DEVELOPMENT OF STRATEGIES FOR PROTEIN BIOSEPARATION, BIOCATALYST DESIGNS AND PROTEIN REFOLDING

By

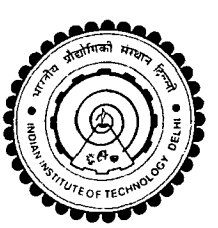
Smita Raghava

Chemistry Department

Submitted

in fulfillment of the requirement of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

MARCH, 2009

Keep your face to the sunshine and you cannot see the shadow.

- Helen Keller

To

My Parents

CERTIFICATE

This is to certify that the thesis entitled "**Development of strategies for protein bioseparation, biocatalyst designs and protein refolding**" being submitted by **Ms. Smita Raghava** to the Indian Institute of Technology Delhi for the award of the degree of Doctor of Philosophy in Chemistry is a record of bonafide research work carried out by her. Ms. Raghava has worked under our supervision, and has fulfilled the requirements for the submission of the thesis which, to our knowledge, has reached requisite standard.

The results contained in this dissertation have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma.

Prof. M.N. Gupta
(Supervisor)
Department of Chemistry
IIT Delhi

Dr. S. K. Khare
(Co-Supervisor)
Department of Chemistry
IIT Delhi

Acknowledgements

This thesis arose in part out of years of research that has been done since I joined IIT. During this journey, I have worked with a great number of people whose contribution in assorted ways to the research and the making of the thesis deserves special mention. It is a pleasure to convey my gratitude to them all in my humble acknowledgement.

In the first place I would like to record my gratitude to **Prof. M. N. Gupta** *for his supervision, advice, and guidance from the very early stage of this research. Above all and the most needed, he provided me unflinching encouragement and support in various ways. In addition, he was always accessible and willing to help. As a result, research life became smooth and rewarding for me. His passion in science exceptionally inspires and enriches my growth as a student and a researcher. I am indebted to him more than he knows.*

I gratefully acknowledge **Dr. S. K. Khare** *for his advice and guidance as my co-supervisor. He has always been a constant source of encouragement during my research work. Thank you, Sir, for always being supportive.*

I am indebted to **Prof. B. Jayaram**, *Head, Department of Chemistry, for providing the necessary facilities during the work.*

I would like to acknowledge **Indian Council of Medical Research** *(ICMR) for supporting my research with the fellowship grants. The generous support from* **Dr. Sandhya Diwakar**, *Scientist E, ICMR, is greatly appreciated.*

Prof. A. K. Singh *&* **Prof. P.S. Pandey**, *Department of Chemistry, and* **Dr. Prashant Mishra**, *Department of Biochemical Engineering and Biotechnology, deserve a special thanks as my research committee members.*

I am much indebted to **Prof. Raghavan Varadarajan**, *Molecular Biophysics Unit, Indian Institute of Science, Bengaluru, for providing the DNA constructs, his advice and guidance. He always kindly granted me his time even for answering some of my unintelligent questions. Thank you sir for your suggestion to attend the Gordon Research Conference on Proteins and your help while I was applying for it. Attending the conference was a stimulating and a memorable experience for me.*

I gratefully thank **Prof. Dinakar Salunke**, *National Institute of Immunology (NII), New Delhi, for helping me with N-terminal sequence analysis.*

Many thanks to **Dr. R. P. Roy**, *NII, for his advice and generous help with analytical ultracentrifugation studies. Thank you, Sir, for being so approachable. Discussions with you have always helped me be more analytical in my approach.*

I gratefully thank **Prof. Viresh Dutta**, *Centre for Energy Studies, IIT Delhi, for providing the titanium dioxide nanoparticles and* **Prof. A.K. Ganguli**, *IIT Delhi, for his generous help with zeta potential measurements.*

My sincere thanks are due to **Dr. Aparna Sharma**, **Dr. Sunita Teotia**, **Dr. Meryam Sardar** *and* **Dr. Ipsita Roy** *who helped me adjust in the lab during my early years and for their valuable suggestions. I thank Dr. Aparna for those free car rides back home at lunch times.*

I would also acknowledge my seniors **Sulakshana Jain**, **Kalyani Mondal** *and* **Shweta Shah** *for their advice and their willingness to share their bright thoughts with me, which were very fruitful for shaping up my ideas and research. I thank* **Anshu** *and* **Ruchi** *for extending their helping hand whenever needed.*

*Collective and individual acknowledgements are also owed to my colleagues **Abir** (Thanks for helping me with HPLC and spectropolarimeter), **Sohel** (Thanks for helping me in chromatography), **Pradeep** (Thanks for being cooperative and of course, the scientific discussions throughout), **Kusum and Veena** (Thanks for helping me with the nitty-gritties of the lab and always listening to me), **Manali, Gulam and Benu** (Thanks for obediently following or trying to follow my orders) for creating a pleasant working atmosphere in the lab.*

*To **Mahendra, Sarah, Malini,** thanks for your willingness to help even in the work not directly related to you.*

*To **Anjani,** thanks for your technical assistances and always telling me, "Don't worry".*

*To Piar Chand (**PC** as we call him), thanks for all the help and support in making the lab a cleaner place to work in.*

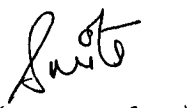
*My six months stint at **Institute of Genomics and Integrative Biology** (IGIB), Delhi, opened my mind and motivated me to join Ph.D. I thank **Prof. P. Usha Sarma** for hiring me to her team. I could never have embarked and started all of this without her prior teachings in molecular biology and biochemistry. Thank you, Ma'am.*

*I wish to thank my friends, **Smriti** and **Samra,** for helping me get through the difficult times, and for all the emotional support and camaraderie.*

*Where would I be without my family? My parents deserve special mention for their inseparable support and prayers. My Father, **Dr. R. C. Raghava,** in the first place is the person who has shown me the joy of intellectual pursuit ever since I was a child. My Mother, **Mrs. Anita Raghava,** is the one who sincerely raised me with her caring and love. To them I dedicate this thesis.*

*Words fail me to express my appreciation to my sister, **Swita** and brother-in-law, **Sushil,** whose love and persistent confidence in me, has taken the load off my shoulder. They have always been one of my best counselors.*

*I thank my brother, **Ishashis,** for being a supportive sibling. Talking with you has always helped release the tension.*



*(**Smita Raghava**)*

Abstract

Proteins are the workhorses of biology. Some of the challenges in protein science are:

- Producing proteins efficiently and economically. The downstream component here is often the costlier one. Hence, the recent efforts have been made to develop more efficient strategies for protein isolation and purification. Bioseparation of proteins has also assumed considerable importance in the context of proteomics.
- Creating efficient biocatalyst designs involving higher stability and greater catalytic efficiency. Also, use of biocatalysts to obtain chiral compounds has emerged as one of the powerful applications of enzymology.
- Recovering activity from inactive proteins, especially inclusion bodies.
- Understanding interactions of proteins with nanoparticles.

These aspects form the subject of studies described in the present thesis.

Chapter 1 is the introductory chapter and reviews the literature on various techniques of protein extraction and purification including the affinity tag-based approaches. Development of strategies for efficient biocatalyst designs has also been reviewed. This chapter also provides an overview of the methods to obtain correctly refolded proteins.

Three phase partitioning (TPP) is a simple technique which involves addition of a salt (mostly ammonium sulfate) and a water miscible organic solvent (mostly t-

butanol) to an aqueous protein solution. **Chapter 2** describes the use of this technique for selective permeabilization of microbial cells. Microbial cells suspended in an aqueous buffer were subjected to TPP where the permeabilized cells formed the interface below the upper organic solvent layer and proteins were released in the lower aqueous layer. It was found that a preincubation step with t-butanol for different time periods before initiating TPP by addition of ammonium sulfate controlled the permeability of the cells. This, in turn, led to selective release of the proteins of a specific size range from the permeabilized cells of *Thermus thermophilus*, *Saccharomyces cerevisiae*, and *Escherichia coli*. Smaller proteins (green fluorescent protein and lipase, molecular weights of 29 kDa and 34 kDa, respectively) were released with preincubation time of 15 min whereas penicillin G acylase (~85 kDa) was released with preincubation time of 30 min. The high molecular weight proteins (alcohol dehydrogenase from *S. cerevisiae* and *T. thermophilus*, molecular weights of 150 kDa and 170 kDa, respectively) were retained even after 60 min of preincubation. The specific activities and electrophoretic analysis of some of the proteins obtained reflected their high purity. Transmission electron microscopy showed structural alterations of *E. coli* membrane during permeabilization. Hence, the strategy described in this chapter outlines a powerful approach to the use of permeabilization for isolating proteins in considerably pure form right at this early upstream step.

The maltose binding protein (MBP) affinity tag has been extensively used for protein purification. **Chapter 3** outlines an affinity precipitation approach to

purify MBP-fusion proteins. Affinity precipitation consists of adding a smart macro-(affinity ligand) to a protein solution and selectively precipitating the affinity complex of the desired proteins. Unlike chromatography, affinity precipitation takes place from free solutions and has a lot less mass transfer limitations during the formation of the affinity complex. Affinity precipitation is also easily scalable. An industrial grade modified starch was used as a 'smart' material and a combination of polyethylene glycol (PEG)/Ca^{2+} was used as a stimulus. The preparation initially used to develop the purification strategy was a fairly pure MBP. It was found to co-precipitate as a complex with the cationic starch with 10 % (wv^{-1}) PEG and 50 mM $CaCl_2$. Overexpressed MBP (in *E. coli*) and MBP-CD4bs (a soluble MBP fusion with a fragment of human immunodeficiency virus protein gp120) fusion protein could be precipitated by using a similar strategy. The precipitated MBP or MBP fusion protein could be selectively dissociated by suspending the precipitate in 1 M NaCl. In the case of overexpressed MBP preparation (about 65 % pure according to the densitometer analysis of SDS-PAGE), 75 % of the loaded protein bound to the polymer (which showed high but not absolute selectivity at the binding stage) and 84 % of the bound protein was eluted. In the case of MBP-CD4bs fusion protein, 38 % of the contaminating proteins could be removed by precipitation with PEG/$CaCl_2$ and 100 % of the fusion protein was recovered. In all cases, the purified proteins showed a single band on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and the expected changes in fluorescence emission

spectra upon binding to maltose. The data from the Scatchard plots yielded a K_a of 4.3 x 10^6 M^{-1} for MBP and 16.0 x 10^6 M^{-1} for MBP-CD4bs. These values of K_a for maltose binding to MBP were obtained was in close agreement with earlier reported data. The method compares favorably with the existing approach of using chromatographic media to purify MBP-tagged proteins.

Alcohol dehydrogenases (ADH) from microorganisms have attracted considerable attention because of their valuable applications in biotransformations. While aldehydes, ketones and alcohols have always been useful in food, pharmaceutical and fine chemical industries, chiral compounds are becoming increasingly important for obtaining chirally pure pharmaceutical products. In this regard, enzymes in general and ADH in particular from thermophiles are of considerable value because of their higher stability during the industrial process.

In **Chapter 4**, purification and characterization of an NAD^+-dependent and zinc containing ADH from *Thermus thermophilus* (TTHADH) is described. The enzyme could be purified with 25-fold purification and 68% yield using a single step purification protocol. The SDS-PAGE showed a single band of molecular weight 43 kDa. Analytical ultracentrifugation showed the molecular mass to be about 170 kDa. This observation indicated the tetrameric nature of enzyme. The N-terminal amino acid sequence of TTHADH appeared to be different from sequences of ADHs from some of the other thermophiles like *Thermoanaerobacter brockii*, *Thermococcus roseus*, *Thermus litoris*, and *Thermococcus essas.* The pH optimum of the purified enzyme was 8.8 and the

temperature optimum was found to be 80°C. Thermal denaturation curves were determined by monitoring the CD values at 222 nm and the T_m was found to be 89°C. The enzyme showed much higher activity towards glycerol as compared to short chain primary and secondary alcohols. This thermostable enzyme was also highly stereospecific in oxidation of glycerol and converted glycerol into D-glyceraldehyde. The unusual specificity of the enzyme towards glycerol has considerable industrial importance. A large amount of glycerol is being produced as a byproduct of biodiesel production. This has led to search for conversions/ biotransformations which can convert glycerol into value added products. The enzyme which converts glycerol into a chiral molecule like D-glyceraldehyde opens up several synthetic opportunities.

Immobilization of enzymes has become an established approach for enhancing their usefulness in biotechnology. Enzymes immobilized on smart polymers can catalyze the reaction in homogeneous phase but can be recovered just like heterogeneous catalysts. Thermosensitive polymers are smart polymers which precipitate as the temperature is raised beyond lower critical solution temperature (LCST). Conjugation of these polymers with proteins results in thermoresponsive biocatalysts and allows their separation from aqueous solutions induced by a small shift of temperature. Covalent attachment of enzymes and other proteins to the smart polymer, poly (N-isopropylacrylamide) [poly (NIPAAm)], has been widely used as a method for the preparation of thermosensitive protein conjugates.

In **Chapter 5**, reversibly soluble-insoluble polymer-enzyme conjugates were prepared by conjugating a copolymer of NIPAAm with 5-mol % of 6-acrylaminohexanoic acid to trypsin by the carbodiimide-NHS (N-hydroxysuccinimide) coupling method. Four bioconjugates with different units of enzyme coupled to the matrix were prepared. Increased enzymatic activity in terms of high effectiveness factor (in the range of 3–5) was found in the conjugates. Kinetic parameters for the bioconjugates and free enzyme were determined. The V_{max}/K_m value of the enzyme significantly increased on immobilization by the factors in the range of 12–28. The bioconjugates also showed stability towards autolysis at 50°C. The optimum pH for the bioconjugates was found to be same as that of free trypsin, i.e., pH 8.2. However, the immobilized preparations showed broader pH optima with large activity surviving at alkaline pH (pH 9–10 region). To investigate the effect of conjugation on accessibility to the active sites of trypsin, soybean trypsin inhibitor (STI, molecular weight of 20,000) was employed. The data showed that STI was unable to totally block the access of BAPNA (low molecular weight amide substrate) to the enzyme active site in conjugates. However, with high molecular weight substrate, casein, both trypsin and the conjugates were completely inhibited by STI, indicating that all enzyme molecules accessible to casein were accessible to STI as well. The bioconjugates showed a considerable increase (1.5-2.5 times) in quantum yield of fluorescence.

The goal of obtaining enzyme forms which show greater stability, higher catalytic efficiency, and reusability has been pursued over several decades. Several approaches such as immobilization, site-directed mutagenesis, and directed evolution have been tried.

A recent biocatalyst design cross-linked enzyme aggregate (CLEA) is based upon cross-linking of a protein which is initially precipitated by adding an organic solvent or a salt. Cross-linked enzyme aggregates (CLEAs) form an effective methodology for enzyme immobilization. The enzyme does not need to be of high purity. The relative simplicity of the procedure ideally lends itself to high-throughput experimentation.

In **Chapter 6**, formation of CLEAs of restriction endonucleases, *Eco*RI and *Bam*HI and a DNA polymerase, *Tth* DNA polymerase directly from the lysate of bacterial cells, has been reported. *Eco*RI and *Bam*HI CLEAs were found to be stable at 25°C for 10 days. The restriction endonuclease CLEAs could be reused five times. CLEAs showed impressive enhancement in resistance to thermal inactivation of enzyme activity. Half-life ($t_{1/2}$) in case of *Eco*RI CLEAs was found to be 5.4 h at 50 °C as compared to 0.2 h of free enzyme and *Bam*HI CLEAs had a half-life ($t_{1/2}$) of 4.0 h at 50 °C as compared to 0.18 h of free enzyme. *Eco*RI CLEAs (unlike most commercially available preparations of the enzyme) had no star activity in presence of organic solvents. The *Tth* DNA polymerase CLEA was found to withstand larger number of cycles (as compared to free enzyme) during DNA amplification by polymerase chain reaction (PCR).

Upon overexpression in *E. coli*, many recombinant proteins form dense, insoluble aggregates called inclusion bodies. It is therefore necessary to develop efficient and scalable procedures for solubilization and refolding of recombinant proteins from inclusion bodies.

In **Chapter 7**, the application of TPP (described in chapter 2) to refold recombinant proteins from inclusion bodies has been reported. The inclusion bodies were solubilized by unfolding the proteins in 8M urea/100 mM dithiothreitol and TPP was carried out. The refolded proteins were obtained as interfacial precipitates and could be dissolved in aqueous buffer. TPP was used to refold twelve different proteins from their inclusion bodies. These proteins were ribonuclease A (RNase A), the first two domains of human CD4 (CD4D12), five insoluble constructs of the protein tyrosine phosphatases (PTPs) from *Drosophila melanogaster* and five highly unstable and aggregation prone mutants of the *E. coli* proteins controller of cell division or death B (CcdB), maltose binding protein (MBP) and thioredoxin (Trx). In each case, the protein refolded by TPP gave either higher refolding yield than the earlier reported method or succeeded where earlier efforts have failed. The RNase A refolded by the TPP method was characterized and compared to protein purified from inclusion bodies by the conventional method. Yields of 40-45 mg l^{-1} were obtained by TPP method in merely 36 hours as compared to 4-8 mg l^{-1} obtained by conventional methods in 3-4 days. In case of CD4D12, the yield obtained was 3-5 mg l^{-1} in 17 hours as compared to 1 mg l^{-1} by conventional method in 24 hours. CcdB mutants, MBP

mutants, Trx mutant and five constructs of PTPs could not be refolded by conventional methods whereas the TPP method of refolding resulted in yields ranging from 30-60 mg l^{-1}. The methodology is scaleable, parallelizable and no concentration step is required subsequent to refolding.

The role of both chaperonins and polymeric additives such as polyethylene glycol (PEG) in protein refolding arises from their binding to hydrophobic structures in the partially unfolded protein molecules, and this inhibits their aggregation. Stimuli-sensitive polymers are presumed to act similarly. In some cases, simultaneous purification along with refolding is also observed during affinity precipitation. In that respect, smart polymers offer an advantage over polymers such as PEG.

In **Chapter 8**, alginate, a calcium-sensitive polymer, has been used to refold and simultaneously purify chemically denatured and thermally denatured α-amylase and Eudragit™ S-100, a pH-sensitive poly (methyl methacrylate), has been used to refold overexpressed CcdB protein that was obtained in the form of inclusion bodies. In case of α-amylase, activity recovery of 80 % was obtained with chemically denatured protein and 70 % with thermally denatured protein. The fluorescence spectra showed refolding, and PAGE showed the absence of any aggregates in the refolded preparation. Initial experiments with wild-type (WT) CcdB showed that Eudragit™ bound and precipitated (upon lowering the pH to 4.0) CcdB quantitatively from the latter's aqueous solution. In this case, the protein could not be not be eluted off the polymer. CcdB was obtained as a

bioconjugate with Eudragit™ S-100 and the bioconjugate showed DNA gyrase inhibition activity and could be recycled. The inclusion bodies of CcdB mutant, CcdB-F17P, were solubilized in 8 M urea/100 mM dithiothreitol. This preparation could be refolded by precipitation with Eudragit™. The fluorescence and CD spectra showed correct protein refolding. Unfortunately, the bioconjugate of CcdB-F17P with Eudragit™ contained contaminating nuclease activity and hence could not be assayed for DNA gyrase inhibition.

Interactions of proteins with particles of micrometre sizes have been extensively studied. This knowledge has already proved valuable in the area of applied biocatalysis and biology. Interactions of proteins with particles of nanodimensions are less studied but data related to this has many obvious applications in the diverse areas of bioreactor design, biosensors, medical devices and microfluidics. Bionanofabrications with inherent component of self-assembly has already generated sufficient interest in bioinspired designs.

In **Chapter 9**, it is shown that simple unmodified titanium dioxide (TiO_2) particles (~10 nm in size) effectively assist refolding of thermally denatured basic proteins α-chymotrypsin, ribonuclease A and papain. The zeta potential of titanium dioxide nanoparticles changed from -36±1.5 mV to -17±1.4 mV when α-chymotrypsin was bound to the particles. This confirmed that interactions between proteins and nanoparticles were electrostatic in nature. The refolded enzymes regained nearly 100% activity in all the three cases and their CD spectra were similar to

corresponding CD spectra of these enzymes in their native form. Dynamic light scattering (DLS) showed that complexes between TiO_2 nanoparticles and denatured proteins reached their maximum sizes in the same time period (i.e., 1 hour) which was optimum for regaining the biological activities during nanoparticle assisted refolding.

Table of Contents