Purification and crystallization of modular polyketide synthase involved in Lasalocid A biosynthesis

Thesis submitted to

Indian Institute of Science Education and Research, Pune

In partial fulfillment of the requirements for the BS-MS Dual Degree Program

by

Saket R. Bagde

Student ID 20121058

Research Mentor: Dr. Chu-Young Kim

Department of Chemistry,

University of Texas at El Paso, TX, USA

CERTIFICATE

This is to certify that this dissertation entitled “Purification and crystallization of modular polyketide synthase involved in Lasalocid A biosynthesis" towards the partial fulfilment of the BS-MS Dual Degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Saket R. Bagde at The University of Texas at El Paso under the supervision of Dr. Chu-Young Kim, Associate Professor, Department of Chemistry, The University of Texas at El Paso, during the academic year 2016-2017.

Dr. Chu-Young Kim

Date: 28-03-2017

Department of Chemistry

The University of Texas at El Paso

DECLARATION

I hereby declare that the matter embodied in the report entitled "Purification and crystallization of modular polyketide synthase involved in Lasalocid A biosynthesis" are the results of the work carried out by me at The University of Texas at El Paso under the supervision of Dr. Chu-Young Kim and the same has not been submitted elsewhere for any other degree.

Saket R. Bagde

Date: 28-03-2017

Abstract

Lasalocid A is a broad spectrum veterinary antibiotic given to cattle. However, it exhibits cross species toxicity and thus there is a need to develop lasalocid A analogs without cross species toxic effects. Lasalocid A is a polyether ionophore and its biosynthetic pathway consists of seven polyketide synthase enzymes (Lsd11-Lsd17), one epoxidase (Lsd18), and one epoxide hydrolase (Lsd19). In this work, polyketide synthase proteins Lsd13, Lsd14, Lsd15 and Lsd16 have been purified and preliminary X-ray diffraction study of Lsd14 crystals has been carried out. Determination of high resolution atomic structure of Lsd14 using X-ray crystallography will provide new structural information that can be used for biosynthetic production of Lasalocid A analogs.

Contents

List of Figures

List of Tables

Acknowledgement

I would like to extend my sincere gratitude towards my project supervisor Dr. Chu-Young Kim. I was able to complete this work not only because of the training I received during this project work at the Department of Chemistry at the University of Texas at El Paso, but also, due to the initial training I have received, back in IISER Pune under the guidance of Dr.Gayathri Pananghat, to whom I am very thankful. I'm grateful to Dr. Irimpan I. Mathews for helping us with diffraction data collection. I would also like to thank members of the lab including Qian Wang, Priyanka Gade, Sutinee Poolsup and Xi Chen for providing me with useful suggestions and keeping a lively and productive workspace.

1. Introduction

The need to come up with new line of antibiotics has increased recently due to the rapid rise of drug resistant bacteria. Structural analogs of existing antibiotics are possible contenders for this new line of drugs. Such analogs can be produced by organic synthesis. However, since natural product compounds often have very complex structure, large scale chemical synthesis of these molecules is not practical. An alternative to total chemical synthesis, or semi-synthesis, of natural product molecules and their analogs is to engineer the natural biosynthetic pathway and then produce the compound of interest in living cells. For this purpose, it is often necessary to understand the mechanism of the enzymes that make up the pathway. Atomic structure determination of the enzymes of interest will provide a wealth of detailed information, including the enzyme's catalytic mechanism, substrate specificity, and protein- protein interaction sites. This information can be used to rationally design enzyme mutants which, when introduced into the pathway, can produce the desired product**.**

1.1. Polyketides

Polyketides are secondary metabolites produced by bacteria, fungi, plants, and animals, and contain multiple β-hydroxyketone or β-hydroxyaldehyde (-$H_2C(=O)CH_2CH(OH)CH_2C(=O)$-) functional groups, which are further derivatized and modified into bioactive natural products with diverse biological activities and pharmacological properties(Robinson, 1991). These molecules find use in diverse roles such as antibiotics (example, Erythromycin)(Washington and Wilson, 1985), antifungals (example, Amphotericin B)(Caffrey et al., 2001), anticancer agents (example, Doxorubicin)(Lomovskaya et al., 1999),cholesterol lowering agents (example Lovastatin)(Singer et al., 1988), membrane components and cytotoxins (figure 1).

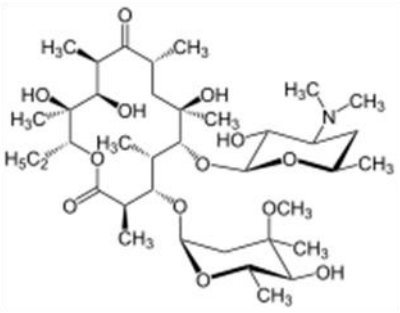
1.2. Polyketide Biosynthesis

Polyketides are synthesized by coupling the coenzyme A (CoA)-derived building blocks (e.g., acetyl-CoA, malonyl-CoA, and methylmalonyl-CoA) into linear chains by iterative Claisen condensation (Robinson, 1991). The resulting β-keto groups are sometimes reductively modified to yield α,β-hydroxyl, an α,β-double bond, or fully-reduced

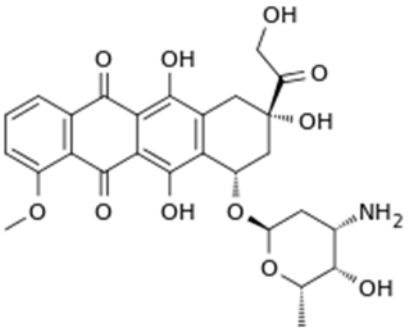
methylene group. These reactions are carried out by complexes involving multiple enzymes, which are called polyketide synthase.

1.3. Polyketide synthase (PKS) enzymes and their functional domains

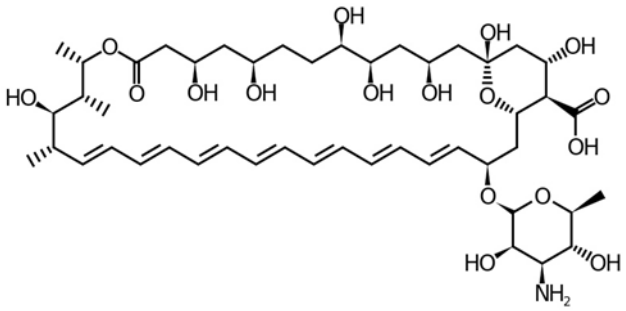
The polyketide synthase enzymes comprise of various catalytic functional units which can either be standalone enzymes coming together to form a productive complex or can



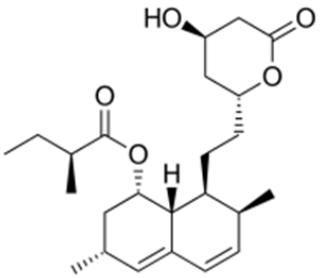
Erythromycin A



Doxorubicin



Amphotericin B



Lovastatin

Figure 1. Structures of representative polyketides.

be parts of a single polypeptide chain. These domains include acyltransferase, acyl carrier protein, ketosynthase, ketoreductase, dehydratase, enoyl reductase and thioesterase (Bevitt et al., 1992; Donadio et al., 1991).

The Acyltransferase (AT) selects the appropriate extender unit to be incorporated in the growing polyketide chain. Acyl carrier protein (ACP) serves as a covalent attachment point for the intermediate during the assembly process. The intermediate is attached to the phosphopantetheine (Ppant) prosthetic group present on the ACP. Ketosynthase (KS) domain

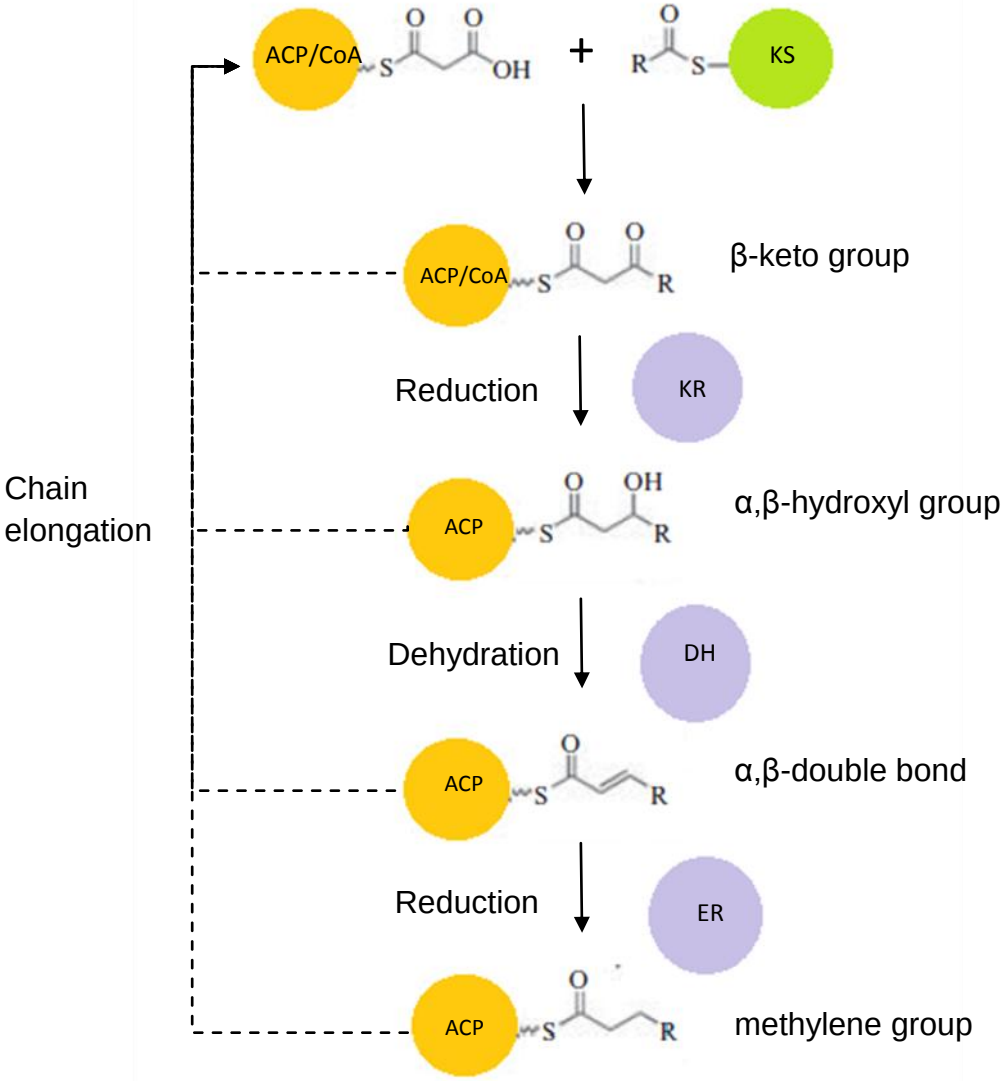


Figure 2. Reaction scheme for polyketide biosynthesis.

joins the building block chosen by AT to the growing polyketide chain using Claisen-condensation reaction. Ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains are the reductive modifying domains which yield respectively α,β-hydroxyl, α,β-double bond, or fully-reduced methylene (figure 2). Thioesterase domain (TE) catalyzes the cyclization

and release of the polyketide chain from the polyketide synthase(Staunton et al., 2001; Weissman, 2009).

1.4. Types of Polyketide synthase

Polyketide synthases have been categorized into three broad categories Type I, Type II and Type III(Weissman, 2009). In Type I polyketide synthases, the catalytic and carrier protein domains are joined together by linker regions and organized into operational units called modules, such that each module typically carries out one round of chain extension and reductive processing (Khosla et al., 2014). Type I PKS can also be iterative, which use the same catalytic units for each cycle, for example Lovastatin polyketide synthase (Staunton, 1998).

Type II polyketide synthases are iterative and consist of discrete catalytically functional enzymes which associate and form complexes to produce the polyketide product upon iterations through a defined number of chain extension cycles (Hertweck et al., 2007).

Type III polyketide synthases are ACP independent and iterative. The extender unit attaches directly to the ketosynthase which catalyze the condensation of extender units onto a starter unit upon several iterative chain extension cycles (Katsuyama and Ohnishi, 2012).

1.5. Lasalocid and its biosynthesis pathway

Lasalocid is an antibacterial agent which is produced by the strains of *Streptomyces lasaliensis*. It is of commercial value and is one of the active components in the feed additives called Bovatec (CVM and Onade, 2006). It is a polyether ionophore and hence can induce ion transport across apolar phase (including lipid bilayer membranes) by forming complexes with monovalent and divalent cations. (Duax et al., 1996; Hilgenfeld and Saenger, 1982).

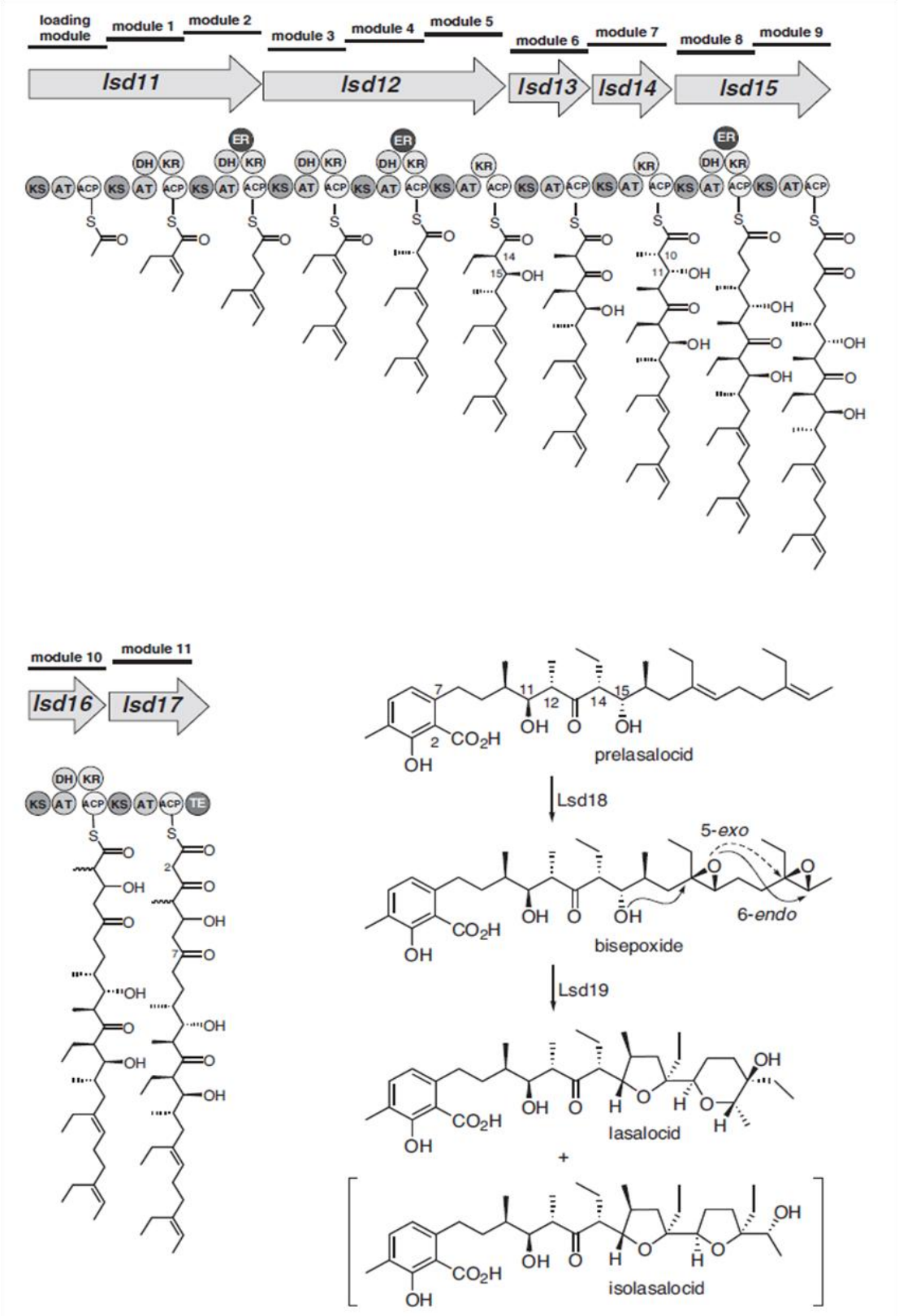


Figure 3. Biosynthesis of Lasalocid A. Organisation of Polyketide synthase for synthesis of Lasalocid (Lsd11-Lsd17) and the predicted structures of precursors bound to the PKS, and epoxidation and ether ring formation by Lsd18 and Lsd19. Figure modified from (Migita et al.)

Lasalocid biosynthesis pathway consists of polyketide synthases (Lsd11- Lsd17), an epoxidase (Lsd18) and an epoxide hydrolase (Lsd19) (Figure 3). The PKS involved in Lasalosid biosynthesis are type I polyketide synthases and consist of one loading module and 11 extension modules placed on seven polypeptides (Lsd11, Lsd12, Lsd13, Lsd14, Lsd15, Lsd16 and Lsd17). Five malonyl-CoA, four methylmalonyl-CoA and three ethylmalonyl-CoA subunits are consumed to produce a linear dodekaketide which further undergoes cyclization, epoxidation and ring opening hydrolysis reactions to give the final polycyclic ether product(Migita et al.; Wong et al., 2015).

Lsd11 contains three modules including the loading module. Lsd12 and Lsd15 contain three and two modules respectively. Lsd13, Lsd14, Lsd16 and Lsd17 contain one module each. Hence this system provides opportunities to study interactions between modules both, which are part of separate polypeptides and single polypeptide chain. Though NMR and X-ray structures for individual domains of other polyketide synthases have been solved, to date there is no high resolution atomic structure of a polyketide synthase full length module. This can be attributed to the large molecular weights of these enzymes (>100kDa) and presence of flexible linkers. The lack of detailed structural information on polyketide biosynthesis puts a major limitation on experiments aimed at engineering the polyketide biosynthetic pathway to produce desired products. Hence, there is a need to solve the atomic structure of the full length polyketide synthase module Lsd14 which is the focus of this work.

2. Materials and Methods

2.1. Construction of expression vector

Custom-synthesized codon optimized DNA sequences for Lsd11, Lsd12, Lsd13, Lsd14, Lsd15, Lsd16 and Lsd17 proteins cloned in pET28a+ vector were provided by GenScript®. Genes for Lsd11, Lsd12, Lsd13 and Lsd17 were introduced into pCold-I vector via restriction-enzyme based cloning and ligation. The pET-Lsd constructs were transformed into chemically competent *E. coli* DH5α cells by heat shock treatment. Positive transformants were selected on LB agar culture plates supplemented with 50µg/ml kanamycin. Single colonies were inoculated in 5ml LB media supplemented

with 50µg/ml kanamycin and incubated at 37°C for 16 hours. Cells were harvested by centrifugation at 6000xg and the resulting cell pellets were used for plasmid DNA extraction using PureLink® Quick Plasmid Miniprep Kit from Invitrogen®. Extracted plasmids were digested using NdeI and EcoRI restriction enzymes obtained from New England Biolabs® Inc. in a single digestion reaction. The gene of interest with 'sticky' ends was gel extracted using PureLink® Quick Gel Extraction Kit from Invitrogen® after separation from the linearized vector DNA upon gel electrophoresis using 0.6% Agarose LE gel. Similarly, linearized pCold-I vector with NdeI-EcoRI 'sticky' ends was obtained. The gene of interest was inserted into pCold-I vector via ligation reaction carried out using Instant Sticky-end Ligase Master Mix from New England Biolabs® Inc. The ligation reaction mixture was used for transformation of One Shot® Mach1™-T1R Chemically Competent *E. coli* cells and positive transformants were selected on LB agar culture plates supplemented with 100µg/ml ampicillin. Positive clones were identified by restriction digestion of extracted DNA plasmid using NdeI and EcoRI to check for insertion of gene of interest.

2.2. Chemical transformation of E. coli

Approximately 100ng of plasmid DNA was added to competent cell aliquot and mixed gently by tapping. The transformation mix was incubated in ice for 30 minutes. Heat shock was given at 42°C for 60 seconds using Eppendorf ThermoMixer® C after which the transformation mix was incubated in ice for 1 minute. 200 µL of pre-warmed (37°C) LB media was added and the mixture was incubated at 37°C with shaking at 220rpm for 1 hour. The transformed cells were spread on selective LB agar plates containing appropriate antibiotic and incubated at 37°C for 12 hours.

2.3. Protein expression

Plasmids (pet28a+) expressing Lsd14, Lsd15 and Lsd16 as N-terminal thrombin cleavable hexa-Histidine tagged constructs and plasmids (pCold-I) expressing Lsd11, Lsd12 , Lsd17 as N-terminal Factor Xa cleavable hexa-Histidine tagged constructs were expressed in BL21(DE3) *E. coli* cells. Expression culture (LB broth, supplemented with 50µg/ml kanamycin for pet28a+ and 100µg/ml ampicillin for pCold-1) was inoculated

with primary culture grown until OD_{600} reached a value of 2 at 37°C, in the ratio 1:50. The expression culture was grown until OD_{600} reached a value of 0.4, after which the temperature of the incubator was switched to 15°C. The cultures were induced with 0.1mM IPTG at OD_{600} value 1.0. The cells were harvested after 24 hours by centrifugation at 6000xg and the cell pellets were stored at -80°C.

2.4. Protein purification

Cells from 1L expression culture were resuspended in 25ml lysis buffer (200mM NaCl, 50mM Tris, 20% Glycerol, pH 8.0; supplemented with 1mM EDTA, 0.5mM PMSF), homogenized using sonication, and centrifuged at 30,000g for 1 hour to clear the cell debris. All the purification steps were carried out at 4°C. Protein of interest was precipitated upon addition of a 100% saturated solution of ammonium sulfate drop by drop, with constant stirring, to a final saturation of 45% for Lsd14 and 50% for Lsd13, Lsd15 and Lsd16. The precipitated protein was harvested upon centrifugation at 30,000g for 30 minutes and was resuspended in 11.25ml IMAC buffer A (300mM NaCl, 20mM Sodium Phosphate, 10% Glycerol, pH 7.0 supplemented with 1mM TCEP) and was applied to IMAC (Immobilised Metal Affinity Chromatography) column charged with nickel ions, pre-equilibrated with IMAC buffer A (two 1ml Histrap FF columns in series were used from GE Healthcare; and one 5ml Histrap HP column from GE Healthcare was used when the purification was upscaled using 5L culture).Protein of interest was eluted with IMAC buffer A containing 250mM imidazole, after washing the column with 10CV and 60CV of IMAC buffer A with 50mM and 100mM imidazole respectively. The protein so obtained was further purified using anion exchange chromatography (AEC). IMAC purified protein was applied to 1ml Hitrap QHP column (GE healthcare; 5ml column was used upon scale up) upon dilution with AEC buffer A (50mM Tris, 10% glycerol, pH 8.5) in the ratio 1:5. Protein was eluted with a gradient of increasing percentage of AEC buffer B (1M NaCl, 50mM Tris, 10% glycerol, pH 8.5) from 34.5%B to 50%B over 8 column volumes after 10 column volume wash with 34.5%B for Lsd15, and 30%B to 50%B over 10 column volumes, after 10 column volume wash with 30%B for Lsd13 and Lsd14, and 20%B to 60%B over 5 column volume for Lsd16. AEC purified protein was further purified with SEC (Size Exclusion Chromatography) using

Superdex200 Increase 10/300 GL column (GE Healthcare) equilibrated with SEC buffers (Lsd13, Lsd14 - 150mM NaCl, 25mM HEPES, pH 7.2; Lsd15 - 50mM NaCl, 25mM HEPES, pH 7.5; Lsd16 - 250mM NaCl, 25mM HEPES pH 7.5) giving the most homogeneous sample.

2.5. Reductive alkylation of purified protein

Reductive alkylation of protein was carried out using Reductive alkylation Kit from Hampton Research. 20 µl of 1M Dimethylamine Borane Complex solution and 40 µl of 1.0 M Formaldehyde was added to 1.0 ml of protein at a concentration of around 1mg/ml. The reaction mixture was mixed gently and incubated at 4°C for 2 hours. After 2 hours incubation, 20 µl of 1M Dimethylamine Borane Complex solution and 40 µl of 1.0 M Formaldehyde was added again and the reaction mixture was incubated further for 2 hours. Further, 10 µl of 1M Dimethylamine Borane Complex solution was added and the reaction mixture was incubated at 4°C for 12 hours. The reaction was stopped by addition of 125 µl of 1M Glycine and the protein was separated from the reaction components using Size exclusion chromatography. Same protocol was used for reductive ethylation of protein using acetaldehyde instead of formaldehyde and reductive isopropylation of protein using acetone.

2.6. Crystallization and data collection

Initial crystallization screening experiments were set up for Lsd13, Lsd14, Lsd15 and Lsd16, using commercially available crystallization screens JCSG+ and Morpheus from Molecular Dimensions and MCSG Crystallization Suite from Anatrace. 0.2µL of protein at three different concentrations between 1-5mg/ml was mixed with 0.2µL of reservoir solution in a 96 well sitting drop crystallization plate using NT8 nanoliter-volume liquid handler robot and the plates were incubated at 18°C.

Diffraction quality crystals for Lsd14 were obtained by mixing 0.5 µL of protein at 1.2-1.5mg/ml with 0.5 µL of reservoir condition (formulation has not been stated as this work has not been published yet) in a sitting drop experiment. Crystals were obtained in hanging drop vapor diffusion set up using 24 well VDX crystallization plates upon mixing 1 µL of protein at 1.2-1.5mg/ml with 1 µL of reservoir condition.

Co-crystallization of Lsd14 with NADPH and NADP was carried out by setting crystallization experiments with protein incubated with 30 times molar excess of NADP or NADPH for 12 hours.100mM stock solutions of NADP/NADPH dissolved in de-ionized water were used and the pH of the solution was not set.

Clusters of Lsd14 crystals were broken to yield single crystals by crushing the crystal between two cover slides. A single cluster was picked up using a cryo loop and was transferred to 2 µL drop of reservoir solution supplemented with cryoprotectant (10-25% ethylene glycol/glycerol). Another coverslip was placed on this drop so that the drop containing the crystal is sandwiched between the two coverslips, which were moved against each other gently so as to break the cluster apart into single crystals suitable for X-ray diffraction studies. The single crystals so obtained were harvested and flash cooled in liquid nitrogen.

Diffraction data set was collected at Stanford Synchrotron Radiation Lightsource on the BL12-2 microfocus beamline at 100K. Data reduction was performed using XDS program suite (Kabsch, 2010).

2.7. Concentration estimation of protein

The concentrations of proteins stated in this work were measured using Eppendorf Biospectrophotometer Basic by measuring absorbance at 280nm considering the theoretical extinction coefficient calculated using ProtParam (Lsd14 - 204935$M^{-1}cm^{-1}$).

3. Results and discussion

3.1. Recombinant expression of polyketide synthase enzymes involved in Lasalocid Biosynthesis

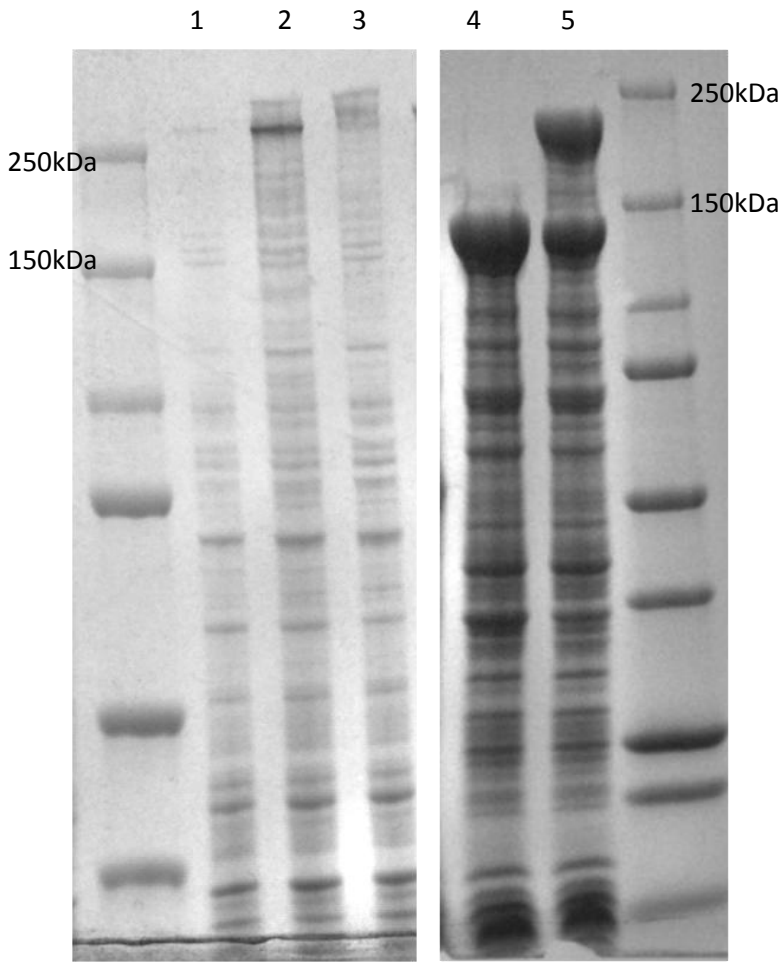


Figure 4. Expression of Lsd11 (515kDa), Lsd12 (572kDa), Lsd17 (134kDa) and co-expression of Lsd16 and Lsd17. Lanes : 1- uninduced sample, 2- Lsd11 expression, 3-Lsd12 expression, 4-Lsd17 expression, 5- Lsd17 and Lsd16 coexpression

The seven polyketide synthase enzymes Lsd11 (515kDa), Lsd12 (572kDa), Lsd13 (119kDa), Lsd14 (174kDa), Lsd15 (390kDa), Lsd16 (200kDa) and Lsd17 (134kDa) were successfully expressed as in E. coli BL21 (DE3) cells, as shown in figure 4, 9 and 10.

The expression levels of full length constructs of Lsd11 and Lsd12 were undetectable when expressed in vector pet28a+. However, over expression of smaller sized bands was observed suggesting possible degradation by proteases or incomplete translation. Increasing the inducer concentration and changing expression temperatures did not increase the expression levels of the full length protein. Expression level of Lsd13 when expressed from pET28a+ vector was also low, leading to low yields of purified protein. Expression of Lsd11, Lsd12 and Lsd13 in pCold-I vector led to increase in expression levels for Lsd11, Lsd12 and Lsd13. However, the expression levels for Lsd11 and Lsd12 were still quite low large for scale purification to be carried out for crystallographic studies.

Lsd17 on the other hand was highly expressed but no protein in the soluble fraction was obtained. Expression with low levels of inducer (IPTG) concentration, expression at low temperatures (12°C, 15°C and 18°C) was tried. Lsd17 is preceded by Lsd16 in the Lasalocid biosynthetic pathway and is expressed as a soluble construct. Hence coexpression of Lsd17 with Lsd16 was also tried. Also, expression of Lsd17 in other *E. coli* expression strains known to increase soluble fraction of overexpressed protein was tried. Expression in BL21AI cells and Lemo(DE3) cells (New England BioLabs) was carried out to allow for tunable expression of the protein in order to obtain soluble and properly folded protein. Lysozyme is a natural inhibitor of T7 RNA polymerase. Hence, the levels of overexpression can be varied in Lemo(DE3) cells by varying the level of lysozyme (Schlegel et al., 2012). The level of lysozyme is modulated by adding L-rhamnose to the expression culture at levels from zero to 2000 µM. Tunable expression in BL21AI cells can be achieved by varying the amount of L-arabinose added to the expression culture as the T7 RNA polymerase expression in BL21AI cells is under the control of the arabinose-inducible *ara*BAD promoter.

An untagged construct of Lsd17 was also expressed to dismiss the possibility of the interference of the N-terminal hexa-histidine tag leading to improper folding of the protein. However, none of the trials mentioned above were successful in getting Lsd17 in the soluble fraction.

Further, co-expression of Lsd17 with *Streptomyces* chaperones may be able to produce soluble Lsd17(Betancor et al., 2008). Also, expression of the protein from the native organism *Streptomyces lasaliensis* can be tried.

The low expression levels of Lsd11, Lsd12 and the lack of soluble Lsd17 protein would be a major bottleneck in the in vitro re-constitution experiments or heterologus biosynthesis experiments to synthesize Lasalocid. The expression-solubility levels of Lsd13, Lsd14, Lsd15 and Lsd16 were satisfactory. Hence the purification of these polyketide synthases for X-crystallographic studies was focused upon.

3.2. Purification of polyketide synthase enzymes Lsd13, Lsd14, lsd15 and Lsd16

The purification procedure for Lsd13, Lsd14, Lsd15 and Lsd16 has been standardized. The steps involved are ammonium sulphate precipitation, immobilized metal affinity chromatography (Ni), anion exchange chromatography and size exclusion chromatography as discussed in detail in section 2.4.

Ammonium sulphate precipitation: The solubility of proteins usually decreases at high salt concentrations, and the protein precipitates out of the solution. This is a result of increased hydrophobic interaction between protein and water due to increase in surface tension of water upon addition of salt, which leads to self association of protein molecules to minimize contact area with water. These interactions will depend on the size (surface area exposed to solvent) and nature of the residues exposed to the solved. Hence, precipitation of the protein out of the solution upon addition of ammonium sulphate can lead to purification of the protein. The precipitated protein can be resolubilized upon resuspension in low salt buffer.

For purifying proteins with low expression levels, ammonium sulphate precipitation also gives an added advantage of concentrating the protein into lower volumes, from a large volume of cell lysate which is used to get higher yield of protein for crystallographic studies.

Immobilized Metal Affinity Chromatography (IMAC)

All the proteins in this study were expressed as N-terminal hexa-Histidine tagged constructs to aid in purification using IMAC. The Histidine residues in the His-tag bind to nickel ions immobilized on a stationary matix. Non-specifically or weakly bound impurity proteins are eluted upon washing the matrix with buffer containing low concentrations of imidazole which binds competitively to the immobilized nickel ions on the matrix. His-tagged protein can then be eluted using higher concentration of imidazole in the elution buffer.

The binding of Lsd13, Lsd14, Lsd15 and Lsd16 to the IMAC column was sub-optimal. The IMAC buffer composition was optimized to get maximal binding of the protein to the IMAC column, and the purity of the protein so obtained. IMAC trials using buffers with 300mM, 500mM and 1M NaCl, at pH 7.0, 7.4 and 8.0 were carried out to screen for

buffer giving maximum binding. Buffer 300mM NaCl, 20mM Sodium Phosphate, 10% Glycerol, pH 7.0 gave the best binding.

Further purification of the protein was necessary to get high quality protein for crystallographic studies. Hence, IMAC purified protein was further subjected to Anion Exchange Chromatography.

Anion Exchange Chromatography (AEC)

Proteins can be separated using ion-exchange chromatography based on the net surface charge. Anion exchange chromatography involves use of a positively charge stationary phase and hence proteins with negative net surface charge can bind to the stationary phase. At a given pH, proteins with different iso-electric point (pI) values will have different surface charge and can be separated using ion-exchange chromatography based on their affinity to the charged stationary phase upon application of a salt gradient. Proteins having high affinity to the ion exchange matrix will elute at high ionic strength.

Size exclusion chromatography (SEC)

AEC purified protein was further subjected to Size exclusion chromatography to separate aggregates and to get a homogenous protein sample for crystallographic studies.

Buffers with pH 7.0, 7.5, 8.0 and 9.0, NaCl concentrations 0mM, 50mM, 150mM, 250mM, 350mM, 500mM, and additives EDTA, DTT, and glycerol, were screened to get homogeneous protein sample marked by a single symmetric peak in the size exclusion chromatogram. Lsd13, Lsd14 and Lsd16 eluted as homogenous samples in buffer with composition 150mM NaCl, 25mM HEPES, pH 7.2 (Lsd13,Lsd14) and 500mM NaCl, 25mM HEPES pH 7.5 (Lsd16) respectively, as seen in the SEC chromatograms in Figures 5,6 and 7. Lsd15 eluted as two merged peaks in the UV chromatogram, one of which being in the void volume of the column, suggesting part of the protein is either aggregated or a higher order oligomeric species. Best separation in SEC was obtained in buffer 50mM NaCl, 25mM HEPES, pH 7.5 for Lsd15. (Figure 8)

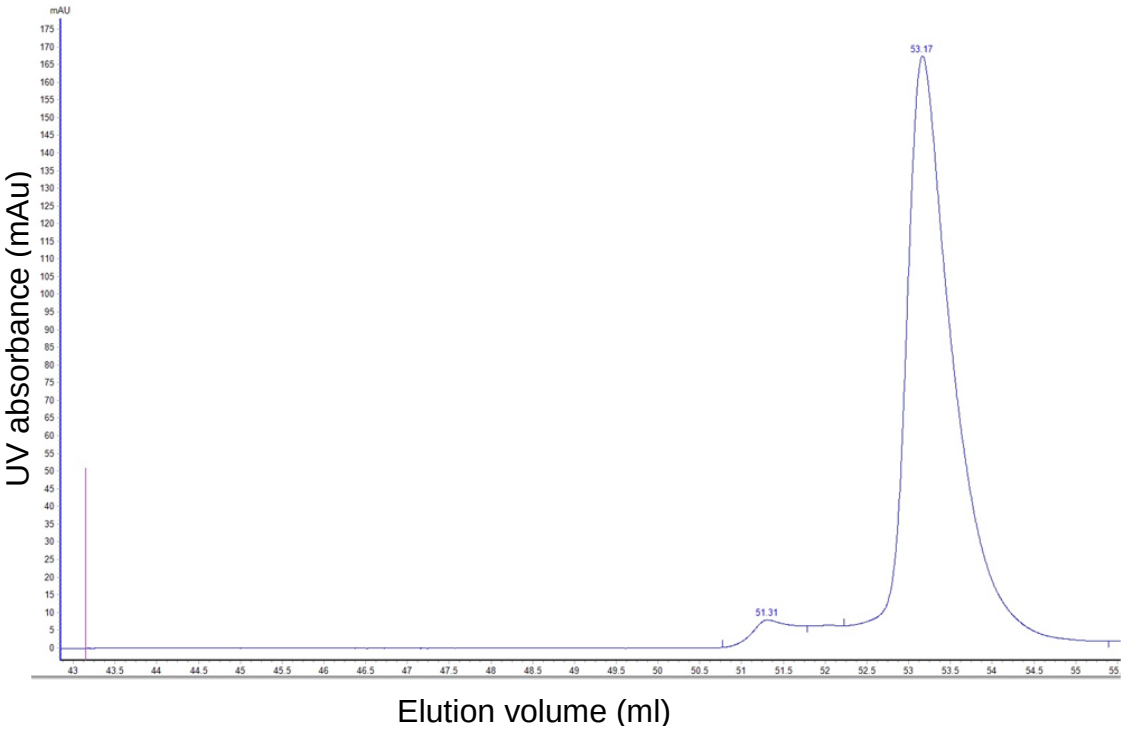


Figure 5. Size exclusion chromatogram of Lsd13

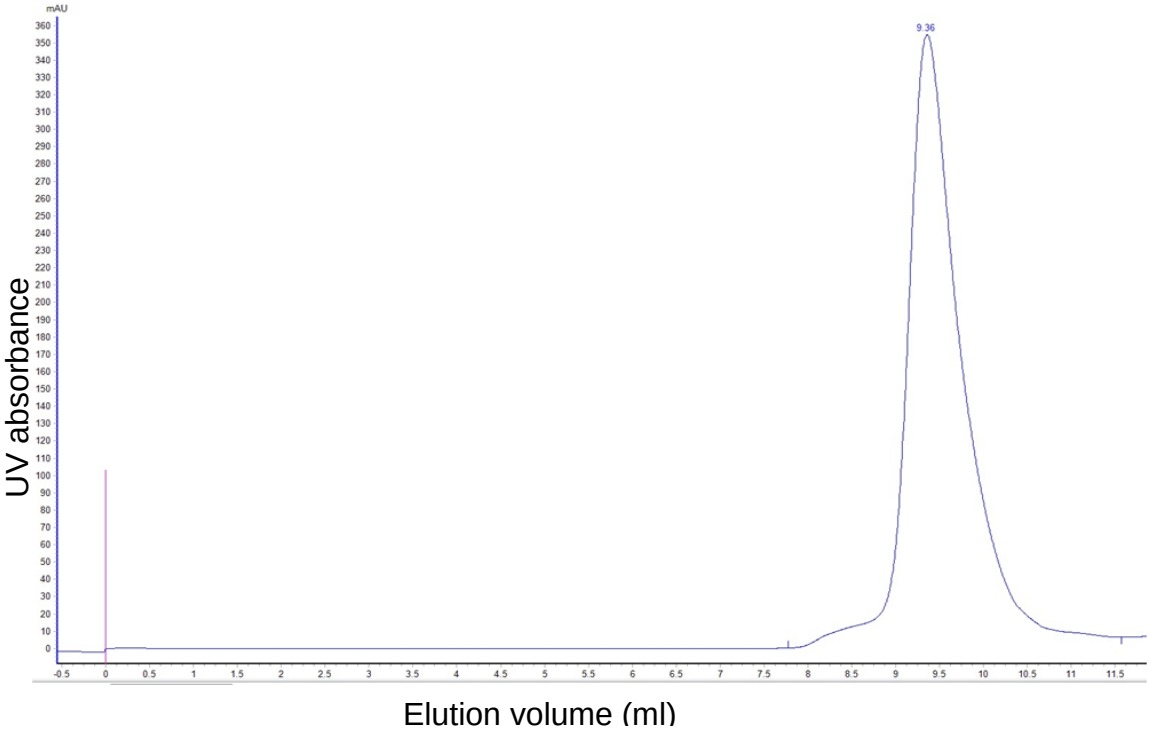


Figure 6. Size exclusion chromatogram of Lsd14

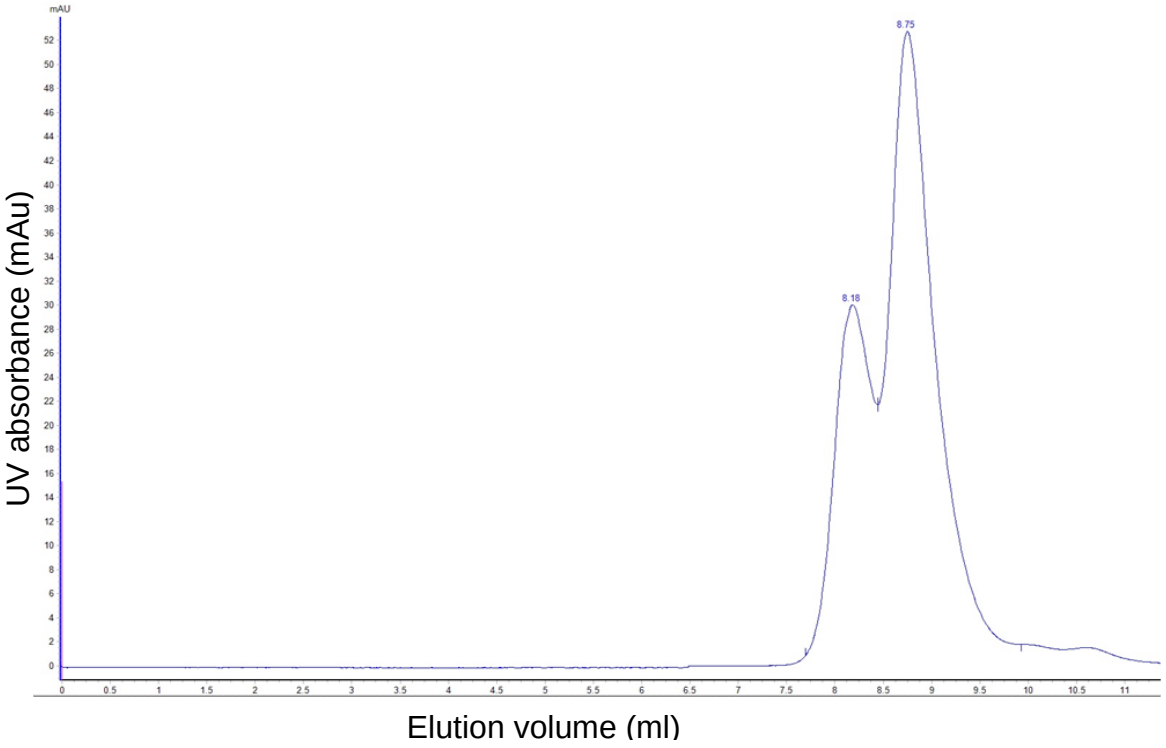


Figure 7. Size exclusion chromatogram of Lsd15

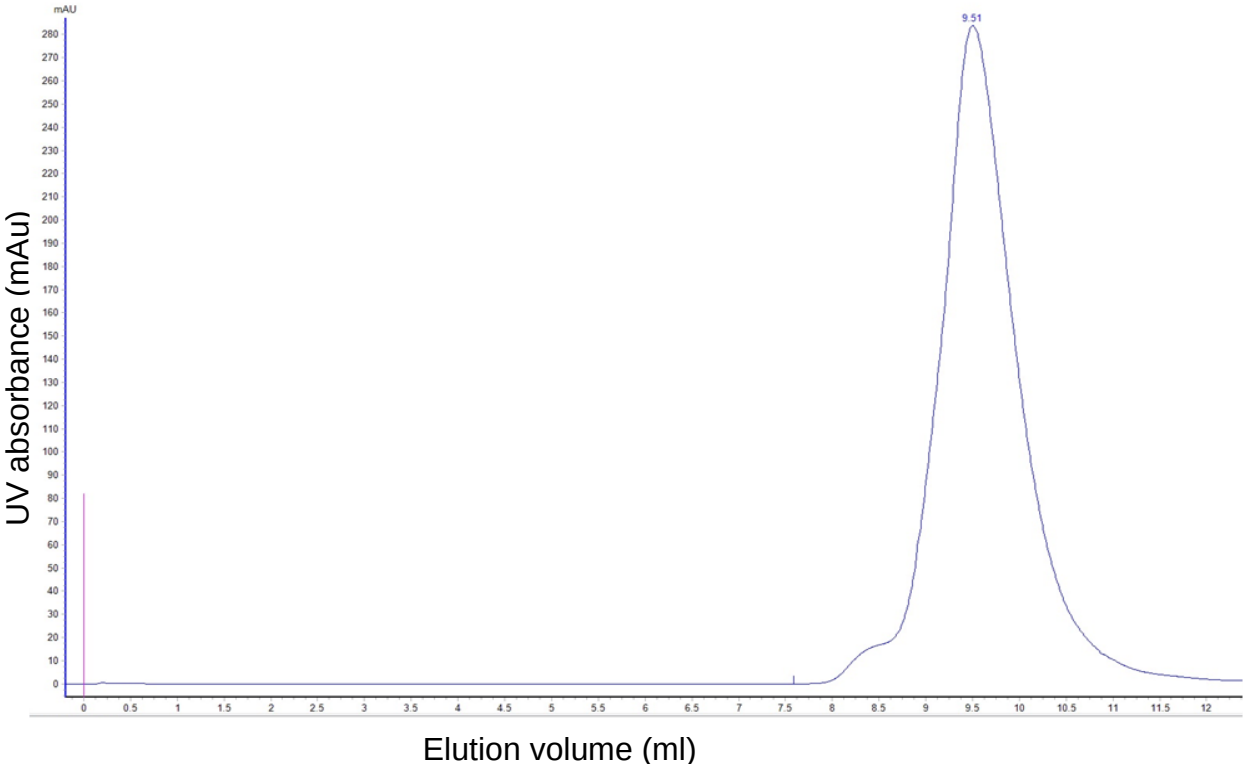


Figure 8. Size exclusion chromatogram of Lsd16

Figure 9 and 10 show purified proteins analyzed by SDS PAGE. High purity levels were obtained for Lsd13, Lsd14 and Lsd16. However, the purity of Lsd15 was suboptimal for crystallization experiments. Further optimization of the purification procedure needs to be carried out for Lsd15 to get a higher purity and homogeneity. Including another affinity tag on the C terminal of the protein (example Strep tag) can aid in purification to give higher purity levels.

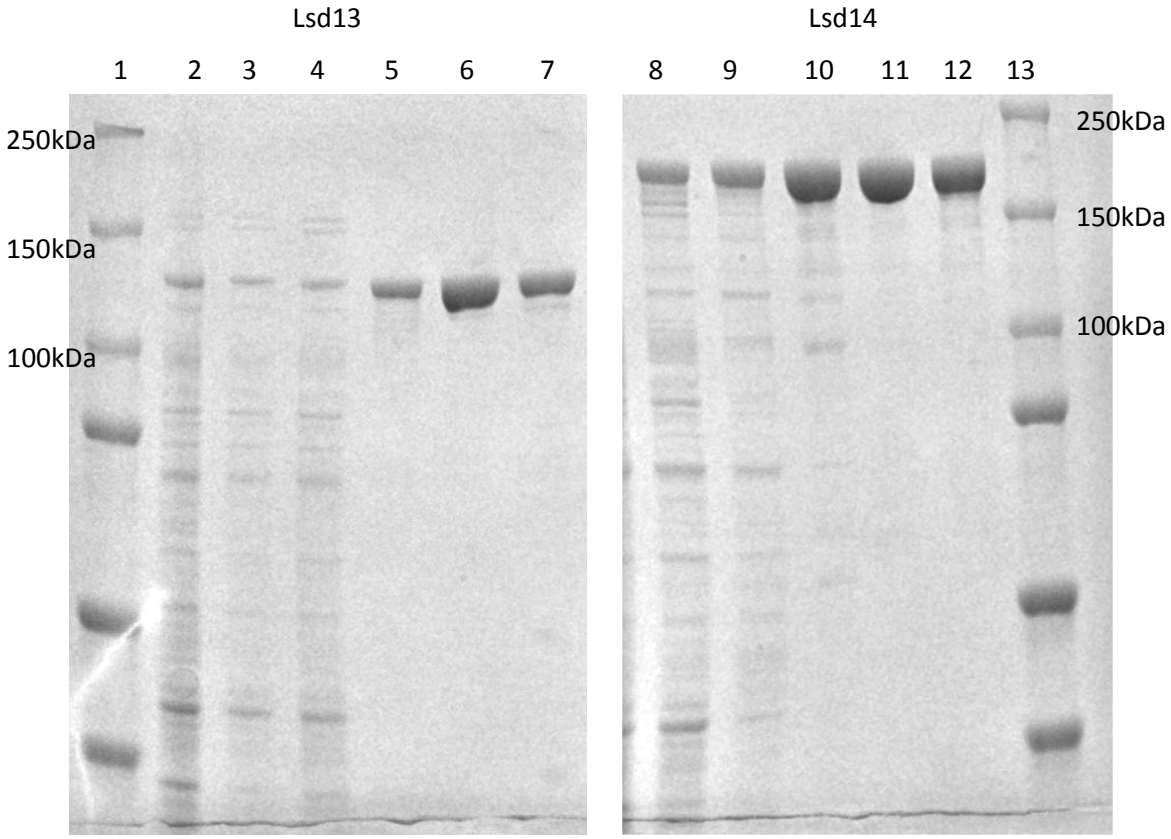


Figure 9. SDS PAGE analysis of purified Lsd13 and Lsd14. Lanes : 1,13 are protein molecular weight markers, 2- Lsd13 total cell lysate, 3- soluble fraction of total cell lysate, 4- IMAC load, 5- IMAC purified protein, 6- AEC purified protein, 7- SEC purified protein. 8-Lsd14 soluble fraction of cell lysate, 9-IMAC load, 10-IMAC purified protein, 11-AEC purified protein, 12-SEC purified protein

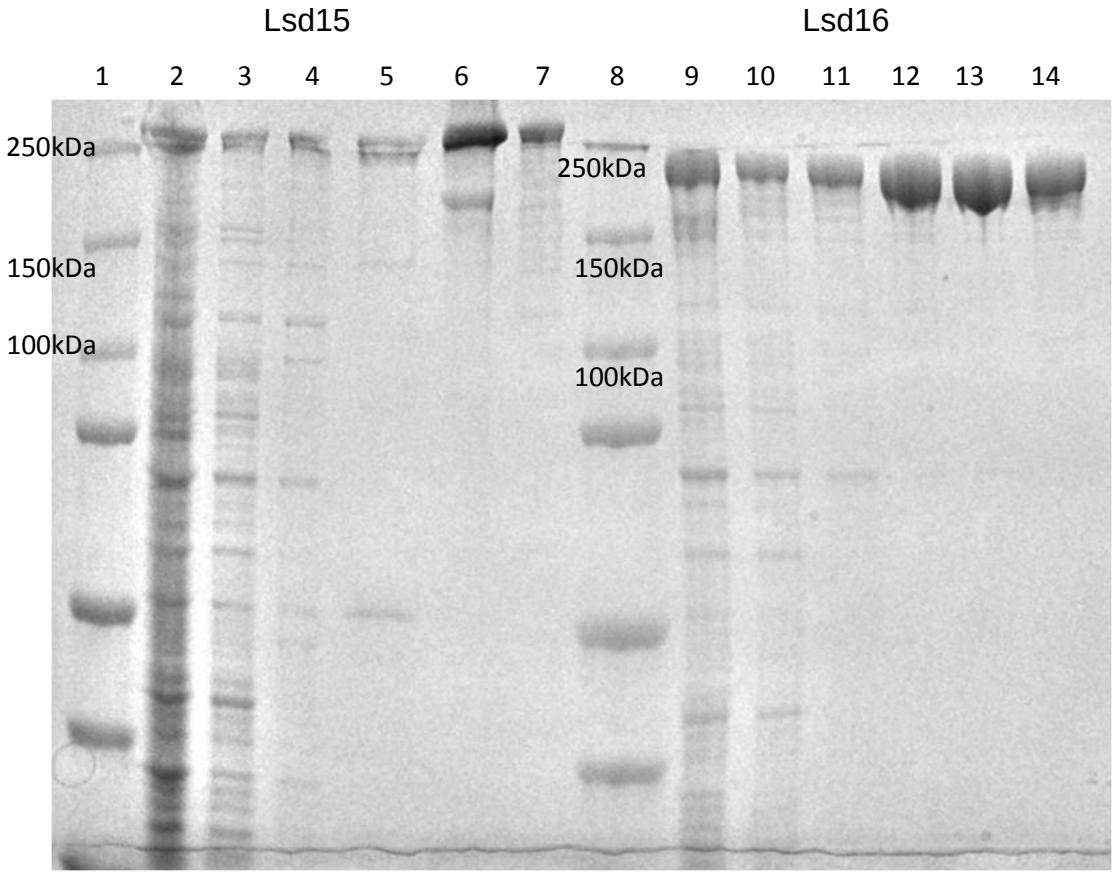


Figure 10. SDS PAGE analysis of purified Lsd15 and Lsd16. Lanes : 1,8 are protein molecular weight markers, 2- Lsd15 total cell lystate, 3- soluble fraction of total cell lysate, 4- IMAC load, 5- IMAC purified protein, 6- AEC purified protein, 7- SEC purified protein. 9-Lsd16 total cell lysate, 10-soluble fraction of cell lysate, 11-IMAC load, 12-IMAC purified protein, 13-AEC purified protein, 14-SEC purified protein

3.4. Crystallization of Lsd14 and preliminary diffraction study

Initial crystallization screening of Lsd13, Lsd15 and Lsd16 using commercial screens did not yield any crystal hits.

Reductive alkylation of lysine residues in Lsd16 was carried out in order to change protein properties (isoelectric point, solubility and hydropathy), which may promote crystallization via improved crystal packing(Kim et al., 2008). However, reductive alkylation of Lsd16 increased the in-homogeneity of the sample as shown in the SEC

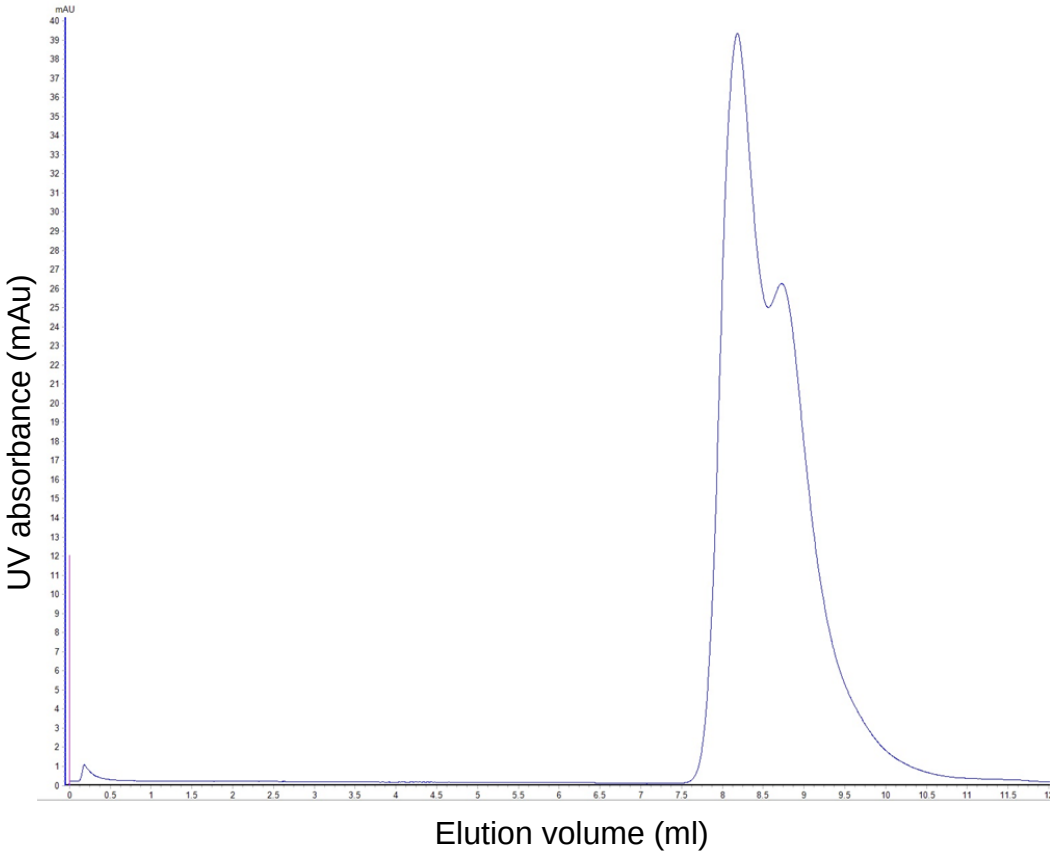


Figure 11. Size exclusion chromatogram of methylated Lsd16.

chromatograms in figure 11, 12 and 13. Hence the alkylated sample could not be used for crystallization experiments.

An initial crystal hit for Lsd14 was obtained (figure 14) which was further optimized to increase the size of the crystals. The optimization experiments involved screening of

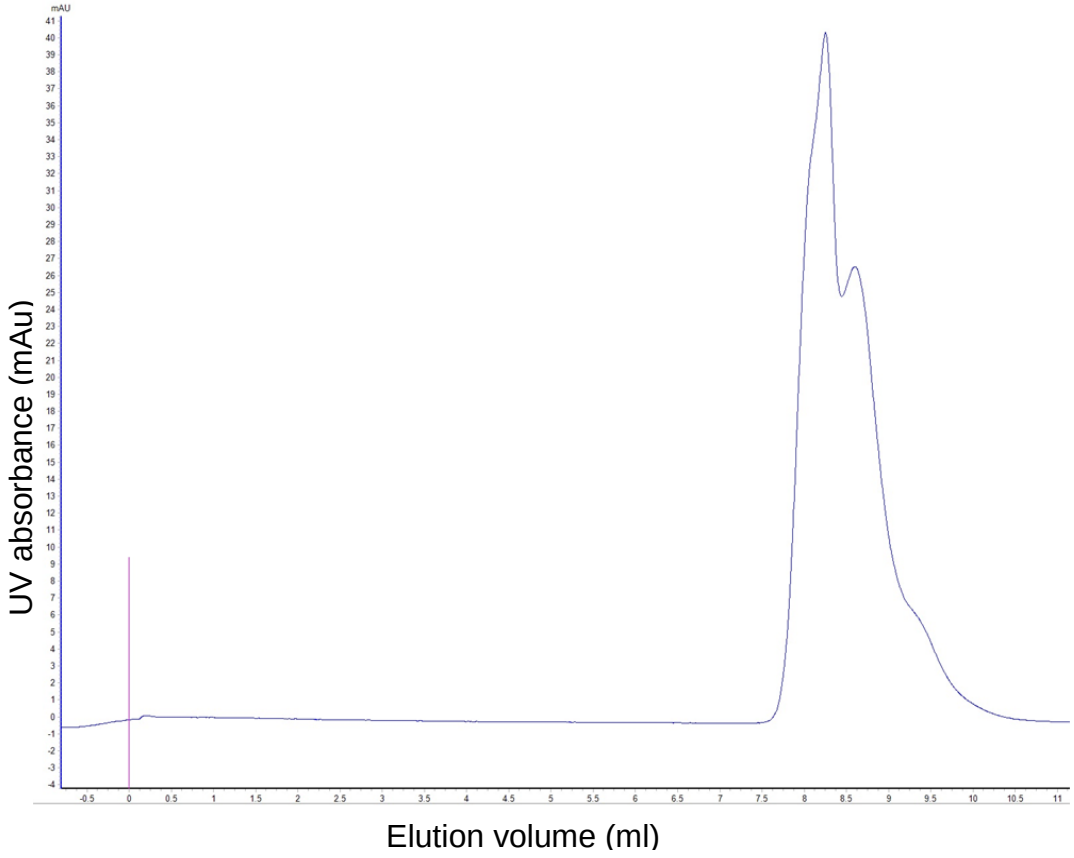


Figure 12. Size exclusion chromatogram of ethylated Lsd16.

different buffers and pH values ranging from 2 to 11 and screening of different salts included in the Hampton Research Stock Option pH kit and Salt kit. Additive screen and Silver bullets screen from Hampton Research were also screened to improve crystal quality. Including LiCl, changing the buffer component to Sodium acetate pH 4.0 in the condition for crystallization and addition of urea as an additive to the crystallization drop improved the size of the crystals. However, the crystals obtained were small (about 20 microns in size) and clustered. Single crystals were obtained upon addition of Silver bullets additive (Hampton Research) having components 0.2% w/v 1,2-Diaminocyclohexane sulfate, 0.2% w/v Diloxanide furoate, 0.2% w/v Fumaric acid, 0.2% w/v Spermine, 0.2% w/v Sulfaguanidine, 0.02 M HEPES sodium pH 6.8 as an additive to the crystallization drop as shown in figure 14 . However, these crystals were not

reproducible. Co-crystallization of Lsd14 NADP and NADPH was carried out since Lsd14 has a ketoreductase domain which uses NADPH to reduce a keto group to a hydroxyl group.

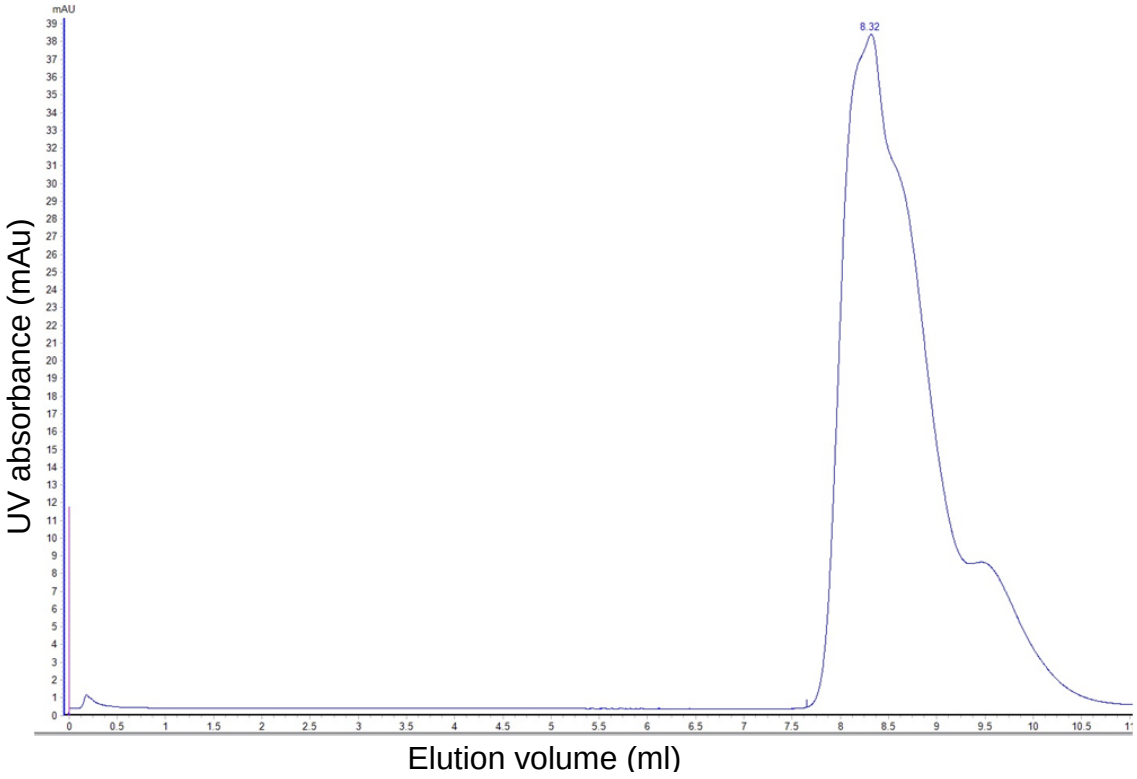


Figure 13. Size exclusion chromatogram of isopropylated Lsd16.

Single crystals were obtained by breaking the clusters as discussed in section 2.6.

The signal to noise ratio is reduced by the background noise contributed in the diffraction images by the beam cross section not intercepted by the crystal. Hence, Lsd14 crystals being small, were subjected to X-ray diffraction at the micro focus beamline BL12-2 at SSRL.

The data collection statistics for diffraction from Lsd14 crystal co-crystallized with NADP have been summarized in table 1. $CC_{1/2}$ indicates the percentage of correlation between intensities of random half data set. (Karplus and Diederichs, 2012). Mean I/σI gives signal to noise ratio and a value greater than equal to 2 is accepted.

The completeness of the dataset is low (85.6%) and hence more crystals will be screened to get a better data set with higher completeness (>90%). The low completeness could be

Table 1. Data Collection statistics (Values in parentheses correspond to last resolution shell)	
Cell dimensions	P1
a,b,c	61.70, 94.38, 111.69
α,β,γ	94.517, 90.586, 108.780
Resolution	39.09-1.6
CC(1/2)	99.6(78.6)
I/σI	10.94 (1.53)
Completeness (%)	85.6(81.5)

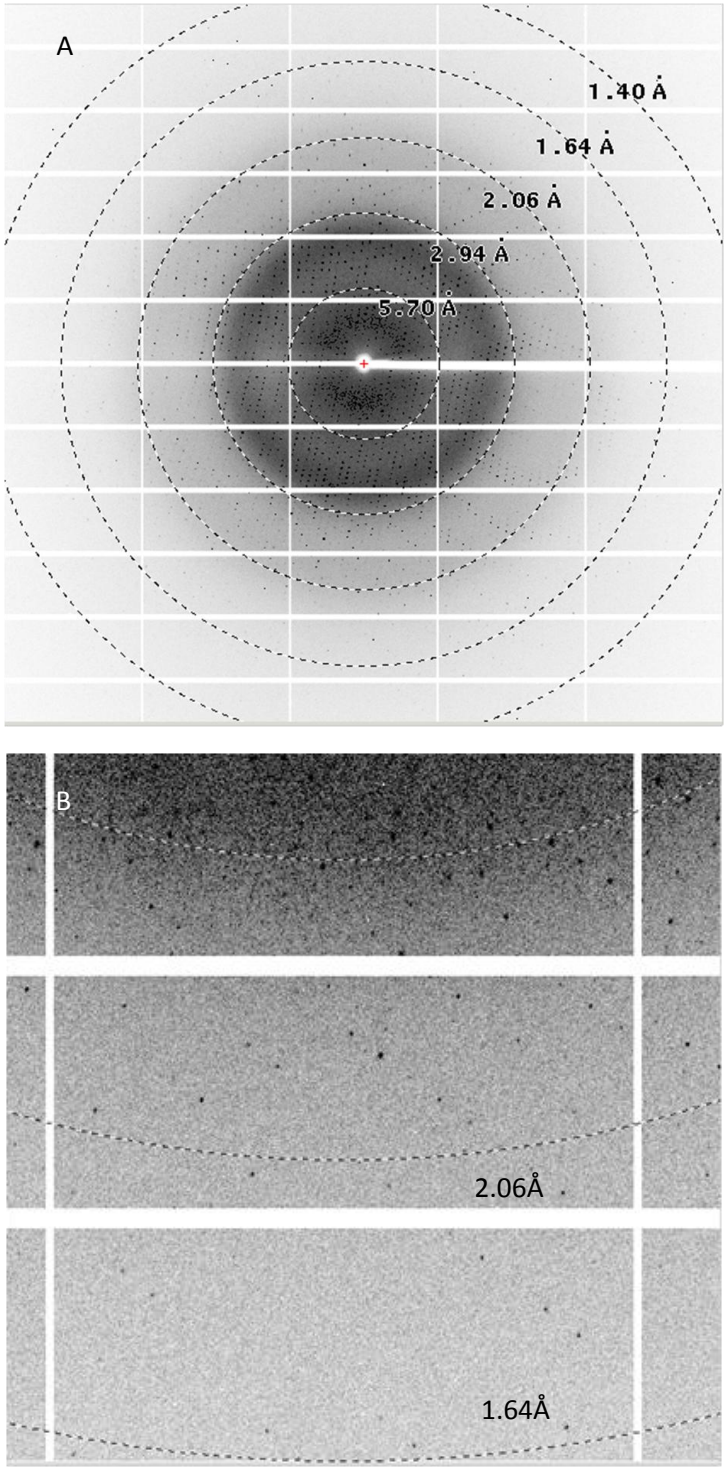


Figure 15. Diffraction images. A. Diffraction image for Lsd14-NADP crystal. B. Zoomed view to show diffraction spots in the highest resolution shell.

due to radiation damage to the crystal and hence, attempts to increase crystal size which can withstand more radiation damage should be carried out. Macroseeding and microseeding experiments will be carried out to increase the size of Lsd14 crystals or to get other crystal forms which may give a better dataset.

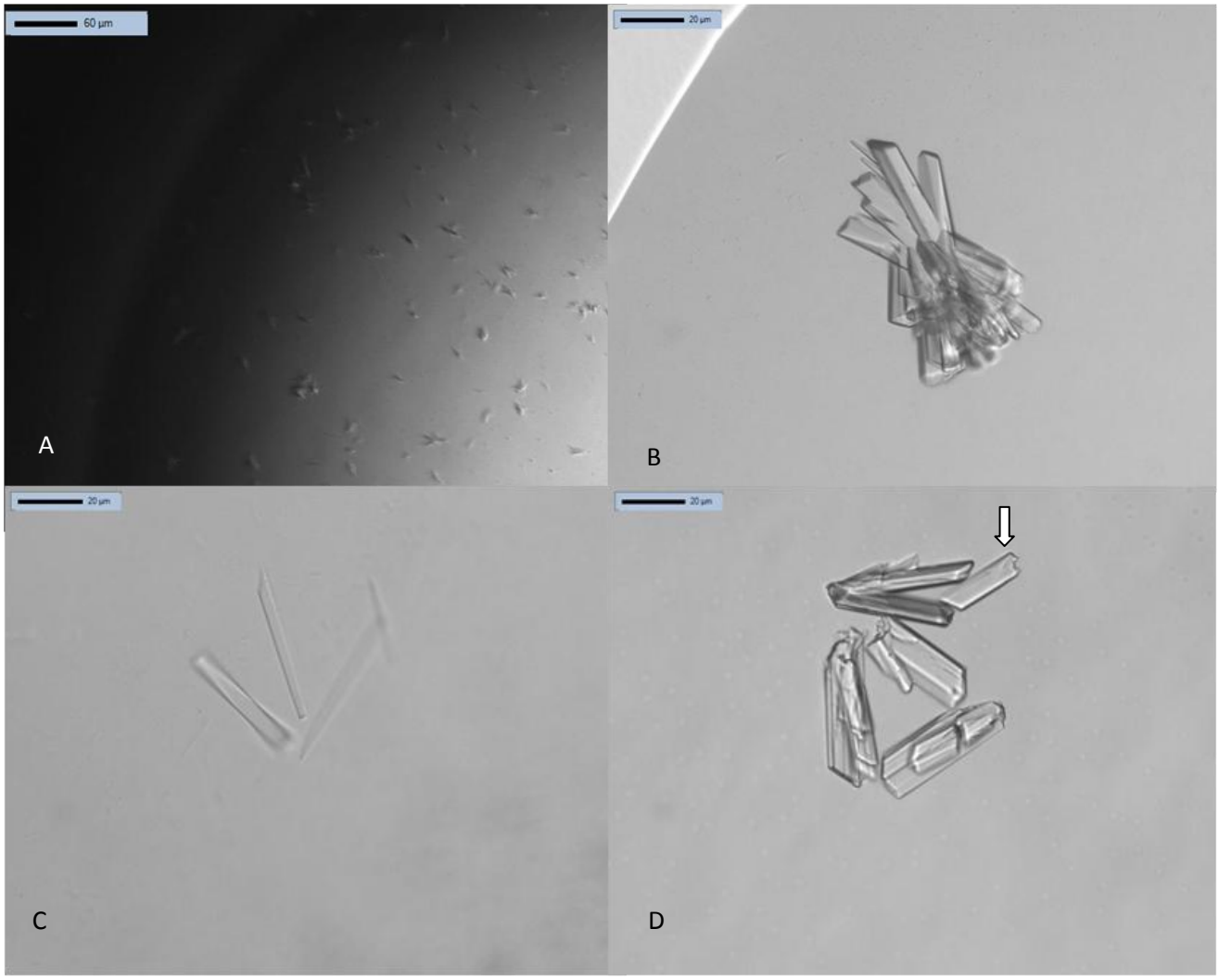


Figure 14. Lsd14 crystals. A. Crystals obtained in initial screening, B. Larger clusters of Lsd14 crystals obtained upon optimization, C. Single crystals obtained with silver bullet additive, D. A cluster of Lsd14 crystals broken to yield single crystal (shown using the arrow).

4. Conclusion and Future directions

This work led to successful crystallization and preliminary diffraction experiments towards atomic structure determination of polyketide synthase Lsd14. Further attempts to improve crystal quality and to get complete dataset should be carried out. Purification procedure to purify other polyketide synthase enzymes in the Lasalocid biosynthesis pathway including Lsd13, Lsd14 and Lsd16 to yield crystallization quality protein has been standardized. Attempts for co-crystallization of Lsd13 and Lsd14 should be carried out. The Lsd13-Lsd14 complex structure will provide structural insights on interaction of different modules in a modular polyketide synthesis pathway. This information will be helpful in designing systems where skipping or inclusion of one or more modules can be used to synthesize polyketide product analogs.

In-vitro reconstitution experiments to synthesize Lasalocid were not carried out due to time constraints and major bottlenecks including low expression of Lsd11 and Lsd12 and solubility problem of Lsd17. Nevertheless, the purification procedure of four out of the nine proteins involved in the biosynthesis pathway has been standardized in this work. The purification procedures of two more proteins in the pathway, Lsd18 and Lsd19, have been standardized by other colleagues in our group. Partial reconstitution of the pathway using purified Lsd11,Lsd12, Lsd13 and Lsd14 can be used to study catalytic activity of Lsd14 and also, substrate loaded enzyme can be obtained which can be used for structural studies. Hence, future work will involve attempts at purifying Lsd11, Lsd12 and Lsd17 so that in-vitro reconstitution experiments can be carried out.

5. References

Betancor, L., Fernández , M.-J., Weissman , K.J., and Leadlay, P.F. (2008). Improved Catalytic Activity of a Purified Multienzyme from a Modular Polyketide Synthase after Coexpression with *Streptomyces* Chaperonins in *Escherichia coli* . ChemBioChem *9*, 2962–2966.

Bevitt, D.J., Cortes, J., Haydock, S.F., and Leadlay, P.F. (1992). 6-Deoxyerythronolide-

B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. Eur. J. Biochem. *204*, 39–49.

Caffrey, P., Lynch, S., Flood, E., Finnan, S., and Oliynyk, M. (2001). Amphotericin biosynthesis in Streptomyces nodosus: deductions from analysis of polyketide synthase and late genes. Chem. Biol. *8*, 713–723.

CVM, and Onade (2006). Lasalocid sodium.

Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., and Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. Science *252*, 675–679.

Duax, W.L., Griffin, J.F., Langs, D.A., Smith, G.D., Grochulski, P., Pletnev, V., and Ivanov, V. (1996). Molecular structure and mechanisms of action of cyclic and linear ion transport antibiotics. Biopolymers *40*, 141–155.

Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007). Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. Nat. Prod. Rep. *24*, 162–190.

Hilgenfeld, R., and Saenger, W. (1982). Structural chemistry of natural and synthetic inoophores and their complexes with cations. (Springer, Berlin, Heidelberg), pp. 1–82.

Kabsch, W. (2010). *XDS*. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 125–132.

Karplus, P.A., and Diederichs, K. (2012). Linking Crystallographic Model and Data Quality. Science (80-.). *336*, 1030–1033.

Katsuyama, Y., and Ohnishi, Y. (2012). Type III Polyketide Synthases in Microorganisms. pp. 359–377.

Khosla, C., Herschlag, D., Cane, D.E., and Walsh, C.T. (2014). Assembly line polyketide synthases: mechanistic insights and unsolved problems. Biochemistry *53*, 2875–2883.

Kim, Y., Quartey, P., Li, H., Volkart, L., Hatzos, C., Chang, C., Nocek, B., Cuff, M., Osipiuk, J., Tan, K., et al. (2008). Large-scale evaluation of protein reductive methylation for improving protein crystallization. Nat. Methods *5*, 853–854.

Lomovskaya, N., Otten, S.L., Doi-Katayama, Y., Fonstein, L., Liu, X.C., Takatsu, T., Inventi-Solari, A., Filippini, S., Torti, F., Colombo, A.L., et al. (1999). Doxorubicin overproduction in Streptomyces peucetius: cloning and characterization of the dnrU ketoreductase and dnrV genes and the doxA cytochrome P-450 hydroxylase gene. J. Bacteriol. *181*, 305–318.

Migita, A., Watanabe, M., Hirose, Y., Watanabe, K., Tokiwano, T., Kinashi, H., and Oikawa, H. Identification of a Gene Cluster of Polyether Antibiotic Lasalocid from Streptomyces lasaliensis.

Robinson, J.A. (1991). Polyketide Synthase Complexes: Their Structure and Function in Antibiotic Biosynthesis. Philos. Trans. R. Soc. B Biol. Sci. *332*, 107–114.

Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D.J., and de Gier, J.-W. (2012). Optimizing Membrane Protein Overexpression in the Escherichia coli strain Lemo21(DE3). J. Mol. Biol. *423*, 648–659.

Singer, I.I., Scott, S., Kazazis, D.M., and Huff, J.W. (1988). Lovastatin, an inhibitor of cholesterol synthesis, induces hydroxymethylglutaryl-coenzyme A reductase directly on membranes of expanded smooth endoplasmic reticulum in rat hepatocytes. Proc. Natl. Acad. Sci. U. S. A. *85*, 5264–5268.

Staunton, J. (1998). Combinatorial biosynthesis of erythromycin and complex polyketides. Curr. Opin. Chem. Biol. *2*, 339–345.

Staunton, J., Weissman, K.J., Fujii, I., Shibuya, M., Ebizuka, Y., Horinouchi, S., Rasmussen, B., Lamzin, V.S., Kunau, W.H., Wierenga, R.K., et al. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.

Washington, J.A., and Wilson, W.R. (1985). Erythromycin: a microbial and clinical perspective after 30 years of clinical use (1). Mayo Clin. Proc. *60*, 189–203.

Weissman, K.J. (2009). Chapter 1 Introduction to Polyketide Biosynthesis. pp. 3–16.

Wong, F.T., Hotta, K., Chen, X., Fang, M., Watanabe, K., and Kim, C.-Y. (2015). Epoxide Hydrolase–Lasalocid A Structure Provides Mechanistic Insight into Polyether Natural Product Biosynthesis. J. Am. Chem. Soc. *137*, 86–89.