aggregation present and will also have negative contribution to the correctness of the produced models [74,86,87]. In practice to be able to only see the contribution of the protein itself, careful measurements of the buffer solution alone, needs to be performed.

3.2.4.1. DAMMIN and GASBOR

There are two major methods to calculate *ab initio* models from scattering data. DAMMIN is shape restoration method which uses a simulated annealing procedure from a single phase dummy atom. This method searches for a compact dummy beads configuration minimizing the discrepancy function applying the non-compactness penalty from a random configuration. In GASBOR protein sequence is taken and chain like ensemble model is created. It also uses more q-angle data compared to DAMMIN and thus requires more high angle data. GASBOR assumes that the folded polypeptide chains is composed of amino acid residues separated by ~ 0.38 nm between adjacent C_{α} atoms in the primary sequence. At a resolution of 0.5 nm, a protein structure can therefore be considered as an assembly of dummy residues (DR) centered at the C_{α} positions [88,84].

Normally 10 to 20 models are generated by the above methods and averaged using the program DAMAVER [89]. From the averaged set a most typical model is selected. Normalized spatial discrepancy (NSD), a parameter with a value of zero for identical structures, is used to check how well they agree between models [90].

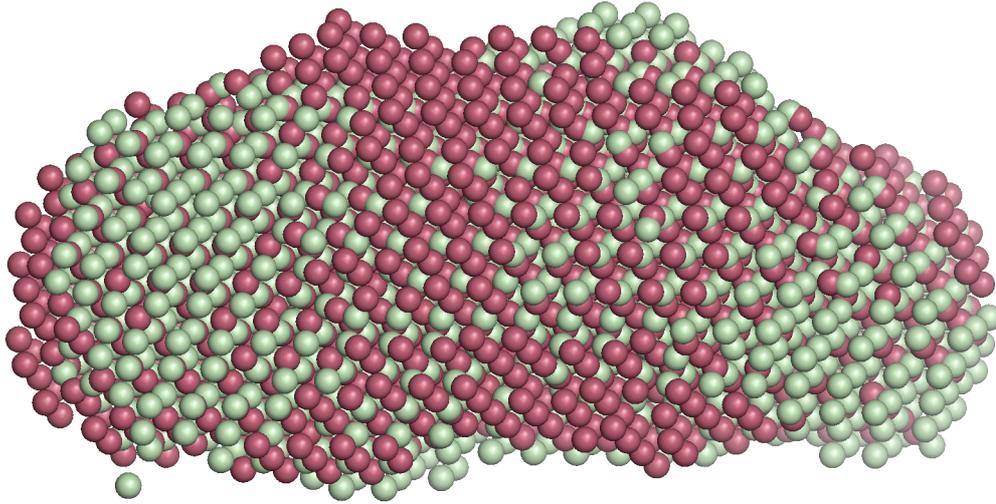


Figure 33: Superimposed GASBOR model of apo (red) and inhibitor bound (green) *Xenopus laevis* leukotriene A4 hydrolase. A wider form of the apo enzyme is visible here (paper IV and V). The models were superimposed by SUPCOMB [91] and image made by Pymol.

3.2.5. Comparison of SAXS data with a crystal structure

The program CRY SOL can generate a simulated scattering curve from coordinates available from a PDB file (i.e. a crystal structure). This simulated curve then can be compared with the experimental scattering and based on the quality of the fit, a chi value is given and this value can be used to judge if there is any discrepancy between the solution and crystal structure. This program also takes into account of the bound water molecules to the protein which might not be available in the crystal structure [92]. Generated *ab initio* models can be superimposed to a crystal structure using the program SUPCOMB [91].

4. CIRCULAR DICHROISM SPECTROSCOPY

Circular dichroism (CD) spectroscopy is a useful technique to explore secondary structure of proteins being in solution. It can be used to obtain information on the secondary structure content of a protein, or as a tool to follow folding experiments or gauge the influence of specific mutations in a protein. CD is observed when optically active matter absorbs left and right handed circularly polarized light slightly differently. This arises due to the presence of structural asymmetry in the sample measured. Protein molecules have chirality and the electronic transitions originating from their peptide backbones and aromatic side chains create different absorption spectra upon interacting with polarized lights [93]. If viewed from a plane and the two waves applied in CD can be visualized as the resultant of two vectors of equal length and are represented as trace-out circles, one rotates clockwise and the other rotates counterclockwise [94].

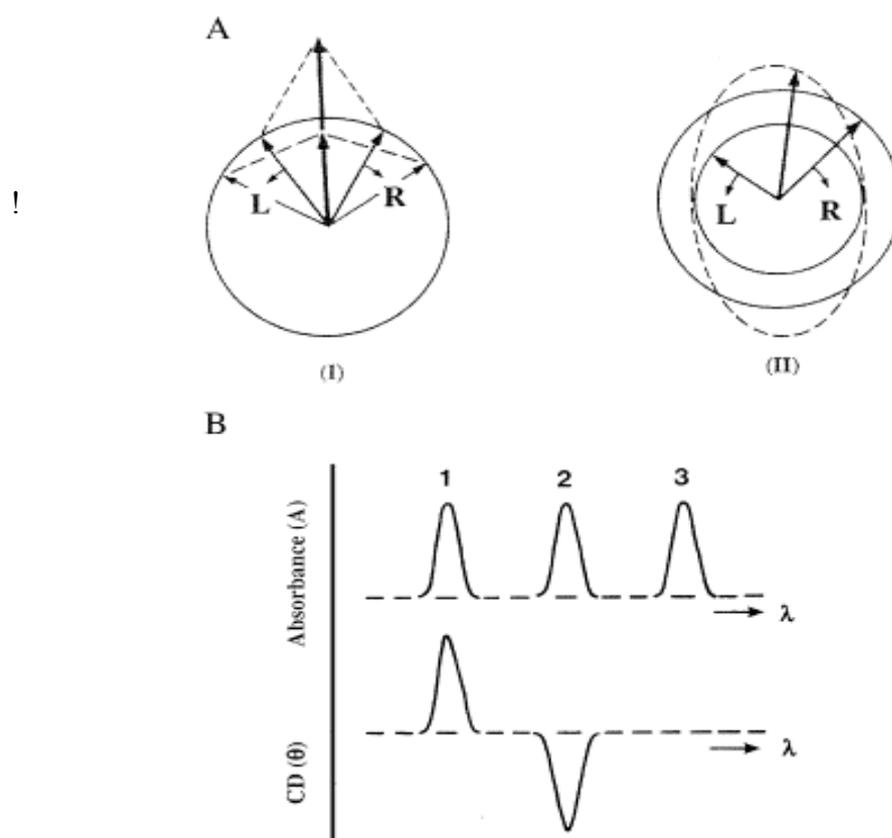


Figure 34: Origin of the CD effect. (A) The left (L) and right (R) circularly polarized components of plane polarized radiation: (I) the two components have the same amplitude and when combined they generate plane polarized radiation; (II) the components are of different magnitude and the resultant (dashed line) is elliptically polarized light. (B) The relationship

between absorption and CD spectra. Band 1 has a positive CD spectrum with L absorbed more than R; band 2 has a negative CD spectrum with R absorbed more than L; band 3 is due to an achiral chromophore. [95].

Compared to NMR and X-ray crystallography CD is a low-resolution technique, however it has a number of advantages. E.g. it allows temperature scans in a wide range of solution conditions, data collection in very short time and only small amounts of sample are required [96,97].

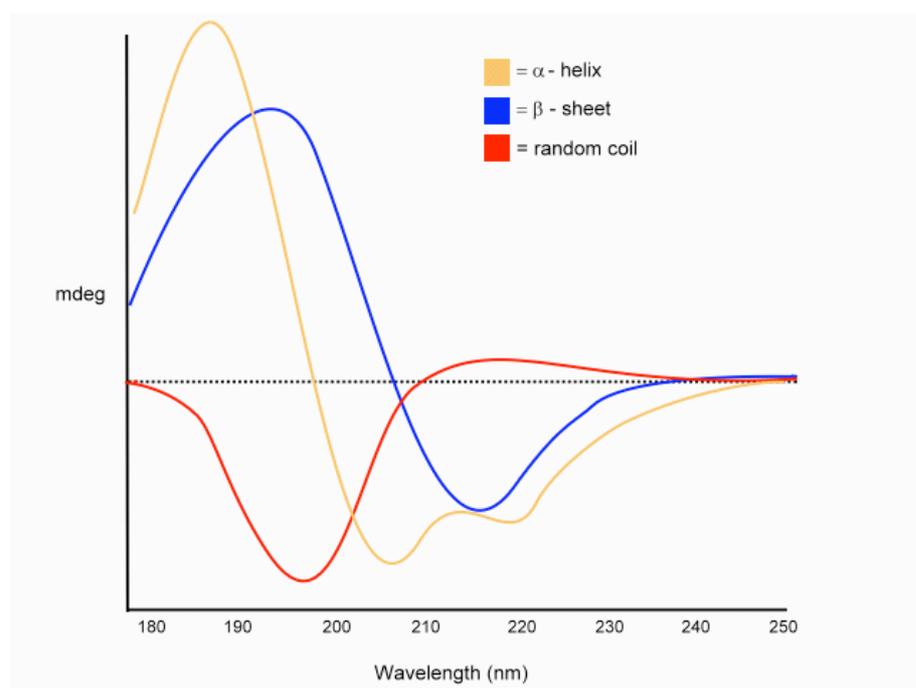


Figure 35: Examples of CD spectra from peptides with different pure secondary structure. Normally for a typical α -helix the negative band is at 222 nm with a negative and positive couplet at about 208 and 190 nm. For a β -sheet, the negative band is at about 215 nm and the positive band at about 198 nm. For random coil the negative band is at about 198 nm and the positive band at about 218 nm.

Secondary and tertiary structures can be monitored by the peptide transitions in the far UV (\sim 190-220 nm) and by the aromatic side chain transitions in the near UV (\sim 270-290 nm), respectively.

4.1 Synchrotron CD Spectroscopy

A conventional table-top CD instrument uses xenon arc lamps as light source and measure as low as around 185 nm with limitations coming from drop of flux from the source. On the other hand synchrotron based CD can provide lower wavelength measurements with high light flux reaching down to ~165 nm. Collecting data at lower wavelength includes additional electronic transitions and thus become richer in information [93,98,99]. These data allows interpreting more accurately the secondary structure content of a protein since this becomes clearer in the lower wavelength region.

The intense light generated by the synchrotron also makes the signal-to-noise ratio higher, catching peptide backbone transitions in the far UV region. In a synchrotron, sample holder with a shorter path-length can be used, as a result removing the nonchiral absorption problem. This creates an additional advantage of measuring the protein in its original buffer rather than having to change to a suitable buffer (e.g. phosphate buffer), which is required to avoid interference with solvent absorption in the UV region when measured in a tabletop CD instrument [100].

4.2 Data Processing

CD is the tangent of an angle (θ) called ellipticity which can be measured in several ways (mainly in degrees or millidegrees). CD is the difference in the absorption of left- and right-circularly polarized light and is defined as:

$$\Delta\varepsilon(\lambda) = \varepsilon_L(\lambda) - \varepsilon_R(\lambda)$$

Where ε_L and ε_R respectively are the extinction coefficients for the left- and right-circularly polarized components at wavelength λ [101].

4.2.1. Delta Epsilon ($\Delta\varepsilon$): The per residue molar absorption units of circular dichroism measured in $M^{-1}cm^{-1}$. $\Delta\varepsilon$ is sometimes referred to as molar circular dichroism. Data peaks are usually in the range of 0-10. This is the most commonly used output data and all of the analysis programs accept these input units except K2D.

4.2.2. Mean Residue Ellipticity (MRE) [θ]: Mean residue ellipticity is the most commonly reported unit and is measured in degrees $\text{cm}^2 \text{ dmol}^{-1} \text{ residue}^{-1}$. Data peaks are usually in their 10,000's and the relationship between [θ] and $\Delta\epsilon$ is shown below:

$$\Delta\epsilon = [\theta] / 3298$$

4.2.3. Theta Machine Units [θ]: Machine units measure the difference in molar extinction coefficients between left and right-handed light, usually between 1 and 100, and need to be corrected to account for the amount of protein used in the sample. This type of data requires to specify the mean residue weight $\text{MRW} = \text{protein mean weight (in atomic mass units/daltons) / number of residues}$ amu for the protein, path length (P) in cm and protein concentration (CONC) in mg/ml. To convert from machine units in millidegrees, to delta epsilons, the following equation is applied [94].

$$\Delta\epsilon = \theta \times \frac{(0.1 * \text{MRW})}{(\text{P} * \text{CONC}) * 3298}$$

4.3. Data analysis and analysis software

A number of methods have been developed for determining protein secondary structure content based on empirical analyses of CD spectra. For the analysis of the spectra of protein samples, using constrained least squares fitting analysis programs where polypeptide based reference sets are supplied, will yield more accurate results. Linear regression fits the spectrum of an unknown protein by comparison to the spectra of a set of fixed standards [102]. Suitable standards include: polypeptides with known conformations, which are essential for the analysis of non-globular polypeptides and fibrous proteins; standard [103,104].

5. SUMMARY OF SPECIFIC RESULTS OBTAINED FROM PROJECTS

Project 1: MID⁹⁶²⁻¹²⁰⁰

The purpose of the project was to obtain structural information on this domain in the first hand by X-ray diffraction. For this purpose the domain was cloned with a cleavable GST-tag and a purification protocol was developed as an older protocol gave crystallization conditions that could not be optimized. The expressed protein from this new clone provided the protein without any tag and it was observed that the protein was not secreted out of the cell into medium which occurs when expressed from the old clone in plasmid pET-26b(+). With the new preparation of protein it was possible to find crystallization conditions and optimize these. Recently we have managed to collect a data set at 2.3 Å resolution at ESRF, however a full structure determination is not yet achieved as the phasing has remained a problem. In order to obtain structural information, other methods have been used. The secondary structure content of the protein was experimentally determined by circular dichroism (both SRCD and table top CD). The prediction of secondary structure content was also done using some prediction servers. From SAXS data was collected both at MAX IV laboratory and ESRF *Ab initio* DAMMIN models could be generated and these show that the domain has an elongated shape and is most likely to form trimers. Most of the results are presented in paper I in this thesis.

As noted this domain has proven to be very difficult to crystallize (according to Xtalpred, the difficulty prediction level is 4), which could be indebted to it being a fibrous protein, and the presence of a large unordered/flexible/coil structure. It was observed that it was not possible to get good quality crystals without the additive 1,8-Diaminooctane and it takes between 3 to 8 months to grow the crystals. The quality of the crystals (in terms of diffraction resolution) is very much variable even between crystals from the same drop and diffraction screening of several crystals was necessary to find the better ones.

In order to obtain phase information, the domain was also crystallized in presence of heavy metals (e.g. tungsten, cobalt, mercury, gold, platinum and rubidium) as there is no

Methionine in the sequence. These crystals diffracted to 4 Å resolution at best and the quality was not good. Recently we have collected anomalous data in presence of iodine, zinc, strontium and bromine but no anomalous signal could be detected so far.

With the goal to co-crystallize this domain with its binding partner IgD a purification attempt was started for IgD. IgD was expressed in CHO cell line in the group of Professor Kristian Riesbeck at Malmö University Hospital. The GST-tagged MID⁹⁶²⁻¹²⁰⁰ was loaded on glutathione resin and used to bind IgD. After washing with binding buffer {100 mM Tris (pH 8.5), 300 mM NaCl and 5 mM Azid} the protein was eluted along with IgD elution buffer {100 mM Tris (pH 8.5), 300 mM NaCl, 5 mM Azid and 10 mM reduced glutathione}. In the following image (figure 36) the purified IgD heavy and light chains along with the uncleaved MID⁹⁶²⁻¹²⁰⁰ are marked for elution samples which are 8 times concentrated.

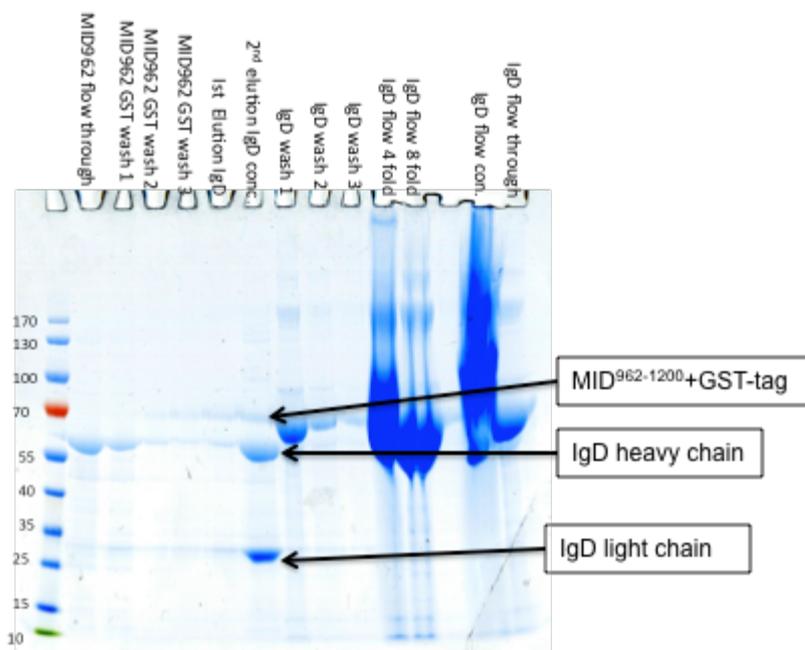


Figure 36: SDS-PAGE image of copurified MID⁹⁶²⁻¹²⁰⁰ along with IgD. The lanes starts with marker and rest are collected fractions at different steps.

The identity of the obtained protein (i.e. IgD) was further confirmed (not shown) by Western blot, which shows that this purification method was successful. However it was

also observed that the yield is quite low and a way to scale up the protein production needs to be further investigated.

Project 2: MID⁷⁶⁴⁻⁹¹³

Here the goal was also to obtain structural information on the adhesin domain of the MID protein. The protein was expressed from plasmid pET-26b(+) and purification protocol was developed (presented in paper II). From the elution profile from gel filtration, it was observed that the protein existed mainly in a hexameric or trimeric form. As for the IgD binding domain, the secondary structure content of this domain MID⁷⁶⁴⁻⁹¹³ was also experimentally determined by circular dichroism. A prediction of secondary structure content was also done by using several prediction servers.

Initial crystallization trials were started with both oligomeric states of the domain, but for the trimeric form, the crystals that were grown were very thin and small and despite many attempts, further optimization was not possible. For the hexameric form it was possible to optimize the crystal quality and a native dataset at 2.7 Å was collected at I911-3 at the MAX IV Laboratory. It was observed that the crystals suffered from some kind of rotational disorder (i.e. anisotropic diffraction if rotated, see figure 22) and it was needed to screen many crystals in order to obtain a crystal that gave reasonable quality diffraction. This inherent disorder, seriously limited the search for heavy metals as is needed for solving the phase problem, as neither this domain contains any methionine residues. It was found that a brief annealing (5-10 sec of thawing the crystal by blocking the cryogenic gas-stream) was necessary to improve the diffraction quality. Crystals of this domain were also grown in the presence of 10 mM KAu(CN)₂. A full dataset was collected of these crystals but no anomalous signal was present and a further search for heavy atoms is required.

SAXS data was collected on several occasions at both the MAX IV laboratory and the ESRF. The trimeric form suffered from aggregation and several attempts to optimize the buffer condition have failed so far. The hexameric form indicated paradoxically the absence of proper folding and a recent attempt of online SAXS data, where the SAXS data is collected right after a gel filtration step, at ESRF also was not successful.

Project 3: Crystallization of *sc*LTA4H with bestatin

In this project the goal was to obtain a structure of *sc*LTA4H in complex with the inhibitor bestatin. As the native form of *sc*LTA4H showed a large rearrangement of the domains, it was necessary to check whether binding of the inhibitor would trigger a change in conformation. As a conformational change could occur it was important that the complex would be solved from co-crystallized complex crystals and not from crystals soaked with the inhibitor. The protein was expressed, optimized and purified following the protocols mentioned in paper III. It was observed that the protein is prone to aggregation and it was necessary to set up drops immediately after purification. Good diffracting crystals grown in the presence of bestatin were obtained from commercial available screens and further optimization in the presence of additives was performed. The final crystals diffracted to 1.95 Å resolution at beamline I911-3 at the MAX IV laboratory, Lund, Sweden. The structure was solved using molecular replacement and subsequently refined. The crystal structure revealed that indeed a conformational change occurs upon inhibitor binding and it is proposed that the enzyme uses an induced fit mechanism to perform its catalytic activity.

Project 4: SAXS of *x*LTA4H

The recent structure determination by crystallography of *x*LTA4H show that this protein is a dimer instead of a monomer as seen in other LTA4H related proteins. In order to check that this is not an artifact caused by the crystal restraints, SAXS studies were used to study the state of the protein in solution. The protein was further purified by size exclusion chromatography (SEC) and data was collected for the inhibitor complexed form at BM29, ESRF, France (paper IV). After initial scaling of the data a GASBOR model (*ab initio*) was built. As the data clearly shows that *x*LTA4H is a dimer in high to low protein concentrations it confirms the crystal structure that this LTA4H is a dimer while the other LTA4H (human and yeast) are monomeric enzymes.

Project 5: SAXS/WAXS studies on *hum*LTA4H, *sc*LTA4H and *x*LTA4H

The conformational change shown by *sc*LTA4H is not observed for *hum*LTA4H, despite a large set of different structures for the latter. As *hum*LTA4H is crystallized in the presence of a metal, required to form crystal contacts between the symmetry related molecules in the crystals, it was thought that the lack of observations for conformational change could be due to the constraints of the crystal lattice. Solution studies by SAXS were initiated to study the behavior of the protein.

SAXS data were collected of these three proteins with or without their inhibitor. It was observed that *x*LTA4H shows a more compact form upon bestatin binding (paper V & figure 33) as judged by the difference in scattering patterns. *Hum*LTA4H did not show any major difference in scattering patterns upon binding to bestatin or the powerful inhibitor SC-57461A and it seems therefore that this enzyme does not display larger conformational changes. Samples of *sc*LTA4H showed an aggregation problem and recent data processing indicated that this protein, like *x*LTA4H, also forms dimers in solution and needs to be further processed by the program SASREF where modeling of multisubunit complex is possible using the information from the crystal structure [105].

6. CONCLUSION AND FUTURE PROSPECTS

The MID protein does contain portions of coiled coil structure and it is predicted to be flexible. These properties might have contributed towards difficulty to obtain good crystals. Recloning of the MID⁷⁶⁴⁻⁹¹³ domain with a cleavable tag is planned, as a tag can be a major impediment for obtaining well diffracting crystals. From sequence analysis it is shown that both the IgD binding domain MID⁹⁶²⁻¹²⁰⁰ as well as the adhesin domain MID⁷⁶⁴⁻⁹¹³ do not contain any methionine. Recently our group was successful in introducing a Met residue into a protein sequence for phasing purposes [106]. A similar approach for the MID domains with introducing methionine residues for phasing purposes is also planned. This will allow collecting Se-met data and solving the phase problem by MAD methods.

Human IgD, unlike other Igs, has only been poorly studied due to the low concentration in normal serum and its great susceptibility to proteolytic degradation. This makes that it is difficult to purify in larger amounts and quantify. The purification method with using MID⁹⁶²⁻¹²⁰⁰ GST-linked resin will be further explored in the future to investigate the possibilities to increase the yield of IgD required for SAXS, crystallization attempts and other studies.

On-line SAXS experiments will be further employed in the case of *scLTA4H*, to avoid the quick aggregation that occurs for this protein. The studies on LTA4H enzymes should be extended to the WAXS regime in order to observe if smaller conformational changes occur.

Finally the structural information obtained from the projects has shown that MID is a trimer and as such a true member of the family of autotransporters. It has been confirmed that it is an elongated protein with predominantly a beta structure and only a limited amount of alpha helical structure. Still a full structure determination is required to obtain more information for this otherwise unusual adhesin.

For LTA4H the work described in this thesis shows that this class of proteins is more variable than originally thought. *ScLTA4H* showed conformational changes not observed before and unexpectedly *xLTA4H* is a dimer both in crystal and also as this

thesis confirms in solution. Implications for the evolution of this class of very interesting proteins carrying two different enzymatic activities are intriguing. Furthermore as *hum*LTA4H is an attractive possible drug design target, these studies will also have an impact on that type of research.

7. Acknowledgements

During little more than 4 years of my doctoral studies the life was full with new experiences and there are many people I would like to thank.

First of all I would like to thank my supervisor Marjolein Thunnissen for bringing me into the structural biology world and mentoring me with great patience.

Thanks to my co-supervisor Professor Kristian Riesbeck as well. My BSc thesis supervisor Mahbub-e-Sobhani for constantly inspiring me to come to research world. Also Professor Sofi and Patrik for creating foundation of knowledge to continue doing research.

A profound gratitude to Tomás Plivelic for introducing me to SAXS data collection and showing initial processing.

Raymond (also for many useful discussions) & Karolina for introducing me to some instrument and techniques.

Thanks to my Masters students Vivek & Arman for helping me while your learning process, and also short time student Manjuma & Saleh.

Kamrul for many recreational gossips and proofreading.

Thanks Renzo for your friendliness. Christopher S. for discussion about data and other people from structural biology group Wael, Oskar, Eva-Christina, Janina, Saraboji, Maria, Martin, Harm for giving company at the department, synchrotron visit & meetings, discussion, etc.

Birendra & Tamim for experimental discussion and other activities.

Stefan, Jennifer & Susanna for filling up the lab and company to recent synchrotron visit.

Course assistant colleagues Adine, Christopher P., Hanna, Tania, Robert, Irem for making a good team effort.

The Bangladeshi community in Lund & Malmö for arranging all recreational activity.

Finally, I would like to thank my wife Mustakima for being around every way possible.

REFERENCES

1. de Vries, S. P., Bootsma, H. J., Hays, J. P., & Hermans, P. W. (2009). Molecular aspects of *Moraxella catarrhalis* pathogenesis. *Microbiology and molecular biology reviews*, 73(3), 389-406.
2. Bootsma, H. J., van Dijk, H. A. N. S., Verhoef, J., Fleer, A., & Mooi, F. R. (1996). Molecular characterization of the BRO beta-lactamase of *Moraxella (Branhamella) catarrhalis*. *Antimicrobial agents and chemotherapy*, 40(4), 966-972.
3. Mazzulli, T. (2005). Resistant microorganisms in head and neck infections. *Journal of otolaryngology*, 34.
4. Johnson, D. M., Sader, H. S., Fritsche, T. R., Biedenbach, D. J., & Jones, R. N. (2003). Susceptibility trends of *Haemophilus influenzae* and *Moraxella catarrhalis* against orally administered antimicrobial agents: five-year report from the SENTRY Antimicrobial Surveillance Program. *Diagnostic microbiology and infectious disease*, 47(1), 373-376.
5. Goldstein, E. J., Murphy, T. F., & Parameswaran, G. I. (2009). *Moraxella catarrhalis*, a human respiratory tract pathogen. *Clinical infectious diseases*, 49(1), 124-131.
6. Hallström, T., Müller, S. A., Mörgelin, M., Möllenkvist, A., Forsgren, A., & Riesbeck, K. (2008). The *Moraxella* IgD-binding protein MID/Hag is an oligomeric autotransporter. *Microbes and infection*, 10(4), 374-381.
7. Vidakovics, M. L. A. P., Jendholm, J., Mörgelin, M., Månsson, A., Larsson, C., Cardell, L. O., & Riesbeck, K. (2010). B cell activation by outer membrane vesicles—a novel virulence mechanism. *PLoS pathogens*, 6(1), e1000724.
8. Leo, J. C., Grin, I., & Linke, D. (2012). Type V secretion: mechanism (s) of autotransport through the bacterial outer membrane. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592), 1088-1101.
9. Samuelsson, M., Jendholm, J., Amisten, S., Morrison, S. L., Forsgren, A., & Riesbeck, K. (2006). The IgD CH1 region contains the binding site for the human respiratory pathogen *Moraxella catarrhalis* IgD-binding protein MID. *European journal of immunology*, 36(9), 2525-2534.
10. Forsgren, A., Brant, M., Möllenkvist, A., Muyombwe, A., Janson, H., Woin, N., & Riesbeck, K. (2001). Isolation and characterization of a novel IgD-binding protein from *Moraxella catarrhalis*. *The Journal of Immunology*, 167(4), 2112-2120.

11. Pearson, M. M., Lafontaine, E. R., Wagner, N. J., Geme, J. W. S., & Hansen, E. J. (2002). A hag mutant of *Moraxella catarrhalis* strain O35E is deficient in hemagglutination, autoagglutination, and immunoglobulin D-binding activities. *Infection and immunity*, 70(8), 4523-4533.
12. Balder, R., Krunkosky, T. M., Nguyen, C. Q., Feezel, L., & Lafontaine, E. R. (2009). Hag mediates adherence of *Moraxella catarrhalis* to ciliated human airway cells. *Infection and immunity*, 77(10), 4597-4608.
13. Bullard, B., S. Lipski, and E. R. Lafontaine. 2007. Regions important for the adhesin activity of *Moraxella catarrhalis* Hag. *BMC Microbiol.* 7:65.
14. Nordström, T., Forsgren, A., & Riesbeck, K. (2002). The immunoglobulin D-binding part of the outer membrane protein MID from *Moraxella catarrhalis* comprises 238 amino acids and a tetrameric structure. *Journal of Biological Chemistry*, 277(38), 34692-34699.
15. Nordström, T., Jendholm, J., Samuelsson, M., Forsgren, A., & Riesbeck, K. (2006). The IgD-binding domain of the *Moraxella* IgD-binding protein MID (MID962-1200) activates human B cells in the presence of T cell cytokines. *Journal of leukocyte biology*, 79(2), 319-329.
16. Riesbeck, K., Tan, T. T., & Forsgren, A. (2006). MID and UspA1/A2 of the human respiratory pathogen *Moraxella catarrhalis*, and interactions with the human host as basis for vaccine development. *ACTA BIOCHIMICA POLONICA-ENGLISH EDITION*-, 53(3), 445.
17. Forsgren, A., Brant, M., & Riesbeck, K. (2004). Immunization with the Truncated Adhesin *Moraxella catarrhalis* Immunoglobulin D—Binding Protein (MID764–913) Is Protective against *M. catarrhalis* in a Mouse Model of Pulmonary Clearance. *Journal of Infectious Diseases*, 190(2), 352-355.
18. Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M., & Goldman, A. (2004). The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel β -roll. *The EMBO journal*, 23(4), 701-711.
19. Tiyawisutsri, R., Holden, M. T., Tumapa, S., Rengpipat, S., Clarke, S. R., Foster, S. J., ... & Peacock, S. J. (2007). *Burkholderia* Hep_Hag autotransporter (BuHA) proteins elicit a strong antibody response during experimental glanders but not human melioidosis. *BMC microbiology*, 7(1), 19.
20. Biegert, A., Mayer, C., Remmert, M., Söding, J., & Lupas, A. N. (2006). The MPI Bioinformatics Toolkit for protein sequence analysis. *Nucleic acids research*, 34(suppl 2), W335-W339.

21. Szczesny, P., & Lupas, A. (2008). Domain annotation of trimeric autotransporter adhesins—daTAA. *Bioinformatics*, 24(10), 1251-1256.
22. Bleves, S., Viarre, V., Salacha, R., Michel, G. P., Filloux, A., & Voulhoux, R. (2010). Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons. *International Journal of Medical Microbiology*, 300(8), 534-543.
23. Cotter, S. E., Surana, N. K., & St Geme III, J. W. (2005). Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends in microbiology*, 13(5), 199-205.
24. Linke, D., Riess, T., Autenrieth, I. B., Lupas, A., & Kempf, V. A. (2006). Trimeric autotransporter adhesins: variable structure, common function. *Trends in microbiology*, 14(6), 264-270.
25. Meng, G., St Geme III, J. W., & Waksman, G. (2008). Repetitive Architecture of the *Haemophilus influenzae* Hia Trimeric Autotransporter. *Journal of molecular biology*, 384(4), 824-836.
26. Łyskowski, A., Leo, J. C., & Goldman, A. (2011). Structure and biology of trimeric autotransporter adhesins. In *Bacterial Adhesion* (pp. 143-158). Springer Netherlands.
27. Edwards, T. E., Phan, I., Abendroth, J., Dieterich, S. H., Masoudi, A., Guo, W., ... & Stewart, L. J. (2010). Structure of a *Burkholderia pseudomallei* trimeric autotransporter adhesin head. *PLoS One*, 5(9), e12803.
28. Möllenkvist, A., Nordström, T., Halldén, C., Christensen, J. J., Forsgren, A., & Riesbeck, K. (2003). The *Moraxella catarrhalis* immunoglobulin D-binding protein MID has conserved sequences and is regulated by a mechanism corresponding to phase variation. *Journal of bacteriology*, 185(7), 2285-2295.
29. Wechsler, M. E., & Israel, E. (2002). Pharmacogenetics of treatment with leukotriene modifiers. *Current opinion in allergy and clinical immunology*, 2(5), 395-401.
30. Samuelsson, B. (1983). Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation. *Science*, 220(4597), 568-575.
31. Hedi, H., & Norbert, G. (2004). 5-Lipoxygenase pathway, dendritic cells, and adaptive immunity. *BioMed Research International*, 2004(2), 99-105.
32. Iovannisci, D. M., Lammer, E. J., Steiner, L., Cheng, S., Mahoney, L. T., Davis, P. H., ... & Burns, T. L. (2007). Association Between A Leukotriene C4 Synthase Gene Promoter

Polymorphism and Coronary Artery Calcium in Young Women The Muscatine Study.
Arteriosclerosis, thrombosis, and vascular biology, 27(2), 394-399.

33. Orning, L., Gierse, J. K., & Fitzpatrick, F. A. (1994). The bifunctional enzyme leukotriene-A4 hydrolase is an arginine aminopeptidase of high efficiency and specificity. *Journal of Biological Chemistry*, 269(15), 11269-11273.
34. Stsiapanava, A., Olsson, U., Wan, M., Kleinschmidt, T., Rutishauser, D., Zubarev, R. A., ... & Haeggström, J. Z. (2014). Binding of Pro-Gly-Pro at the active site of leukotriene A4 hydrolase/aminopeptidase and development of an epoxide hydrolase selective inhibitor. *Proceedings of the National Academy of Sciences*, 111(11), 4227-4232.
35. Haeggström, J. Z., Tholander, F., & Wetterholm, A. (2007). Structure and catalytic mechanisms of leukotriene A4 hydrolase. *Prostaglandins & other lipid mediators*, 83(3), 198-202.
36. Haeggström, J. Z. (2004). Leukotriene A4 hydrolase/aminopeptidase, the gatekeeper of chemotactic leukotriene B4 biosynthesis. *Journal of Biological Chemistry*, 279(49), 50639-50642
37. Thunnissen, M. M., Nordlund, P., & Haeggström, J. Z. (2001). Crystal structure of human leukotriene A4 hydrolase, a bifunctional enzyme in inflammation. *Nature Structural & Molecular Biology*, 8(2), 131-135.
38. Tholander, F., Kull, F., Ohlson, E., Shafqat, J., Thunnissen, M. M., & Haeggström, J. Z. (2005). Leukotriene A4 hydrolase, insights into the molecular evolution by homology modeling and mutational analysis of enzyme from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 280(39), 33477-33486.
39. Helgstrand, C., Hasan, M., Uysal, H., Haeggström, J. Z., & Thunnissen, M. M. (2011). A Leukotriene A4 Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding. *Journal of Molecular Biology*, 406(1), 120-134.
40. Van Den Berg, S., Löfdahl, P. Å., Härd, T., & Berglund, H. (2006). Improved solubility of TEV protease by directed evolution. *Journal of biotechnology*, 121(3), 291-298.
41. Stsiapanava A, Tholander F, Kumar R, Qureshi A, Niegowski D, Hasan M, Thunnissen M, Haeggström J. Z. and Rinaldo-Matthis A (2014). Product Formation Controlled By Substrate Dynamics In Leukotriene A4 Hydrolase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(2), 439–446.
42. Jancarik, J. A. K. S., & Kim, S. H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of applied crystallography*, 24(4), 409-411.

43. McPherson, A., Koszelak, S., Axelrod, H., Day, J., Williams, R., Robinson, L., ... & Cascio, D. (1986). An experiment regarding crystallization of soluble proteins in the presence of beta-octyl glucoside. *Journal of Biological Chemistry*, 261(4), 1969-1975.
44. Yamada, H., Tamada, T., Kosaka, M., Miyata, K., Fujiki, S., Tano, M., ... & Kuroki, R. (2007). 'Crystal lattice engineering,' an approach to engineer protein crystal contacts by creating intermolecular symmetry: Crystallization and structure determination of a mutant human RNase 1 with a hydrophobic interface of leucines. *Protein science*, 16(7), 1389-1397.
45. Loris, R., Marianovsky, I., Lah, J., Laeremans, T., Engelberg-Kulka, H., Glaser, G., ... & Wyns, L. (2003). Crystal structure of the intrinsically flexible addiction antidote MazE. *Journal of Biological Chemistry*, 278(30), 28252-28257.
46. McPherson, A. (1999). *Crystallization of biological macromolecules*. Cold Spring Harbor Laboratory Press.
47. Luft, J. R., & DeTitta, G. T. (1999). A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography*, 55(5), 988-993.
48. Stura, E. A., & Wilson, I. A. (1990). Analytical and production seeding techniques. *Methods*, 1(1), 38-49.
49. Stura, E. A., & Wilson, I. A. (1991). Applications of the streak seeding technique in protein crystallization. *Journal of crystal growth*, 110(1), 270-282.
50. Bergfors, T. (2003). Seeds to crystals. *Journal of structural biology*, 142(1), 66-76.
51. Chayen, N. E., & Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, 5(2), 147-153.
52. Incardona, M. F., Bourenkov, G. P., Levik, K., Pieritz, R. A., Popov, A. N., & Svensson, O. (2009). EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *Journal of synchrotron radiation*, 16(6), 872-879.
53. Popov, A. N., & Bourenkov, G. P. (2003). Choice of data-collection parameters based on statistic modelling. *Acta Crystallographica Section D: Biological Crystallography*, 59(7), 1145-1153.
54. Dauter, Z. (1999). Data-collection strategies. *Acta Crystallographica Section D: Biological Crystallography*, 55(10), 1703-1717.

55. Heras, B., & Martin, J. L. (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallographica Section D: Biological Crystallography*, 61(9), 1173-1180.
56. Dobrianov, I., Caylor, C., Lemay, S. G., Finkelstein, K. D., & Thorne, R. E. (1999). X-ray diffraction studies of protein crystal disorder. *Journal of crystal growth*, 196(2), 511-523.
57. Karplus, P. A., & Diederichs, K. (2012). Linking crystallographic model and data quality. *Science*, 336(6084), 1030-1033.
58. Wlodawer, A., Minor, W., Dauter, Z., & Jaskolski, M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *Febs Journal*, 275(1), 1-21.
59. Joyce, M. G., Radaev, S., & Sun, P. D. (2010). A rational approach to heavy-atom derivative screening. *Acta Crystallographica Section D: Biological Crystallography*, 66(4), 358-365.
60. Bella, J., & Rossmann, M. G. (1998). A general phasing algorithm for multiple MAD and MIR data. *Acta Crystallographica Section D: Biological Crystallography*, 54(2), 159-174.
61. Svergun, D. I., & Koch, M. H. (2003). Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, 66(10), 1735.
62. Tsutakawa, S. E., Hura, G. L., Frankel, K. A., Cooper, P. K., & Tainer, J. A. (2007). Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. *Journal of structural biology*, 158(2), 214-223.
63. Vestergaard, B., Sanyal, S., Roessle, M., Mora, L., Buckingham, R. H., Kastrop, J. S., ... & Ehrenberg, M. (2005). The SAXS solution structure of RF1 differs from its crystal structure and is similar to its ribosome bound cryo-EM structure. *Molecular cell*, 20(6), 929-938.
64. Wang, S., Overgaard, M. T., Hu, Y., & McKay, D. B. (2008). The Bacillus subtilis RNA Helicase YxiN is Distended in Solution. *Biophysical journal*, 94(1), L01-L03.
65. Hura, G. L., Budworth, H., Dyer, K. N., Rambo, R. P., Hammel, M., McMurray, C. T., & Tainer, J. A. (2013). Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. *Nature methods*, 10(6), 453-454.
66. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., & Svergun, D. I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *Journal of the American Chemical Society*, 129(17), 5656-5664.

67. Currie, M. A., Cameron, K., Dias, F. M., Spencer, H. L., Bayer, E. A., Fontes, C. M., ... & Jia, Z. (2013). Small angle X-ray scattering analysis of *Clostridium thermocellum* cellulosome N-terminal complexes reveals a highly dynamic structure. *Journal of Biological Chemistry*, 288(11), 7978-7985.
68. Guilbaud, J. B., & Saiani, A. (2011). Using small angle scattering (SAS) to structurally characterise peptide and protein self-assembled materials. *Chemical Society Reviews*, 40(3), 1200-1210.
69. Gabel, F. (2011). Combining Small-angle Neutron and X-ray Scattering for Studying Protein Denaturation. *Neutron News*, 22(3), 20-23.
70. Schnablegger, H., & Singh, Y. (2011). The SAXS guide: Getting acquainted with the principles. *Anton Paar GmbH*, 2.
71. Petoukhov, M. V., & Svergun, D. I. (2007). Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Current opinion in structural biology*, 17(5), 562-571.
72. Neutze, R., & Moffat, K. (2012). Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. *Current opinion in structural biology*, 22(5), 651-659.
73. Receveur-Bréchet, V., & Durand, D. (2012). How random are intrinsically disordered proteins? A small angle scattering perspective. *Current protein & peptide science*, 13(1), 55.
74. Whitten, A. E., & Trehwella, J. (2009). Small-angle scattering and neutron contrast variation for studying bio-molecular complexes. In *Micro and Nano Technologies in Bioanalysis* (pp. 307-323). Humana Press.
75. Mylonas, E., & Svergun, D. I. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *Journal of applied crystallography*, 40, s245-s249.
76. Orthaber, D., Bergmann, A., & Glatter, O. (2000). SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *Journal of Applied Crystallography*, 33(2), 218-225.
77. Fischer, H., Oliveira Neto, M. D., Napolitano, H. B., Polikarpov, I., & Craievich, A. F. (2009). Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *Journal of Applied Crystallography*, 43(1), 101-109.

78. Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly reviews of biophysics*, 40(03), 191-285.
79. Rambo, R. P., & Tainer, J. A. (2013). Super-resolution in solution X-ray scattering and its applications to structural systems biology. *Annual review of biophysics*, 42, 415-441.
80. Lobanov, M. Y., Bogatyreva, N. S., & Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, 42(4), 623-628.
81. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, 25(4), 495-503.
82. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., ... & Svergun, D. I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography*, 45(2), 342-350.
83. Rambo, R. P., & Tainer, J. A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, 95(8), 559-571.
84. Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal*, 76(6), 2879-2886.
85. Franke, D., & Svergun, D. I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 42(2), 342-346.
86. Mertens, H. D., & Svergun, D. I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of structural biology*, 172(1), 128-141.
87. Jacques, D. A., & Trewthella, J. (2010). Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls. *Protein Science*, 19(4), 642-657.
88. Svergun, D. I., Petoukhov, M. V., & Koch, M. H. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical journal*, 80(6), 2946-2953.
89. Volkov, V. V., & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 36(3), 860-864.

90. Patel, T. R., Reuten, R., Xiong, S., Meier, M., Winzor, D. J., Koch, M., & Stetefeld, J. (2012). Determination of a molecular shape for netrin-4 from hydrodynamic and small angle X-ray scattering measurements. *Matrix Biology*, 31(2), 135-140.
91. Kozin, M. B., & Svergun, D. I. (2001). Automated matching of high-and low-resolution structural models. *Journal of Applied Crystallography*, 34(1), 33-41.
92. Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRY SOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, 28(6), 768-773.
93. Wallace, B. A. (2009). Protein characterisation by synchrotron radiation circular dichroism spectroscopy. *Quarterly reviews of biophysics*, 42(04), 317-370.
94. Greenfield, N. J. (2007). Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols*, 1(6), 2876-2890.
95. Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1751(2), 119-139.
96. Miles, A. J., & Wallace, B. A. (2006). Synchrotron radiation circular dichroism spectroscopy of proteins and applications in structural and functional genomics. *Chemical Society Reviews*, 35(1), 39-51.
97. Greenfield, N. J., & Hitchcock-Degregori, S. E. (1993). Conformational intermediates in the folding of a coiled-coil model peptide of the N-terminus of tropomyosin and $\alpha\alpha$ -tropomyosin. *Protein Science*, 2(8), 1263-1273.
98. Janes, R. W., & Cuff, A. L. (2005). Overcoming protein denaturation caused by irradiation in a high-flux synchrotron radiation circular dichroism beamline. *Journal of synchrotron radiation*, 12(4), 524-529.
99. Hussain, R., Jávorfí, T., & Siligardi, G. (2011). Circular dichroism beamline B23 at the Diamond Light Source. *Journal of synchrotron radiation*, 19(1), 132-135.
100. Wallace, B. A. (2000). Conformational changes by synchrotron radiation circular dichroism spectroscopy. *Nature Structural & Molecular Biology*, 7(9), 708-709.
101. Sreerama, N., & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods in enzymology*, 318-350.

102. Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic acids research*, 32(suppl 2), W668-W673.
103. Manavalan, P., & Johnson Jr, W. C. (1987). Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Analytical biochemistry*, 167(1), 76-85.
104. Sreerama, N., & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical biochemistry*, 287(2), 252-260.
105. Petoukhov, M. V., & Svergun, D. I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical journal*, 89(2), 1237-1250.
106. Singh B, Al Jubair, T, Förnvik K, Thunnissen MM & Riesbeck K. Crystallization and X-ray diffraction analysis of a novel surface-adhesin protein: protein E from *Haemophilus influenzae*. (2012) *Acta Cryst F68*, 222-226.

SAXS measurements and preliminary crystallographic investigations show an elongated three-fold organization of MID⁹⁶²⁻¹²⁰⁰

Mahmudul Hasan¹, Kristian Riesbeck² and Marjolein MMGM Thunnissen¹

¹Department of Biochemistry and Structural Biology, Kemicentrum, Getingevägen 60, Lund University, 22100 Lund, Sweden.

²Medical Microbiology, Department of Laboratory Medicine Malmö, Lund University, Skåne University Hospital, 20502 Malmö, Sweden.

ABSTRACT

The gram-negative bacterium *Moraxella catarrhalis* is an effective human B-cell mitogen. This activity is mediated by an IgD binding outer-membrane protein: MID. Here we describe preliminary crystallization and solution scattering experiments on the Ig-D binding domain MID⁹⁶²⁻¹²⁰⁰ from MID of *Moraxella cararhallsis*. The crystals belong to space group P321 with cell dimensions $a = 56.5 \text{ \AA}$, $b = 56.5 \text{ \AA}$, $c = 241.9 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$ and $\gamma = 120^\circ$. A native dataset to 2.3 \AA was collected. Assuming that there is a dimer in the asymmetric unit the volume of Matthews (V_M) is $2.04 \text{ \AA}^3/\text{Dalton}$ with a solvent content of 39.93 %. Small angle X-ray scattering experiments subsequently showed that this domain forms trimers in solution. From *ab initio* low-resolution molecular envelope determinations based on the scattering data it can be seen that the domain has a fibrous elongated shape. The studies were complemented with circular dichroism and secondary structure predictions.

Keywords: MID, crystallization, circular dichroism, SAXS studies, OCA family

1. INTRODUCTION

Moraxella catarrhalis is an important cause of lower respiratory tract infections, particularly in adults with COPD [1] where it is the second major cause of exacerbations that lead to lost work time, hospital admissions, respiratory failure and sometimes death. In COPD and also in immuno-compromised hosts, the bacterium can cause a variety of severe infections including pneumonia, endocarditis, sepsis, and meningitis [2]. Because *M. catarrhalis* has long been considered a harmless commensal, relatively little is known about its pathogenic characteristics and virulence factors, although developments in this field of research have accelerated over the past decade.

The outer membrane proteins (OMPs) of *M. catarrhalis* are amongst the most important virulence factors for this bacterium. They can be divided into two groups: adhesins and proteins involved in nutrient acquisition. The most studied adhesins of *M. catarrhalis* are UspA1, UspA2 and UspA2H, which have been shown to be involved in the adhesion of *M. catarrhalis* to Chang conjunctival and HEp-2 laryngeal epithelial cells. Both UspA1 and UspA2 are involved in fibronectin binding [3]. Another adhesin is a 200 kDa OMP which is responsible for the IgD binding of *M. catarrhalis* namely the *Moraxella* IgD-binding protein MID [4], also known as HAG [5].

M. catarrhalis is an effective human B-cell mitogen [6]. It can directly stimulate B-cells without any recognition of T-cells and it can therefore be classified as a T-cell independent antigen. Only very few gram-negative bacteria have the capacity to bind Ig in a non-immune fashion, while more examples are known for gram-positive bacteria. It has been known for more than two decades that *M. catarrhalis* binds soluble IgD in a non-immune fashion. Since the soluble and membrane bound forms of IgD have many similarities it is likely that the mitogenic signal is carried through the membrane bound form of Ig-D on the surface of the B-cell. Binding is proposed to be to the CH1 domain of IgD [7].

From MID-deficient *Moraxella* mutants it could be seen that MID is the only outer membrane protein responsible for IgD binding. This protein binds specifically to both soluble as well as membrane bound IgD on B-cells. However MID does not bind to other immunoglobulins such as IgG, IgM, IgA, or IgE myeloma proteins. Apart from

IgD binding, MID is also an important adhesin capable of agglutinating human erythrocytes and binds to type II alveolar epithelial cells [8]. In contrast, *M. catarrhalis* isolates with low MID expression levels do not agglutinate erythrocytes and have a 50% lower adhesive capacity.

The capacity for IgD binding of MID has been located in an IgD binding domain of 239 amino acids (MID⁹⁶²⁻¹²⁰⁰) [9]. Early ultracentrifugation experiments and gel electrophoresis suggested that MID⁹⁶²⁻¹²⁰⁰ is a tetramer and that this multimeric form binds IgD 20-fold more efficient than the monomeric form [10]. The truncated MID⁹⁶²⁻¹²⁰⁰ is efficiently attracted to a standard IgD serum and purified myeloma IgD but not to IgG, IgM, or IgA myeloma sera. MID⁹⁶²⁻¹²⁰⁰ in itself can activate human B-cells independent from T-cells while in contrast a mutated form of MID lacking the IgD binding domain is not able to activate B-cells [11].

MID belongs to the oligomeric autotransporter group of proteins (The oligomeric coiled-coil adhesin or OCA family) and it shares sequence similarities with the *Y. enterocolitica* OMP YadA and the *H. influenza* Hia adhesin [12]. These proteins are representatives of a group of non-fimbrial adhesins, which are found in proteobacteria. In this group of proteins there are many prominent pathogenic determinants including *M. catarrhalis* UspA1 and A2, YadA, Hia and Hsf and more. Autotransporter proteins do not require energy or accessory factors to secrete themselves through the bacterial outer-membrane. Proteins in this family form in general trimeric complexes that are anchored to the outer membrane with a “head-rod-anchor” architecture. Commonly the head is located at the N-terminal end of the sequence while the anchor region is at the C-terminal end. The membrane bound anchor region forms a porin-like 12 stranded β -barrel [13]. The central part of these autotransporters is a flexible fibrous rod formed by long α -helices that projects the head from the cell-surface and varies greatly in length between the different OCA family members. The head forms the bulkier part of the structure and is responsible for the functional properties of the proteins such as eg. binding to proteins of the extra-cellular matrix or autoagglutination. Several structures of head domains have been solved by X-ray crystallography e.g. one from YadA [14] and one from Hia [15]. Both show a trimeric arrangement but each of these domains has a fundamentally different fold.

For most members of the OCA family the binding domains are located at the N-terminal end of the sequence. These members are in general between 45 and 80 kDa big [16]. MID belongs to a sub-class within the family that is larger and has a size of around 200 kDa. In the case of MID the two functional domains are located in the middle of the sequence while so-called Hep-Hag repeats that are characteristic for a b-roll structure can be found in both the N-terminal and C-terminal regions (see figure 1). Several short Haemmagglutinin motifs (HIM) can also be located in the sequence. However it has been showed that the adhesive properties of MID are not located in these Hep-Hag regions and the specific role of these domains has to be investigated further [17]. These findings indicate that MID is having different structure-function relationships than the related OCA proteins. From electron microscopy work it has been seen that MID forms indeed a fibrous protein with a globular head like the other OCA transporters however by using immuno electron microscopy it could be observed that MID is folding back upon itself [18]. The protein is membrane anchored in the C-terminal region and has an unique double folded stalk protruding approximately 100 nm in length. A globular head consisting of the functional domains for adhesion and IgD binding is resided at the very tip of the protein, distal to the cell membrane. MID is about double the size of YadA and its stalk is longer and more fibril-like.

In order to obtain more information how MID⁹⁶²⁻¹²⁰⁰ does recognize IgD and as tetrameric arrangement is not observed for any other autotransporter a more thorough investigation into the oligomeric organization is required. Therefore structural investigations using crystallography, CD spectroscopy and small angle scattering (SAXS) experiments were started. Secondary structure prediction databases were used to assess the structural composition of this domain since the sequence of MID⁹⁶²⁻¹²⁰⁰ is unique and no homologues can be found in databases for which structural information is available.

2. MATERIALS AND METHODS

2.1. Cloning of GST and His-tagged MID⁹⁶²⁻¹²⁰⁰

The template sequence of MID⁹⁶²⁻¹²⁰⁰ was amplified using Taq DNA polymerase from the previous clone [10] incorporating NcoI site on the forward primer

(CCATGGTTGACCAAACCAAAGGCTTA) and XhoI site in the reverse primer (CTCGAGCTATGGGGTTTGGGCTTGGG). The amplified template of MID⁹⁶²⁻¹²⁰⁰ was purified from 1.2% agarose gel by extraction kit (Qiagen). Plasmid pETM-30 (a kind gift from Dr. G. Stier, EMBL, Hamburg) was purified using plasmid extraction kit (Qiagen) by following their protocol. Both the amplified sequence and the pETM-30 plasmid were digested by enzyme NcoI and XhoI and the products were then ligated using DNA ligase enzyme. The ligated product was transformed into host cell and cells were then plated on agar media containing kanamycin and chloramphenicol. The plasmid from the grown single colony was purified and sequenced by 3130 Genetic analyzer sequencer (Applied Biosystems) using Big Dye enzyme. All the enzymes and reagents used for cloning were purchased from Fermentas and their protocols were followed.

2.2. Expression and Purification of GST-(His)₆-tagged MID⁹⁶²⁻¹²⁰⁰

The GST-(His)₆-tagged protein was expressed in BL21 (DE3) pLysS cells by growing 1-liter culture in a 3 liter baffled flask at 37 °C to OD 0.5, lowering the temperature to 20 °C and inducing expression with 1 mM IPTG at OD 0.7–0.8, followed by overnight incubation at 20 °C. The cells were then harvested by centrifuging at 6000 rpm and the pellets were frozen. A frozen cell pellet from 0.5 liters of expression was resuspended in 50 ml of 100 mM Tris-HCl (pH 8.5), 300 mM NaCl, 10 mM NaN₃ containing complete mini EDTA-free protease inhibitor (Roche) and lysed by sonication for 2x5 min. The insoluble cell debris was removed by centrifugation at 30,000g for 30 min, and the lysate filtered through a 0.45-µm syringe filter. The cleared lysate was mixed with 5 ml of glutathione Sepharose 4B (GE Healthcare) that had been pre-equilibrated with 100mM Tris-HCl (pH 8.5), 300mM NaCl, 5mM NaN₃ and was incubated for 60 min on a shaking platform at room temperature. The lysate/resin mix was transferred to a gravity flow column (BioRad EconoPack), and the resin was allowed to drain. After washing with several column volumes of 100 mM Tris-HCl (pH 8.5), 300 mM NaCl, 5 mM NaN₃, the bound protein was eluted with 3 column volumes of 100 mM Tris-HCl (pH 8.5), 300 mM NaCl, 5 mM NaN₃, 10 mM reduced glutathione (AppliChem).

The eluted MID⁹⁶²⁻¹²⁰⁰ protein was mixed with TEV protease produced in the lab (40 µg for each estimated mg of protein) following the protocol described previously [19] and

incubated at room temperature overnight. To remove the cleaved affinity tag, uncleaved protein and TEV protease, the protein solution was adjusted to 500 mM NaCl and 10 mM imidazole and applied in a gravity flow column containing 5 ml Ni²⁺Sepharose FF (GE healthcare) resin that was pre-equilibrated with 100 mM Tris-HCl (pH 8.5), 300 mM NaCl and 10 mM imidazole. After 10 min of incubation, the gravity flow column was unplugged and the resin was allowed to drain. This flow-through contained the cleaved MID⁹⁶²⁻¹²⁰⁰.

MID⁹⁶²⁻¹²⁰⁰ was further purified by injecting in a Superdex 200 10/300 (GE Healthcare) size exclusion chromatography column pre-equilibrated with filtered and degassed 30 mM Tris-HCl (pH 8.5) and 150 mM NaCl with a flow rate of 0.3 ml/min. The trimeric form of the protein eluted at ~13.5 ml and corresponding peaks were pooled and concentrated using Amicon Ultra Centrifugal Filter (Millipore) with cut off value of 5 kDa.

2.3. Crystallization

An initial small crystal hit was obtained from the PACT screen (Molecular Dimensions) by sitting drop vapor diffusion method with a protein to reservoir volume ratio of 200:200 nl using a Mosquito nanolitre pipetting robot (TTP Labtech). These crystals were small and needed to be optimized. After setting up experiments with the additive screen (Hampton Research) better quality crystals were obtained in presence of 1, 8-Diaminooctane. MID⁹⁶²⁻¹²⁰⁰ at concentration 9 mg/ml was mixed with 1, 8-Diaminooctane to a final concentration of 50 mM in the mix and the PACT screen was repeated. Good quality crystals grew after 4 months in 2 mM ZnCl₂, 100 mM Tris-HCl (pH 8.0) and 20% PEG 6000. All crystallizations were performed at 20° C.

2.4. Diffraction data collection and processing

As the crystallization themselves conditions provide cryo-protection, crystals could be cooled without ice formation, straight from the crystallization drops. Data were collected from a single MID⁹⁶²⁻¹²⁰⁰ crystal at beamline ID29 at the ESRF, Grenoble, France, using a MD2 diffractometer with a Pilatus 6M Detector (Dectris) [20]. The wavelength used for data collection was 1.2825 Å. The data were collected to a

resolution of 2.3 Å. In total 1800 frames of each 0.1° of data were collected. The data were indexed, integrated and scaled using the program XDS [21] while further data manipulations were carried out using programs from the CCP4 package [22]. Further details on crystal parameters and data quality can be found in table 1.

2.5. SAXS data collection

Data were collected on MID⁹⁶²⁻¹²⁰⁰ at 3-9 mg/ml concentrations with an exposure time of 300 second at station I911-4 at the MAX IV Laboratory [23]. For all analysis, the used wavelength was 0.91 Å and the beam size was 0.3x0.3 mm². All samples were placed in quartz capillary maintaining temperature at 10° C, with a distance of 2 m from the detector (two-dimensional MAR165 CCD) and covered a scattering vector range ($q=4\pi\sin\theta/\lambda$) from 0.01 to 0.325 Å⁻¹. No-beam, empty cell, water and buffer (before and after protein sample) scattering data were collected. The no-beam, empty cell and buffer data were subtracted from protein scattering data and obtained 2D images were reduced to one-dimensional scattering profiles using software Bli711 [24].

The monodispersity of the protein samples was measured by dynamic light scattering using a Zetasizer Nano-ZS (Malvern) instrument and data were analyzed using their software. Protein concentrations were determined by A_{280} measurements using a NanoDrop spectrophotometer (Thermo Scientific) and molar extinction coefficient (10430 M⁻¹cm⁻¹) calculated in ProtParam [25].

2.6. SAXS data processing

The datasets collected from the different concentrations of protein were compared to each other in order to examine the presence of aggregation and high and low concentration of data were merged to one file. Initial data scaling along with the experimental radius of gyration (R_g) and the forward scattering intensity $I(0)$ were calculated from data at low q values in the range of $qR_g < 1.3$, using the Guinier approximation: $\ln I(q) \approx \ln I(0) - R_g^2 q^2/3$ using the software Primus [26]. A pair distribution function was calculated by GNOM [27]. *Ab initio* model calculations for the molecular envelop of MID⁹⁶²⁻¹²⁰⁰ were performed by the program DAMMIN [28]. The initial 15 models were averaged using DAMAVER [29]. Further details about data

processing can be found in table 2.

2.7. Circular Dichroism data collection

Synchrotron radiation circular dichroism (SRCD) spectra for MID⁹⁶²⁻¹²⁰⁰ at concentration of 8 mg/ml were collected at beamline CD1 of the ASTRID storage ring at the University of Aarhus, Denmark. All sample and baseline spectra were collected in quartz cell, maintaining the temperature at 20° C, measuring over the wavelengths 280 to 165 nm decreasing 1 nm in every step. Each spectrum was collected 3 times. The data were averaged and buffer data subtracted from sample using CDtool [30].

CD spectra of MID⁹⁶²⁻¹²⁰⁰ at a concentration of 0.1 mg/ml and buffer were recorded using a Jasco J-815 (Jasco) and were averaged using manufacturer's software. MID⁹⁶²⁻¹²⁰⁰ was purified using 20 mM sodium phosphate (pH 8.0) as running buffer using size exclusion chromatography. All sample and baseline spectra were collected in quartz cells (path length 1 cm) over the wavelengths 260 to 190 nm with decreasing 1 nm in every step while keeping the temperature at 20° C. The secondary structure content of MID⁹⁶²⁻¹²⁰⁰ was analyzed at the DichroWeb server [31].

2.8. Secondary Structure predictions

The sequence of MID⁹⁶²⁻¹²⁰⁰ was used as a query in a number of secondary structure programs, Porter [32], PredictProtein/PROF Proteus-2 [33], PsiPRED [34], SOPma [35] and SSPro [36] to be able to predict the presence of secondary structure elements in it. The server MUSTER [33] was used to thread the sequence of the domain against the library of known folds. The program GLOBE [37] which is part of the PredictProtein server was used to make an assessment of the overall shape of the protein.

3. RESULTS & DISCUSSION

3.2. Crystallographic

Preliminary crystallization experiments were started in order to solve the structure of the IgD binding domain of MID⁹⁶²⁻¹²⁰⁰. The first attempts to crystallize this domain were with using different screens. Optimization of relevant hits gave an improvement of crystal-size and single crystals with dimensions of about 0.17 mm x 0.03 mm x 0.02 mm were obtained in condition containing 2 mM ZnCl₂, 100 mM Tris-HCl (pH 8.0) and 20% PEG 6000. The crystal belonged to space group P321 with cell-dimensions $a = 56.5 \text{ \AA}$, $b = 56.5 \text{ \AA}$, $c = 241.9 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$ and $\gamma = 120^\circ$. The dataset was 99.6% complete with an Rmerge of 9.1%. Further details on the crystal parameters and data processing statistics can be found in table 1.

From the cell dimensions and the size of the protein (242 amino acids) it seems that there is dimer in the asymmetric unit. This gives rise to a $V_M = 2.04 \text{ \AA}^3/\text{dalton}$ and a solvent content of 39.93% (calculated using a molecular weight of 24688 dalton for the domain). These values are at the lower end of the normal observed range (between 1.8 and $4.2 \text{ \AA}^3/\text{dalton}$ for the V_M and between 27 and 78% for the average solvent content in protein crystals) [Matthews, 1968]. More than two copies of the monomer in the asymmetric unit are thus highly unlikely while only one copy in the asymmetric unit would lead to an unusual high solvent content.

3.3. SAXS

The scattering profiles at different concentrations showed monodisperse distribution with no signs of aggregation. The Kratky plot (figure 5) indicates a folded state and not a multi-domain protein with a flexible linker. The pair distribution function indicated a very elongated shape of the molecule (figure 3). The *ab initio* model obtained from using the program DAMMIN showed that a good fitting for a trimeric state of MID⁹⁶²⁻¹²⁰⁰ could be obtained, while a tetrameric state of MID⁹⁶²⁻¹²⁰⁰ was not supported by the data. The envelopes calculated show a fibrous elongated shape (figure 4), which is a typical feature of trimeric autotransporter proteins. The P(r) curves shows a clear secondary bump at increasing distance after the peak (figure 3), a feature consistent with

a solution structure in which two or more domains are distended from each other rather than condensed into a compact assembly.

3.4. Circular Dichroism data deconvolution

All calculations for estimating the secondary structure content for MID⁹⁶²⁻¹²⁰⁰ were performed using Dichroweb. The deconvolution of the spectra indicated a large fraction of beta structure. A typical solution with a good NRMSD (normalized root-mean-square deviation) indicated a secondary structure content of 11% α -helix, 32% β -sheet, 25% turns and 34% unordered or coil structures.

3.5. Secondary structure prediction

A series of calculations was performed to obtain more information on the structural content of MID⁹⁶²⁻¹²⁰⁰. The series of secondary structure predictions gave a united picture of a protein with about 30 to 38% beta structure and 16 to 20% alpha structure (figure 7). A set of further programs was used to probe the structure of MID⁹⁶²⁻¹²⁰⁰ further. In order to obtain more information about the globularity of the domain, a prediction program for globularity (GLOBE) from the PredictProtein suite was used. The results from this program indicate that MID⁹⁶²⁻¹²⁰⁰ does not seem to fold as globular protein but rather an elongated fold. In order to probe whether there was any relationships to known fold, the fold recognition server MUSTER [38] was used to investigate whether MID⁹⁶²⁻¹²⁰⁰ shared a common fold with other proteins. From sequence homology searches alone, using BLAST or FASTA, did not point to a particular fold and no distant homologues or any other distantly related proteins could be found.

4. CONCLUSIONS

The data presented in this paper seem to indicate that the MID⁹⁶²⁻¹²⁰⁰ domain has an elongated form, containing predominantly beta structure in a trimeric arrangement. From excluded volume value calculations based on the SAXS data it can be concluded that MID⁹⁶²⁻¹²⁰⁰ forms a trimer in solution. In general protein molecules Porod (i.e., measurement of a hydrated protein) volumes in nm³ are approximately two to three

times larger than those expected given the molecular mass in kDa [39]. The Porod volume obtained for MID⁹⁶²⁻¹²⁰⁰ is 138 nm³ while the predicted molecular weight from Expasy is 25 kDa. This indicates that the molecule is not likely to be a dimer or tetramer and a trimeric arrangement fits best. As autotransporters are trimeric proteins, this is an observation that is consistent with MID being a member of this family of proteins.

The *Ab initio* envelope model based on the same data clearly indicates an elongated shape of the molecule which is similar as other members of the family of Autotransporters [40]. The CD and structure predictions using an array of programs both point to a structure with predominantly β -strand secondary structure elements, which is also seen in other members of the family. Furthermore the Kratky plot (figure 5) based on the SAXS data indicates the presence of an unfolded region, which is consistent with the CD data that also indicates a presence of unfolded structure. During the data processing of the SAXS data a wide variation of the radius of gyration values was observed which can be a characteristic for elongated and flexible molecules.

As there is no significant homology to proteins whose structure is deposited in the protein databank, the structure will have to be solved using ab-initio phasing methods. As the sequence of this domain does not contain any Methionine residues, which could be modified to Se-Met, a search for a suitable heavy atom derivative is ongoing.

REFERENCES

1. LaFontaine, E. R., Snipes, L. E., Bullard, B., Brauer, A. L., Sethi, S., & Murphy, T. F. (2009). Identification of domains of the Hag/MID surface protein recognized by systemic and mucosal antibodies in adults with chronic obstructive pulmonary disease following clearance of *Moraxella catarrhalis*. *Clinical and Vaccine Immunology*, *16*(5), 653-659.
2. Verduin, C. M., Hol, C., Fleer, A., van Dijk, H., & van Belkum, A. (2002). *Moraxella catarrhalis*: from emerging to established pathogen. *Clinical microbiology reviews*, *15*(1), 125-144.
3. Tan, T. T., Nordström, T., Forsgren, A., & Riesbeck, K. (2005). The respiratory pathogen *Moraxella catarrhalis* adheres to epithelial cells by interacting with fibronectin through ubiquitous surface proteins A1 and A2. *Journal of Infectious Diseases*, *192*(6), 1029-1038.
4. Forsgren, A., Brant, M., Möllenkvist, A., Muyombwe, A., Janson, H., Woin, N., & Riesbeck, K. (2001). Isolation and characterization of a novel IgD-binding protein from *Moraxella catarrhalis*. *The Journal of Immunology*, *167*(4), 2112-2120.
5. Pearson, M. M., Lafontaine, E. R., Wagner, N. J., Geme, J. W. S., & Hansen, E. J. (2002). A hag mutant of *Moraxella catarrhalis* strain O35E is deficient in hemagglutination, autoagglutination, and immunoglobulin D-binding activities. *Infection and immunity*, *70*(8), 4523-4533.
6. Banck, G. (1982). *Staphylococcus aureus* cowan I and *Branhamella catarrhalis* as B lymphocyte mitogens. Culture conditions for optimal DNA synthesis and selective stimulation of human B lymphocytes. *Journal of immunological methods*, *51*(3), 279-286.
7. Samuelsson, M., Jendholm, J., Amisten, S., Morrison, S. L., Forsgren, A., & Riesbeck, K. (2006). The IgD CH1 region contains the binding site for the human respiratory pathogen *Moraxella catarrhalis* IgD-binding protein MID. *European journal of immunology*, *36*(9), 2525-2534.
8. Riesbeck, K., Tan, T. T., & Forsgren, A. (2006). MID and UspA1/A2 of the human respiratory pathogen *Moraxella catarrhalis*, and interactions with the human host as basis for vaccine development. *ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-*, *53*(3), 445.
9. Nordström, T., Forsgren, A., & Riesbeck, K. (2002). The immunoglobulin D-binding part of the outer membrane protein MID from *Moraxella catarrhalis* comprises 238 amino acids and a tetrameric structure. *Journal of Biological Chemistry*, *277*(38), 34692-34699.

10. Hallström, T., Müller, S. A., Mörgelin, M., Möllenkvist, A., Forsgren, A., & Riesbeck, K. (2008). The *Moraxella* IgD-binding protein MID/Hag is an oligomeric autotransporter. *Microbes and infection*, *10*(4), 374-381
11. Nordström, T., Jendholm, J., Samuelsson, M., Forsgren, A., & Riesbeck, K. (2006). The IgD-binding domain of the *Moraxella* IgD-binding protein MID (MID962-1200) activates human B cells in the presence of T cell cytokines. *Journal of leukocyte biology*, *79*(2), 319-329.
12. Desvaux, M., Parham, N. J., & Henderson, I. R. (2004). The autotransporter secretion system. *Research in microbiology*, *155*(2), 53-60.
13. Meng, G., St Geme III, J. W., & Waksman, G. (2008). Repetitive Architecture of the *Haemophilus influenzae* Hia Trimeric Autotransporter. *Journal of molecular biology*, *384*(4), 824-836.
14. Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M., & Goldman, A. (2004). The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel β -roll. *The EMBO journal*, *23*(4), 701-711.
15. Yeo, H. J., Cotter, S. E., Laarmann, S., Juehne, T., St Geme, J. W., & Waksman, G. (2004). Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. *The EMBO journal*, *23*(6), 1245-1256.
16. Cotter, S. E., Surana, N. K., & St Geme III, J. W. (2005). Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends in microbiology*, *13*(5), 199-205.
17. Bullard, B., Lipski, S., & Lafontaine, E. R. (2007). Regions important for the adhesin activity of *Moraxella catarrhalis* Hag. *BMC microbiology*, *7*(1), 65.
18. Forsgren, A., Brant, M., & Riesbeck, K. (2004). Immunization with the Truncated Adhesin *Moraxella catarrhalis* Immunoglobulin D—Binding Protein (MID764–913) Is Protective against *M. catarrhalis* in a Mouse Model of Pulmonary Clearance. *Journal of Infectious Diseases*, *190*(2), 352-355.
19. Van Den Berg, S., Löfdahl, P. Å., Härd, T., & Berglund, H. (2006). Improved solubility of TEV protease by directed evolution. *Journal of biotechnology*, *121*(3), 291-298.
20. de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., ... & Mueller-Dieckmann, C. (2012). ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *Journal of synchrotron radiation*, *19*(3), 455-461.

21. Kabsch, W. (2010). Xds. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 125-132.
22. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., ... & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), 235-242.
23. Labrador, A., Cerenius, Y., Svensson, C., Theodor, K., & Plivelic, T. (2013, March). The yellow mini-hutch for SAXS experiments at MAX IV Laboratory. In *Journal of Physics: Conference Series* (Vol. 425, No. 7, p. 072019). IOP Publishing.
24. Knaapila, M., Svensson, C., Barauskas, J., Zackrisson, M., Nielsen, S. S., Toft, K. N., ... & Cerenius, Y. (2009). A new small-angle X-ray scattering set-up on the crystallography beamline I711 at MAX-lab. *Journal of synchrotron radiation*, 16(4), 498-504.
25. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExpASY server. In *The proteomics protocols handbook* (pp. 571-607). Humana Press.
26. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, 36(5), 1277-1282.
27. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, 25(4), 495-503.
28. Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal*, 76(6), 2879-2886.
29. Volkov, V. V., & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography*, 36(3), 860-864.
30. Lees, J. G., Smith, B. R., Wien, F., Miles, A. J., & Wallace, B. A. (2004). CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Analytical biochemistry*, 332(2), 285-289.
31. Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, 18(1), 211-212.
32. Mirabello, C., & Pollastri, G. (2013). Porter, PaleAle 4.0: high-accuracy prediction of protein secondary structure and relative solvent accessibility. *Bioinformatics*, 29(16), 2056-2058.

33. Rost, B., Yachdav, G., & Liu, J. (2004). The predictprotein server. *Nucleic acids research*, 32(suppl 2), W321-W326.
34. Buchan, D. W., Minnici, F., Nugent, T. C., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic acids research*, 41(W1), W349-W357.
35. Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences: CABIOS*, 11(6), 681-684.
36. Montgomerie, S., Cruz, J. A., Shrivastava, S., Arndt, D., Berjanskii, M., & Wishart, D. S. (2008). PROTEUS2: a web server for comprehensive protein structure prediction and structure-based annotation. *Nucleic acids research*, 36(suppl 2), W202-W209.
37. Rost, B., Yachdav, G., & Liu, J. (2004). The predictprotein server. *Nucleic acids research*, 32(suppl 2), W321-W326.
38. Wu, S., & Zhang, Y. (2008). MUSTER: improving protein sequence profile-profile alignments by using multiple sources of structure information. *Proteins: Structure, Function, and Bioinformatics*, 72(2), 547-556.
39. Rambo, R. P., & Tainer, J. A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, 95(8), 559-571.
40. Yeo, H. J., Cotter, S. E., Laarmann, S., Juehne, T., St Geme, J. W., & Waksman, G. (2004). Structural basis for host recognition by the Haemophilus influenzae Hia autotransporter. *The EMBO journal*, 23(6), 1245-1256.

ACKNOWLEDGEMENTS

This study was financially supported by Research School in Pharmaceutical Science (FLÄK) and Royal Physiographic Society in Lund,

The authors would like to thank Dr. Tomás S. Plivelic at MAX IV for assistance with the SAXS measurements and Dr. Petri Kursula for kindly collecting data at ASTRID, Denmark and Maria Håkansson at Crystallization Facility, MAX IV laboratory.

Table 1: Crystal parameters and Data processing statistics

Space group	P321
a (Å)	56.5
b (Å)	56.5
c (Å)	241.9
α (°)	90
β (°)	90
γ (°)	120
Wavelength (Å)	1.2825
Maximum resolution (Å)	2.3
V_{Matthews} (Da/Å ³) dimer in the asymmetric unit	2.04
Solvent content (%)	39.93
Total observations	127719
Unique reflections	20968
Rmerge (%)	9.1 (70.2)
Completeness (%)	99.6 (98.4)
Multiplicity	6.1
Average (I)/(Σ I)	14.17 (2.51)

a The corresponding data for the highest resolution shells are given in parentheses.

$R_{\text{merge}} = \frac{\sum_h \sum_i |I_i(h) - I(h)|}{\sum_h \sum_i I_i(h)}$, where $I_i(h)$ is the *i*th measurement of reflection *h* and $I(h)$ is the *i*th measurement of reflection *h* and $I(h)$ is the weighted mean of all measurements of *h*.

Table 2: SAXS data collection and scattering derived parameters.

Data-collection parameters	
Beamline	I911-4, MAX IV
Wavelength (Å)	0.91
q range (Å ⁻¹)	0.01-0.325
Exposure time (min)	5
Concentration range (mg ml ⁻¹)	3-9
Temperature (K)	283.15
Structural parameters	
$I(0)$ (cm ⁻¹) [from P(r)]	0.2045
R_g (Å) [from P(r)]	46.9
D_{\max} (Å)	154
Porod volume estimate (Å ³) [†]	118495
Dry volume calculated from sequence (Å ³)	30324
Molecular-mass determination	
Partial specific volume (cm ³ g ⁻¹)	0.724
Contrast ($\Delta\rho \times 10^{10}$ cm ⁻²)	
Molecular mass M_r [from excluded volume [‡]]	75111
Calculated monomeric M_r from sequence	25060.4
Software employed	
Primary data reduction	<i>Bli711</i>
Data processing	<i>Primus</i>
<i>Ab initio</i> analysis	<i>DAMMIN</i>
Validation and averaging	<i>DAMAVER</i>
Rigid-body modeling	N/A
Computation of model intensities	N/A
Three-dimensional graphics representations	<i>PyMOL</i>

[†] Calculated using datporod with GNOM file as an input.

[‡]Molecular mass calculated from excluded volume is more accurate for an elongated particle.

FIGURES



Figure 1. MID⁹⁶²⁻¹²⁰⁰ crystals grown by sitting drop vapor diffusion method.

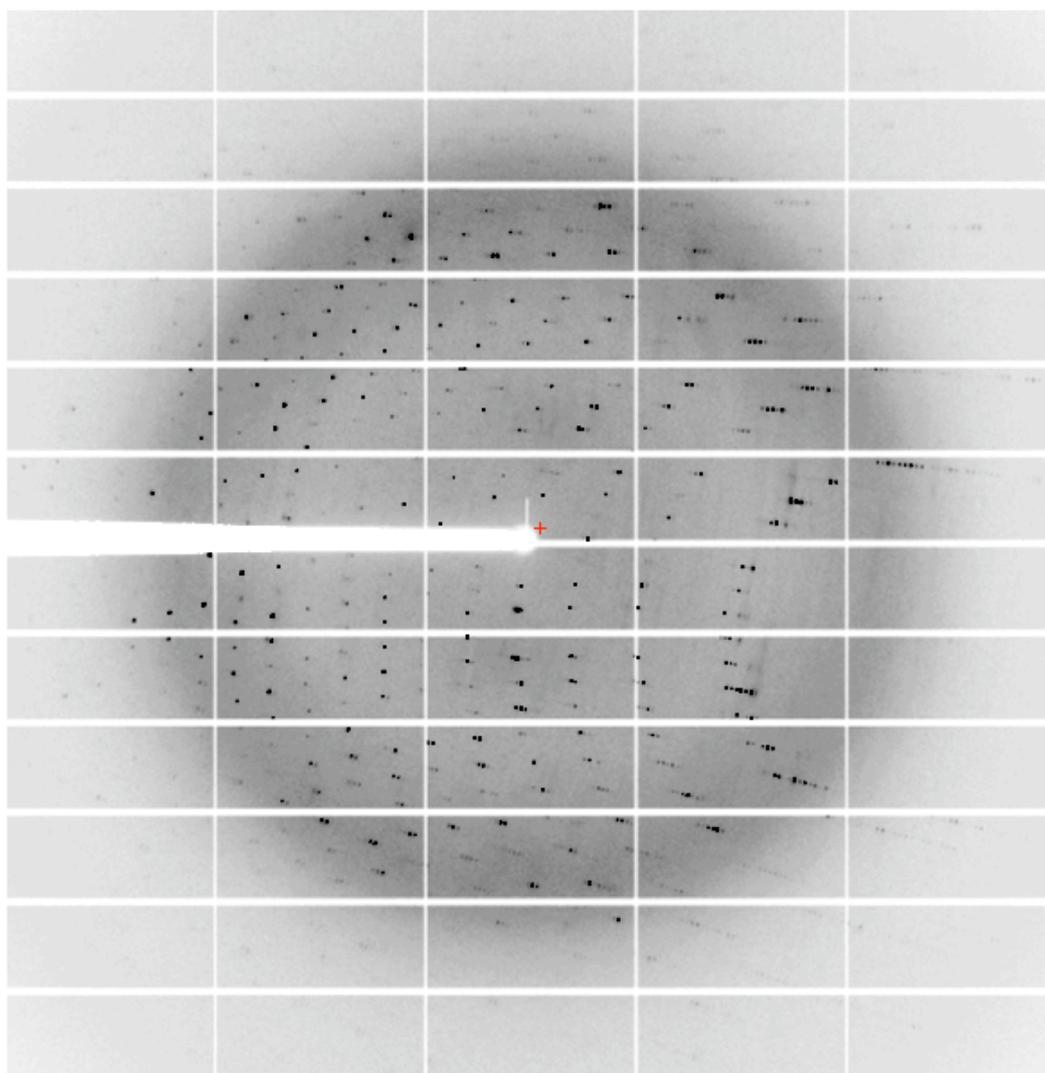
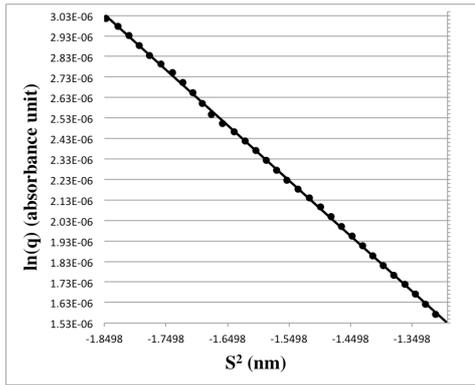
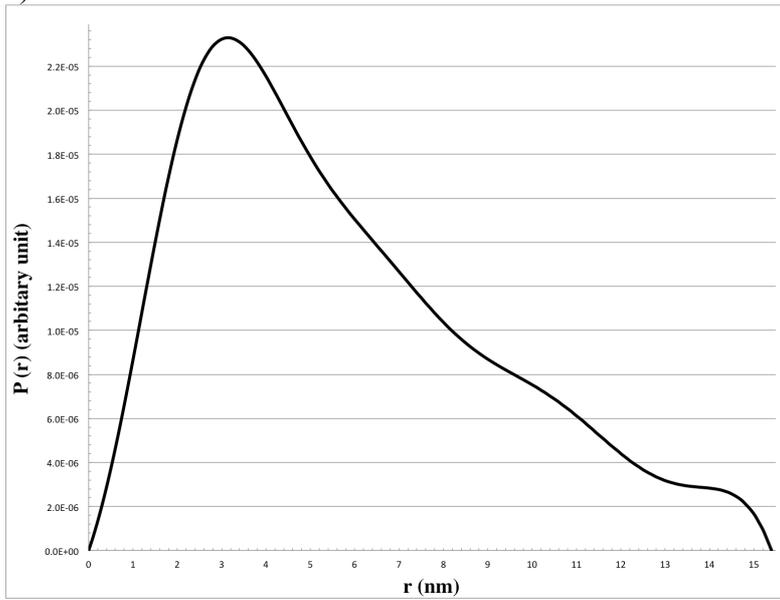


Figure 2: Diffraction image collected at ID29 at the ESRF, Grenoble. Wavelength = 1.2825 Å, sample to detector distance = 363.06 mm, rotation = 0.1°, exposure time = 0.045 sec per frame.



a)



b)

Figure 3: a) Guinier region of scattering data of MID⁹⁶²⁻¹²⁰⁰ and fit. b) Pair distribution function of MID⁹⁶²⁻¹²⁰⁰ calculated (D_{\max} 154 Å) using GNOM indicating elongated shape.

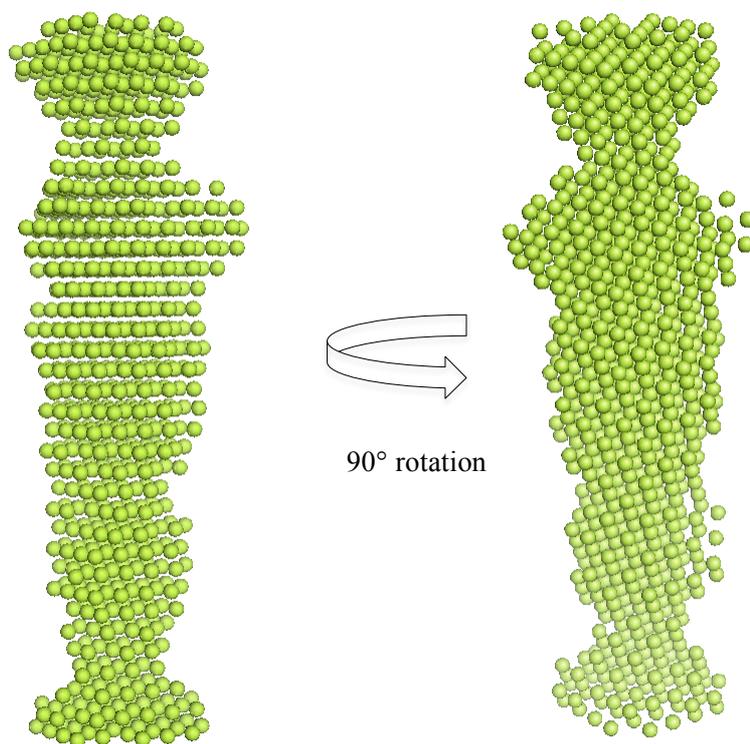


Figure 4: *Ab initio* model of MID⁹⁶²⁻¹²⁰⁰ from SAXS.

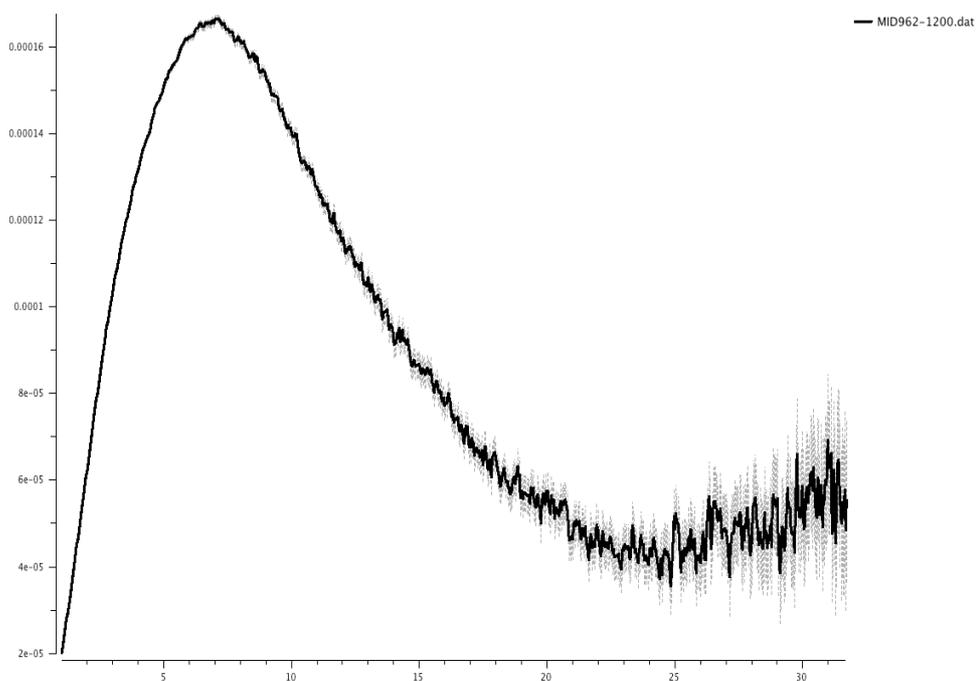


Figure 5: Kratky plot of MID⁹⁶²⁻¹²⁰⁰.

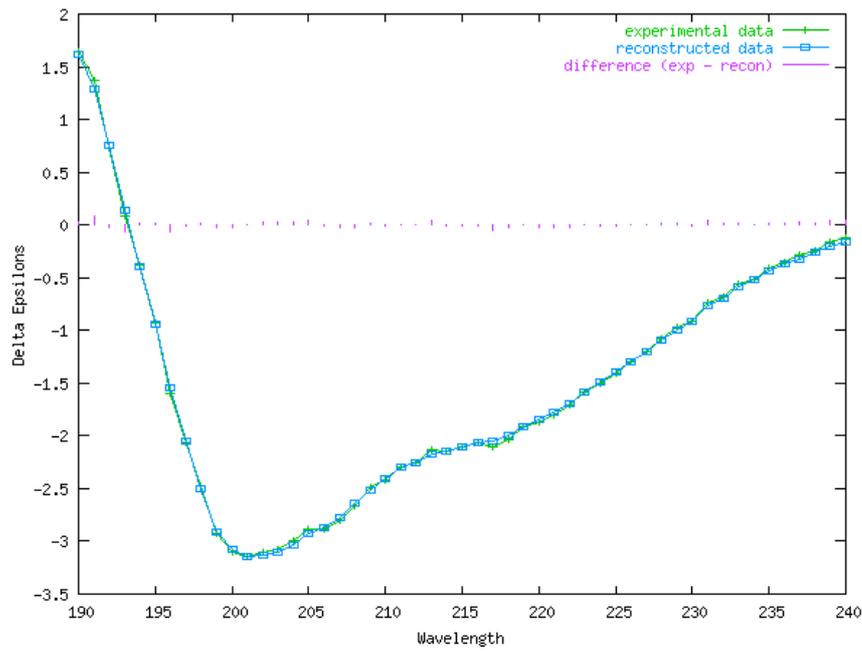


Figure 6: SRCD Spectra of $MID^{962-1200}$. The original spectra is colored in green and fit in blue. The discrepancy between spectra and fit is shown as bar colored in magenta.

Crystallization, circular dichroism and SAXS studies on the adhesin domain MID⁷⁶⁴⁻⁹¹³ from *Moraxella catarrhalis*

Mahmudul Hasan¹, Vivek Srinivas¹, Kristian Riesbeck² and Marjolein Thunnissen¹

¹Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Getingevägen 60, Lund University, 22100 Lund, Sweden.

²Medical Microbiology, Department of Laboratory Medicine Malmö, Lund University, Skåne University Hospital, 20502 Malmö, Sweden.

ABSTRACT

Moraxella catarrhalis is widely recognized human-restricted gram-negative bacterium for which it has become clear that it is a true pathogen of both the upper and lower respiratory tract. After *Haemophilus influenzae* and *Streptococcus pneumoniae* it is the third most common cause of otitis media in children. It is found to be second most common cause of chronic obstructive pulmonary disease after *H. influenzae*. The colonization to human respiratory tract is mediated by a 150-residue adhesin domain (MID⁷⁶⁴⁻⁹¹³), which is a part of the 2139 residue long protein. The protein belongs to trimeric autotransporter family and secreted via the type V secretion pathway. We present here results from preliminary crystallization, circular dichroism and small-angle X-ray scattering results of the adhesin domain from *M. catarrhalis*. The crystals belong to space group C2221 with cell dimensions $a = 42.8 \text{ \AA}$, $b = 51.6 \text{ \AA}$, $c = 303.2 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. A native dataset to 2.7 \AA was collected. Assuming that there is a dimer in the asymmetric unit the volume of Matthews (VM) is $2.16 \text{ \AA}^3/\text{Dalton}$ with a solvent content of 43.17 %. Circular dichroism measurements revealed that the protein contains a large β -sheet content (35%) and only 9% α -helix content which is also observed from secondary structure predictions. The CD spectra also indicate a certain degree of disorder (36%) that is also seen in Kratky plot as measured by SAXS.

Keywords: MID, trimeric autotransporter, adhesin, crystallization, circular dichroism

1. INTRODUCTION

Moraxella catarrhalis is a human-restricted gram-negative bacterium which has been widely recognized as a commensal bacterium. However during the last decades it has become apparent that it is a true pathogen of both the upper and lower respiratory tract. After *Haemophilus influenzae* and *Streptococcus pneumoniae* it is the third most common cause of otitis media in children. It is found to be second most common cause of chronic obstructive pulmonary disease after *H. influenzae*. [1].

One of the key virulence factors of *M. catarrhalis* is a 200-kDa protein named as *Moraxella* IgD binding protein (MID), which belongs to the trimeric autotransporter family [2]. The bacterium is capable to colonize on the lung epithelial cells via a 150-residue long (MID⁷⁶⁴⁻⁹¹³) domain, which is part of the 2139 residue long protein [3]. This domain is capable of binding to human alveolar epithelial cells and causes agglutination of human erythrocytes [4]. It was shown that antibodies generated against the MID⁷⁶⁴⁻⁹¹³ portion of the total protein could effectively inhibit adherence to the human alveolar epithelial cells [5], both in in vitro as well as in mouse model studies [6]. Cell isolates from *M. catarrhalis* expressing MID at a high concentration bound more effectively to epithelial cells [7]. The increase of antibiotic resistance of *Moraxella catarrhalis* [8] has generated much interest in the identification of potential vaccine as the bacterium contains β -lactamases and is resistant against ampicillin [9].

Like many other trimeric autotransporters [10,11] MID forms an extended lollipop type structure and becomes anchored on the outer membrane [12] by a beta-sheet porin like anchoring domain. Adhesin domains often sit on an extended like structure and this helps them to avoid recognition by the host [13]. It has been found that most of the adhesin proteins form trimers and this organization is required for normal folding, facilitates stability and helps maintaining adhesive activity [14].

In order to get insight about how MID⁷⁶⁴⁻⁹¹³ make interactions with human cells, structural information on this adhesin domain is important. Hence crystallization, circular dichroism (CD) and small-angle x-ray scattering (SAXS) studies were started. Secondary structure prediction databases were used to assess the structural composition

of this domain since the sequence of MID⁷⁶⁴⁻⁹¹³ is unique and no close homologues can be found in databases for which structural information is available.

2. MATERIALS AND METHODS

2.1. Cloning, Expression and purification of MID⁷⁶⁴⁻⁹¹³

The His₆-tagged protein was cloned into plasmid pET-26b(+) as mentioned previously [1]. The protein was expressed in BL21 (DE3) pLysS cells by growing 1-liter culture in a 3 liter baffled flask at 37 °C to OD 0.5, lowering the temperature to 20 °C and inducing expression with 1 mM IPTG at OD 0.7–0.8, followed by overnight incubation at 20 °C. The cells were then harvested by centrifuging at 6000 rpm and the pellets were frozen.

A frozen cell pellet from 1 liter of expression was resuspended in 100 ml of 100 mM Tris–HCl (pH 8.0), 300 mM NaCl, 10 mM MgCl₂, 10 mM NaN₃ containing complete mini EDTA-free protease inhibitor (Roche). The cells were lysed by sonication for 4x5 min with each time 3 sec pulse and 3 sec interval. The insoluble cell debris was removed by centrifugation at 30,000g for 30 min, and the lysate filtered through a 0.45- μ m syringe filter. The cleared lysate was then loaded on 1 ml HisTrap column (GE Healthcare) that was pre-equilibrated with binding buffer containing 50 mM Tris–HCl (pH 8.0), 200 mM NaCl, 10 mM MgCl₂ and 20 mM Imidazole and connected to a Äkta purifier (Amersham Biosciences). The column was washed with 5 volumes of binding buffer and the bound His-tagged protein was eluted with a gradient of 100% reaching in 20 minutes having elution buffer composed of 50 mM Tris–HCl (pH 8.0), 200 mM NaCl, 10 mM MgCl₂ and 250 mM Imidazole.

MID⁷⁶⁴⁻⁹¹³ was further purified by injecting in a Superdex 200 10/300 (GE Healthcare) size exclusion chromatography column pre-equilibrated with filtered and degassed 30 mM Tris–HCl (pH 8.0) and 200 mM NaCl and 20 mM MgCl₂ with a flow rate of 0.3 ml/min. A higher oligomeric form of the protein eluted at ~11.5 ml and the trimeric form of the protein eluted at ~15.5 ml. Corresponding higher oligomeric fractions were pooled and concentrated using Amicon Ultra Centrifugal Filter (Millipore) with cut off value of 3 kDa.

The monodispersity of the protein samples was measured by dynamic light scattering using a Zetasizer Nano-ZS (Malvern) instrument and data were analyzed using their software. Protein concentrations were determined by A_{280} measurements using a NanoDrop spectrophotometer (Thermo Scientific) and molar extinction coefficient ($4470 \text{ M}^{-1}\text{cm}^{-1}$) calculated in ProtParam [6].

2.2. Crystallization

An initial crystal hit was obtained from the HR2-110 screen (Hampton Research) using the sitting drop vapor diffusion method with a protein to reservoir volume ratio of 200:200 nl using a Mosquito nanolitre pipetting robot (TTP Labtech). These crystals were small and needed to be optimized. In experiments with a reservoir solution, containing 200 mM LiSO_4 , 100 mM Tris-HCl (pH 8.5) and PEG 4000 (20%) and additive screen (Hampton Research) and a protein to precipitant ratio 4:1, better quality crystals were obtained in presence of 20 mM MgCl_2 . In order to improve further, crystals were harvested and crushed using seed beads (Hampton Research) and serially diluted microseed solutions [2] were prepared. Optimization plates were set up with 280 nl MID⁷⁶⁴⁻⁹¹³ (8-10 mg/ml), 90 nl reservoir and 30 nl microseed solution using the Mosquito pipetting robot. Three-dimensional crystals were obtained after 1 to 2 weeks with conditions containing 220-320 mM LiSO_4 , 100 mM Tris-HCl (pH 7.5-8.5) and PEG 4000 (19-22%). All crystallization plates were incubated at 20° C and imaged automatically using a Minstrel HT (Rigaku Corporation).

2.3 Diffraction data collection and processing

Since the crystallization conditions themselves provide cryo-protection, crystals could be cooled without ice formation, straight from the crystallization drops. Data were collected at beam station I911-3 at MAX IV lab, Lund, Sweden, using a MD2 diffractometer with a Mar225 CCD Detector from a single MID⁷⁶⁴⁻⁹¹³ crystal [15]. The wavelength used for data collection was 0.9718 Å. The data were collected to a resolution of 2.7 Å. In total 800 frames of each with oscillation range of 0.25° of data were collected. The data were indexed, integrated and scaled using the program XDS [16] while further data manipulations were carried out using programs from the CCP4

package [17]. Further details on crystal parameters and data quality can be found in table 1.

2.4. Circular Dichroism data collection and processing

CD spectra of MID⁷⁶⁴⁻⁹¹³ at a concentration of 0.1 mg/ml and buffer were recorded using a Jasco J-815 (Jasco) and were averaged using manufacturer's software. MID⁷⁶⁴⁻⁹¹³ was purified using 25 mM sodium phosphate (pH 8.0) and 100 mM NaF as running buffer using size exclusion chromatography. All sample and baseline spectra were collected in quartz cells (path length 1 cm) over the wavelengths 260 to 190 nm with decreasing 1 nm in every step while keeping the temperature at 20° C. Total 8 accumulations were collected and averaged which was further subtracted with buffer. The secondary structure content of MID⁷⁶⁴⁻⁹¹³ was analyzed and calculated at the DichroWeb server [18].

2.5. SAXS data collection and processing

For SAXS analysis, MID⁷⁶⁴⁻⁹¹³ was a second time purified by repeating the above-mentioned protocol and data were collected with concentration between 2-10 mg/ml. Data were collected at BioSAXS beamline BM29 (ESRF, Grenoble, France) with X-ray beam at wavelength $\lambda = 1.0 \text{ \AA}$, and the distance from the sample to detector (Pilatus 1 M, Dectris Ltd.) was 2.85 m, covering a scattering vector range ($q = 4\pi\sin\theta/\lambda$) from 0.004 to 0.45 \AA^{-1} [19].

All data were collected at 283 K in a sample cell of 1.8 mm diameter quartz capillary and 10 frames of two-dimensional images were recorded for each buffer or sample, with an exposure time of 2 second per frame. The 2D images were averaged and reduced to one-dimensional scattering profiles and the scattering of the buffer was subtracted from the sample profile using the software on site. Initial data scaling along with the experimental radius of gyration (Rg) and the forward scattering intensity I(0) were calculated from data at low q values in the range of $qRg < 1.3$, using the Guinier approximation: $\ln(I(q)) \approx \ln(I(0)) - Rg^2 q^2/3$ using the software Primus [20].

2.6. Secondary Structure predictions

The sequence of MID⁷⁶⁴⁻⁹¹³ was used as a query in a number of secondary structure programs, Porter [21], PredictProtein/PROF Proteus-2 [22,23], PsiPRED [24], SOPma [25] and SSPro [26] to be able to predict the presence of secondary structure elements in it. The servers LoMETS [27] and MUSTER [28] were used to thread the sequence of the domain against the library of known folds. The metaserver metaPrDOS [29] was used to predict whether the protein contained disordered regions.

3. RESULTS AND DISCUSSION

Preliminary crystallization experiments were started in order to solve the structure of the adhesin domain MID⁷⁶⁴⁻⁹¹³. In the first attempts commercially available screens were used to find crystallization hits. Primary hits were obtained from conditions # 9, 17 and 40 of the HR2-110 screen (Hampton Research) and conditions #12, 49 and 50 of the JCSG+ screen (Molecular Dimension). Optimization of condition #17 of the HR2-110 gave an improvement of crystal-size and single crystals with dimensions of 0.12 mm x 0.06 mm x 0.02 mm were obtained. A native dataset using a single crystal was collected at the I911-3 station at the MAX-II ring of the MAX IV Laboratory. All X-ray experiments were performed using a liquid-nitrogen stream (100 K), with each crystal being vitrified in the cryostream with no added cryoprotectant. The crystallization conditions included PEG 4000, which appears to be adequate to protect the crystals from damage arising from cooling. The crystal belonged to space group C222₁ with cell-dimensions $a = 42.8 \text{ \AA}$, $b = 51.6 \text{ \AA}$, $c = 303.2 \text{ \AA}$, $\alpha, \beta \text{ \& } \gamma = 90^\circ$. The dataset was 95.6% complete with an Rmerge = 7.5%. Further details on the crystal parameters and data processing statistics can be found in table 1.

From the cell dimensions and the size of the protein (172 amino acids) it seems that there are two monomers in the asymmetric unit. This gives rise to a VM = 2.16 $\text{\AA}^3/\text{dalton}$ and a solvent content of 43.17% (calculated using a molecular weight of 18043 dalton for the domain). These values are at the lower end of the normal observed range (between 1.8 and 4.2 $\text{\AA}^3/\text{dalton}$ for the VM and between 27 and 78% for the average solvent content in protein crystals) [Matthews, 1968]. More than two copies of the monomer in the asymmetric unit are thus highly unlikely.

As indicated by the prediction program [4], deconvolution of the CD spectra indicated that a large fraction of the protein possesses beta structure (figure 3). A typical solution with a best NRMSD (normalized root-mean-square deviation) indicated 9% α -helix, 35% β -sheet, 19% turns and 36% unordered structures. The results do show that recombinant human MID⁷⁶⁴⁻⁹¹³ is folded, as expected, into a mainly β -sheet structure. The Kratky plot obtained from the SAXS data shows an indication of the presence of unfolded structure, which is consistent with unorderedness indicated by CD spectroscopy. This is in line with the secondary structure predictions for this domain. An array of programs was used (see figure 3) that all show that the protein has around 35 to 50 beta structure and only 8 to 10% helical content. Results from the disorder prediction metaserver metaprDOS, show that the protein is predominantly folded but that the N and C terminal ends might be disordered as well as a higher proportion for disorder for a region in the later 1/3 of the protein (see figure 6).

The protein mainly exists in a hexameric or trimeric form as seen from the size exclusion chromatography. From SAXS data (not shown) it was observed that the hexameric form had folding problems in solution and the interpretation of these scattering curves was not reliable for detailed analysis. Therefore, for the SAXS studies only data for the corresponding trimeric form was used and as here there was a presence of aggregation, the radius of gyration or further *ab initio* envelope model calculations were not possible and a more reliable method of purification has to be obtained in order to improve these measurements.

REFERENCES

1. de Vries, S. P., Bootsma, H. J., Hays, J. P., & Hermans, P. W. (2009). Molecular aspects of *Moraxella catarrhalis* pathogenesis. *Microbiology and molecular biology reviews*, 73(3), 389-406.
2. Bootsma, H. J., van Dijk, H. A. N. S., Verhoef, J., Fleer, A., & Mooi, F. R. (1996). Molecular characterization of the BRO beta-lactamase of *Moraxella* (*Branhamella*) *catarrhalis*. *Antimicrobial agents and chemotherapy*, 40(4), 966-972.
3. Forsgren, A., Brant, M., Karamehmedovic, M., & Riesbeck, K. (2003). The immunoglobulin D-binding protein MID Mazzulli, T. (2005). Resistant microorganisms in head and neck infections. *Journal of otolaryngology*, 34.
4. Bullard, B., S. Lipski, and E. R. Lafontaine. 2007. Regions important for the adhesin activity of *Moraxella catarrhalis* Hag. *BMC Microbiol.* 7:65.
5. Forsgren, A., Brant, M., & Riesbeck, K. (2004). Immunization with the Truncated Adhesin *Moraxella catarrhalis* Immunoglobulin D—Binding Protein (MID764–913) Is Protective against *M. catarrhalis* in a Mouse Model of Pulmonary Clearance. *Journal of Infectious Diseases*, 190(2), 352-355.
6. Forsgren, A., Brant, M., & Riesbeck, K. (2004). Immunization with the Truncated Adhesin *Moraxella catarrhalis* Immunoglobulin D—Binding Protein (MID764–913) Is Protective against *M. catarrhalis* in a Mouse Model of Pulmonary Clearance. *Journal of Infectious Diseases*, 190(2), 352-355.
7. Forsgren, A., Brant, M., Möllenkvist, A., Muyombwe, A., Janson, H., Woin, N., & Riesbeck, K. (2001). Isolation and characterization of a novel IgD-binding protein from *Moraxella catarrhalis*. *The Journal of Immunology*, 167(4), 2112-2120.
8. Mazzulli, T. (2005). Resistant microorganisms in head and neck infections. *Journal of otolaryngology*, 34.
9. Bootsma, H. J., Van Dijk, H., Vauterin, P., Verhoef, J., & Mooi, F. R. (2000). Genesis of BRO β -lactamase-producing *Moraxella catarrhalis*: evidence for transformation-mediated horizontal transfer. *Molecular microbiology*, 36(1), 93-104.
10. Cotter, S. E., Surana, N. K., & St Geme III, J. W. (2005). Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends in microbiology*, 13(5), 199-205.

11. Linke, D., Riess, T., Autenrieth, I. B., Lupas, A., & Kempf, V. A. (2006). Trimeric autotransporter adhesins: variable structure, common function. *Trends in microbiology*, *14*(6), 264-270.
12. Hallström, T., Müller, S. A., Mörgelin, M., Möllenkvist, A., Forsgren, A., & Riesbeck, K. (2008). The *Moraxella* IgD-binding protein MID/Hag is an oligomeric autotransporter. *Microbes and infection*, *10*(4), 374-381.
13. Kline, K. A., Fälker, S., Dahlberg, S., Normark, S., & Henriques-Normark, B. (2009). Bacterial adhesins in host-microbe interactions. *Cell host & microbe*, *5*(6), 580-592.
14. Cotter, S. E., Surana, N. K., Grass, S., & Geme, J. W. S. (2006). Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *Journal of bacteriology*, *188*(15), 5400-5407.
15. Ursby, T., Unge, J., Appio, R., Logan, D. T., Fredslund, F., Svensson, C., ... & Thunnissen, M. M. (2013). The macromolecular crystallography beamline I911-3 at the MAX IV laboratory. *Journal of synchrotron radiation*, *20*(4), 0-0.
16. Kabsch, W. (2010). Xds. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 125-132.
17. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., ... & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 235-242.
18. Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, *18*(1), 211-212.
19. Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., ... & McSweeney, S. (2013). Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of synchrotron radiation*, *20*(4), 0-0.
20. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, *36*(5), 1277-1282.
21. Mirabello, C., & Pollastri, G. (2013). Porter, PaleAle 4.0: high-accuracy prediction of protein secondary structure and relative solvent accessibility. *Bioinformatics*, *29*(16), 2056-2058.

22. Rost, B., Yachdav, G., & Liu, J. (2004). The predictprotein server. *Nucleic acids research*, 32(suppl 2), W321-W326.
23. Montgomerie, S., Cruz, J. A., Shrivastava, S., Arndt, D., Berjanskii, M., & Wishart, D. S. (2008). PROTEUS2: a web server for comprehensive protein structure prediction and structure-based annotation. *Nucleic acids research*, 36(suppl 2), W202-W209.
24. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic acids research*, 41(W1), W349-W357.
25. Geourjon, C., & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences: CABIOS*, 11(6), 681-684.
26. Cheng, J., Randall, A. Z., Sweredoski, M. J., & Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic acids research*, 33(suppl 2), W72-W76.
27. Wu, S., & Zhang, Y. (2007). LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic acids research*, 35(10), 3375-3382.
28. Wu, S., & Zhang, Y. (2008). MUSTER: improving protein sequence profile-profile alignments by using multiple sources of structure information. *Proteins: Structure, Function, and Bioinformatics*, 72(2), 547-556.
29. Ishida, T., & Kinoshita, K. (2008). Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics*, 24(11), 1344-1348.

Acknowledgement

Financial support: FLÄK, Royal Physiographic Society in Lund.

The authors would like to thank to people at BioSAXS beamline BM29 (ESRF), I911-3 and Maria Håkansson at Crystallization Facility, MAX IV laboratory.

Table 1: MID⁷⁶⁴⁻⁹¹³ crystal parameters and data processing statistics.

Space group	C2221
a (Å)	42.8
b (Å)	51.6
c (Å)	303.2
α (°)	90
β (°)	90
γ (°)	90
Wavelength (Å)	0.9718
Maximum resolution (Å)	2.7
V_{Matthews}(Da/Å³) 2 Monomers in the asymmetric unit	2.16
Solvent content (%)	43.17
Total observations	60284
Unique reflections	16963
Rmerge (%)	7.5 (42.3)
Completeness (%)	95.6 (78.8)
Multiplicity	3.6 (2.5)
Average (I)/(ΣI)	8.04 (1.33)

a The corresponding data for the highest resolution shells are given in parentheses.

$R_{\text{merge}} = \frac{\sum_h \sum_i |I_i(h) - I(h)|}{\sum_h \sum_i I_i(h)}$, where $I_i(h)$ is the *i*th measurement of reflection *h* and $I(h)$ is the *i*th measurement of reflection *h* and $I(h)$ is the weighted mean of all measurements of *h*.

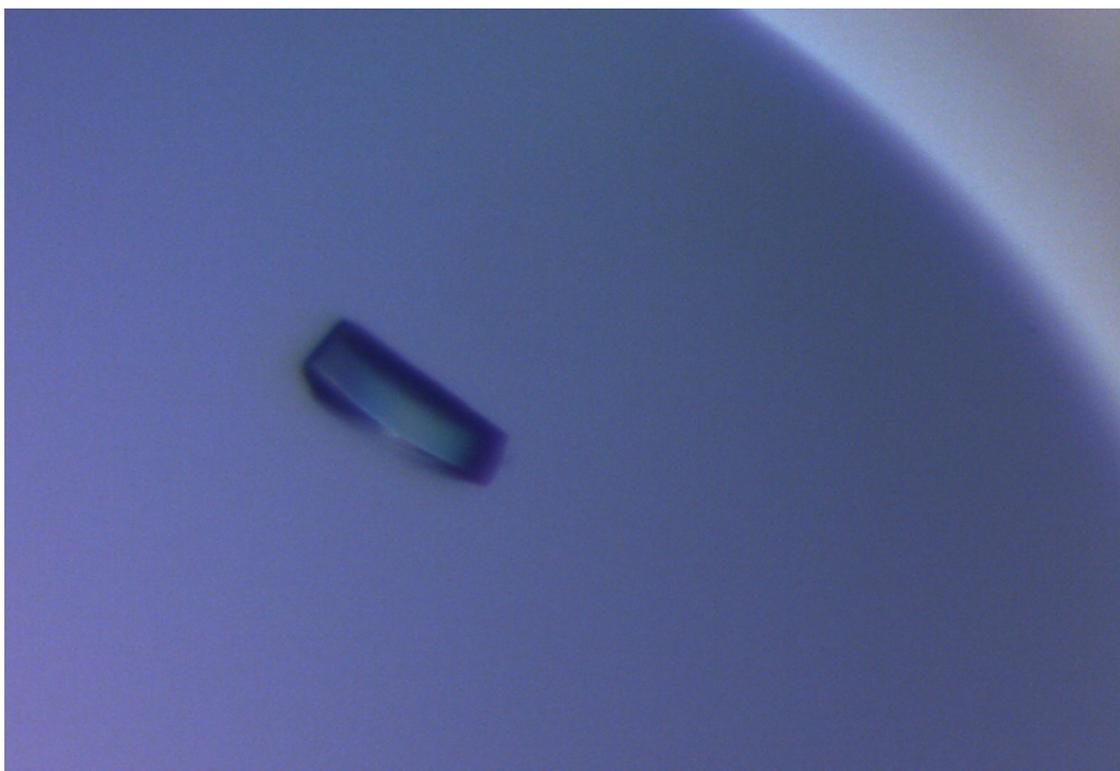


Figure 1: Image of crystal of MID⁷⁶⁴⁻⁹¹³ obtained from optimization screen by sitting drop vapor diffusion method.

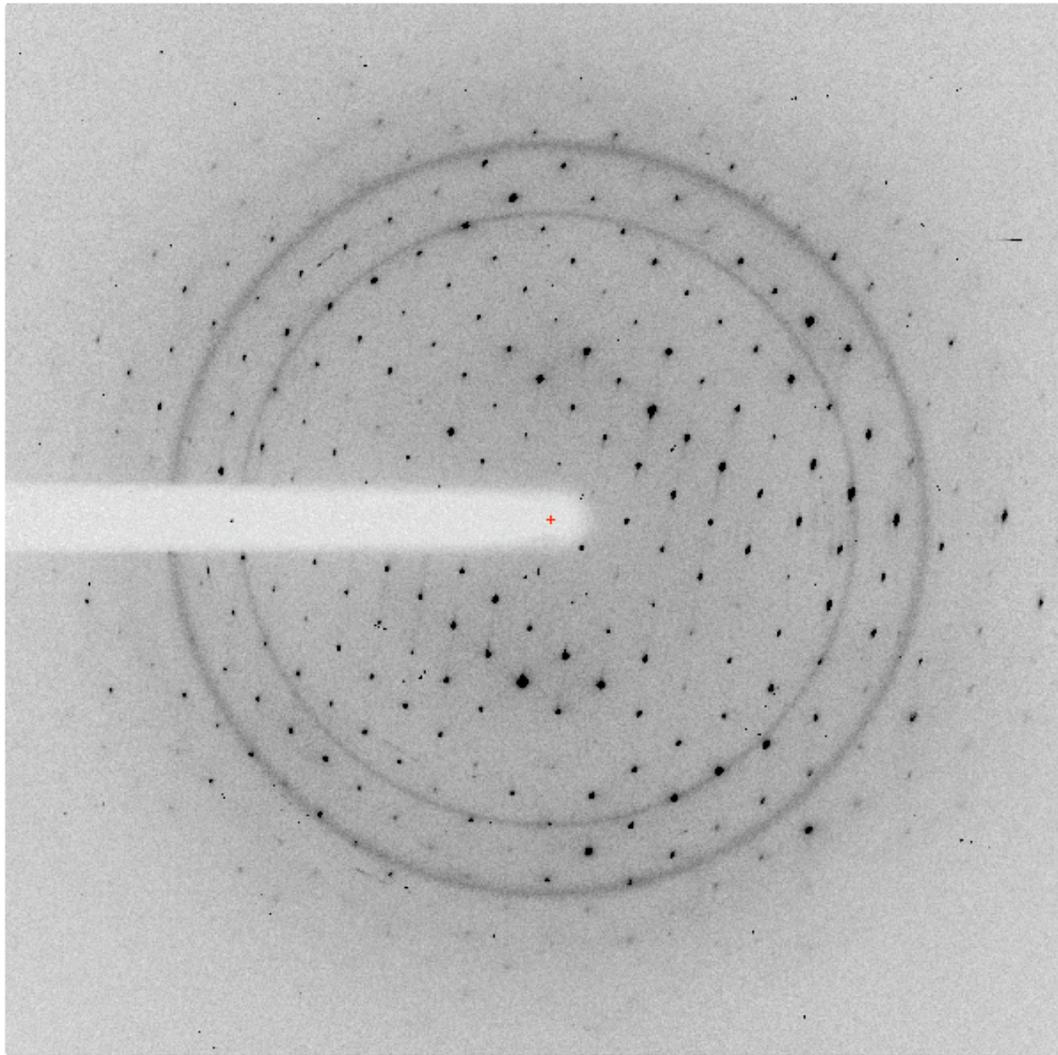


Figure 2: Diffraction image collected at I911-3 at the MAX IV laboratory, Lund, Sweden. Wavelength = 0.9718 Å, distance sample to detector = 301.67 mm, rotation = 0.25°, exposure time = 35 sec per frame. Data processing statistics are available at table 1.

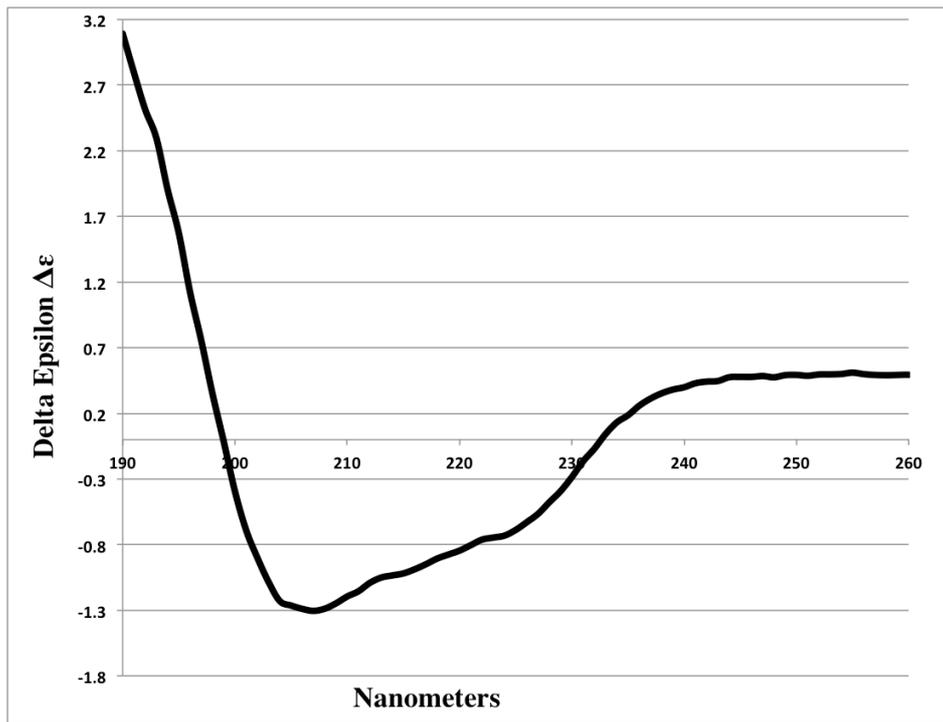


Figure 3: Normalized circular dichroism spectra of MID⁷⁶⁴⁻⁹¹³.

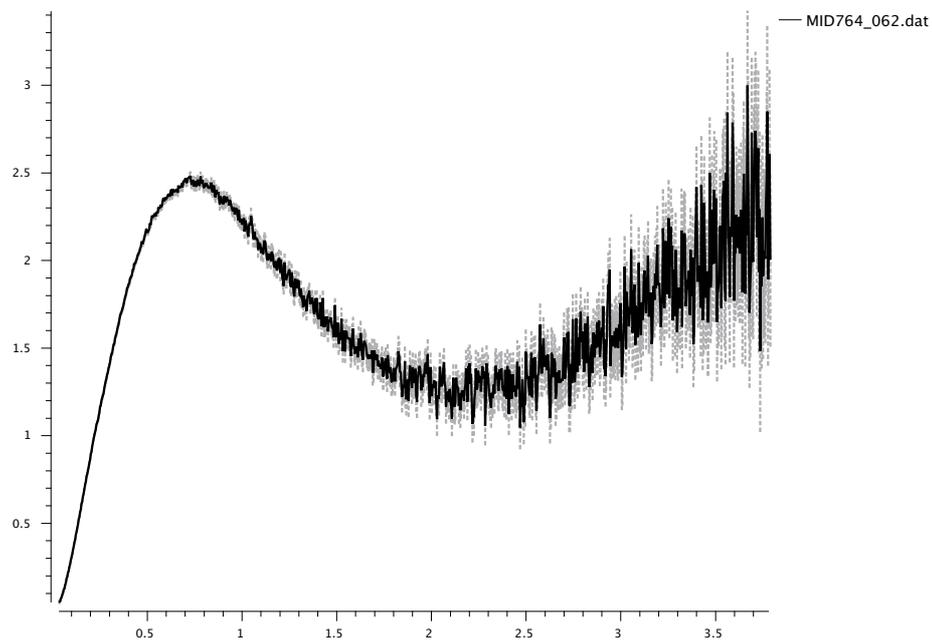


Figure 4: Kratky plot of MID⁷⁶⁴⁻⁹¹³ plotted using Primus. Failure to reach to plateau is indication of unordered region present in the structure.



A Leukotriene A₄ Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding

Charlotte Helgstrand¹, Mahmudul Hasan¹, Hüseyin Uysal¹,
Jesper Z. Haeggström² and Marjolein M. G. M. Thunnissen^{1*}

¹Centre of Molecular Protein Science, Lund University, Getingevägen 60, SE 22100 Lund, Sweden

²Department of Medical Biochemistry and Biophysics, Division of Chemistry 2, Karolinska Institute, SE 171-77 Stockholm, Sweden

Received 31 August 2010;
received in revised form
24 November 2010;
accepted 30 November 2010
Available online
10 December 2010

Edited by G. Schulz

Keywords:

leukotriene A₄ hydrolase;
aminopeptidase;
epoxide hydrolase;
catalysis;
crystallography

Vertebrate leukotriene A₄ hydrolases are bifunctional zinc metalloenzymes with an epoxide hydrolase and an aminopeptidase activity. In contrast, highly homologous enzymes from lower organisms only have the aminopeptidase activity. From sequence comparisons, it is not clear why this difference occurs. In order to obtain more information on the evolutionary relationship between these enzymes and their activities, the structure of a closely related leucine aminopeptidase from *Saccharomyces cerevisiae* that only shows a very low epoxide hydrolase activity was determined. To investigate the molecular architecture of the active site, the structures of both the native protein and the protein in complex with the aminopeptidase inhibitor bestatin were solved. These structures show a more spacious active site, and the protected cavity in which the labile substrate leukotriene A₄ is bound in the human enzyme is partially obstructed and in other parts is more solvent accessible. Furthermore, the enzyme undergoes induced fit upon binding of the inhibitor bestatin, leading to a movement of the C-terminal domain. The main triggers for the domain movement are a conformational change of Tyr312 and a subtle change in backbone conformation of the PYGAMEN fingerprint region for peptide substrate recognition. This leads to a change in the hydrogen-bonding network pulling the C-terminal domain into a different position. Inasmuch as bestatin is a structural analogue of a leucyl dipeptide and may be regarded as a transition state mimic, our results imply that the enzyme undergoes induced fit during substrate binding and turnover.

© 2010 Elsevier Ltd. All rights reserved.

*Corresponding author. E-mail address: marjolein.thunnissen@biochemistry.lu.se.

Present addresses: C. Helgstrand, Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark; H. Uysal, Department of Medical Biochemistry and Biophysics, Division of Medical Inflammation Research, Karolinska Institute, SE 171-77 Stockholm, Sweden.

Abbreviations used: scLTA₄H, leukotriene A₄ hydrolase from *Saccharomyces cerevisiae*; GST, glutathione S-transferase; PEG, polyethylene glycol; Tris, tris(hydroxymethyl)aminomethane; ES, 2-(N-morpholino)ethanesulfonic acid; MMT buffer, L-malic acid; MES, Tris buffer.

Introduction

The zinc metalloenzyme leukotriene A₄ (LTA₄) hydrolase (LTA4H) catalyses the hydrolysis of the unstable epoxide LTA₄ (5*S*-*trans*-5,6-oxido-7,9-*trans*-11,14-*cis* eicosatetraenoic acid) into LTB₄ (5*S*,12*R*-dihydroxy-6,14-*cis*-8,10-*trans*-eicosatetraenoic acid), a potent chemotactic agent and mediator of inflammation.^{1,2} The LTA4Hs found in vertebrates are bifunctional proteins, since they also possess an anion-dependent Arg-aminopeptidase activity.^{3–5} The zinc atom is essential for both activities, which take place in overlapping active sites. The human enzyme has been studied extensively, and the key elements for the catalytic activities have been identified by a combination of site-directed mutagenesis and crystallography techniques.⁶

LTA4H is part of a class of multifunctional proteins, also known as moonlighting proteins, which have more than one distinct function within a single polypeptide chain.⁷ These proteins may contain dual catalytic activities or combinations of noncatalytic functions with enzymatic activities. Moonlighting proteins are a diverse class of proteins that can be found in all phyla, and their functions address a wide range of cellular and physiological processes.⁸ To perform multiple functions, these proteins can utilize separate binding sites for different substrates or assume different roles upon complex formation with cofactors or other proteins. Products of gene fusions, splice variants, or proteins showing promiscuous enzymatic activity are not considered to be part of this class. It has been speculated that in order to obtain a moonlighting function, a protein needs to have an innate compatibility to obtain this function.⁹ According to this hypothesis, additional mutations are required to develop the full multifunctional protein. LTA4H is a somewhat unique moonlighting protein, since its two activities are exerted via distinct but yet overlapping active sites. Thus, within one active centre, certain amino acids are common for the two activities, whereas others are specifically used in only one of the two catalytic mechanisms.⁶ The function of the epoxide hydrolase activity of LTA4H

is well established and provides a potent chemotactic factor, LTB₄, during the initiation of inflammation. Recently, data were presented, suggesting that the aminopeptidase activity of LTA4H cleaves and inactivates a proinflammatory peptide, Pro-Gly-Pro, during the resolution of inflammation.¹⁰ Hence, the two activities of LTA4H seem to play specific and functionally opposite roles during two separate phases of an inflammatory reaction.

LTA4H is a member of the M1 family of metalloproteases,¹¹ characterized by a common Zn binding signature, HEXXH, in which the His residues function as Zn binding ligands and the Glu serves as the general base catalyst in the peptide hydrolysis. The zinc-binding site is completed by a Glu residue located 18 amino acids downstream of the HEXXH motif. All members of the family exhibit the aminopeptidase activity; however, the epoxide hydrolase activity is not conserved despite considerable sequence conservation. This activity can only be detected among vertebrates, including birds, frogs, and fish.^{12–16} Thus, although aminopeptidase 1 from *Caenorhabditis elegans* is highly homologous at the amino acid level to human LTA4H, no LTA₄ hydrolysis could be detected¹⁷ (see Table 1 for homology percentages). These observations raise questions on the evolutionary relationship of these enzymes, how the presence of the epoxide hydrolase activity was established in these proteins, and whether conclusions drawn for other moonlighting proteins play a role for LTA4H.

Apart from the structure of human LTA4H,²¹ structures of other members of the family have recently become available. The most homologous is the structure of a cold-active aminopeptidase from *Colwellia psychrethraea*,²⁰ while more distant structures are represented by those of the Tricorn interacting factor F3¹⁹ and the aminopeptidases N from *Escherichia coli*,^{18,22} *Neisseria meningitidis*,²³ and *Plasmodium falciparum*.²⁴ All these enzymes are built from several domains coming together in a hook-like formation with the active site positioned in a deep cavity between the domains. They also exhibit sequential and structural conservation for the two N-terminal domains, while there is considerable

Table 1. Homology of selected related proteins to human LTA4H

Enzyme	Species	Identity (%)	Similarity (%)	Residue overlap
LTA4 hydrolase ¹⁸	<i>S. cerevisiae</i>	40.9	68.8	616
Aminopeptidase 1 ¹³	<i>C. elegans</i>	36	63.5	636
Cold-active aminopeptidase ¹⁵	<i>C. psychrethraea</i>	33.9	62.3	623
Aminopeptidase N ¹⁹	<i>P. falciparum</i>	26.5	53.5	465
Aminopeptidase N ²⁰	<i>N. meningitidis</i>	24.8	50.0	468
Aminopeptidase N ^{17,21}	<i>E. coli</i>	22.9	54.3	494
Tricorn interacting factor F3 ¹⁶	<i>T. acidophilum</i>	22.0	54.5	451

Homology comparisons were calculated by using FASTA at the EBI with BLOSUM50 as similarity matrix. Sequence for *hum*LTA4H from Haeggström *et al.*, 1990.³

variation for the C-terminal domains. Human LTA₄H and the cold active protein have three domains, whereas the other structurally characterized proteins have an additional middle domain. In addition, the all-helical C-terminal domain is larger in nonvertebrate enzymes and only partially overlaps with that of the human enzyme. Several highly conserved sequential elements (such as the PYGA-MEN sequence) are positioned in the walls of the cavity, and these have been indicated as important for the aminopeptidase activity in the human enzyme as well as for other M1-family zinc metalloproteases.²⁵ The binding site of the peptidyl substrates is very similar in all the proteins, while the additional binding site for the lipid LTA₄ in the human enzyme is formed by an elongation of the central cavity by a narrow L-shaped pocket lined with hydrophobic residues. This pocket can also be seen in the cold-adapted enzyme from *C. psychrethraea*.²⁰ However, because of rearrangements of loops that form the walls of the lipid binding site and the presence of large side chains, the size of the binding site is smaller in this enzyme as compared to human LTA₄H. It was postulated that this was the probable reason for why the cold-adapted enzyme did not exhibit an LTA₄ hydrolase activity in spite of the fact that all other elements required for hydrolysis of the epoxide are present.

A leucine aminopeptidase from *Saccharomyces cerevisiae*, a close homologue of human LTA₄H (for simplicity abbreviated as scLTA₄H), has been used in previous investigations to study the evolutionary relationships between the LTA₄-converting enzymes. This protein is 40.9% identical (68.8% similar) to human LTA₄H and has a low and unspecific epoxide hydrolase activity^{26,27} that converts LTA₄ into the main product 5S,6S-DHETE (5S,6S-dihydroxy-7,9-*trans*-11,14-*cis*-eicosatetraenoic acid), as well as small amounts of LTB₄ and Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ (5S,12R-dihydroxy-6,10-*trans*-8,14-*cis*-eicosatetraenoic acid). In addition, the enzyme has a different aminopeptidase specificity compared to the human enzyme, since it prefers Leu as a side chain at the P1 position instead of Arg. While the presence of LTA₄ inhibits the aminopeptidase activity of the human enzyme, for scLTA₄H it acts as a strong activator for the aminopeptidase activity, in a manner suggesting the presence of an LTA₄ binding site.²⁷

From modelling studies, sequence alignment, and mutagenesis experiments, it was postulated that scLTA₄H would have a more spacious binding site for LTA₄, and this would impair the precise formation of the product LTB₄.²⁸ Within these studies, some catalytically important residues for the epoxide hydrolase reaction (Glu316 and Arg627) could be identified, but Asp425 (Asp375 in human LTA₄H) seemed to be located too far away from the lipid binding site to make a direct contact with the substrate. In mutagenesis studies, this residue could

indeed be mutated without loss of activity. Mutagenesis efforts to convert scLTA₄H into a protein with properties more like the human enzyme were only partially successful. It was thus possible, by changing 4 residues, to alter the protein from a Leu into an Arg aminopeptidase. However, the epoxide hydrolase activity was only weakly enhanced with no gain in substrate specificity. Therefore, it seems that the introduction of the epoxide hydrolase activity within the framework of the M1 class of aminopeptidases is the result of an overall reshaping of the binding pocket by subtle alterations in the overall structure of the protein.

To address the question on what impact evolutionary reshaping of the lipid substrate binding site has on catalytic activity, we have solved the structure of scLTA₄H both on its own and in complex with bestatin, an inhibitor of the M1 family of metalloproteases.

Results and Discussion

Overall native scLTA₄H structure

Although scLTA₄H was crystallised several years ago,²⁹ its structure determination remained elusive. The crystals were pseudo-merohedrally twinned and a reliable molecular replacement solution could not be obtained. The crystals were also fragile, and heavy-atom derivative searches destroyed the reflection patterns. In order to overcome these difficulties, it was decided to alter the purification

Table 2. Data processing and refinement statistics

	P3 ₂ 2 ₁ form	P2 ₁ 2 ₁ 2 ₁ form	Bestatin complex
Space group	P3 ₂ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
<i>a</i> (Å)	159.6	70.8	58.6
<i>b</i> (Å)	159.6	98.1	99.9
<i>c</i> (Å)	76.8	99.2	112.7
Maximum resolution (Å)	2.8	2.4	1.95
Total observations	112,679	59,123	192,141
Unique reflections	26,809	26,435	48,055
<i>R</i> _{merge} (%) ^a	8.3 (35.2)	8.5 (24.3)	5.8 (42.1)
Completeness (%) ^a	98.9 (95.4)	96.0 (93.7)	98.9 (97.6)
Multiplicity ^a	4.1 (3.6)	2.2 (2.1)	4.0 (3.9)
Average <i>I</i> /(ΣI) ^a	7.6 (2.0)	5.7 (2.1)	15.31 (3.25)
<i>Refinement</i>			
Refinement range (Å)	26–2.8	15–2.4	28.7–1.95
<i>R</i> _{crys} (%)	17.2	19.2	19.0
<i>R</i> _{free} (%)	23.5	26.5	24.7
RMSD bond length (Å)	0.008	0.007	0.001
RMSD bond angles (°)	1.130	0.987	1.112
<i>Ramachandran plot statistics</i>			
Most favored regions (%)	94.9	95.8	96.9
Disallowed regions (%)	0	0	0

^aThe corresponding data for the highest-resolution shell are given in parentheses.

scheme of the protein and recrystallise the protein into another crystal form. This approach was successful; by recloning the protein, expressing it as a GST fusion protein, and using a different purification approach, a new crystal form was obtained. Molecular replacement using the structure of the human enzyme gave a clear solution, and subsequent rebuilding and refining yielded a structure with good statistics (see Table 2). Unambiguous density can be seen for the full-length protein, apart from a surface loop (226–230). During the refinement of the crystal form from the GST-affinity purified protein, a glutathione molecule was identified. It is located in the crystallographic interface between symmetry-related protein molecules, and it is covalently bound to Cys147. One of its carboxyl groups interacts with His97, while the other carboxyl group makes interactions with Lys601 and Lys604 from a crystallographic neighbouring molecule. Using the *P*₃₂₁ structure as a model for molecular replacement using PHASER,³⁰ it was possible to obtain a clear solution for the previous twinned *P*₂₁₂₁ crystal form using data from a crystal that had a twin fraction of only 0.07. The two structures superpose with each other with an RMSD

of 0.490 Å for 4412 atom pairs (out of 5092), and only small conformational differences are seen in the loops 144–149 and 96–103. Both loops are involved in glutathione binding and also make several crystal contacts in the *P*₃₂₁ form. Loop 96 to 103 is flexible in the *P*₂₁₂₁ form. Since data for the *P*₂₁₂₁ form extend to 2.4 Å instead of 2.8 Å in the case of the *P*₃₂₁ form, and only very minor changes can be seen between the two crystal forms, all following results and discussion are based on the *P*₂₁₂₁ form.

ScLTA4 is built from three different domains (an all-β N-terminal domain, residues 40–252; a mixed α-β catalytic domain, residues 253–506; and an all helical C-terminal domain, residues 507–671) (Fig. 1). In comparison with the human enzyme, scLTA4H has an extension at the N terminus, some small deletions (3–4 residues) in the N-terminal domain, and two larger insertions (6 and 12 residues, respectively) located in loop areas between the helices in the C-terminal domain.

The catalytic domain is the most conserved (56% identity, 76% similarity at the amino acid level) and shows only minor structural changes with an RMSD of 0.80 Å for 210 C^α positions. In contrast, the C-terminal domain is only 26.1% identical (46.0%

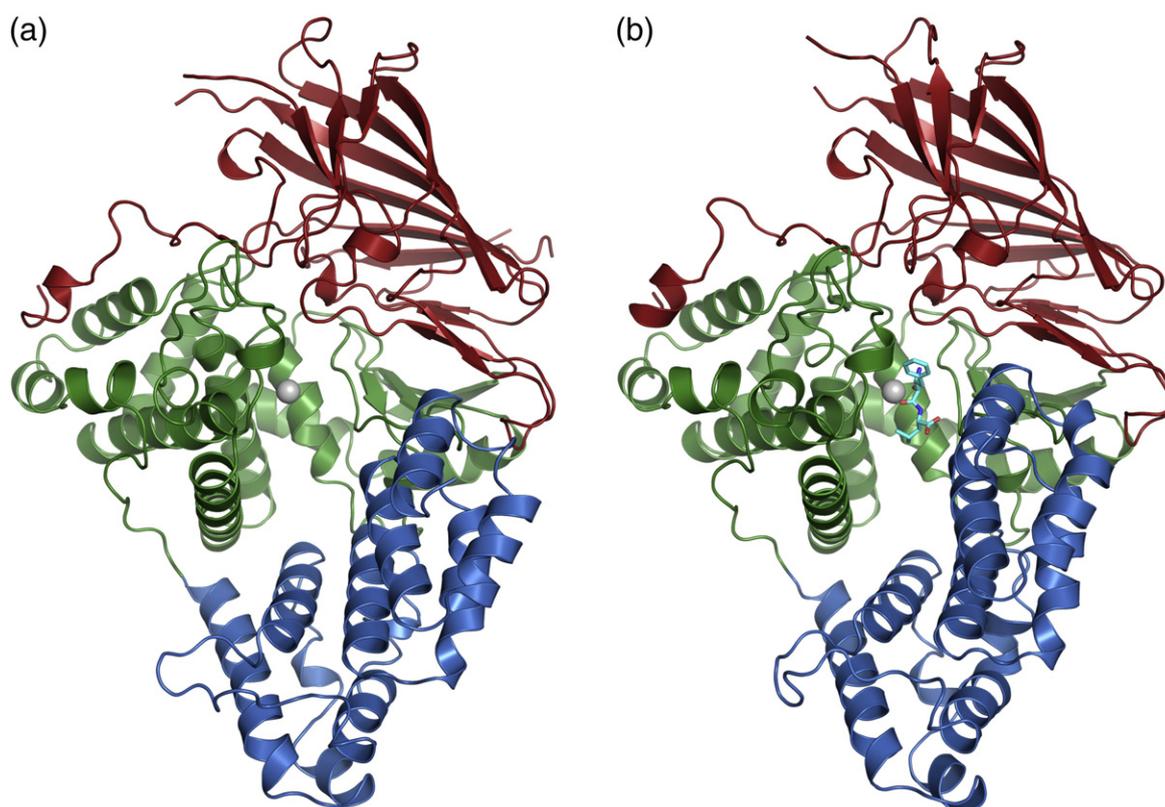


Fig. 1. Cartoon presentation of scLTA4H. (a) Structure of the enzyme alone; (b) structure of the enzyme in complex with the bestatin inhibitor. All figures are produced with the program Pymol³¹ if not indicated differently. Both cartoons are coloured as follows: red, N-terminal domain; green, catalytic domain; and blue, C-terminal domain. The bestatin molecule is drawn as a stick model in aqua, while the active-site zinc molecules are drawn as grey spheres.

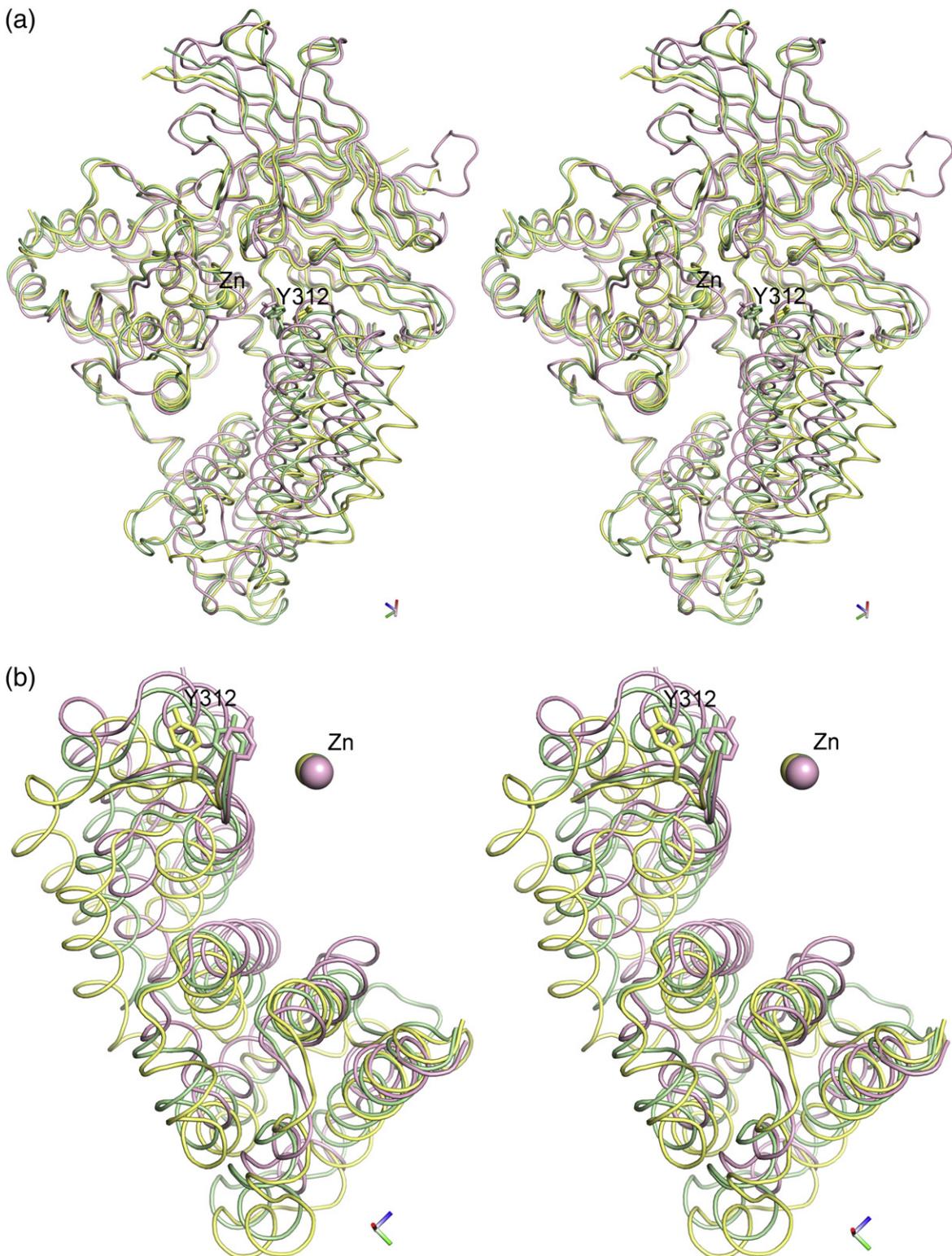


Fig. 2. Stereo figures of the conformational change of the different domains in scLTA4H compared to the human LTA4H. The proteins are drawn as ribbons. (a) Overview of the whole protein. (b) Close-up of the C-terminal domain (front view). (c) Close-up of the C-terminal domain (top view). In all three sets, yellow indicates scLTA4H, green represents scLTA4H in complex with bestatin, and pink represents humLTA4H in complex with bestatin. Residue Tyr312 is shown as a stick model and the zinc ion as a sphere in panels (b) and (c). Numbering is according to scLTA4H.

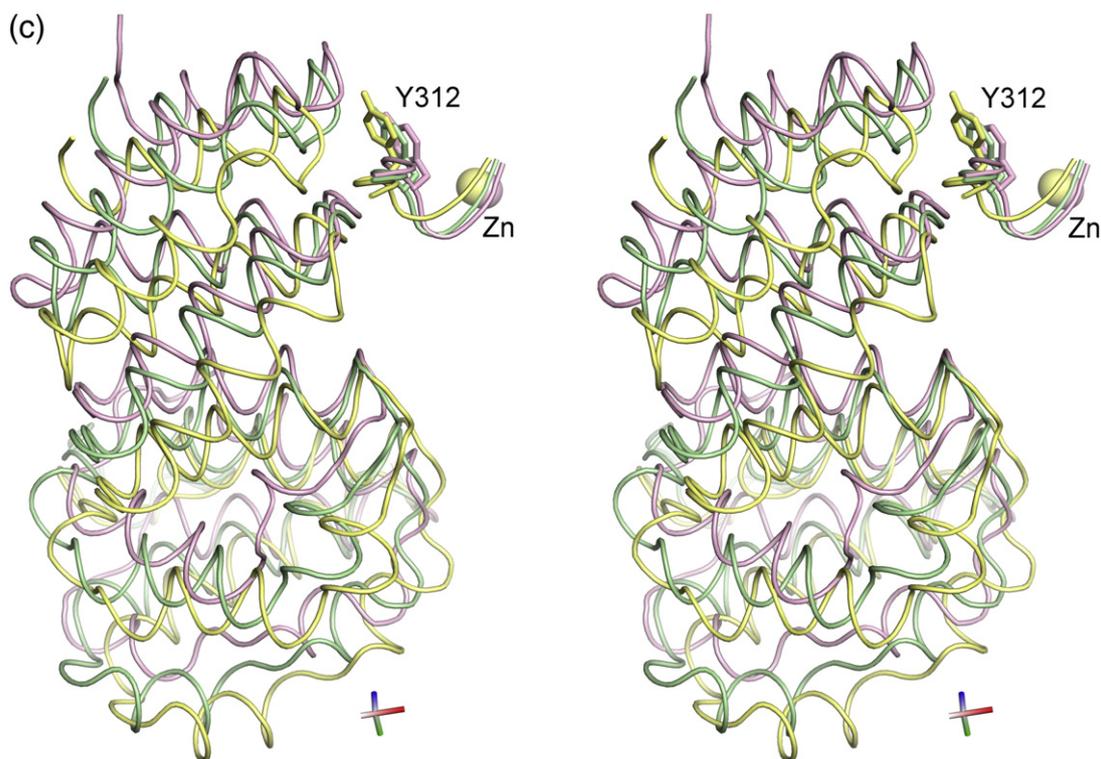


Fig. 2 (legend on previous page)

similar), and here the RMSD is 1.28 Å for 119 C^α positions. The larger insertions that are found in this domain are situated in the loop areas that connect the helices. The largest insertion of 12 residues forms a long meandering loop with no secondary structure elements. For the N-terminal domain (33.8% identity, 50% similarity), there are only minor structural differences, with the RMSD being 0.67 Å for 144 C^α positions. The extension at the N terminus forms an extended loop with a small helix that sits on top of the C-terminal helical lobe of the catalytic domain. All further deletions and insertions are located in loop areas.

The C-terminal domain shows not only the largest local rearrangements but also its position relative to the other two domains has been altered in comparison with human LTA₄H (Fig. 2a and b). While the N-terminal and catalytic domains have a similar position with respect to each other, the C-terminal domain has been rotated 25°. The main impact of this rotation is that the cavity surrounding the catalytic Zinc has become wider and more accessible to solvent. The opening of the active site is not due to a crystallographic artefact, since both crystal forms of the native enzyme show the same rotation of the C-terminal domain, in spite of the fact that the observed glutathione molecule in the P3₂21 form is bound to a residue (Cys147) near the opening of the cavity.

Inhibitor binding causes a domain movement and closure of the active site

The large conformational change of the C-terminal domain suggested that scLTA₄H might undergo an induced fit upon substrate or inhibitor binding. In order to explore this possibility, the structure of the complex between scLTA₄H and the aminopeptidase inhibitor bestatin was solved. To avoid possible crystal-induced bias, crystals of the complex were obtained by co-crystallisation rather than a soaking procedure. The structure could be solved using the native scLTA₄H as a model, and after refinement, unambiguous density could be seen for the whole model. $F_o - F_c$ difference density maps clearly showed the presence of the bestatin molecule in the active site (Fig. 3a). The conformation of the inhibitor bound to scLTA₄H is almost identical to what is seen in the complex with the human enzyme,²¹ with similar interactions to the zinc ion by the hydroxyl and carbonyl oxygen atoms. The hydroxyl oxygen also makes an interaction to Tyr429, while the carbonyl oxygen makes an additional H-bond to Glu316. The phenyl ring of bestatin is bound in a hydrophobic pocket lined with Tyr312, Met315, Phe359, and the aliphatic part of the side chain of Glu186. This pocket is closed off at the bottom by the side chain of Phe424 (Fig. 3b, schematic picture). The main-chain amides of

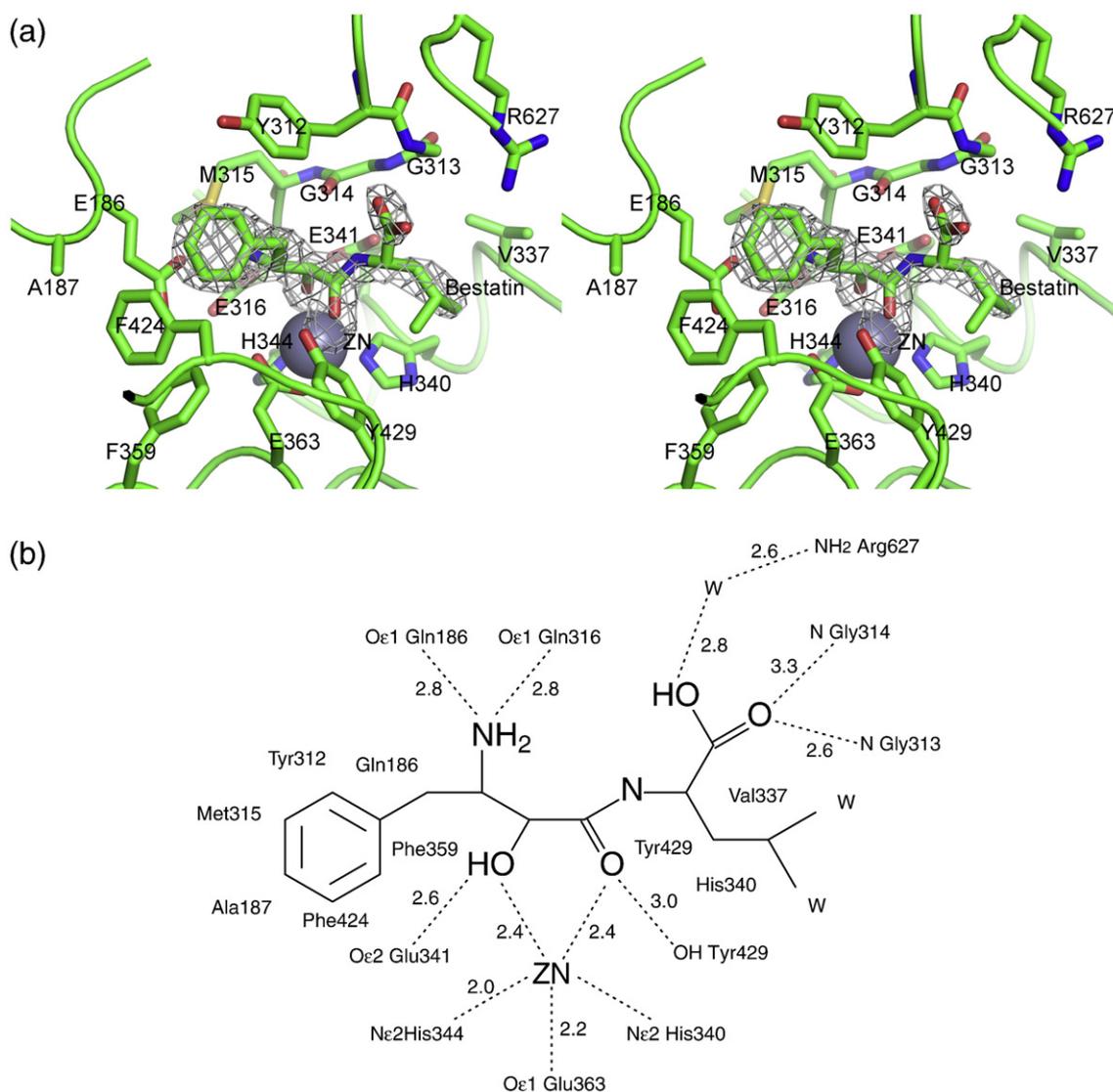


Fig. 3. (a) Bestatin binding in *scLTA4H*. Simulated omit $F_o - F_c$ density calculated using Phenix³² at 2.5σ level is shown for the bestatin molecule. Sequence numbering is according *scLTA4H*. (b) Schematic figure of bestatin binding in *scLTA4H*. Hydrogen bonds are depicted in dashed lines.

Gly313 and Gly314 make hydrogen bonds to one of the carboxyl oxygens of bestatin. Glu186 and Glu316 bind the amine nitrogen at the other end of the bestatin molecule.

The most obvious effect of the binding of bestatin is the closing of the wide open cavity present in the native *scLTA4H*. When the overall structures of the native and complexed enzyme are compared, it can be seen that the C-terminal domain has rotated 13.4° and the overall arrangement is more like the human enzyme. The large cavity seen in the native *scLTA4H* has become smaller and the distance from one wall of the cavity to the other has decreased, on average, 3 to 4 Å with a maximum difference of 6 Å in width near the position where

the phenyl ring of the bestatin is located (Fig 2a and b and Supplementary movie).

From an analysis using the protein domain motion server DynDom,³³ the main moving domain consists of a loop between residues 305 and 314 and the C-terminal domain between residues 510 and 671. The bending residues are 305–306, 313–314, and 510–511. The domain movement is primarily rotational with a total rotation of 13.4° and only a small translation of 0.4 Å. Loop 305 to 314 is part of one of the fingerprint sequences (block PR00756C). This particular fingerprint, which contains the “PYGA/GMEN” peptide substrate specificity motif, is involved in peptide substrate recognition.^{25,34} The Gly residues 313 and 314 are involved in the

backbone recognition of the peptidyl substrate molecule,²⁵ while the side chains of residues 312 and 316 are part of the binding pocket for the side chain of the P1 residue of the peptidyl substrate. This S1 pocket is also part of the binding site for the lipid substrate in the epoxide hydrolase active site in human LTA4H.

The key residue for the conformational changes appears to be Tyr312, which moves 2.6 Å at the C^α atom position and 4.6 Å at the tip of the hydroxyl group upon recognition of the phenyl group of bestatin (Fig. 4). This change in conformation breaks a hydrogen bond to the side chain of Gln167. The side chain of Gln167, in turn, moves in position and forms a hydrogen bond to the backbone carbonyl oxygen of Tyr657, which has moved 3.4 Å, and the side chain of Ser309. A second trigger is formed by the backbone of residues Gly313 and Gly314, which also change positions such that hydrogen bonds are made to atoms in the bestatin molecule. Hydrogen bonds between the backbone atoms of Gly313 and Gly314 and the side chain of Arg627 in the native *sc*LTA4H are broken upon inhibitor binding, and the guanidinium group of Arg627 moves 1.8 Å towards the inhibitor with which it makes an interaction mediated through a water molecule. The overall effect of these changes is that the C-terminal domain is pulled forward and the outer helices in the structure move as rigid elements about 4.2 Å. Many more small alterations in the interface that is mainly formed between the N-terminal and

C-terminal domains occur (see Fig. 5 for interactions in this interface in the open and closed form). In contrast, no major changes can be seen in the other interface, which is mainly between the other end of the C-terminal and catalytic domains.

In the structurally related aminopeptidase N from *E. coli*, Met260, which is at an equivalent position as Tyr312 in *sc*LTA4H, also changes conformation upon bestatin binding.^{18,22} However, in this enzyme, Met260 in the native form occupies the S1 pocket in which the phenyl ring of bestatin binds. Upon binding of the inhibitor, the residue moves away and space becomes available for the phenyl ring. Furthermore, the large domain movements, as seen in *sc*LTA4H, are not observed in this protein nor are the adjustments of the Gly residues of the PYGA/GMEN fingerprint that are made for optimal hydrogen bonding to the inhibitor. For another related structure, the Tricorn interacting factor F3, three different states of the protein involving domain movements have been observed.¹⁹ These states are characterized by a difference in position of the fourth domain, which is equivalent to the C-terminal domain of LTA4H, leading to a widening (maximum 5 Å) of the cleft between the domains. These three conformations were labelled as open, intermediate, and closed forms. The cleft in the F3 enzyme is much bigger, but the domain changes observed are similar to those in *sc*LTA4H. Since there is no inhibitor/substrate-bound form of F3, it is difficult to assess the origin

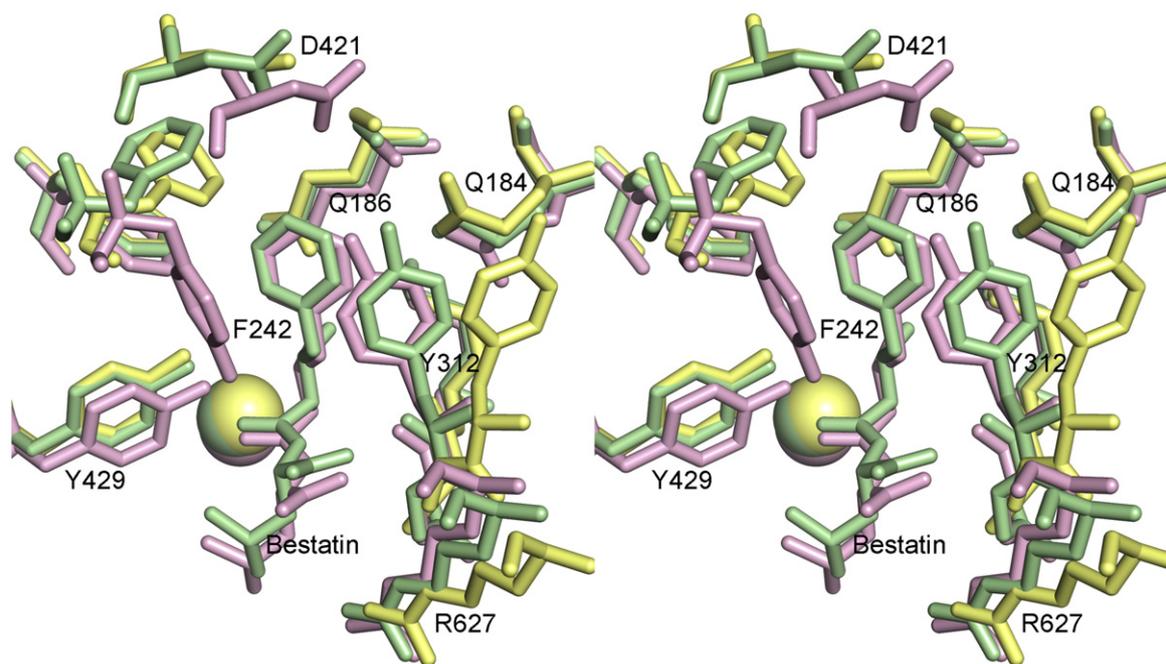


Fig. 4. Conformational changes of the residues forming the bestatin binding pocket. *Sc*LTA4H is shown in yellow, *sc*LTA4H in complex with bestatin in green, and human LTA4H in pink. Amino acid sequence numbering is the same as in *sc*LTA4H.

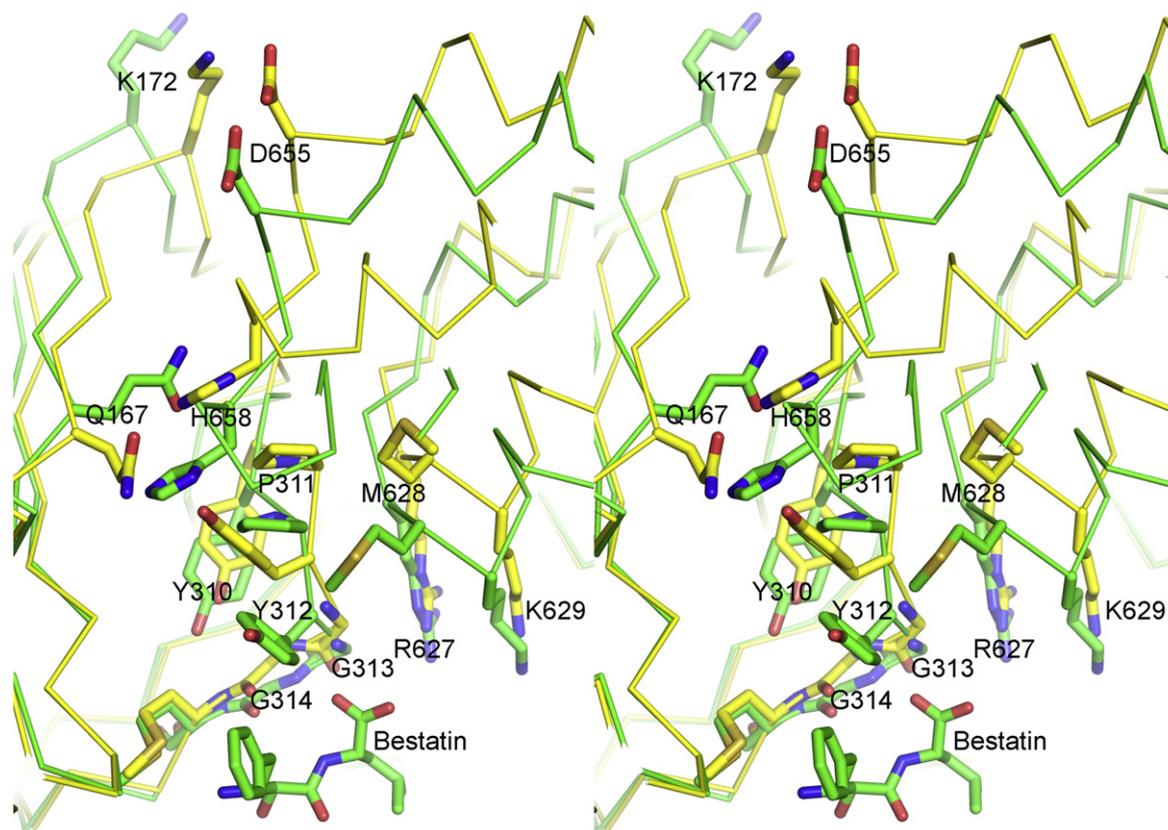


Fig. 5. Stereo figure showing the conformational changes in the interface that mediate the domain movements between the open and closed forms of *scLTA4H*. Open *scLTA4H* is in yellow, and closed *scLTA4H* in complex with bestatin is in green. The bestatin molecule is shown in green. For key residues, the side chains are depicted as stick models.

of the observed differences. The F3 enzyme contains an insertion right in front of the PYGA/GMEN fingerprint, and Tyr312 is replaced by an Ala. However, the position of Tyr351, a putative catalytic residue (Tyr429 in *scLTA4H*, Tyr383 in human LTA4H), is rotated nearly 3 Å away from the zinc atom, suggesting that none of the observed structures of the F3 enzyme is productive and that further conformational changes have to occur to activate this protein.¹⁹

Absence of critical Pro residues may explain flexibility of the C-terminal domain

For the human enzyme, no changes in the relative position of the C-terminal domain have been observed in previously solved structures. Although most human LTA4H structures in the PDB database contain inhibitors or substrate analogues, the structure of the R563A mutant (PDB code 1SQM)³⁵ was obtained without the presence of an inhibitor. This structure only shows minimal differences compared to the human LTA4H in complex with bestatin (RMSD). Some minor rearrangements are observed near the site of mutation, but there are no overall changes in the domain architecture or significant

changes in conformation of residues involved in inhibitor binding. One reason for the seemingly increased rigidity of human LTA4H might be the presence of two Pro residues in the first bending region at the equivalent positions of Val307 and Asp308. The presence of two rigid Pro residues just at the bending sites might prohibit the domain movement as seen in the yeast protein. The cold-adapted aminopeptidase from *C. psychrerythrae* has also prolines at these positions. The structure of this protein without inhibitor is very similar to the human enzyme complexed with bestatin, and no opening of the active site can be seen.²⁰ Interestingly, in most members of the LTA4H family, a double or, in some cases, a single proline is found at these positions. Only in LTA4H homologues from *S. cerevisiae*, *C. glabrata*, *Z. rouxii*, and *V. polyspora*, no proline residues are present at these positions and instead a Val-Asp is found.

Comparing the substrate binding pocket in the human and yeast enzymes

The difference in peptidase specificity between human LTA4H (Arg) and *scLTA4H* (Leu) can be explained by changes in both the size and nature of

the pocket. The different position of the side chain of Phe424, as discussed above, makes the pocket wider and shallower and thus better suited for aliphatic side chains. Furthermore, a loop in which Asp421 (Asp375 in human LTA4H) is located has changed conformation due to a deletion at position 416. In comparison with human LTA4H, this Asp is placed at a larger distance from the zinc ion and is engaged in a salt bridge with His189. Therefore, the direct charge compensation that this residue provides for optimal binding of the arginine side chain of Arg-Ser-Arg in human LTA4H²⁵ is not possible in the yeast enzyme. In human LTA4H, Asp375 is an essential residue for the epoxide hydrolase activity,³⁶ but this is not the case for scLTA4H²⁸ whose low epoxide hydrolase activity is not affected by mutation of this residue.

The very low level of LTA₄ hydrolase activity exhibited by scLTA4H can be explained by the fact that the extension of the cavity, which in the human enzyme is the protected binding site for the hydrophobic tail of the labile substrate LTA₄, has disappeared. This is mainly due to a local rearrangement. A single residue insertion at position 416 in the yeast structure changes the conformation of the whole loop between residues 412 and 425. Phe424 changes in position and the side chain rotates around the C^α-C^β bond by 180° and thereby fills part of the binding site. The phenyl ring is involved in aromatic stacking interactions with Phe359, and further hydrophobic interactions are made with Trp356. In the human enzyme, the equivalent residue of Phe424 is a Tyr (Tyr378), and a rotamer position as seen in the yeast enzyme is not possible for this residue due to steric hindrance from atoms in loop 412–425. In the human enzyme, this side chain is instead involved in a hydrogen bond with the catalytic residue Tyr383. The rearrangement of Phe424 in scLTA4H also changes the environment of the zinc atom. In the human enzyme, the binding pocket for the substrates is most narrow near the zinc ion and a bottleneck is formed. This is mainly due to the position of the side chain of Tyr378 that forms one side of the pocket. This narrow passage is not present in scLTA4H due to the change in position of the side chain of the equivalent Phe424 (Fig. 6a and b) and instead the passage is wider and more water accessible.

The change in conformation of Phe424 in scLTA4H is interesting. In mutagenesis studies of both the yeast and human enzyme, it was shown that this residue plays a role in the epoxide hydrolase substrate binding. A mutation of the equivalent Tyr378 into a Phe enhances the k_{cat} of the epoxide hydrolase activity of human LTA4H by almost a factor of 3, but the precise product specificity is lost.³⁸ The observed substrate suicide inhibition linked to Tyr378 is also abolished when this residue is mutated into a Phe.³⁹ In scLTA4H, a

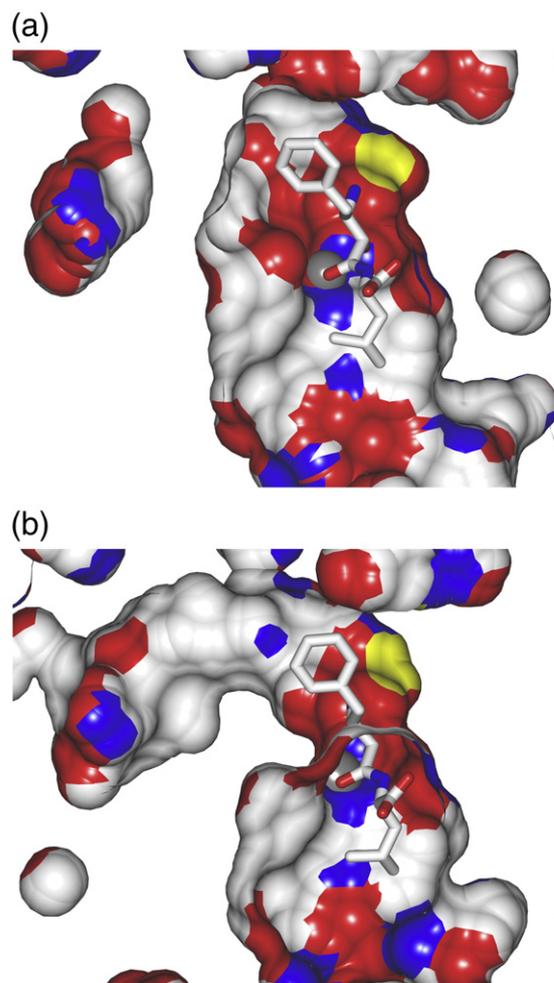


Fig. 6. Surface presentations for the active site near the Zn ion in scLTA4H and human LTA4H, both in complex with bestatin. The figure was prepared using the program CCP4MG.³⁷ Colouring is according to atom type: white, carbon; red, oxygen; blue, nitrogen; and yellow, sulfur atoms. (a) The active-site surface of scLTA4H. (b) The active-site surface of human LTA4H. In both figures, the bestatin molecule is drawn in sticks and the position of the Zn ion is depicted by a dark grey sphere.

mutation of Phe424 into Tyr means an increase in the epoxide hydrolase activity as well as an improved specificity.⁴⁰ It was postulated that this increase of activity and specificity was mainly due to the introduced hydroxyl group when mutating from Phe to Tyr. In human LTA4H, the hydroxyl group provides a hydrogen bond to Tyr383, but the moiety itself has no direct interactions with the substrate molecule. In that conformation, however, the phenyl moiety makes stacking interactions with the phenyl ring of bestatin.²¹ Therefore, it seems that the effects on activity and specificity are regulated by the position of this residue, which in human LTA4H is partly maintained by the hydrogen bond between

the hydroxyl groups of Tyr378 and Tyr383. As discussed above, the conformation of the Tyr378 in the human enzyme makes the active-site pocket narrow near the zinc atom, while in the yeast enzyme the channel near the zinc is considerably broader.

However, other changes around the active site play a role as well. In the cold-adapted aminopeptidase, the equivalent residue to Phe424 is also a Phe, but no large structural differences between the position of this residue and that of the Tyr378 in the human LTA4H can be seen²⁰ and the specificity of this protein has not changed. In *scLTA4H*, new stacking interactions between Phe424 and Phe359 occur, while in the cold-adapted protein the residue at position 359 is not a Phe but instead a Leu. Moreover, just like human LTA4H, the cold-adapted enzyme does not have an insertion at position 416, which means that a similar conformation as the yeast enzyme would also introduce some steric hindrance from residues at the equivalent positions of 420 to 421.

The change in size of the active site pocket is not only due to local rearrangements and changes in side-chain conformations but also due to the overall domain movement seen in the open and closed forms of *scLTA4H*. From an analysis using the DynDom server,³³ it can be seen that the C-terminal domain in *scLTA4H* complexed with bestatin is not as closed as in human LTA4H (*scLTA4H* + bestatin, 65% closed; human LTA4H + bestatin, 83% closed, in reference to the uncomplexed *scLTA4H*). Thus, in the *scLTA4H*-bestatin complex, the overall conformation of the C-terminal domain is in between the positions it has in the native *scLTA4H* and the human LTA4H. One of the causes for this difference might be the substitution of a Tyr for a Phe at position 310 (*scLTA4H* numbering). Residue 310 is part of the main rotating areas positioned between bending residue 305–306 and 313–314. The additional hydroxyl group requires more space, and this is provided by the translation of loop 308–312. The loop 624–629 from the C-terminal domain of *scLTA4H* has also been displaced slightly. This means that the S1 pocket is somewhat larger in the *scLTA4H* and the domains cannot come as close as in the human enzyme.

The structures of *scLTA4H* with and without bestatin suggest that during evolution, the introduction of the epoxide hydrolase activity in this class of enzymes has been a process involving subtle structural changes. The presence of the essential catalytic residues for the epoxide hydrolase activity within *scLTA4H* makes it plausible that this class of related aminopeptidases was predisposed to obtain a moonlighting epoxide hydrolase function. This is collaborated by the fact that *scLTA4* has a low epoxide hydrolase activity. Since all the required catalytic residues for the epoxide hydrolase activity are present in *scLTA4H*, the main adaptation seems

to be a gradual process in which these residues are positioned such that their relative positions are compatible with a turnover of the lipid leukotriene substrate. At the same time, the substrate binding site has been adapted to provide a shielded hydrophobic environment in which the labile allylic epoxide of LTA₄ is protected from unwanted premature cleavage resulting in biologically inactive products. In the yeast enzyme, the pocket is both too limited in size and too accessible to solvent to provide this environment. Thus, as in the case for the directed laboratory evolution experiments on serum paraoxonase and carbonic anhydrase II,⁹ it seems that an innate compatibility combined with gradual accumulation of mutations leads to acquisition of an additional function.

The fact that *scLTA4H* displays an induced fit is an interesting observation. Many members of the M1 family seem to display a certain degree of induced fit (see paragraph above), a feature that has never been observed for human LTA4H. It is attractive to speculate that a preformed pocket functioning through a lock-and-key mechanism is better suited to supply the protective and precise environment required for hydrolysis of LTA₄ into LTB₄. Further studies are required to fully understand the adaptation of this intriguing class of bifunctional enzymes.

Materials and Methods

Materials

Reagents for determination of protein concentration were obtained from Bio-Rad, and bestatin hydrochloride and other chemicals were purchased from Sigma-Aldrich unless indicated otherwise. All chemicals used were of the highest quality commercially available.

Cloning of GST and His-tagged *scLTA4H*

The full-length sequence of *scLTA4H* was cloned into the expression vectors petM-11 and petM-30 (a kind gift of Dr. G. Stier, EMBL Hamburg), which add an N-terminal (His)₆-tag and GST-(His)₆-tag, respectively, to the expressed recombinant proteins. The affinity tags can be removed by proteolytic cleavage by tobacco etch virus protease.

A standard PCR reaction was carried out for amplification of the *scLTA4H* gene with the primers 5'-ccatgggtgctctttcaattgagcagag-3' (forward) and 5'-ggcgccgctcaagacctaaatctgtt-3' (reverse) and *Pfu* Turbo DNA polymerase (Promega). The forward primer introduced an NcoI restriction site, whereas the reverse primer introduced a NotI restriction site downstream of the stop codon. The amplified DNA fragment of *scLTA4H* was purified from a 1% agarose gel and cloned into a pCR-Blunt II-TOPO cloning vector using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) according to the

manufacturer's instructions. Six colonies from an agar plate were amplified and purified by a plasmid miniprep (Qiagen). The plasmids were digested with EcoRI (Fermentas) and analysed on a 1% agarose gel for insert. Positive clones were labelled as pCR-scLTA4H and further verified by automated DNA sequencing using a BigDye Terminator version 3.0 DNA sequencing kit (Applied Biosystems). The plasmids pCR-scLTA4H, pETM-11, and pETM-30 were purified on a midiprep scale and double-digested with NcoI and NotI (Fermentas). Digestion reactions were run on a 1% agarose gel. The DNA fragment encoding scLTA4H and linear vector fragments of pETM-11 and pETM-30 were purified from the agarose gel using a gel extraction kit (Qiagen). The purified fragment of scLTA4H was ligated into pETM-11 and pETM-30 using T4 DNA ligase (Fermentas) and transformed into competent *E. coli* TOP-10 cells (Invitrogen). Plasmids from six colonies of each construct were miniprep purified and analysed by double digestion using NcoI and NotI. Positive clones were subjected to expression analysis. All six clones were confirmed to contain the correct insert and transformed to BL21 (DE3) RIL cells. Small-scale expression tests showed no variation in protein expression among the different clones.

Expression and purification of GST-(His)₆-tagged scLTA4H

In large-scale expressions, both the (His)₆-tagged and GST-(His)₆-tagged proteins were expressed very well (>25 mg protein purified per liter of culture), but using GST affinity purification as the first purification step gave higher-purity protein. For the subsequent work, the GST-(His)₆-tagged protein was used. The protein was expressed in BL21 (DE3) RIL cells by growing 2×1-liter cultures at 37 °C to OD 0.5, lowering the temperature to 20 °C and inducing expression with 1 mM IPTG at OD 0.7–0.8, followed by 4–6 h of incubation at 20 °C. The cells were harvested by centrifugation at 6000 rpm and the pellets were frozen.

A frozen cell pellet from 2 liters of expression was resuspended in 50 ml of 60 mM Tris–HCl (pH 7.5), 100 mM NaCl, 10 mM NaN₃ containing complete mini EDTA-free protease inhibitor (Roche) and lysed by sonication for 2×5 min. The insoluble cell debris was removed by centrifugation at 30,000g for 30 min, and the lysate filtered through a 0.45-μm syringe filter.

The cleared lysate was mixed with 5 ml of glutathione Sepharose 4B (GE Healthcare) that had been pre-equilibrated with 30 mM Tris–HCl (pH 7.5), 50 mM NaCl, 5 mM NaN₃ and was incubated for 60 min on a shaking platform at room temperature. The lysate/resin mix was transferred to a gravity flow column (BioRad EconoPack), and the resin was allowed to drain. After washing with several column volumes of 30 mM Tris–HCl (pH 7.5), 50 mM NaCl, 5 mM NaN₃, the bound protein was eluted with 3 column volumes of 30 mM Tris–HCl (pH 7.5), 50 mM NaCl, 5 mM NaN₃, 10 mM reduced glutathione.

The eluted scLTA4H protein was mixed with AcTEV protease (Invitrogen; 1 μl for each estimated mg of protein) and incubated at 4 °C overnight. To remove the cleaved affinity tag and uncleaved protein, the protein solution was adjusted to 500 mM NaCl and 10 mM imidazole and mixed with 5 ml of Ni²⁺ Sepharose FF

(GE healthcare) that had been pre-equilibrated with 30 mM Tris–HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole, 5 mM N₃. After 60 min of incubation on a shaking platform at room temperature, the mix was transferred to a gravity flow column and the resin was allowed to drain. This flow-through contained the cleaved scLTA4H and was dialyzed overnight against 2 liters of 20 mM Tris–HCl (pH 8.0).

The protein solution was further diluted with 5 volumes of 20 mM Tris–HCl (pH 8.0) and loaded on a 5-ml HiTrap Q FF ion exchange column (GE Healthcare) pre-equilibrated in 20 mM Tris–HCl (pH 8.0). The bound protein was eluted by a steep salt gradient of 0–100% 20 mM Tris–HCl (pH 8.0), 500 mM NaCl over 5 ml, and the eluted fractions were pooled and concentrated to >3 mg/ml. The protein preparation was judged to be >90% pure by SDS-PAGE.

scLTA4H was further purified by injection on a Superdex 200 10/300 gel filtration column (GE Healthcare) pre-equilibrated with 20 mM Tris–HCl (pH 7.5) with a flow of 0.3 ml/min. It is worth noting that a bimodal size distribution was always observed during the gel filtration, with one monomer peak and one multimer peak representing 20–60% of total protein amounts. Additional aggregation of protein from the monomer peak was observed already after 3 days of storage at 4 °C, and gel-filtration was always performed immediately before setting up crystallisation drops. Fractions from the monomer peak were pooled and concentrated to 2–12 mg/ml.

Crystallisation of the P₃21 crystal form

Using the PACT screen (Molecular Dimensions) and vapour diffusion with scLTA4H at a concentration of 6.6 mg/ml, initial crystal hits were identified with MMT buffer (L-malic acid, MES, Tris buffer; pH 5–6.5) with 25% polyethylene glycol (PEG) 1500 and in Hepes [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; pH 7.5] with 20% PEG 6000. Further refinement of the crystallisation conditions revealed that scLTA4H (2–6 mg/ml) would crystallise in Hepes or MMT buffer between pH 7 and 9 with 12–18% PEG 6000 or 15–20% PEG 1500. Crystals made this way were extremely small, flaky, and mixed with heavy precipitate. To obtain diffraction quality crystals, hanging drops of 2 μl of protein (6 mg/ml) and 1 μl of mother liquor [0.1 M MMT buffer (pH 8), 12–14% PEG 6000] were allowed to equilibrate against the mother liquor for 18–24 h before the drops were streak seeded from earlier crystals. Crystals of acceptable size and morphology generally appeared overnight in the seeded drops. All crystallisations were performed at 20 °C.

Purification and crystallisation of the P₂1₂1 crystal form

The purification and crystallisation of the P₂1₂1 form has been described elsewhere.²⁹ The crystals were grown in liquid–liquid diffusion experiments using melting point capillaries in which 5 μl of protein (6 mg/ml) in 10 mM Tris (pH 8.5) was layered on 5 μl of 14% PEG 8000 and 100 mM Tris (pH 7.3). Rod-shaped crystals appeared in 8 to 10 weeks at room temperature.

Crystallisation of the complex between scLTA4H and bestatin

For crystallising the complex between scLTA4H and bestatin, the GST-(His)₆-tagged protein was purified as described above. The protein was crystallised in the presence of the inhibitor using the hanging drop vapour diffusion. In the initial experiments, microcrystals were found in condition #41 [0.085 M Hepes-Na (pH 7.5), 8.5% (v/v) isopropanol, 17% (w/v) PEG 4000, and 15% anhydrous glycerol] from the Crystal Screen Cryo (Hampton Research). After additive and pH screening, good-quality crystals were obtained in 15% glycerol, 16.95% PEG 4000, 8.5% isopropanol, and 0.1 M MMT buffer (pH 7.0) with 0.1 M betaine monohydrate as an additive. The ratios of protein (9.6 mg/ml) with bestatin (6 mM), reservoir, and additive solutions in the final crystallisation drops were 2:1:0.5.

Data collection and processing

Data for the P₃₂₁ form of scLTA4H were collected at station I911-2, and the data for the P₂₁₂₁ scLTA4H at station I711 at MAX lab, Lund, Sweden, both on Mar165 CCD detectors (Rayonix, MarResearch, USA). Data for the enzyme inhibitor complex were collected at beam station I911-3 at MAX lab, Lund, Sweden, on a Mar225 CCD detector. All data were collected at 100 K. Prior to data collection, the crystals were transferred to appropriate cryo-solutions before being flash-cooled in a stream of nitrogen gas at 100 K. The P₃₂₁ crystals were transferred to a drop of ultimate cryoprotectant (8% glycerol, 8% ethylene glycol, 9% sucrose, and 2% glucose) for less than 1 min and flash frozen in the cryo-cooling stream at the data collection station.

The P₂₁₂₁ crystals were transferred in a stepwise manner to a cryoprotecting solution consisting of 20% l-(+)-2,3-butanediol, 8% PEG 8000, 50 mM Tris (pH 7.3), and 2.5 mg/ml scLTA4H.²⁹ For the inhibitor complex, the crystals were soaked in a 1:1 mixture of reservoir and UCP solutions and, after 1 min, transferred to the cryogenic stream of nitrogen. Complete data sets were obtained from single crystals. Data processing and reduction for the P₃₂₁ data set was performed with the program MOSFLM⁴¹ and programs from the CCP4-suite⁴²; the data for the P₂₁₂₁ crystal form were processed by Denzo and Scalepack⁴³, while the data for the complex were processed with XDS.⁴⁴

Molecular replacement and refinement

In case of the P₃₂₁ crystal form of scLTA4H, the phasing problem was solved by molecular replacement using PHASER³⁰ with the structure of the human enzyme (PDB code 1hs6)²¹ as the searching model. Only protein atoms were included in the calculations.

Subsequent refinement was carried out using Phenix.³² $2F_o - F_c$ and $F_o - F_c$ electron density maps were used in COOT⁴⁵ for model inspection and rebuilding. At position Cys155, clear $F_o - F_c$ density indicated the presence of an attached glutathione molecule remaining from the purification procedure. The glutathione could be built unambiguously and was refined using full occupancy.

The structure of the P₂₁₂₁ crystal form was solved by molecular replacement with PHASER³⁰ using the structure of the P₃₂₁ form of scLTA4H as the searching model. The crystal form has cell dimensions in which *b* and *c* have almost the same values, and this gave rise to pseudo-merohedral twinning. Calculations by XTRIAGE from the Phenix package⁴⁶ made it possible to determine the twin fractions of data sets from different crystals; by using data from a crystal with the lowest twin fraction (0.07), a solution for the molecular replacement problem was obtained. The structure was refined by Phenix.³² For the P₂₁₂₁ form, one l-(+)-2,3-butanediol molecule and one tris(hydroxymethyl)aminomethane molecule were identified in difference density maps.

In case of the inhibitor complex, using the P₂₁₂₁ structure as a search model gave a clear signal in molecular replacement calculations in PHASER.³⁰ Rigid-body refinement with the different domains of the protein as rigid objects was used to refine the clear domain movement that had occurred. Subsequent refinement by Phenix³² and inspection and manual model building in $2F_o - F_c$ and $F_o - F_c$ electron density maps in Coot⁴⁵ were followed. Clear difference electron density indicated the presence of the bestatin molecule near the Zn ion.

Validation by MolProbity⁴⁷ indicated areas of the structure that needed further rebuilding for all the different structures.

Detailed statistics for the refinement of the different structures can be found in Table 2.

Accession numbers

The coordinates of the structures are deposited at the PDB with the following accession codes: scLTA4H P₃₂₁ crystal form, 2xpy; scLTA4H P₂₁₂₁ crystal form, 2xpz; and the bestatin-scLTA4H complex, 2xq0.

Acknowledgements

The formulation of a ultimate cryoprotectant solution recipe was generously provided by Mark Pusey. This work was supported by an EMBO Long-Term Fellowship, the Swedish Research Council, CIDaT (Vinnova), and Karolinska Institutet.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2010.11.059](https://doi.org/10.1016/j.jmb.2010.11.059)

References

1. Samuelsson, B. (1983). Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation. *Science*, **220**, 568–575.

2. Ford-Hutchinson, A. W. (1990). Leukotriene B₄ in inflammation. *Crit. Rev. Immunol.* **10**, 1–12.
3. Haeggström, J. Z., Wetterholm, A., Vallee, B. L. & Samuelsson, B. (1990). Leukotriene A₄ hydrolase: an epoxide hydrolase with peptidase activity. *Biochem. Biophys. Res. Commun.* **173**, 431–437.
4. Minami, M., Ohishi, N., Mutoh, H., Izumi, T., Bito, H., Wada, H. *et al.* (1990). Leukotriene A₄ hydrolase is a zinc-containing aminopeptidase. *Biochem. Biophys. Res. Commun.* **173**, 620–626.
5. Orning, L., Gierse, J. K. & Fitzpatrick, F. A. (1994). The bifunctional enzyme leukotriene A₄ hydrolase is an arginine aminopeptidase of high efficiency and specificity. *J. Biol. Chem.* **269**, 11269–11273.
6. Haeggström, J. Z. (2004). Leukotriene A₄ hydrolase/Aminopeptidase, the gatekeeper of chemotactic leukotriene B₄ biosynthesis. *J. Biol. Chem.* **279**, 50639–50642.
7. Jeffery, C. J. (2004). Moonlighting proteins: complications and implications for proteomics research. *Drug Design Today: Targets*, **3**, 71–78.
8. Huberts, D. H. E. W. & van der Klei, I. J. (2010). Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta*, **1803**, 520–525.
9. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M. C. Q., Roodveldt, C. & Tawfik, D. S. (2005). The ‘evolvability’ of promiscuous protein functions. *Nat. Genet.* **37**, 73–76.
10. Snelgrove, R. J., Jackson, P. L., Hardison, M. T., Noerager, B. D., Andrew Kinloch, A., Gaggari, A. *et al.* (2010). A critical role for LTA₄H in limiting chronic pulmonary neutrophilic inflammation. *Science*, **330**, 70–74.
11. Barret, A. (1998). Family M1 of membrane alanyl aminopeptidase. In *Handbook of Proteolytic Enzymes* (Barret, A. J., Rawlings, N. D. & Woessner, J. F., eds.), pp. 994–996, Academic Press, London.
12. Green, F. A. (1987). In vivo generation of 5-lipoxygenase products in frogs and toads. *Biochem. Biophys. Res. Commun.* **148**, 1533–1539.
13. Green, F. A., Herman, C. A., Herman, R. P., Claesson, H. E. & Hamberg, M. (1987). Monohydroxyeicosatetraenoic acid and leukotriene production by the inflammatory cells of *Xenopus laevis*. *J. Exp. Zool.* **243**, 211–215.
14. Habenicht, A. J., Goerig, M., Rothe, D. E., Specht, E., Ziegler, R., Glomset, J. A. & Graf, T. (1989). Early reversible induction of leukotriene synthesis in chicken myelomonocytic cells transformed by a temperature-sensitive mutant of avian leukemia virus E26. *Proc. Natl Acad. Sci. USA*, **86**, 921–924.
15. Pettitt, T. R., Rowley, A. F., Barrow, S. E., Mallet, A. I. & Secombes, C. J. (1991). Synthesis of lipoxins and other lipoxygenase products by macrophages from the rainbow trout, *Oncorhynchus mykiss*. *J. Biol. Chem.* **266**, 8720–8726.
16. Knight, J., Holland, J. W., Bowden, L. A., Halliday, K. & Rowley, A. F. (1995). Eicosanoid generating capacities of different tissues from the rainbow trout, *Oncorhynchus mykiss*. *Lipids*, **30**, 451–458.
17. Baset, H. A., Ford-Hutchinson, A. W. & O’Neill, G. P. (1998). Molecular cloning and functional expression of a *Caenorhabditis elegans* aminopeptidase structurally related to mammalian leukotriene A₄ hydrolases. *J. Biol. Chem.* **273**, 27978–27987.
18. Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F. *et al.* (2006). Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from *Escherichia coli* and conformational change of methionine 260 involved in substrate recognition. *J. Biol. Chem.* **281**, 33664–33676.
19. Kyrieleis, O. J., Goettig, P., Kiefersauer, R., Huber, R. & Brandstetter, H. (2005). Crystal structures of the tricorn interacting factor F3 from *Thermoplasma acidophilum*, a zinc aminopeptidase in three different conformations. *J. Mol. Biol.* **349**, 787–800.
20. Bauvois, C., Jacquamet, L., Huston, A. L., Borel, F., Feller, G. & Ferrer, J. L. (2008). Crystal structure of the cold-active aminopeptidase from *Colwellia psychrerythraea*, a close structural homologue of the human bifunctional leukotriene A₄ hydrolase. *J. Biol. Chem.* **283**, 23315–23325.
21. Thunnissen, M. M. G. M., Nordlund, P. & Haeggström, J. Z. (2001). Crystal structure of human leukotriene A₄ hydrolase, a bifunctional enzyme in inflammation. *Nat. Struct. Biol.* **8**, 131–135.
22. Addlagatta, A., Gay, L. & Matthews, B. W. (2006). Structure of aminopeptidase N from *Escherichia coli* suggests a compartmentalized, gated active site. *Proc. Natl Acad. Sci. USA*, **103**, 13339–13344.
23. Nocek, B., Mulligan, R., Bargassa, M., Collart, F. & Joachimiak, A. (2008). Crystal structure of aminopeptidase N from human pathogen *Neisseria meningitidis*. *Proteins*, **70**, 273–279.
24. McGowan, S., Porter, C. J., Lowther, J., Stack, C. M., Golding, S. J., Skinner-Adams, T. S. *et al.* (2009). Structural basis for the inhibition of the essential *Plasmodium falciparum* M1 neutral aminopeptidase. *Proc. Natl Acad. Sci. USA*, **106**, 2537–2542.
25. Tholander, F., Muroya, A., Roques, B. P., Fournié-Zaluski, M. C., Thunnissen, M. M. G. M. & Haeggström, J. Z. (2008). Structure-based dissection of the active site chemistry of leukotriene A₄ hydrolase: implications for M1 aminopeptidases and inhibitor design. *Chem. Biol.* **15**, 920–929.
26. Nars, F., Bécam, A. M. & Herbert, C. J. (1996). The sequence of 12.8 kb from the left arm of chromosome XIV reveals a sigma element, a pro-tRNA and six complete open reading frames, one of which encodes a protein similar to the human leukotriene A₄ hydrolase. *Yeast*, **12**, 493–499.
27. Kull, F., Ohlson, E. & Haeggström, J. Z. (1999). Cloning and characterization of a bifunctional leukotriene A₄ hydrolase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **274**, 34683–34690.
28. Tholander, F., Kull, P., Ohlson, E., Thunnissen, M. M. G. M. & Haeggström, J. Z. (2005). Leukotriene A₄ hydrolase: insights to the molecular evolution by homology modeling and mutational analysis of enzyme from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **280**, 33477–33486.
29. Andersson, B., Kull, F., Haeggström, J. Z. & Thunnissen, M. M. G. M. (2003). Crystallization and X-ray diffraction data analysis of leukotriene A₄ hydrolase from *Saccharomyces cerevisiae*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 1093–1095.
30. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser

- crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
31. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
 32. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
 33. Hayward, S. & Berendsen, H. J. C. (1998). Systematic analysis of domain motions in proteins from conformational change; New results on Citrate Synthase and T4 Lysozyme. *Proteins, Structure, Function and Genetics*, **30**, 144–154.
 34. Thunnissen, M. M. G. M., Andersson, B., Samuelsson, B., Wong, C. H. & Haeggstrom, J. Z. (2002). Crystal structures of leukotriene A4 hydrolase in complex with captopril and two competitive tight-binding inhibitors. *FASEB J.* **16**, 1648–1650.
 35. Rudberg, P. C., Tholander, F., Andberg, M., Thunnissen, M. M. G. M. & Haeggstrom, J. Z. (2004). Leukotriene A4 hydrolase: identification of a common carboxylate recognition site for the epoxide hydrolase and aminopeptidase substrates. *J. Biol. Chem.* **279**, 27376–27382.
 36. Rudberg, P. C., Tholander, F., Thunnissen, M. M. G. M., Samuelsson, B. & Haeggström, J. Z. (2002). Leukotriene A4 hydrolase: selective abrogation of leukotriene B4 formation by mutation of aspartic acid 375. *Proc. Natl Acad. Sci. USA*, **99**, 4215–4220.
 37. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P. *et al.* (2004). Developments in the CCP4 molecular-graphics project. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2288–2294.
 38. Mueller, M. J., Andberg, M. B., Samuelsson, B. & Haeggström, J. Z. (1996). Leukotriene A4 hydrolase, mutation of tyrosine 378 allows conversion of leukotriene A4 into an isomer of leukotriene B4. *J. Biol. Chem.* **271**, 24345–24348.
 39. Mueller, M. J., Blomster, M., Oppermann, U. C. T., Jörnvall, H., Samuelsson, B. & Haeggström, J. Z. (1996). Leukotriene A4 hydrolase: protection from mechanism-based inactivation by mutation of tyrosine-378. *Proc. Natl Acad. Sci. USA*, **93**, 5931–5935.
 40. Kull, F., Ohlson, E., Lind, B. & Haeggström, J. Z. (2001). Leukotriene A4 hydrolase from *Saccharomyces cerevisiae*, formation of leukotriene B4 and identification of catalytic residues. *Biochemistry*, **40**, 12695–12703.
 41. Leslie, A. G. W. (1992). “Recent changes to the MOSFLM package for processing film and image plate data”, Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26.
 42. Collaborative Computational Project N (1994). The CCP4 suite—programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
 43. Otwinowski, Z. (1993). In *Proceedings of the CCP4 Study Weekend Data Collection and Processing* (Saywer, L., Isaacs, N & bailey, S, eds.), pp. 80–86, Warrington: Daresbury Laboratories.
 44. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800.
 45. Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D*, **66**, 486–501.
 46. Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). Characterization of X-ray data sets. *CCP4 Newsl*, **42**; contribution 10.
 47. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J. *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 12–21.

SUPPLEMENTAL DATA:

Table S1

Fig. S1 – Fig. S6

Table S1. Description of mutants of *scLTA4H* and their relation to *humLTA4H*

Mutation(s)	Residue(s) in <i>humLTA4H</i>	Assigned role in <i>humLTA4H</i>	Type of mutation and rationale for mutant selection.	Residue conservation (1)
Wild-type <i>scLTA4H</i>	-	-	-	About 40% overall identity to <i>humLTA4H</i> .
E316Q E316A E316D	E271	Yes (19).	Test of catalytic function. Human residue involved in epoxide ring opening and binding of α -amino groups of peptides. In <i>scLTA4H</i> possibly general base in epoxide hydrolysis	Belongs to conserved GXMEN motif.
R627A R627K	R563	Yes (26).	Test of catalytic function. Human residue involved in binding of substrate carboxy group.	Conserved within mammalian LTA4H:s and occurs in some lower organisms.
E186Q E186A	Q136	No	Human mimic/ test of catalytic function. Position suggesting involvement in binding of α -amino groups of peptides. Possible general base in epoxide hydrolysis.	Gln most common within protein family, yeast and rat LTA4H has a Glu.
D421N	D375	Yes (25).	Test of catalytic function. Human residue shown to be general base in LTB ₄ formation without an identified role in peptidase activity.	Fully conserved within mammalian LTA4H:s, common in lower organisms.
F424Y	Y378	Yes (see above).	Human mimic/ test of catalytic function. Human residue involved in suicide inhibition and substrate alignment. Included in 4M.	All mammalian LTA4H:s has a Tyr and most lower organisms a Phe.
D422V	D376	No	Human mimic/ test of catalytic function. Possible domain interaction affecting shape of LTA ₄ binding pocket. Located next to the general base required for LTB ₄ production in human (D375).	All mammalian LTA4H:s have a Val residue whereas lower organisms have an Asp.
Q412V	V367	No	Human mimic/ test of catalytic function. Possibly affects nature and depth of LTA ₄ binding pocket. Included in 4M.	Val residue in all mammalian LTA4H:s, most lower organisms has a Gln.
N417del (deletion of N417)	Not present	-	Human mimic. Loop residue located close to catalytic residue. Could possibly affect catalytic residues.	Loop insertion absent in all mammalian LTA4H.
N417del/D422V	Not present/D376	No	D422V and N417del combined. Possibility to affect position of general base required for LTB ₄ production in human (D375).	(Combination of mutants.)
3M (= Q412V, T418I and F424Y)	V367, I372 and Y378	Yes, for Y378 (see above).	Human mimic. Includes 3 residues that differ in human and yeast and directly face the putative binding pocket.	(Combination of mutants.)
4M (= E186Q, Q412V, T418I and F424Y)	Q136, V367, I372 and Y378	Yes, for Y378 (31, 32).	Human mimic. Includes all four residues that differ in human and yeast and directly face the putative binding pocket.	(Combination of mutants.)
5M (as 4M plus D422V)	Q136, V367, I372, D376 and Y378	Yes, for Y378 (see above).	Human mimic. Includes the 4M mutations and in addition D422V.	(Combination of mutants.)
6M (as 5M plus N417del)	Q136, V367, not present, I372, D376 and Y378	Yes, for Y378 (see above).	Human mimic. Includes the 4M mutations and in addition N417del and D422V.	(Combination of mutants.)

Figure S1. Alignment between human and *scLTA4H*.

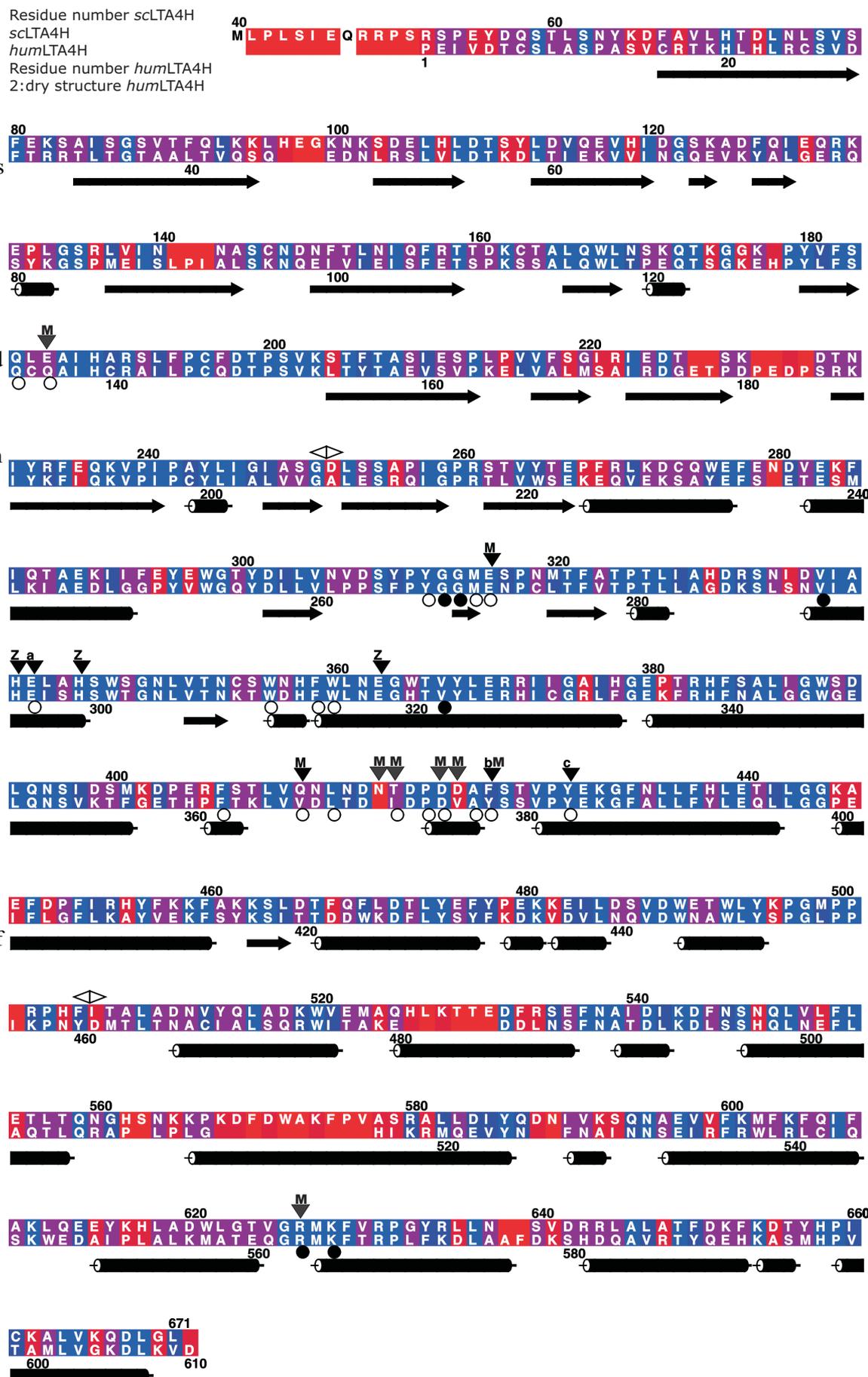
The degree of conservation is indicated by colour code with values according to Zvelebil et al. (28).

Blue represents the highest degree of conservation. Mutated residues are marked with an M. Letters 'a' and 'c' represent Glu-341 and Tyr-429,

the general base and proton donor in the peptidase reaction, whereas 'b' represents Phe-429 and 'Z' the zinc binding ligands. Residues lining the deeper part of the putative binding pocket are indicated with open circles while filled circles indicate residues close to the upper part of the pocket. Domain borders are marked with open horizontal

triangles. The secondary structure of *humLTA4H* is presented underneath the aligned sequences using cylinders and arrows to symbolize helices and β -strands, respectively.

The figure was generated with Alscript (29).



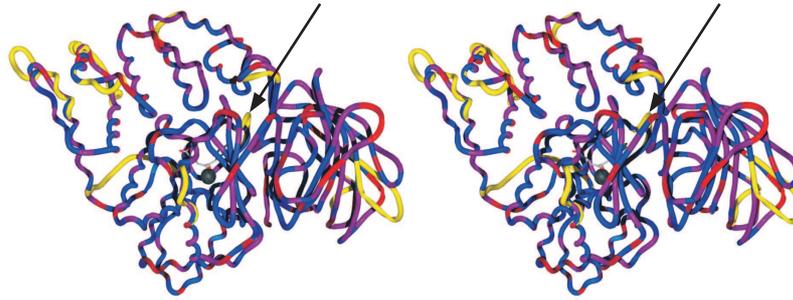


Figure S2. Stereo diagram of the modeled structure of *scLTA4H*.

Stereo diagram of the modeled regions of *scLTA4H* (before energy minimization) superimposed on its template *humLTA4H*. The picture depicts the C- α trace of LTA4H and modeled regions of *scLTA4H*. The model is coloured according to the same scheme as in S1. Inserted and truncated segments of the model *scLTA4H* are shown in yellow. The position of the loop located close to the active site, which is discussed in the text, is indicated with an arrow. The corresponding loop is visible in the close-up in figure 1. The figure was generated with Swiss PDB viewer (30) and povray (www.povray.org).

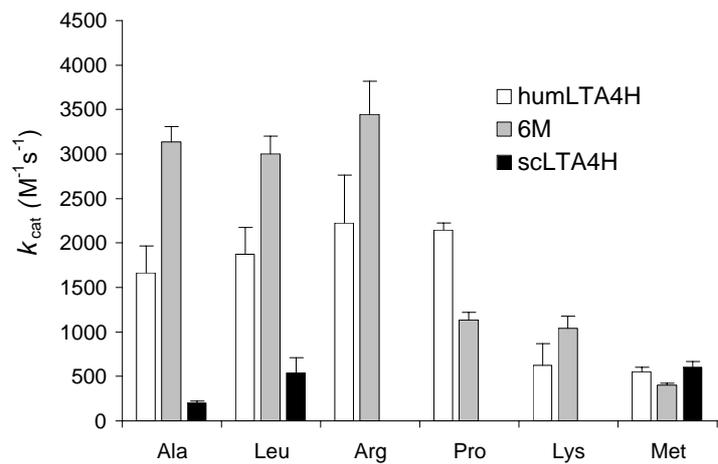


Figure S3. k_{cat}/K_m values of humLTA4H (white bars), 6M (gray bars) and scLTA4H (black bars) for hydrolysis of six different peptide substrates.

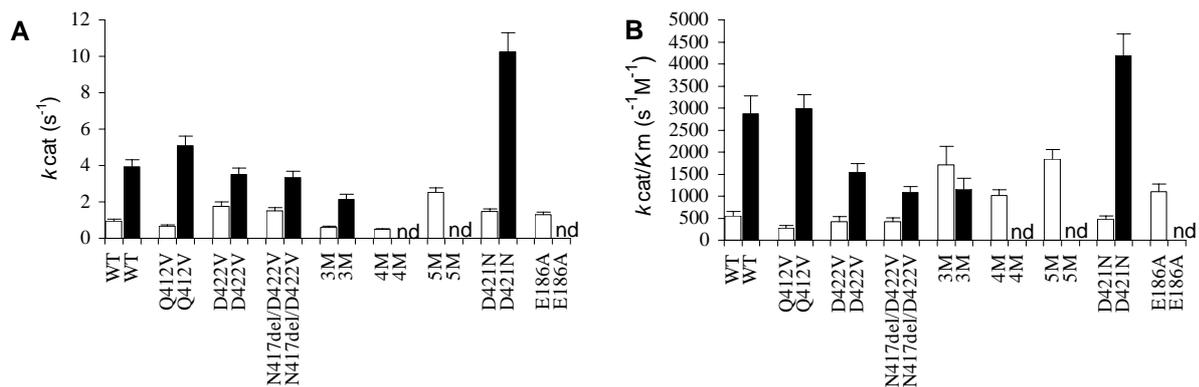


Figure S4. Effect of LTA_4 treatment on the kinetic parameters for peptide hydrolysis by wild-type and various mutants of *scLTA4H*. Apparent k_{cat} (panel A) and k_{cat}/K_m (panel B) values for Leu-*p*-NA hydrolysis by LTA_4 -treated (black bars) and untreated (white bars) enzymes. The error bars indicate standard errors and “nd” denotes “not determined”.

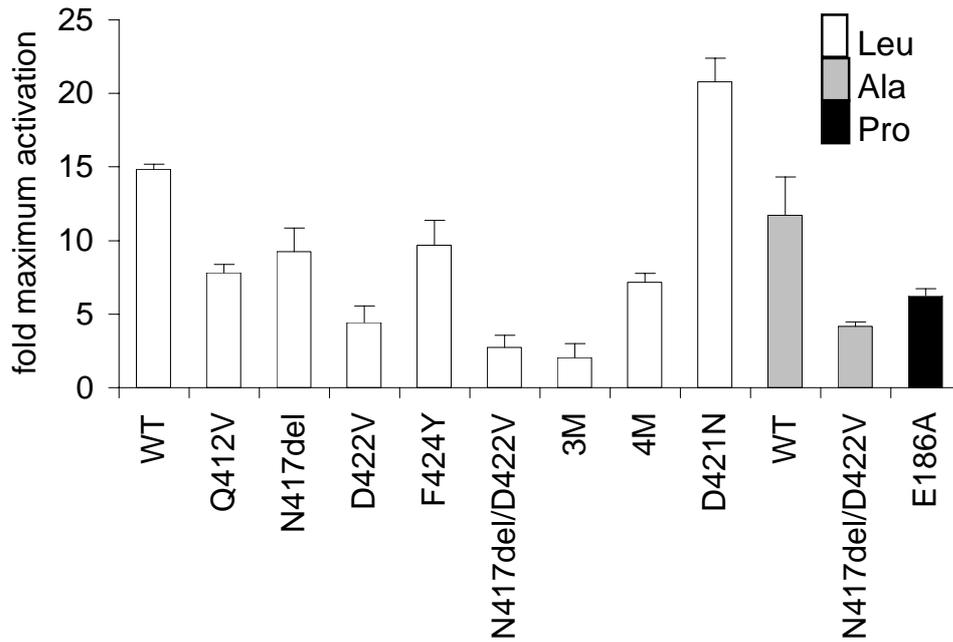


Figure S5. *LTA₄*-induced stimulation of the peptidase activity of wild type and mutated *scLTA4H*. The specific peptidase activity of enzymes susceptible to *LTA₄* stimulation was assayed with different peptide substrates and increasing amounts of *LTA₄*, until saturation was achieved. The maximum increase of the specific Leu-, Ala- and Pro-*p*-NA aminopeptidase activities are presented. The bars denote mean value \pm SE.

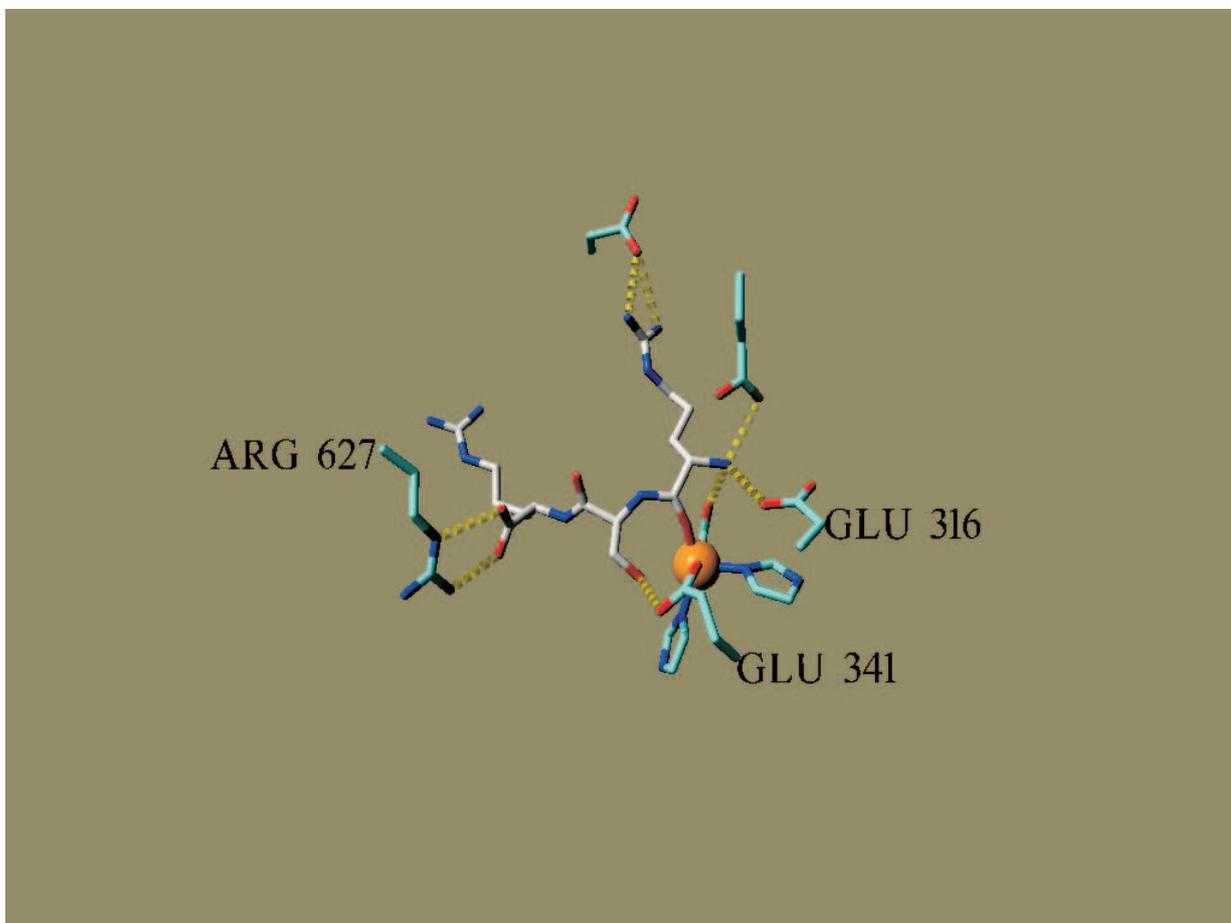


Figure S6. *Modeled binding conformation of an Arg-Ser-Arg peptide.* The figure presents certain key-interactions between the substrate (white carbons) and *scLTA4H* (turquoise carbons with the Zn-ion as an orange sphere). Only the functional groups of the Zn-binding ligands, i.e. His-340, His-344 and Glu-363, are presented. Arg-627 and Glu-316 together aligns, via binding of the C- and N-termini of the substrate, the scissile bond of the tripeptide substrate with the Zn ion and the general base catalyst Glu-341. In addition, Asp-421 and Glu-186 (both unlabeled in the figure) interacts with the guanido group and the free amine of the P1 residue, respectively. The binding conformation was obtained by docking using the Autodock3 software (33) followed by energy minimization of the lowest energy conformation using the Amber99 force field implementation of the YASARA software (34, 15). The figure was created with YASARA (www.yasara.org) and PovRay (www.povray.org).



Product formation controlled by substrate dynamics in leukotriene A₄ hydrolase



Alena Stsiapanava^a, Fredrik Tholander^a, Ramakrishnan B. Kumar^{a,1}, Abdul Aziz Qureshi^{a,2}, Damian Niegowski^a, Mahmudul Hasan^b, Marjolein Thunnissen^b, Jesper Z. Haeggström^a, Agnes Rinaldo-Matthis^{a,*}

^a Department of Medical Biochemistry and Biophysics, Scheeles väg 2, Karolinska Institutet, 17177 Stockholm, Sweden

^b Centre of Molecular Protein Science Lund University, Box 124, S-22100 Lund, Sweden

ARTICLE INFO

Article history:

Received 6 September 2013

Received in revised form 21 November 2013

Accepted 3 December 2013

Available online 11 December 2013

Keywords:

Leukotriene A₄ hydrolase

X-ray crystallography

Xenopus laevis

Small angle X-ray scattering

ABSTRACT

Leukotriene A₄ hydrolase/aminopeptidase (LTA₄H) (EC 3.3.2.6) is a bifunctional zinc metalloenzyme with both an epoxide hydrolase and an aminopeptidase activity. LTA₄H from the African claw toad, *Xenopus laevis* (xLTA₄H) has been shown to, unlike the human enzyme, convert LTA₄ to two enzymatic metabolites, LTB₄ and another biologically active product Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ (5(*S*),12*R*-dihydroxy-6,10-*trans*-8,14-*cis*-eicosatetraenoic acid). In order to study the molecular aspect of the formation of this product we have characterized the structure and function of xLTA₄H. We solved the structure of xLTA₄H to a resolution of 2.3 Å. It is a dimeric structure where each monomer has three domains with the active site in between the domains, similar as to the human structure. An important difference between the human and amphibian enzyme is the phenylalanine to tyrosine exchange at position 375. Our studies show that mutating F375 in xLTA₄H to tyrosine abolishes the formation of the LTB₄ isomeric product Δ^6 -*trans*- Δ^8 -*cis*-LTB₄. In an attempt to understand how one amino acid exchange leads to a new product profile as seen in the xLTA₄H, we performed a conformer analysis of the triene part of the substrate LTA₄. Our results show that the Boltzmann distribution of substrate conformers correlates with the observed distribution of products. We suggest that the observed difference in product profile between the human and the xLTA₄H arises from different level of discrimination between substrate LTA₄ conformers.

© 2013 Published by Elsevier B.V.

1. Introduction

The catalytic mechanism, as well as the 3D structure of the human LTA₄H (hLTA₄H) has been characterized using a combination of activity assays, site-directed mutagenesis and crystallography techniques [1,2]. As an epoxide hydrolase, the enzyme catalyzes the hydrolysis of

the allylic epoxide LTA₄ (5*S*-*trans*-5,6-oxido-7,9-*trans*-11,14-*cis*-eicosatetra-enoic acid) to form LTB₄ (5*S*,12*R*-dihydroxy-6,14-*cis*-8,10-*trans*-eicosatetraenoic acid), a potent chemotactic agent and a mediator of inflammation [3] (Fig. 1). In addition to the epoxide hydrolase activity, LTA₄H also possesses an anion-dependent aminopeptidase activity [4] that, in recent studies have shown to have a role in resolution of inflammation. LTA₄H has been reported to be a soluble cytosolic monomeric enzyme with a molecular mass of 69 kDa. It has been detected in almost all types of mammalian cells and characterized from different organisms including human, guinea pig, rat, mouse and the African claw toad, *Xenopus laevis* [5].

LTA₄H has two overlapping active sites, where the epoxide hydrolase and the peptidase activities take place. Both activities are dependent on a Zn atom that is coordinated by two His and one Glu residue from the Zn binding motif (HEXXH-X18-E) [6]. The Zn binding motif is a consensus sequence that is a signature of the M1 family of metalloproteases to which LTA₄H belongs [7].

A characteristic feature of the hLTA₄H is suicide inactivation, where LTA₄ covalently modifies a tyrosine at position 378 resulting in a loss of both the epoxide hydrolase and the aminopeptidase activity [8]. During

* Corresponding author at: Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden. Tel.: +46 852487612; fax: +46 8 7360439.

E-mail addresses: alena.stsiapanava@ki.se (A. Stsiapanava), Fredrik.Tholander@ki.se (F. Tholander), ramakrishnan.balakrishnan-kumar@ki.se (R.B. Kumar), aszizq@dbb.su.se (A.A. Qureshi), damian.niegowski@ki.se (D. Niegowski), Mahmudul.Hasan@biochemistry.lu.se (M. Hasan), Marjolein.Thunnissen@biochemistry.lu.se (M. Thunnissen), Jesper.Haeggstrom@ki.se (J.Z. Haeggström), Agnes.Rinaldo-Matthis@ki.se (A. Rinaldo-Matthis).

¹ Present address: Department of Biosciences and Nutrition, Karolinska Institutet School of Technology and Health, KTH—Royal Institute of Technology, Novum, SE-141 83 Huddinge, Sweden.

² Present address: Department of Biochemistry and Biophysics, Stockholm University, Svante Arrhenius väg 16C, 106 91 Stockholm, Sweden.

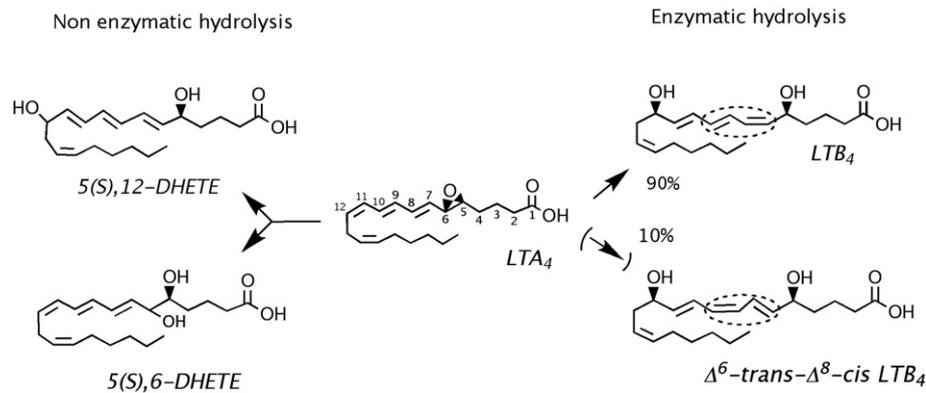


Fig. 1. Schematic drawing of enzymatic and non-enzymatic hydrolysis of LTA₄. LTA₄H from *Xenopus laevis* catalyzes the formation of LTB₄ as well as Δ^6 -trans- Δ^8 -cis-LTB₄. LTA₄ is spontaneously hydrolyzed in aqueous solution to 5(S),12-DHETE and 5(S),6-DHETE. The two dihedral bonds that are changed in the two enzymatic product isomers are highlighted with dashed lines. The parenthesis indicates an apparent formation of Δ^6 -trans- Δ^8 -cis-LTB₄ from LTA₄.

the study of suicide inactivation in hLTA₄H, Haeggström et al. discovered that Y378F mutant was able to form a novel isomeric LTB₄ product, later identified as Δ^6 -trans- Δ^8 -cis-LTB₄ [9].

Sequence analysis showed that mentioned tyrosine at position 375 is a phenylalanine in the xLTA₄H as well as in lower vertebrates species such as fish and insects. The analysis of the catalytic activity of the xLTA₄H purified from natural source showed a reduced sensitivity to suicide inactivation [10,11]. The same study showed that the xLTA₄H demonstrated an ability to convert LTA₄ into two products, LTB₄ and Δ^6 -trans- Δ^8 -cis-LTB₄ (5S,12R-dihydroxy-6,10-trans-8,14-cis-eicosatetraenoic acid), with a relative formation of the LTB₄ and the second product of 90 and 10%, respectively [10]. The new product was seen to be able to bind the BLT1 receptor and able to induce contractions in guinea pig lungs, however not as pronounced as LTB₄ [11].

In order to understand the molecular basis for the formation of the two isomeric LTB₄ products as seen in the *X. laevis* enzyme but not in the human enzyme, we have cloned, expressed and kinetically characterized the wild type and the F375Y mutant of xLTA₄H. XILTA₄H is 68% sequence identical to the human enzyme (Fig. 2). Steady state kinetics assays were used to determine the catalytic behavior as well as suicide inactivation properties. HPLC was employed to analyze the formation of Δ^6 -trans- Δ^8 -cis-LTB₄ and X-ray crystallography was used to study the structure of xLTA₄H. We used computational conformer

analysis to identify the equilibrium population of different LTA₄ conformers.

2. Materials and methods

2.1. Materials

All chemicals were obtained from common commercial sources. LTA₄ methyl ester (Cayman) in tetrahydrofuran was saponified with 1 M LiOH (6% (v/v)) for 48 h at 4 °C.

2.2. Cloning

At the time of cloning, the full-length sequence of xLTA₄H was not available but was derived from several shorter EST clone sequences. Different I.M.A.G.E. clones (<http://imageconsortium.org/>) [12] were screened by PCR and clone 6322848 (I.M.A.G.E. ID) was found to contain a cloning vector with a full length xLTA₄H insert. The identified plasmid was purified and a pair of LTA₄H specific primers was used to amplify the LTA₄H encoding sequence by PCR. The upstream primer encoded an N-terminal His-tag, a Sal I restriction enzyme cleavage sites, and an ATG start codon (upstream primer, 5'-GGT GGG TTA GGA AGC CGC ATG CAT CAT GAG GCA CTG ATA TGG-3') and the downstream primer

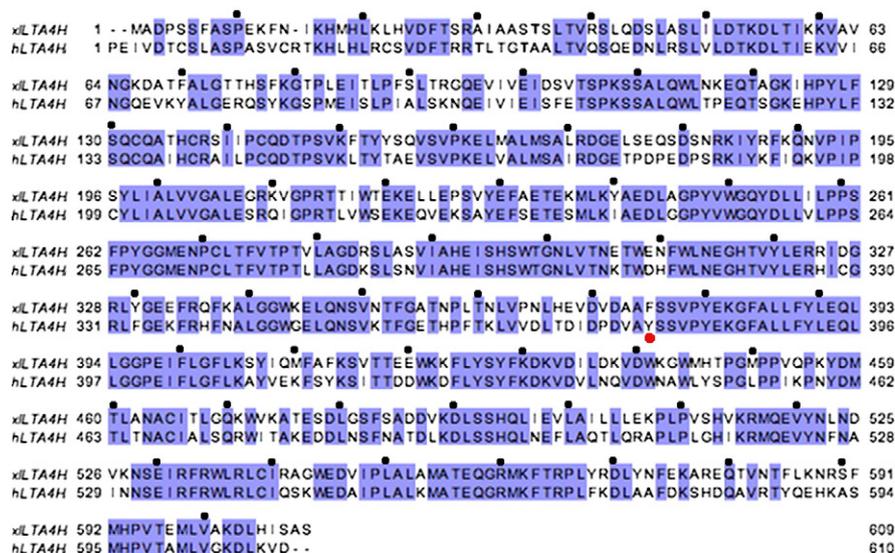


Fig. 2. Sequence alignment comparing xLTA₄H with hLTA₄H. The sequences are 68% sequence identical. The red circle indicates the position of the catalytically important Phe375 in xLTA₄H that is a Tyr in the human counterpart. Black dots are added every 10th amino acid to facilitate amino acid counting. The first amino acid of hLTA₄H is omitted.

encoded an Sph I restriction enzyme cleavage site and a stop codon (downstream primer, 5'-GGG GTC GAC ATG CAT CAC CAC CAT CAC CAT GCA GAT CCC AGC TCC TTT GC-3'). Obtained PCR products were ligated into the expression vector pT3 to yield an expression construct for xLTA4H (pT3-xLTA4H). The mutant xLTA4H F375Y mutant was prepared using site directed mutagenesis kit (QuikChange, Stratagene, La Jolla, CA). Both plasmid constructs were transformed into the competent *Escherichia coli* JM101 cells. The correctness of the full-length sequences was verified by DNA sequencing (SEQLAB, Göttingen, Germany).

2.3. Protein expression and purification

Both xLTA4H wild type and the mutant F375Y with an N-terminal His tag, were expressed in *E. coli* JM101 cells and purified to homogeneity by Ni-sepharose Fast Flow column (Amersham Biosciences) immobilized metal affinity chromatography (IMAC) column and ion exchange MonoQ column. The fractions containing the protein were desalted using a YM-30 centricon (Amicon Millipore) and eluted in 50 mM Tris-HCl pH 7.8. The protein concentration was determined using Bradford assay [13], and the purity of the protein was checked by SDS-PAGE on a PhastSystem (GE Healthcare) utilizing 10–15% gradient gels. Protein bands were visualized with Coomassie Brilliant Blue.

2.4. Activity measurements

For both the wild type and the mutant enzyme, the peptidase assay was performed as described previously [14]. In brief, 2 µg of LTA₄H was incubated with 150 µl of 1 mM substrate (alanine-4-nitroanilide, Ala-p-NA) in assay buffer (50 mM Tris-HCl pH 7.8 and 100 mM KCl) for 10 min at room temperature. 4-nitroaniline production was detected on a Multiscan spectrophotometer (Labsystems) at 405 nm. The enzyme activity was calculated using of extinction coefficient of 9500 M⁻¹ cm⁻¹ for 4-nitroaniline.

For the epoxide hydrolase activity, 2 µg of xLTA4H was incubated with 48 µM LTA₄ in 100 µl of 50 mM Tris-HCl pH 7.8 at room temperature for 15 s. Reaction was stopped with 2 volumes of methanol followed by 1 volume of water. 378 pmol prostaglandin (PG)B1 was used as internal standard. The samples were analyzed on HPLC with 3.9 × 150 mm column (Nova-Pak Waters, C₁₈). The column was eluted with a mobile phase of acetonitrile/methanol/water/acetic acid in ratio 30:35:35:0.1 (v/v) at flow rate 0.66 ml/min. Absorbance was monitored

at 270 nm (Fig. 3). To obtain Michaelis–Menten parameters for the two reactions, the same assays were used as described above. The K_M and V_{max} for the aminopeptidase reaction were determined by varying the substrate Ala-p-NA from 0.01 to 3 mM. To obtain Michaelis–Menten parameters for the epoxide hydrolase reaction, the LTA₄ substrate was varied from 0.1 to 80 µM (Table 1).

2.5. Suicide inactivation

The effects of repetitive additions of LTA₄ or LTA₄-methyl ester (ME), were studied for both the epoxide and the peptidase activities of the enzyme, essentially using the same methods as used in previous studies [15]. In brief, for epoxide hydrolase activity, 1 µg of enzyme was diluted in 10 µl 50 mM Tris-HCl pH 7.8 and treated repeatedly with 1–4 doses of 20 µM LTA₄-ME or 20 µM LTA₄. Both epoxides inhibit the enzyme equally well (data not shown). For the aminopeptidase activity, 1–2 µg of enzyme was diluted in 50 µl buffer and treated with of 1–4 doses of 20 µM LTA₄. The pretreated samples were then checked for both activities as described above. The experiments were done in duplicates for epoxide hydrolase activity measurements and in triplicates for aminopeptidase activities, respectively.

2.6. Crystallization of xLTA4H

For crystallization of the protein, the limited proteolysis technique was used. The purified xLTA4H enzyme (10 mg/ml) in 25 mM Tris-HCl pH 7.8, 20 mM NaCl, 1 mM TCEP, 2% (v/v) glycerol, was incubated with chymotrypsin for 16 h at 273 K in weight ratio of 4:1, xLTA4H: protease, respectively. The proteolytic digestion of xLTA4H was analyzed using SDS-PAGE and then purified using GE Superdex 200 10/300 to separate xLTA4H from chymotrypsin. The xLTA4H was eluted as one homogenous peak from the size exclusion chromatography and concentrated with Amicon Centricon concentrator (10 K MWCO) (Merck Millipore) to 4 mg/ml. The protein was incubated with 0.8 mM bestatin for 1 h at 273 K. Nascent crystallization conditions were found using JCSC + suite screen (QIAGEN) by sitting drop method at 295 K. The nascent hit (100 mM Bis-Tris pH 5.5, 25% (w/v) PEG 3350) was optimized and three crystals appeared after six months at 100 mM Bis-Tris pH 6.5, 20% PEG 3350. These crystals were cryoprotected using 25% (v/v) glycerol in reservoir solution. Cryoloops were used to mount the crystals and the mounted crystals were then vitrified by liquid nitrogen.

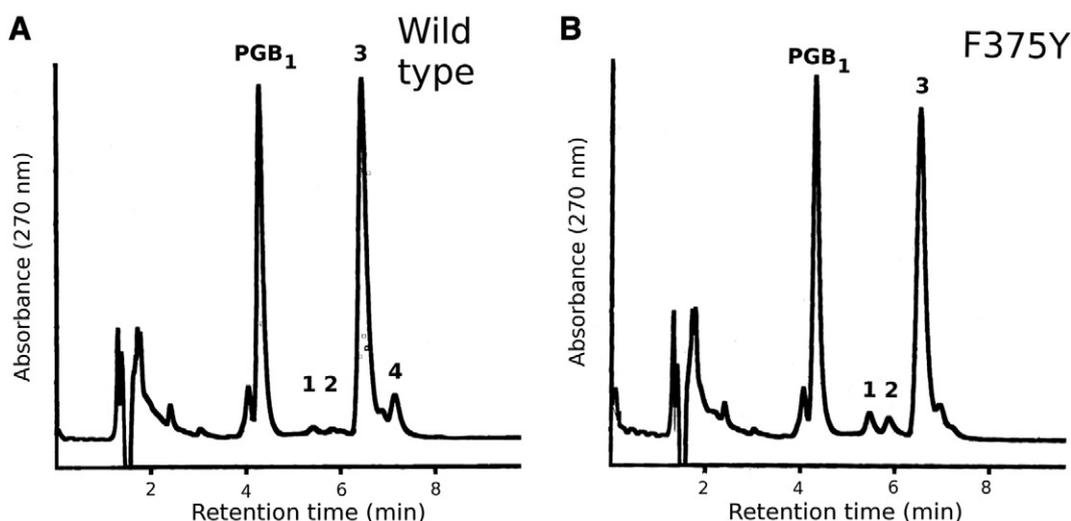


Fig. 3. HPLC analysis of products formed by wild type in (A) and F375Y mutant in (B) of xLTA4H incubated with LTA₄. Prostaglandin (PG) B1 was used as an internal standard. In (A), Peak 1 and 2 correspond to the two non-enzymatic hydrolysis products of LTA₄, the third peak eluted with synthetic LTB₄ whereas the fourth peak had the same HPLC retention time and UV spectrum as Δ^6 -trans- Δ^8 -cis-LTB₄. In (B) The mutant F375Y did not produce the peak corresponding to Δ^6 -trans- Δ^8 -cis-LTB₄.

Table 1
Steady state kinetic parameters for the epoxide hydrolase and aminopeptidase activity of LTA4H. We measured the activities of *Xenopus laevis* wild type and mutant F375Y, human wild-type enzyme and mutant Y378F, obtained at room temperature, pH 7.8 and with substrate incubations for 15 s.

Enzyme	Epoxide hydrolase			Aminopeptidase		
	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($\text{s}^{-1}/\mu\text{M}$)	k_{cat} (s^{-1})	K_M	k_{cat}/K_M (s^{-1}/mM)
X.L. LTA4H wt	20.1 ± 2	45 ± 2	0.45	4.3 ± 0.2	2.6 ± 0.5	1.7
X.L. LTA4H F375Y	3.1 ± 1.1	4.0 ± 2.2	0.8	3.44 ± 0.2	0.87 ± 0.06	3.96
Human LTA4H wt ^a	0.85	5.8	0.15	0.21	0.47	0.45
Human LTA4H (Y378F) ^a	2.5	23.1	0.11	0.19	3.30	0.06

^a [35].

2.7. Data collection and processing

Diffraction data from a monoclinic crystal was collected at beamline ID14-4 (ESRF, Grenoble, France) [16], equipped with Q315r ADSC CCD detector. 200 images with 1° oscillation angle were collected for 2.3 Å resolution data set. Data were indexed, integrated and scaled using *autoPROC* software package [17]. The data-collection statistics are presented in Table 2.

2.8. Structure solution and refinement

The initial phase of xLTA4H was determined by molecular replacement using *Molrep* with the hLTA4H structure, PDB ID: 3b7t without ligands and waters as a template structure. There were two monomers in the asymmetric unit (Fig. 4) and the solvent content of crystal was estimated to 42.7%. The structure was refined in *Buster* [18] and manual building of the model was conducted with *Coot* [19]. After rigid-body refinement and several cycles of restrained refinement, autoNCS was performed. Amino acids 326–331, 346–366, 474–489 residues in

monomer A, as well as N- and C-terminal residues, residues 105–141, 172–187, 199–212 and 571–587 in monomer B were disordered with just fragments of supporting electron density and could not be built into the primary model. Iterative process, consisting of building of lacking structure fragments in *Buccanier* [20] followed by restrained refinement of overall model in *Buster* allowed us to build both chains in the monomers in the final model. After every cycle, the geometry of overall model was inspected in validation application of *Coot* and at *Molprobrity* server. Poor rotamers, Ramachandran outliers and residues with bad angels were manually fixed in *Coot*. The introduction of TLS groups to restrained refinement resulted in lowering of *R*_{work}/*R*_{free} from 0.28/0.31 to 0.25/0.28. The further refinement resulted in final *R*_{work} = 0.22 and *R*_{free} = 0.25, see Table 2.

Initial *mFo-DFc* difference map in the active site of both monomers was interpreted as two bestatin molecules and two zinc ions that were included in the model. Final validation of the model quality with respect to the data was done in *Sfcheck* [21]. The geometric parameters of the model were checked using validation tools of *Coot* and the *MolProbrity* server [22]. The coordinates and structure factors were deposited in the Protein Data Bank (<http://www.rcsb.org>) under accession code 4gaa.

2.9. SAXS data collection and processing

For SAXS analysis, xLTA4H was further purified using size exclusion chromatography with 30 mM Tris (pH 7.8) and 100 mM NaCl in the running buffer. xLTA4H at 1.9 mg/ml concentration was incubated with 50 mM bestatin for at least half an hour before the data collection.

Data were collected at BioSAXS beamline BM29 (ESRF, Grenoble, France) with X-ray beam at wavelength $\lambda = 1.0$ Å, and the distance from the sample to detector (Pilatus 1 M, Dectris Ltd.) was 2.85 m, covering a scattering vector range ($q = 4\pi\sin\theta/\lambda$) from 0.004 to 0.45 Å⁻¹. All data were collected at 283 K in a sample cell of 1.8 mm diameter quartz capillary and 10 frames of two-dimensional images were recorded for each buffer or sample, with an exposure time of 2 second per frame. The 2D images were averaged and reduced to one-dimensional scattering profiles and the scattering of the buffer was subtracted from the sample profile using the software on site.

Initial data scaling along with the experimental radius of gyration (*R*_g) and the forward scattering intensity *I*(0) were calculated from data at low *q* values in the range of $qR_g < 1.3$, using the Guinier approximation: $\ln(I(2\theta)) \approx \ln(I(0)) - R_g^2 q^2/3$ using the software *Primus* [23]. Pair distribution function was calculated by *Gnom* [24]. Ab initio rigid body models were calculated using *Gasbor* [25]. The initial models were averaged using *Damaver* [26].

2.10. Conformational study of substrate LTA₄

For the analysis of LTA₄ conformers, LTA₄ was divided into sub-structures (Fig. 5). Only C4–C13 triene sub-structures that demonstrated changes in configuration in the product as compared with the substrate, were used in the analysis. The remaining part of LTA₄ has the character of a saturated alkane and its dihedrals are therefore flexible and likely to adapt a staggered conformation. The four triene

Table 2
Data collection and refinement statistics of LTA4H from *Xenopus laevis* in complex with bestatin.

Data collection	
Wavelength (Å)	0.939
Resolution range ^a (Å)	29.3–2.3 (2.268–2.261)
Unit-cell parameters (Å, °)	$a = 222.12$, $b = 52.17$, $c = 109.90$, $\alpha = \gamma = 90$, $\beta = 111.58$
Space group	C121
Measured reflections	217,021
Unique reflections	54,861
Completeness (%) ^a	99.1 (99.6)
$R_{\text{merge}}^{\text{a,b}}$	0.072 (0.531)
$\langle I/\sigma(I) \rangle^{\text{a}}$	12.7 (2.9)
Average multiplicity	4.0 (3.9)
Wilson B-factor	39.53
Refinement	
Reflections used in working set	3762
Reflections used in test set	194
Maximum resolution (Å)	2.26
$R_{\text{work}}/R_{\text{free}}^{\text{d}}$	0.22/0.25
No. of non-H atoms	9831
No. of protein atoms	9692
No. of ligand atoms, ions	46
No. of water molecules	93
Average B factor (Å ²)	89.94
R.m.s. deviations from ideal	
Bond lengths (Å)	0.009
Bond angles (°)	1.04
PDB ID	4gaa

^a Values for the highest resolution shell are given in parentheses.

^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity of this reflection.

^c $R = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|$.

^d R_{free} [36] was monitored with 5% of the reflection data excluded from refinement.

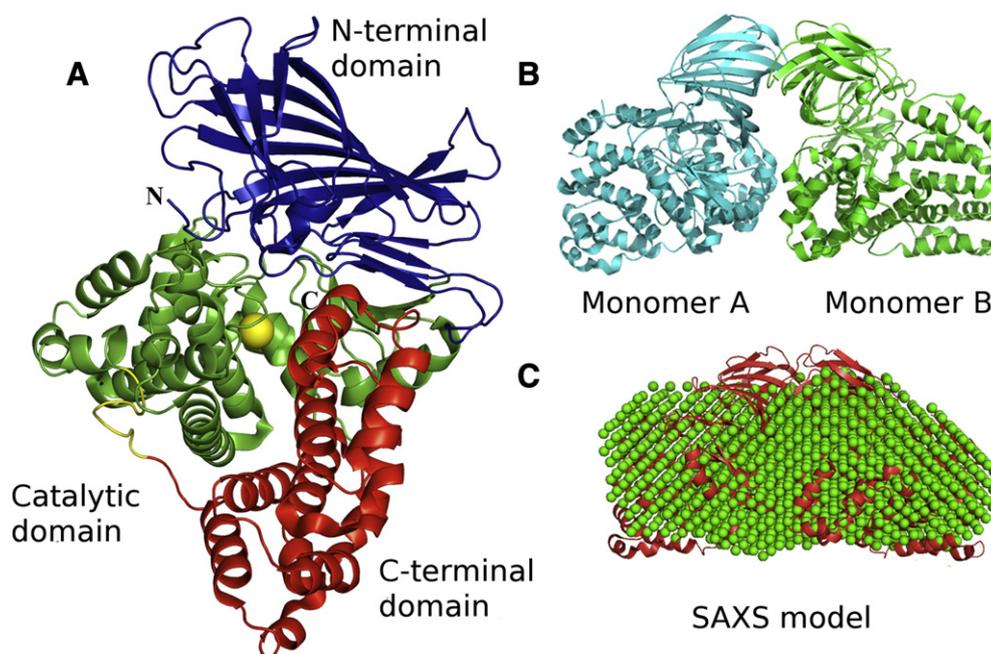


Fig. 4. In (A), the crystal structure of xLTA4H wild type enzyme, solved to 2.3 Å resolution. The overall structure has 3 domains and the active site in between the domains. In (B) the two molecules in the asymmetric unit and in (C) ab initio averaged model of xLTA4H (green) generated by GASBOR superimposed to cartoon crystal structure of xLTA4H (red).

conformers of LTA₄ C4–C13 were generated (i.e. the internal single bond dihedrals of the triene system were set to all combinations of *s-cis* (0°) and *s-trans* (180°), while the double bonds were kept in *trans-trans-cis* configuration) and subsequently energy minimized (Merck Molecular Force Field 94 with a Poisson–Boltzmann solvation model) using Szybki (www.eyesopen.com). The Boltzmann distribution for the conformers was calculated assuming that the four conformers were in equilibrium. After energy minimization, most examined dihedral angles remained close to either 0° (*cis*) or 180° (*trans*). However, the dihedrals of single bonds in the original *s-cis* configuration were distorted by approximately 45–55° during energy minimization, which was caused by steric repulsions between opposing protons. The derived force field (MM94) energies of the conformers were used to analyze their Boltzmann distribution (Fig. 5).

2.11. Molecular docking

AutoDock Vina [27] interfaced from Yasara Structure (www.yasara.com) was used for the docking. Yasara was used to prepare the structures for docking; hydrogens were added, the hydrogen bonding network optimized [28], pKa values assigned [29], the structures were energy minimized using different force fields [30] and water molecules removed. Docking was focused on the active site and only docking poses binding in a productive conformation were considered, i.e. with the proper functional group interacting with relevant catalytic residues. Only the two LTA₄ conformers that give rise to experimentally observed enzymatic products, i.e. *all-s-trans*-LTA₄ and *s-cis-s-trans*-LTA₄, were analyzed. During docking all dihedrals were allowed to rotate except the triene moiety, which was kept frozen to preserve each specific LTA₄ conformer.

3. Results and discussion

3.1. General properties of xLTA4H

LTA4H from *X. laevis* has been studied here to increase our understanding of the formation of the product Δ^6 -*trans*- Δ^8 -*cis*-LTB₄. This product is formed in addition to LTB₄ when the xLTA4H wild type enzyme but not the mutant F375Y, is incubated with LTA₄ (Fig. 1). We

have cloned, expressed and purified the wild type and mutant F375Y. The structure was solved for the wild type enzyme. The catalytic turnover of the xLTA4H wild type enzyme was approximately 20 times higher than for the human enzyme (Table 1).

3.2. Phe375 is involved in the formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄

To determine the role of F375 in the formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ in xLTA4H, we characterized the catalytic parameters of both the epoxide hydrolase and aminopeptidase activity. Steady-state kinetics of both the epoxide hydrolase and aminopeptidase reactions of the recombinant enzymes (wild type and F375Y), showed a higher K_M and k_{cat} for the wild type enzyme (Table 1) ($K_M^{LTA_4} = 45 \mu M$ and $K_M^{Ala-p-NA} = 2.6 \text{ mM}$ and $k_{cat}^{LTA_4} = 20.1$ and $k_{cat}^{Ala-p-NA} = 4.3 \text{ s}^{-1}$) as compared with the mutant F375Y ($K_M^{LTA_4} = 4 \mu M$ and $K_M^{Ala-p-NA} = 0.87 \text{ mM}$ and $k_{cat}^{LTA_4} = 3.1$ and $k_{cat}^{Ala-p-NA} = 3.4 \text{ s}^{-1}$). Both enzymes displayed a similar catalytic efficiency measured as (k_{cat}/K_M). The higher K_M in the wild-type enzyme as compared with the mutant form reflects a role of phenylalanine at position 375 in binding and recognition of the substrate. The higher k_{cat} in the wild type as compared with the mutant form, as seen for the epoxide hydrolase activity, suggests a more rapid turnover at the expense of substrate selectivity. These results agree with previously obtained results for the human enzyme, wild type and Y378F mutant. Furthermore, a reduced k_{cat} in the hLTA4H wild type enzyme as compared with the hLTA4H Y378F has been described as being due to suicide inactivation where LTA₄ becomes covalently attached to Y378 during the catalytic turnover [31].

3.3. Mutant F375Y is more sensitive to suicide inactivation than the wild type enzyme

Suicide inactivation is a well-characterized feature of hLTA4H and is thought to be an important mechanism for the overall regulation of the LTB₄ biosynthetic pathway in vivo. Studies of hLTA4H have shown that, during suicide inactivation, LTA₄ binds covalently with its C6 or C12 to the amino acid Y378 and inactivates the enzyme.

We measured suicide inactivation of the two enzymatic activities of xLTA4H. For the epoxide hydrolase activity, the wild type enzyme was inhibited by 20% when treated with LTA₄-ME as compared to an almost

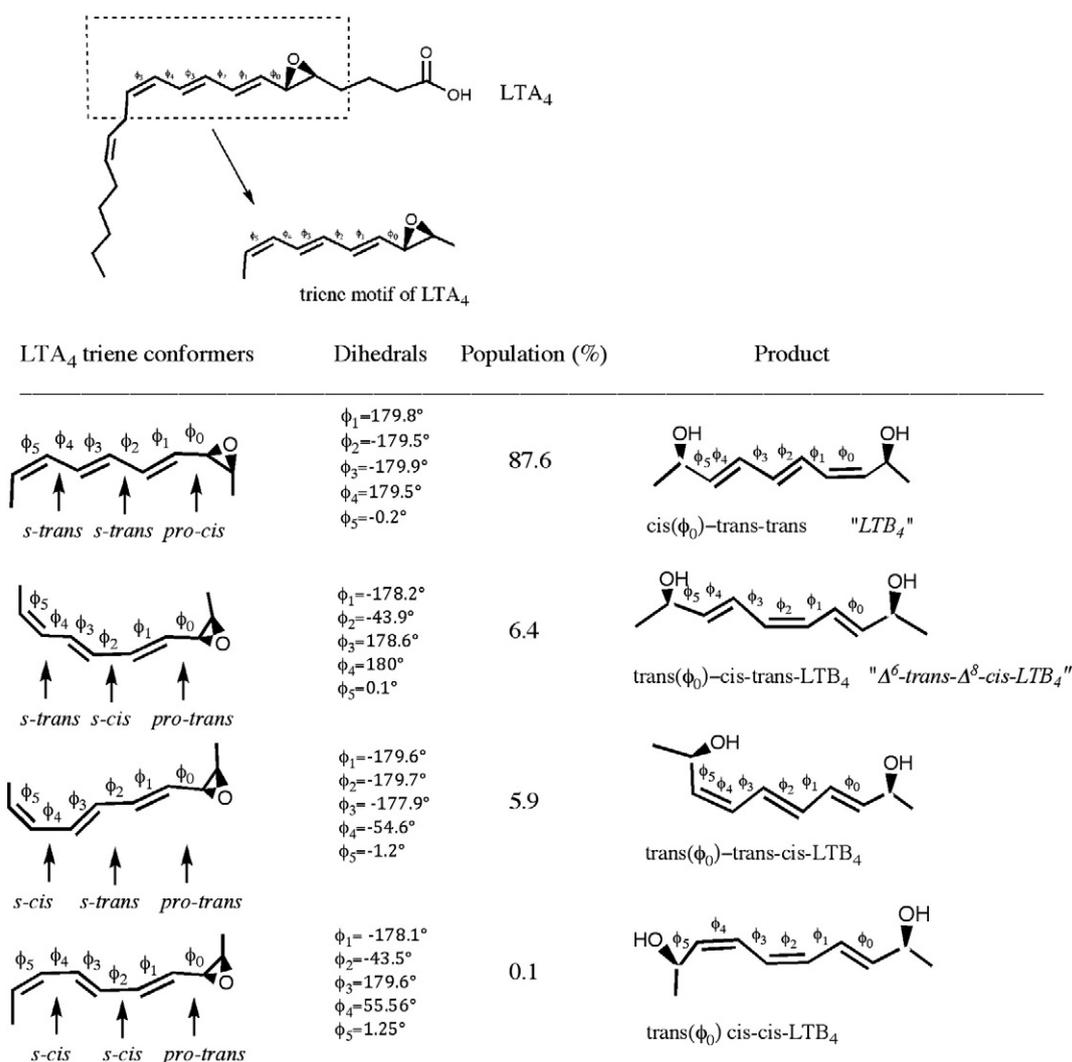


Fig. 5. Conformer analysis of the LTA₄ triene unit. Since most of the LTA₄ is flexible and will not change dihedral angles during catalysis, we focused on the substructure corresponding to the triene C4–C13 of LTA₄. The LTA₄ triene conformers that generate the products LTB₄ and Δ^6 -trans- Δ^8 -cis-LTB₄, are the most populated trienes present in a ratio of 87.6% and 6.4%, respectively.

80% loss of activity of the treated mutant F375Y. The aminopeptidase activity showed the same results as for epoxide reaction, where the wild type enzyme displayed a suicide inactivation of 15% whereas the activity of the mutant F375Y was more sensitive to LTA₄ treatment and showed a reduction in activity of 90%. The observed suicide inactivation, although low, in the wild-type xLTA4H can be explained by LTA₄ being covalently attached to other residues such as Tyr383. Studies by Murphy et al. have shown that LTA₃ (a double bond isomer of LTA₄ lacking the isolated C12–C13 double bond) is able to bind covalently to Tyr383 in hLTA4H [11].

3.4. Enzymatic formation of Δ^6 -trans- Δ^8 -cis-LTB₄

To analyze the product profile of xLTA4H and the mutant F375Y, we used reverse-phase HPLC analysis (Fig. 3). Enzyme samples of wild type xLTA4H, incubated with LTA₄, revealed four peaks, where peak 1 and 2 co-chromatographed with the two non-enzymatic hydrolysis products of LTA₄, the 5(S),6-DHETE and 5(S),12-DHETE. The third peak eluted with synthetic LTB₄ whereas the fourth peak had the same HPLC retention time and UV spectrum as Δ^6 -trans- Δ^8 -cis-LTB₄. The formation of Δ^6 -trans- Δ^8 -cis-LTB₄ appeared in a ratio 1:10 relative to LTB₄. The mutant F375Y did not produce the peak corresponding to Δ^6 -trans- Δ^8 -cis-LTB₄ (Fig. 3).

3.5. Substrate channel in crystal structure of xLTA4H

To get more insights into the formation of Δ^6 -trans- Δ^8 -cis-LTB₄ we determined the X-ray structure of xLTA4H to a resolution of 2.3 Å. The overall structure has 3 domains and the active site is between the domains (Fig. 4A). We used limited proteolysis to crystallize the protein [32] and the final model consists of residues 2–607 in monomer A and residues 2–606 in monomer B. One amino acid is missing in each N-terminal and two and three amino acids are missing in the C-terminal of monomer A and B, respectively. For statistics of data collection and structure refinement see Table 2. The xLTA4H was solved as a dimer in the asymmetric unit (Fig. 4B). The dimeric state of xLTA4H was also seen when analyzing the protein in solution as observed from SAXS analysis (Fig. 4C). However to further understand the physiologic relevance of xLTA4H being active as a dimer (human enzyme has been reported as monomeric), further investigations have to be performed.

The xLTA4H shares 68% sequence identity with the human enzyme and comparing the human structure with monomer A and B of the xLTA4H structure, an rmsd (C α comparisons) of 0.6 and 1.2 Å, respectively, is obtained. Calculating rmsd between the two monomers, a value of 0.3 Å is obtained.

Most amino acid differences between the two structures are situated on the surface and the active sites are similar having an L-shaped cavity

composed of hydrophobic residues, except for amino acids that are important for the hydrolase (Asp372), aminopeptidase (Glu293, Tyr380) and both activities (Glu268, Arg560, Lys562). Both the epoxide hydrolase and the aminopeptidase activities are dependent on a zinc ion that is coordinated by His292, His296 and Glu315.

The protein was co-crystallized with the protease inhibitor bestatin that was found in each active site of the two monomers.

The main difference in the active sites of the human and the amphibian enzymes is the tyrosine to a phenylalanine exchange at position 375, closely situated to Y380 and to where the LTA₄ triene is thought to bind. Y383 (Y380 in xLTA4H) is important for LTB₄ formation where it creates a hydrogen bond to Y378 in the human enzyme structure. Y378 has been suggested to have an important role to stabilize Y383 for efficient catalysis in the human enzyme. The xLTA4H with a Phe in position 375 is not able to form a hydrogen bond to Y380 and has a less stabilizing effect on the positioning of Y380.

The active site of LTA4H is narrow and for LTA₄ to form Δ^6 -*trans*- Δ^8 -*cis*-LTB₄, a large active site is needed. The formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ requires bond rotations involving large propeller-like movement of the LTA₄ chain and the steric hindrance imposed by the active site of the enzyme would prevent the Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ to form. In addition, the reaction intermediate has a carbocation delocalized over the triene system inducing a double bond character, which would further restrict the dihedral transition from *s-trans* to *s-cis* of the C6–C7 bond that would be required to form the proper product configuration. Thus, the possibility that LTA₄ changes conformation to the one compatible with formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ when bound to the active site is less likely. To seek another explanation, we explored the possibility that conformer analysis of LTA₄ would give us clues about how the Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ product is formed. The hypothesis being that an LTA₄ conformer, able to generate the isomeric product, would exist already in solution before entering the active site. In order to study this we did conformer analysis of LTA₄.

3.6. Conformer selectivity of substrate controls product formation in LTA4H

It is known that triene systems can adopt different conformations of their single bonds (i.e. *s-cis* or *s-trans*) and that these conformers thermally equilibrate [33]. We performed an in silico analysis of the equilibrium distribution of LTA₄ conformers. Since most of the LTA₄ is very flexible, we focused on the substructure corresponding to the triene C4–C13 of LTA₄, the part of the molecule that is changed during the catalytic reaction (Fig. 5).

The Boltzmann distribution of the energy-minimized LTA₄ conformers shows that the dominating conformation (with a population of 87.6%) has all internal single bond dihedrals of the triene in *s-trans* configuration. This corresponds to the conformation that gives rise to LTB₄. The second most populated conformer (6.4%) corresponds to the conformation that forms *trans-cis-trans*-LTB₄, which corresponds to the Δ^6 -*trans*- Δ^8 -*cis*-LTB₄. The two conformers with the lowest abundance, 5.9% and 0.1%, exhibited dihedral angles that would give rise to LTB₄ isomers that have not been observed experimentally. These LTA₄

conformers are probably not favorably bound to the enzyme and therefore no corresponding LTB₄ isomer can be formed.

The results of the computational study show that the two most populated conformers appeared approximately at the same ratio, 87.6% and 6.4%, as have been observed for the formation of the products, LTB₄ and Δ^6 -*trans*- Δ^8 -*cis*-LTB₄, formed at ratio of 90 and 10%. This indicates that the substrate LTA₄ conformer distribution might determine the product outcome.

Analyzing HPLC traces of non-enzymatic LTA₄ hydrolysis products [34] peaks additional to peak 1 and 2 can be observed, which could correspond conformers obtained from non-enzymatic hydrolysis. However further experiments need to be done, such as studying the effect of conformer distribution after changing the conformer equilibrium, to ascertain that the observed peaks correspond to LTA₄ conformers.

Taken together, the narrow binding pocket of LTA4H that hinders bond rotations of the triene system, suggests that the final double bond configuration of LTB₄ is dictated by the configuration that LTA₄ has, when it is first bound to the active site.

The most likely explanation for the formation of the Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ side product in xLTA4H, and in hLTA4H mutant Y378F, is then that the conformer, giving rise to this product, exists already in solution, or forms during the entrance to the active site. LTA₄ and LTA₄ conformer compatible with the formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄ have small difference in shape and charge and they occupy similar positions in space (Fig. 6A). However, LTA4H favors the “proper” substrate conformer for turnover, and an F in position 375 is somehow less restrictive in discrimination between substrate conformers allowing the enzyme to bind two different LTA₄ and we see two products formed (Fig. 6B). Mutating F375 to a tyrosine increases the selection and only one substrate conformer binds in a productive manner as in the human enzyme (Fig. 6C). In an evolutionary perspective, the human enzyme has adapted to its environment, not by increasing its catalytic efficiency (the amphibian and the human k_{cat}/K_M are similar), but instead by increasing its selectivity of substrate conformer. There are many examples of enzymes that have evolved by having a better product outcome induced by a change in catalysis, but xLTA4H, seems to be an enzyme that provides product diversity by controlling substrate conformer selectivity.

Abbreviations

LTA ₄	leukotriene A ₄ , 5(<i>S</i>)- <i>trans</i> -5,6-oxido-7,9- <i>trans</i> -11,14- <i>cis</i> -eicosatetraenoic acid
LTB ₄	leukotriene B ₄ , 5(<i>S</i>),12 <i>R</i> -dihydroxy-6,14- <i>cis</i> -8,10- <i>trans</i> -eicosatetraenoic acid
Δ^6 - <i>trans</i> - Δ^8 - <i>cis</i> -LTB ₄	5(<i>S</i>),12 <i>R</i> -dihydroxy-6,10- <i>trans</i> -8,14- <i>cis</i> -eicosatetraenoic acid
LTA ₄ -ME	LTA ₄ methyl ester
LTA4H	leukotriene A4 hydrolase
xLTA4H	<i>Xenopus laevis</i> LTA4H
SAXS	small-angle X-ray scattering
Ala-p-N	alanine-4-nitroanilide

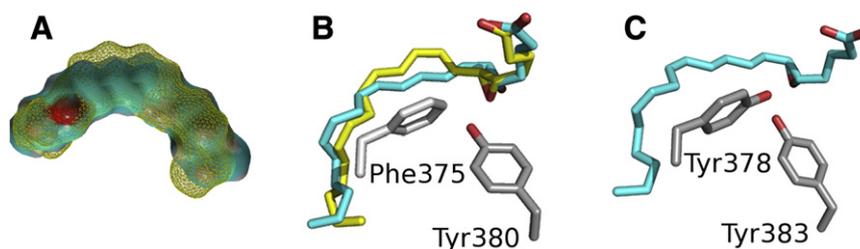


Fig. 6. In (A), a molecular surface and charge calculations of LTA₄ (cyan), and the LTA₄ conformer (in yellow), compatible with formation of Δ^6 -*trans*- Δ^8 -*cis*-LTB₄. The differences in shape and charge are small and they occupy similar positions in space. In (B) stick models of the same two molecules where LTA₄ in cyan and LTA₄ conformer, in yellow, are docked into the binding pocket of xLTA4H with a Phe in position 375 and in (C) the LTA₄ is docked in the active site of hLTA4H that has a Tyr in position 378.

Acknowledgements

The authors thank the Michaela Mårback for the technical assistance, the personnel at the beamlines of ID14-4 and BM29 of the ESRF synchrotron facility for the assistance with data collection as well as Martin Moche at the Protein Science Facility unit at the Dept. of Medical Biochemistry and Biophysics (Karolinska Institutet) for the assistance with structure refinement. This study was supported by grants from the Alfred Österlunds stiftelse (MT and MH), the Swedish Research Council (A.R.M. with grant 2301373 and 2301393 and J.Z.H. with grant 10350), Vinnova, CIDaT and a Distinguished Professor Award from the Karolinska Institutet.

References

- [1] M.M. Thunnissen, P. Nordlund, J.Z. Haeggstrom, Crystal structure of human leukotriene A (4) hydrolase, a bifunctional enzyme in inflammation, *Nat. Struct. Biol.* 8 (2001) 131–135.
- [2] A. Rinaldo-Matthis, J.Z. Haeggstrom, Structures and mechanisms of enzymes in the leukotriene cascade, *Biochimie* 92 (2010) 676–681.
- [3] B. Samuelsson, Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation, *Science* 220 (1983) 568–575.
- [4] M. Minami, N. Ohishi, H. Mutoh, T. Izumi, H. Bito, H. Wada, Y. Seyama, H. Toh, T. Shimizu, Leukotriene A4 hydrolase is a zinc-containing aminopeptidase, *Biochem. Biophys. Res. Commun.* 173 (1990) 620–626.
- [5] J.Z. Haeggstrom, Structure, function, and regulation of leukotriene A4 hydrolase, *Am. J. Respir. Crit. Care Med.* 161 (2000) S25–S31.
- [6] B.L. Vallee, D.S. Auld, Zinc coordination, function, and structure of zinc enzymes and other proteins, *Biochemistry* 29 (1990) 5647–5659.
- [7] J.Z. Haeggstrom, A. Wetterholm, B.L. Vallee, B. Samuelsson, Leukotriene A4 hydrolase: an epoxide hydrolase with peptidase activity, *Biochem. Biophys. Res. Commun.* 173 (1990) 431–437.
- [8] L. Orning, J. Gierse, K. Duffin, G. Bild, G. Krivi, F.A. Fitzpatrick, Mechanism-based inactivation of leukotriene A4 hydrolase/aminopeptidase by leukotriene A4. Mass spectrometric and kinetic characterization, *J. Biol. Chem.* 267 (1992) 22733–22739.
- [9] M.J. Mueller, M. Blomster, U.C. Oppermann, H. Jornvall, B. Samuelsson, J.Z. Haeggstrom, Leukotriene A4 hydrolase: protection from mechanism-based inactivation by mutation of tyrosine-378, *Proc. Natl. Acad. Sci. U. S. A.* 93 (1996) 5931–5935.
- [10] F. Stromberg-Kull, J.Z. Haeggstrom, Purification and characterization of leukotriene A4 hydrolase from *Xenopus laevis* oocytes, *FEBS Lett.* 433 (1998) 219–222.
- [11] F. Stromberg, M. Hamberg, U. Rosenzvist, S.E. Dahlen, J.Z. Haeggstrom, *Eur. J. Biochem.* 238 (1996) 599–605.
- [12] G. Lennon, C. Auffray, M. Polymeropoulos, M.B. Soares, The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression, *Genomics* 33 (1996) 151–152.
- [13] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [14] A. Wetterholm, J.F. Medina, O. Radmark, R. Shapiro, J.Z. Haeggstrom, B.L. Vallee, B. Samuelsson, Recombinant mouse leukotriene A4 hydrolase: a zinc metalloenzyme with dual enzymatic activities, *Biochim. Biophys. Acta* 1080 (1991) 96–102.
- [15] M.J. Mueller, M. Andberg, J.Z. Haeggstrom, Analysis of the molecular mechanism of substrate-mediated inactivation of leukotriene A4 hydrolase, *J. Biol. Chem.* 273 (1998) 11570–11575.
- [16] A.A. McCarthy, S. Brockhauser, D. Nurizzo, P. Theveneau, T. Mairs, D. Spruce, M. Guijarro, M. Lesourd, R.B. Ravelli, S. McSweeney, A decade of user operation on the macromolecular crystallography MAD beamline ID14-4 at the ESRF, *J. Synchrotron Radiat.* 16 (2009) 803–812.
- [17] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, Data processing and analysis with the autoPROC toolbox, *Acta Crystallogr. D* 67 (2011) 9.
- [18] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O.S. Smart, C. Vonrhein, T.O. Womack, BUSTERversion X.Y.Z, Global Phasing Ltd., Cambridge, United Kingdom, 2011.
- [19] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 486–501.
- [20] K. Cowtan, Fitting molecular fragments into electron density, *Acta Crystallogr. D Biol. Crystallogr.* 64 (2008) 83–89.
- [21] A.A. Vaguine, J. Richelle, S.J. Wodak, SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model, *Acta Crystallogr. D Biol. Crystallogr.* 55 (1999) 191–205.
- [22] V.B. Chen, W.B. Arendall, J.J. Headd III, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 12–21.
- [23] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *J. Appl. Crystallogr.* 36 (2003) 1277–1282.
- [24] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr.* 25 (1992) 495–503.
- [25] D.I. Svergun, M.V. Petoukhov, M.H.J. Koch, Determination of domain structure of proteins from X-ray solution scattering, *Biophys. J.* 80 (2001) 2946–2953.
- [26] V.V. Volkov, D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering, *J. Appl. Crystallogr.* 36 (2003) 860–864.
- [27] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455–461.
- [28] E. Krieger, R.L. Dunbrack Jr., R.W. Hooft, B. Krieger, Assignment of protonation states in proteins and ligands: combining pKa prediction with hydrogen bonding network optimization, *Methods Mol. Biol.* 819 (2012) 16.
- [29] E. Krieger, J.E. Nielsen, C.A. Spronk, G. Vriend, Fast empirical pKa prediction by Ewald summation, *J. Mol. Graph. Model.* 25 (2006) 481–486.
- [30] E. Krieger, G. Koraimann, G. Vriend, Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field, *Proteins* 47 (2002) 393–402.
- [31] J. McGee, F. Fitzpatrick, Enzymatic hydration of leukotriene A4. Purification and characterization of a novel epoxide hydrolase from human erythrocytes, *J. Biol. Chem.* 260 (1985) 12832–12837.
- [32] A. Dong, X. Xu, A.M. Edwards, C. Chang, M. Chruszcz, M. Cuff, M. Cymborowski, R. Di Leo, O. Egorova, E. Evdokimova, E. Filippova, J. Gu, J. Guthrie, A. Ignatchenko, A. Joachimiak, N. Klostermann, Y. Kim, Y. Korniyenko, W. Minor, Q. Que, A. Savchenko, T. Skarina, K. Tan, A. Yakunin, A. Yee, V. Yim, R. Zhang, H. Zheng, M. Akutsu, C. Arrowsmith, G.V. Avvakumov, A. Bochkarev, L. Dahlgren, S. Dhe-Paganon, S. Dimov, L. Dombrovski, P. Finerty Jr., S. Flodin, A. Flores, S. Graslund, M. Hammerstrom, M.D. Herman, B.S. Hong, R. Hui, I. Johansson, Y. Liu, M. Nilsson, L. Nedyalkova, P. Nordlund, T. Nyman, J. Min, H. Ouyang, H.W. Park, C. Qi, W. Rabeh, L. Shen, Y. Shen, D. Sukumard, W. Tempel, Y. Tong, L. Tresagues, M. Vedadi, J.R. Walker, J. Weigelt, M. Welin, H. Wu, T. Xiao, H. Zeng, H. Zhu, In situ proteolysis for protein crystallization and structure determination, *Nat. Methods* 4 (2007) 1019–1021.
- [33] A.M. Turek, J. Saltiel, T.R. Krishna, G. Krishnamoorthy, Resolution of conformer-specific all-trans-1,6-diphenyl-1,3,5-hexatriene UV absorption spectra, *J. Phys. Chem. A* 116 (2012) 5353–5367.
- [34] J.Z. Haeggstrom, A. Wetterholm, M. Hamberg, J. Meijer, R. Zipkin, O. Radmark, Enzymatic formation of 5,6-dihydroxy-7,9,11,14-eicosatetraenoic acid: kinetics of the reaction and stereochemistry of the product, *Biochim. Biophys. Acta* 958 (1988) 469–476.
- [35] M.J. Mueller, M.B. Andberg, B. Samuelsson, J.Z. Haeggstrom, Leukotriene A4 hydrolase, mutation of tyrosine 378 allows conversion of leukotriene A4 into an isomer of leukotriene B4, *J. Biol. Chem.* 271 (1996) 24345–24348.
- [36] A.T. Brunger, Assessment of phase accuracy by cross validation: the free R value. Methods and applications, *Acta Crystallogr. D Biol. Crystallogr.* 49 (1993) 24–36.

SAXS studies on leukotriene A4 hydrolases reveal conformational differences upon inhibitor binding

Mahmudul Hasan¹, Agnes Rinaldo-Matthis², Jesper Z. Haeggström² and Marjolein Thunnissen¹

¹Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Getingevägen 60, Lund University, 22100 Lund, Sweden.

²Department of Medical Biochemistry and Biophysics, Scheeles väg 2, Karolinska Institutet, 17177 Stockholm, Sweden

ABSTRACT

Vertebrate leukotriene A₄ hydrolases are zinc metalloenzymes with an epoxide hydrolase and aminopeptidase activity belonging to the M1 family of aminopeptidases. Bestatin, an amino peptidase inhibitor, can inhibit both activities. The human enzyme produces LTB₄, a powerful mediator of inflammation and is implicated in a wide variety of rheumatoid diseases. The yeast homolog *scLTA₄H* contains only a rudimentary epoxide hydrolase activity. Both the structure of the human enzyme and recently the structure of *scLTA₄H* have been solved to investigate the molecular architecture of their active sites both with and without inhibitor bestatin.

The structure of native *scLTA₄H* shows an open active site. Upon inhibitor binding, a domain shifts occur and the final binding pocket for the aminopeptidase substrate is formed. In the human enzyme the LTA₄ binding site is a narrow preformed hydrophobic channel, which protects the labile substrate when bound to the enzyme and no indications of induced fit have been observed. Many members of the M1 family however seem to display a certain degree of induced fit, a feature, which however, has never been observed for *humLTA₄H*. In order to investigate whether the behavior of the human enzyme is due to crystal-system restrictions, small angle X-ray scattering studies have been performed on three members of the family, *humLTA₄H*, *scLTA₄H* and *Xenopus* LTA₄H. These studies seem to suggest that *humLTA₄H* functions according to a lock-and-key mechanism while the other two members display conformational changes, more fitting with an induced fit mechanism.

Keywords: leukotriene A₄ hydrolase, SAXS, Domain movement

1. INTRODUCTION

The human zinc metalloenzyme leukotriene A₄ (*humLTA*₄) hydrolase (LTA4H) catalyzes the hydrolysis of the unstable epoxide LTA₄ (5*S*-*trans*-5,6-oxido-7,9-*trans*-11,14-*cis* eicosatetraenoic acid) into LTB₄ (5*S*,12*R*-dihydroxy-6,14-*cis*-8,10-*trans*-eicosatetraenoic acid). LTB₄ is a powerful mediator of inflammation and present in many autoimmune diseases. The *humLTA*4H has been studied extensively and the key elements for the catalytic activities have been identified by a combination of site directed mutagenesis and crystallography techniques [1,2].

From sequence comparisons it can be seen that related proteins of *humLTA*4H can be found in many organisms from bacteria to humans. All members of this family of LTA4H related proteins exhibit the aminopeptidase activity, however the second property (i.e. the epoxide hydrolase activity) is not conserved despite considerable sequence conservation. This activity can only be detected among vertebrates, including birds, frogs and fish [2]. Thus highly homologous enzymes from lower organisms, for example the enzyme from yeast, lack the epoxide hydrolase activity functions [3]. LTA4H from *S. cerevisiae* (*scLTA*4H) contains a rudimentary epoxide hydrolase activity and its structure has recently been solved to investigate the molecular architecture of the active site both with and without the aminopeptidase inhibitor bestatin in order to obtain information on the reasons why this protein lacks the epoxide hydrolase activity as all required active site aminoacids are present in the sequence.

Upon solving both the apo and inhibitor bound structure of *scLTA*4H, we discovered that the apo structure makes a large domain movement (Figure 1) compared with the human enzyme, which creates an open active site. This is in contrast with the human enzyme where the substrate-binding pocket contains a narrow hydrophobic channel, which is suggested to protect the labile substrate LTA₄ upon binding. The structure of the complex of *scLTA*4H with the inhibitor bestatin shows that a rearrangement of the structure and the active site closes upon binding of the inhibitor. This points to an induced fit mechanism where the full active site is only formed upon binding of the substrate. The lack of leukotriene A₄ hydrolase activity within the yeast homologue seems to be an obstruction of the protected cavity which is no longer present, while other parts of the binding pocket have become more accessible [4].

The fact that *scLTA4H* displays this type of induced fit mechanism is an interesting observation, which requires further studies. Many members of the M1 family of metalloproteases of which the LTAH proteins are a member, such as Aminopeptidase N [5,6] or Tricorn protease [7], seem to display a certain degree of induced fit. This feature has however never been observed for *humLTA4H*, despite a number of structures of different mutants, both in apo form and in complex with many different inhibitors. It is attractive to speculate that a preformed pocket functioning according to a lock and key mechanism is better suited to supply the protective and precise environment required for the precise hydrolysis of LTA4 into LTB4.

However as the human enzyme is always crystallized in the presence of YbCl_3 [8] which is required as it is involved in several crystal contacts that are essential for crystal formation, the lack of an observation of conformational changes could be an artifact. It could be that the crystal form itself locks the protein in a specific state and this prohibits the observation of conformational changes. Thus in order to investigate the mechanism of these proteins further, studies in solution were performed, as these are independent from possible artifacts caused by crystal contacts. Therefore small angle scattering studies were performed in order to study possible larger conformation changes occurring when inhibitor binding occurs.

Three proteins from the LTA4H family of enzymes were used in this study, *humLTA4H*, *scLTA4H* and the LTA4H from *Xenopus leavis* (*xLTA4H*). The last enzyme displays both the aminopeptidase and the leukotriene hydrolase activities, albeit without the product specificity that *humLTA4H* possesses [9]. Recently the structure of *xLTA4H* was solved [10] and it seems that the reduced product specificity can be correlated to a heightened substrate dynamics within the binding pocket of LTA4. In addition the structure showed, that in contrast with other LTA4H related proteins, *xLTA4H* forms a dimer in the asymmetric unit, and subsequent SAXS studies showed that the enzyme also behaves as a dimer in solution. Since the crystal structure of *xLTA4H* was only obtained in the presence of bestatin and no native structure is available, this protein was included in the present SAXS studies.

For *humLTA4H*, both bestatin and SC-57461A were used as inhibitors. SC-57461A is a mimic of LTA4 and one of the most powerful inhibitors of *humLTA4H* [11]. It is

supposed to fill the complete LTA4 pocket, while bestatin only partially occupies this pocket [12].

2. MATERIALS AND METHODS

2.1. Expression and purification of *hum*LTA4H, *sc*LTA4H and *x*LTA4H

Previously established protocols were followed to purify *hum*LTA4H [13], *sc*LTA4H [4] and *x*LTA4H [10]. For SAXS studies, *hum*LTA4H was further purified by injecting the protein in a Superdex 200 10/300 (GE Healthcare) size exclusion chromatography column with filtered and degassed running buffer containing 30 mM Tris-HCl (pH 7.9) and 100 mM NaCl with a flow rate of 0.3 ml/min. Fractions were pooled and concentrated using Amicon Ultra Centrifugal Filter (Millipore) with cut off value of 10 kDa.

The monodispersity of the protein samples was measured by dynamic light scattering using a Zetasizer Nano-ZS (Malvern) instrument and data were analyzed using software from the manufacturer. Protein concentrations were determined by A_{280} measurements using a NanoDrop spectrophotometer (Thermo Scientific) and specific molar extinction coefficients for the different proteins (*hum*LTA4H $104,905 \text{ M}^{-1}\text{cm}^{-1}$, *sc*LTA4H $104,655 \text{ M}^{-1}\text{cm}^{-1}$ and *x*LTA4H $103,165 \text{ M}^{-1}\text{cm}^{-1}$) calculated in ProtParam [14].

2.2. SAXS data collection and processing

For SAXS analysis all protein data were collected with concentrations between 0.5-10 mg/ml. Data were collected at the BioSAXS beamline BM29 at ESRF, Grenoble, France at a wavelength $\lambda = 1.0 \text{ \AA}$, and the distance from the sample to detector (Pilatus 1 M, Dectris Ltd.) was 2.85 m, covering a scattering vector range ($q = 4\pi\sin\theta/\lambda$) from 0.004 to 0.45 \AA^{-1} . All data were collected at 283 K in a sample cell of 1.8 mm diameter quartz capillary and 10 frames of two-dimensional images were recorded for each buffer or sample, with an exposure time of 2 second per frame.

At the MAX IV Laboratory, data were collected with an exposure time of 300 second at station I911-4. For all experiments, the used wavelength was 0.91 \AA and the beam size

was 0.3x0.3 mm². All samples were placed in a quartz capillary maintaining the temperature at 10° C, with a distance of 2 m between sample and the detector (a two-dimensional MAR165 CCD) thus covering a scattering vector range ($q=4\pi\sin\theta/\lambda$) from 0.01 to 0.325 Å⁻¹. No-beam, empty cell, water and buffer (before and after protein sample) scattering data were collected. The no-beam, empty cell and buffer data were subtracted from protein scattering data and the obtained 2D images were reduced to one-dimensional scattering profiles using software Bli711 [15]. The 2D images obtained from ESRF were averaged and reduced to one-dimensional scattering profiles and the scattering of the buffer was subtracted from the sample profile using the software on site [16].

Initial data scaling along with the experimental radius of gyration (R_g) and the forward scattering intensity $I(0)$ were calculated from data at low q values in the range of $qR_g < 1.3$, using the Guinier approximation: $\ln I(q) \approx \ln I(0) - R_g^2 q^2/3$ using the software Primus [17].

2.3. Scattering curve prediction by Crysol

To generate predicted scattering profiles for the crystal structure of the known proteins, PDB files of apo *scLTA4H* (2xpz) and inhibitor bestatin bound (2xq0) were downloaded from www.rcsb.org. PDB file 1hs6 for bestatin bound *humLTA4H*, 3u9w for SC-57461A bound and 4gaa for bestatin bound *xLTA4H* were also downloaded. The same PDB files were used to compare with their respective experimental scattering curve. Version 2.7 [18] was used in all calculations with parameters of maximum order of harmonics 50, order of Fibonacci grid 18 and electron density of the solvent 0.334 e/Å³. The same values of decision-making parameters were used to compare the scattering curve of ligand bound *humLTA4H* and *xLTA4H* to their respective crystal structure.

2.4. Pair distance distribution function calculation

For pair distance distribution function calculations, the whole scattering curve were used and were calculated by the Fourier inversion of the scattering intensity $I(q)$ using the program Gnom [19]. All the data were normalized so that the sum of all

probability values is in unity. For apo *x*/LTA4H 367 GNOM data points and for bestatin bound form 362 points were used. For inhibitor bound *hum*LTA4H 397 data points and apo form 477 points were used.

2.5. Inhibitor preparation and data collection

Bestatin was purchased from Sigma Aldrich and 50 mM stock solution was prepared. Buffer scattering curve without inhibitor (at various concentration) and with inhibitor were collected and compared to investigate if there were any difference present. SC-57461A 20 mM solution was prepared by dissolving the powder first in ammonia solution.

3. RESULTS

The experimental scattering curve obtained for *hum*LTA4H and *x*/LTA4H were initially scaled using PRIMUS and all these data showed no signs of aggregation. Radius of gyration (R_g) calculated for both the apo and bestatin bound *x*/LTA4H from the curve (36.6 ± 3 Å) is quite indicative of dimeric state of the molecule and is consistent between high and low concentration (0.5-12 mg/ml). The R_g value 27.6 ± 3 of *hum*LTA4H obtained from the curve is similar as the predicted value obtained from the crystal structure. When both the apo and inhibitor (bestatin and SC-57461A) bound scattering curves were compared with their respective crystal structure (see materials and methods section) a chi value of 1.74 was obtained, which indicates a good agreement between them. The curves of *sc*LTA4H showed R_g of 37.6 ± 3 with presence of instability indicating formation of unstable dimer or aggregation.

When the scattering profiles of the apo *hum*LTA4H and bestatin-bound or SC-57461A bound LTA4H, normalized for buffer content, are compared no major changes could be observed. These data indicate that *hum*LTA4H does not undergo any larger conformational changes upon inhibitor binding.

As in the crystal structure of the bestatin complex of *hum*LTA4, bestatin only occupied the LTA4 pocket partially, SC-57461A was used as an example of an inhibitor that should fill the pocket completely and serve as a better mimic of LTA4 itself. No

difference in behavior was observed for either of these inhibitors. While smaller conformational changes will go undetected in a SAXS experiment and need to be addressed in comparing scattering curves from carefully recorded WAXS experiments, these initial experiments show that the assumption that the *hum*LTA4H has a preformed substrate binding pocket, still holds.

We have observed that *sc*LTA4H might form dimers in solution (figure 5) similarly as seen for *x*LTA4H, which forms dimers both in crystal form and in solution. The formation of the dimer of *x*LTA4H is quite robust and consistent for very low to high concentrations (2-12 mg/ml) of protein. Since limited proteolysis was used during *x*LTA4H crystallization, *ab initio* rigid body modeling confirmed that the dimer formation is also consistent in solution and not just a crystallographic artifact. *x*LTA4H shows a change in scattering profiles in the higher/wide angular region ($>1 \text{ nm}^{-1}$) and a decrease in Porod volume of approximately 20 nm^3 but no change in R_g or D_{max} was observed (figure 4).

4. DISCUSSION AND CONCLUSIONS

Our solution SAXS studies show that *hum*LTA4H does not seem to undergo major conformational changes upon inhibitor binding which is consistent with our previous observations based on the crystal structures that this protein functions by a lock and key mechanism rather than induced fit. In contrast a change in scattering profiles is detected for the *sc*LTA4H and *x*LTA4H, indicating that conformational changes might be present. It is also observed that like in crystal structure *x*LTA4H forms dimer in solution. Similarly *sc*LTA4H forms dimer in solution, which has not been observed in the crystal structure. Taken together, *Xenopus* and *sc*LTA4H have a more compact form, with a decrease in flexibility, upon inhibitor binding, which reflects the changes observed in the crystal forms of apo and bestatin complexed forms of *sc*LTA4H.

As proteins are dynamic entities, and specific enzymatic mechanisms seem to arise from interplay between the substrate ligand and the protein, the family of related LTA4H proteins is a very attractive target for further studies. As the *x*LTA4H studies showed, substrate mobility could influence the product specificity [10]. In addition, mobility of the protein itself is also a source for both substrate and product specificity. The exact

interplay of the labile substrate LTA4, *hum*LTA4h and the product LTB4 and the role of the dynamics is an important issue as *hum*LTA4H is a target for drug discovery programs [20]. As previously speculated, a lock and key mechanism as observed for the *hum*LTA4H might be essential to ensure that only the biologically active LTB4 is formed from the short-lived and labile epoxide LTA4. In contrast *sc*LTA4H does not have these requirements and can allow for a larger mobility as provided by an induced fit mechanism. Interestingly other LTA4 converting enzymes, such as LTC4 synthase function as full membrane proteins and the lipidic environment these enzymes function in, might help to protect LTA4 (for review see [21]).

Further studies, in particular WAXS as well as computational studies such as combined quantum mechanics/ molecular mechanics (QM/MM) methods are therefore required to obtain more insight in the requirements of the different LTA4 related proteins and the impact on the catalytic mechanisms. This will lead to a better understanding of the double activity as well as the evolution of this intriguing family of proteins.

REFERENCES:

1. Samuelsson, B. (1983). Leukotrienes: mediators of immediate hypersensitivity reactions and inflammation. *Science*, 220(4597), 568-575.
2. Haeggström, J. Z. (2004). Leukotriene A4 hydrolase/aminopeptidase, the gatekeeper of chemotactic leukotriene B4 biosynthesis. *Journal of Biological Chemistry*, 279(49), 50639-50642.
3. Kull, F., Ohlson, E., & Haeggström, J. Z. (1999). Cloning and characterization of a bifunctional leukotriene A4 hydrolase from *Saccharomyces cerevisiae*. *Journal of biological chemistry*, 274(49), 34683-34690.
4. Helgstrand, C., Hasan, M., Uysal, H., Haeggström, J. Z., & Thunnissen, M. M. (2011). A Leukotriene A4 Hydrolase-Related Aminopeptidase from Yeast Undergoes Induced Fit upon Inhibitor Binding. *Journal of Molecular Biology*, 406(1), 120-134.
5. Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., ... & Yoshimoto, T. (2006). Crystal structure of aminopeptidase N (proteobacteria alanyl aminopeptidase) from *Escherichia coli* and conformational change of methionine 260 involved in substrate recognition. *Journal of biological chemistry*, 281(44), 33664-33676.
6. Addlagatta, A., Gay, L., & Matthews, B. W. (2006). Structure of aminopeptidase N from *Escherichia coli* suggests a compartmentalized, gated active site. *Proceedings of the National Academy of Sciences*, 103(36), 13339-13344.
7. Kyrieleis, O. J., Goettig, P., Kiefersauer, R., Huber, R., & Brandstetter, H. (2005). Crystal Structures of the Tricorn Interacting Factor F3 from *Thermoplasma acidophilum*, a Zinc Aminopeptidase in Three Different Conformations. *Journal of molecular biology*, 349(4), 787-800.
8. Thunnissen, M. M., Nordlund, P., & Haeggström, J. Z. (2001). Crystal structure of human leukotriene A4 hydrolase, a bifunctional enzyme in inflammation. *Nature Structural & Molecular Biology*, 8(2), 131-135.
9. Strömberg, F., Hamberg, M., Rosenqvist, U., Dahlén, S. E., & Haeggström, J. Z. (1996). Formation of a novel enzymatic metabolite of leukotriene A4 in tissues of *Xenopus laevis*. *European Journal of Biochemistry*, 238(3), 599-605.
10. Stsiapanava A, Tholander F, Kumar R, Qureshi A, Niegowski D, Hasan M, Thunnissen M, Haeggström J. Z. and Rinaldo-Matthis A (2014). Product Formation Controlled By Substrate Dynamics In Leukotriene A4 Hydrolase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(2), 439–446.

11. Penning, T. D., Russell, M. A., Chen, B. B., Chen, H. Y., Liang, C. D., Mahoney, M. W., ... & Smith, W. G. (2002). Synthesis of potent leukotriene A4 hydrolase inhibitors. Identification of 3-[methyl [3-[4-(phenylmethyl) phenoxy] propyl] amino] propanoic acid. *Journal of medicinal chemistry*, 45(16), 3482-3490.
12. Thunnissen, M. M., Andersson, B., Samuelsson, B., Wong, C. H., & Haeggström, J. Z. (2002). Crystal structures of leukotriene A4 hydrolase in complex with captopril and two competitive tight-binding inhibitors. *The FASEB Journal*, 16(12), 1648-1650.
13. Blomster, M., Wetterholm, A., Mueller, M. J., & Haeggström, J. Z. (1995). Evidence for a catalytic role of tyrosine 383 in the peptidase reaction of leukotriene A4 hydrolase. *European Journal of Biochemistry*, 231(3), 528-534.
14. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (pp. 571-607). Humana Press.
15. Knaapila, M., Svensson, C., Barauskas, J., Zackrisson, M., Nielsen, S. S., Toft, K. N., ... & Cerenius, Y. (2009). A new small-angle X-ray scattering set-up on the crystallography beamline I711 at MAX-lab. *Journal of synchrotron radiation*, 16(4), 498-504.
16. Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E., ... & McSweeney, S. (2013). Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *Journal of synchrotron radiation*, 20(4), 0-0.
17. P.V.Konarev, V.V.Volkov, A.V.Sokolova, M.H.J.Koch and D. I. Svergun (2003). PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J Appl Cryst.* 36, 1277-1282.
18. Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRY SOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, 28(6), 768-773.
19. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, 25(4), 495-503.
20. Çaliskan, B., & Banoglu, E. (2013). Overview of recent drug discovery approaches for new generation leukotriene A4 hydrolase inhibitors. *Expert opinion on drug discovery*, 8(1), 49-63.
21. Rinaldo-Matthis, A., & Haeggström, J. Z. (2010). Structures and mechanisms of enzymes in the leukotriene cascade. *Biochimie*, 92(6), 676-681.

Acknowledgement

Financial support: Alfred Österlund's foundation, Swedish Chemistry Society, Chemistry Society in Lund, Royal Physiographic Society in Lund.

The authors would like to thank to people at beamline BM29 (ESRF), especially Adam Round of helpful discussion about data processing, and Tomás Plivelic and Christopher Söderberg at I911-4, MAX IV laboratory.

Table 1: Crysol predicted values from apo and bestatin bound scLTA4H.

Parameters	scLTA4H Native (2XPZ) (Å)	scLTA4H complexed with bestatin (2XQ0) (Å)
Electron Rg:	25.74	25.53
Envelope Rg:	25.90	25.61
Shape Rg:	25.72	25.61
Shell Rg:	33.34	32.99
Shell volume:	0.3388E+05	0.3395E+05
Envelope volume:	0.1064E+06	0.1054E+06
Displaced volume:	0.9127E+05	0.9124E+05
Dry volume:	0.8798E+05	0.8801E+05
Shell width:	3.000	3.000
Molecular Weight:	0.7259E+05	0.7262E+05
Number of residuals:	627	625
Envelope surface:	0.1022E+05	0.1023E+05
Envelope radius:	47.60	48.21
Envelope diameter:	91.08	92.10
Average atomic rad.:	1.619	1.619
Center of the excess electron density:	0.283 0.314 0.139	0.061 -0.375 0.279



Figure 1: Superimposed scLTA4h crystal structure of native (Colored yellow, PDB code 2XPZ) to the complex with bestatin (colored blue, PDB code 2XQ0).

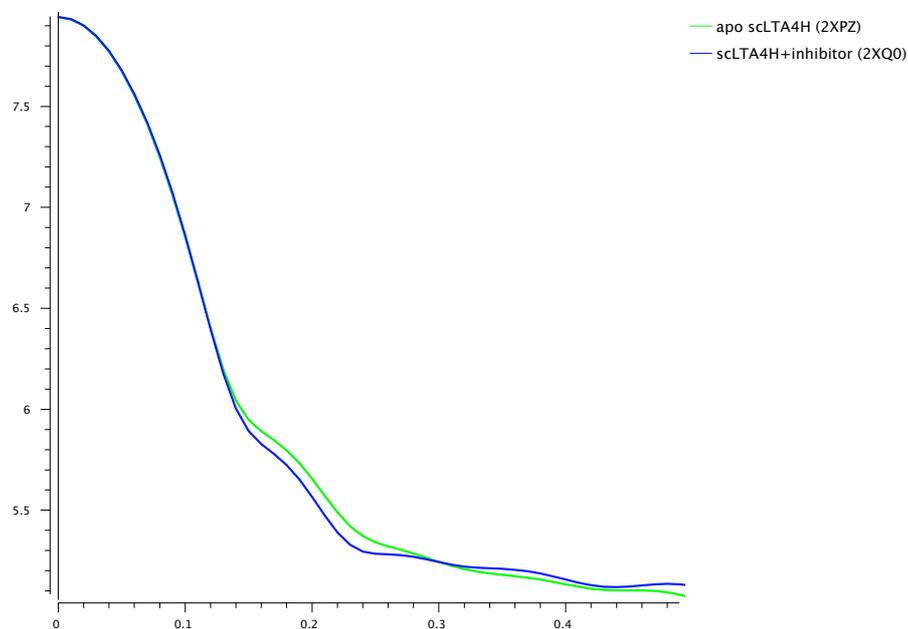


Figure 2: Overlaid predicted scattering curve from crystal structure of apo (PDB entry 2XPZ) and inhibitor bound (PDB entry 2XQ0) Yeast LTA4H calculated using program Crysol [17]. Curves are plotted by SASPLOT.

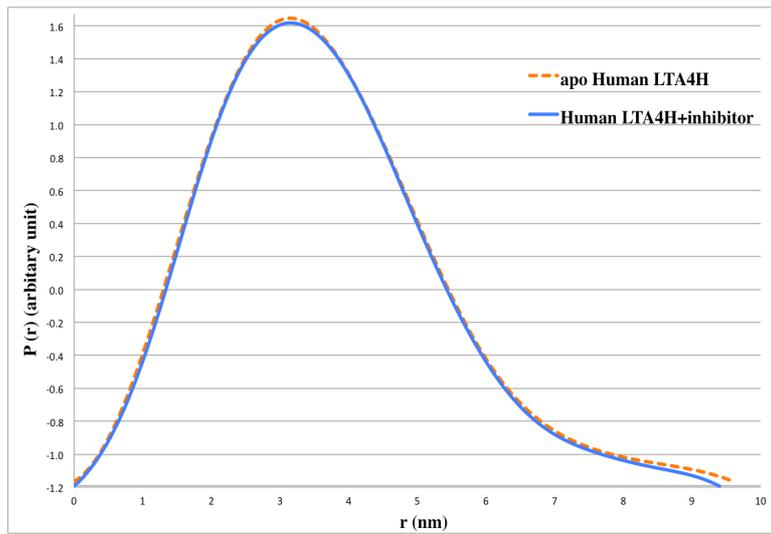
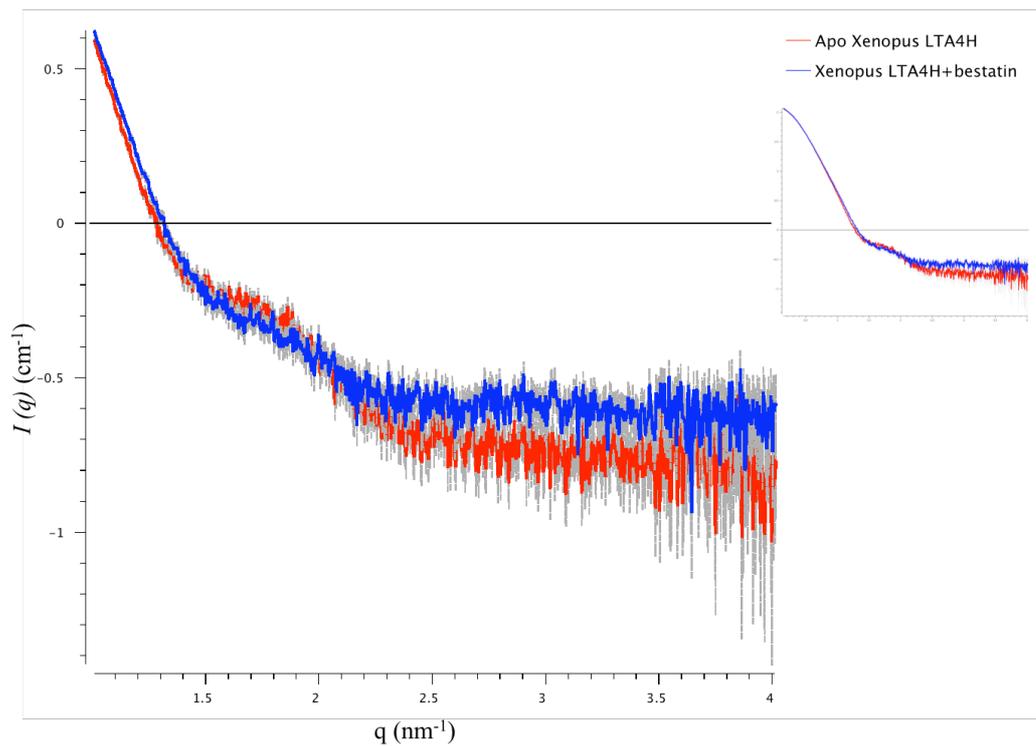
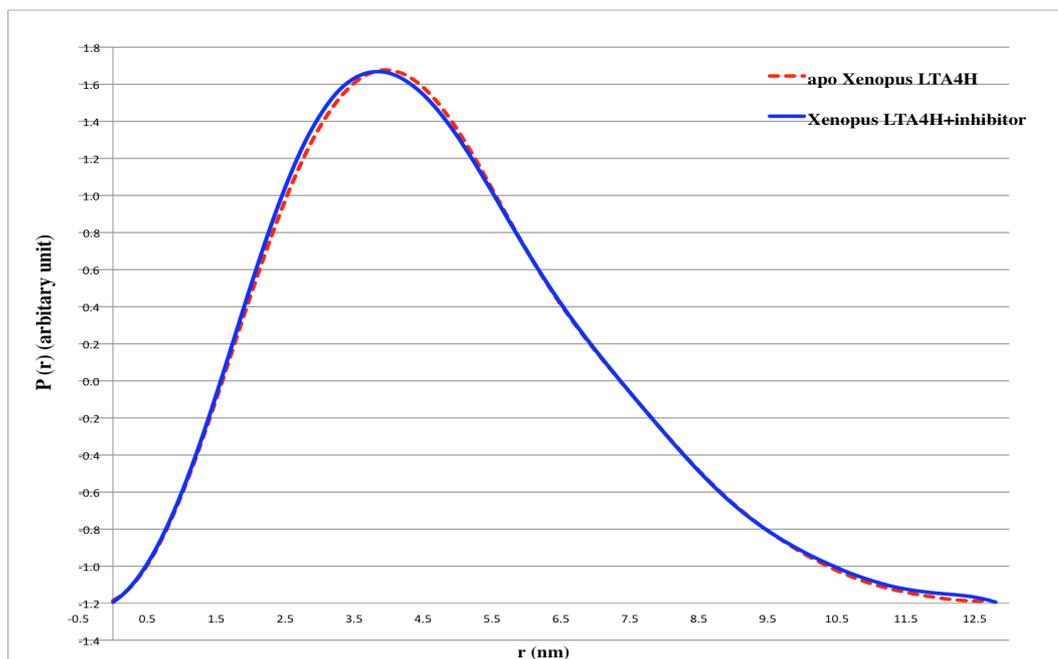


Figure 3: Overlaid pair distance distribution $[P(r)]$ function of apo human LTA4H to inhibitor bound. No difference was observed between the curves.



(a)



(b)

Figure 4: (a) Overlaid experimental scattering curve of apo and inhibitor bound *Xenopus* LTA4H. Whole curve is shown in inset. (b) Overlaid pair distance distribution function $[p(r)]$ of apo *Xenopus* LTA4H and inhibitor bound. The molecule is dimer in solution and a clear difference between the two curves (in lower, middle and higher q angle) is seen both in scattering curve and $p(r)$ function.

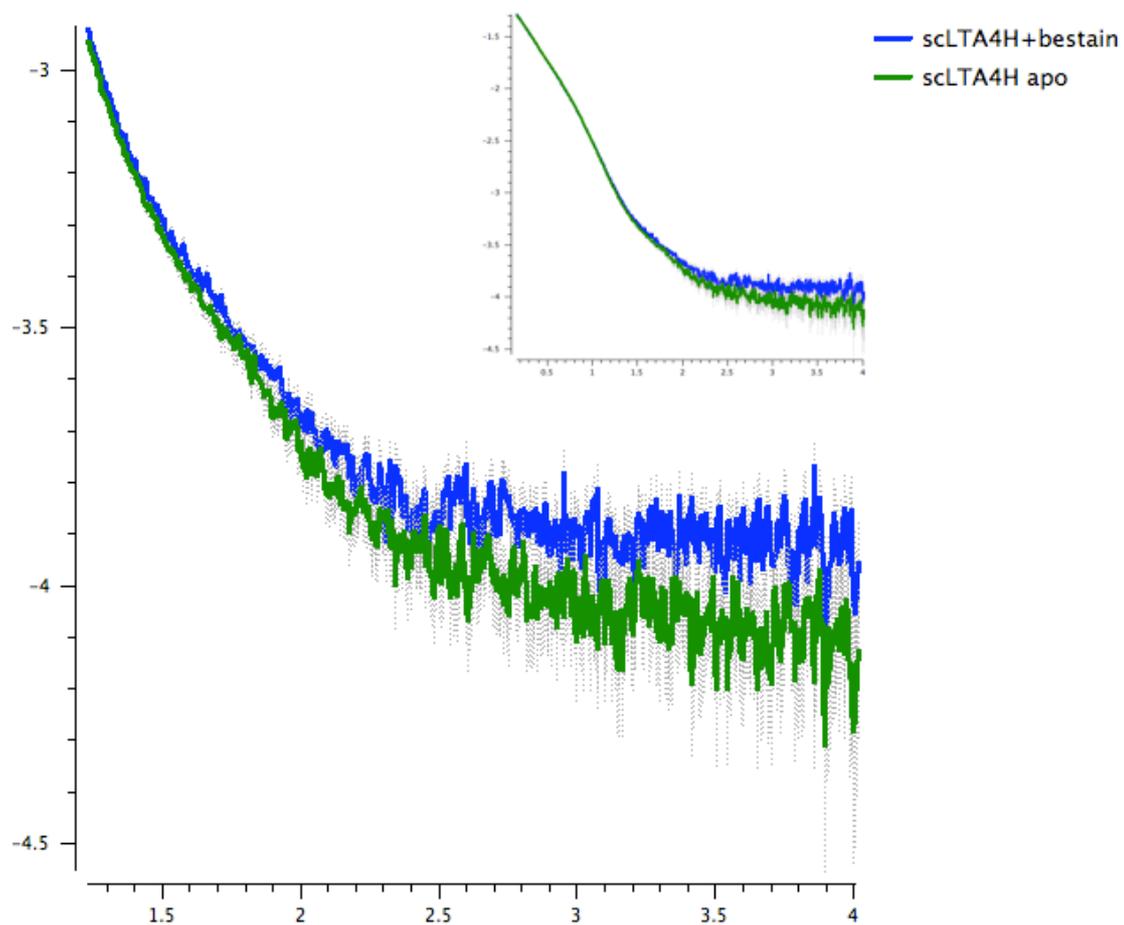


Figure 5: Overlaid experimental scattering curve of apo and inhibitor bestatin bound *scLTA4H*. Whole curve is shown in inset. Unlike crystal structure, the molecule is dimer in solution and a clear difference between the two curves (in middle and higher q angle) is seen. X- and Y-axis are same as in figure 4 (a)