Studies on Assembly and Regulation of Cysteine Regulatory Complex of Bacterial Pathogens

THESIS

SUBMITTED TO THE
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE FACULTY OF SCIENCE

CSIR-IMTECH

RAHISUDDIN

CSIR-INSTITUTE OF MICROBIAL TECHNOLOGY
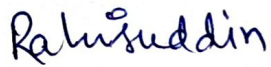SECTOR 39-A, CHANDIGARH, INDIA
2020

Certificate

The research work embodied in this thesis entitled **"Studies on Assembly and Regulation of Cysteine Regulatory Complex of Bacterial Pathogens"** has been carried out by Mr. Rahisuddin under the supervision of Dr. S. Kumaran at the CSIR-Institute of Microbial Technology, Sector 39A, Chandigarh. This work is original and has not been submitted in part or full for any other degree or diploma to any other University.

28/12/2020

Dr. S. Kumaran

Supervisor

Rahisuddin

This thesis is dedicated

to my family…

Acknowledgments

"So which of the favors of your Lord would you deny?" …………Al Quran

With all humility, I acknowledge The Almighty Allah and owe my happiness and success to his kind providence.

I feel immense bliss in writing my sincerest gratitude to every person who was a part of my endeavor and has helped and supported me directly or indirectly in achieving this greatest accomplishment of my life. I dedicate this piece of research to my parents and my wife, who supported and believed me during this arduous journey.

Firstly, I wish to express my deepest and sincere gratitude to my Ph.D. supervisor, Dr. S. Kumaran, who was abundantly helpful and offered invaluable support and guidance throughout the Ph.D. tenure whenever I needed it the most. I thank him for painstakingly correcting my thesis. I couldn't have imagined having a better supervisor and mentor for my Ph.D. I admire him for his patience, motivation, and immense knowledge. As a mentor, he gave me enormous freedom and opportunities and contributed a lot of his intellect and time whenever I needed it. He always encouraged and helped me to achieve my thesis on time. This work would not have been possible without his constant support and expert guidance. He has guided me in all imaginable ways and every conceivable aspect of my research work. His scientific inventiveness, analytical power, and trouble-shooting skills always helped me during my entire Ph.D. tenure. I want to thank him for believing in me and for teaching me to have faith in myself. I'll always cherish the time spent with him and the long conversations we had on random topics. His friendly nature could make anyone feel so comfortable in the lab. He always allowed me to visit my parents and

wife. Most sincere thanks to him for providing constant support and encouragement.

Great admiration and respect for the former director, Dr. Anil Koul, acting director, Dr. Manoj Raje, and present director, Dr. Sanjeev Khosla for maintaining excellent research infrastructure at IMTECH. Special thanks for all the support extended by him to promote the eco-friendly campus. I thank our warden, Dr. Deepak Sharma for maintaining a wonderful living atmosphere in the hostel. I am deeply thankful to Dr. Charu Sharma, Dr. Balwinder Singh, and Dr. S. Karthikayan, coordinator JNU –IMTECH Ph.D. program, and the staff for taking care of academic formalities during my entire tenure.

A special thanks to the Ph.D. coursework faculties especially Dr. Ashwani Kumar, Dr. Ashish Ganguli, Dr. Pawan Gupta, Dr. Javed Agrewala, Dr. Beena Krishnan, Dr. Barnali Choudhari, Dr. Pradip Sen, and more who helped me to understand the concepts to a deep and to perform quality research. I also want to thank Dr. Krishan Gopal and Dr. S. Karthikayan for maintaining the X-ray facility and providing solutions for the instrument-related problems with patience. I thank them for the help he provided me in looking at my data a few times after the X-ray data collection. A Special thanks to Dr. Srikrishna Subramanian for maintaining a 3D server facility in the protein center. The 3D facility made my life so simple.

I thank my lab seniors, Dr. Abhishek, with whom I worked for almost one year at the beginning of my tenure. He has an excellent work ethic. He graced me with several biophysical techniques and I believe I learned most from him. The way he works in the lab has always inspired me. He used to come early to the lab almost every day on time, and he inspired me to do the same. I have always found him to be a good at heart person even though he had been quite moody at times. Dr. Ravi, another senior at the time I joined the lab, is the coolest and decent person I would ever know. I especially want to thank him for getting me familiar with the software used in crystallography. We shared many

memories from coffee time, pizza time. I will cherish the memories of the parties we used to have with you guys.

I especially thank my present lab member, Neha, for being there in the time of need. We worked on projects and papers together, and it was a pleasure working with her. All the time when we had long arguable discussions regarding some techniques or results or planning an experiment, it just added on to my knowledge. It was great working with you, and I wish you a bright future ahead. I also want to thank my lab members Madhuri, Narender, and Nancy, for being a part of this journey.

I would like to express my gratitude to my Seniors; Dr. Zeeshan Bhai, Dr. Faraz Bhai, Dr. Shoiab Bhai, Dr. Salman Bhai, Dr. Bhagi Bhai Dr. Sajid Nadeem Bhai Dr. Sajid Zilli Bhai, Prashant Sir, Dr. Sourav, Dr. Anil Patidar, and more who always gave me the right suggestions and having fun with them. I especially thank Dr. Rehan Bhai (My roommate for one year), who helped me a lot during my admission time and whose conversations with me always inspired me. I especially thank Naushad (Chacha) and Hilal, my junior-cum friends, who remained part of my IMTECH journey from having Saturday's dinner to going outside.

I would also like to thank my batchmates; Yogita, Vaidhavi, Urvashi, Komal, Shelly, Priyanka, Naveen, Deep Jyoti (DJ), Harpreet, Rajesh, Vinod, Alka, Kanti, Jitu, Srajan, Poushali, Debhargya, Drishti, Heena for being a part of this journey. I would like to thank my juniors and friends; Lucky, Jagriti, Anu, Madhumita, Sonal, Harsh, Sumit, Shukla Ji, Manish, Yachna, and others who always supported me directly or indirectly.

A special thanks to my bachelors' group from Aligarh Muslim University; Waqar Bhai, Anas Bhai, Zohaib Bhai, Sadique Bhai for sharing good times and having discussions on a lot of topics whenever I visit Aligarh.

I warmly thank all the technical staff, bioinformatics center for the excellent internet facilities, stores, and purchase section, the accounts division, and

administration especially Rajinder Sir, PTM, Library, guest house for their timely help and cooperation. Special thanks to Mr. Deepak Bhatt for maintaining the DNA sequencing facility efficiently.

I would like to thank all the labs on our third-floor GNRPC, especially Dr. Deepak Sharma's lab members for calling us out to be a part of their lab parties.

I owe much gratitude to my parents, sister, brothers, and my extended family. No words can express my appreciation for the unadulterated and unconditional love that my parents have given me. They have always been with me through the ups and downs of my life and when I needed them most. I thank them for all of the unforgettable sacrifices that they made on my behalf.

At last, I do not know how to begin by saying thank you to my soul mate, my wife, and my best friend, Samreen. I thank you for everything, for being so understanding, for being my strength and a constant source of encouragement, and for having gratitude for my parents throughout my Ph.D. tenure. I am genuinely thankful for having you in my life. I wish my child, Haaziq, very successful and long life ahead. I also thank Farhan, Saba, Farha for sharing good memories with them throughout my tenure.

I would also like to thank everyone who supported me in writing and encouraged me to strive towards my goal. Finally, I gratefully acknowledge the financial support from DBT, INDIA.

Rahisuddin

Symbols and abbreviations used in the study

Mw	Molecular weight		**UV**	Ultraviolet
α	Alpha		**OD**	Optical density
β	Beta		**pmol**	Pico mole
γ	Gamma		**RNA**	Ribonucleic acid
λ	Lambda		**rpm**	Revolution per minute
µg	Microgram		**RT**	Room temperature
µL	Microliter		**s**	second
µM	Micromolar		**SDS**	Sodium dodecyl sulfate
µm	Micrometer		**OD**	Optical density
mL	Milliliter		**L**	Liter
mg	Milligram		**°C**	Degree Celsius
mM	Millimolar		**CS**	Cysteine Synthase
ng	Nanogram		**CRC**	Cysteine Regulatory Complex
nM	Nanomole		**SAT**	Serine Acetyl Transferase
%	Percent		**SV**	Sedimentation Velocity
aa	Amino acid		**SEC**	Size exclusion chromatography
cm	Centimeter		**UV**	Ultraviolet
kDa	Kilo Daltons		**CD**	Circular Dichroism
DNA	Deoxyribonucleic acid		**PDB**	Protein Data Bank
EDTA	Ethylene Diamine tetra acetic acid		**Tris**	Trishydroxylmethyl-aminomethane
et al.	et alia (and others)		**Kb**	Kilobase
etc.	Et cetera (and so forth)		**S$_{20,w}$**	Normalized Sedimentation coefficient to water at 20 ºC

Table of Contents

Symbols and abbreviations used in the study

Chapter 1

Review of Literature

Chapter 2

Materials and Methods

Chapter 3

Biochemical, structural and kinetic characterization of serine acetyl transferase from *Mycobacterium tuberculosis*

Chapter 4

Characterization of thermodynamic and kinetic properties of CRC assembly.

Chapter 5

Understanding key regulatory features of CRC by solution and structural approaches

Chapter 1

Review of Literature

Review of Literature

1.1 Complex biological systems

Complex systems, made up of multiple components or parts, perform new functions that cannot be performed by any one of the constituents by itself. The new emergent properties arise out of specific interactions between components that work in a coordinated manner to carry out emergent new functions (Buldyrev et al., 2010). Complex systems, natural or manmade, share many designs or patterns that are built to predict structural and functional dynamics in response to spatio-temporal changes in the environment (Ma'ayan, 2017). From the mechanistic point of view, emergent properties of complex biological systems such as multiple protein assemblies and protein-DNA assemblies can be traced to coordinated and specific interaction patterns among the constituents. Components of a complex system are highly interdependent and this interdependency translates into nonlinear output and emergence of new properties. The cooperative interaction between components confers "self-regulating" abilities to the complex which "self-corrects" through multiple mechanisms. One of the most well-characterized "self-correcting" mechanisms in biology is correction through feedback inhibition (Gerhart~ & Pardee$, 1962).

Often biological complexes are integrated into the respective pathway with multiple feedback loops in order to constantly interact with the environment and make adjustments (**Figure 1.1**). A small change in the initial value can cause a large change in the overall output of the complex through the feedback loop. The scale and magnitude of feedback induced sensitivity are the key response balance mechanisms to adjust to the changes in the environment. Therefore, studying and uncovering design principles of coordinated interactions between multiple components within the complex will allow us to understand the molecular origins of emerging properties. Further, the understanding of emerging properties can enhance our understanding of how life operates at the cellular level and reveal mechanisms that can be repurposed in vitro for carrying out many useful tasks. Biochemical signaling networks and multi-cascade-multi-enzyme complexes are a few of many natural complex systems that depend on "self-sustaining" feedback loops for processing environmental cues (Wei et al., 2020).
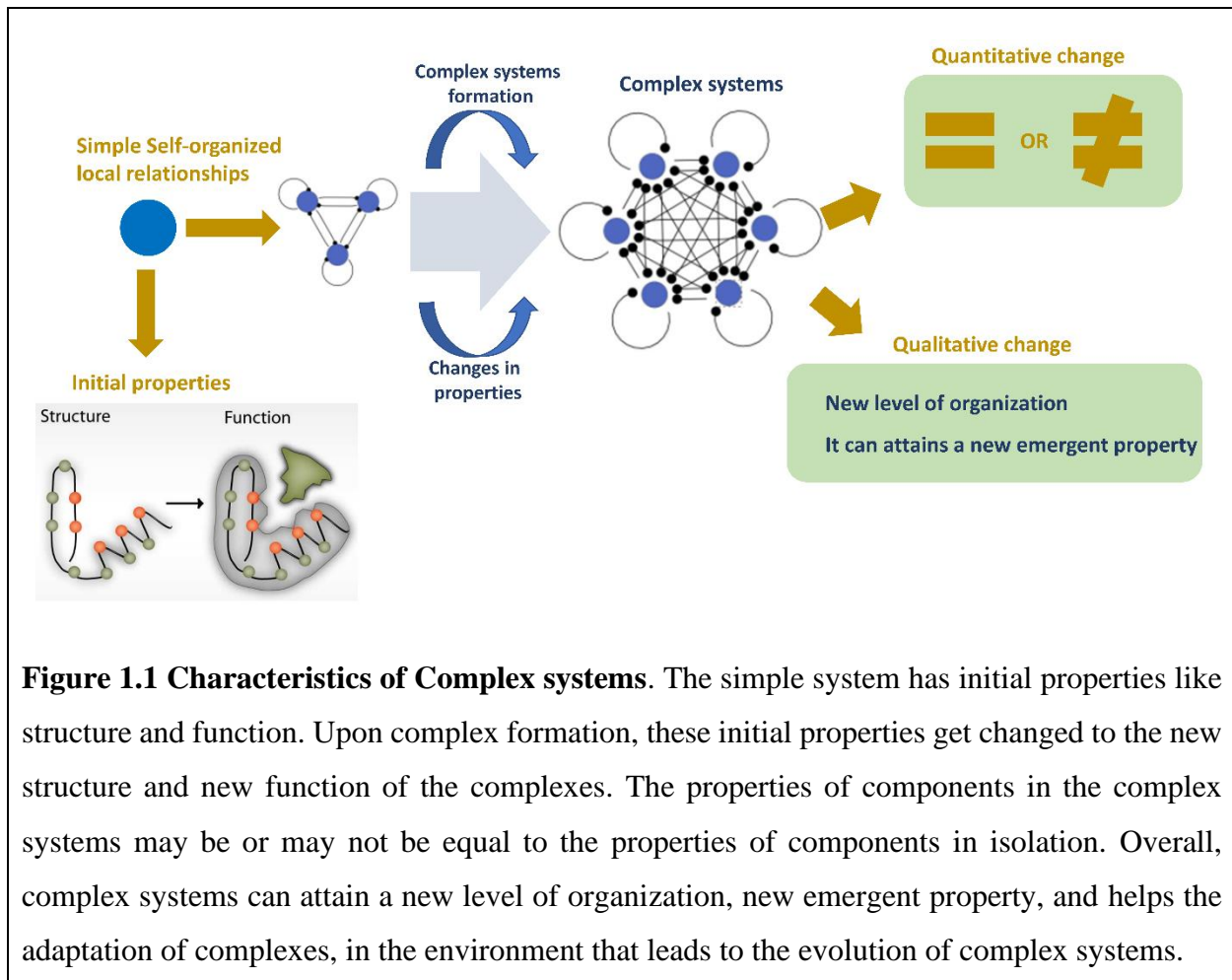
Figure 1.1 Characteristics of Complex systems. The simple system has initial properties like structure and function. Upon complex formation, these initial properties get changed to the new structure and new function of the complexes. The properties of components in the complex systems may be or may not be equal to the properties of components in isolation. Overall, complex systems can attain a new level of organization, new emergent property, and helps the adaptation of complexes, in the environment that leads to the evolution of complex systems.

One example of a complex system whose emergent properties have been studied extensively is cellular automata. In a cellular automaton, a grid of cells, each having one of the finitely many states, evolves according to a simple set of rules (Goldenberg & Efroni, 2001). These rules guide the "interactions" of each cell with its neighbors. Although the rules are only defined locally, they have been shown capable of producing globally interesting behavior. When emergence describes the appearance of unplanned order, it is spontaneous order (in the social sciences) or self-organization (in physical sciences). Spontaneous order can be seen in herd behavior, whereby a group of individuals coordinates their actions without centralized planning. Self-organization can be seen in the global symmetry of certain crystals, for instance, the apparent radial symmetry of snowflakes, which arises from purely local attractive and repulsive forces both between water molecules and between water molecules and their surrounding environment.

1.2 Multiple enzyme complex systems

Enzymes that catalyze different steps of a pathway or different pathways physically associate to form multi-enzyme complexes that exhibit emergent properties that none of the component enzymes alone possess (Proschel et al., 2015; Y. H. P. Zhang, 2011). Emergent properties which are mostly regulatory or responsive, in nature, control the structure and functional dynamics of component enzymes in response to changes in cellular conditions (Walde et al., 2014). Specific protein-protein interactions between enzymes drive the complex formation and guide the assembly and orientation of functional and regulatory sites. Coordinated and cooperative interactions between monomeric units lead to new emergent properties that control the underlying cellular processes. Complex cellular processes such as signal transduction and metabolism are controlled by multi-protein assemblies which form and dissociate dynamically to amplify and transmit the signals received at the cell surface or changes in the cellular environment (Brunton et al., 2004).

Studies on few well-characterized multi-enzyme assemblies such as tryptophan synthase, fatty acid synthase, and pyruvate dehydrogenase complex show that the efficiency of final product formation is increased in the complex (Domingo et al., n.d.; Hydesg et al., 1988; Lane et al., 1984; Leibundgut et al., 2008; Sweetlove & Fernie, 2018). The increased product formation efficiency results from improved stabilities and catalytic efficiencies of component enzymes of the complex (**Figure 1.2**). The stability of the complex itself is sensitive to changes in the levels of metabolites inside the cell and hence behave like sensors of metabolite concentration by assembling and dissociating in response to changes of the metabolites(Srivastava & Bernhard, 2015). Further, the structure and function of complexes are differently sensitive to both natural and synthetic inhibitors as compared to the sensitivities of component enzymes (Vinitsky et al., 1992). It has been observed that component enzymes after complex formation become more resistant to allosteric inhibitors.
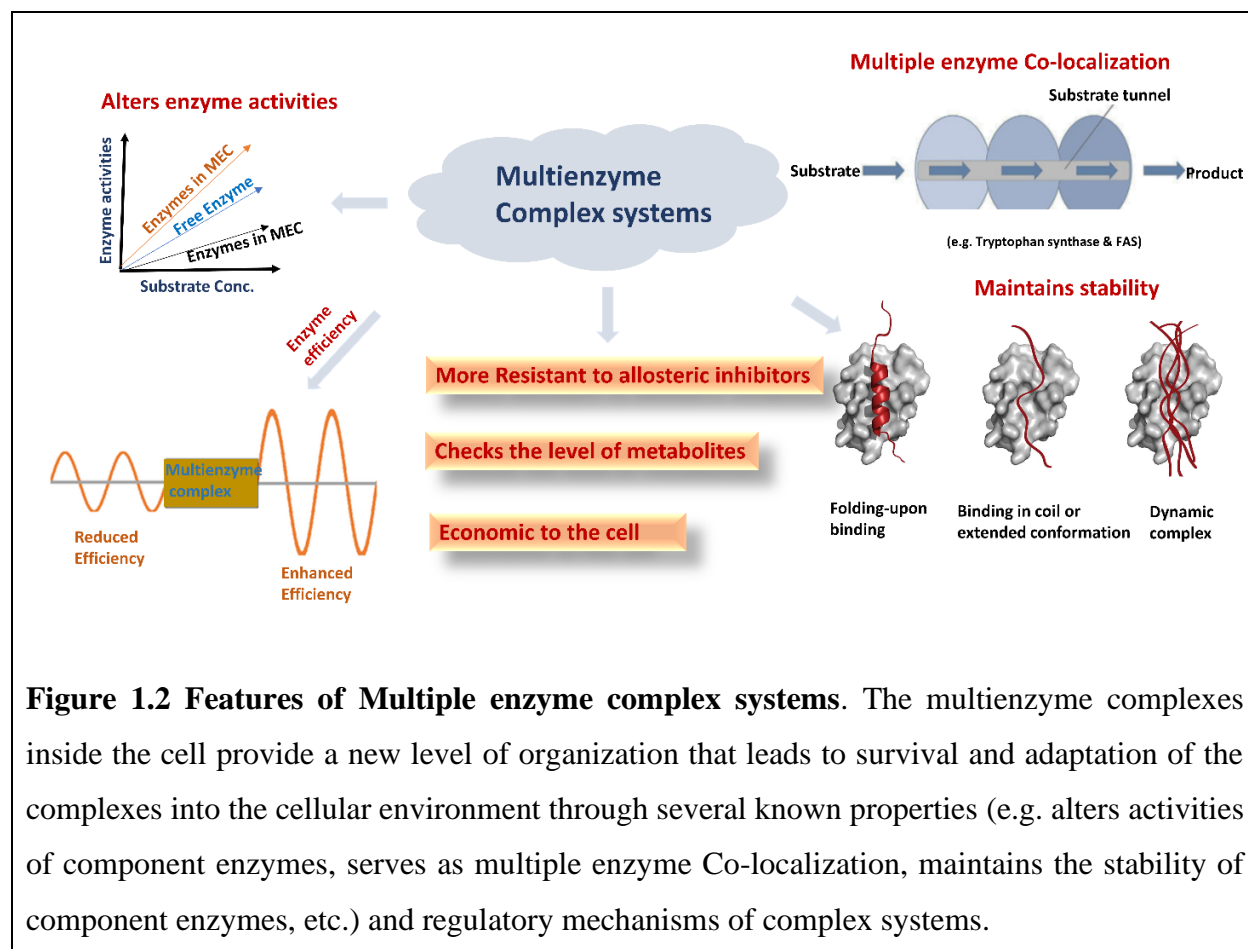
Figure 1.2 Features of Multiple enzyme complex systems. The multienzyme complexes inside the cell provide a new level of organization that leads to survival and adaptation of the complexes into the cellular environment through several known properties (e.g. alters activities of component enzymes, serves as multiple enzyme Co-localization, maintains the stability of component enzymes, etc.) and regulatory mechanisms of complex systems.

1.3 *Substrate/metabolite channeling, a new emergent property of multi-enzyme assemblies*

Assembly rules of multi-enzyme entities guide reaction centers of component enzymes to align channels or tunnels that increase the efficiency of transfer of metabolites/intermediates between enzymes and subsequent enhancement in the final product formation (**Figure 1.2**) (X. Huang et al., 2001; Y. Zhang & Hess, 2017). Tryptophan Synthase (TS), a multi-enzyme complex composed of two enzymes that catalyze the last two steps of tryptophan synthesis, is one of the well-characterized examples of substrate channeling (Dunn, 2012). The metabolite indole is transferred from one enzyme to another via interconnecting ~ 25 Å long tunnel. Metabolic tunneling between two active sites increases the effective local concentration of indole and also prevents toxic and labile indole molecules from diffusing away and be lost.

FAS, another multi-enzyme complex that consists of two identical 272 kDa multifunctional polypeptides catalyzes fatty acid synthesis. During the repeated fatty acid chain elongation step, substrates are handed over from one functional domain to the next to facilitate the sequential

actions of ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains. The growing fatty acid chain is released by the action of a thioesterase (TE) upon reaching a carbon chain length of 16 carbons (palmitic acid) (Chakravarty et al., 2004). Like in the TS, intermediates are efficiently passed onto the next active site, increasing their availability and enhancing the rate of reactions.

The Pyruvate Dehydrogenase Complex (PDH), a heterotrimeric complex, decarboxylates pyruvate to produce acetyl-CoA. Pyruvate, an important metabolite of the glycolysis cycle, is moved to mitochondria by pyruvate translocase for acetyl-CoA synthesis (Krivoruchko et al., 2015; van Rossum et al., 2016). PDH, localized inside the mitochondrial matrix, synthesizes acetyl-CoA from pyruvate in a sequence of steps. PDH consists of pyruvate dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2), and dihydrolipoamide dehydrogenase (E3), three enzymes that convert pyruvate into acetyl-CoA (Chandrasekhar et al., 2013; Hiromasa et al., 2004; Patel et al., 2014; Smolle et al., 2006). Structural and analytical approaches provided insight into the arrangements of different subunits and active sites. 48 subunits of E2 and 12 units of E3 form the core, which is surrounded by 60 subunits of the E1 enzyme. Half-life reactivity studies examined by steady-state and fast-kinetic studies showed evidence for product transfer between different subunits (Nemeria et al., 2004; Seifert et al., 2006). One of the multienzyme complexes which are formed by the two enzymes, Serine acetyltransferase (SAT) and Cysteine synthase (CS), is Cysteine regulatory complex (CRC), and both SAT, as well as CRC, are the main topics of study in this thesis.

1.4 The Cysteine regulatory complex

The CRC was initially purified from *Salmonella typhimurium* by Kredich in 1969 from the crude extract of cells following traditional methods like using a homogenizer, ammonium sulfate precipitation followed by density gradient centrifugation. The supernatant was assayed for CS activity and it was deduced from a previous study that out of the total CS activity, 5% activity was because of the CRC complex. The complex was purified from gel filtration chromatography and its molecular weight was experimentally determined to be ~ 309 kDa. In the presence of O-acetyl serine (OAS) at a concentration of 10^{-4} to 10^{-3} M, the complex dissociated reversibly into 1 molecule of active SAT (Mw ~160 kDa) and 2 molecules of CS (Mw ~68 kDa) (Kredich et al., 1969; Kredich & Tomkins, 1966a). Even though the stoichiometry for this complex was determined to be 1 molecule of hexameric SAT binding to 2 molecules of dimeric CS, this

stoichiometry was different for the CRCs isolated from other systems. In published studies of the CS-SAT peptide complexes, it was shown that 1 molecule of dimeric CS can bind to two molecules of C-terminal SAT peptide at each of its active sites (Salsi, Bayden, et al., 2010; Wang & Leyh, 2012). Formation of CRC has been tested by assembling the purified proteins (SAT and CS) and also by purifying them as CRC (Berkowitz et al., 2002a; Campanini et al., 2005a; Droux et al., 1998; Francois et al., 2006; Kumaran et al., 2009; Mino et al., 2000). *In vitro* characterization of the interaction between these two enzymes has been studied in many systems using many different biochemical and biophysical techniques like pull-down assay, size exclusion chromatography, fluorescence spectroscopy, yeast two-hybrid studies, surface plasmon resonance, isothermal titration calorimetry, etc. (Berkowitz et al., 2002b; Bonner et al., 2005; Campanini et al., 2005a; Mino et al., 2000; Zhao et al., 2006). These studies have established that the last 10-20 amino acids at the C terminal of the SAT protein are sufficient for complex formation. Full-length protein-protein interactions for SAT-CS have been extensively characterized for *Arabidopsis thaliana, Glycine max, Haemophilus influenzae,* and *Escherichia coli* using a variety of experimental approaches. The consensus of the CRC stoichiometry has been shown to be a ratio of 1:2 i.e., one hexamer of SAT binding to two dimers of CS (*E coli, Salmonella, Haemophilus*). This may not be true globally as one homo tetrameric SAT can bind to two CS dimers (spinach) and one SAT trimer can also bind to three CS dimers (soybean) as shown by biophysical solution studies (Droux et al., 1998; Kumaran et al., 2009). The lack of structural evidence for CRC in the absence of crystal structures may indicate that the CRC stoichiometry depends on the oligomeric state of SAT. The molecular weight of CRC determined by sedimentation equilibrium studies ranged from 295 kDa (*E. coli*) to 310 kDa (*Salmonella*) to 330 kDa (*Glycine max*).

1.5 Structural and molecular features of component enzymes of CRC

Serine acetyltransferase (SAT)

SAT, (EC 2.3.1.30) carries out the first step of cysteine biosynthesis. SAT is a member of the bacterial O-acetyl transferases subfamily and possesses a left-handed β-helix, formed by repeating hexapeptide motifs, a repeating signature in which residue i is aliphatic, i + 1 is usually glycine, and i + 4 is a small residue, thus has [LIV]-[GAED]-X2-[STAV]-X motif (Gorman & Shapiro, 2004a; Pye et al., 2004a). The entire known SATs share the structural homology with a variety of acyltransferases (Thoden et al., 2009, 2012). The C-terminal tail of the SAT is highly conserved, flexible, unstructured, and susceptible to proteolytic cleavage. (Mino et al., 2000). The crystal

structures of Serine acetyltransferases have been resolved as both hexamer (**PDB ID: 1S8O**) and trimer (**PDB ID: 1T3D**) (Gorman & Shapiro, 2004b; Kumar et al., 2014; Olsen et al., 2004) from different organisms. The hexamers are formed by two dimers of trimers arranged in a head-to-head orientation at their N-terminal (**Figure 1.3**)**.** Both forms share the same topology and similar three-dimensional structures. Structurally the cross-section of the trimer displays triangular shape symmetry along the 3-fold axis. The three parallel β strands of monomers are folded into a helix by hydrophobic interactions to form the symmetric trimer. The first crystal structure reported for the SAT was resolved from *E. coli.* The structure is composed of two domains: α helical domain consisting of eight α helices and β helical domain consisting of fourteen β strands constructing the five coils of the helix. The crystal structure of the SAT enzyme shows that it exists in trimeric as well as hexameric oligomeric states in different species (e.g., *Entamoeba histolytica, Escherichia coli* respectively). The *St*SAT and *Hi*SAT exclusively exist as stable hexamers (Gorman & Shapiro, 2004b; Olsen et al., 2004). Whereas, the *Gm*SAT is reported to function as either trimeric or hexameric proteins depending upon the protein concentrations (Yi et al., 2013a). However, the *Eh*SAT exists exclusively as a trimer. It is suggested that the trimeric nature of *Eh*SAT is attributed
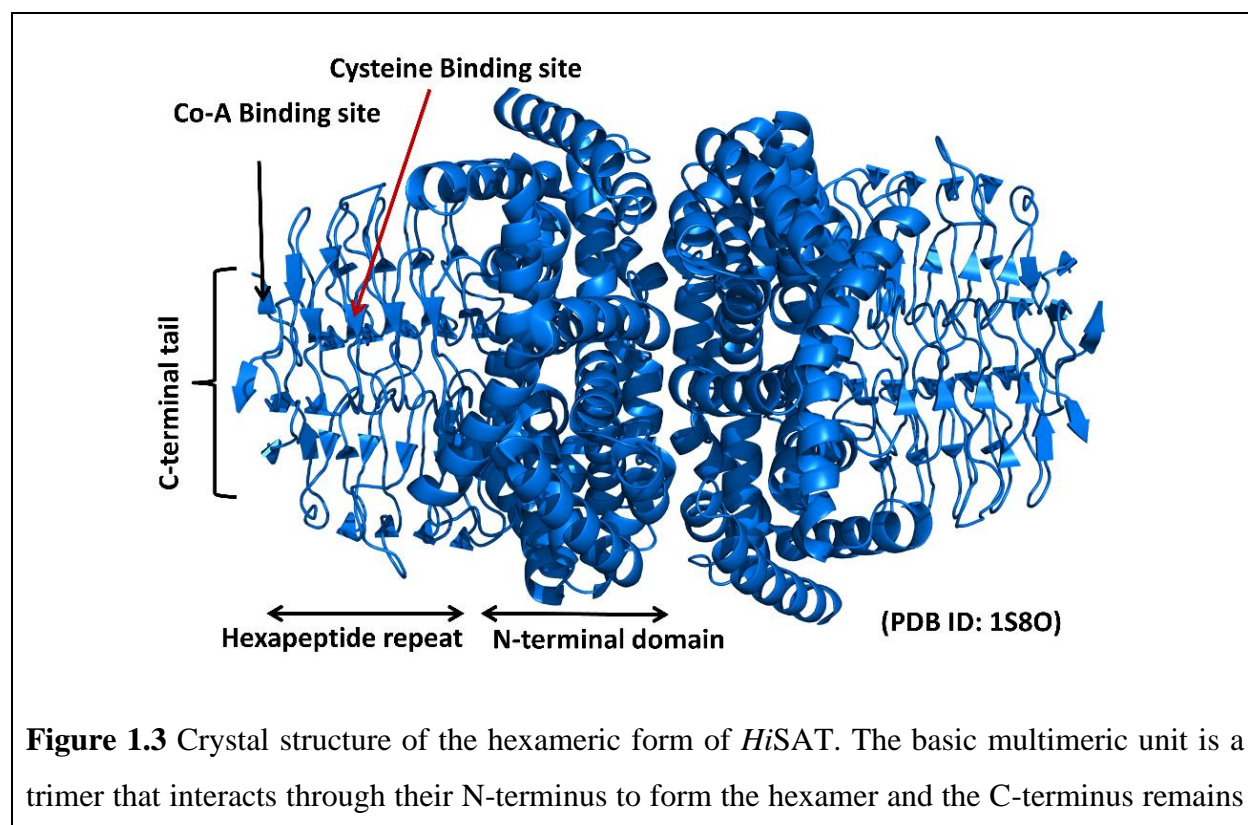
Figure 1.3 Crystal structure of the hexameric form of *Hi*SAT. The basic multimeric unit is a trimer that interacts through their N-terminus to form the hexamer and the C-terminus remains

free to interact with CS enzyme and with different metabolites present inside the cell. C-terminus has a hexapeptide repeat motif that consists of an active site and cysteine binding site.

to its N-terminal domain which shows dissimilar N-terminal sequence compared to *Ec*SAT/*Hi*SAT (Kumar et al., 2011a). It has also been shown that the hydrophobic surface area of *Ec*SAT /*Hi*SAT in this region is large in comparison to *Eh*SAT, which attributes to the formation of the trimer.

Cysteine synthase (CS)

Cysteine synthase (EC 2.5.1.47) has a two-domain structure, an N-terminal (1-145 residues) and a C-terminal (146-311 residue) domain. CS across species share ~ 60% similarity in amino acid sequences and share a similar topology of folds and similar 3-dimensional structure. All the CS are typical α, β structures with almost 70% α helices and 30% β sheets with β folds in the center of each domain flanked by α folds. The active site is present 20 Å deep into the channel, flanked by the periphery of the N-terminal and C-terminal domain. A shallow pocket is created by the contribution of residues from α-helix2 (Asparagine loop or "TSGNT loop" in *Hi*CS) and creates
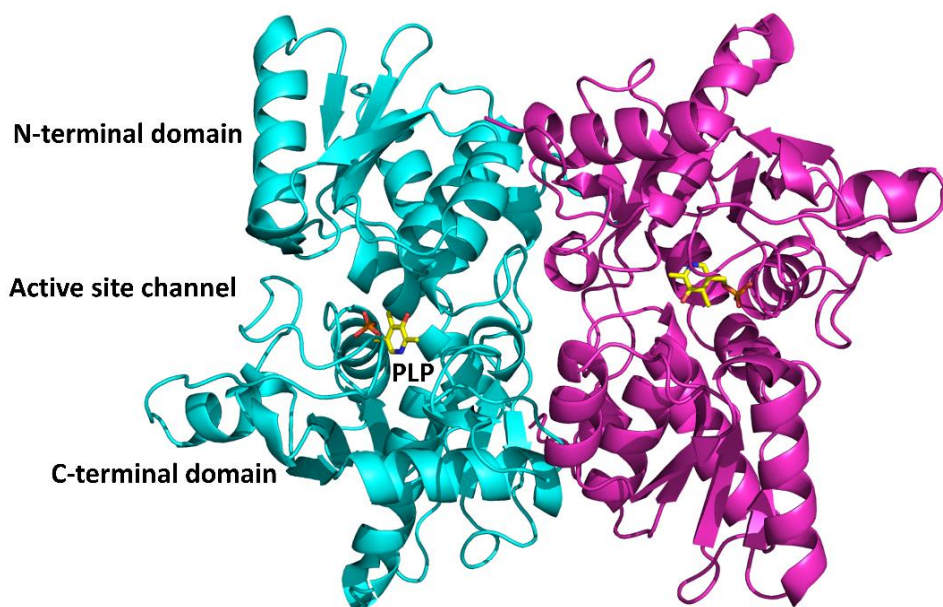
Figure 1.4 Crystal structure of *Salmonella typhimurium* CS dimer enzyme at 2.2 Å resolution. The subunits consist of N-terminal and C-terminal domains and one PLP moiety is found buried inside the active site between the N-terminal and C-terminal domain of each subunit.

an amenable environment for the binding of the α-carboxylate group of OAS through hydrogen bonds and electrostatic interactions (Burkhard et al., 1998; Rabeh & Cook, 2004). CS exists as a homodimer in the solution and has a molecular weight of ~66 kDa. The structure of *Hi*CS is solved as a monomer in the asymmetric unit, whereas the *St*CS is a dimer that displays the arrangement of two monomers.

The monomers are arranged such that their active site exposes the bulk solvent for the accessible entry of substrates. The CS has a cofactor pyridoxal 5′-phosphate (PLP) bound to a (lysine) K42 as an internal aldimine at the active site. PLP is stabilized by 8-9 hydrogen bonds mostly contributed by the residues surrounding the active site (Burkhard et al., 1998). The large anionic charge of the phosphate moiety is neutralized by the dipole of helix 7 in the C-terminal domain, one of the three-helix dipoles directed toward the active site (Cook et al., 1992). The active site mutant of *St*CS (K42) does not hamper cofactor binding to the enzyme (Rege et al., 1996) and protein still appears yellow and was found to be present as external Schiff base with amino acid methionine, present in the growth medium of the *S. typhimurium.* The crystal structure of covalently bound L-methionine as an external aldimine to the PLP in the K41A mutant of CS shows that external aldimine formation induces a large conformational change in the N-terminal domain of the protein. Methionine mimics the action of the substrate OAS during catalysis and results in the closure of the active site (Burkhard et al., 1999a).

1.6 Kinetic and regulatory features of component enzymes of CRC

The kinetic studies on *St*SAT (*S. typhimurium*) have suggested a ping-pong reaction mechanism (Leu & Cook, 1994). SAT catalyzes two half-reactions; in the first half of the reaction, an acetyl-enzyme intermediate is formed at the active site, and in the second half, L-serine binds and condenses with acyl-enzymes intermediate to form OAS. However, studies on *Ec*SAT have strongly supported the steady-state random order mechanism which involves the formation of the productive ternary complex without any covalent enzyme-substrate complex (Hindson & Shaw, 2003). The active site of SAT is located towards the C-terminal ends with the residues from the monomer interfaces contributing to the active site. Each trimer forms three catalytic sites.

The activity of SAT is regulated by cysteine the end product of the biosynthetic pathway by the competitive and non-competitive inhibition mechanisms with an inhibition constant of KI ~1.0 μM (Hindson, 2003; Kredich et al., 1969; Kumar et al., 2011a; Leu & Cook, 1994). The comparison of acetyl-CoA and cysteine bound structure from, *Ec*SAT and *Hi*SAT reveal that

cysteine binds to the same site as serine, explaining the feedback inhibition of SAT. The crystal structures of *Ec*SAT and *Hi*SAT show that cysteine negatively regulates the activity of SAT by directly competing with serine and induces allosteric conformational changes, which reduces the binding affinity of acetyl-CoA (Olsen et al., 2004; Pye et al., 2004b). The structure of *Hi*SAT suggests that His-154 functions as a general base whereas His-189 and Asp-139 contribute to catalysis (Guan et al., 2008). Cysteine accommodates itself into the small cleft formed between the two adjacent monomers, a site very near to the catalytic center and binding of cysteine buries the acetyl-CoA binding site near to the active site. However, such an effect is not observed in *Eh*SAT and was found to be insensitive to the cysteine inhibition. The loop region in the third coil of the beta-helix interacts with the C-terminal region and upon cysteine binding, it buries the acetyl-CoA binding site, therefore indirectly or allosterically affects the activity of the enzyme. However, the *Eh*SAT has an additional 8 residues in the third coil and also C-terminal region folds differently in *Eh*SAT, which results in fewer interactions and this leaves the C-terminal region of *Eh*SAT highly disordered thus making the acetyl-CoA site accessible to the bulk solvent (Kumar et al., 2011a).
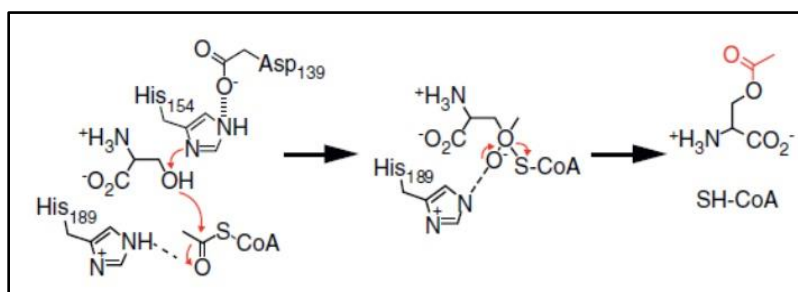
Figure 1.5 The *Hi*SAT enzyme catalysis mechanism. *Hi*SAT structure suggested that His154 functions as a general base with His189 and Asp139 contributing to catalysis. His154 as the acceptor of a proton from the β-hydroxyl of serine as the tetrahedral intermediate is formed following a nucleophilic attack on the thioester carbonyl of acetyl-CoA. His189 contributes to substrate binding and orientation for efficient catalysis and Asp139 interacts with His154 to enhance basicity (Guan et al., 2008).

CS a PLP dependent enzyme catalyzes the penultimate reaction step of cysteine biosynthesis. CS catalyzes a beta-elimination reaction in which the α, β-acetoxy group of O-acetyl-L-serine (OAS) is eliminated and replaced by a bisulfide group to give L-cysteine and acetate as products (Rabeh & Cook, 2004). The versatile property of PLP chemistry makes it an important cofactor in diverse reactions ranging from transamination, decarboxylation, deamination, and racemization. PLP binds to its substrate, then acts as an electrophile and stabilizes the carbanion intermediates formed during the catalysis (**Figure 1.6)**. CS exists in two isoforms encoded by genes cysK and cysM, also known as CS-A and CS-B. Both the isoforms share approximately ~30% amino acid sequence similarity and display different substrate specificity for the sulfide donor group. CS-A utilizes the reduced bisulfide (HS-) whereas CS-B in addition to reduced bisulfide (HS-), can also utilize thiosulfate to synthesize cysteine thiosulfonate. However, the conversion of cysteine thiosulfonate to cysteine is not reported in *E. coli*. (Agren et al., 2008; Nakamura et al., 1984a; Sekowska et al., 2000a). CS-B preferentially utilized under anaerobic conditions for cysteine biosynthesis in *E. coli* and *S. typhimurium* (Mino & Ishikawa, 2003). Structural studies on the CS-B shows the difference in the types of amino acids in the active tunnel widen the active site entrance to accommodate a broad range of nucleophiles (Sulfur donor group) (Chattopadhyay et al., 2007; Claus et al., 2005). An important study from *M. tuberculosis* shows that the synthesis of cysteine by CS-B is very different from CS-A. CS-B instead of utilizing OAS shows specificity for O-phosphoserine (Agren et al., 2008). CS-B does not directly utilize reduced Sulphur (HS-), instead, it depends on the product of another gene *cysO* which is thiocarboxylated at the C-terminus and serves as the vehicle for the intermediate adduct of cysteine bound to the C-terminus of CysO which is later cleaved by Zn2+ dependent protease into the final product cysteine (Jurgenson et al., 2008). The fact that this metalloprotease is specific for the cleavage the gene coding for the CysM, CysO and the metalloprotease is located on the same site in the genome. The CS-A is present in excess over CS-B, therefore, serves as the major enzyme for the cysteine biosynthesis. Studies using various isotopic and spectroscopy techniques, the kinetic mechanism of CS-A and CS-B has been established to follow the ping-pong Bi-Bi reactions mechanism. The CS is a tryptophan synthase beta-subunit like pyridoxal 5´ phosphate (PLP) dependent enzyme that gives characteristic fluorescence spectra at 507 nm upon excitation at a λ max of 412 nm. Even though CS do not share a significant sequence identity, with the beta subunit of tryptophan synthase, the considerable similarity is seen in the overall folds of both the enzymes.
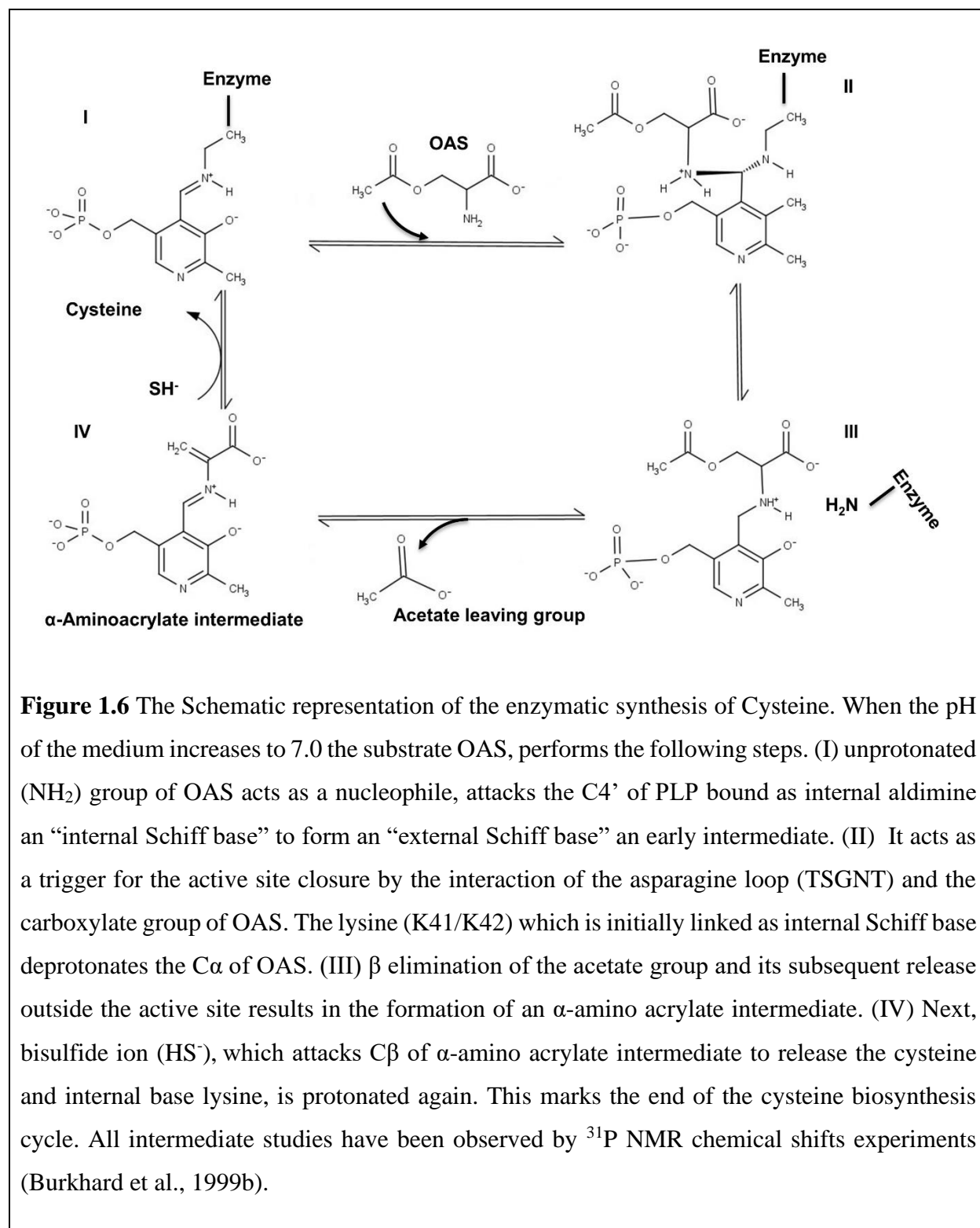
Figure 1.6 The Schematic representation of the enzymatic synthesis of Cysteine. When the pH of the medium increases to 7.0 the substrate OAS, performs the following steps. (I) unprotonated (NH_2) group of OAS acts as a nucleophile, attacks the C4' of PLP bound as internal aldimine an "internal Schiff base" to form an "external Schiff base" an early intermediate. (II) It acts as a trigger for the active site closure by the interaction of the asparagine loop (TSGNT) and the carboxylate group of OAS. The lysine (K41/K42) which is initially linked as internal Schiff base deprotonates the Cα of OAS. (III) β elimination of the acetate group and its subsequent release outside the active site results in the formation of an α-amino acrylate intermediate. (IV) Next, bisulfide ion (HS^-), which attacks Cβ of α-amino acrylate intermediate to release the cysteine and internal base lysine, is protonated again. This marks the end of the cysteine biosynthesis cycle. All intermediate studies have been observed by ^{31}P NMR chemical shifts experiments (Burkhard et al., 1999b).

1.7 Salient features of Cysteine regulatory complex

Both SAT and CS physically associate to form CRC complex under normal conditions. SAT is known to be less stable but protein-protein interaction between SAT and CS enhances SAT stability resulting in the increased catalytic activity (Mino et al., 2001). Therefore, one of the physiological advantages of CRC is assumed to be enhancing SAT turnover even though improvements are only moderate (Kumaran et al., 2009). The other physiological significance is proposed to be metabolic channeling but no study has thus far shown any conclusive evidence for metabolic channeling in CRC. High resolution structural and analytical biochemistry approaches are needed to collect substantial evidence whether OAS, the product of SAT is channeled to the active site of CS or not. CRC also acts as the sensor of metabolites. It checks the concentration level of OAS, NAS, and Sulphate inside the cell.

The catalytic activity of SAT is sensitive to cysteine, the end product of the pathway as well as an allosteric inhibitor of SAT but it is not known whether SAT in complex with CS in CRC will exhibit a similar magnitude of allosteric inhibition to cysteine. Structural and biochemical studies showed that the last 10 amino acids of the C-terminal of SAT bind into the active site tunnel of CS. This finding led to the assumption that binding of SAT C-terminal may block access to OAS, the substrate of CS (Campanini et al., 2005b). Steady-state kinetics studies of CS, performed in the presence of either SAT C-terminal or full-length SAT, confirmed the observations of the structural studies and clearly show that binding of SAT C-terminal inhibits cysteine synthesis activity of CS (Droux et al., 1998; Schnell et al., 2007). Therefore, protein-protein interactions between CS and SAT have a negative effect on the cysteine flux in the cell. SAT is the natural inhibitor of CS and the physiological significance of this protein-protein interaction relationship remains unknown. In the complex, cysteine biosynthesis activity reduces significantly to ~75 % (Droux et al., 1998), and SAT activity is expected to be modulated to balance the negative consequences of protein-protein interactions on CS activity (Saito et al., 1995).

The activity of the SAT within the bi-enzyme complex was observed to increase only 2-4-fold (Kumaran et al., 2009). Several studies have attempted to address the nature and biochemical features of protein-protein interactions between CS and SAT (Francois et al., 2006; Schnell et al., 2007). These studies used various biochemical and biophysical methods to understand the molecular signatures of CRC formation in plants and bacteria (Bogdanova and Hell, 1997; Jez & Dey, 2013; Kredich et al., 1969; Williams et al., 2009). Both thermodynamic and kinetic studies

have been used to quantitate the affinity of the interaction between these two enzymes. The affinity of CS towards full-length SAT as well as binding to the C-terminal tail (last 10 residues) of SAT has been extensively investigated for plant and bacterial systems (Kumaran & Jez, 2007; Wirtz et al., 2010). Results of all studies converge towards a high-affinity model for CS and SAT

Figure 1.7 Characteristics features of CRC complex and its component enzymes, SAT and CS are illustrated. The new emergent properties of CRC are highlighted in the bottom-middle box.
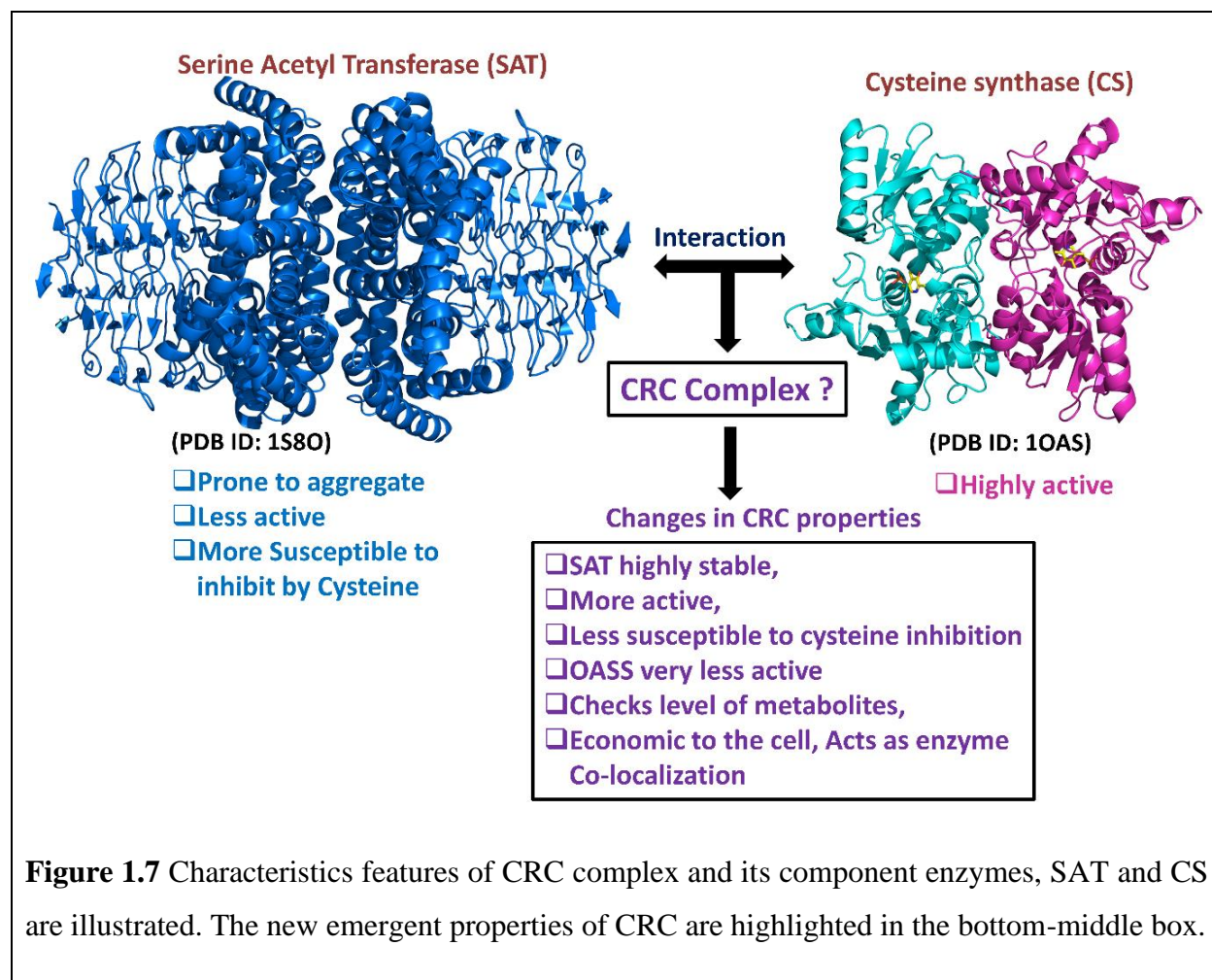
interaction. The affinity of binding of one molecule of CS dimer to one SAT oligomer (homohexamer or a trimer depending on the species) is found to be in the range of 0.1 to 4 nM (Berkowitz et al., 2002b; Kumaran et al., 2009). Similarly, equilibrium binding experiments have been used to determine the affinities of CS with the C-terminal tail of the SAT. Results of these experiments also showed that the C-terminal of SAT binds to the CS active site with a high affinity which is in the range of 1.0 to 10.0 nM (Francois et al., 2006). However, the affinity of CS to SAT C-terminal for *Haemophilus influenzae* was significantly lower (~ 900 nM) than the other species, and the basis for this low-affinity interaction is not known (Campanini et al., 2005a). In summary,

interactions between SAT and CS from across the species are shown to be high-affinity interactions, suggesting that SAT is a high-affinity natural inhibitor of the CS. On the contrary, steady-state kinetic studies that have characterized CS from plants and bacteria have shown that the substrate affinity of CS (Michaelis-Menten Constant Km) is in the range of 1.0 to 3.0 mM, almost 5-6 log-fold lower than the affinity of the natural inhibitor, SAT (Warrilow & Hawkesford1, 2000; Wirtz et al., 2004). Signatures of protein-protein and protein-peptide (CS and SAT C-terminal) have revealed that most of the binding free energy is derived from the C-terminal. The crystal structures have also provided the evidence that peptide binds to the active site of each monomer and the last isoleucine is locked in a series of hydrogen bonds and hydrophobic interactions with active site residues of CS (Francois et al., 2006; Mino et al., 1999; Schnell et al., 2007). It should also be noted that the activity of SAT is sensitive to cysteine concentration. Cysteine, the end-product of the biosynthetic pathway, inhibits SAT activity suggesting that SAT is an important regulatory enzyme that is at the branch point of nitrogen and sulfate metabolisms (Hussain et al., 2009).

1.8 Regulation of CRC assembly

A few studies analyzed the assembly state of CRC by biochemical methods and proposed a CRC model in which one hexameric SAT is bound to two dimers of CS (Kredich et al., 1969). Studies also showed that CRC dissociates in the presence of OAS and it has also suggested that CRC dissociation is favored under sulfur limiting conditions as the OAS levels will increase when sulfur levels are low (Wang & Leyh, 2012). To elucidate the mechanism of *St*CRC assembly, disassembly, and regulatory features of CRC in detail, our lab undertook a systematic study using high-resolution analytical approaches like analytical ultracentrifugation velocity studies and kinetic approaches. The study showed that CRC exists in at least two major assembly states, with CS and its cognate substrate OAS controlling the assembly and disassembly of CRC (Kaushik et al., 2017). As the CS concentration is increased at fixed SAT concentration, CRC assembly is promoted and it forms CRC1 and CRC2 complexes. The low molecular weight ~310 kDa CRC1 complex is composed of one SAT hexamer and two CS dimers whereas the high-molecular-weight ~450 kDa CRC2 complex is composed of one SAT hexamer and four CS dimers (Kaushik et al., 2017).

Figure 1.8 A dynamic model for CRC association and dissociation. Cartoon representations of CS (**PDB ID: 1OAS**) and SAT (**PDB ID: 1S80**), generated from structural coordinates of homologous of *Salmonella typhimurium* and *Hemophilus influenzae* enzymes respectively were used to describe CRC assembly and disassembly cycle. (F) shows forward reaction and (R) shows reverse reactions. **Step1**, F: free SAT hexamer and free CS dimer enzymes interact to form CRC1 complex (in presence of a low concentration of OAS and an abundant amount of CS). Two of six high-affinity binding sites of SAT are occupied. R: CRC1 complex dissociates when OAS levels rise (high OAS concentration and low amount of CS). **Step2**, F: two more CS dimer binds to CRC1 complex to form CRC2 complex (in presence of a low concentration of OAS and an abundant amount of CS). Four of six high-affinity binding sites of SAT are occupied in the CRC2 complex state. R: CRC2 dissociates when OAS levels increase. Complete dissociation to free enzymes occurs when OAS levels increase significantly high (>8 mM) under sulfate limiting conditions (diagonal arrows). **Step3**, F: two more CS dimer binds to CRC2 complex to form CRC3 complex (in presence of a low concentration of OAS and an abundant amount of CS). All six high-affinity binding sites of SAT are occupied in the CRC3 complex (CRC3 may be favored only if: 1-CS levels rise very high with a significant reduction in OAS

level or high positive cooperativity between adjacent sites in SAT. R: two free sites within the CRC2 complex pose a steric hindrance problem for additional CS binding or negative cooperativity that's why the CRC3 complex assembly is not formed (Kaushik et al., 2017).

The association and dissociation cycles of SAT-CS is correlated with changes in the OAS concentrations, indirectly to changes in sulfur flux. The high concentration of OAS dissociates the CRC complex into free SAT and free CS in a stepwise manner. First, it dissociates the 1:4 CRC2 complex into 1:2 CRC1 complex then further increase in OAS concentration leads the CRC1 complex to dissociate into free enzymes. The second metabolite which also regulates CRC assembly formation is the end product of the pathway that is cysteine. Cysteine inhibits the SAT enzyme activity resulting in less OAS product formation ultimately stabilizes the CRC assembly. However, there are no reports of cysteine effect on the oligomeric state of the CRC complex.

1.9 Specificity mechanism of Substrate recruitment by CS

Although many structural and biochemical studies contributed to the understanding of CRC assembly and modulation of enzyme activities of both enzymes, many other important questions remain unresolved. For example, substrate or metabolite channeling within CRC and the quaternary structure of CRC remained unknown for a long time. Another counter-intuitive property that puzzled many in the field is how CS recruits its substrate inside the cell as it has a very high affinity for its natural inhibitor SAT as compared to its affinity for its substrate, OAS. CS displays 4-6 log fold higher affinity towards SAT and therefore, it was not known how the low-affinity substrate, OAS will gain access to the active site of CS inside the cell. In the absence of any metabolite channeling and relatively high concentrations of CS in the cell, promotes an idea that supports the existence of an unknown mechanism by which CS may bind the low-affinity substrate in the presence of a high-affinity natural inhibitor. Our lab used an integrated approach using a variety of high resolution structural and analytical methods and discovered a non-canonical "competitive-allostery mechanism by which CS selectively binds its substrate while facilitating the dissociation of the bound inhibitor (Singh et al., 2017). We resolved a series of binary and ternary high-resolution crystal structures and showed that both substrate and inhibitor co-localize within the same active site of CS. The in-coming substrate, upon binding to the entrance of the active site channel, ~20 Å away from the reaction center, invokes allosteric changes at the site of

binding that propagate to the reaction center where the C-terminal ILE of natural inhibitor (SAT) is tightly bound and breaking interactions between the substrate/inhibitor binding loop and the bound inhibitor. We demonstrated that the substrate-induced facilitated dissociation of inhibitor does not follow the classical competitive-dissociation model but rather follows the non-classical competitive dissociation model in which the dissociation rates are insensitive to substrate concentrations (Singh et al., 2017). In the follow-up study, a systematic study from our lab traced the molecular origins of the "competitive-allosteric" mechanism by demonstrating that M120 residue located at the entrance is the key substrate sensor and is connected to the inhibitor/substrate
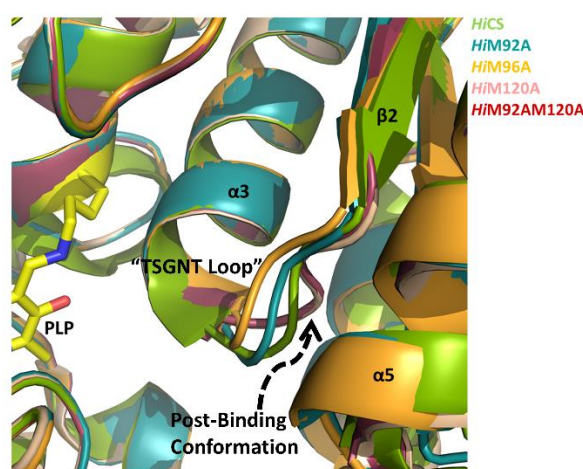
Figure 1.9 Structure alignment of wild type CS to different methionine mutants from the *Hemophilus influenzae* showing the "TSGNT loop" shift in the post inhibitor binding conformation in M120A and M92AM120A mutants. In the wild type CS, the pre-inhibitor binding loop conformation is responsible for the high affinity of wild type CS towards the natural inhibitor.

binding loop through an allosteric network (Kaushik et al., 2020). Mutation of M120 reduces the binding affinity of the substrate to CS and we showed that M120 facilitates CS to selectively bind substrate even when the inhibitor is bound. In that study, we generated different Methionine mutants to know the structural changes in the CS enzyme. We found that in M120 mutant the "TSGNT loop" moves toward the post inhibitor binding conformation reducing the affinity about 10^{-2} folds for the natural inhibitor in the M120A mutant. Secondly, the affinity of M120A towards

the OAS reduced about 10^{-3} folds in the M120A mutant. Combining the two factors, M120 contributes 10^{-5} folds affinity in favor of the 'substrate recruitment' and 'inhibitor dissociation' mechanism.

1.10. Enzymes involved in Cysteine biosynthesis pathway of Mycobacterium tuberculosis

The formation of cysteine is the final step in the sulfur assimilation pathway and it serves as the sulfur donor for the synthesis of all sulfur-containing metabolites in the cell (Sekowska et al., 2000b). This includes the biosynthesis of cofactors like thiamine, biotin, S-adenosyl methionine, molybdopterin, lipoic acid, providing sulfur to iron-sulfur clusters, several antibiotics, and the cellular redox-glutathione or mycothiol (*Mycobacterium tuberculosis*). The classical cysteine biosynthesis in bacteria, including *Mycobacterium,* begins with the assimilation of sulfate from the surroundings. The sulfate is transported into the cell by sulfate permeases or transporters, CysTWA, SubI, and ABC transporter complex in the membrane (Wooff et al., 2002). The inorganic sulfate is unreactive and therefore, it is activated by the enzyme ATP sulfurylase (CysDN), resulting in the activated form of the sulfate, adenosine-5- phosphosulfate (APS). The activated APS lies at a metabolic branch point from where it enters two different physiological routes. In the first route, APS is phosphorylated at the 3'- hydroxyl by APS kinase, (CysC) to form 3'-phosphoadenosine-5'-phosphosulfate (PAPS), the universal sulfate donor for sulfotransferases. In the second route, APS is reduced by APS reductase (CysH) to form sulfite which is further reduced to sulfide (S^{2-}) by sulfite reductase (nirA) (H37Rv). This sulfide is then incorporated into cysteine by the enzymatic action of CS; which is commonly referred to as the 'classical or conventional route of cysteine biosynthesis'.

The third route of sulfate reduction is considered as the 'classical or conventional route of cysteine biosynthesis'. In this route, first, acetyl-CoA-dependent enzyme serine acetyltransferase (SAT or CysE) (Rv2335) condenses an acetyl group from acetyl CoA on the hydroxyl group of serine, to form OAS. In the second step, OAS undergoes a β-replacement reaction to be finally converted to cysteine, via the elimination of acetate and addition of S^{2-} by the final enzyme of this pathway, PLP dependent CS-A (Rv2334) (Johnson et al., 2004; Tai et al., 1993). This conventional pathway of cysteine biosynthesis is found to exist in most eubacteria and plants (Bogdanova and Hell, 1997; Rege et al., 1996).

In summary, maintenance of intracellular cysteine concentration is essential for the survival and virulence of the pathogen. Therefore, *Mycobacterium* has evolved three different routes to

ensure a continuous and steady supply of cysteine. Complete inhibition of cysteine metabolism will require a fundamental understanding of all the branches of biosynthesis. Besides, the biochemistry and enzymology of the cysteine biosynthesis pathway in *Mycobacterium* are poorly understood.

1.11 Structural, kinetic, and regulatory properties of Mycobacterium tuberculosis CS

The genome of *M. tuberculosis* H37Rv encodes three genes which are annotated as CSs. While CS-A (CysK1) and CS-B (CysM) are both involved in the biosynthesis of cysteine via different substrates, the role of CysK2 (Rv0848) has been recently identified as an S-sulfocysteine synthase (Steiner et al., 2014). The primary role of CysK2 is catalyzing the formation of S-sulfocysteine from the substrate, O-phosphoserine (OPS), and sulfur donor, thiosulfate, in a PLP dependent reaction. But it can also catalyze the formation of L-cysteine by utilizing OPS and sulfur donor, sulfide. Therefore, CysK2 provides an additional pathway for synthesis of cysteine during dormancy in *Mtb*, utilizing OPS and sulfide as its substrates (Steiner et al., 2014); which is in correlation with its overexpression, during hypoxia and oxidative stress conditions in infected macrophages (Rahbari et al., 2017). Reports also show that cysteine can also be synthesized from methionine via the reverse transsulfuration pathway (Wheeler et al., 2005).

The crystal structure of *Mtb* CS-A with a four-residue peptide derived from the C-terminal of *Mtb* SAT has been solved (Schnell et al., 2007). The crystal structure of *Mt*CS with the *Mt*SAT derived C-terminal peptide (DFSI), addresses whether *Mt*CS interaction with *Mt*SAT would be similar to other known CS-SAT complexes and also reveals the binding conformation of the SAT derived peptide with CS. The peptide is bound to the active site of *Mt*CS, situated between the two domains of *Mt*CS, such that it hinders complete domain rotation, and *Mt*CS is trapped in an open conformation. The SAT peptide and *Mt*CS interaction are mediated by hydrogen bonds and hydrophobic interactions. The C-terminal isoleucine (Ile) residue in SAT, is the only conserved residue having a common binding mode with similar SAT peptide-CS complexes from *H. influenzae* and *A. thaliana* (Francois et al., 2006; B. Huang et al., 2005). The Ile residue is buried in the *Mt*CS active site via hydrogen bonding and hydrophobic interactions with the side chain of the CS active site. The other residues in the *Mt*SAT peptide are mostly not conserved and have different binding modes when compared to SAT from other studied organisms like *H. influenzae* and *S. typhimurium*.

Figure 1.10 The Crystal structure of *Mt*CS with the C-terminal DFSI peptide of *Mt*SAT. The last amino acid of the peptide is buried inside the active site pocket through hydrogen bonding and hydrophobic interactions.
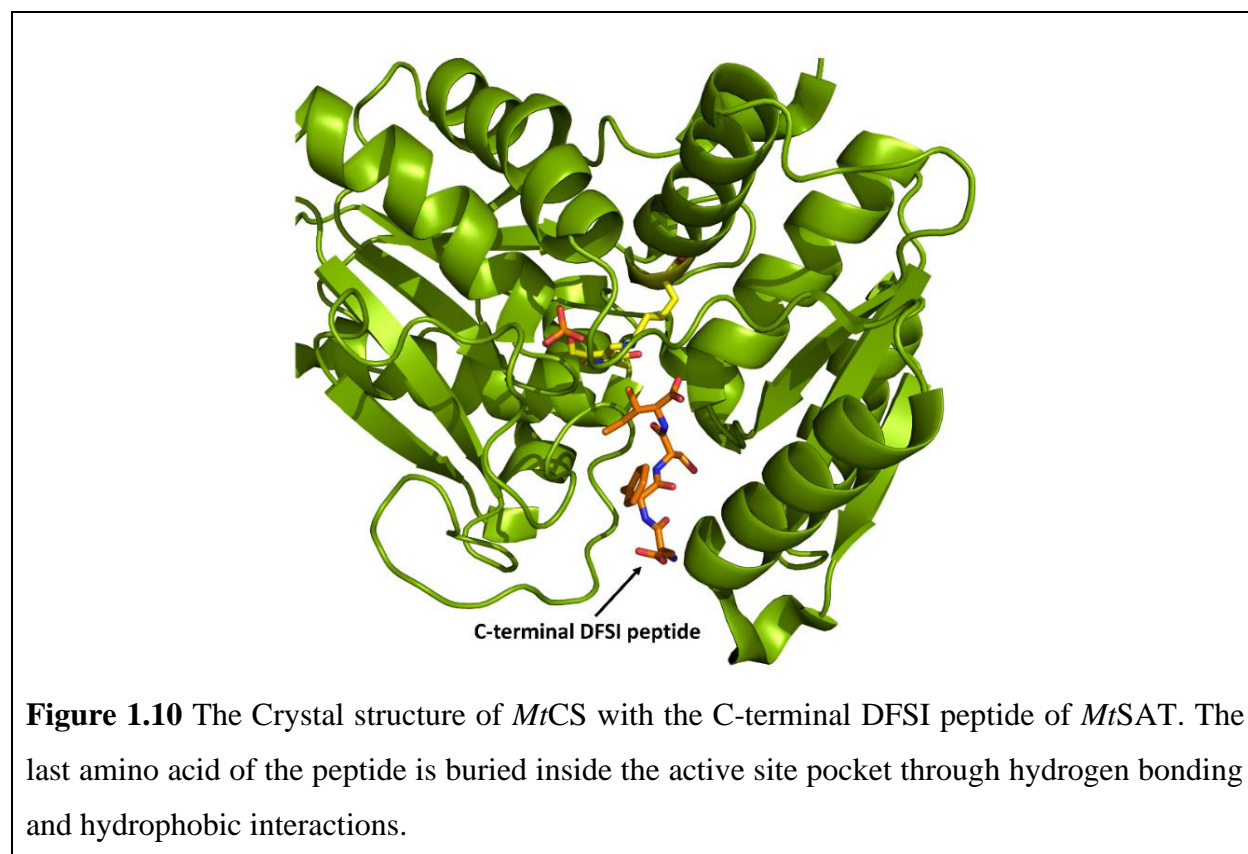
1.12 Serine acetyltransferase of Mycobacterium tuberculosis is not studied well so far

Mtb employs a battery of sophisticated, networked sensors to monitor redox and stress-specific signals. The understanding of whole redox machinery in *Mtb* may probably help design some strategy to combat this pathogen. The final steps of cysteine biosynthesis are conserved in the *Mtb* as the bioinformatics analyses of the *Mtb* genome shows the presence of both CS and SAT sequences. However, analyses of primary sequences of SAT and CS of *Mtb* shows that while the sequence, structure, and function of *Mt*CS is highly similar to those of other bacteria, the sequence of *Mt*SAT (Rv2335 gene locus) is very different from that of other bacteria. While multiple studies have characterized the assembly state, structure, and kinetic properties of *Mt*CS, no detailed characterization of *Mt*SAT has been carried out till now. Only two preliminary studies have attempted to characterize *Mt*SAT so far (Qiu et al., 2013, 2014) *Mt*SAT shows only about 30% identity in amino acid sequence to serine acetyltransferases from other bacteria and plants. The cysE gene in *Mtb* is 690 bp in size that encodes 229 amino acids enzyme. We have carried out multiple sequence alignment (MSA) of the *Mtb*SAT enzyme with other SATs of different organisms **(Figure 1.11 A)**. We have found almost 98 amino acids from the N- terminal end are

missing in *Mt*SAT. Structural studies have shown that protein-protein interaction between the N-terminal helix domain of SAT stabilized both trimer and hexameric oligomeric states (Gorman & Shapiro, 2004b) **(Figure 1.11 B)**. The N-terminal amino acid sequence variation in *Entamoeba histolytica* SAT (*Eh*SAT) changes the oligomeric assembly of the enzyme to exist in trimer only as compared to the other SATs (Kumar et al., 2011a). Therefore, *Mt*SAT with missing ~98 N-terminal residues is expected to be less stable and in a different oligomeric state as compared to SAT from other bacteria as the N-terminal is the key contributor of oligomerization. In other bacteria, the end product cysteine inhibits the SAT activity ultimately leads to a decrease in its production by feedback inhibition. Cysteine binds to the Coenzyme A binding site of SAT and competes for SAT binding by competitive inhibition. So, studies on *Mt*SAT activity and oligomerization state are important to reveal interesting facts about the biochemistry and assembly states of *Mt*SAT.
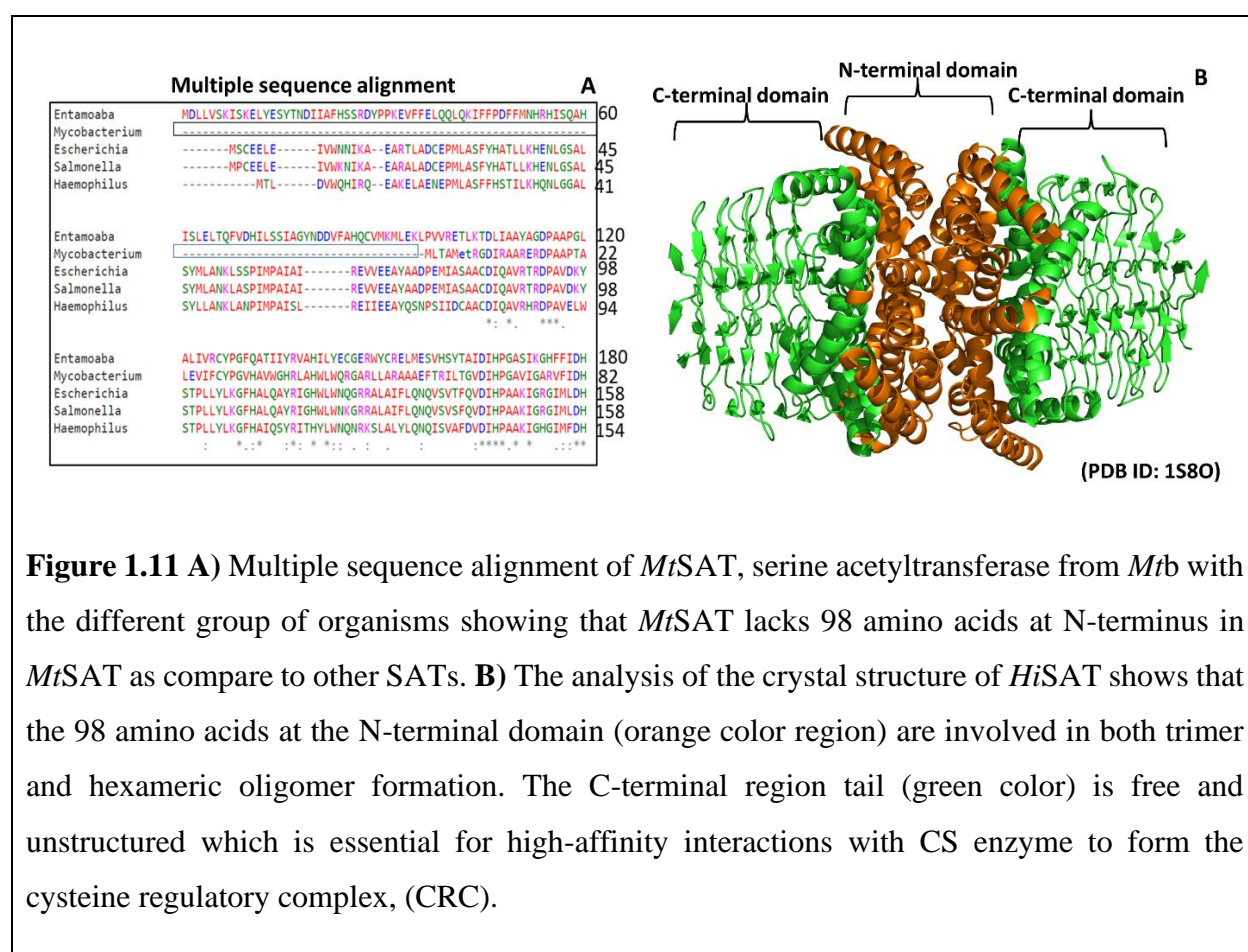
Figure 1.11 A) Multiple sequence alignment of *Mt*SAT, serine acetyltransferase from *Mt*b with the different group of organisms showing that *Mt*SAT lacks 98 amino acids at N-terminus in *Mt*SAT as compare to other SATs. **B)** The analysis of the crystal structure of *Hi*SAT shows that the 98 amino acids at the N-terminal domain (orange color region) are involved in both trimer and hexameric oligomer formation. The C-terminal region tail (green color) is free and unstructured which is essential for high-affinity interactions with CS enzyme to form the cysteine regulatory complex, (CRC).

1.13 The Cysteine regulatory complex of Mycobacterium tuberculosis

The significance of the enzymes of the classical cysteine biosynthesis pathway in *Mtb* is intriguing. Both the genes, CS-A and SAT, have been observed to be critical for cysteine biosynthesis in other bacteria (Fiegler & Brückner, 2006; Kredich & Becker, 1971; Salsi, Campanini, et al., 2010), whereas, CS-B pathway is active under anaerobic conditions (Tai et al., 1993), accounting for only 20% of the cysteine synthase activity (Nakamura et al., 1984b). However, similar observations have not been made in *Mtb*, where CS-A and SAT are not upregulated in stress conditions, like the genes of the other cysteine biosynthetic pathways. Due to the presence of two other pathways of cysteine biosynthesis in *Mycobacteria*, which utilize more stable substrates like OPS and thiosulfate (Steiner et al., 2014) probably the genes of the conventional pathway, including SAT and CS, maybe non-essential in *Mycobacteria* as the substrates of this pathway are more prone to oxidative damage. However, it is important to note that, although the CS-B (CysM) mediated pathway might be more important during oxidative stress conditions of *Mtb*, the CysM mutant strains are attenuated in macrophages, but they still survive (Rengarajan et al., 2005), clearly indicating the role of another route of cysteine metabolism. Till now, no reports have been published on whether the CRC complex exists in *Mtb* or not due to the difficulty in the purification of the *Mt*SAT enzyme of the cysteine synthase pathway. The enzyme activities of the CRC complex are also not studied so far. In *Mtb*, it is not clear whether the *Mt*SAT and *Mt*CS interact with each other to form the CRC complex and at what ratio they form the CRC complex. Further, no structural evidence is available for the CRC complex.

1.14 Scope and objective of the research

Enzymes of the cysteine biosynthetic pathway are essential for the survival and virulence of many pathogenic organisms (Hatzios and Bertozzi, 2011). Cysteine is the sole sulfur source for all components of the cell, including redox components. Both *Salmonella* and *Mycobacterium* have complex redox pathways that require a constant flux of cysteine from which thioredoxin, mycothiol, and other redox metabolites are synthesized (Paritala & Carroll, 2013). Interruption of the supply chain of cysteine biosynthetic pathways would, therefore, have a direct but negative effect on the ability of these pathogens to maintain redox homeostasis. *Mycobacterium,* in particular*,* has evolved with extraordinary abilities to adapt to a fast-changing hostile host-environment. Several studies have shown that proper maintenance of redox homeostasis is essential for survival and reactivation (Banuls et al., 2015).

Multienzyme complexes such as CRC substantiate the notion of "primary blocks" of metabolic systems (Kurganov, 1986). It is expected that the existence of metabolic systems forms the material basis metabolic control and therefore, studying the molecular features of these assemblies should allow us to understand the integration of cell metabolism with other processes. Further, only a few multienzyme complexes have been characterized in detail due to difficulties associated with purifying and assembling these complexes. Research from our lab has expertise in purifying, assembling, and characterizing CS, SAT, and CRC of *Salmonella typhimurium* (Kaushik et al., 2017). While continuing our efforts to study regulatory features of cysteine biosynthesis in bacteria, we also aim to provide the first detailed characterization of *Mt*SAT and *Mt*CRC of *Mycobacterium tuberculosis*. Thesis objectives have been designed to provide a comparative as well as a comprehensive study on CRC and its two component enzymes from *Salmonella* and *Mycobacterium*.

Thesis Title and Objectives

"Studies on Assembly and Regulation of Cysteine Regulatory Complex of Bacterial Pathogens"

- Biochemical, structural, and kinetic characterization of serine acetyltransferase from *Mycobacterium tuberculosis*
- Characterization of thermodynamic and kinetic properties of CRC assembly
- Understanding key regulatory features of CRC by solution and structural approaches

Our first objective was to characterize the *Mt*SAT enzyme using high resolution analytical and structural approaches. We cloned, expressed, purified the *Mt*cysE gene, and studied its assembly state and catalytic properties in detail. In the second objective, we studied and compared kinetic and thermodynamic signatures of CRC complexes from *Mycobacterium tuberculosis* and *Salmonella typhimurium* to understand species-specific-features. In the third objective, we examined key regulatory features such as feedback inhibition in *St*SAT and dynamic disassembly of CRC complexes using multiple approaches.

Chapter 2

Materials & Methods

Materials

Chemicals

All general lab chemicals and reagents were of analytical grade and purchased from various commercial sources like Merck, Hi-Media, SRL, Fisher-Scientific, and Sigma. All oligo-nucleotides were desalted and of analytical quality and procured from Integrated DNA Technologies (IDT) or Sigma (USA).

Bacterial strains

The strains used in this study with their relevant genotypes are listed in **Table 2.1**. *Escherichia coli (E. Coli)* was grown in Luria Bertani Broth (LB) and maintained on (LB) agar plates supplemented with suitable antibiotics as necessary. All bacterial cultures were grown at 37 °C with shaking at 200 rpm for liquid cultures unless mentioned otherwise. For routine storage, plates were maintained at 4°C and for longer terms, at -80°C in 50% glycerol for liquid culture stocks.

Table 2.1 Bacterial strains used in the study

Strain	Genotype	Source
***E.coli* DH5αT1**	*F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20, φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK–mK+), λ–*	Life technologies
***E.coli* TOP10**	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-*	Life technologies
Arctic Express (DE3)	*E. coli B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr]*	Agilent Technologies
BL21(DE3)-pLysS	*F- ompT hsdSB(rB- mB-) gal dcm (DE3) pLysS (CamR)*	Novagen
***E.coli* BL21(DE3)**	*B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)*	Novagen

E.coli **(DE3)**	**Rosetta**	*F- ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE (CamR)*	EMD millipore
BL21(DE3)		*E. coli B, F– ompT lon, with a λ prophage carrying the T7 RNA polymerase*	Novagen

Growth Media

Various bacterial growth media used in the study are described in **Table 2.2**. Media were prepared using R.O. water or distilled water and were autoclaved for 15 min at 121°C at 15 pounds per square inch pressure. Stock solutions of glucose, non-autoclavable drugs, and chemicals such as (IPTG) were sterilized by filtration through 0.22 µM filters (Millipore, USA) before adding to media. Unless mentioned otherwise, media were prepared following procedures as described in (Sambrook and Russell, 2001; Sambrook et al., 1989).

Table 2.2 Growth media and their composition

Media	Composition
Luria Bertani (LB medium)	1% Tryptone, 0.5% yeast Extract & 1% NaCl in D.D. water (pH 7.0)
LB Agar	LB medium supplemented with 1.5% agar
Super Broth (SOB)	35g Tryptone, 20g yeast extract, 5g NaCl. Make the volume up to 1litre.
Terrific broth (TB)	12 g Tryptone, 24 g Yeast Extract, 4 mL glycerol. Makeup volume 1litre with distilled water and autoclave.

Plasmids

pET28a+: A pBR322 derived plasmid. It is an *E.coli* expression vector containing a T7 promoter and a kanamycin resistance gene as a selection marker. This vector is used for the expression of proteins fused with 6-x His-tag at their N-terminals.

Antibiotics

All the antibiotics used in the studies were obtained from Gold biotech USA, Himedia, or Sigma Chemicals, and their stocks were prepared as described below (**Table 2.3**)

Table 2.3 List of Antibiotics used in the study

Antibiotic	Stock Concentration	Working Concentration
Kanamycin	50 mg/mL in MQ or D.D. water	50 µg/mL
Ampicillin	100 mg/mL in MQ or D.D. water	100 µg/mL
Chloramphenicol	35 mg/mL in Ethanol (100%)	17 µg/mL
Tetracycline	15 mg/mL in MQ or D.D. water	15 µg/mL

Stock solutions of antibiotics were sterilized through 0.22 μM filters (Millipore) and stored in aliquots at -20°C.

Primers

To amplify the CysE gene of *Mycobacterium tuberculosis H37Rv*, and the 216 bp of N-terminus of CysE gene of *Salmonella typhimurium,* oligonucleotides (primers) were designed on Vector NTI software. Oligonucleotide primers used in this work were commercially obtained from Integrated DNA Technologies (IDT) Inc., Coralville, IN, USA, or Sigma Chemicals, India. They were either of the desalted grade or High-Affinity Purified (HAP) grade.

Table 2.4 List of primers used to make StMtSAT gene construct

Primers (5'-3')	Sequence
*Mt*SAT_NheI_F	ATTAGCTAGCATGCTGACGGCCATGCGG
*Mt*SAT_HindIII_R	TAATAAGCTTTCAGATCGAGAAGTCCTCGCCGTG
*St*SAT216_NdeI_F	ATTACATATGCCGTGTGAAGAACTGGAAATCGTCTG
*ST*SAT216_NheI_R	TAATGCTAGCGGCATAGGCTTCTTCAACTACTTCACG

Buffers and solutions for recombinant DNA work

All buffers and solutions were prepared either in sterile Milli Q™ (Millipore Corporation, Germany) and/or autoclaved for 15 min at 121°C at 15 pounds per square inch pressure. Buffers with heat-labile constituents were filter sterilized.

Table 2.5 Buffer and solutions for recombinant DNA studies

Buffers/ Solutions	Composition
TE	1 mM EDTA, 10 mM Tris-HCl (pH 8.0)
50x TAE	242 g Tris base, 57.1 mL Glacial acetic acid, 100 mL 0.5 M EDTA (pH 8.0), deionized water to make final volume to 1 L.
0.5 M EDTA [pH 8.0]	18.61 g of disodium ethylene diamine tetraacetate.2H2O (SRL, India) was added to 80 mL of double-distilled water, and pH was raised to 8.0 using NaOH pellets. EDTA is usually dissolved at pH 8.0. The volume was adjusted to 100 mL. Sterilized by autoclaving and stored at room temperature.
6x DNA gel loading buffer	10 mM Tris-HCl (pH 7.6), 0.03% Bromophenol blue, 0.03% Xylene cyanol FF, 60% glycerol, 60 mM EDTA.

Table 2.6 Solutions for Alkaline-lysis miniprep

Solution I (Resuspension Solution)	a) Glucose 50 mM
	b) Tris.Cl (pH 8.0) 25 mM
	c) EDTA 10 mM
	(all solutions are made in deionized water).
Solution II (Lysis Solution)	a) NaOH 0.2 N

b) SDS 1.0 %

Solution III (Neutralization Solution)   a) 5 M Potassium acetate 60.0 mL

b) Glacial acetic acid 11.5 mL

c) Deionized water 28.5 mL

Table 2.7 Reagents for analysis of proteins (Solutions for SDS-PAGE)

Acrylamide (30%)	29 g Acrylamide, 1 g N, N'-Methylene bisacrylamide. Volume made to 100 mL with deionized water.
Lower Tris (pH 8.8)	18.17 g Tris, made to 100 mL with deionized water after adjusting pH 8.8 with 6 N HCl.
Upper Tris (pH 6.8)	12.11 g Tris, made to 100 mL with deionized water after adjusting pH 6.8 with 6 N HCl.
5x Sample buffer	0.15 M Tris-Cl (pH 6.8), 5%SDS, 25%Glycerol, 12.5% β-mercaptoethanol, 0.06% Bromophenol blue. Volume made to 10 mL with deionized water.
Laemmli buffer (Laemmli, 1970)	3 g Tris, 14.4 g Glycine, 1 g SDS. Volume made up to 1 L with deionized water.
Gel staining solution	10% Acetic acid, 45%Methanol, 0.25% Coomassie brilliant blue R250, and 45% deionized water.
Gel destaining solution	10%Acetic acid, 40%Methanol, 50% Deionized water

The composition of the stacking and resolving gels was the same as specified by Sambrook and Russel (2001).

Methods

Polymerase chain reaction (PCR)

Table 2.8 Composition of different ingredients used in the PCR protocol

Constituents of Reaction Mix	Concentration of Stock solutions	Volume used	Final concentration
Template DNA	500 ng/µl	1.0 µl	500 ng
Taq buffer	10 X	5.0 µl	1X
Taq Polymerase	10,000 units/ml	0.5 µl	5 units
dNTPs	10 mM	5.0 µl	1.0 mM
Forward primer	20 pm/ µl	1.0 µl	0.4 pm
Reverse primer	20 pm/ µl	1.0 µl	0.4 pm
$MgCl_2$	10 mM	5.0 µl	1.0 mM
MilliQ		31.5 µl	
Total volume		50.0 µl	

Table 2.9 Steps in the PCR program

Steps	Temperature and Time
Initial denaturation	95˚C for 5 minutes
Final denaturation	95˚C for 30 seconds
Primer annealing	56˚C/Tm for 40 seconds
Extension	72˚C for 1 min 30 seconds
No. of cycles	30 cycles
Final extension	72˚C for 10 minutes

Site-directed mutagenesis

The protocol used is based on a combination of the Stratagene Quikchange protocol. The gene construct in pET28a was used as a template for amino acid substitutions by PCR-based site-directed mutagenesis. After thermocycling, the original template DNA was eliminated by treating with DpnI at 37 °C for 1 hour, and 4 µl of the PCR products were transformed into 100 µl of *E. coli* TOP10 chemical competent cells and plated on Kanamycin resistance LB plate. 5 colonies were picked up, mini-prepped, and purified plasmid DNA was sequenced using an ABI capillary DNA Sequencer to confirm the mutation.

PCR protocol ingredients:

0.5 µl forward primer, 2.5 pmoles/µl

0.5 µl reverse primer, 2.5 pmoles/µl

0.25 µl 40 mM dNTP mix, (10 mM each)

1.25 µl 10x PfuUltra Buffer (contains Mg^{2+} ions)

1 µl Template DNA, 2 ng/µl

0.25 µl PfuUltra Hotstart (Stratagene)

8.25 µl sterile H2O

12.5 µl total

Finally, PCR products were transformed on appropriate antibiotics containing LB plates

Agarose Gel Electrophoresis

Agarose gels were made in 1X TAE buffer (Tris base-4.84 g, Acetic Acid-1.142 mL, EDTA-2mL of 0.5 M EDTA/L). Agarose was dissolved in 1X TAE by boiling. Ethidium bromide (0.5 µg/mL) was added to agarose and then poured into the gel cast. Combs were inserted and the gel was allowed to solidify. The combs were then removed and the gel along with the cast was placed in a gel tank containing 1X TAE buffer. DNA samples were mixed with 6x loading dye (bromophenol blue and glycerol) and then loaded onto the gel. The DNA bands were visualized under a UV transilluminator and the further image was captured by Gel Doc.

DNA Extraction from Agarose Gel

DNA fragments were eluted out from low melting agarose. After the gel fragment had resolved, the gel slice containing the band was excised out, placed in an Eppendorf, and heated to 55˚C with the binding buffer in 1:1 ratio, provided in the Gel Extraction Kit (Thermo Scientific) until the gel slice dissolved. The dissolved DNA was then transferred to a spin column and centrifuged at

13,000 rpm (5000g) for 1 minute. The flow-through was discarded and 700 µL of wash buffer (containing ethanol) was added to the column. It was again centrifuged at 13,000 rpm for 1 min and the flow-through was discarded. The empty column was centrifuged for 2 min at high speed to remove the residual buffer. The column was then placed in a fresh 1.5 mL Eppendorf tube and DNA was eluted in a volume of 20-50 µL of elution buffer or autoclaved MilliQ water.

Restriction Enzyme Digestion of DNA

Purified DNA samples were digested with restriction endonucleases in the NEB buffer provided along with the enzyme in a reaction volume of 40-70 µL at 37˚C for 3-4 hours. The samples were then analyzed on 0.8% agarose gel by electrophoresis. Restriction digestion reaction set up was as follows:

Table 2.10 Components of Restriction Enzyme Digestion Reaction

Sr. No.	Component	Volume
1.	PCR Product	3.0 µL (90 ng)
2.	NEB Buffer (10x)	3.5 µL
3.	NheI	1.0 µL
4.	HindIII	1.0 µL
5.	BSA	3.5 µL
6.	MilliQ	23 µL
7.	Total	35 µL

Ligation

The ligation mix consisted of digested vector and insert in the molar ratio of 1:3 in ligation buffer (50 mM Tris-Cl, 10 nM $MgCl_2$, 10 nM DTT, 1 mM ATP pH 7.8) and 1U of T4 DNA ligase in a final reaction volume of 20 µL. Ligation was carried out at 16˚C for 16 hours and the ligated DNA was used for transformation in DH5α *E. coli* competent cells.

Preparation of Competent Cells

E. coli competent cells (DH5α, BL-21(DE3), TOP10, etc.) were prepared by the method by Cohen *et.al,* (1972) with slight modifications. A single colony from a streaked plate of desired strains of *E. coli* was inoculated into 5 ml of LB media and allowed to grow for overnight at 37˚C in shaking at 200 RPM. One percent (1%) of this primary culture was inoculated into 100 mL of LB medium and grown till 0.3-0.4 OD at 600 nm. The culture was then chilled in ice water for 30 minutes and centrifuged at 4000 rpm for 10 minutes at 4 ˚C. The cells were then resuspended in 30 ml of chilled 0.1M $CaCl_2$ and incubated on ice for 1 hour. Cells were again centrifuged at 4000 rpm, for 10 minutes at 4˚C and resuspended very gently in 8 mL of resuspension solution (0.1 M $CaCl_2$ + 20% glycerol) and incubated at 4˚C for 3-4 hours. Tubes with 200 µL aliquots were made and stored at -80˚C for further use.

Transformation

Competent *E. coli* cells were transformed with the ligation mix or with plasmid DNA as described by Sambrook, *et.al.,* (1990). Competent cells were thawed on ice. The ligation mix or plasmid DNA was added to 200 µL of competent cells, thoroughly mixed by gentle tapping, and incubated on ice for 30 minutes. The cells were subjected to heat shock at 42˚C for 90 seconds and immediately incubated on ice for 5 minutes. To these cells, 800 µL of LB media was added and the cells were revived by shaking at 37 ˚C for 1 hour. The transformed cells were centrifuged at 8000 rpm for 5 minutes and then resuspended in 200 µL of residual LB media. These were then plated on LB-agar plates containing appropriate antibiotics and incubated at 37 ˚C for 16 hours.

Isolation of Plasmid DNA

The plasmid DNA was isolated using the alkaline lysis method. All reagents used for isolation were taken from the Thermo Scientific Plasmid DNA Purification Kit, already provided above. A single colony was inoculated into 10 ml of LB media and grown overnight with shaking at 37 ˚C at 220 rpm. These overnight cultures were harvested by centrifugation at 6000 rpm for 5 min at room temperature. The cell pellets were resuspended in 250 µl of re-suspension buffer (50 mM Tris, pH 8.0, 10 mM EDTA, glucose, 100 µg/mL RNase A) and then 250 µL of lysis buffer (0.2 N NaOH and 1% SDS) was added to it. The cells were gently mixed by inverting the tube 8-10 times till the solution became clear. Then 350 µL of Neutralization solution (3 M potassium acetate pH 5.5) was added to the tube and gently mixed by inverting and the cell debris formed

was removed by centrifugation at 13000 rpm for 5 min. The recovered supernatant was transferred to a spin column and centrifuged at high speed for 1 min. The flow-through was discarded and 500 µl of wash buffer (50% ethanol) was added to the column. The column was again centrifuged for 1 min and the wash step was repeated. The column was centrifuged for 2 min to remove the residual wash buffer. DNA was eluted in a volume of 30-50 µL TE (Tris EDTA pH 7.5) in a fresh 1.5 mL Eppendorf tube and quantified using Nanodrop (Thermo Scientific).

Cloning of *StMt*SAT *gene construct*

The MtSAT gene was first amplified by using the primers provided in **Table 2.4.** Similarly, the 216 bp gene sequence was amplified by using the primers. Both genes were subjected to restriction enzyme digestion as provided in **Table 2.10.** After restriction digestion, the genes were ligated according to the method mentioned in the ligation section. Further, the *StMt*SAT gene construct was cloned into the pET28a+ vector between the restriction sites NheI & HindIII having 6x His Tag at N-terminus. This clone was transformed into the Arctic host cells to express and purify the protein.

DNA Sequencing

Sequencing reaction for all clones was set up according to the mentioned protocol. The samples were subjected to sequencing PCR and clones were confirmed by automated DNA sequencing using the ABI Prism Automated DNA sequencer at the Institute of Microbial Technology, Chandigarh.

Table 2.11 The sequencing reaction composition.

REACTION	QUANTITY
Terminator Ready Reaction Mix	1.0 µL
Sequencing Buffer (5X)	2.0 µL
DNA Template	150-300 ng
Sequencing Primer	3.2 pmol
Water	As required
Total Volume	10.0 µL

Table 2.12 DNA sequencing PCR protocol

Initial denaturation	96˚C for 1 min
Final denaturation	96˚C for 10 seconds
Primer annealing	50˚C for 5 seconds
Extension	60˚C for 4 min
No. of cycles	24 cycles
Hold	4 ˚C

Protein expression and purification

The DNA coding for CS protein (cysK gene) from *Salmonella typhimurium* strain LT2 and (cysK1 gene) from *Mycobacterium tuberculosis* was PCR amplified and cloned into the expression vector pET28a+ by Dr. Mary Krishna and Dr. Shrijita Benergy respectively. These clones were transformed in BL-21(λDE3) *E. coli* competent cells using the heat shock method on the LB Agar plates containing 50 µg Kanamycin. A single colony was picked and grown in 10 mL LB medium overnight (1⁰ culture) having Kanamycin antibiotic. This 1⁰ culture was used as inoculum for 800 mL LB media. Cells were further grown till O.D. (Optical density) reaches 0.4 to 0.6 at 37 ⁰C. Protein was expressed at 20 ⁰C by adding 1 mM IPTG (Isopropyl-β-d-thiogalactopyranoside) for 14 hrs. The cultures were harvested by centrifugation at 6000 rpm for 15 min to obtain the cell pellet. The pellet was resuspended in 30 mL of Lysis buffer A (50 mM Tris-Cl pH 8.0, 300 mM NaCl, 10% glycerol, 0.1% Tween-20) and lysed by sonicating on a sonicator for 30 min with 8 sec "on" and 12 sec "off" pulses. The sonicated supernatant was recovered by centrifugation at 13k rpm for 30 min. Two to four mL of nickel-nitrilotriacetic (Ni-NTA) agarose beads affinity resin (Qiagen) was equilibrated in Buffer B (50 mM Tris-Cl pH 8.0, 300 mM NaCl, and 10 % glycerol). The clarified supernatant was loaded onto the equilibrated resin and incubated for 10-15 minutes. After this, the resin was washed with 5 mL of wash buffer containing 40 mM imidazole. Fractions were eluted using 300 mM imidazole and analyzed on a 12% SDS PAGE gel. An intense band with a molecular weight of ~35 kDa was observed which matched with the expected molecular

weight of CS. The fractions containing protein were pooled and further purified using gel filtration chromatography.

A similar protocol was followed for the purification of *St*SAT with slight modifications. The *St*SAT gene was previously cloned by Dr. Mary Krishna in pET 28a+ vector and transformed into Rosetta gami cells. A single colony was picked and grown in 10 mL LB medium overnight (1⁰ culture) having Kanamycin and Chloramphenicol antibiotic. This 1⁰ culture was used as inoculum for 2⁰ culture of 800 mL of LB media. Cells were further grown till O.D. (Optical density) reaches 0.4 to 0.6 at 37 ⁰C. These were then induced with 0.5 mM IPTG and grown for 14 hours post-induction at 20 °C. Cells were harvested by centrifugation at 6000 rpm for 10 min and lysed in 30 mL of lysis buffer (50 mM Tris-Cl pH 8.0, 300 mM NaCl, 10% glycerol, 0.1% Tween-20, 250 mM ammonium sulfate, 0.1 mM EDTA, 0.5 mM DTT). The cell lysate was sonicated very mildly at 7 seconds "on" pulse and 8 sec "off" pulses for 30 min. The clarified lysate was recovered with centrifugation at 13k rpm for 30 min at 4 °C and the recovered supernatant was loaded onto Ni-NTA resin for affinity purification. The protein-containing resin was washed with buffer containing 40 mM of imidazole. Finally, the protein was eluted from the resin by using 300 mM imidazole, subjected to quick dialysis, and loaded onto the HiLoad 16/600 Superdex S-200 gel filtration column and gel filtration chromatography was performed.

For, purification of *Mt*SAT protein, the *Mt*SAT (CysE) gene was cloned into a pET 28a+ vector and this vector was successfully transformed into pLys cells, grown in LB agar plate containing 50µg/mL kanamycin and 35µg/mL chloramphenicol. One single colony was picked and inoculated in 10 mL Luria Bertani media overnight at 37⁰ C. Secondary culture of 800 mL SB media containing appropriate antibiotics, was inoculated by adding 1% inoculum of 1⁰ (primary) culture at 37⁰ C. The culture was induced with 0.1 mM IPTG when OD at 600 nm reached 0.4-0.5 and protein was expressed at 30⁰ C for 14 hrs. Cells were then pelleted down at 6000 rpm for 10 min and resuspended the pellet in the lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl, 20 % Glycerol) and sonicated. The lysed cells were then centrifuged at 13000 rpm for 30 min and the supernatant was loaded on pre-equilibrated Ni-NTA beads. After binding for 1 hour, beads were washed with washing buffer (50 mM Tris pH 8.0, 500 mM NaCl, 20 % Glycerol, 20 mM Imidazole). The *Mt*SAT protein then further eluted with elution buffer (50 mM Tris pH 8.0, 500 mM NaCl, 20 % Glycerol, 300 mM Imidazole). Eluted *Mt*SAT protein then concentrated and

loaded onto pre-equilibrated superdex 16/600 S200 column and collected the purified protein and used for the experiments or stored at -80 ⁰C.

SDS-PAGE analysis

Polyacrylamide gel electrophoresis (SDS-PAGE) was performed in denaturing conditions. For a 12 % PAGE, resolving gel consisted of 30 % (W/V) acrylamide: bisacrylamide (29:1), 1.5 M Tris-Cl pH 8.8, 10 % SDS, 10 % APS, double distilled water, and TEMED for polymerization. The 5% stacking gel was made similarly with 1.0 M Tris-Cl pH 6.8. Samples were boiled in reducing sample buffer (50 mM Tris-Cl pH 6.8, 0.1 % bromophenol blue, 5 % glycerol, 1 % SDS, and 2.5 % β mercaptoethanol) for 10-15 minutes. These were centrifuged at 13000 rpm for 2 minutes before loading onto the gel. The gel was electrophoresed in running buffer (25 mM Tris-Cl pH 8.3, 250 mM glycine, 0.1% SDS) at constant current (30 mA) followed by staining with Coomassie brilliant blue G-250 (0.05% w/v) in 45% methanol and 10% acetic acid for 1 hour and destained in methanol: glacial acetic acid: water (4:1:5).

Size Exclusion Chromatography (SEC)

The size-exclusion column used to purify the Proteins was Superdex S200 16/600 (GE Healthcare) on an Äkta Explorer 100 system (Amersham Pharmacia Biotech). The column was equilibrated with 1.5 column volume of gel filtration buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 10 % glycerol). Samples were loaded on to the column and the purified protein was eluted at a flow rate of 0.8 mL/min, by monitoring at 280 nm. Elution fractions were collected and run on 12% SDS-PAGE to confirm purity. The proteins were concentrated using an Amicon centrifugation device with a 30-kDa molecular mass cut-off. Aliquots of the protein preparations were flash-frozen in liquid nitrogen and stored at -80 °C.

For approximate molecular weight determination, Analytical size exclusion chromatography experiments were carried out using a Superdex S200 GL 10/300 column (GE Healthcare) at 5 °C. The column was equilibrated with 2 column volumes (CV) of gel filtration buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 10 % glycerol) before the run. Elution was carried out at a flow rate of 0.3 mL/min and monitored at 280 nm. Blue dextran 2000 was used to determine the void volume of the column and a mixture of known protein standards Ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), and ribonuclease A (14

kDa) were applied to the column and their elution positions were used for estimating the approximate molecular weight of *St*CS, *Mt*CS, *St*SAT, *Mt*SAT, *St*CRC, and *Mt*CRC.

PROTEIN CONCENTRATION DETERMINATION

Protein concentrations were estimated by measuring the UV-visible absorbance spectra of the purified proteins between wavelength range of 200 nm to 500 nm in a Cary Series UV-Vis Spectrophotometer (Agilent Technologies) and absorbance at 280 nm was used for concentration estimation. Molar extinction coefficients used for protein concentration determination were; 19940 $M^{-1}cm^{-1}$ for *St*CS, 11500 $M^{-1}cm^{-1}$ for *Mt*CS, 26990 $M^{-1}cm^{-1}$ for *St*SAT 30840 $M^{-1}cm^{-1}$ for *Mt*SAT obtained from Prot-Param program of the Expasy server (https://web.expasy.org/protparam/). The ratio of absorbance at 280/412 nm was taken as a measure of the CS protein's quality.

Steady-State Kinetic Studies for CS

The enzyme assay for *St*CS, *Mt*CS were carried out using the acid ninhydrin assay for cysteine (Gaitonde, 1967). The O-acetylserine (OAS, Sigma chemicals) and Na_2S dissolved in 0.1 M HEPES pH 7.0. were used as substrates. CS catalyzes the formation of cysteine, which can be detected by monitoring the absorption at 560 nm. The enzyme reactions were performed at 30 ° C in 0.1M HEPES, at pH 7.0. The total volume of the reaction was 150 µL and the volume of substrate did not exceed 5%. A maximum of 4.5 µl of the enzyme was added to the reaction. This assay is based on the specific reaction of ninhydrin with cysteine under strongly acidic conditions. OAS concentrations were taken in a range of 0.1 mM to 10 mM while the Na_2S concentration was fixed at 3 mM. The components were mixed and the reaction was initiated by the addition of CS (100 ng). The reaction was allowed to proceed at 30 °C for 25 min. The reaction was terminated by the addition of 5% TCA (Tri-chloroacetic acid), centrifuged at 13k rpm and 125 µL of supernatant was transferred to a new tube. To this tube was added 125 µL of glacial acetic acid and 125 µL of acid ninhydrin reagent (250 mg of ninhydrin dissolved in 2 ml of concentrated HCl and 3 ml of glacial acetic acid). After mixing, samples were boiled for 10 min at 100 °C in the preheated water bath. Samples were cooled to room temperature and diluted with 625 µl of chilled 95% ethanol. Absorbance was recorded at 560 nm. In most cases, plots of initial velocity vs. substrate concentrations were fit to the Michaelis-Menten equation (1) to obtain K_M and V_{MAX}. Kinetic assays were verified by comparison of values in the standard curve for cysteine drawn

beforehand in the same conditions. The goodness of fit was assessed and the average r^2 (0.94-0.96) values were found to be satisfactory.

$$Vo = (Vmax * [S])/(Km + [S]) \qquad\qquad \text{(Eq. 1)}$$

Steady-State Kinetic Studies for SAT

The serine acetyltransferase activity of the *St*SAT and *Mt*SAT protein was determined by the Ellman's reagent DTNB (5,5'-dithiobis-(2-nitrobenzoic acid)) reaction protocol, it reacts with the CoA group and gives increased absorbance at 412 nm (Hindson et al., 2000; Kredich & Tomkins, 1966b). Tris, 50 mM buffer was used to preparing a 200 uL reaction mixture containing 1 mM DTNB reagent and assayed in the 96 well plate formats. Serine concentration was varied from 0.1 mM to 8.0 mM when treated as the ligand, where acetyl-CoA concentration was kept constant at 0.2 mM. TSE buffer containing 2 mM serine and enzyme was used as one control and acetyl-CoA, an enzyme in TSE buffer was used as another control. A total of 100 ng of protein was used to initiate the reaction for 10 minutes at 25°C. Absorbance readings were monitored with a microplate reader (Tecan scientific instruments) at 412 nm. Initial velocities were plotted against the concentration of ligand monitored for 10 minutes. Initial velocity vs. substrate concentrations was fit to the Michaelis Menten Equation (Eq. 1) to obtain K_M and V_{MAX}. The rate of product formation was calculated using the molar extinction coefficient of DTNB (14000 $M^{-1}cm^{-1}$).

Bioinformatics Analysis

The multiple sequence alignment was performed for Serine acetyltransferase (SAT) where 5 homologous protein sequences of SAT protein including *Mycobacterium tuberculosis* were taken from the UniProt database. The alignment was performed by using ClustalW, software freely available at the Swiss-EMBL site**.**

Circular Dichroism measurements (CD)

CD measurements were carried out with a JASCO spectropolarimeter (Jasco, Tokyo, Japan) equipped with a Peltier-type temperature controller (PTC-348W). Far-UV spectra were obtained in a quartz cuvette with a 10 mm light path-length and each spectrum obtained was an average of 7 scans. The ellipticity of protein CD spectra is reported as mean residue ellipticity (MRE) in deg/cm²/dmol units. For all cases, the spectrum of buffer was subtracted to obtain the real protein spectra. Deconvolution of far UV-CD spectra was performed using spectra manager (Jasco)

software provided by the manufacturer. The following formula was used to convert the CD signal to mean residual ellipticity $[\theta]^{MRE}$.

$[\theta]^{MRE} = [\theta \text{ (Degrees)} * MRW] / [10 * \text{concentration (mg/mL)} * \text{path length (cm)}]$

MRW (mean residual weight) = (Total molecular weight of the protein/ total number of amino acids).

$[\theta]MRE$ = Mean residual ellipticity (deg cm2 dmol-1)

θobs = Raw ellipticity (mdeg).

Analytical ultracentrifuge (AUC)

Sedimentation velocity experiments were performed using an Optima XL-I analytical ultracentrifuge equipped with absorbance optics with an An-50Ti 8 place rotor (Beckman Inc.). Purified CS, SAT, and CRC (either collected from SEC elution or assembled from SEC-purified CS and SAT) were extensively dialyzed in the standard buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 5% glycerol). Sedimentation velocity studies were carried out at 40,000 rpm (129048 x g) at 20 °C using two-channel charcoal-filled centerpieces with Sapphire glass windows. Samples were loaded onto the two-sector centerpiece (400 uL reference cell and 390 uL sample cell). The velocity data were collected by scanning samples at a wavelength of 280 nm with a spacing of 0.003 cm and an average of 3 scans per step. The partial specific volumes at 20 ºC were determined from amino acid composition and solvent density was calculated using SEDNTERP (Stafford, 1992). Assembly states of CS, SAT, and CRC complexes were analyzed by direct curve fitting of sedimentation boundaries using Sedfit (Schuck, 2000). Fit to data was selected based on the root mean square deviations (RMSD) of below 0.008. Sedimentation coefficients were corrected to standard conditions at 20 °C in water, $S_{20,w}$. The Data analysis was also performed using DCDT+ software (Philo, 2006) resulting in a similar finding to support the data.

Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry experiments were performed using a VP-ITC calorimeter (Microcal, Inc). Proteins used for the ITC were GFC purified with extensively equilibrated GFC column with buffer (20 mM HEPES and 20 mM NaCl). All samples and buffers were degassed at room temperature before use. *Mycobacterium tuberculosis* CS (10 μL/injection) was added using an automated-controlled 250 μL microsyringe at an interval of 3-4 min into the cell containing *Mt*SAT. Control experiments were performed by injecting the same amount of buffer into the cell

to calculate the heat of dilution for each injection. Data obtained from titrations were analyzed either using a single-site binding model equation (2) or two sets of sites using equation (3), or a three-site model using equation (4).

$$Qitot = VoEtot\{\frac{(K1P)\Delta H1}{1+K1P}\} \qquad\qquad \text{(Eq. 2)}$$

$$Qi\,Total = Vo\,Etot\,\{\frac{[\Delta H1K1P+(\Delta H1+\Delta H2)K1K2P2]}{1+K1P+K1K2P2}\} \qquad \text{(Eq. 3)}$$

$$Qitot = VoEtot\,\{[\Delta H1K1P + (\Delta H1 + \Delta H2)\,K1K2P2] + (\Delta H1 + \Delta H2 +$$
$$\Delta H3)\,K1K2K3P3\}/(1 + K1P + K1K2P2 + K1K2K3P3)\} \qquad \text{(Eq. 4)}$$

Where Qitot is total heat after the ith injection, V0 is the volume of the calorimetric cell, K1, K2, and K3 are the observed equilibrium constants for each site, P is the concentration of free CS dimer, and $\Delta H1$, $\Delta H2$, and $\Delta H3$ are the corresponding enthalpy changes. The corresponding Kobs and ΔH were obtained by fitting the experimental data to either of the models using Origin 7 ITC data analysis software provided by the instrument manufacturer (Microcal, Inc.). The observed binding constants were converted to change in free energy (ΔG) using ΔG= -RTln Kobs, where R is the gas constant (1.9872 cal $K^{-1}mol^{-1}$) and T is the absolute temperature. Changes in entropy (ΔS) were calculated using $\Delta G = \Delta H - T\Delta S$.

SAXS Experiments

The SAXS experiments were performed to analyze the structural organization of the *Mt*SAT and *St*CRC complex in solution. The measurements were carried out at BM 29 beamline at ESRF at 4 °C using 20 µL of the protein sample solution. The sample (1- 5 mg/mL) yielded SAXS data devoid of inter particular interaction or aggregation. The scattering data sets were analyzed using the PRIMUSQT (D. Franke et al., 2017) integrated suite of programs in ATSAS v2.7.8. The radius of gyration (Rg) and forward scattering intensity I_0 were obtained from Guinier plot analysis. Kratky plot [I(Q)Q2 versus Q] was generated to know whether the protein is properly folded or not. Indirect Fourier transformation of scattering data over measured Q range was computed to obtain a pairwise distribution function of interatomic vectors. This analysis also provided an estimation of the maximum linear dimension (Dmax) of the scattering shape of a protein in solution along with Rg and I0 values.

Model fitting into SAXS envelope

The GNOM program was used to generate the P(R) plot and its output forms the basis of low-resolution ab initio shape construction using simulated annealing optimization of a dummy atom set, as implemented in the DAMMIF and DAMMIN program (Daniel Franke & Svergun, 2009). The DAMMIF runs gave 10 independent models and were averaged using the DAMAVER suite of programs. The program minimizes the normalized spatial discrepancy (NSD) between the models, thus developing the most favorable average model. After refinement, for the low bead occupancy positions and loosely connected average atom model, DAMFIT was used. This generated the most ideal and depictive envelope, that agrees with particle-excluded volume. Finally, the structural model was superimposed onto the refined envelope using the SUPCOMB program. The experimental and theoretical scattering data were compared using the CRYSOL program. The chi-square value thus generated indicates similarity between the three-dimensional model and the low-resolution envelope. PyMOL program was used for visualization and generation of all superimposed structures. Curve-fitting and data–plotting were performed using Origin software.

Crystallization and structure determination

A search for finding the right crystallization conditions was carried out by utilizing the Wiz 1, 2, Wiz 3, 4, (Jena Bioscience), Structure screen 1 (molecular dimension), Helix (molecular dimension), Nextel (Qiagen) each of which has 96 different buffer conditions for crystal growth. We proceeded with the crystallization of the *St*SAT protein via the sitting drop vapor diffusion method. Each drop contained 1.0 µL of protein (30 mg/mL) and 1.0 µL of reservoir solution, equilibrating against 70 µL of precipitant solution. The crystals grew slowly and within a month rhombohedral *St*SAT crystals appear in Wiz 1, 2 screen in the condition of 1.0 M Potassium sodium tartrate, 0.1 M MES sodium hydroxide pH 6.0. The *St*SAT crystals were flash-frozen in liquid nitrogen using mother liquor with 20 % glycerol as cryoprotectant. The X-ray diffraction data for *St*SAT was collected on the in-house MAR345 image plate detector mounted on a Rigaku MicroMax-007HF microfocus rotating Cu kα anode as an X-ray generator operated at 40 kV and 30 mA. The data was collected by the oxford cryo stream on 100 K and for high-resolution data collection, diffraction was also carried out at BM 30 beamline of ESRF at Grenoble, France. A total of 300 frames were collected with one-degree oscillation exposed for two minutes for each image obtained and crystal to detector distance was set to 200 mm. The diffraction data collected

was processed, integrated by imosflm, and scaled using a scala suit (CCP4 crystallography suit). The structures were solved by molecular replacement using the program PHASER and CCP4 suite (McCoy et al., 2007). The model was subsequently refined using Phenix, Refmac5, and the fit of the model to the electron density evaluated by COOT (Adams et al., 2010; Emsley & Cowtan, 2004; Murshudov Alexe et al., 1997). Pymol is used to generate good quality high-resolution structure figures (The PyMOL Molecular Graphics System, Schrodinger).

Chapter 3

Biochemical, structural, and kinetic characterization of serine acetyl transferase from Mycobacterium tuberculosis

Research Problem

SAT is a key enzyme in cysteine biosynthesis and it has been projected to be a potential drug target for developing anti-TB compounds due to its key role in cysteine production and also lack of any SAT homolog in humans (Brunner et al., 2016; Paritala & Carroll, 2013). SAT catalyzes the formation of OAS from L-serine and acetyl Co-A and its activities are regulated by feedback inhibition as well as by binding to cysteine synthase (CS) (Benoni et al., 2017). Only very recently, SAT from *Mtb* has been identified and its function is characterized (Qiu et al., 2013). However, unlike SAT from other bacteria, very limited structural and biochemical details of *Mt*SAT are available and no systematic or detailed characterization study of *Mt*SAT is available (Kumar et al., 2011a; Pye et al., 2004b). Further, it is not known whether *Mtb* adopts similar regulatory mechanisms in controlling cysteine biosynthesis like in other bacteria and plants or not. Two of the most well-characterized regulatory mechanisms of SAT are feedback inhibition and regulation through protein-protein interaction by forming cysteine regulatory complex (CRC) (Kumaran et al., 2009; Salsi, Campanini, et al., 2010). In the absence of details of structural, assembly, and kinetic properties of *Mt*SAT, the key features of the cysteine synthesis step and regulatory mechanisms of *Mtb* remain unknown.

SAT from *Salmonella typhimurium*, *Hemophilus influenzae*, and *Escherichia coli* exists as homo-hexamer while SAT of *Glycine max* exists in both trimeric and hexameric states (Gorman & Shapiro, 2004b; Pye et al., 2004b; Yi et al., 2013a). Recent structural studies show that the SAT of *Entamoeba histolytica* is a trimer but detailed solution studies are needed to support structural information (Kumar et al., 2011b). The trimer is the unit building block and two trimers assemble in a head-to-head manner to form the hexamer. Each monomer of SAT is defined by an N-terminal α-helical domain and a C-terminal left-handed β-sheet domain. The N-terminal domain of each monomer mediates several non-covalent interactions with two other monomers to form the core trimeric state and again, the N-terminals of two trimers interact to form a trimer-trimer interface in the hexameric state. The C-terminal consists of three functional regions; the active site, substrate/product/metabolite binding feed-back inhibition site, and an unstructured C-terminal tail necessary for binding to the active site of CS to form the CRC (Feldman-Salit et al., 2009; Mino et al., 2000). The sequence of *Mt*SAT, annotated as Rv2335, is a ~ 24 kDa (690 base pairs & 229 amino acids) protein, much smaller in size than SAT from other bacteria and plants which are 30 kDa - 36 kDa (250-310 amino acids) in size (Qiu et al., 2013). The multiple sequence alignment

(MSA) of the *Mt*SAT sequence with known SAT sequences shows that *Mt*SAT is missing the first ~98 N-terminal residues (**Figure 3.1A**). Since the N-terminal α-helical domain is important for protein-protein interaction between monomers, the truncated N-terminal of *Mt*SAT implies that the quaternary state(s) of *Mt*SAT will be different than that of other SATs. Since cysteine biosynthesis is tightly linked to redox-mediated defense mechanisms of *Mtb*, (Bhat et al., 2012; Paulsen & Carroll, 2013; Poole, 2015) the different quaternary states of *Mt*SAT may of physiological significance to this pathogen. Therefore, detailed characterization of functional assembly or quaternary states of *Mt*SAT is essential to understand the organization roles of cysteine biosynthesis of *Mtb*.

In this objective, we used multiple analytical and biophysical methods to characterize the assembly state(s) and catalytic properties of each quaternary/assembly states under different solution conditions. Our results show that *Mt*SAT exists in three different assembly states, trimer, hexamer, and dodecamer with average molecular weights of ~72 kDa, ~144 kDa, and ~288 kDa respectively. Interestingly, each quaternary state is different with respect to the regulatory feedback and protein-protein interaction mechanisms. We propose a multi-ensemble CRC model for the regulation of cysteine biosynthesis in *Mtb*.

Results

Bioinformatics analyses of Mycobacterium tuberculosis Serine Acetyl Transferase (MtSAT)

The cysE gene in *Mtb* is 690 bp in size that encodes 229 amino acids enzyme serine acetyltransferase (*Mt*SAT). SATs of other different bacteria are much larger in amino acid sequence having an amino acid sequence in the range of 250-310 amino acids and a molecular weight between 30 kDa to 36 kDa. The sequence similarity between the different groups of bacteria is about 30%. We performed the multiple sequence alignment (MSA) of the *Mt*SAT protein with other SATs proteins of different organisms including bacteria *E. coli, H. influenzae, S. typhimurium*, and unicellular parasite *E. histolytica* (**Figure 3.1 A**) by using ClustalW software freely available at the Swiss-EMBL site. We found that almost 98 amino acids are missing at the N- terminal end of *Mt*SAT as compare to other SAT proteins. Structural analysis of previously solved crystal structure of the hexameric form of *Hi*SAT shows that the ~98 N-terminal amino acids are mainly involved in the N-terminal domain protein-protein interactions to form the hexameric SAT oligomer (**Figure 3.1 B**) (Gorman & Shapiro, 2004b).
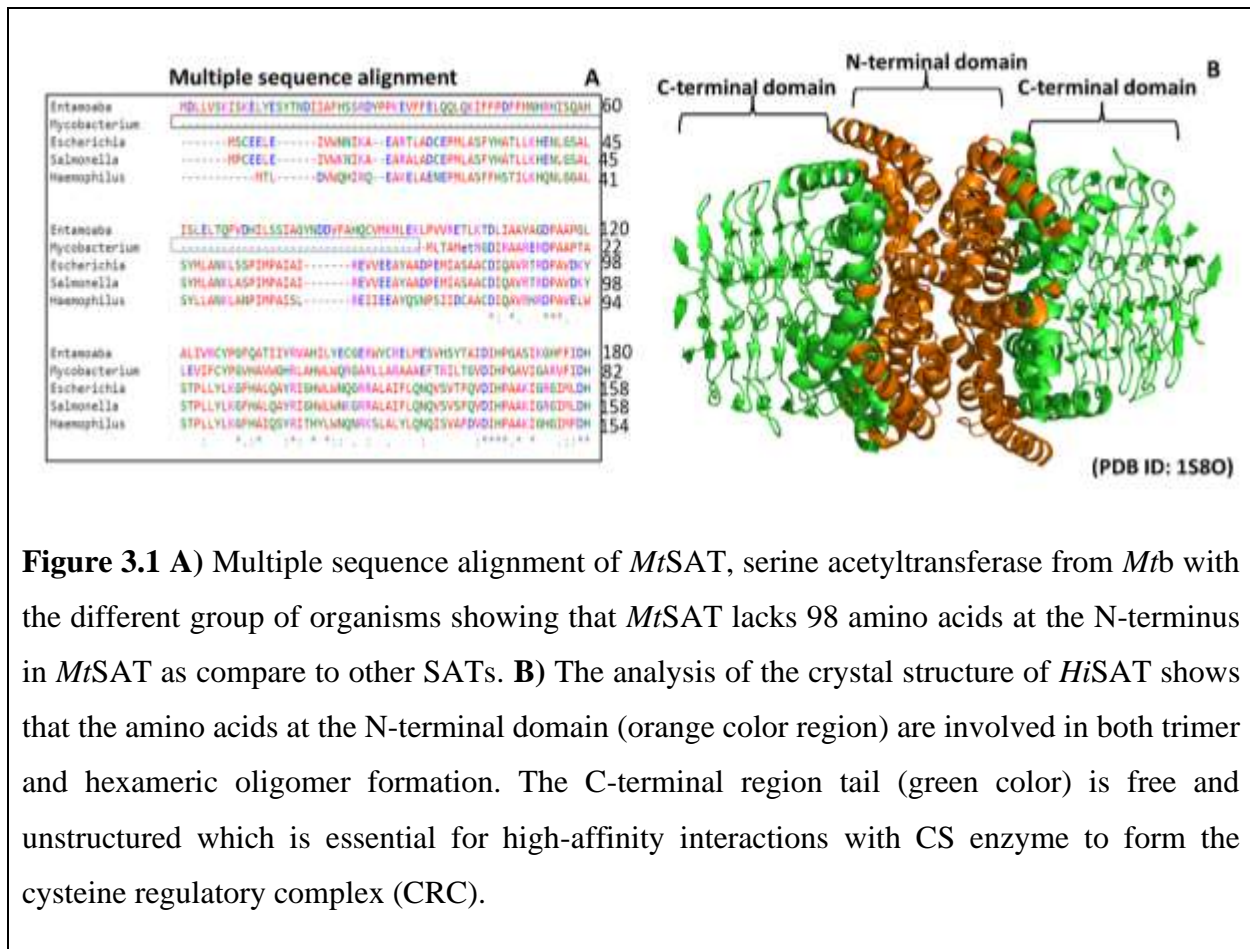
Figure 3.1 A) Multiple sequence alignment of *Mt*SAT, serine acetyltransferase from *Mt*b with the different group of organisms showing that *Mt*SAT lacks 98 amino acids at the N-terminus in *Mt*SAT as compare to other SATs. **B)** The analysis of the crystal structure of *Hi*SAT shows that the amino acids at the N-terminal domain (orange color region) are involved in both trimer and hexameric oligomer formation. The C-terminal region tail (green color) is free and unstructured which is essential for high-affinity interactions with CS enzyme to form the cysteine regulatory complex (CRC).

MtSAT gene cloning and expression

The cysE gene (SAT) is 690 bp and is previously cloned by Dr. Srijita into the pET28a vector. The gene sequence was confirmed by gene sequencing. Sequencing data shows that the gene is cloned between the restriction sites NheI and XhoI and has the 6x His tag coding segment at the N-terminus of the gene. We transformed this clone into different host cells including BL21 (DE3), BL21(DE3) pLysS, Rosetta (DE3), Arctic Express (DE3) cells in an Agar plate containing appropriate antibiotics. We standardized the expression conditions for the *Mt*SAT protein under which we can get the maximum soluble fraction of the protein. To achieve the maximum soluble expression, we screened different types of media, expressed at different temperatures, and varied buffer conditions for lysis and purification. Finally, we found that expressing *Mt*SAT in the BL 21 (DE3) pLysS cells at 30-degree temperature for 14 hrs and induced with 0.1 mM IPTG yields the desired soluble fraction for analytical experiments (**Figure 3.2 A**). Due to the lack of N-terminal 98 amino acids sequence stretch, protein shows a very high propensity to aggregate and

therefore stringent purification protocol was required. After lots of expression standardization, the protein was expressed in a soluble fraction **(Figure 3.2 A, lane 5)**. It is expressed as His-tagged *Mt*SAT, therefore, purified through Nickel affinity chromatography in the first step, and in the second step, the size-exclusion chromatography method was used to further purify the *Mt*SAT protein. In size exclusion chromatography, three peaks were observed for *Mt*SAT protein and PAGE analyses confirmed all three peaks were *Mt*SAT protein **(Figure 3.2 B)**. The first, second, and third peaks were eluted at 10.0 mL, 12.0 mL, and 14.0 mL respectively. Standards run in this column shows that a 10.0 mL peak fraction corresponds to a molecular weight of ~300 kDa. The second and third peak fractions, eluted at 12.0 and 14.0 mL correspond to ~150 kDa and ~75.0
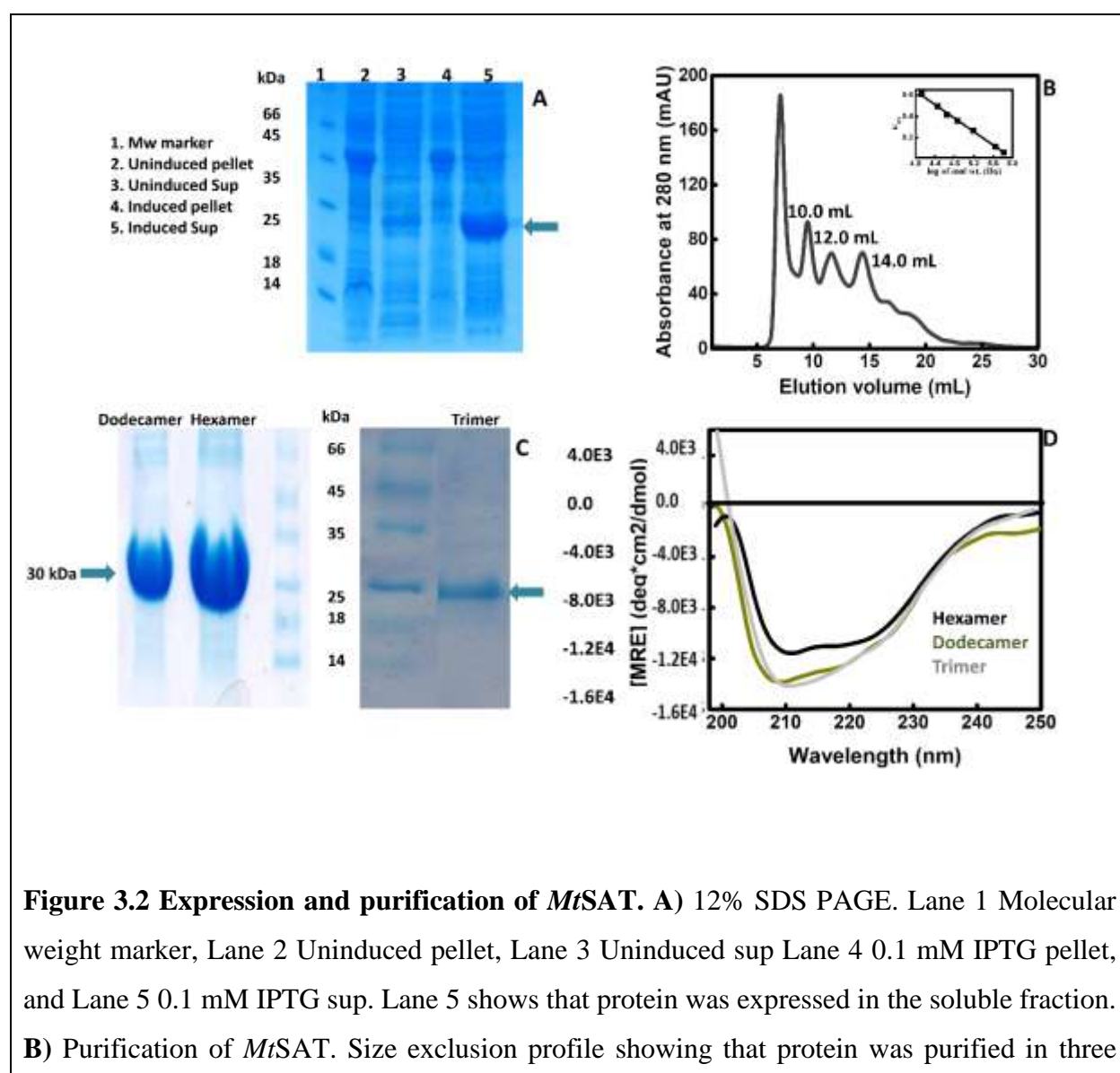
Figure 3.2 Expression and purification of *Mt*SAT. A) 12% SDS PAGE. Lane 1 Molecular weight marker, Lane 2 Uninduced pellet, Lane 3 Uninduced sup Lane 4 0.1 mM IPTG pellet, and Lane 5 0.1 mM IPTG sup. Lane 5 shows that protein was expressed in the soluble fraction. **B)** Purification of *Mt*SAT. Size exclusion profile showing that protein was purified in three

different oligomeric states and elutes at 10.0 mL, 12.0 mL, and 14.0 mL. **C)** 12 % SDS PAGE for purified *Mt*SAT fractions. The dodecamer, hexamer, and trimer fractions show that the proteins are more than 90 % pure. **D)** Circular dichroism spectrum showing that the dodecamer, hexamer, and trimer proteins were properly folded.

kDa respectively. Based on the theoretical molecular weight (also confirmed by mass spectrometry) for the monomeric subunit (26 kDa), the ~300 kDa, ~150 kDa, and ~75 kDa peak fractions may correspond to dodecameric, hexameric, and trimeric oligomeric states of the protein. It can be further inferred that the existence of multiple oligomeric states is a consequence of N-terminal truncation. The 12 % SDS PAGE was run for all three purified fractions. The SDS PAGE shows that all the purified *Mt*SAT fractions are more than 90 % pure (**Figure 3.2 C**). The Circular Dichroism experiments show that the eluted fractions were well folded and have typical αβ structures (**Figure 3.2 D**).

Analytical ultracentrifugation studies of Dodecamer, Hexamer, and Trimer fractions
Characterization of the dodecameric fraction

The different eluted fractions of *Mt*SAT protein were subjected to oligomeric characterization by the more rigorous AUC analysis method. The size exclusion method has a limitation as it cannot resolve different oligomers in equilibrium. Since the protein elutes as three different oligomers, each fraction must have all three oligomers present, although to different extents. Therefore, sedimentation velocity characterization of each fraction will reveal the extent of equilibrium. First the 10.0 mL dodecameric fraction was characterized. In the AUC sedimentation velocity experiment, the dodecameric fraction of *Mt*SAT shows multiple species as expected. So, we hypothesized that the different oligomeric species of dodecamer might occur due to variation at the N-terminal of *Mt*SAT protein. Also, from the size exclusion chromatography, it was observed that these oligomeric *Mt*SAT species were in equilibrium with each other. To understand the properties of this equilibrium better, we performed concentration-dependent velocity experiments on the dodecamer fraction. SEC purified 10.0 mL dodecamer fraction was concentrated and experiments were set up at different total protein concentrations (6.0 μM, 20.0 μM, and 40.0 μM). Velocity profiles and analyses of experiments are shown (**Figure 3.3 A-C**). At lower protein

concentrations (6.0 μM), the 10.0 mL fraction display a major population of hexamer as compared to the dodecamer and another higher oligomer (Hexamer 50%, dodecamer 36%, and higher oligomer 14%) (**Figure 3.3 A**). Analyses of velocity profiles performed at the concentration of 20.0 μM, both hexameric and dodecameric fractions have almost equal distribution, suggesting that increasing total protein concentration pushes the equilibrium to the right, towards higher molecular weight species (Hexamer 44 %, dodecamer 45 %, and higher oligomer 10 %) (**Figure 3.3 B**). On the other hand, the fraction of higher oligomeric species has not changed, indicating that this could non-specific aggregation present in the SEC fraction. When the AUC velocity experiments were performed at 40.0 μM of total protein concentration, the dodecameric form dominates over the hexameric form (Hexamer 44 %, dodecamer 49 %, and higher oligomer 7 %), consistent with mass balance theory the higher oligomers are favored at high concentrations
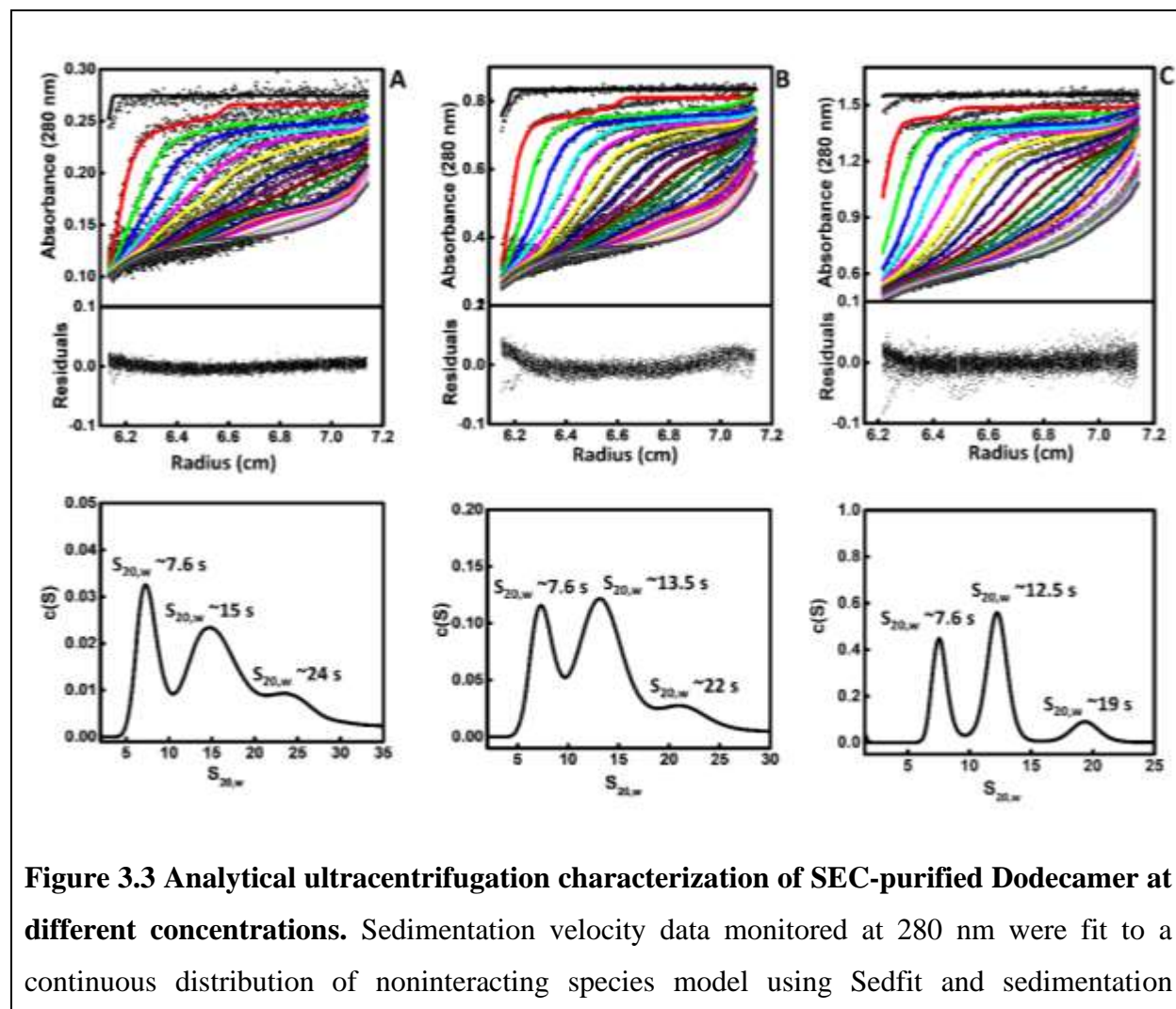
Figure 3.3 Analytical ultracentrifugation characterization of SEC-purified Dodecamer at different concentrations. Sedimentation velocity data monitored at 280 nm were fit to a continuous distribution of noninteracting species model using Sedfit and sedimentation

coefficient distribution c(s) as a function of the normalized sedimentation coefficient ($S_{20,w}$). **A), B), & C)** Velocity data of dodecameric *Mt*SAT at different concentrations (6.0 µM, 20.0 µM, and 40.0 µM) were fit (-) to a continuous distribution of noninteracting species model, and residuals are shown at the bottom of raw and fit data. Sedimentation coefficient distribution c(s) analyses at 6.0 µM show that the Dodecamer fraction consists of three species: first hexameric major species (50 %, $S_{20,w}$ ~ 7.6 s, and Mw ~ 156 kDa) and second dodecameric less predominant species (36%, $S_{20,w}$ ~ 15 s, and Mw ~ 300 kDa) and third minor high molecular weight species (14%, $S_{20,w}$ ~ 24 s Mw ~ 550 kDa). B) Increasing the dodecameric concentration from 6.0 µM to 40.0 µM shows that there is an increase in the dodecameric species as compare to hexameric species.

(**Figure 3.3 C**). From the results of concentration, we estimate the equilibrium constants for the transition to dodecameric state using the following equation

$Kd_{[D-H]}$ = (Fraction of Dodecamer * total protein concentration)/Fraction of Hexamer * Total protein concentration)

Results suggest having predominantly dodecameric above 40.0 µM and predominantly hexameric fraction below 40.0 µM. In summary, we designed and executed the first study to characterize the multi-oligomeric properties of *Mt*SAT and determined the boundaries of each oligomeric species.

Characterization of Hexameric fraction

Having characterized the composition and equilibrium boundaries of dodecameric fraction, we extended a similar approach to characterize the 12.0 mL hexameric fraction. The basic design of the experiment was once again to quantitatively analyze the composition of SEC purified 12.0 mL hexameric fraction. First, we collected SEC purified 12.0 mL fraction and performed velocity experiments at three different protein concentrations (6.0 µM, 12.0 µM, and 40.0 µM). Velocity profiles and analyses are shown (**Figure 3.4 A-C**). The sedimentation velocity experiments show that as we increase the hexamer concentration from 6.0 µM to 40.0 µM, the oligomerization of hexamer does not change. Results of sedimentation velocity experiments show that this fraction primarily dominated by the trimeric state (75%) which more or less sediments as a single species with a sedimentation coefficient of ~ 4.1 ± 0.3 s, which in turn corresponds to the molecular weight of ~75 ± 5 kDa (**Figure 3.4 A-C**). The hexameric form (20%) was also observed in equilibrium

with the trimeric form which sediments at ~7.6 s that corresponds to the molecular weight of ~150 kDa. The other minor peak (5%) was also observed at a higher sedimentation coefficient of ~12.5 showing the large aggregates with an estimated molecular weight of ~454 kDa. It is important to note that this higher oligomeric aggregate was also present in the SEC purified dodecameric fraction but its fraction remains constant. This also suggests that the high molecular weight aggregate contribution to the overall equilibrium would small as compared to contributions from both hexameric and dodecameric fractions. Analyses of equilibrium constants using the velocity studies confirm that hexameric species dominate between concentration ranges of 6.0 to 40.0 μM.

Figure 3.4 Analytical ultracentrifugation characterization of SEC-purified Hexamer at different concentrations. Sedimentation velocity data monitored at 280 nm were fit to a continuous distribution of noninteracting species model using Sedfit and sedimentation coefficient distribution c(s) as a function of the normalized sedimentation coefficient ($S_{20,w}$). **A), B), & C**) Velocity data of hexamer at different concentrations (6.0 μM, 12.0 μM, and 40.0 μM) were fit (-) to a continuous distribution of noninteracting species model, and residuals are shown at the bottom of raw and fit data. Sedimentation coefficient distribution c(s) analyses show that

by increasing the hexamer concentration from 6.0 µM to 40.0 µM we find the trimeric oligomer state is the predominant species while the hexameric state is the less predominant species in the concentration range.

Analytical ultracentrifugation of *Mt*SAT Trimeric fraction

Similarly, we characterized the 14.0 mL trimeric fraction in a concentration-dependent manner. We collected and concentrated SEC purified 14.0 mL fraction and performed velocity experiments at three different protein concentrations (6.0 µM, 20.0 µM, and 40.0 µM). Velocity profiles and analyses are shown in (**Figure 3.5 A-C**). The sedimentation velocity experiments show that as we increase the total protein concentration from 6.0 µM to 40.0 µM, the fraction of trimer does not change. The protein sediments as a predominantly single species with a sedimentation coefficient,
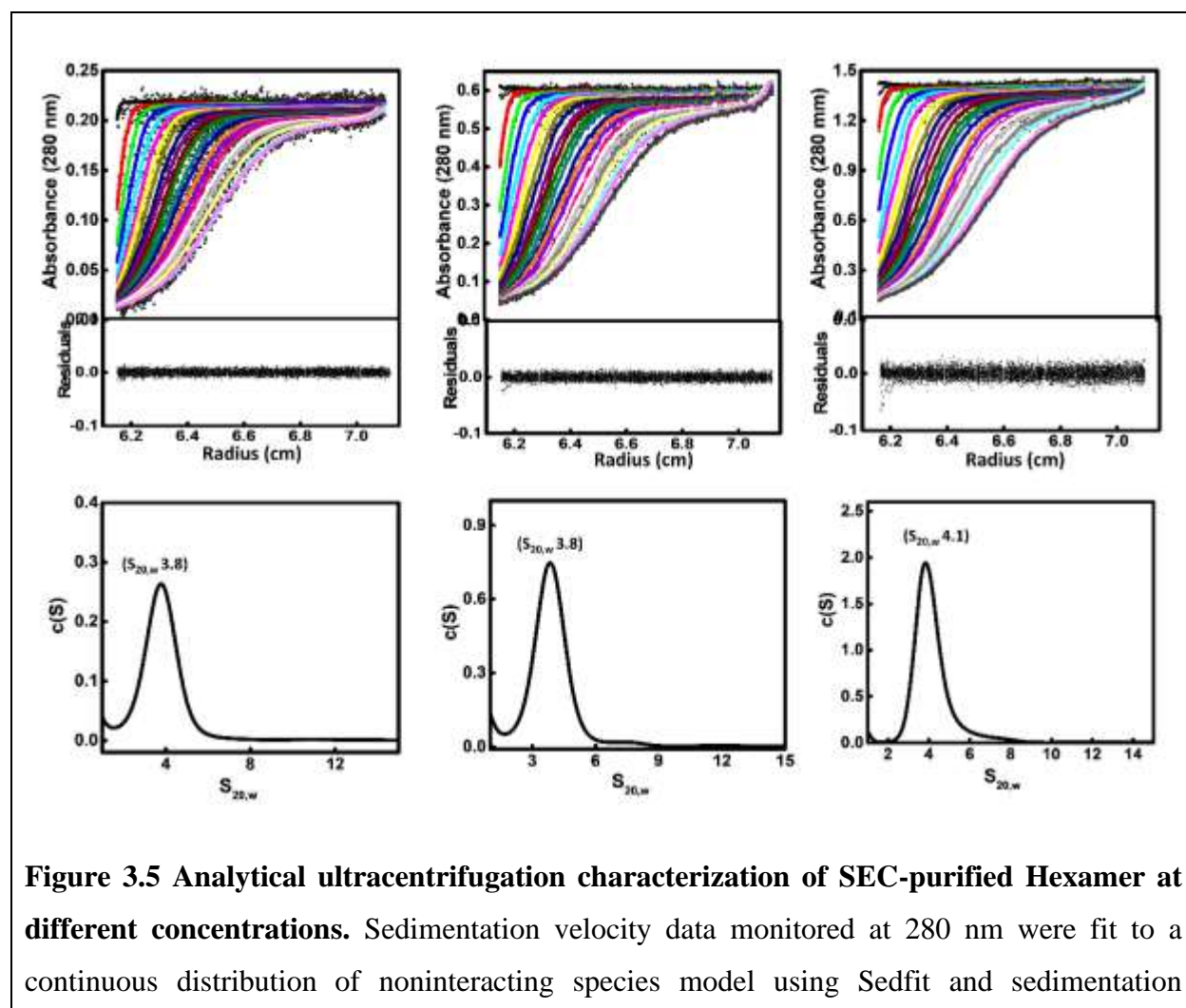
Figure 3.5 Analytical ultracentrifugation characterization of SEC-purified Hexamer at different concentrations. Sedimentation velocity data monitored at 280 nm were fit to a continuous distribution of noninteracting species model using Sedfit and sedimentation

coefficient distribution c(s) as a function of the normalized sedimentation coefficient ($S_{20,w}$). **A), B), & C**) Velocity data of hexamer at different concentrations (6.0 µM, 20.0 µM, and 40.0 µM) were fit (-) to a continuous distribution of noninteracting species model, and residuals are shown at the bottom of raw and fit data. Sedimentation coefficient distribution c(s) analyses show that by increasing the trimer concentration from 6.0 µM to 40.0 µM there is no change in the trimeric oligomer.

$S_{20,w}$ value of ~ 3.9 ± 0.1 s, an insignificant fraction of higher-order oligomers were also observed at 20 µM. Hexameric and dodecameric states were expected to populate at higher concentrations but not observed. One of the reasons we believe is that rate of dissociation is very high as compared to the rate of association which is directing the trimeric form to be a more stable oligomeric form followed by hexamer and dodecamer.

Catalytic properties of different oligomeric fractions of SEC

Mtb has successfully evolved to survive inside the macrophages where the pH ranges from 6.5 to 4.5 units. Therefore, we first used single-point activity assays to examine the pH-dependent activities of all three SEC purified fractions. It was found that all three fractions (10.0 mL, 12.0 mL, and 14.0 mL) were active but maximum activities for all three fractions were found to be at pH 7.5. Normalization of activity per monomer shows that hexameric fraction displays more activity per monomer at all pH, followed by dodecameric and trimeric fractions. It might be possible that cooperative interactions between monomers within the higher-order oligomers may be different to some extent in different oligomers and cooperative interactions within the hexameric state are better optimized for higher catalytic activity. We have performed the detailed characterization of equilibrium boundaries and kinetic parameters in the next section for quantifying and comparing the activity differences of all three oligomers.
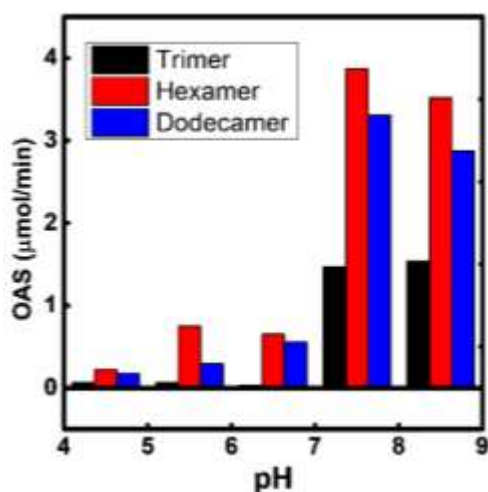
Figure 3.6 The single point activity of hexamer, dodecamer, and trimer fractions was performed at 5.0 mM L-serine concentration, and 1.0 μg total enzyme concentrations. The single point activity shows that different oligomeric fractions show different activities from pH 3.5 to 8.5. All the fractions show the highest activity at pH 7.5 and hexamer shows the highest activity among all the fractions.

Catalytic properties of the hexamer, dodecamer, and trimer

To quantify the catalytic power of each oligomeric state, we carried out detailed steady-state kinetic studies of all three fractions. We started with the hexameric fraction first. We determined initial velocities of the hexameric fraction at varying substrate concentrations and then analyzed the initial velocities as a function of substrate concentrations to obtain kinetic parameters, turnover rate (k_{cat}) and apparent substrate affinity, Michaelis-Menten constant (K_M). Analyses of Initial velocities as a function of substrate concentration and fit to the data are shown in (**Figure 3.7 A**). The L-serine concentration was varied while fixing Acetyl CoA concentration to 0.2 mM. The turnover rate, $k_{cat,}$ and the apparent substrate affinity, K_M of hexamer were found to be ~ 62 s^{-1} and ~1.6 mM respectively (**Table 1**). Similarly, steady-state kinetics of the dodecameric and the trimeric fraction was also performed (**Figure 3.7 B & C**). The dodecameric fraction shows higher k_{cat} than the trimeric fraction (47 s^{-1} and ~ 9 s^{-1} respectively) and lower K_m (1.6 mM and 1.9 mM respectively). Similar to results of single-point activity studies, a comparison of steady-state kinetic parameters of three oligomeric states show that hexameric fraction is more active both in

terms of turnover rate and substrate affinity. On the other hand, the trimeric fraction shows the lowest turnover rate and highest K_M than the other two oligomers. It might be possible that the active site within hexamer is well optimized for optimal substrate binding and catalytic activities.
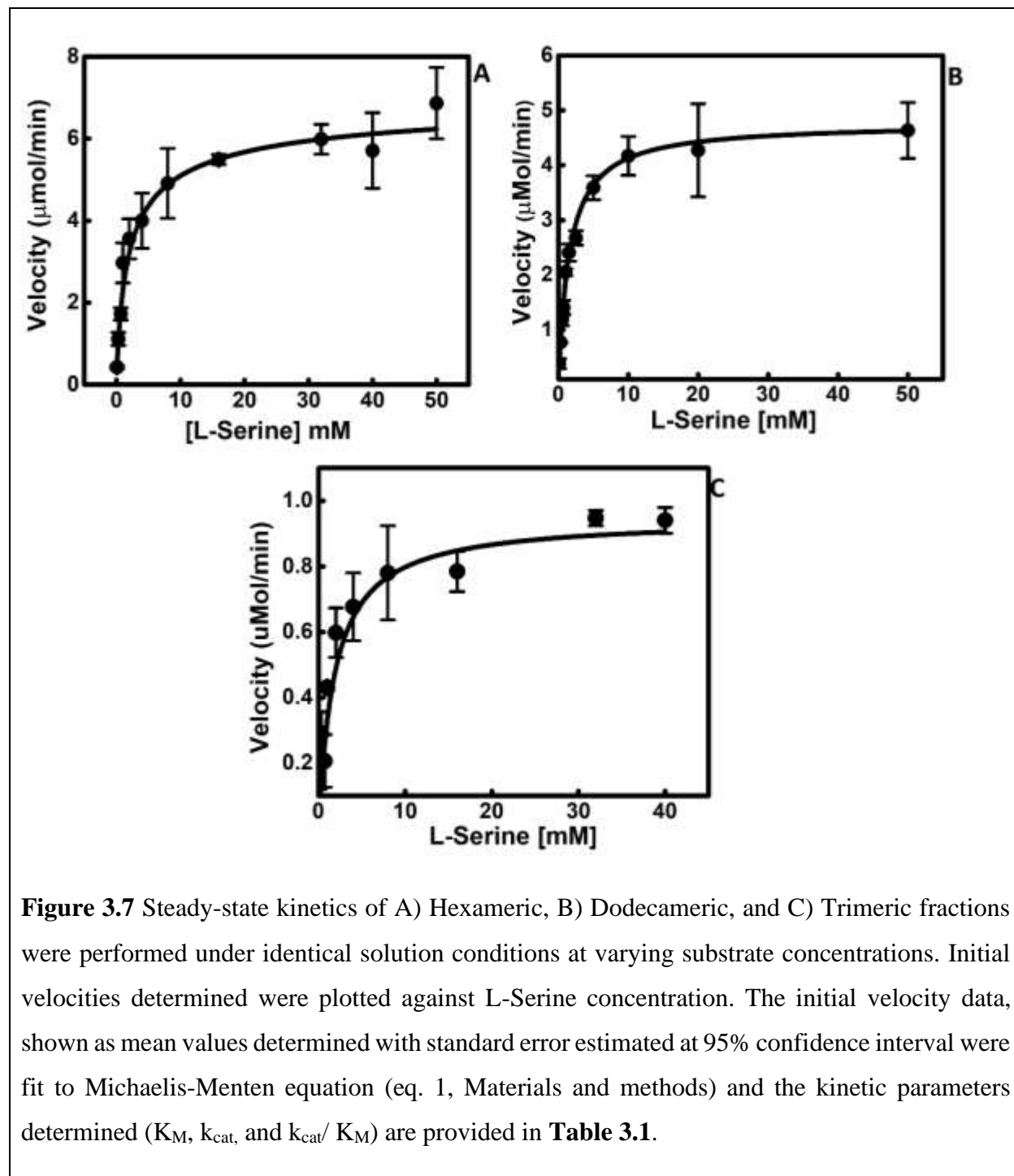
Figure 3.7 Steady-state kinetics of A) Hexameric, B) Dodecameric, and C) Trimeric fractions were performed under identical solution conditions at varying substrate concentrations. Initial velocities determined were plotted against L-Serine concentration. The initial velocity data, shown as mean values determined with standard error estimated at 95% confidence interval were fit to Michaelis-Menten equation (eq. 1, Materials and methods) and the kinetic parameters determined (K_M, k_{cat}, and k_{cat}/K_M) are provided in **Table 3.1**.

Table 3.1 Steady-state kinetic parameters for Hexamer, Dodecamer, and trimer.

Protein Type	K_m (mM) (Apparent affinity)	k_{cat} (sec^{-1}) (Turn over)	k_{cat} / K_m (sec^{-1} mM^{-1}) (Catalytic Efficiency)
Hexamer	1.6 ± 0.3	62 ± 5	39 ± 7
Dodecamer	1.6 ± 0.6	47 ± 1	29 ± 10
Trimer	1.9 ± 0.4	9 ± 1	5 ± 1

SAXS Experiments for the structure determination of MtSAT

To gain structural insight into the mechanistic and regulatory aspects of *Mt*SAT, we made lots of attempts to crystallize using a variety of crystallizing conditions. Due to multiple equilibria and associated heterogeneity, our crystallization attempts were not successful. Therefore, we employed the SAXS method to gain more information on the assembly state and get low-resolution structural information of the *Mt*SAT enzyme. The SAXS data collection was carried out at 4 °C at BM 29 beamline of ESRF using 20 μL of the protein sample solution. Analyses of scattering profiles of data collected at 2.0 mg/mL showed no inter-particular interaction or aggregation (**Figure 3.9 A inset**). The radius of gyration (Rg) was found to be ~4.5 nm from Guinier plot analysis. This analysis also. Further, the GNOM program was used to generate the p(r) distribution plot which provided an estimation of the maximum linear dimension (Dmax) value of ~10 nm of the scattering shape of a protein in solution. The output of GNOM forms the basis of low-resolution *ab initio* shape construction using simulated annealing optimization of a dummy atom set, as implemented in the DAMMIF. The DAMMIF runs gave 10 independent models and these were averaged using the DAMAVER suite of programs. After refinement, for the low bead occupancy positions and loosely connected average atom model, DAMFIT was used. This generated the most ideal and depictive envelope, that agrees with particle-excluded volume. The final model of *Mt*SAT shows that it is elongated in shape with a longitudinal axis of ~10 nm. In other bacterial systems like *Haemophilus influenzae*, the SAT exists as an elongated form which is also found in the *Mt*SAT enzyme
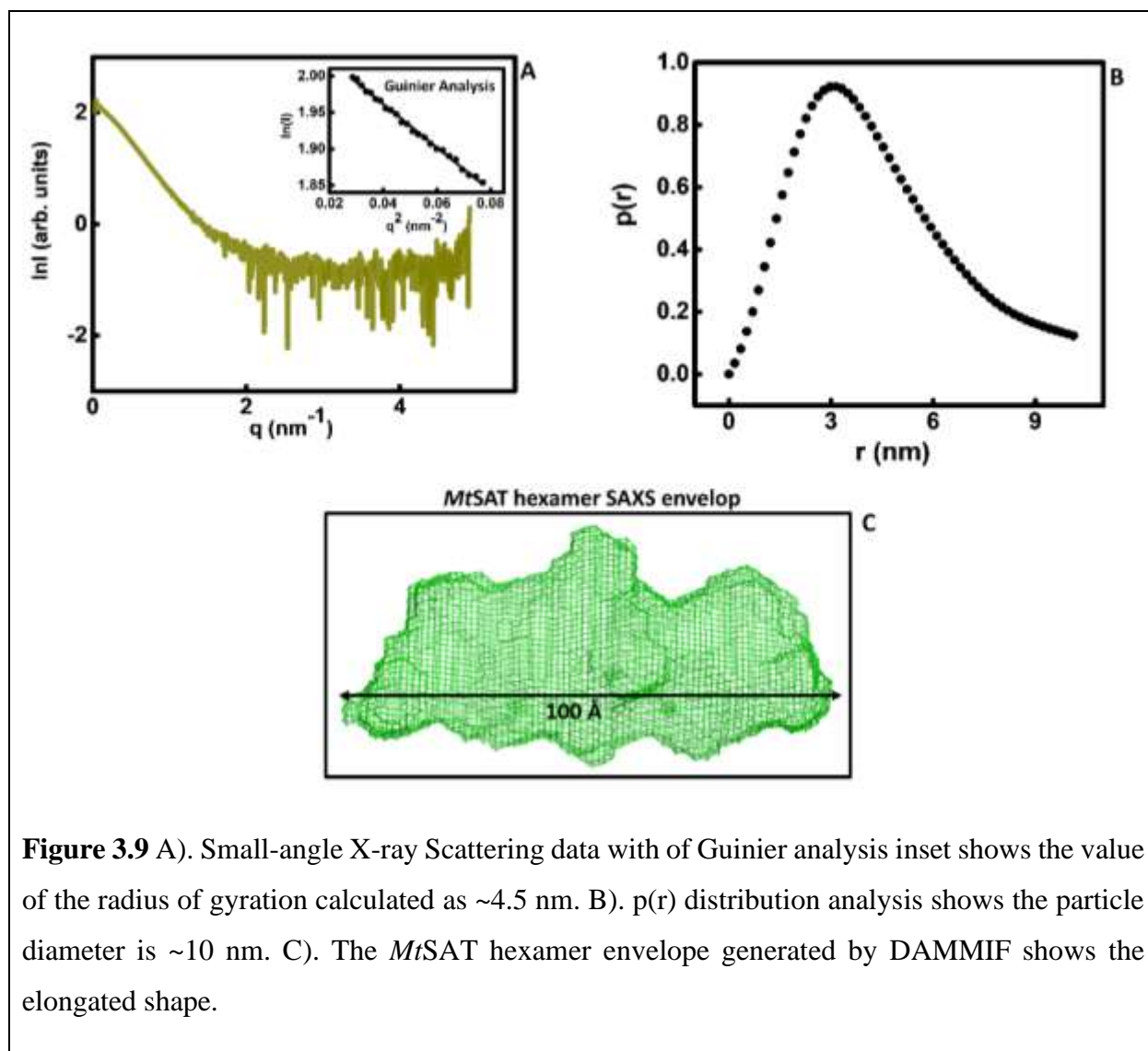
Figure 3.9 A). Small-angle X-ray Scattering data with of Guinier analysis inset shows the value of the radius of gyration calculated as ~4.5 nm. B). p(r) distribution analysis shows the particle diameter is ~10 nm. C). The *Mt*SAT hexamer envelope generated by DAMMIF shows the elongated shape.

Discussion

The oligomerization and regulation of multienzyme complexes are one of the most crucial mechanisms to survive the pathogens in adverse environmental conditions like high salt concentrations, extremely low and high pH ranges, lower and higher temperature. The CRC complex consists of two enzymes; serine acetyltransferase (SAT), which synthesizes OAS, and CS, which synthesizes cysteine from OAS. The SAT enzyme, in different bacteria, exists as a hexamer while the CS enzyme exists as a dimer in solution. However, SAT is found to exist as a trimer and hexamer in *Glycine max* depending upon the concentration of enzyme in solution while in *Entamoeba histolytica* it solely exists as in trimer only in solution which was suggested due to the variation of N-terminal amino acids. The SAT of *Mycobacterium tuberculosis* shows it has ~ 98 amino acid stretch missing at its N-terminus. This is striking because Mycobacterial enzymes always behave differently from the other organism's enzymes. *Mt*SAT enzyme has a molecular weight of ~ 25 kDa which is the lowest molecular weight of SAT enzyme studied to date. So, we cloned, expressed, and purified this enzyme to study the oligomeric assembly and catalytic properties of the *Mt*SAT enzyme. Due to the absence of an amino acid stretch at the N-terminal region, this enzyme was highly prone to be aggregated and so the purification of this enzyme took lots of standardization and optimization of purification conditions. Finally, we were able to purify this enzyme and found that the *Mt*SAT enzyme exists in three different oligomers; dodecamer, hexamer, and trimer. This is striking because in other bacterial SAT enzymes to date only trimer and hexameric SAT oligomers have been discovered and the presence of a dodecameric form of *Mt*SAT might give it an extra advantage to survive in an adverse oxidizing environment inside the host. The existence of *Mt*SAT in three oligomeric states might be due to the reasons for its N-terminal amino acid truncation in the *Mtb*. Further, we characterized the three oligomers by AUC experiments to know the concentrations dependent oligomerization of different oligomers. AUC velocity experiments showed that the dodecameric fraction can exist in two oligomeric forms; dodecamer, and hexamer. The concentration-dependent AUC experiments show that as we increase the concentration of dodecamer, there is an increase in the dodecameric fraction as compared to the hexameric fraction which shows the concentration-dependent behavior of dodecamer. However, the hexameric fraction shows that it predominantly exists in the trimeric form while also having hexameric form in the equilibrium and this equilibrium does not get changed even at high hexamer concentration. Further, the trimeric fraction shows that it does not

exist in equilibrium with other oligomeric forms even at high protein concentrations. As *Mtb* evolved to survive inside the macrophages where the pH ranges from 4.5 to 6.5 units therefore, we performed the activity experiments to know whether these oligomers show activity or not at different pHs. We found that all three fractions are showing the activity at pHs ranging from pH 3.5 to pH 8.5. However, the extent of activities is different at pHs. All oligomers displayed maximum activity at pH 7.5 and hexameric fraction showed the highest activity, followed by dodecamer and trimer. The hexameric form of SAT is mostly used by other bacterial species, so it might be possible that the hexameric state of *Mt*SAT is the predominant form of the enzyme in *Mtb* under normal conditions inside the host cell. Further, the detailed steady-state kinetic characterization of different oligomers showed that hexamer has the highest catalytic turnover rate, followed by dodecamer and trimer. These results clearly show that ~ 98 N-terminal amino acid stretch of *Mt*SAT are mainly responsible for the presence of multiple oligomeric species, the concentration-dependent oligomerization of dodecamer, and differences in catalytic efficiencies. We have also studied the low-resolution structural features of the hexameric state by SAXS and showed that hexamer has an elongated structure in line with the other SAT's elongated shape.

Chapter 4

Characterization of thermodynamic and kinetic properties of CRC assembly

Research Problem

The first experimental evidence for the existence of the CRC complex in *Salmonella* was provided by Kredich in 1966. He purified the CRC complex from the crude cell extract and assayed the CS activity in the supernatant and found that 5% of CS activity was associated with CRC complex and the rest of 95% came from free CS. By size exclusion chromatography approach, he estimated the molecular weight of the CRC complex as ~309 kDa and also determined the stoichiometry of the complex by dissociating the CRC complex in the presence of OAS into one SAT hexamer and two CS dimers (Kredich et al., 1969; Kredich & Tomkins, 1966a). Many studies followings the pioneering work of Kredich provided more support to his discovery and reported the presence of CRC in other bacteria as well as in plants. However, detailed information on the structure of CRC, assembly, and disassembly mechanisms, and regulatory features remain unknown. Only very recently, attempts from our lab and few other groups have started providing additional information on dynamic assembly-dissociation features of CRC.

This study aims to answer two major questions. First to address whether *Mycobacterium* cysteine biosynthesis follows similar regulatory features like that of other bacteria. Although *Mycobacterium* CS is both structurally and biochemically characterized well, no such details are available for *Mt*SAT. Therefore, no information on the existence, structure, assembly, and regulatory details of CRC in *Mycobacterium* is available. Further, dissecting the complex dynamics of CRC, in general, is essential for understanding the physiological relevance as well as design principles of the regulatory circuit of cysteine biosynthesis. Few of the known features are modulation of catalytic efficiencies of component enzymes in the CRC and ligand sensitive dissociation dynamics of the complex. Studies from our lab and others showed the activity of SAT increases but that of CS decreases upon complex formation. Similarly, our previous work showed how CRC is dynamically dissociated (Kaushik et al., 2017).

Here we provide a first insight into how *Mt*SAT of *Mycobacterium* is different from other bacterial and plant SATs and how differences in the assembly properties of *Mt*SAT manifest in the *Mt*CRC. We provide experimental evidence for striking differences between *Mt*CRC and *St*CRC and argue that the difference in *Mt*SAT is the major contributor.

Results

Purification of CRC complex from Mycobacterium tuberculosis

For purification of the *Mt*CRC complex, we first purified two-component enzymes, *Mt*SAT and *Mt*CS of the *Mt*CRC complex. As discussed in chapter 3, the first enzyme of the cysteine regulatory complex is SAT which synthesizes the OAS from the L-serine and Acetyl CoA. The *Mt*SAT was eluted in three different fractions 10.0 mL, 12.0 mL, and 14.0 mL and the standards run in the column show the molecular weight of eluted fractions as ~300 kDa ~150 kDa, and ~75 kDa respectively. The most common oligomeric state of SAT is hexamer so we started the complex formation of *Mt*CRC by mixing the *Mt*SAT hexamer with *Mt*CS enzyme in an appropriate ratio. The *Mt*CS enzyme was previously cloned by Dr. Srijita, a former Ph.D. student of our lab, and this clone was used for purification of the *Mt*CS enzyme. The *Mt*CS was purified by the Ni-NTA affinity chromatography and further purified by size exclusion chromatography (SEC). The *Mt*CS eluted at ~80.0 mL which shows a molecular weight of ~70 kDa. From the molecular weight (35 kDa) of the monomeric protein, we deduced that the *Mt*CS exists as a homodimer in the solution (**Figure 4.1 A**).

For purification of *Mt*CRC, the SAT enzyme was purified in similar purification conditions as *Mt*CS. The hexameric form of the *Mt*SAT enzyme was extensively dialyzed and mixed with an appropriate ratio of *Mt*CS enzyme and kept for ~2 hrs at 4⁰ C and allowed to form the CRC complex. The *Mt*SAT and *Mt*CS enzyme interacted with each other and formed the *Mt*CRC complex. This *Mt*CRC complex was further purified by SEC. In the SEC, the *Mt*CRC complex eluted at ~59.0 mL (**Figure 4.1 B**). The extra amount of *Mt*CS has eluted as a sperate peak at ~ 80.0 mL.

From the standards run in the column (inset of **Figure 4.1 A**), we estimated the approximate molecular weight of the *Mt*CRC complex to be ~440 kDa. This translates into an *Mt*CRC complex that may consist of one *Mt*SAT hexamer in the center and four *Mt*CS dimers bound to the C-terminal tail of four monomers of SAT hexamer. We recently showed that CRC from *Salmonella* can assemble into two distinct complexes, *St*CRC1 and *St*CRC2 in which one *St*SAT hexamer bound to two *St*CS dimers and four *St*CS dimers respectively. However, we find that CRC of *Mycobacterium* consists of one *Mt*SAT hexamer which is ~150 kDa and it interacts with four ~70 kDa *Mt*CS dimers to form the ~440 kDa molecular weight assembly of *Mt*CRC complex in the ratio of 1:4 CRC complex.

Calculation of *Mt*CRC complex stoichiometry

Molecular mass of *Mt*SAT monomer = 26 kDa

Molecular mass of a *Mt*SAT hexamer (26 kDa*6) = 156 kDa

The molecular mass of the complex = ~440 kDa

Mol. mass of *Mt*CRC complex – Mol. mass of *Mt*SAT hexamer = (~440 kDa - ~156 kDa) = ~284 kDa

Molecular mass of *Mt*CS monomer = 35 kDa

Hence the number of *Mt*CS monomers present along with a hexamer of *Mt*SAT in the complex will be: 284 / 35 = 8.11 (~8 monomers or 4 dimers)

Stoichiometry of the *Mt*CRC complex = No. of monomers of *Mt*SAT / No. of monomers of *Mt*CS = 6 / 8

So, the stoichiometry of the *Mt*CRC is 1:4 or one mole of SAT hexamer bound with four moles of CS dimer.

Further, the 12 % SDS PAGE was performed of the eluted *Mt*CRC complex to know the identities of component enzymes of *Mt*CRC (**Figure 4.1 C**). In the SDS PAGE, the *Mt*CRC complex separates into two bands one showing the molecular at ~35 kDa that corresponds to the *Mt*CS enzyme and the other band showing the molecular weight of ~30 kDa corresponding to the *Mt*SAT enzyme. From the SDS PAGE analysis, it is confirmed that the *Mt*CRC complex consists of the *Mt*SAT enzyme and MtCS enzyme. Overall, we confirmed the existence of the *Mt*CRC complex in the *Mycobacterium tuberculosis*.
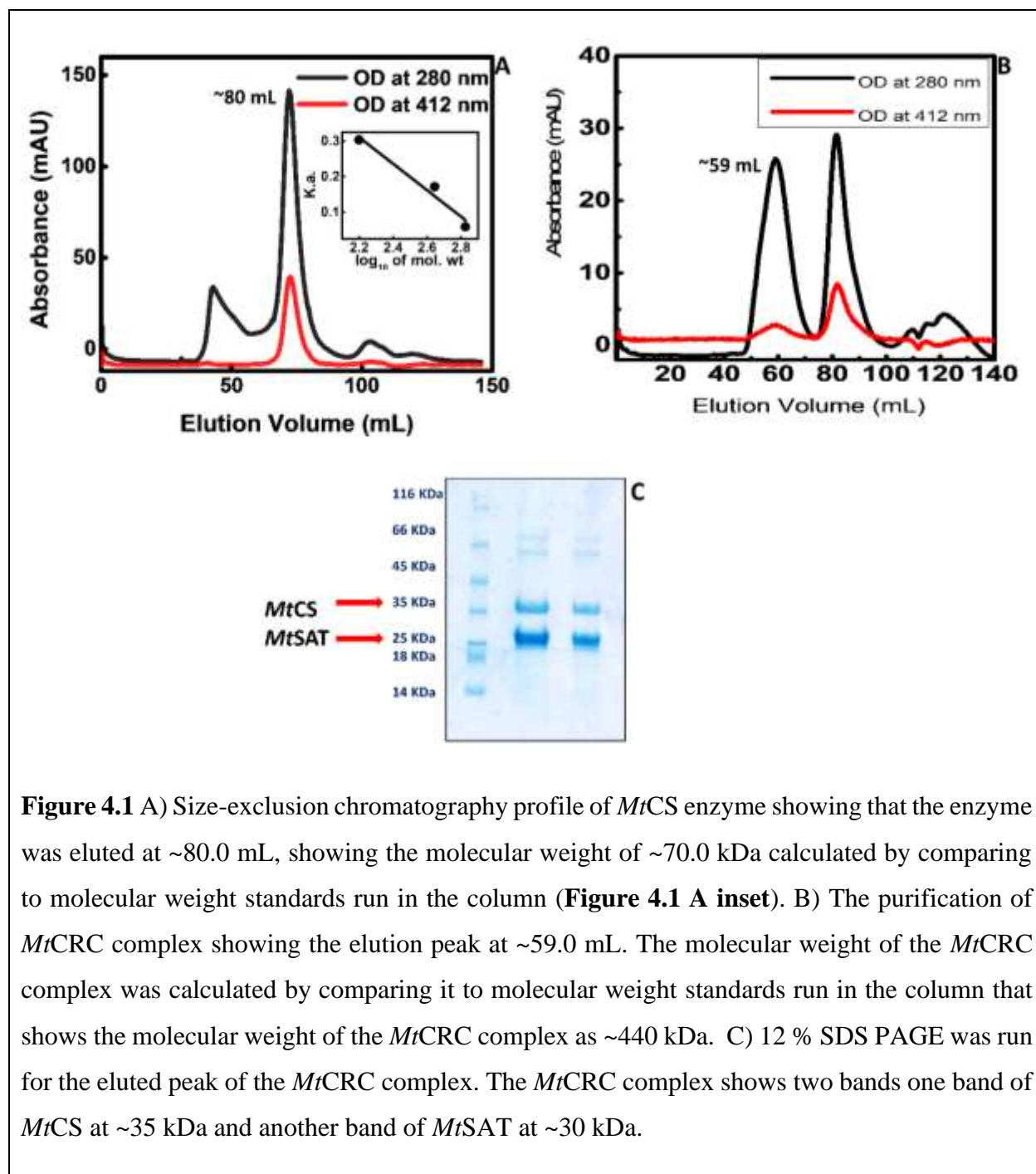
Figure 4.1 A) Size-exclusion chromatography profile of *Mt*CS enzyme showing that the enzyme was eluted at ~80.0 mL, showing the molecular weight of ~70.0 kDa calculated by comparing to molecular weight standards run in the column (**Figure 4.1 A inset**). B) The purification of *Mt*CRC complex showing the elution peak at ~59.0 mL. The molecular weight of the *Mt*CRC complex was calculated by comparing it to molecular weight standards run in the column that shows the molecular weight of the *Mt*CRC complex as ~440 kDa. C) 12 % SDS PAGE was run for the eluted peak of the *Mt*CRC complex. The *Mt*CRC complex shows two bands one band of *Mt*CS at ~35 kDa and another band of *Mt*SAT at ~30 kDa.

AUC sedimentation velocity experiments for MtCRC complex

After the estimation of the stoichiometry of the *Mt*CRC complex by size exclusion chromatography, we analyzed the eluted *Mt*CRC complex by analytical ultracentrifuge, a more rigorous and very sensitive method. The AUC velocity raw and fit data are shown in (**Figure 4.2**

A) and the residuals of fit and raw data are shown at the bottom. The velocity data were fit to the sedfit software and plotted as a continuous distribution of $S_{20,w}$ against the normalized sedimentation coefficient to water at 20 ºC ($S_{20,w}$). AUC velocity profiles are biphasic with one very slow-moving species along with a major species. Analyses of AUC velocity profiles show that almost single *Mt*CRC complex that sediments at the $S_{20,w}$ value of 11.01 ± 0.2 that corresponds to the molecular weight of 413 ± 10 kDa but the distribution is tailed towards right indicating that another high molecular weight species may also exist (**Figure 4.2 B**). In the previous study of *St*CRC from our lab, we have shown through the AUC velocity experiments that the *St*CRC exists in two different oligomeric forms lower molecular weight *St*CRC1 and higher molecular weight *St*CRC2. However, in the case of the *Mt*CRC complex, we observed only the *Mt*CRC2 complex as the major fraction. The observed stoichiometry of the *Mt*CRC complex was confirmed by AUC velocity experiments show which showed the stoichiometry as 1:4 (1 *Mt*SAT hexamer: 2 *Mt*CS dimers). Additional experiments are needed to characterize the higher molecular weight species observed along with the *Mt*CRC2 complex.

Figure 4.2 Analytical ultracentrifugation characterization of SEC-purified *Mt*CRC complex. A) Sedimentation velocity data monitored at 280 nm were fit to a continuous distribution of noninteracting species model using Sedfit and sedimentation coefficient distribution c(s) as a function of the normalized sedimentation coefficient ($S_{20,w}$). Residuals are shown at the bottom

of raw and fit data. B) Sedimentation coefficient distribution c(s) analyses of *Mt*CRC show that the eluted fraction consists mainly single peak of *Mt*CRC having $S_{20,w}$ value of 11.01 ± 0.2, and molecular weight of 440 kDa. The stoichiometry of *Mt*SAT hexamer to *Mt*CS dimer binding was found to be 1:4.
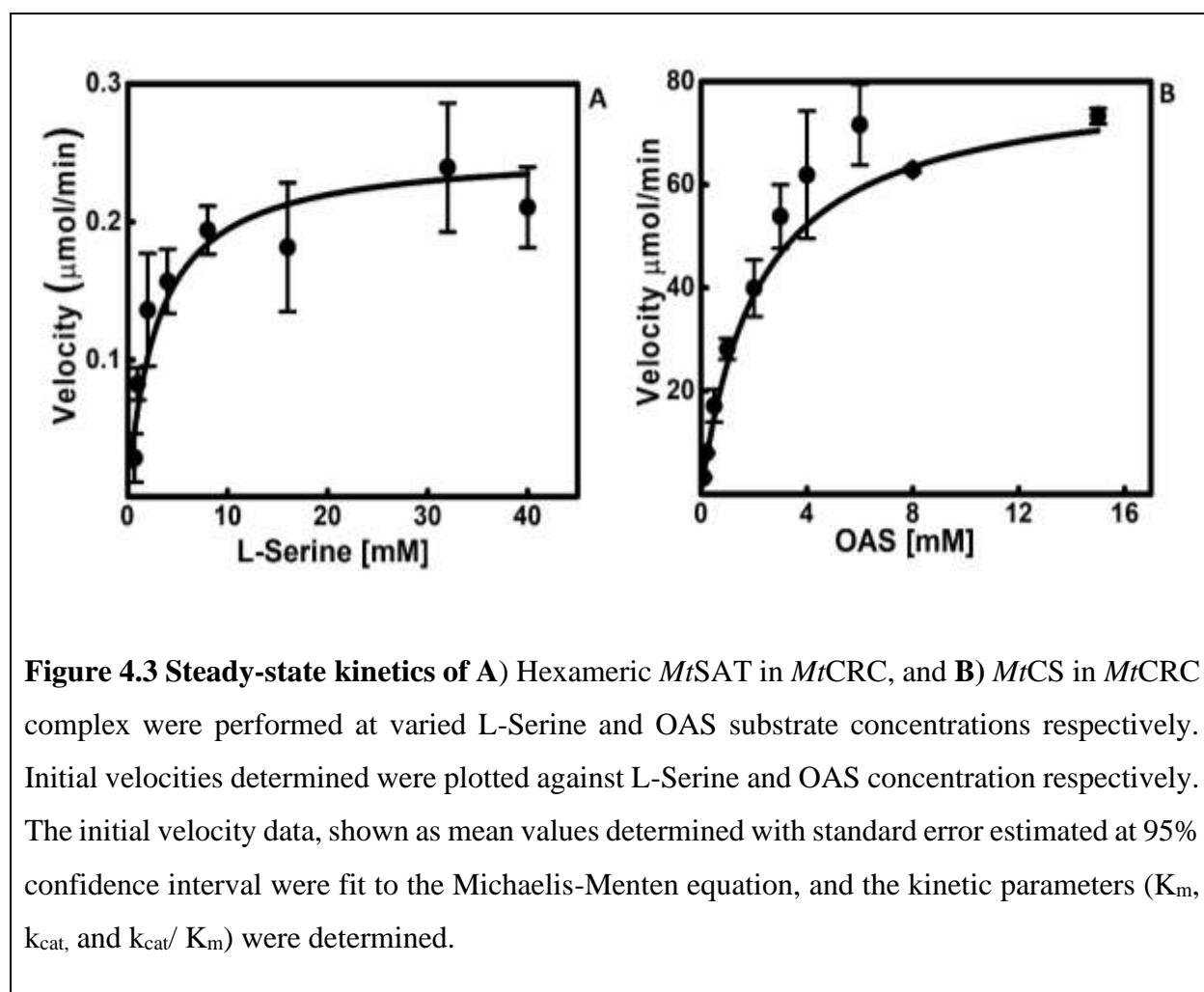
Kinetics of MtCRC complex

The enzyme activities of free SAT, free CS, SAT in CRC, and CS in CRC has been done from several bacteria and unicellular organisms. In all cases, the free SAT shows the lesser activity as compared to the activity of the SAT in the CRC. The opposite is true for CS as the CS in CRCs shows only ~30 % or less activity as compared to the activity of free CS. However, in *Mtb*, it is not reported how the enzymes within the CRC complex function. So, we performed the kinetic experiments to determine the enzyme activities of component enzymes of *Mt*CRC.

First, we purified the *Mt*CRC complex by size exclusion chromatography and assayed for *Mt*SAT activity. Initial velocities were measured by changing substrate concentrations as described in methods and initial velocities were plotted against the concentration of L-serine. The plot showed a hyperbolic trend expected for Michaelis-Menten kinetics and therefore, data were fitted to the Michaelis-Menten equation (Eq. 1 materials and methods). We determined the kinetic parameters; k_{cat} was found to be 2.4 ± 0.2 s^{-1} which is ~26 folds less than that of free *Mt*SAT (62 s^{-1}) (**Table 4.1**). The catalytic efficiency of *Mt*SAT in *Mt*CRC shows the value of 1.0 ± 0.1 s^{-1}mM^{-1} which is about 39 folds lower than the free *Mt*SAT. These observations are in contrast to activity enhancements reported for other bacterial and plant SAT enzymes in CRC. It shows that somehow the activity of SAT is compromised by complex formation only in *Mycobacterium*. But in other bacteria, the SAT activity increases after complex formation (e.g., *Haemophilus influenzae* SAT). As *Mt*SAT lacks the first 98 amino acids at N-terminus as compared to other SATs, so this might be the reason that it behaves differently in *Mtb* than other SATs. Also, the *Mt*SAT in *Mt*CRC displays lower affinity (3.0 ± 0.3 mM) towards L-serine as compared to the free form *Mt*SAT (~1.6 mM). Lower affinity could also be the reason for the lower catalytic efficiency of *Mt*SAT in the *Mt*CRC complex. Overall, the results show that the *Mt*SAT in the complex is less active as compared to the free form.

After characterizing the *Mt*SAT activity in the *Mt*CRC complex, we determined the kinetic parameters of *Mt*CS in the *Mt*CRC complex. Initial velocities were estimated and velocity data

were plotted against the concentration of OAS. Then the kinetic parameters k_{cat} and K_m were calculated by fitting the initial velocity data to the Michaelis-Menten model (Eq. 1 materials and methods). The *Mt*CS enzyme shows K_m and k_{cat} values of 2.2 ± 0.14 mM, and 444 s^{-1} respectively. The catalytic efficiency of *Mt*CS was calculated to be 202 ± 14 $s^{-1}mM^{-1}$. Surprisingly, the catalytic efficiency of *Mt*CS upon complex formation is not reduced as much observed in other CS enzymes and somehow it is able to recruit the OAS at the *Mt*CS active site. This is in striking contrast to the behavior of other bacterial and plant CS in CRC. In all cases reported so far, whether, in bacteria or plants, the cysteine synthesis catalytic efficiency of CS is severely reduced to 70 % or more when CS is in complex with SAT or in the CRC. *Mycobacterium* catalytic efficiencies of free *Mt*CS and *Mt*CS within the *Mt*CRC are highly comparable, very surprising, and require detailed investigation.

Figure 4.3 Steady-state kinetics of A) Hexameric *Mt*SAT in *Mt*CRC, and **B)** *Mt*CS in *Mt*CRC complex were performed at varied L-Serine and OAS substrate concentrations respectively. Initial velocities determined were plotted against L-Serine and OAS concentration respectively. The initial velocity data, shown as mean values determined with standard error estimated at 95% confidence interval were fit to the Michaelis-Menten equation, and the kinetic parameters (K_m, k_{cat}, and k_{cat}/ K_m) were determined.
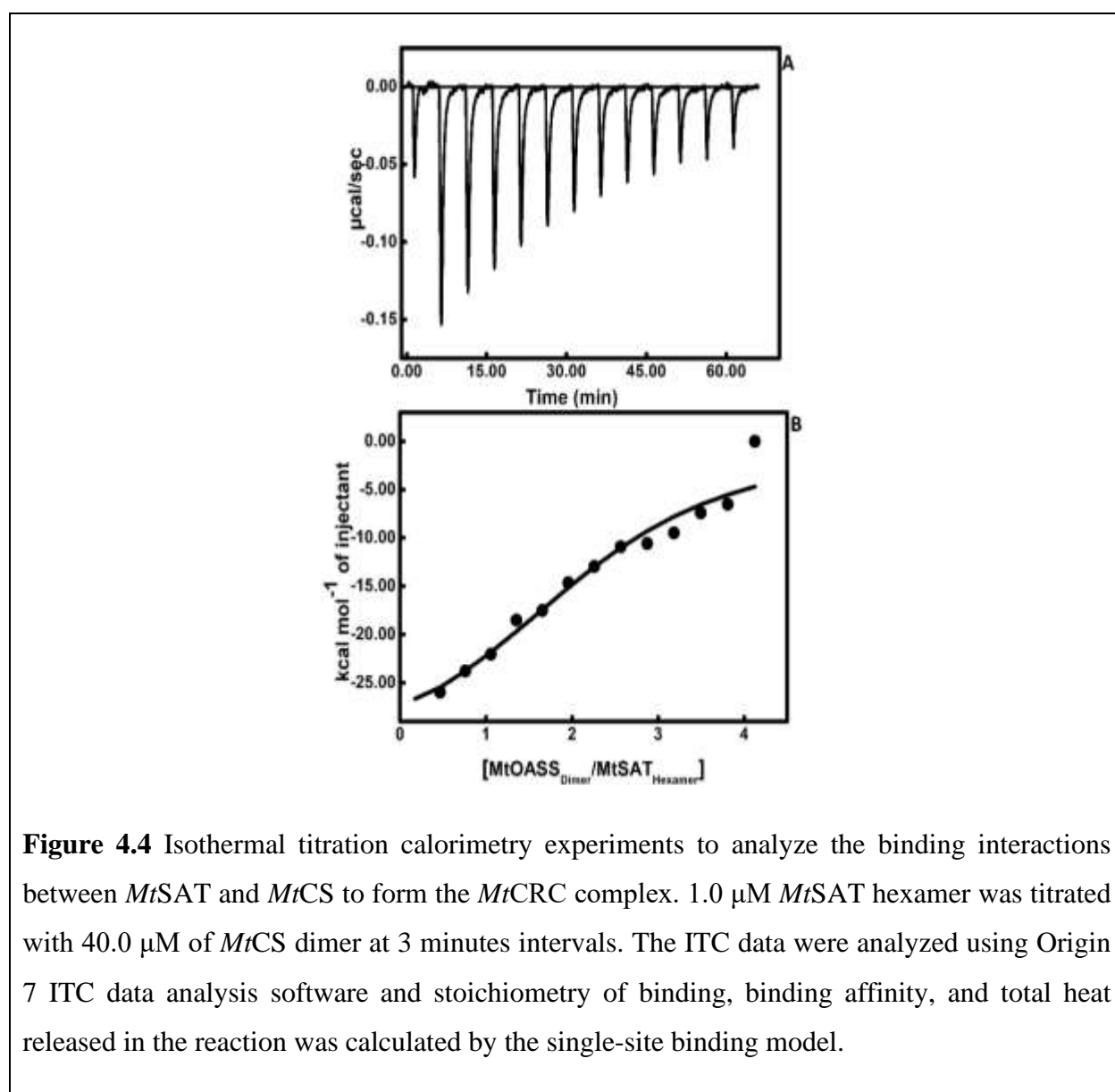
Table 4.1 Steady-state kinetic parameters for *Mt*SAT and *Mt*CS in the *Mt*CRC complex.

Protein Type	K_m (mM) (Apparent affinity)	k_{cat} (sec^{-1}) (Turn over)	k_{cat} / K_m (sec^{-1} mM^{-1}) (Catalytic Efficiency)
MtSAT in MtCRC	3.0 ± 0.3	2.4 ± 0.2	1.0 ± 0.1
MtCS in MtCRC	2.2 ± 0.14	444 ± 12	204 ± 14

Characterization of thermodynamics of interactions between MtSAT and MtCS

Previously, the binding of full-length SAT and C-terminal tail of SAT with CS from bacteria like *Salmonella typhimurium, Escherichia coli,* and *Haemophilus influenzae* shows that they interact with very high affinity in the range of 900 nM to 1 nM. The binding of SAT to CS to form the CRC complex occurs in a stepwise manner. The first two CS dimers bind with very high affinity with SAT as compared to the third and fourth CS dimers. However, protein-protein interaction between *Mt*SAT and *Mt*CS has not been studied before. We performed the Isothermal Titration Calorimetry (ITC) experiments to obtain thermodynamic signatures of *Mt*SAT and *Mt*CS enzymes. SEC purified proteins were extensively dialyzed in the experimental buffer, degassed before performing the ITC experiments. *Mt*CS enzyme dimer (syringe concentration 40.0 μM) was injected into *Mt*SAT hexamer (1.0 μM) and titrated at an interval of 3 minutes **(Figure 4.4)**. The ITC data were analyzed using Origin 7 ITC data analysis software provided by the instrument manufacturer (Microcal, Inc.). The ITC experiments showed that the data could be fitted to a single-site binding model and therefore, we did not use any complex models. We calculated the stoichiometry of *Mt*CS to *Mt*SAT binding to be 4 *Mt*CS dimers to 1 *Mt*SAT hexamer and determined the binding affinity as 0.7 μM of the reaction, and total heat released in the binding interaction $(1.7 \pm 0.3) *10^4$ Cal/mol or 1.7 Kcal/mol. We also tried to fit the data by a two-site binding model but it could not be fitted. It shows that all the four *Mt*CS dimers bind to *Mt*SAT hexamer with similar affinities and not in a stepwise manner as observed for other bacteria and plants. The affinity between *Mt*SAT and *Mt*CS binding was in the micromolar range (Kd ~0.7 μM) which shows they interact with high affinity to form the *Mt*CRC complex. The binding affinity between *Hi*SAT and *Hi*CS also showed the micromolar range (Kd ~0.9 μM). However, the binding

affinity between the *St*SAT and *St*CS was found to be in the nanomolar range (Kd ~1.0 to 10.0 nM). From the thermodynamic experiments, it can be deduced that the *Mt*SAT and *Mt*CS binds with very high-affinity interaction to form the *Mt*CRC complex. However, the regulation of activities of component enzymes are very different and strikingly in contrast with observations of CRC from other bacteria.

Figure 4.4 Isothermal titration calorimetry experiments to analyze the binding interactions between *Mt*SAT and *Mt*CS to form the *Mt*CRC complex. 1.0 μM *Mt*SAT hexamer was titrated with 40.0 μM of *Mt*CS dimer at 3 minutes intervals. The ITC data were analyzed using Origin 7 ITC data analysis software and stoichiometry of binding, binding affinity, and total heat released in the reaction was calculated by the single-site binding model.

Purification of CRC complex from Salmonella typhimurium

To compare and study other latent mechanistic and regulatory features of CRC, we focused on the assembly, purification, and kinetic characterization of the *St*CRC complex. The clones of *St*SAT

and *St*CS were created by Dr. Mary, a former Ph.D. student of our lab. pET28a vector, cloned with genes was transformed into the Rosetta and BL21 (DE3) expression host cells respectively (materials and methods). For assembling the *St*CRC complex, we mixed SEC purified *St*SAT enzyme with *St*CS. The *St*SAT enzyme was eluted at ~68 mL which corresponds to the molecular weight of ~180 kDa estimated by the standards run in the column. From the molecular weight of the monomeric *St*SAT (30 kDa), it was confirmed that the *St*SAT eluted as a hexamer (**Figure 4.5 A**). Similarly, *St*CS enzyme was purified that elutes at ~80 mL and corresponds to a molecular weight of ~70 kDa. From the monomeric molecular weight of the *St*CS (35 kDa), the enzyme was confirmed as a dimer (**Figure 4.5 B**).

The *St*CRC complex was assembled by mixing *St*SAT hexamer and *St*CS dimer in an appropriate ratio and kept for 2 hrs at 4⁰ C. This complex was then further purified by the size exclusion chromatography. The *St*CRC complex elutes in two oligomeric forms, one higher molecular weight (*St*CRC2) at ~57 mL and the second lower molecular weight (*St*CRC1) at ~63 mL (**Figure 4.5 C**). The estimated molecular weight of *St*CRC2 and *St*CRC1 was found to be ~460 kDa and ~320 kDa respectively calculated by the standards run in the SEC column. From the molecular weight of the *St*SAT hexamer and *St*CS dimer, we have estimated the stoichiometry of the complexes as follows.

Calculation of *St*CRC1 and StCRC2 complex stoichiometry

Actual molecular mass of *St*SAT monomer = 30 kDa

Molecular mass of a *St*SAT hexamer (30*6) = 180 kDa

Actual molecular mass of *St*CS monomer = 35 kDa

Molecular mass of a *St*CS dimer (35*2) = 70 kDa

Molecular mass of the *St*CRC1 complex = 320 kDa

Mol. mass of complex – Mol. mass of StSAT hexamer = (320 kDa - 180 kDa) = 140 kDa

Hence the number of *St*CS monomers present along with a hexamer of *St*SAT in the *St*CRC1 complex will be: 140 / 35 = 4 (~4 monomers or 2 dimers)

Stoichiometry of the *St*CRC1 complex = No. of monomers of *St*SAT / No. of monomers of *St*CS = 6 / 4

So, the stoichiometry of the *St*CRC1 complex is 1:2 or one mole of *St*SAT hexamer bound with two moles of *St*CS dimer.
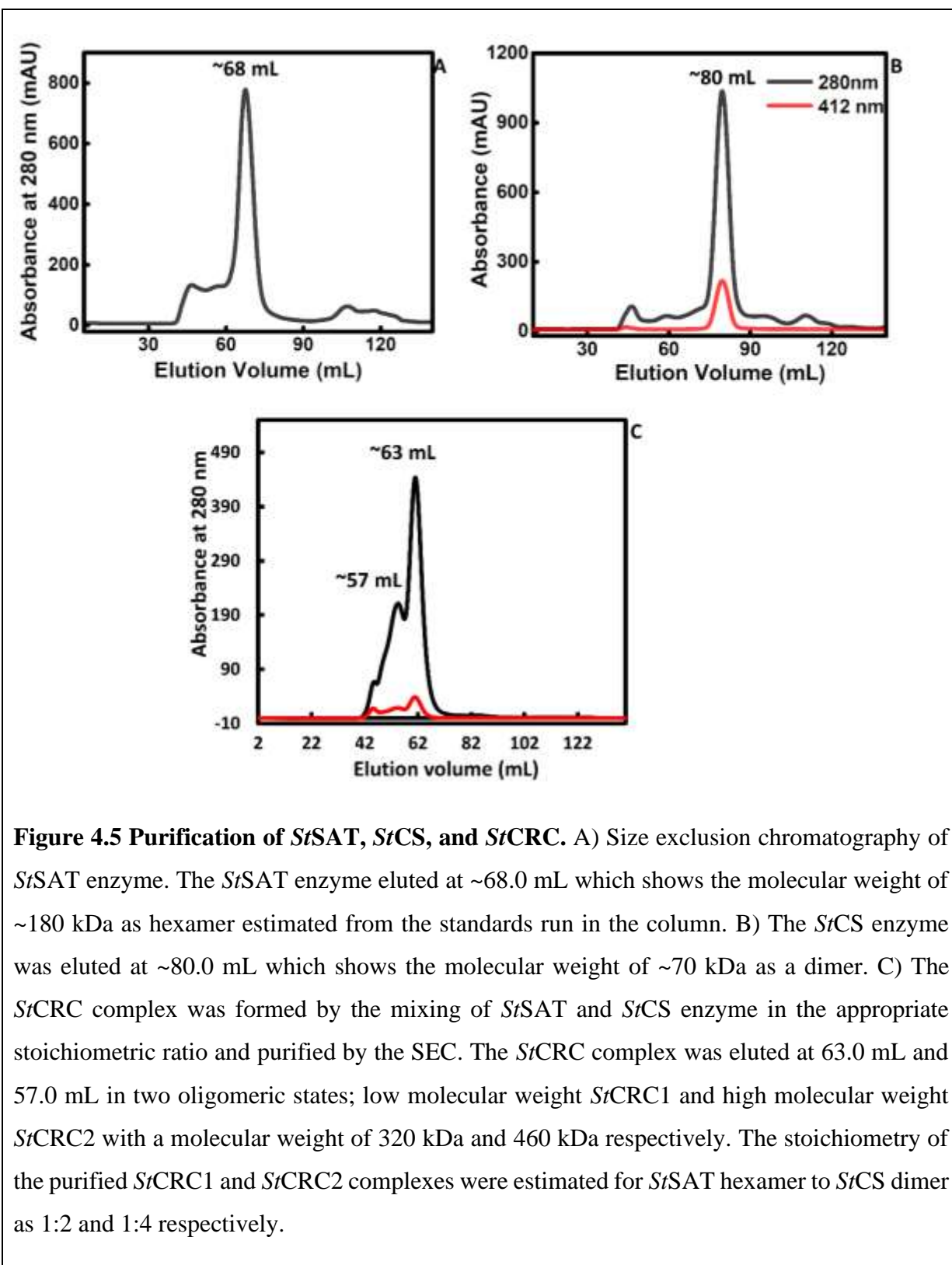
Figure 4.5 Purification of *St*SAT, *St*CS, and *St*CRC. A) Size exclusion chromatography of *St*SAT enzyme. The *St*SAT enzyme eluted at ~68.0 mL which shows the molecular weight of ~180 kDa as hexamer estimated from the standards run in the column. B) The *St*CS enzyme was eluted at ~80.0 mL which shows the molecular weight of ~70 kDa as a dimer. C) The *St*CRC complex was formed by the mixing of *St*SAT and *St*CS enzyme in the appropriate stoichiometric ratio and purified by the SEC. The *St*CRC complex was eluted at 63.0 mL and 57.0 mL in two oligomeric states; low molecular weight *St*CRC1 and high molecular weight *St*CRC2 with a molecular weight of 320 kDa and 460 kDa respectively. The stoichiometry of the purified *St*CRC1 and *St*CRC2 complexes were estimated for *St*SAT hexamer to *St*CS dimer as 1:2 and 1:4 respectively.

Similarly, for StCRC2 complex

Mol. mass of complex – Mol. mass of *St*SAT hexamer = (460 kDa - 180 kDa) = 280 kDa

Hence the number of *St*CS monomers present along with a hexamer of *St*SAT in the *St*CRC2 complex will be: 280 / 35 = 8 (~8 monomers or 4 dimers)

Stoichiometry of the *St*CRC2 complex = No. of monomers of *St*SAT / No. of monomers of *St*CS = 6 / 8

So, the stoichiometry of the *St*CRC2 complex is 1:4 or one mole of *St*SAT hexamer bound with four moles of *St*CS dimer.

In the size exclusion chromatography, we found that these two eluted *St*CRC complexes were present in the equilibrium with each other.

Catalytic properties of StSAT enzyme in presence of different stoichiometric ratios of StCS

It is known that after complex formation the SAT activities are enhanced while CS activities decrease. To know whether, in *Salmonella typhimurium,* the SAT activity also changes, we performed the activity experiments for *St*SAT in presence of *St*CS dimers at different stoichiometric ratios of *St*SAT hexamer to *St*CS dimers ($StSAT_{hexamer}$: $StCS_{dimer}$; 1:2, 1:4, and 1:6). The L-serine concentration was varied while fixing Acetyl CoA concentration to 0.2 mM. First of all, we measured the activity of free *St*SAT. The initial velocity data were fitted to the Michaelis-Menten equation and kinetics parameters were estimated (**Table 4.2**). The activity of free *St*SAT shows the K_m, k_{cat}, and k_{cat}/K_m values of 0.86 mM, 158 s^{-1}, and 184 s^{-1} mM^{-1} respectively (**Figure 4.6 A**). Then the activity of *St*SAT has performed in presence of two dimers of *St*CS in the form of the *St*CRC1 complex. The activity of *St*SAT in the *St*CRC1 complex shows the K_m, k_{cat}, and k_{cat}/K_m values of 0.46 mM, 193 s^{-1}, and 420 s^{-1} mM^{-1} respectively (**Figure 4.6 B**). The catalytic turnover rate and k_{cat}/K_m value shows that *St*SAT in the *St*CRC1 complex is more efficient in catalytic activity as compared to the free form *St*SAT. Then the *St*SAT activity was performed in presence of four *St*CS dimers in the form of an *St*CRC2 complex. The activity of *St*SAT in the *St*CRC2 complex form shows K_m, k_{cat}, and k_{cat}/K_m values of 0.48 mM, 212 s^{-1}, and 442 s^{-1} mM^{-1} respectively (**Figure 4.6 C**). The catalytic turnover rate and enzyme efficiency of *St*SAT in the *St*CRC2 complex increases very significantly. Then the activity of *St*SAT in presence of 6 *St*CS

dimers in the form of *St*CRC3 complex was performed. The activity of *St*SAT in the *St*CRC3 complex shows K_m, K_{cat}, and K_{cat}/K_m values of 0.47 mM, 208 s^{-1}, and 443 s^{-1} mM^{-1} respectively (**Figure 4.6 D**). In the *St*CRC3 complex, the catalytic turnover rate and enzyme efficiency do not increase further beyond the *St*CRC2 complex. However, an insignificant decrease of *St*SAT enzyme turnover rate was observed in the *St*CRC3 complex.
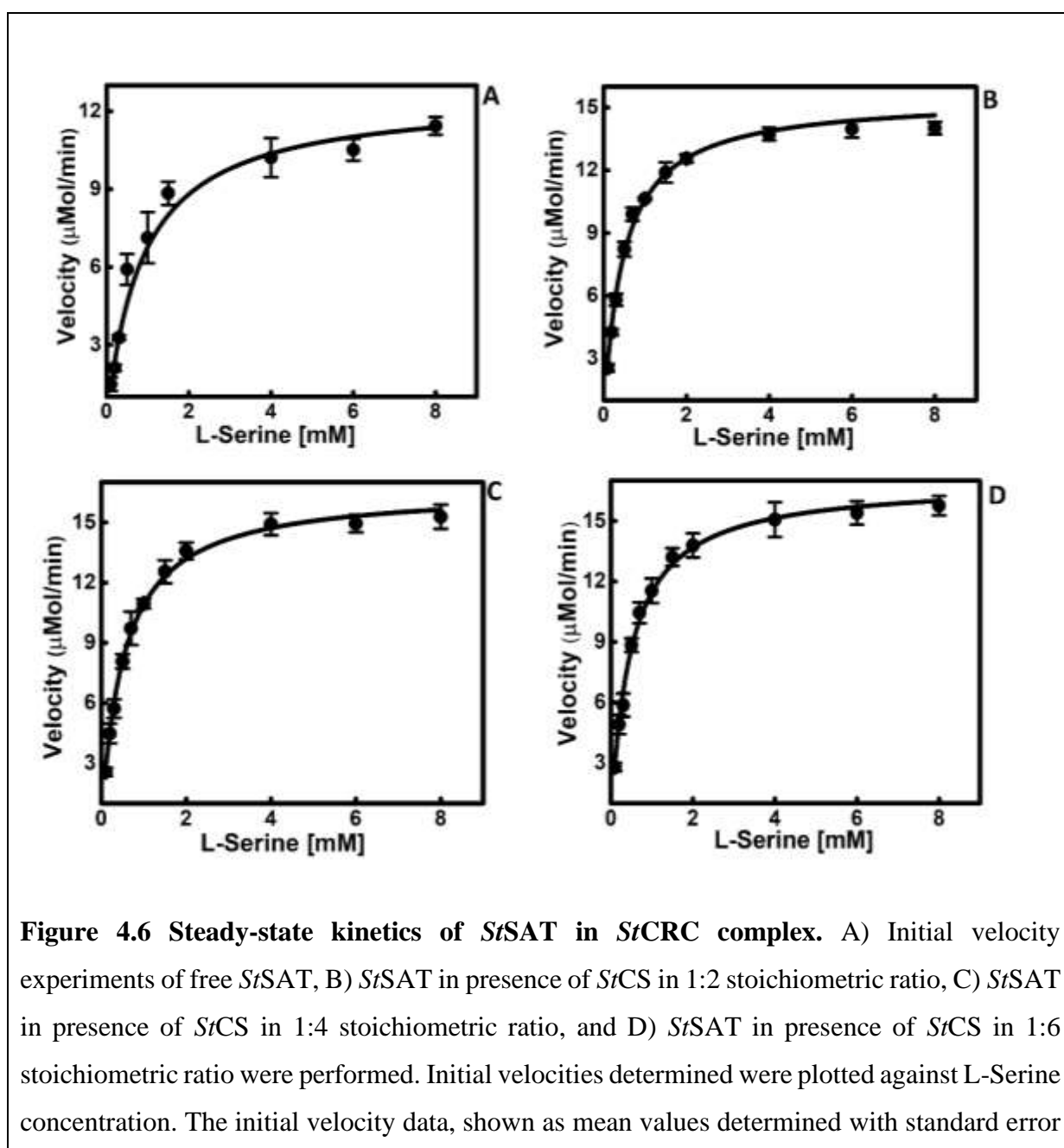
Figure 4.6 Steady-state kinetics of *St*SAT in *St*CRC complex. A) Initial velocity experiments of free *St*SAT, B) *St*SAT in presence of *St*CS in 1:2 stoichiometric ratio, C) *St*SAT in presence of *St*CS in 1:4 stoichiometric ratio, and D) *St*SAT in presence of *St*CS in 1:6 stoichiometric ratio were performed. Initial velocities determined were plotted against L-Serine concentration. The initial velocity data, shown as mean values determined with standard error

estimated at 95% confidence interval were fit to the Michaelis-Menten equation and kinetic parameters (K_m, k_{cat}, and k_{cat}/K_m) were determined.
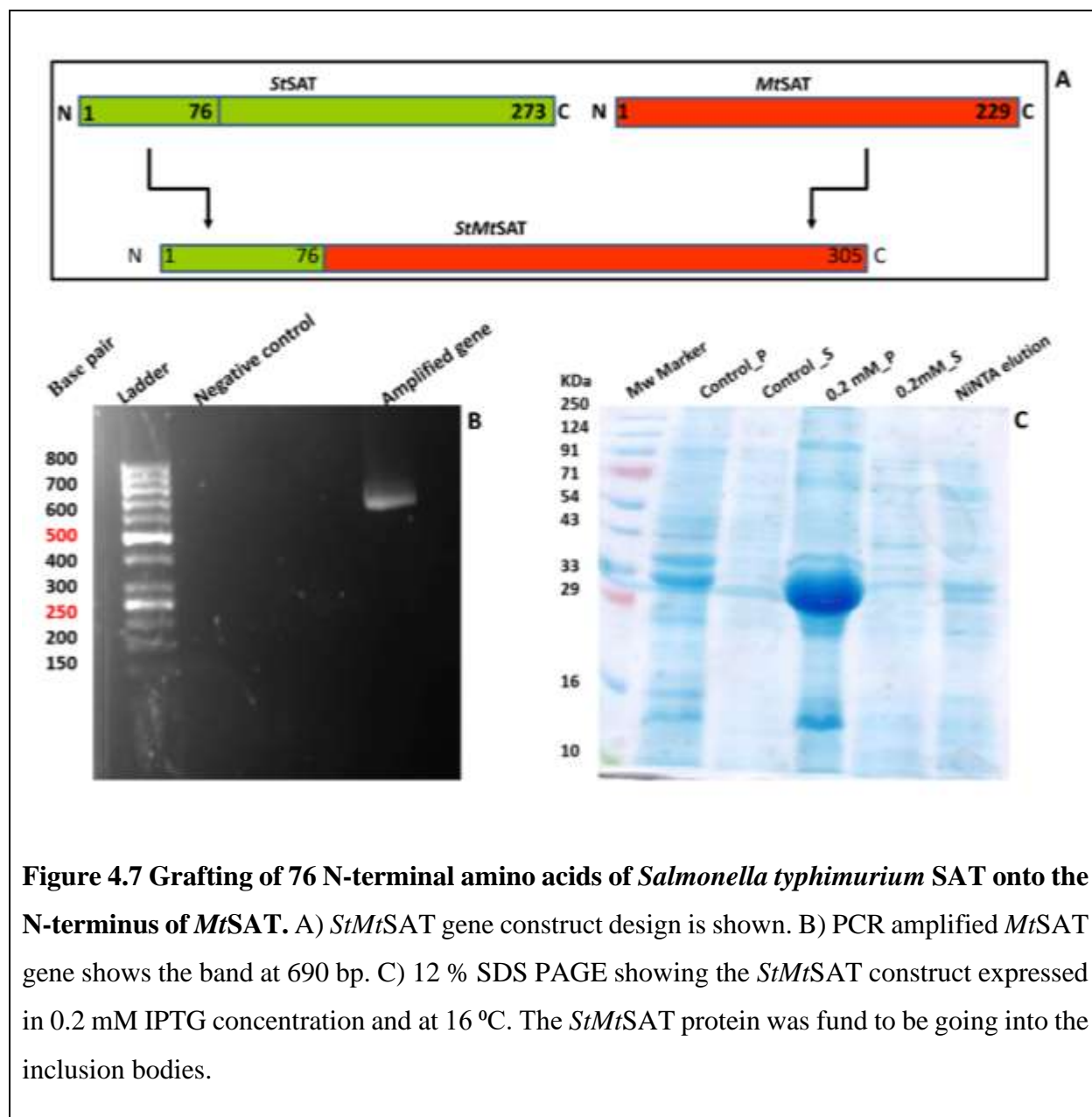
Table 4.2 Steady-state kinetic parameters for *St*SAT in free form as well as in different stoichiometric ratios with *St*CS.

Protein Type	Km (mM) (Apparent affinity)	k_{cat} (sec^{-1}) (Turn over)	k_{cat} / K_m (sec^{-1} mM^{-1}) (Catalytic Efficiency)
Free *St*SAT$_{hexamer}$	0.86 ± 0.06	158 ± 3.0	184 ± 13
*St*SAT$_{hexamer}$ to *St*CS$_{dimer}$ (1:2)	0.46 ± 0.02	193 ± 4.0	420 ± 20
*St*SAT$_{hexamer}$ to *St*CS$_{dimer}$ (1:2)	0.48 ± 0.02	212 ± 3.0	442 ± 19
*St*SAT$_{hexamer}$ to *St*CS$_{dimer}$ (1:2)	0.47 ± 0.02	208 ± 3.0	443 ± 20

Grafting of 76 amino acids of N-terminus of StSAT onto the N-terminus of MtSAT

The multiple sequence alignment of the SAT enzymes from different organisms shows that *Mt*SAT lacks a stretch of ~ 98 amino acids at the N-terminus with respect to other SATs and ~76 amino acids with respect to the *Salmonella typhimurium* SAT. The structural analyses of *Haemophilus influenzae* (**PDB ID: 1S80**) SAT, shows that the N-terminal region is involved in the hexameric assembly formation. The previous study of *Entamoeba histolytica* SAT (*Eh*SAT) shows that due to the differences in the N-terminal amino acid, it exists only in trimeric form. So, we hypothesized that the N-terminal region of *Mt*SAT might have a very important and crucial role in the regulation of SAT oligomerization and its activity. Results presented in this thesis validate our intuition that *Mt*SAT exists in multi-oligomeric states, very different than other SATs due to the N-terminal truncation. We extended our initial hypothesis to an idea that if we transfer the mission N-terminal residues from *St*SAT to *Mt*SAT, *Mt*SAT might start behaving like *St*SAT. Therefore, we designed a construct of SAT enzyme in which 76 N-terminal amino acids coding sequence from the *Salmonella typhimurium SAT* was joined onto the N-terminus of *Mycobacterium tuberculosis* SAT

(Figure 4.7 A). For constructing the *StMt*SAT gene, we amplified the *Mt*SAT full gene and N-terminus 216 bp gene segment of *St*SAT. The amplified gene of *Mt*SAT is shown in **Figure 4.7 B**. Both full-length *Mt*SAT gene and 216 bp N-terminus gene fragment of *St*SAT were subjected to restriction enzyme digestion and further purified by agarose gel.

Figure 4.7 Grafting of 76 N-terminal amino acids of *Salmonella typhimurium* SAT onto the N-terminus of *Mt*SAT. A) *StMt*SAT gene construct design is shown. B) PCR amplified *Mt*SAT gene shows the band at 690 bp. C) 12 % SDS PAGE showing the *StMt*SAT construct expressed in 0.2 mM IPTG concentration and at 16 ºC. The *StMt*SAT protein was fund to be going into the inclusion bodies.

In the first step, the two gene fragments were ligated and in the second step, this *StMt*SAT ligated gene construct was cloned into the pET28a+ vector between the restriction sites NdeI and HindIII.

The cloned gene was successfully transformed into different host cells. After a lot of standardizing the different expression conditions like temperate, IPTG concentration, and media, finally, we get the *StMt*SAT gene construct to be expressed in Arctic host cells. However, we could not get the *StMt*SAT gene construct in the soluble fraction. The 12 % SDS PAGE is shown for *StMt*SAT protein expressed at 0.2 mM IPTG and 16 ℃ that shows *StMt*SAT protein is expressing in the inclusion bodies (**Figure 4.7 C**).

Discussion

The existence of the Cysteine Regulatory complex (CRC) multienzyme assembly in *Salmonella typhimurium* was first evidenced by Kredich in 1966. The CRC complex was further found to exist in other bacterial and plant systems. The CRC complex was shown to dissociate into the SAT and CS enzyme in presence of OAS, the product of the SAT enzyme. In several other *in vitro* studies, it was shown that free SAT and free CS can form the CRC complex. However, the dynamics of association and dissociation of the CRC complex and whether the six C-terminal tail of SAT hexamer can bind to six CS dimers was not clear until recently. Recently, our lab has discovered that in *Salmonella,* CRC can exist as *St*CRC1 and *St*CRC2 states in equilibrium and this equilibrium can be shifted from *St*CRC2 to *St*CRC1 state by Na_2S ligand and from *St*CRC1 to free *St*SAT and free *St*CS by OAS. It is also shown that in the CRC complex, SAT has 2-3 times higher activity as compared to the free SAT and CS has ~70 % lower activity in CRC as compared to free CS. In *Mycobacterium tuberculosis,* CS is both structurally and biochemically characterized well but no such details are available for *Mt*SAT. In our SEC and AUC experiments, it's clear that in *Mtb*, CRC exists in a 1:4 stoichiometric ratio where one *Mt*SAT hexamer interacts with four *Mt*CS dimers. Surprisingly, the *Mt*CRC1 species could not form in the *Mtb* which might be due to missing ~98 amino acids at the N-terminal region. We further, characterized the interactions between the *Mt*SAT and *Mt*CS enzymes by thermodynamic experiments. From the thermodynamic experiments, we found that both enzymes interact with each other with high affinity in the micromolar range. The high-affinity interactions between the *Mt*SAT and *Mt*CS enzyme show a very stable *Mt*CRC complex exists in the *Mtb*. Further, the binding stoichiometry shows all four CS dimers can bind to the one SAT hexamer in a single-step mechanism. In other bacteria like *Salmonella typhimurium*, it was found that *St*CS dimers bind to the *St*SAT hexamer in a two-step process; in the first step, two CS dimers bind with the SAT hexamer to form the *St*CRC1 complex with very high affinity in the nanomolar range while the third and fourth *St*CS dimers bind to *St*SAT hexamer with comparatively low affinity to form the *St*CRC2 complex. In other bacterial species, SAT enzyme activity enhances 2-3 times in the CRC complex while the CS activity reduces to 70 % in the CRC complex. So, we analyzed the catalytic features of *Mt*SAT and *Mt*CS enzymes of *Mtb*. The *Mt*SAT enzyme shows ~ 26 folds lower turnover rate as compared to free *Mt*SAT. This is striking and shows somehow *Mt*SAT activity is minimally required by the *Mycobacterium* when it forms the *Mt*CRC complex. We have also calculated the catalytic

properties of the *Mt*CS enzyme in the *Mt*CRC complex. Again, to our surprise, we found no significant change in the *Mt*CS activity in the *Mt*CRC complex. This is also striking and completely opposite to the 70 % reduction in the activity of bacterial CS in the CRC. In general, all CS enzymes show lower activity when they form the CRC complex due to the less availability of the active site. However, in *Mtb*, the almost same activity of CS enzyme was observed which shows that OAS substrate somehow manages to reach the *Mt*CS's active site which needs further investigation to know how the OAS finds the *Mt*CS's active site in the *Mt*CRC complex. For direct comparison, we have also characterized the catalytic properties of *Salmonella* SAT enzyme in presence of *St*CS dimers. We find that as we increase the stoichiometry of *St*CS dimer to *St*SAT hexamer there is an increase in the catalytic turnover rate of *St*SAT enzyme. Here we also observed a 2 to 3 times more turnover rate as observed in other bacterial CRC systems. The higher catalytic activity *St*SAT upon complex formation may be due to changes in the catalytic site or stability of the enzyme in the complex. Together, our studies demonstrate that the first ~98 amino acids play a crucial role in *Mt*SAT enzyme activity as well as their ability to interact with *Mt*CS to form the *Mt*CRC complex. Further, to know the role of these N-terminal amino acids, we made a protein construct, where we grafted the first ~76 N-terminal amino acids of *Salmonella* SAT to the N-terminus of *Mt*SAT protein and tried to purify. However, we were not able to purify this construct in the soluble fraction. Results described here provide the first line of evidence that regulatory features underlying sulfate and cysteine metabolism in *Mtb* likely to be different as compared to other bacterial enzymes.

Chapter 5

Understanding key regulatory features of CRC by solution and structural approaches

Research Problem

The Multienzyme complexes, such as the CRC complex, substantiate the notion of "primary blocks" of metabolic systems (Kurganov, 1986). The existence of multienzyme metabolic systems forms the material basis for metabolic regulation control. Molecular features of these assemblies allowed us to understand the hierarchical integration of cell metabolism with other processes. Further, only a few multienzyme complexes have been characterized in detail due to difficulties associated with purifying, assembling, quantitatively measuring properties of these complexes (Fu et al., 2016). The dynamics of assembly and dissociation of Cysteine Regulatory Complex (CRC) in response to the environmental milieu and how CRC controls fluxes are central to understand cysteine mediated redox biology of the pathogenic bacteria. There are plenty of reports in which the effect of different substrates, intermediates, and products directly or indirectly controls multienzyme complex association and dissociation. OAS, the product of the SAT and substrate of the CS, plays a key role in regulating the assembly and disassembly of CRC (Kaushik et al., 2017). A high concentration of OAS dissociates CRC complex into free SAT and free CS enzymes in a stepwise manner. First, it dissociates the high molecular weight 1:4 CRC2 assembly into low molecular weight 1:2 CRC1 assembly, then further increase in OAS concentration leads CRC1 assembly to dissociate into free SAT and CS enzymes.

SAT continues to catalyze the formation of OAS even under sulfur scarcity conditions, leading to accumulation of OAS as S^{2-} is not available for conversion of OAS into cysteine by CS. The first effect of high OAS concentration is that it competes for the C-terminus of SAT to the active site of CS, leading to dissociation of SAT from the CRC complex and decreasing the SAT activity. The decrease in SAT activity reduces OAS concentration inside the cell. The second effect of a high OAS concentration is that OAS non-enzymatically converts into the N-acetyl serine (NAS). The NAS dissociates cysB, transcription repressor and regulator of sulfate metabolism genes, leading to the upregulation of sulfur transporting enzymes and restoring the sulfate levels.

Another most important self-regulation of sulfate and cysteine metabolism is the feedback inhibition of SAT by the end product, cysteine. However, no detailed study is available to understand how cysteine regulates the SAT activity. A structural study showed that cysteine binds to the C-terminal part, close to the active site. It was assumed that cysteine binding induces conformational changes to the active site, preventing substrate binding. We report high resolution structural and biochemical studies that provide insights into cysteine mediated self-regulation.

First, we examined the effect of ligands on the stability dynamics of CRC complexes from *Salmonella* and then resolve the crystal structure of the feedback-inhibited complex.

Results

Ligand mediated StCRC1 and StCRC2 complex association and dissociation

A previous study from our lab showed that both OAS and Na_2S, substrates of CS, dynamically dissociate the *St*CRC complex in a concentration-dependent manner. In this work, we examined the sensitivity of each complex to ligands. We analyzed the SEC purified *St*CRC1 fraction by velocity experiments. Velocity profiles show that along with a major fraction of *St*CRC1, a small fraction of *St*CRC2 species is also observed. This suggested that SEC purified *St*CRC1 fraction exists in the equilibrium with *St*CRC2. Similarly, we analyzed the AUC velocity profiles of the *St*CRC2 fraction. As expected, along with a major population of *St*CRC2, a small fraction of *St*CRC1 was also observed, consistent with principles of laws of mass action. To check the ligand sensitivity of the *St*CRC1 complex, we analyzed the velocity profiles in the presence of Na_2S.
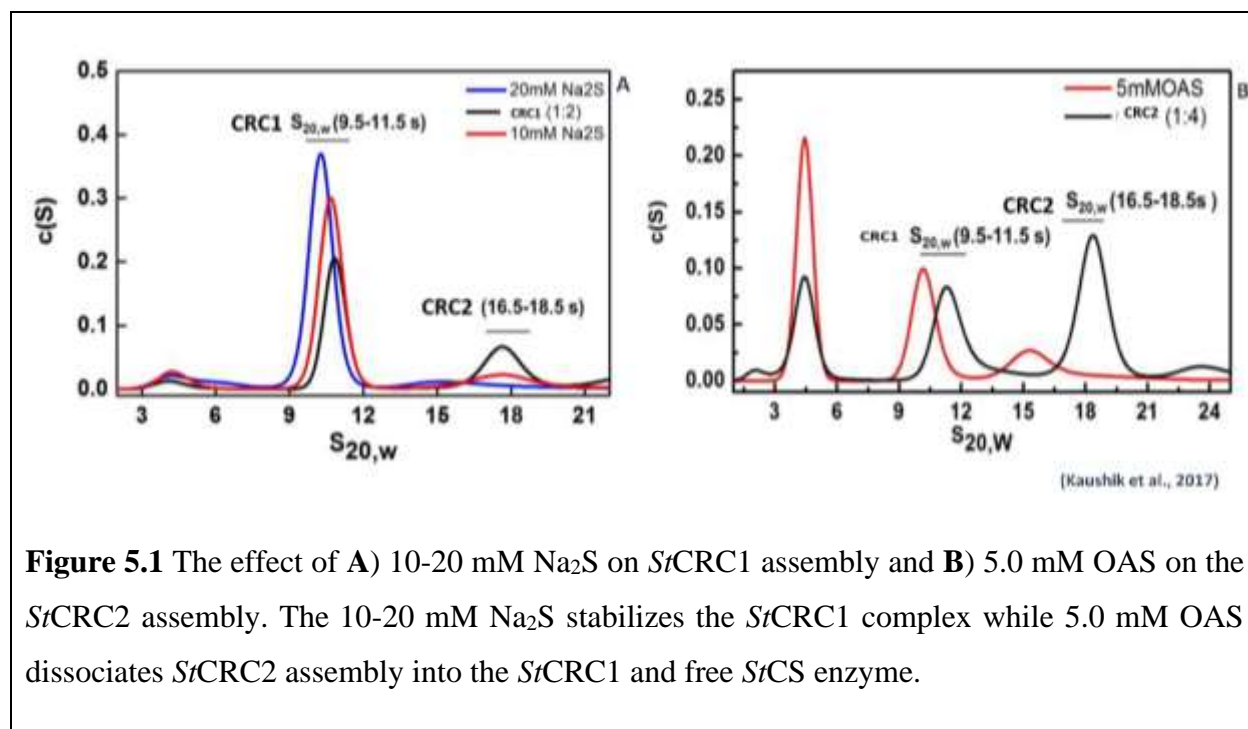
Figure 5.1 The effect of **A**) 10-20 mM Na_2S on *St*CRC1 assembly and **B**) 5.0 mM OAS on the *St*CRC2 assembly. The 10-20 mM Na_2S stabilizes the *St*CRC1 complex while 5.0 mM OAS dissociates *St*CRC2 assembly into the *St*CRC1 and free *St*CS enzyme.

Results show that Na_2S in the concentration range 10-20 mM do not affect the equilibrium. We then examined the stability of *St*CRC2 in the presence of 10-20 mM Na_2S and results show that the net oligomeric state is pushed towards *St*CRC1 and this state remains in the predominant

oligomeric form of *St*CRC1 (**Figure 5.1 A**). On the contrary, 5.0 mM OAS reduces the *St*CRC2 assembly to *St*CRC 1 assembly and free *St*CS enzymes (**Figure 5.1 B**).

Effect of cysteine on oligomerization and regulation of the StCRC1 and StCRC2 complex association and dissociation

Cysteine is the end product of the sulfur metabolism and cysteine synthesis pathway. The catalytic efficiencies of SAT from a different group of bacteria have been found to be reduced in the presence of cysteine (Hindson, 2003; Kumar et al., 2011a; Olsen et al., 2004). The effect of cysteine is opposite to the effect of CS. The stability of SAT increases when CS binds to form the CRC complex and the catalytic efficiency of SAT within the CRC is higher than that of free SAT (Yi et al., 2013b). However, in the absence of detailed studies, the mechanistic and regulatory features of feedback inhibition of SAT by cysteine is not clear. Further, it was not known until we reported about the two types of CRC1 and CRC2 complexes. Therefore, the role of cysteine on assembly dynamics and feedback regulation on both complexes need to be studied. As *St*CRC in size exclusion chromatography elutes in two oligomer forms one at ~63.0 mL of *St*CRC1 and another at ~57.0 mL of *St*CRC2 of molecular weight ~320 kDa and ~460 kDa respectively, we first purified the *St*CRC complex in the presence of 5.0 mM cysteine to know the effect of cysteine on the *St*CRC assembly. The *St*SAT and *St*CS enzymes were mixed in an appropriate stoichiometric ratio and kept for 2 hrs at 4 ºC. The column was equilibrated with a buffer of 50 mM Tris pH 8.0, 100 mM NaCl, 10 % Glycerol containing 5.0 mM cysteine. Then the *St*CRC complex was loaded onto the pre-equilibrated column and purified *St*CRC complex in the presence of cysteine. Surprisingly, the second oligomer *St*CRC2, which was observed in the absence of cysteine in SEC, disappeared in the presence of cysteine. This suggests that cysteine favors the equilibrium to the left by stabilizing *St*CRC1 while promoting the dissociation of *St*CRC2. We determined the molecular weight of this *St*CRC1 complex in the presence of cysteine and from the standards run in the column, we found it to be ~ 320 kDa (**Figure 5.2 A**). To check the effect of cysteine on different *St*CRC complexes, we systematically increased the *St*CS dimer stoichiometry to *St*SAT enzyme. In the first case, the stoichiometry of *St*SAT hexamer to *St*CS dimer was kept 1:2, and purified the formed *St*CRC complex by SEC. To form the *St*CRC2 and *St*CRC3 complex, the stoichiometry of *St*SAT hexamer to *St*CS dimer was kept 1:4 and 1:6 respectively. Only the *St*CRC1 complex was formed in all three cases, and the rest of the *St*CS is eluted as free *St*CS dimer at ~ 15.0 mL (**Figure 5.2 B**). So, from the SEC experiment, it is confirmed that cysteine

stabilizes the *St*CRC1 oligomeric form over the *St*CRC2 complex from the size exclusion chromatography.
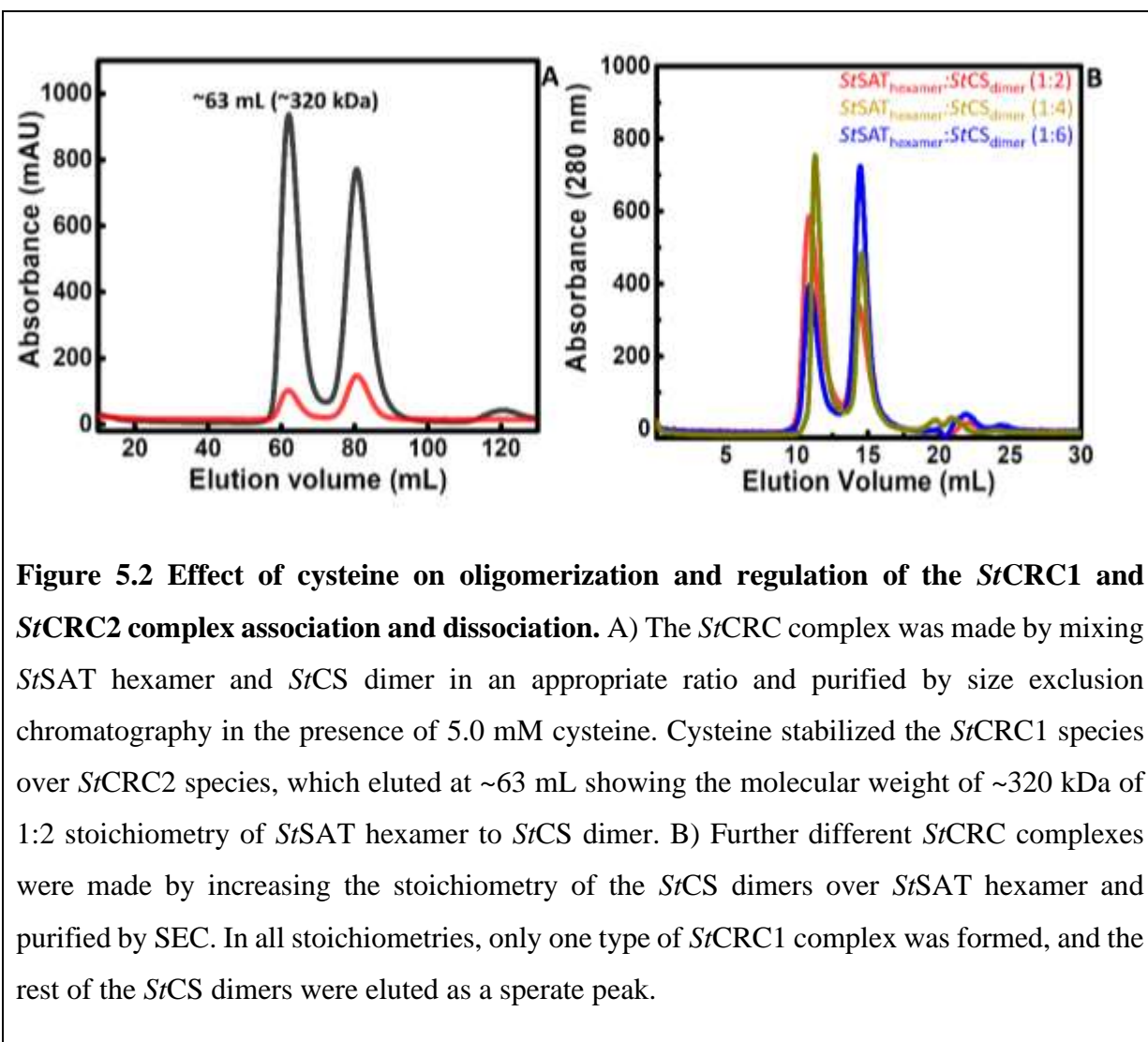
Figure 5.2 Effect of cysteine on oligomerization and regulation of the *St*CRC1 and *St*CRC2 complex association and dissociation. A) The *St*CRC complex was made by mixing *St*SAT hexamer and *St*CS dimer in an appropriate ratio and purified by size exclusion chromatography in the presence of 5.0 mM cysteine. Cysteine stabilized the *St*CRC1 species over *St*CRC2 species, which eluted at ~63 mL showing the molecular weight of ~320 kDa of 1:2 stoichiometry of *St*SAT hexamer to *St*CS dimer. B) Further different *St*CRC complexes were made by increasing the stoichiometry of the *St*CS dimers over *St*SAT hexamer and purified by SEC. In all stoichiometries, only one type of *St*CRC1 complex was formed, and the rest of the *St*CS dimers were eluted as a sperate peak.

AUC velocity characterization of *St*CRC complex in the presence of cysteine

For an extensive analysis of the *St*CRC complex, we performed the AUC velocity experiments of eluted fractions in the absence and presence of cysteine. The sedfit software was used to fit the AUC velocity data, and the concentration distribution of sedimentation coefficients c(S) was plotted against the normalized sedimentation coefficient ($S_{20,w}$) value. The AUC experiment results show that in the absence of cysteine, the *St*CRC complex exists in two different states, *St*CRC1 and *St*CRC2, and some higher molecular weight soluble oligomers (**Figure 5.3 A**). AUC analyses show one predominant peak and two less predominant peaks. The first peak of *St*CRC1

has an $S_{20,w}$ value of 9.5 ± 0.2 s that corresponds to the molecular weight of 314 ± 5 kDa (71 %). The second less predominant peak of *St*CRC2 has an $S_{20,w}$ value of 12.5 ± 0.2 s and corresponds to the molecular weight of 420 ± 6 kDa (22 %). The third high molecular weight soluble oligomer minor peak is also observed with a significantly less population (7%). However, when the same experiments were performed in the presence of cysteine, results show that both *St*CRC2 and higher
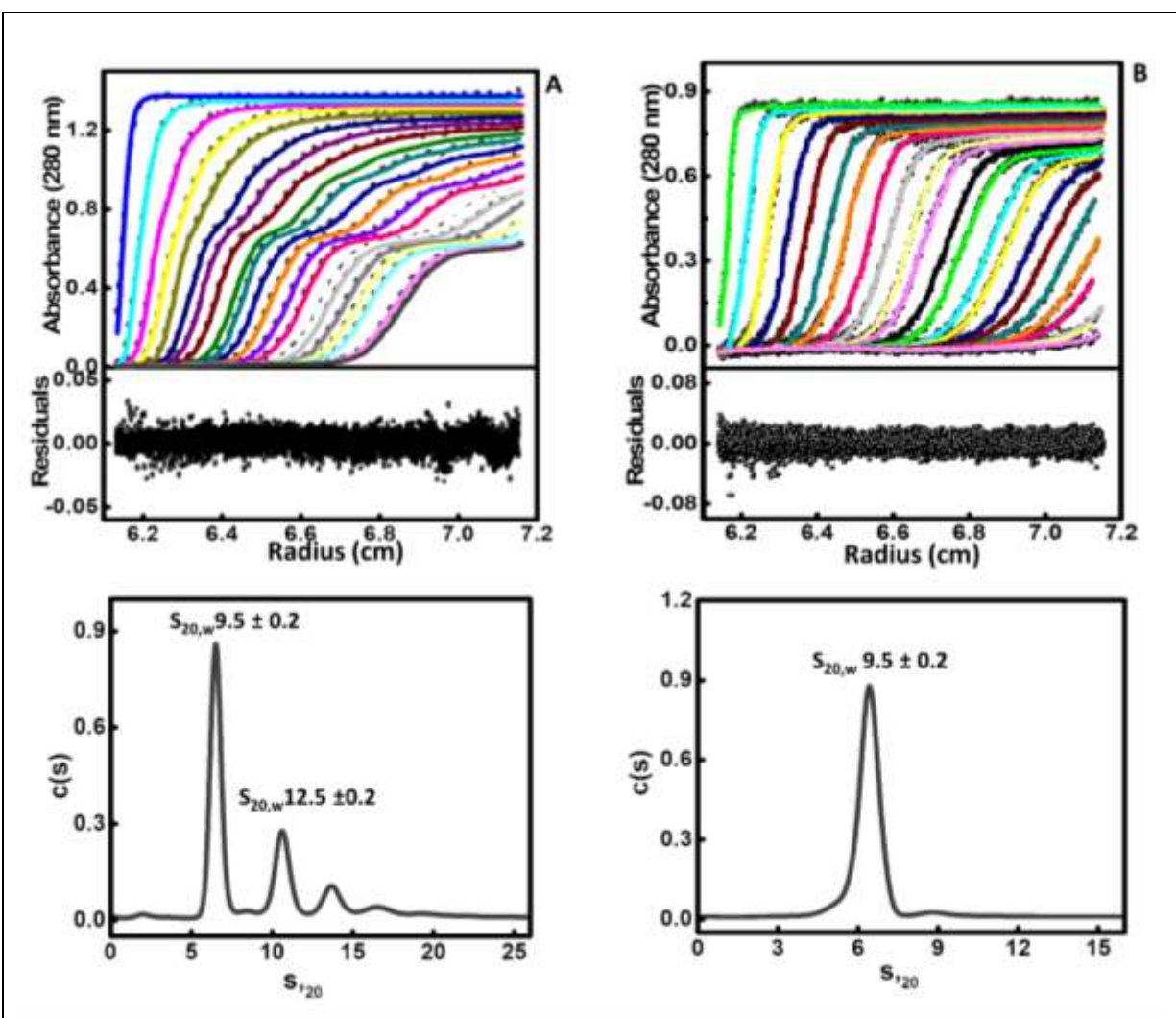
Figure 5.3 Analytical ultracentrifugation characterization of *St*CRC complex in the absence and presence of cysteine. Sedimentation velocity data monitored at 280 nm were fit to a continuous distribution of noninteracting species model using Sedfit and sedimentation coefficient distribution c(s) as a function of the normalized sedimentation coefficient ($S_{20,w}$) was plotted. **A)** Velocity data of *St*CRC in the absence of cysteine shows one prominent peak

of *St*CRC1 ($S_{20,w}$ ~9.5 ± 0.2 s, Mw ~314 ± 5 kDa, 71 %), second less predominant peak of *St*CRC2 ($S_{20,w}$ ~12.5 ± 0.2 s, Mw ~420 ± 6 kDa, 22 %), and third high molecular weight soluble fraction peak (7%). **B)** Velocity data of *St*CRC in the presence of 5.0 mM cysteine shows only one prominent peak of *St*CRC1 ($S_{20,w}$ ~9.5 ± 0.2 s, Mw ~314 ± 5 kDa); the second *St*CRC2 and third higher oligomeric peaks were disappeared.

soluble oligomeric peaks disappeared leaving only *St*CRC1 oligomeric species with $S_{20,w}$ value of 9.5 ± 0.2 s (**Figure 5.3 B**). From the AUC data analysis, it is confirmed that cysteine directly stabilizes the *St*CRC1 complex and prevents further higher oligomerization.

Catalytic properties of Salmonella typhimurium SAT (StSAT) in presence of cysteine

Cysteine is a known feedback inhibitor of SAT enzyme in a different group of bacteria (Kumar et al., 2011a). It inhibits the SAT activity by competitive inhibition where it binds to the Acetyl CoA binding site. When the concentration of cysteine increases more than the required level inside the cell it reduces the SAT activity by 2-3 folds. However, in *Salmonella typhimurium*, there are no reports available on whether cysteine can inhibit the *St*SAT activity or not. So, we performed the activity experiments to determine the catalytic properties of *St*SAT in presence of cysteine. We determined initial velocities of the hexameric fraction at varying substrate concentrations and then analyzed the initial velocities as a function of substrate concentrations to obtain kinetic parameters, turnover rate (k_{cat}), and apparent substrate affinity (K_M). Plots of Initial velocity determined as a function of L-serine concentration and the fitted data are shown (**Figure 5.4**). The L-serine concentration was varied while fixing Acetyl CoA concentration to 0.2 mM. The turnover rate (k_{cat}), apparent substrate affinity (K_m), and enzyme efficiency (k_{cat} / K_m) of *St*SAT were found to be ~54 s^{-1} and 0.52 mM ± 0.05 mM, and ~104 mM^{-1} s^{-1} respectively. In the previous chapter, we determined the catalytic features of the free *St*SAT enzyme where we found the k_{cat} and enzyme efficiency value of 158 s^{-1} and 183 mM^{-1} s^{-1} respectively. When we compared the turn over rate and enzyme efficiency of free *St*SAT and the *St*SAT in 1.0 μM of cysteine, we found that the *St*SAT loses the turnover rate by about 3 folds while the enzyme efficiency was reduced to half. In bacteria like *Entamoeba histolytica,* the mode of SAT enzyme inhibition is suggested to work through competitive inhibition. However, it is not clear whether the cysteine inhibits *St*SAT activity through competitive inhibition or allosteric inhibition as no crystal structure is available

for *St*SAT structure in free form or cysteine bound. In the next section, we describe the results of a new crystal structure of *St*SAT in the cysteine bound form and discuss how cysteine inhibits the catalytic activity of the *St*SAT enzyme.
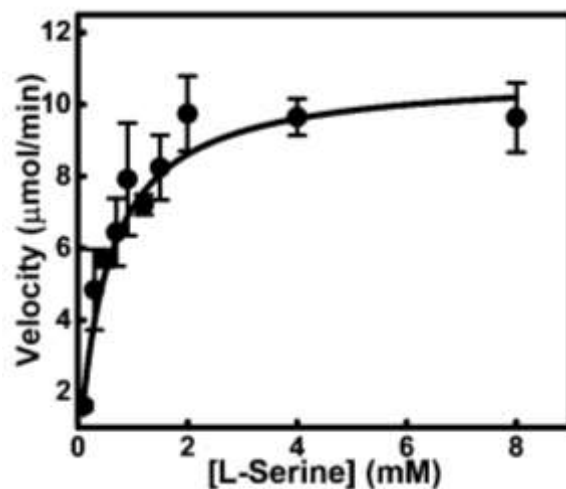
Figure 5.4 *St*SAT enzyme activity inhibition by cysteine feedback inhibitor. The *St*SAT activity experiments were performed in triplicates under identical solution conditions at fixed Acetyl CoA concentration (0.2 mM) and varying L-serine concentrations in presence of 1.0 µM of L-Cysteine. Initial velocities determined were plotted against L-Serine concentration. The initial velocity data, shown as mean values determined with standard error estimated at 95% confidence interval were fit to Michaelis-Menten equation (eq. 1, Materials and methods) and the parameters (K_m, k_{cat}, and k_{cat}/K_m) are determined.

Crystallization and structure determination of Salmonella typhimurium SAT

The crystal structure of *St*SAT is not available. Initial attempts to crystalize the *St*SAT protein was not successful as concentrating the protein beyond the 6 to 8 mg/mL promoted aggregation. We next considered crystallizing *St*SAT in the presence of cysteine as it is one of the ligands of the *St*SAT and also cysteine shows aggregation preventing properties as observed from the results of *St*CRC purification (**Figure 5.2 & 5.3**). In the presence of cysteine, we found that *St*CRC1 can be concentrated up to 30 mg/mL. So, we tried purifying our *St*SAT enzyme in the presence of cysteine in the hope to get more concentrated *St*SAT protein. In the presence of 1.0 mM of cysteine, the

*St*SAT enzyme gets concentrated up to 30 mg/mL. Then we started crystallizing the *St*SAT enzyme in the presence of cysteine. We got diffraction quality crystals within one month and diffracted the same. We collected the X-ray diffraction data at the home source, integrated the images, scaled, and obtained the structure solution by using the coordinates of *Escherichia coli* SAT (*Ec*SAT; **PDB ID:1T3D**) to find electron density phases. The data collection and refinement statistics are provided in **Table 5.1**.

The crystal structure of *St*SAT in the presence of cysteine shows that protein is crystallized in the H3 space group, with two monomers within one asymmetric unit. Crystal structures of SAT from other bacteria showed that the C-terminal tail of SAT is disordered due to the absence of electron density for the C-terminal amino acids. In *Escherichia coli*, the electron density of the last 11 amino acids is missing (Pye et al., 2004b). Similarly, in the structure of SAT from *Entamoeba histolytica, Glycine max, Haemophilus influenzae*, electron densities for the last C-terminal 37, 31, and 27 amino acids respectively were missing (Gorman & Shapiro, 2004b; Kumar et al., 2011a; Yi et al., 2013a). In *St*SAT resolved here in our study, we observed that the A chain does not have electron densities for the last 29 amino acids while the B chain does not have the electron density of the last 121 amino acids (**Figure 5.5**). This is the first structure of the SAT, to the best of our knowledge that monomers from opposite trimers display very different structural properties. In the SAT structure of *Salmonella*, the A chain consists of two domains; an N-terminal domain from 2-140 amino acids consisting of 8 α helices, while the C-terminal domain (141 to 244 amino acids) consisting of 6 β stands. In the B chain, the N-terminal domain consists of 4 to 140 amino acids with 8 α helices while the C-terminus domain has an electron density of only 12 amino acids (141 to 152).
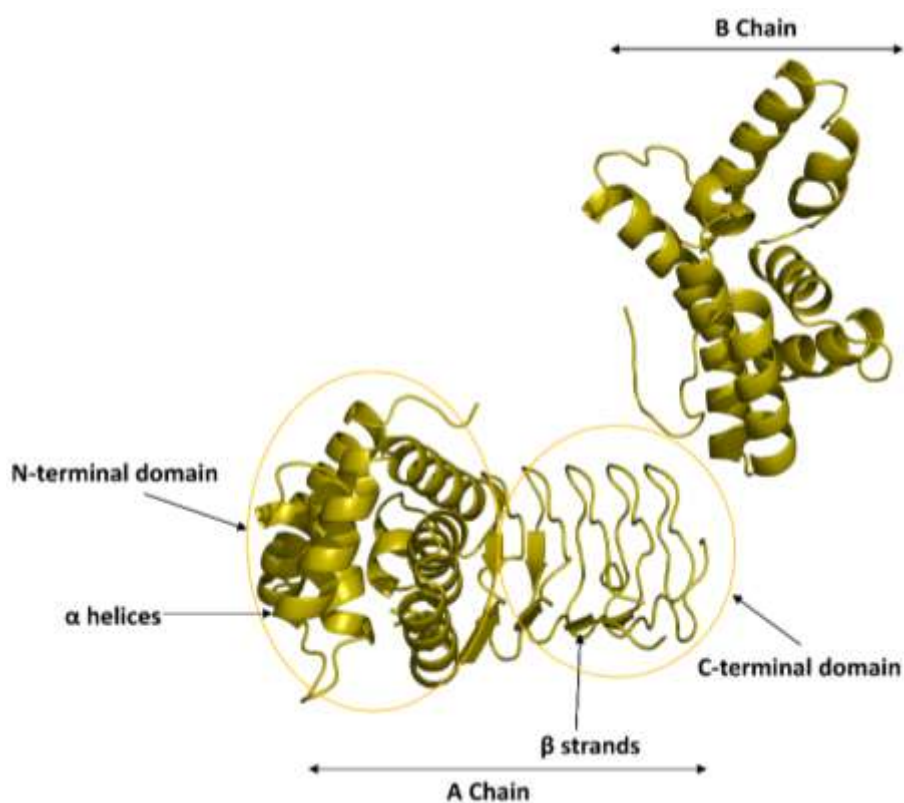
Figure 5.5 The crystal structure of *St*SAT in presence of cysteine showing two monomers (A chain and B chain) in the asymmetric unit. In the *St*SAT structure, A chain consists of 2-244 amino acids and both N-terminal domain (from 2 to 140 amino acids) consisting 8 α helices and C-terminus domain (from 141 to 244 amino acids) consisting 6 β strands were observed. B chain consists of 4 to 152 amino acids with only the N-terminal domain (from 4 to 140 amino acids) consisting of 8 α helices, was observed while only C-terminal domain (from 141 to 152 amino acids) consisting of 12 amino acids only.

However, in the C-terminal domain of chain A, the two β-strands, β-1 and β-3 were missing as compared to the other SAT structures. From the studies on the SATs from other bacteria it is known that the cysteine inhibits SAT activity and, in our case, it makes the *St*SAT enzyme more soluble while inhibiting the activity by three-fold. In other bacterial SATs, the electron density of free form, as well as the cysteine bound form of SAT, shows that only 20 to 30 amino acids are unstructured. But in the structure of *St*SAT, most of the C-terminal domain of the B chain is missing suggesting the role of 1.0 mM cysteine concentration on one of the chains. The *St*SAT

hexamer is highly asymmetric with respect to its two head-to-head locked trimers where monomers of one trimer are well structured like that of in the SAT from other bacteria but most of the C-terminal of monomers of another trimer is unstructured. This very striking as almost the entire C-terminal domain of the B chain is disordered. We can infer that it might be possible the absence of cysteine ligand, both the trimers of *St*SAT have unstructured C-terminal domains which cause the *St*SAT enzyme more prone to aggregation and hence, less solubility. If we increase the cysteine concentration > 1.0 mM then it might be possible that the *St*SAT becomes more soluble which would help the C-terminal domain of monomers of the other trimer to get into folded conformation.

Structural studies have revealed that ligand-binding can induce the folding of unstructured domains or regions into structured domains or motifs (Majewski et al., 2019). In other bacterial and plant SATs, the canonical hexameric architecture consists of two equal trimers joined with head-to-head polarity while protruding their C-terminal domains outwards to form interactions with the CS enzymes. So, we generated the hexameric symmetry by applying symmetry operations around the asymmetric unit (**Figure 5.6 A**). Chain A translates into one trimer and chain B translates into another trimer and two trimers are joined through their respective N-terminal helices
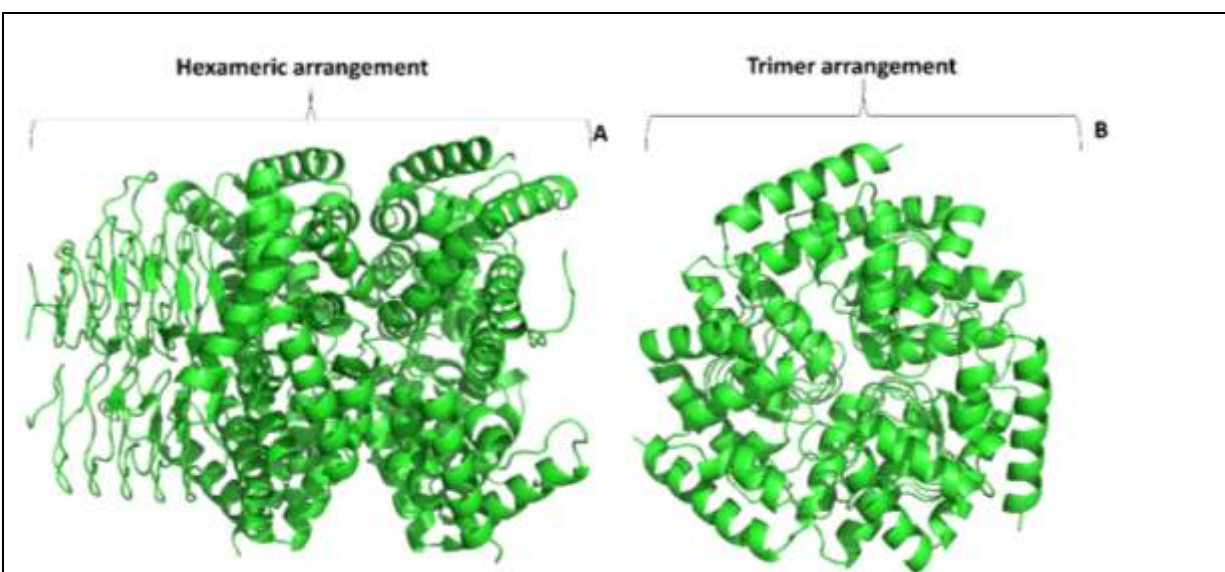
Figure 5.6 A) Hexameric arrangement of the *St*SAT enzyme shows that one trimer has a well-folded C-terminal domain while the other trimer has an unstructured C-terminal domain. B) The trimeric N-terminal domain view of the trimer shows that the three monomeric subunits have 3-fold axis symmetry.

and belong to different hexamers. The model structure of *St*SAT in hexameric and trimeric form shows that one trimer becomes ordered and another remains unordered (**Figure 5.6 A**). The front view of the trimer shows that three subunits are arranged in the three-fold rotating symmetry axis (**Figure 5.6 B**).

In the crystal structure of *St*SAT, we observed the presence of electron density for cysteine ligand at the N-terminal domain near the end of the 8th α helix of one trimer. The Simulated annealing **2Fo-Fc** omit maps for L-cysteine ligand contoured at 1.0 σ clearly shows the presence of cysteine bound to SAT (**Figure 5.7 A**). The cysteine ligand in the structure interacts with Glu-135, Ser-55, and Phe-131 residues of A chain with a distance of 2.71 Å, 3.72 Å, and 3.63 Å respectively. The alignment of the A chain with the B chain of *St*SAT structure reveals that both chains have an almost similar fold in the N-terminal domain. It was observed that the first β strand which is present in the A chain was absent in the B chain where the loop conformation was observed instead of the β strand. It might be possible that because the 8th α helix is in between the N-terminal domain and C-terminal domain and binding of cysteine might be stabilizing the first β

Figure 5.7 A) The Simulated annealing **2Fo-Fc** omit maps of L-cysteine ligand contoured at 1.0 σ. The cysteine ligand interacts with Glutamic acid 135, Serine 55, and Phenylalanine 131 amino acids near the 8th α helix of the A chain with a distance of 2.71 Å, 3.72 Å, and 3.63 Å respectively. B) The alignment of A chain and B chain of *St*SAT structure shows B chain does not have 1st β strand at the C-terminal domain as compared to A chain. The cysteine after binding
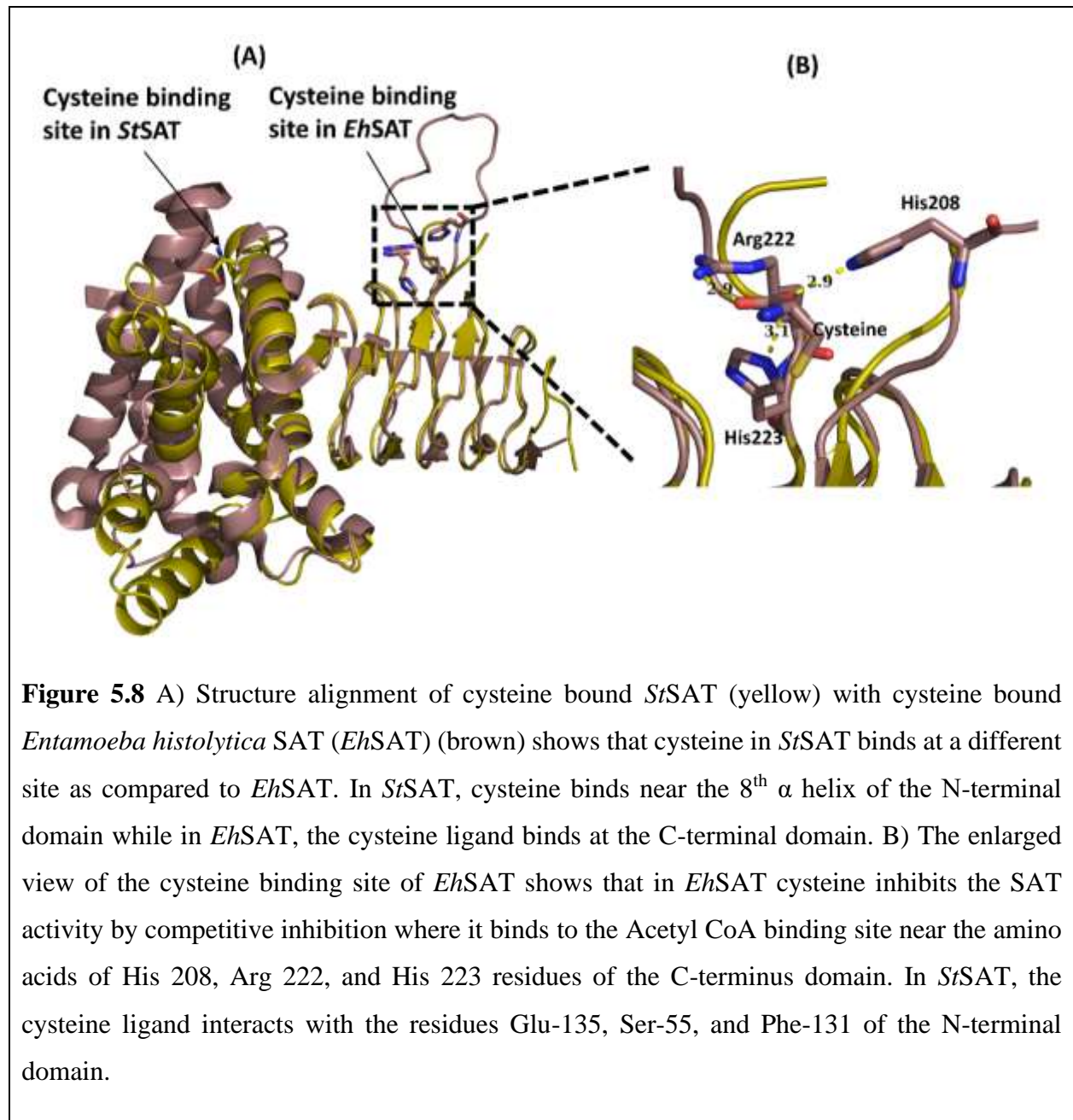
to the 8th α helix, of the A chain it stabilizes the first 12 amino acids of the C-terminal domain into the 1st β strand which further stabilizes the remaining 5 β strands of the C-terminal domain into the well-folded C-terminal domain. Also, we observed electron density of the first two residues at the beginning of the N-terminal domain of A chain which was absent in B chain.

strand of the C-terminal domain of A chain that further stabilizes the other 5 β strands of the C-terminal domain in the ordered conformation (**Figure 5.7 B**). We have also observed electron density of the first two residues at the beginning of the N-terminal domain of A chain which was absent in B chain. Overall, the binding of cysteine to the A chain of *St*SAT enzyme helps in the folding of the C-terminus.

Further, we found that in other bacterial SATs like *Entamoeba histolytica* the cysteine inhibits SAT activity by competitive inhibition mechanism (Kumar et al., 2011a). So, we aligned the cysteine bound *St*SAT and *Eh*SAT structure and found that in *Eh*SAT, cysteine binds to the Acetyl Co-A binding site at the C-terminal domain whereas in the case of *St*SAT cysteine bounds at the 8th α-helix of N-terminal domain (**Figure 5.8 A**). In *Eh*SAT, the cysteine inhibits SAT activity by binding at the competitive site with the C-terminal amino acids His 208, Arg 222, and His223 with distances of 2.9 Å, 2.9 Å, and 3.1 Å respectively (**Figure 5.8 B**). The binding site of cysteine in the structure of *St*SAT is far away from the active site and very different from the cysteine binding site observed in the structure of *Eh*SAT. These results provide clear evidence that feedback inhibition by cysteine in *St*SAT is allosteric, not competitive. Therefore, we propose an allosteric mode of feedback inhibition mechanism in *Salmonella typhimurium* SAT. We hypothesized that the binding of cysteine to this allosteric site might be inducing some conformational changes at the C-terminal domain which further alters the active site of *St*SAT and therefore allosterically modulate the Acetyl Co-A binding site reducing the *St*SAT activity.

Figure 5.8 A) Structure alignment of cysteine bound *St*SAT (yellow) with cysteine bound *Entamoeba histolytica* SAT (*Eh*SAT) (brown) shows that cysteine in *St*SAT binds at a different site as compared to *Eh*SAT. In *St*SAT, cysteine binds near the 8th α helix of the N-terminal domain while in *Eh*SAT, the cysteine ligand binds at the C-terminal domain. B) The enlarged view of the cysteine binding site of *Eh*SAT shows that in *Eh*SAT cysteine inhibits the SAT activity by competitive inhibition where it binds to the Acetyl CoA binding site near the amino acids of His 208, Arg 222, and His 223 residues of the C-terminus domain. In *St*SAT, the cysteine ligand interacts with the residues Glu-135, Ser-55, and Phe-131 of the N-terminal domain.

Table 5.1 Crystallographic Data Collection and Refinement Statistics of *St*SAT

Wavelength	1.541 Å	R-free	0.2711 (0.2933)
Resolution range	33.98 Å - 2.4 Å (2.486 - 2.4) Å	Number of non-hydrogen Atoms	3276
Space group	H 3	macromolecules	2995
Unit cell (A, B, C) Å	79.46 79.46 216.09	Protein residues	390
Total reflections	296670 (29178)	RMS (bonds)	0.018
Unique reflections	19909 (1986)	RMS (angles)	2.29
Completeness (%)	100 (98.46)	Ramachandran favoured (%)	96.61
Avg [I/Sigma(I)]	22.44 (3.91)	Ramachandran allowed (%)	2.6
Wilson B-factor	21.61	Ramachandran outliers (%)	0.78
R_{merge}	0.308 (1.066)	Rotamer outliers (%)	0
R_{meas}	0.329 (1.167)	Clash score	6.44
$CC_{1/2}$	0.94	Average B-factor	30.69
Reflections used in refinement	19608 (1990)	macromolecules	29.63
Reflections used for R-free	918 (81)	solvent	45.247
R-work	0.2413 (0.2971)	Ligands	30

SAXS analysis for determining the oligomeric states of StCRC1 and StCRC2 complexes

All our attempts to crystallize *St*CRC has not resulted in any diffractive quality crystals due to the inherent properties of *St*CRC1 and *St*CRC2 complexes to be present in the equilibria with each other. Therefore, we employed the Small-angle X-ray scattering to understand the low-resolution structural parameters like oligomeric states, approximate molecular diameter, and molecular envelope and to cross verify the results from AUC. Like the sedimentation velocity approach,

SAXS allows us to estimate the molecular weight and assembly state from the scattering properties of molecules in the solution. Four data sets of SAXS were collected, two for *St*CRC1 (1.0 & 2.0 mg/mL) and two for *St*CRC2 complex (1.0 & 2.0 mg/mL). The radius of gyration (Rg), extracted from Guinier plot analysis and normalized pair-distribution function [p(r)] for *St*CRC1 and *St*CRC2 are shown in (**Figure 5.9 A & B**). Analysis of the Guinier plot (inset) shows that the protein-protein complex is free from aggregation and inter-particle interference effects. The porod volume estimated from SAXS experiment for the *St*CRC1 (1.0 & 2.0 mg/mL) and *St*CRC2 (1.0 & 2.0 mg/mL) are ~33 nm^3 and ~146 nm^3 respectively. The volume estimated for *St*CRC1 and *St*CRC2 translates to the molecular weight of ~ 320 kDa and ~460 kDa respectively which is
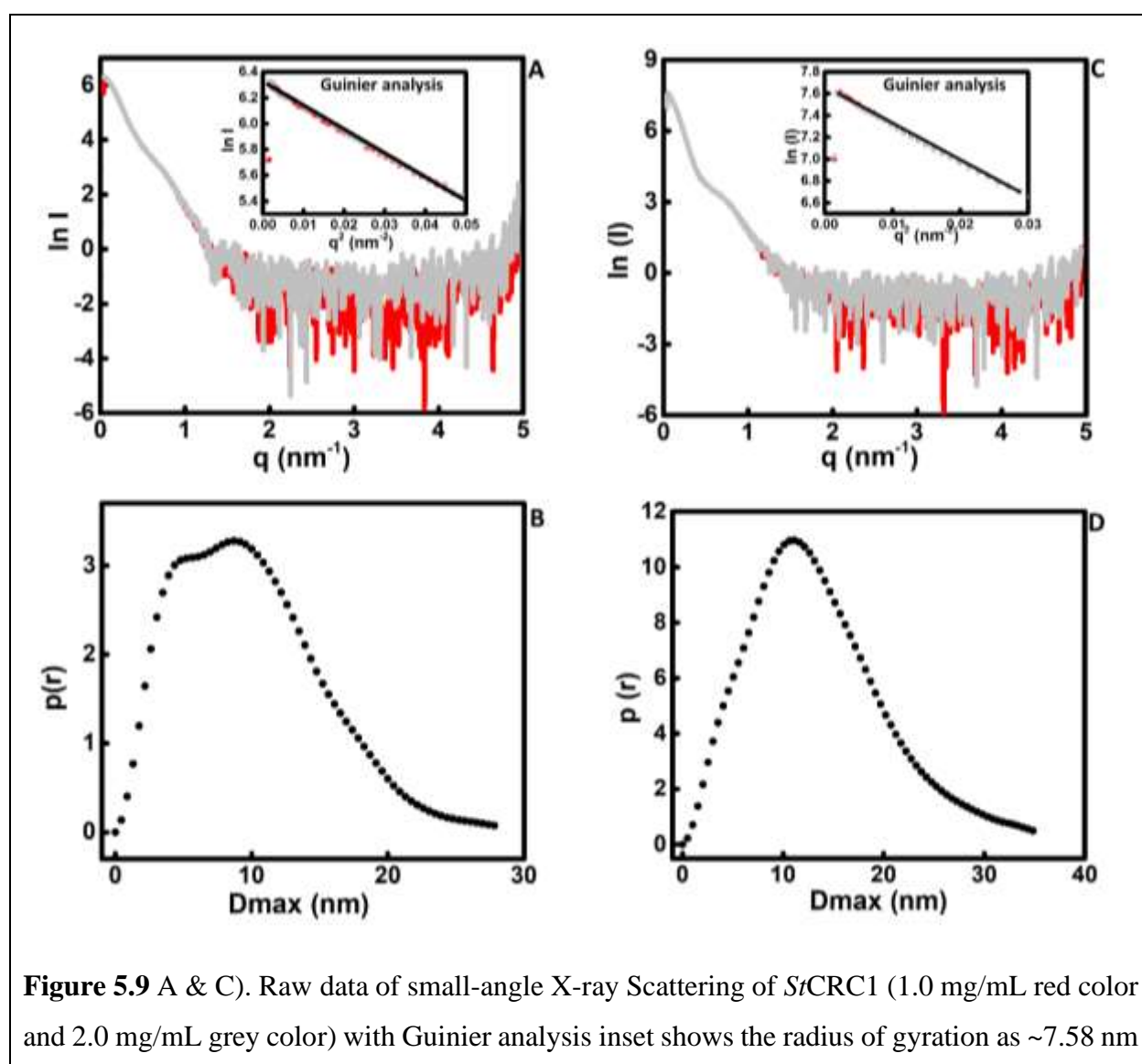
Figure 5.9 A & C). Raw data of small-angle X-ray Scattering of *St*CRC1 (1.0 mg/mL red color and 2.0 mg/mL grey color) with Guinier analysis inset shows the radius of gyration as ~7.58 nm

that matches with the Rg calculated by p(r) distribution (7.65 nm). The p(r) distribution analysis shows the approximate diameter of *St*CRC1 as ~26.2 nm. B & D). Raw small-angle X-ray Scattering data of *St*CRC2 (1.0 mg/mL red color and 2.0 mg/mL grey color) with Guinier analysis inset shows the radius of gyration as 10.4 nm that matches with the Rg calculated by p(r) distribution (~10.7 nm). The p(r) distribution analysis shows the approximate diameter of StCRC1 as ~34 nm.

similar to the molecular weight determined by sedimentation velocity experiments. The Rg obtained for *St*CRC1 (1.0 & 2.0 mg/mL) and *St*CRC2 (1.0 & 2.0 mg/mL) from the Guinier analysis (~7.58 nm and ~10.4 nm respectively) matches closely with the Rg obtained from the [p(r)] distribution analysis (~7.65 nm and ~10.7 respectively).

SAXS analysis also estimated the maximum linear dimension (Dmax) for *St*CRC1 and *St*CRC2 as ~28 nm and ~34 nm respectively (**Figure 5.9 C & D**) which are in line with the results of AUC experiments where *St*CRC2 sediments with high $S_{20,w}$ value of ~12 s as higher molecular weight species as compared to *St*CRC1 which sediments with $S_{20,w}$ of ~9.5 s as lower molecular weight complex. The hexameric model of *St*SAT which we generated by applying the symmetric operation shows the diameter of 10.2 nm while the *St*CS structure in dimer form shows the diameter of 8.0 nm. By comparing the size of *St*SAT hexamer and *St*CS dimer and measuring the size of *St*CRC1 and *St*CRC2 complex theoretically, it can be inferred that our SAXS results are in line with the under-solution studies of the *St*CRC1 and *St*CRC2 complex.

Discussion

Multienzyme complexes are nature's design to regulate metabolic fluxes in organisms. They respond to changes in the levels of metabolites of the respective pathway and to metabolites and components of other pathways or processes that are spatiotemporally connected. They respond by increasing the rate of flux through metabolite channeling and enhancing catalytic efficiencies allosterically, self-regulating their assembly and dissociation with respect to metabolite concentrations, and increasing or decreasing catalytic efficiencies of component enzymes upon the physical association, etc. Cysteine regulatory complex, CRC, is a multienzyme complex that is linked to the regulation of sulfate and cysteine metabolism. However, only a few features of regulation are characterized to date and most details of the molecular mechanism of self-regulation remain unknown. We studied some of the structural and regulatory features of CRC from *Salmonella typhimurium* (*St*CRC) and report our findings here. *St*CRC consists of *St*SAT, a hexamer, and *St*CS, a dimer in solution. Both interact to form *St*CRC with very high affinity (Kd ~ 10.0 nM). The substrates and products of *St*SAT and *St*CS enzymes are known to regulate the *St*CRC complex. Na$_2$S, the substrate of CS enzyme, preferentially stabilizes the *St*CRC1 complex over *St*CRC2 complex; OAS, the second substrate of CS enzyme, is known to dissociate the *St*CRC2 complex into the *St*CRC1 and free *St*CS enzymes (Kaushik et al., 2017); NAS, the homolog of OAS, regulates the transcription of sulfate metabolizing enzymes; and the cysteine, the product of cysteine CS enzyme, directly regulates the SAT enzyme activity. However, the role of cysteine on CRC was not studied before. In the present work, we examined the effect of cysteine on the two assembly states of the *St*CRC complex, *St*CRC1, and *St*CRC2. By employing different analytical methods, we show that cysteine preferentially stabilizes the *St*CRC1 complex over the *St*CRC2 complex even at a high stoichiometric ratio of *St*CS dimers to *St*SAT hexamer. It is interesting to note that as the cysteine concentration is increased to 5.0 mM both *St*CRC2, as well as higher soluble oligomers, disappear, leaving only *St*CRC1. This suggests that cysteine is an effective dissociator of higher-order CRC but it is not able to dissociate CRC1 at the concentration ranges examined here. It is possible that increasing cysteine concentration further may destabilize the *St*CRC1 but with such a high stoichiometric ratio, cysteine to CRC would not be physiologically relevant. We propose that the dissociation mechanism of cysteine may be different than that of OAS and sulfide. It is possible that cysteine may not directly compete with the C-terminal tail of SAT for the active site of CS as is the case for OAS. Since cysteine is shown to be

the feedback inhibitor of SAT, we assume that cysteine may dissociate CRC by binding and remodeling the C-terminal of SAT rather than competing with SAT C-terminal for binding to the active site of CS. Further experiments would be needed to confirm the proposed model.

We examined the feedback inhibition of *St*SAT by cysteine (1.0 μM) and found that the turnover rate of *St*SAT is reduced to one third while enzyme efficiency is reduced to half of the uninhibited *St*SAT enzyme. To elucidate the structure and inhibition mechanism of *St*SAT by cysteine, we crystallized the *St*SAT protein in presence of cysteine. The *St*SAT enzyme crystallized as a dimer in the asymmetric unit where we observed highly asymmetric features of two trimers of *St*SAT hexamer. The C-terminal parts of monomers of the one trimer are fully unstructured and the structure of the other trimer is similar to that of SAT from other bacteria. Further structural analyses revealed a bound cysteine at the N-terminal domain near the end of the 8th α helix of each monomer of one trimer which has electron densities for most of their C-terminal. The monomers of another trimer, which have no cysteine ligand-bound show no electron density for the last 120 residues of the C-terminal. We also studied the low-resolution structural features of *St*CRC by SAXS and show that structural parameters estimated are consistent with the results of AUC.

There are two major outcomes of this structural study. First, the highly asymmetric nature of two trimers of the hexamer. This is the first structure of the SAT, to the best of our knowledge that the C-terminal domain of one of the trimers is missing. The second is the allosteric binding site of cysteine. The cysteine binding site mapped from our structure is different and away from the active site of the SAT, which is also a cysteine binding site in other SATs. It is possible that cysteine allosterically modulates the structural properties of the C-terminal of SAT. By modulating the structural properties of the C-terminal, cysteine is able to regulate both catalytic and protein-protein interactions as a structured C-terminal is necessary for both activities. However, the cysteine-mediated allosteric effect may have opposite outcomes on both activities. While a structured C-terminal and well defined active site can enhance the catalytic activity but reduce the affinity for binding to CS. These results are consistent with our proposed model that cysteine mediated dissociation mechanism of CRC may be different than that of OAS.

Summary

Summary

The presence of multienzyme complexes inside the cells makes the organisms to adapt and survive in the highly changing environment. The association and dissociation of multienzyme complexes is the basic mechanism through which they generate the new properties that are absent in the component enzymes. One of the multienzyme complexes whose regulation of association and dissociation was not studied much since its discovery in 1966 by Kredich is Cysteine regulatory complex (CRC). The CRC complex consists of two enzymes of the cysteine biosynthesis pathway; Serine acetyltransferase (SAT) and Cysteine synthase (CS). Both enzymes show different properties when they exist in free form like; SAT is less stable, less active, and less resistant to allosteric inhibitors while CS is highly active. The SAT after complex formation becomes more stable, more active, more resistant to allosteric inhibitors while CS becomes very less active. Recently, through multiple techniques, in our lab, it was discovered that CRC of *Salmonella typhimurium* exists in two oligomeric forms; CRC1 and CRC2 and these two forms exist in equilibrium with each other. Further, it was shown that this equilibrium can be shifted from CRC2 to CRC1 by Na₂S and from CRC1 to free SAT and free CS by OAS. However, till now it's not reported whether the CRC complex exists in *Mycobacterium tuberculosis* and how cysteine metabolite regulates the SAT enzyme in *Salmonella typhimurium*. So, we framed three objectives to study these features. In the first objective, we characterized the SAT enzyme of *Mtb*. First, we found, through the multiple sequence alignment, that *Mt*SAT protein does not have first ~98 N-terminal amino acids as compared to other bacterial SATs. Since *Mt*SAT lacks ~ 98 N-terminal amino acids, the protein is prone to aggregate. Therefore, highly stringent purification conditions were required to purify the *Mt*SAT enzyme. The purification results showed that the *Mt*SAT enzyme exists in three different oligomeric forms; dodecamer, hexamer, and trimer. This is striking as a normally only hexameric and trimeric form of SAT is seen in other bacterial species. We further characterized the different oligomers by AUC velocity experiments in a concentration-dependent manner where we found that the dodecameric form exists in equilibrium with hexameric form and increasing the concentration of dodecamer results in an increased fraction of dodecamer and the reduced fraction of hexamer. The hexameric fraction also shows that it exists in trimeric and hexameric form but increasing the concentration of hexamer does not change the equilibrium between two oligomers. However, the trimeric fraction does not show the other oligomeric fractions in the equilibrium even at higher concentrations which shows that the dissociation rate

plays a major role among the three oligomers as compared to the association rate resulting in the highly stable trimer. All three oligomers showed activity at pH ranging from 3.5 to 8.5, although, to a different extent. Further detailed steady-state kinetic characterization of all three fractions shows that hexameric form has the highest catalytic turnover rate followed by dodecamer and trimer. Further, the low-resolution structural characterization of hexamer by small-angle X-ray scattering experiments showed that hexamer has an elongated form similar to the SATs of other bacterial species.

In the second objective, we purified the CRC complex of *Mycobacterium* by mixing *Mt*SAT hexamer and *Mt*CS dimer in an appropriate stoichiometric ratio. In the SEC experiments, the stoichiometry of the *Mt*CRC complex was found to be one *Mt*SAT hexamer to four *Mt*CS dimers which was further confirmed by AUC velocity experiments. We have confirmed the purified CRC complex assembly by SDS PAGE analysis where we found the two bands, one at ~ 30 kDa of *Mt*SAT, and the second at ~ 35 kDa of *Mt*CS dimers. Further, we have characterized the binding interactions of the *Mt*CRC complex through the ITC experiments where we found that all four *Mt*CS dimers bind to *Mt*SAT hexamer in a single binding step with a high affinity. In the *Salmonella typhimurium St*CRC complex, it was observed that four *St*CS dimers bind to *St*SAT hexamer in two phases; in the first phase two *St*CS dimers bind to *St*SAT hexamer with very high affinity to form *St*CRC1 complex, and in the second phase, two more *St*CS dimers interact to *St*CRC1 complex with comparatively less affinity to form the *St*CRC2 complex. The occurrence of a single phase of four *Mt*CS dimer binding to one *Mt*SAT hexamer may be due to reasons of missing ~98 N-terminal amino acids in *Mt*SAT. Further, the catalytic characterization of *Mt*SAT and *Mt*CS in the *Mt*CRC complex showed that after complex formation the *Mt*SAT's catalytic turnover rate reduces to ~26 folds as compared to the free form of *Mt*SAT while the *Mt*CS's catalytic turnover rate does not change much to the free *Mt*CS enzyme. We compared the activity results of *Mt*SAT with the activity of *St*SAT where we find, *St*SAT in *St*CRC2 complex has 2-3 times more catalytic turn over rate as compared to free *St*SAT which is in line with other bacterial SAT.

In the third objective, we studied the effect of cysteine metabolites on the CRC complex of *Salmonella* and tried to understand the inhibition mechanism of cysteine on the *St*SAT enzyme. In our SEC experiments, we found that cysteine preferentially stabilizes the *St*CRC1 complex over the *St*CRC2 complex even at a high stoichiometric ratio of *St*CS dimers to *St*SAT hexamer. These

results were further confirmed by AUC velocity experiments which show that in absence of cysteine the *St*CRC complex exists in *St*CRC1, *St*CRC2, and some higher soluble oligomers which are reduced to highly homogenous species of *St*CRC1 complex in presence of cysteine. The steady-state kinetic experiments in the presence of cysteine show that the catalytic turnover rate is reduced by one-third as compared to the SAT's catalytic turnover rate in absence of cysteine. Further, the *St*SAT crystal structure was solved in presence of cysteine where we found the highly asymmetric nature of the *St*SAT enzyme. In the asymmetric unit, two chains of *St*SAT were observed; one with the N-terminal domain and C-terminal domain while another with only the N-terminal domain. After applying the symmetric operations, we generated the hexamer from the asymmetric unit where we found that one trimer does not have a C-terminal domain while the other consists of it, and the N-terminal domain of both trimers have three-fold axis symmetry. One cysteine ligand was found to be present at the end of the N-terminal domain of one trimer that consists C-terminal domain which shows that in presence of cysteine the C-terminal domain somehow gets properly folded. Further, the active site comparison of cysteine bound *St*SAT with the cysteine bound *Eh*SAT structure shows that in the *Eh*SAT cysteine binds to the Acetyl Co-A binding site present at the C-terminal domain while in *St*SAT cysteine binds at the end of the N-terminal domain suggesting a different mode of enzyme inhibition in *St*SAT. Further, the low-resolution structural features of *St*CRC1 and *St*CRC2 complexes were studied by small-angle X-ray scattering which shows that *St*CRC2 has ~ 3 nm higher radius of gyration and ~ 6 nm higher diameter than the *St*CRC1 complex in solution, confirming the results obtained from AUC studies.

In this thesis work, we have revealed and characterized, for the first time, the existence of the Cysteine regulatory complex from *Mycobacterium tuberculosis* and also discovered a new cysteine recognition site in the *Salmonella typhimurium* SAT. However, extensive studies are required to decipher the role of the *Mt*CRC complex in *Mycobacterium* pathogenesis. The existence of a novel cysteine binding site at the N-terminal domain of *St*SAT trimer opens the new domain to understand the allosteric inhibition mechanism in the *St*SAT enzyme which needs further confirmation.

Bibliography

References

Adams, P. D., Afonine, P. v., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 213–221. https://doi.org/10.1107/S0907444909052925

Agren, D., Schnell, R., Oehlmann, W., Singh, M., & Schneider, G. (2008). Cysteine synthase (CysM) of Mycobacterium tuberculosis is an O-phosphoserine sulfhydrylase: Evidence for an alternative cysteine biosynthesis pathway in Mycobacteria. *Journal of Biological Chemistry*, *283*(46), 31567–31574. https://doi.org/10.1074/jbc.M804877200

Banuls, A. L., Sanou, A., van Anh, N. T., & Godreuil, S. (2015). Mycobacterium tuberculosis: Ecology and evolution of a human bacterium. In *Journal of Medical Microbiology* (Vol. 64, Issue 11, pp. 1261–1269). Microbiology Society. https://doi.org/10.1099/jmm.0.000171

Benoni, R., Beck, C. M., Garza-Sánchez, F., Bettati, S., Mozzarelli, A., Hayes, C. S., & Campanini, B. (2017). Activation of an anti-bacterial toxin by the biosynthetic enzyme CysK: Mechanism of binding, interaction specificity and competition with cysteine synthase. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-09022-6

Berkowitz, O., Wirtz, M., Wolf, A., Kuhlmann, J., & Hell, R. (2002a). Use of biomolecular interaction analysis to elucidate the regulatory mechanism of the cysteine synthase complex from Arabidopsis thaliana. *Journal of Biological Chemistry*, *277*(34), 30629–30634. https://doi.org/10.1074/jbc.M111632200

Berkowitz, O., Wirtz, M., Wolf, A., Kuhlmann, J., & Hell, R. (2002b). Use of biomolecular interaction analysis to elucidate the regulatory mechanism of the cysteine synthase complex from Arabidopsis thaliana. *Journal of Biological Chemistry*, *277*(34), 30629–30634. https://doi.org/10.1074/jbc.M111632200

Bhat, S. A., Singh, N., Trivedi, A., Kansal, P., Gupta, P., & Kumar, A. (2012). The mechanism of redox sensing in Mycobacterium tuberculosis. In *Free Radical Biology and Medicine* (Vol. 53, Issue 8, pp. 1625–1641). https://doi.org/10.1016/j.freeradbiomed.2012.08.008

Bogdanova and Hell. (1997). *Cysteine synthesis in plants: protein-protein interactions of serine acetyltransferase from Arabidopsis thaliana. The Plant Journal 11(2), 251-262*.

Bonner, E. R., Cahoon, R. E., Knapke, S. M., & Jez, J. M. (2005). Molecular basis of cysteine biosynthesis in plants: Structural and functional analysis of O-acetylserine sulfhydrylase from Arabidopsis thaliana. *Journal of Biological Chemistry*, *280*(46), 38803–38813. https://doi.org/10.1074/jbc.M505313200

Brunner, K., Maric, S., Reshma, R. S., Almqvist, H., Seashore-Ludlow, B., Gustavsson, A. L., Poyraz, Ö., Yogeeswari, P., Lundbäck, T., Vallin, M., Sriram, D., Schnell, R., & Schneider, G. (2016). Inhibitors of the cysteine synthase CysM with antibacterial potency against dormant mycobacterium tuberculosis. *Journal of Medicinal Chemistry*, *59*(14), 6848–6859. https://doi.org/10.1021/acs.jmedchem.6b00674

Brunton, V. G., MacPherson, I. R. J., & Frame, M. C. (2004). Cell adhesion receptors, tyrosine kinases and actin modulators: A complex three-way circuitry. In *Biochimica et Biophysica Acta - Molecular Cell Research* (Vol. 1692, Issues 2–3, pp. 121–144). https://doi.org/10.1016/j.bbamcr.2004.04.010

Buldyrev, S. v., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, *464*(7291), 1025–1028. https://doi.org/10.1038/nature08932

Burkhard, P., Rao, G. S. J., Hohenester, E., Schnackerz, K. D., Cook, P. F., & Jansonius, J. N. (1998). *Three-dimensional Structure of O-acetylserine Sulfhydrylase from Salmonella typhimurium*.

Burkhard, P., Tai, C.-H., Ristroph, C. M., Cook, P. F., & Jansonius, J. N. (1999a). *Ligand Binding Induces a Large Conformational Change in O-Acetylserine Sulfhydrylase from Salmonella typhimurium*. http://www.idealibrary.com

Burkhard, P., Tai, C.-H., Ristroph, C. M., Cook, P. F., & Jansonius, J. N. (1999b). *Ligand Binding Induces a Large Conformational Change in O-Acetylserine Sulfhydrylase from Salmonella typhimurium*. http://www.idealibrary.com

Campanini, B., Speroni, F., Salsi, E., Cook, P. F., Roderick, S. L., Huang, B., Bettati, S., & Mozzarelli, A. (2005a). Interaction of serine acetyltransferase with O -acetylserine sulfhydrylase active site: Evidence from fluorescence spectroscopy . *Protein Science*, *14*(8), 2115–2124. https://doi.org/10.1110/ps.051492805

Campanini, B., Speroni, F., Salsi, E., Cook, P. F., Roderick, S. L., Huang, B., Bettati, S., & Mozzarelli, A. (2005b). Interaction of serine acetyltransferase with O -acetylserine sulfhydrylase active site: Evidence from fluorescence spectroscopy . *Protein Science*, *14*(8), 2115–2124. https://doi.org/10.1110/ps.051492805

Chakravarty, B., Gu, Z., Chirala, S. S., Wakil, S. J., & Quiocho, F. A. (2004). 44 15567-15572 BIOCHEMISTRY Downloaded at CSIR-Institute of Microbial Technology. In *PNAS November* (Vol. 2). www.ebi.ac.ukclustalW

Chandrasekhar, K., Wang, J., Arjunan, P., Sax, M., Park, Y. H., Nemeria, N. S., Kumaran, S., Song, J., Jordan, F., & Furey, W. (2013). Insight to the interaction of the dihydrolipoamide acetyltransferase (E2) core with the peripheral components in the Escherichia coli pyruvate dehydrogenase complex via multifaceted structural approaches. *Journal of Biological Chemistry*, *288*(21), 15402–15417. https://doi.org/10.1074/jbc.M113.466789

Chattopadhyay, A., Meier, M., Ivaninskii, S., Burkhard, P., Speroni, F., Campanini, B., Bettati, S., Mozzarelli, A., Rabeh, W. M., Li, L., & Cook, P. F. (2007). Structure, mechanism, and conformational dynamics of O-acetylserine sulfhydrylase from Salmonella typhimurium: Comparison of A and B isozymes. *Biochemistry*, *46*(28), 8315–8330. https://doi.org/10.1021/bi602603c

Claus, M. T., Zocher, G. E., Maier, T. H. P., & Schulz, G. E. (2005). *Structure of the O-Acetylserine Sulfhydrylase Isoenzyme CysM from Escherichia coli †, ‡*. https://doi.org/10.1021/bi050485

Cook, P. F., Hara, S., Nalabolu, S., & Schnackerz, K. D. (1992). pH Dependence of the Absorbance and 31P NMR Spectra of O-Acetylserine Sulfhydrylase in the Absence and Presence of O-Acetyl-L-Serine. *Biochemistry*, *31*(8), 2298–2303. https://doi.org/10.1021/bi00123a013

Domingo, G. J., Chauhan, H. J., Lessard, I. A. D., Fuller, C., & Perham, R. N. (n.d.). 1136±1146 (1999) q FEBS 1999. In *J. Biochem* (Vol. 266).

Droux, M., Ruffet, M.-L., Douce, R., & Job, D. (1998). Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants Structural and kinetic properties of the free and bound enzymes. In *Eur. J. Biochem* (Vol. 255).

Dunn, M. F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. In *Archives of Biochemistry and Biophysics* (Vol. 519, Issue 2, pp. 154–166). https://doi.org/10.1016/j.abb.2012.01.016

Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, *60*(12 I), 2126–2132. https://doi.org/10.1107/S0907444904019158

Feldman-Salit, A., Wirtz, M., Hell, R., & Wade, R. C. (2009). A Mechanistic Model of the Cysteine Synthase Complex. *Journal of Molecular Biology*, *386*(1), 37–59. https://doi.org/10.1016/j.jmb.2008.08.075

Fiegler, H., & Brückner, R. (2006). Identification of the serine acetyltransferase gene of Staphylococcus xylosus. *FEMS Microbiology Letters*, *148*(2), 181–187. https://doi.org/10.1111/j.1574-6968.1997.tb10286.x

Francois, J. A., Kumaran, S., & Jez, J. M. (2006). Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the Arabidopsis cysteine synthase complex. *Plant Cell*, *18*(12), 3647–3655. https://doi.org/10.1105/tpc.106.047316

Franke, D., Petoukhov, M. v., Konarev, P. v., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M., & Svergun, D. I. (2017). ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from

macromolecular solutions. *Journal of Applied Crystallography*, *50*, 1212–1225. https://doi.org/10.1107/S1600576717007786

Franke, Daniel, & 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. https://doi.org/10.1107/S0021889809000338

Fu, J., Yang, Y. R., Dhakal, S., Zhao, Z., Liu, M., Zhang, T., Walter, N. G., & Yan, H. (2016). Assembly of multienzyme complexes on DNA nanostructures. *Nature Protocols*, *11*(11), 2243–2273. https://doi.org/10.1038/nprot.2016.139

Gaitonde, M. K. (1967). A Spectrophotometric Method for the Direct Determination of Cysteine in the Presence of Other Naturally Occurring Amino Acids. In *Biochem. J* (Vol. 104).

Gerhart~, J. C., & Pardee$, A. B. (1962). The Enzymology of Control by Feedback Inhibition*. In *THE JOURNAL OP BIOLCI~ICAL CHEMISTRY* (Vol. 237, Issue 3). http://www.jbc.org/

Goldenberg, J., & Efroni, S. (2001). Using cellular automata modeling of the emergence of innovations. *Technological Forecasting and Social Change*, *68*(3), 293–308. https://doi.org/10.1016/S0040-1625(00)00095-0

Gorman, J., & Shapiro, L. (2004a). Structure of serine acetyltransferase from Haemophilus influenzae Rd. *Acta Crystallographica Section D: Biological Crystallography*, *60*(9), 1600–1605. https://doi.org/10.1107/S0907444904015240

Gorman, J., & Shapiro, L. (2004b). Structure of serine acetyltransferase from Haemophilus influenzae Rd. *Acta Crystallographica Section D: Biological Crystallography*, *60*(9), 1600–1605. https://doi.org/10.1107/S0907444904015240

Guan, R., Roderick, S. L., Huang, B., & Cook, P. F. (2008). Roles of histidines 154 and 189 and aspartate 139 in the active site of serine acetyltransferase from Haemophilus influenzae. *Biochemistry*, *47*(24), 6322–6328. https://doi.org/10.1021/bi800075c

Hatzios and Bertozzi. (2011). *The Regulation of Sulfur Metabolism in Mycobacterium tuberculosis*. *12*(4). https://doi.org/10.1371/journal

Hindson, V. J. (2003). Serine acetyltransferase of Escherichia coli: substrate specificity and feedback control by cysteine. In *Biochem. J* (Vol. 375).

Hindson, V. J., Moody, P. C. E., Rowe, A. J., & Shaw, W. v. (2000). Serine Acetyltransferase from Escherichia coli Is a Dimer of Trimers* Downloaded from. In *THE JOURNAL OF BIOLOGICAL CHEMISTRY* (Vol. 275, Issue 1). http://www.jbc.org/

Hindson, V. J., & Shaw, W. v. (2003). Random-order ternary complex reaction mechanism of serine acetyltransferase from Escherichia coli. *Biochemistry*, *42*(10), 3113–3119. https://doi.org/10.1021/bi0267893

Hiromasa, Y., Fujisawa, T., Aso, Y., & Roche, T. E. (2004). Organization of the Cores of the Mammalian Pyruvate Dehydrogenase Complex Formed by E2 and E2 Plus the E3-binding Protein and Their Capacities to Bind the E1 and E3 Components. *Journal of Biological Chemistry*, *279*(8), 6921–6933. https://doi.org/10.1074/jbc.M308172200

Huang, B., Vetting, M. W., & Roderick, S. L. (2005). The active site of O-acetylserine sulfhydrylase is the anchor point for bienzyme complex formation with serine acetyltransferase. *Journal of Bacteriology*, *187*(9), 3201–3205. https://doi.org/10.1128/JB.187.9.3201-3205.2005

Huang, X., Holden, H. M., & Raushel, F. M. (2001). *CHANNELING OF SUBSTRATES AND INTERMEDIATES IN ENZYME-CATALYZED REACTIONS*. www.annualreviews.org

Hussain, S., Ali, V., Jeelani, G., & Nozaki, T. (2009). Isoform-dependent feedback regulation of serine O-acetyltransferase isoenzymes involved in l-cysteine biosynthesis of Entamoeba histolytica. *Molecular and Biochemical Parasitology*, *163*(1), 39–47. https://doi.org/10.1016/j.molbiopara.2008.09.006

Hydesg, C., Ahmedliii, S. A., Padlans, E. A., Milesli, E. W., & Daviess, D. R. (1988). Three-dimensional Structure of the Tryptophan Synthase a & Multienzyme Complex from Salmonella typhimurium*. In *THE JOURNAL OF BIOLOGICAL CHEMISTRY* (Vol. 263, Issue 33).

Jez, J. M., & Dey, S. (2013). The cysteine regulatory complex from plants and microbes: What was old is new again. In *Current Opinion in Structural Biology* (Vol. 23, Issue 2, pp. 302–310). https://doi.org/10.1016/j.sbi.2013.02.011

Johnson, C. M., Huang, B., Roderick, S. L., & Cook, P. F. (2004). Kinetic mechanism of the serine acetyltransferase from Haemophilus influenzae. *Archives of Biochemistry and Biophysics*, *429*(2), 115–122. https://doi.org/10.1016/j.abb.2004.06.006

Jurgenson, C. T., Burns, K. E., Begley, T. P., & Ealick, S. E. (2008). Crystal structure of a sulfur carrier protein complex found in the cysteine biosynthetic pathway of Mycobacterium tuberculosis. *Biochemistry*, *47*(39), 10354–10364. https://doi.org/10.1021/bi800915j

Kaushik, A., Ekka, M. K., & Kumaran, S. (2017). Two Distinct Assembly States of the Cysteine Regulatory Complex of Salmonella typhimurium Are Regulated by Enzyme-Substrate Cognate Pairs. *Biochemistry*, *56*(18), 2385–2399. https://doi.org/10.1021/acs.biochem.6b01204

Kaushik, A., Rahisuddin, R., Saini, N., Singh, R. P., Kaur, R., Kaul, S., & Kumaran, S. (2020). Molecular Mechanism of Selective Substrate Engagement and Inhibitor Dis-engagement of Cysteine Synthase. *Journal of Biological Chemistry*, jbc.RA120.014490. https://doi.org/10.1074/jbc.ra120.014490

Kredich, N. M., & Becker, M. A. (1971). *[197] Cysteine Biosynthesis: Serine Transacetylase and O-Acetylserine Sulfhydrylase (Salmonella typhimurium) 1*.

Kredich, N. M., Becker, M. A., & Tomkins, G. M. (1969). Purification and Characterization of Cysteine Synthetase, a Bifunctional Protein Complex, from Salmonella typhimurium *. In *THE JOURNAL OF BIOLOGICAL CHEMISTRY* (Vol. 244, Issue 9). http://www.jbc.org/

Kredich, N. M., & Tomkins, G. M. (1966a). The Enzymic Synthesis of L-Cysteine in Escherichia coli and Salmonella typhimurium. In *THE JOURNAL OF BIOLOGICAL CHEVISTRY* (Vol. 241, Issue 21). http://www.jbc.org/

Kredich, N. M., & Tomkins, G. M. (1966b). The Enzymic Synthesis of L-Cysteine in Escherichia coli and Salmonella typhimurium. In *THE JOURNAL OF BIOLOGICAL CHEVISTRY* (Vol. 241, Issue 21). http://www.jbc.org/

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., & Nielsen, J. (2015). Microbial acetyl-CoA metabolism and metabolic engineering. In *Metabolic Engineering* (Vol. 28, pp. 28–42). Academic Press Inc. https://doi.org/10.1016/j.ymben.2014.11.009

Kumar, S., Kumar, N., Alam, N., & Gourinath, S. (2014). Crystal structure of serine acetyl transferase from Brucella abortus and its complex with coenzyme A. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1844*(10), 1741–1748. https://doi.org/10.1016/j.bbapap.2014.07.009

Kumar, S., Raj, I., Nagpal, I., Subbarao, N., & Gourinath, S. (2011a). Structural and biochemical studies of serine acetyltransferase reveal why the parasite Entamoeba histolytica cannot form a cysteine synthase complex. *Journal of Biological Chemistry*, *286*(14), 12533–12541. https://doi.org/10.1074/jbc.M110.197376

Kumar, S., Raj, I., Nagpal, I., Subbarao, N., & Gourinath, S. (2011b). Structural and biochemical studies of serine acetyltransferase reveal why the parasite Entamoeba histolytica cannot form a cysteine synthase complex. *Journal of Biological Chemistry*, *286*(14), 12533–12541. https://doi.org/10.1074/jbc.M110.197376

Kumaran, Hankuil, Y., Krishnan, H. B., & Jez, J. M. (2009). Assembly of the cysteine synthase complex and the regulatory role of protein-protein interactions. *Journal of Biological Chemistry*, *284*(15), 10268–10275. https://doi.org/10.1074/jbc.M900154200

Kumaran, S., & Jez, J. M. (2007). Thermodynamics of the interaction between O-acetylserine sulfhydrylase and the C-terminus of serine acetyltransferase. *Biochemistry*, *46*(18), 5586–5594. https://doi.org/10.1021/bi7001168

Kurganov, B. I. (1986). *The Role of Multienzyme Complexes in Integration of Cellular Metabolism* (Vol. 119).

Laemmli. (1970). *Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Naure VOL. 227 AUGUST 15, 1970.*

Lane, A. N., Paul, C. H., & Kirschner, K. (1984). The mechanism of self-assembly of the multi-enzyme complex tryptophan synthase from Escherichia coli. *The EMBO Journal*, *3*(2), 279–287. https://doi.org/10.1002/j.1460-2075.1984.tb01797.x

Leibundgut, M., Maier, T., Jenni, S., & Ban, N. (2008). The multienzyme architecture of eukaryotic fatty acid synthases. In *Current Opinion in Structural Biology* (Vol. 18, Issue 6, pp. 714–725). https://doi.org/10.1016/j.sbi.2008.09.008

Leu, L.-S., & Cook, P. F. (1994). Kinetic Mechanism of Serine Transacetylase from Salmonella typhimurium. In *Biochemistry* (Vol. 33).

Ma'ayan, A. (2017). Complex systems biology. *Journal of the Royal Society Interface*, *14*(134). https://doi.org/10.1098/rsif.2017.0391

Majewski, M., Ruiz-Carmona, S., & Barril, X. (2019). An investigation of structural stability in protein-ligand complexes reveals the balance between order and disorder. *Communications Chemistry*, *2*(1). https://doi.org/10.1038/s42004-019-0205-5

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(4), 658–674. https://doi.org/10.1107/S0021889807021206

Mino, K., Hiraoka, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2000). Characteristics of serine acetyltransferase from escherichia coli deleting different lengths of amino acid residues from the C-terminus. *Bioscience, Biotechnology and Biochemistry*, *64*(9), 1874–1880. https://doi.org/10.1271/bbb.64.1874

Mino, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (2001). Increase in the stability of serine acetyltransferase from Escherichia coli against cold inactivation and proteolysis by forming a bienzyme complex. *Bioscience, Biotechnology and Biochemistry*, *65*(4), 865–874. https://doi.org/10.1271/bbb.65.865

Mino, K., & Ishikawa, K. (2003). Characterization of a novel thermostable O-acetylserine sulfhydrylase from Aeropyrum pernix K1. *Journal of Bacteriology*, *185*(7), 2277–2284. https://doi.org/10.1128/JB.185.7.2277-2284.2003

Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A., & Nakanishi, K. (1999). Purification and characterization of serine acetyltransferase from escherichia coli partially truncated at the C-terminal region. *Bioscience, Biotechnology and Biochemistry*, *63*(1), 168–179. https://doi.org/10.1271/bbb.63.168

Murshudov Alexe, G. N., Vagin, A., & Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. In *Acta Cryst* (Vol. 53).

Nakamura, T., Iwahashi, H., & Eguchi, Y. (1984a). Enzymatic Proof for the Identity of the S-Sulfocysteine Synthase and Cysteine Synthase B of Salmonella typhimurium Downloaded from. In *JOURNAL OF BACTERIOLOGY* (Vol. 158, Issue 3). http://jb.asm.org/

Nakamura, T., Iwahashi, H., & Eguchi, Y. (1984b). Enzymatic Proof for the Identity of the S-Sulfocysteine Synthase and Cysteine Synthase B of Salmonella typhimurium Downloaded from. In *JOURNAL OF BACTERIOLOGY* (Vol. 158, Issue 3). http://jb.asm.org/

Nemeria, N., Baykal, A., Joseph, E., Zhang, S., Yan, Y., Furey, W., & Jordan, F. (2004). Tetrahedral intermediates in thiamin diphosphate-dependent decarboxylations exist as a 1′,4′-imino tautomeric form of the coenzyme, unlike the Michaelis complex or the free coenzyme. *Biochemistry*, *43*(21), 6565–6575. https://doi.org/10.1021/bi049549r

Olsen, L. R., Huang, B., Vetting, M. W., & Roderick, S. L. (2004). Structure of serine acetyltransferase in complexes with CoA and its cysteine feedback inhibitor. *Biochemistry*, *43*(20), 6013–6019. https://doi.org/10.1021/bi0358521

Paritala, H., & Carroll, K. S. (2013). *New Targets and Inhibitors of Mycobacterial Sulfur Metabolism §*.

Patel, M. S., Nemeria, N. S., Furey, W., & Jordan, F. (2014). The pyruvate dehydrogenase complexes: Structure-based function and regulation. In *Journal of Biological Chemistry* (Vol.

289, Issue 24, pp. 16615–16623). American Society for Biochemistry and Molecular Biology Inc. https://doi.org/10.1074/jbc.R114.563148

Paulsen, C. E., & Carroll, K. S. (2013). Cysteine-mediated redox signaling: Chemistry, biology, and tools for discovery. In *Chemical Reviews* (Vol. 113, Issue 7, pp. 4633–4679). https://doi.org/10.1021/cr300163e

Philo, J. S. (2006). Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. *Analytical Biochemistry*, *354*(2), 238–246. https://doi.org/10.1016/j.ab.2006.04.053

Poole, L. B. (2015). The basics of thiols and cysteines in redox biology and chemistry. In *Free Radical Biology and Medicine* (Vol. 80, pp. 148–157). Elsevier Inc. https://doi.org/10.1016/j.freeradbiomed.2014.11.013

Proschel, M., Detsch, R., Boccaccini, A. R., & Sonnewald, U. (2015). Engineering of metabolic pathways by artificial enzyme channels. In *Frontiers in Bioengineering and Biotechnology* (Vol. 3, Issue OCT). Frontiers Media S.A. https://doi.org/10.3389/fbioe.2015.00168

Pye, V. E., Tingey, A. P., Robson, R. L., & Moody, P. C. E. (2004a). The structure and mechanism of serine acetyltransferase from Escherichia coli. *Journal of Biological Chemistry*, *279*(39), 40729–40736. https://doi.org/10.1074/jbc.M403751200

Pye, V. E., Tingey, A. P., Robson, R. L., & Moody, P. C. E. (2004b). The structure and mechanism of serine acetyltransferase from Escherichia coli. *Journal of Biological Chemistry*, *279*(39), 40729–40736. https://doi.org/10.1074/jbc.M403751200

Qiu, J., Ma, Y., Owusu, L., Jiang, T., & Xin, Y. (2014). Functional analysis of serine acetyltransferase from Mycobacterium smegmatis. *Journal of Basic Microbiology*, *54*(7), 670–677. https://doi.org/10.1002/jobm.201300858

Qiu, J., Wang, D., Ma, Y., Jiang, T., & Xin, Y. (2013). Identification and characterization of serine acetyltransferase encoded by the mycobacterium tuberculosis rv2335 gene. *International Journal of Molecular Medicine*, *31*(5), 1229–1233. https://doi.org/10.3892/ijmm.2013.1298

Rabeh, W. M., & Cook, P. F. (2004). Structure and mechanism of O-acetylserine sulfhydrylase. In *Journal of Biological Chemistry* (Vol. 279, Issue 26, pp. 26803–26806). https://doi.org/10.1074/jbc.R400001200

Rahbari, M., Rahlfs, S., Jortzik, E., Bogeski, I., & Becker, K. (2017). H2O2 dynamics in the malaria parasite Plasmodium falciparum. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal

Rege, V. D., Kredich, N. M., Tai, C.-H., Karsten, W. E., Schnackerz, K. D., & Cook, P. F. (1996). *A Change in the Internal Aldimine Lysine (K42) in O-Acetylserine Sulfhydrylase to Alanine Indicates Its Importance in Transimination and as a General Base Catalyst †* (Vol. 35). https://pubs.acs.org/sharingguidelines

Rengarajan, J., Bloom, B. R., & Rubin, E. J. (2005). 8327-8332 MICROBIOLOGY Downloaded at CSIR-Institute of Microbial Technology. In *PNAS* (Vol. 102, Issue 23). www.pnas.orgcgidoi10.1073pnas.0503272102

Saito et al. (1995). *Molecular Cloning and Characterization of Plant Serine Acetyltransferase Playing a Regulatory Role in Cyateine Biosynthesis from Watermelon. J. Biol. Chem.-1995-Saito-16321-6.*

Salsi, E., Bayden, A. S., Spyrakis, F., Amadasi, A., Campanini, B., Bettati, S., Dodatko, T., Cozzini, P., Kellogg, G. E., Cook, P. F., Roderick, S. L., & Mozzarelli, A. (2010). Design of O-acetylserine sulfhydrylase inhibitors by mimicking nature. *Journal of Medicinal Chemistry*, *53*(1), 345–356. https://doi.org/10.1021/jm901325e

Salsi, E., Campanini, B., Bettati, S., Raboni, S., Roderick, S. L., Cook, P. F., & Mozzarelli, A. (2010). A two-step process controls the formation of the bienzyme cysteine synthase complex. *Journal of Biological Chemistry*, *285*(17), 12813–12822. https://doi.org/10.1074/jbc.M109.075762

Sambrook and Russell. (2001). *Molecular cloning : a laboratory manual*.

Sambrook et al. (1989). *Molecular cloning : a laboratory manual*.

Schnell, R., Oehlmann, W., Singh, M., & Schneider, G. (2007). Structural insights into catalysis and inhibition of O-acetylserine sulfhydrylase from Mycobacterium tuberculosis: Crystal

structures of the enzyme α-aminoacrylate intermediate and an enzyme-inhibitor complex. *Journal of Biological Chemistry*, *282*(32), 23473–23481. https://doi.org/10.1074/jbc.M703518200

Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal*, *78*(3), 1606–1619. https://doi.org/10.1016/S0006-3495(00)76713-0

Seifert, F., Golbik, R., Brauer, J., Lilie, H., Schröder-Tittmann, K., Hinze, E., Korotchkina, L. G., Patel, M. S., & Tittmann, K. (2006). Direct kinetic evidence for half-of-the-sites reactivity in the E1 component of the human pyruvate dehydrogenase multienzyme complex through alternating sites cofactor activation. *Biochemistry*, *45*(42), 12775–12785. https://doi.org/10.1021/bi061582l

Sekowska, A., Kung, H.-F., & Danchin, A. (2000a). Sulfur Metabolism in Escherichia coli and Related Bacteria: Facts and Fiction JMMB Review. In *Sulfur Metabolism in Enterobacteria 145 J. Mol. Microbiol. Biotechnol* (Vol. 2, Issue 2). www.caister.com/bacteria-plant

Sekowska, A., Kung, H.-F., & Danchin, A. (2000b). Sulfur Metabolism in Escherichia coli and Related Bacteria: Facts and Fiction JMMB Review. In *Sulfur Metabolism in Enterobacteria 145 J. Mol. Microbiol. Biotechnol* (Vol. 2, Issue 2). www.caister.com/bacteria-plant

Singh, A. K., Ekka, M. K., Kaushik, A., Pandya, V., Singh, R. P., Banerjee, S., Mittal, M., Singh, V., & Kumaran, S. (2017). Substrate-Induced Facilitated Dissociation of the Competitive Inhibitor from the Active Site of O-Acetyl Serine Sulfhydrylase Reveals a Competitive-Allostery Mechanism. *Biochemistry*, *56*(37), 5011–5025. https://doi.org/10.1021/acs.biochem.7b00500

Smolle, M., Prior, A. E., Brown, A. E., Cooper, A., Byron, O., & Lindsay, J. G. (2006). A new level of architectural complexity in the human pyruvate dehydrogenase complex. *Journal of Biological Chemistry*, *281*(28), 19772–19780. https://doi.org/10.1074/jbc.M601140200

Srivastava, D. K., & Bernhard, S. A. (2015). *Metabolite Transfer via Enzyme-Enzyme Complexes Concentration of Enzymes and Metabolites in the Cell* (Vol. 10). www.sciencemag.org

Stafford. (1992). Boundary Analysis in Sedimentation Transport Experiments: A Procedure for Obtaining Sedimentation Coefficient Distributions Using the Time Derivative of the Concentration Profile. In *ANALYTICAL BIOCHEMISTRY* (Vol. 203).

Steiner, E. M., Bo¨th, D., Lo¨ssl, P., Vilaplana, F., Schnell, R., & Schneider, G. (2014). CysK2 from Mycobacterium tuberculosis is an O-phospho-L-serine-dependent S-sulfocysteine synthase. *Journal of Bacteriology*, *196*(19), 3410–3420. https://doi.org/10.1128/JB.01851-14

Sweetlove, L. J., & Fernie, A. R. (2018). The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. In *Nature Communications* (Vol. 9, Issue 1). Nature Publishing Group. https://doi.org/10.1038/s41467-018-04543-8

Tai, C.-H., Nalabolu, S. R., Jacobson, T. M., Minter, D. E., & Cook'pll, P. F. (1993). Kinetic Mechanisms of the A and B Isozymes of 0-Acetylserine Sulfhydrylase from Salmonella typhimurium LT-2 Using the Natural and Alternative Reactants? In *Biochemistry* (Vol. 32).

Thoden, J. B., Cook, P. D., Schäffer, C., Messner, P., & Holden, H. M. (2009). Structural and functional studies of QdtC: An N-acetyltransferase required for the biosynthesis of dTDP-3-acetamido-3,6-dideoxy-R-D-glucose. *Biochemistry*, *48*(12), 2699–2709. https://doi.org/10.1021/bi802313n

Thoden, J. B., Reinhardt, L. A., Cook, P. D., Menden, P., Cleland, W. W., & Holden, H. M. (2012). Catalytic mechanism of perosamine N-acetyltransferase revealed by high-resolution X-ray crystallographic studies and kinetic analyses. *Biochemistry*, *51*(16), 3433–3444. https://doi.org/10.1021/bi300197h

van Rossum, H. M., Kozak, B. U., Niemeijer, M. S., Duine, H. J., Luttik, M. A. H., Boer, V. M., Kötter, P., Daran, J. M. G., van Maris, A. J. A., & Pronk, J. T. (2016). Alternative reactions at the interface of glycolysis and citric acid cycle in Saccharomyces cerevisiae. *FEMS Yeast Research*, *16*(3). https://doi.org/10.1093/femsyr/fow017

Vinitsky, A., Michaud, C., Powers, J. C., & Orlowski',$, M. (1992). Inhibition of the Chymotrypsin-like Activity of the Pituitary Multicatalytic Proteinase Complex+. In *Biochemistry* (Vol. 31).

Walde, P., Umakoshi, H., Stano, P., & Mavelli, F. (2014). Emergent properties arising from the assembly of amphiphiles. Artificial vesicle membranes as reaction promoters and regulators. *Chemical Communications*, *50*(71), 10177–10197. https://doi.org/10.1039/c4cc02812k

Wang, T., & Leyh, T. S. (2012). Three-stage assembly of the cysteine synthase complex from Escherichia coli. *Journal of Biological Chemistry*, *287*(6), 4360–4367. https://doi.org/10.1074/jbc.M111.288423

Warrilow, A. G. S., & Hawkesford1, M. J. (2000). Cysteine synthase (O-acetylserine (thiol) lyase) substrate specificities classify the mitochondrial isoform as a cyanoalanine synthase. In *Journal of Experimental Botany* (Vol. 51, Issue 347).

Wei, X., Han, P., & You, C. (2020). Facilitation of cascade biocatalysis by artificial multi-enzyme complexes — A review. In *Chinese Journal of Chemical Engineering*. Chemical Industry Press. https://doi.org/10.1016/j.cjche.2020.05.022

Wheeler, P. R., Coldham, N. G., Keating, L., Gordon, S. v., Wooff, E. E., Parish, T., & Hewinson, R. G. (2005). Functional demonstration of reverse transsulfuration in the Mycobacterium tuberculosis complex reveals that methionine is the preferred sulfur source for pathogenic mycobacteria. *Journal of Biological Chemistry*, *280*(9), 8069–8078. https://doi.org/10.1074/jbc.M412540200

Williams, R. A. M., Westrop, G. D., & Coombs, G. H. (2009). Two pathways for cysteine biosynthesis in Leishmania major. *Biochemical Journal*, *420*(3), 451–462. https://doi.org/10.1042/BJ20082441

Wirtz, M., Birke, H., Heeg, C., Müller, C., Hosp, F., Throm, C., König, S., Feldman-Salit, A., Rippe, K., Petersen, G., Wade, R. C., Rybin, V., Scheffzek, K., & Hell, R. (2010). Structure and function of the hetero-oligomeric cysteine synthase complex in plants. *Journal of Biological Chemistry*, *285*(43), 32810–32817. https://doi.org/10.1074/jbc.M110.157446

Wirtz, M., Droux, M., & Hell, R. (2004). O-acetylserine (thiol) lyase: An enigmatic enzyme of plant cysteine biosynthesis revisited in Arabidopsis thaliana. *Journal of Experimental Botany*, *55*(404), 1785–1798. https://doi.org/10.1093/jxb/erh201

Wooff et al. (2002). *Functional genomics reveals the sole sulphate transporter of the Mycobacterium tuberculosis complex and its relevance to the acquisition of sulphue in vivo. Molecular Microbiology (2002) 43(3), 653-663.*

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013a). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *Journal of Biological Chemistry*, *288*(51), 36463–36472. https://doi.org/10.1074/jbc.M113.527143

Yi, H., Dey, S., Kumaran, S., Lee, S. G., Krishnan, H. B., & Jez, J. M. (2013b). Structure of soybean serine acetyltransferase and formation of the cysteine regulatory complex as a molecular chaperone. *Journal of Biological Chemistry*, *288*(51), 36463–36472. https://doi.org/10.1074/jbc.M113.527143

Zhang, Y. H. P. (2011). Substrate channeling and enzyme complexes for biotechnological applications. In *Biotechnology Advances* (Vol. 29, Issue 6, pp. 715–725). https://doi.org/10.1016/j.biotechadv.2011.05.020

Zhang, Y., & Hess, H. (2017). Toward Rational Design of High-efficiency Enzyme Cascades. In *ACS Catalysis* (Vol. 7, Issue 9, pp. 6018–6027). American Chemical Society. https://doi.org/10.1021/acscatal.7b01766

Zhao, C., Moriga, Y., Feng, B., Kumada, Y., Imanaka, H., Imamura, K., & Nakanishi, K. (2006). On the interaction site of serine acetyltransferase in the cysteine synthase complex from Escherichia coli. *Biochemical and Biophysical Research Communications*, *341*(4), 911–916. https://doi.org/10.1016/j.bbrc.2006.01.054

Publications

Publications

1. Kaushik, A., **<u>Rahisuddin, R</u>**., Saini, N., Singh, R. P., Kaur, R., Kaul, S., & Kumaran, S. (2020). Molecular Mechanism of Selective Substrate Engagement and Inhibitor Disengagement of Cysteine Synthase. *Journal of Biological Chemistry*, jbc-RA120.

2. **<u>Rahisuddin, R</u>**., Neha Saini, S. Kumaran. Distinct Assembly states of Serine acetyltransferase and Cysteine regulatory complexes of Mycobacterium tuberculosis. (***to be communicated***)

3. Ravi Pratap Singh$^\$$, Saini, N.$^\$$, Gaurav Sharma, **<u>R. Rahisuddin</u>**, Abhishek Kaushik, S. Kumaran*. Moonlighting Biochemistry of Cysteine Synthase: A Species-specific Global Regulator. **Equal contribution (Manuscript to be communicated)**

4. Saini, N., **<u>R. Rahisuddin</u>**, Vinod Choudhry, and S. Kumaran*. Rational Design of Decoys and models for High-Throughput Virtual Screening of TNF Inhibitors **(Manuscript to be communicated)**

PDB DEPOSITED

7C35, 7CM8, 7DJQ

Annexure

Molecular Mechanism of Selective Substrate Engagement and Inhibitor Dis-engagement of Cysteine Synthase

Abhishek Kaushik, R. Rahisuddin, Neha Saini, Ravi P. Singh, Rajveer Kaur, Sukirte Kaul and S. Kumaran*

G. N. Ramachandran Protein Center, Institute of Microbial Technology (IMTECH), Council of Scientific and Industrial Research (CSIR), Sector 39-A, Chandigarh, India 160036

[*]**Corresponding Author**: S. Kumaran

Institute of Microbial Technology (IMTECH)

Council of Scientific and Industrial Research (CSIR)

Sector 39-A, Chandigarh, India

Email: *skumaran@imtech.res.in*

Tel: +91 -0172 -6665474

FAX: 91-172-2690585.

Running Title: Ligand Discriminative Mechanisms of Enzymes

Keywords: Competitive allostery; Substrate selectivity; Cysteine synthesis; X-ray crystallography; Fluorescence