Structural and Functional Studies of Cinnamoyl CoA Reductase (CCR) from *Leucaena leucocephala*

A THESIS
SUBMITTED TO THE
UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

BY
Prashant Dhanajirao Sonawane

Under the guidance of
DR. B. M. KHAN

PLANT TISSUE CULTURE DIVISION
NATIONAL CHEMICAL LABORATORY
PUNE- 411008
INDIA
July 2013

Dedicated to My Uncle

Late Mr. Ramesh

CONTENTS

ACKNOWLEDGEMENTS

I take this opportunity to gratefully acknowledge my research guide ***Dr. B. M. Khan*** *for his invaluable guidance, unending support and keen interest during the course of investigation. He has given me the freedom to think and work, and I shall cherish my learning experience under him. Although this eulogy is insufficient, I preserve an everlasting gratitude for him. I truly feel privileged to have joined his research group. I sincerely thank for the care and affection that I received from him and his family in the entire period.*

I would like to thank my research co-guide ***Dr. C. G. Suresh****, Department of Biochemical Sciences, NCL for his valuable comments and suggestions throughout this research work.*

I am immensely grateful to ***Dr. Sushama Gaikwad****, for her constant support, encouragement, valuable comments and precious suggestion during the progress and completion of thesis work.*

I wish my sincere thanks to ***Dr. S. K. Rawal****, Former Head, PTC Division, NCL, for his comments and suggestions during the progress of this work.*

I would like to extend my thanks to ***Dr. D. C. Agarwal, Dr. Urmil Mehta*** *for allowing me to use their lab facilities and all other members of the scientific and supporting staff (Mr. Mahale, Mr. Jathar and Mr. Khamkar, of the Plant Tissue Culture Division, Biochemical Sciences Division and NCL Library. My sincere thanks to Mrs. S. Kendurkar for suggestions and encouragement. Mr. Suryaprasad deserves special thanks for using his technical expertise for timely repairing of the equipments.*

My sincere thanks to: ***Dr. M. J. Kulkarni****, Biochemical Sciences Division, NCL, for permission to use MALDI-TOF/MS facility,* ***Dr. Monisha Fernandes****, Organic Chemistry Division, NCL, for permission to use CD facility,* ***Dr. Suresh Bhat****, Polymer and Advanced Materials Division, NCL, for help in DLS facility,* ***Dr. Guruswamy Kumarswamy****, Polymer and Advanced Materials Division, NCL, for help in using the SAXS facility.*

I express my deep feelings and love for my lab seniors Dr. Sameer, Dr. Arun, Dr. Abhilash, Dr. Noor, Dr. Pallavi, Dr. Ruby, Dr. Santosh Kumar, Dr. Sumita and Dr.

R.J. Santoshkumar, Dr. Ansary, Dr. Imran, Dr. Shashidhara, Dr. Asad for their encouragement and help during start of this work.

I would like to take this opportunity to express my love and thank my lab mates Dr. Rishi, Somesh, Krunal, Parth, Neha, Shakeel, Uma, Poonam, Kannan for their good company and pleasant atmosphere in the lab and for countless things they have done for me and for always being there with me whenever I needed them.

I would like to mention a special thanks to my friends Dr. Rishi, Somesh, Trupti Kad for a wonderful time inside and outside the lab and for always being there for my support.

I take this opportunity to thank ***Dr. Rishi Kishore Vishwakarma*** *for affection, listening and sharing personal as well as professional things and valuable suggestions and pumping confidence in me during my testing times.*

Dr. Sameer deserves special thanks for initiating work on CCR and their constant support and suggestions.

I thank my trainees Neerja, Pratik and Henna for their valuable help during my work period.

I would also like to express my deep felt gratitude to my friends Dr. Rishi, Somesh, Krunal, Trupti, Sudhanshu, Parth, Shakeel, Ruby, Santosh R.J., Neha, Pooja, Uma, Manas, Manisha, Pushkar, Dr. Fazal, Dr. Santosh, Ezaz, Devdutta (DD), Prabhakar, Kapil, Dr. Bhuban, Dr. Prasad, Raju, Poonam Lapalikar, Rita, Ravi, Prasad Abnave, Yashwant, Sandip, Vitthal, Kannan, Neha Khandelwal, Abhishek, Shamim, Pallavi, Madhurima, Sayali, Sonali, Avinash, Nishant, Manas, Deepak, Tulika, Ruby Singh, Priyabrat, Parul, Priya and Manojkumar Sharma for the support, best wishes and lighter moments shared together throughout my work period.

Special mention is a must for Rishi, Somesh and Parth for being my extended family during the course of my work. Those were really memorable days when, on weekend we all including Neha, Pooja, Krunal and Ruby gathered for dinner. They have given me a perfect family atmosphere and I never ever felt that I am away from my home.

I take this opportunity to thank my M.Sc. friends Manoj (Munnya), Ravikiran (Bhai), Pravin (Povya), Parimal (Parya), Nilesh, Sagar (Shaggy), Adhish (chotu), Ameya,

Vivek (Adhyaksha), Nitin, Rohit (Robert), Prasad (Dadya) for affection, wishes and encouragement.

I also extend my thanks to my childhood friends, Amol (Buklya), Raju, Santosh (Pappu), Vikas, Ravi (Bhai), Vivek, Ashu, Krushnakant (Shendya) for their constant encouragement, belief and support throughout my work.

I owe my deepest gratitude to my parents, ***Aai*** *and* ***Nana****, and my aunty* ***Alka*** *for their blessing, constant support and unconditional affection, which helped me in building my career.*

I would like to take this opportunity to extend my sincere thanks to my brothers, ***Sachin*** *(bhau),* ***Kishore*** *(aaba),* ***Amit*** *(Ambu) and* ***Mayuresh*** *(Sonya) who have been a constant source of inspiration for me during my struggle through the thick and thins of my PhD work. This work would not have been possible without their constant support and sacrifice.*

Special thanks to ***Smita*** *Vahini (Rani),* ***Monali*** *Vahini (Manu) for their help, suggestions, constant support and encouragement.*

I find no words for my cute nephews, ***Sarthak*** *(Pillu) and* ***Shourya*** *(Pandurang, Ryan, Chiku, Shavi, Nelson and so many) for making me alive and refreshing every time when I look at them.*

Thanks to all who directly or indirectly helped during my Ph. D tenure and I might have forgotten to mention those names here.

I am grateful to Dr. Saurav Pal, Director, NCL for providing necessary facilities and permitting me to submit my findings in the form a thesis.

CSIR, India is duly acknowledged for fellowship and grants.

Last but not the least I thank ***almighty God*** *for giving me the strength and courage at every step of my life.*

Prashant D. Sonawane

राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा मार्ग पुणे - 411 008.भारत
NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411 008 India.

Dr. Bashir M. Khan
Head
Plant Tissue Culture Division

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled **"Structural and Functional Studies of Cinnamoyl CoA Reductase (CCR) from *Leucaena leucocephala*"** submitted by **Prashant D. Sonawane** for the degree of Doctor of Philosophy, was carried out under my supervision at the Plant Tissue Culture Division, National Chemical Laboratory, Pune. Materials obtained from other sources have been duly acknowledged in the thesis.

Dr. C. G. Suresh
(Research Co-guide)

Dr. B. M. Khan
(Research guide)

Off : 091 20 2590 2220
Res. : 091 20 2589 3251
Fax : 091 20 2590 2645
e-mail bm.khan@ncl.res.in
URL : www.ncl-india.org.
091 20 2590 2785

DECLARATION

I hereby declare that the thesis entitled **"Structural and Functional Studies of Cinnamoyl CoA Reductase (CCR) from *Leucaena leucocephala*"** has been carried out at Plant Tissue culture Division, National Chemical Laboratory, Pune, under the guidance of **Dr. B. M. Khan**. The work is original and has not been submitted in part or full by me for any other degree or diploma to any other university. I further declare that the materials obtained from other sources have been duly acknowledged in the thesis.

Sonawane

(Prashant D. Sonawane)

Date: July, 2013

Place: Plant Tissue Culture Division
National Chemical Laboratory (NCL)
Pune, Maharashtra, India
Pin- 411008

LIST OF ABBREVIATIONS USED

AA	Amino acid
ANS	1-anilino-8-naphthalenesulfonate
Arg	Arginine
Asp	Aspartic acid
CA	Citraconic acid
bp	Base pairs
CD	Cirular dichroism
C3H	Coumarate 3-hydroxylase
C4H	Cinnamate 4-hydroxylase
CAD	Cinnamyl alcohol dehydrogenase
Cald5H/F5H	Coniferaldehyde 5-hydroxylase/Ferulate 5-hydroxylase
CCoAOMT	Caffeoyl coenzyme A 3-O-methyltransferase
CCR	Cinnamoyl CoA reductase
cDNA	Complementary DNA
cm	Centimeter
COMT	Caffeate O-methyltransferase
CV	Column Volumes
Cys	Cysteine
Da	Dalton
DEAE	Diethyl aminoethyl
DEPC	Diethylpyrocarbonate
DLS	Dynamic light scattering
DNA	Deoxyribose nucleic acid
EDTA	Ethylene Diamine Tetra Acetic acid

GdnHCl	Guanidium hydrochloride
GT	Glycosyltransferase
g /L	Grams per litre
g	Gram
G	Guaiacyl
h	Hour(s)
HCT	p-hydroxycinnamoyl-CoA: quinate shikimate p-hydroxycinnamoyl-transferase (HCT)
FPLC	Fast performance liquid chromatography
His	Histidine
Ile	Isoleucine
IPTG	Isopropyl ß-D-thiogalactoside
Kb	Kilobase pairs
kDa	Kilo Daltons
Kg	Kilogram
K	Kelvin
K_a	Association constant
K_{cat}	Catalytic Efficiency
K_{cat}/ K_m	Specificity constant
K_d	Dissociation constant
kJ	Kilo joules
K_m	Michaelis-Menten constant
K_{sv}	Stern-Volmer (dynamic) quenching constant
L	Litre
Leu	Leucine

Ll-CCRH1	*Leucaena leucocephala* cinnamoyl CoA reductase 1
Lys	Lysine
LB	Luria-Bertani
m	Meter
MALDI-TOF	Matrix assisted laser desorption ionization-Time of flight
mg	Milligram
min	Minute(s)
mL	Millilitre
µL	Micro liter
µg	Microgram
MRE	Mean residue ellipticity
mM	Millimolar
mRNA	Messenger RNA
MS	Mass spectrometery
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NAI	N-acetyl imidzole
NBS	N-bromosuccinimide
NEM	N-ethylmaleimide
nM	Nano molar
nm	Nanometer
NRMSD	Normalized root mean square deviation
O/N	Overnight
OD	Optical density
O-MT	O-methyltransferase

PAL	Phenylalanine ammonia lyase
PCR	Polymerase Chain Reaction
PDB	Protein data bank
Pg	Phenylglyoxal
Phe	Phenylalanine
pHMB	p-hydroxymercurybenzoate
pI	Isoelectric point
pKa	Acid dissociation constant
PMSF	Phenyl methyl sulphonyl fluoride
PVDF	Polyvinlylidene difluoride
RMSD	Root mean square deviation
RNA	Ribose nucleic acid
RT	Room temperature
s	Second(s)
S	Syringyl
SAD	Sinapyl alcohol dehydrogenase
SAXS	Small angle X-ray scattering
SDS	Sodium dodecyl sulphate
Ser	Serine
Sp.	Species
T_m	Melting temperature
Trp	Tryptophan
Tyr	Tyrosine
PAGE	Polyacrylamide gel electrophoresis
SMQ	Sterile Milli Q

U	Units
UV	Ultraviolet
v/v	Volume / Volume
w/v	Weight / Volume
WRK	Woodward's reagent K
V_{max}	Rate of an enzyme catalyzed reaction, maximum reaction rate
α	Alpha
β	Beta
λ	Lamda
%	Percentage
°C	Degree Celsius
µg	Microgram
µg/L	Micrograms per liter
µL	Microlitre
µm	Micrometer
µM	Micromolar
4CL	4-Coumarate coenzyme A ligase

Abstract

In plants, biosynthesis of monolignols is a specialized branch of phenylpropanoid metabolism, a complex series of branching biochemical reactions responsible for synthesis of variety of products like lignin, flavonoids and hydroxycinnamic acid conjugates. Cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the first committed step in monolignol biosynthesis and considered as a first regulatory point in lignin formation. CCR carries out the NADPH dependent reduction of various hydroxycinnamoyl CoA esters to corresponding hydroxycinnamaldehydes and vice versa. Previous studies on CCR from various plants were basically focused on isolation, cloning, molecular characterization and downregulation aspects. Although CCR is one of the most investigated enzyme of lignin biosynthesis pathway, its three dimensional structure remains to be determined. The lack of the three dimensional crystal structure of CCR has precluded a clarification of functional active site residues involved in substrate binding and catalysis. Also, little has been known about the significance of cofactor (NADPH) and its interaction with CCR enzyme. Thus, the functional features of CCR enzyme, like mechanism of catalysis and multiple substrate specificity are still unresolved. On the other hand, absolutely no reports are available on CCR conformation, intrinsic fluorescence and characterization of trp microenvironment. In this context, the present study is aimed at understanding the structural and functional aspects of Cinnamoyl CoA reductase from *Leucaena leucocephala* (Ll-CCRH1)

Chapter 1: General introduction

This chapter gives information on Lignin, its biosynthesis and role of *L. leucocephala* in paper and pulp industry in India. Background of research done on cinnamoyl CoA

reductase (CCR), a key regulatory enzyme in lignin biosynthesis has been dealt in detail. A thorough literature survey of work done on CCR in the area of genetic engineering with lignin biosynthesis pathway genes with regards to the current status of research in this area has been presented. Structure-Functional studies on CCR from various plants have been discussed in detail and need of more research on CCR has been elaborated. Finally, the scope of the present study and objectives of the thesis work have been discussed.

Chapter 2: **Biochemical Characterization of recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) from *L. leucocephala***

This chapter deals with the detailed biochemical characterization of recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) from *L. leucocephala*. Recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) protein from *L. leucocephala* was overexpressed in *E. coli* BL21 (DE3) strain and purified to apparent homogeneity. Optimum pH for forward and reverse reaction was found to be 6.5 and 7.8 respectively. The enzyme was most stable around pH 6.5 at 25 °C for 90 min. The enzyme showed K_{cat}/K_m for feruloyl, caffeoyl, sinapoyl, coumaroyl CoA, coniferaldehyde and sinapaldehyde as 4.6, 2.4, 2.3, 1.7, 1.9 and 1.2 (x 10^6 M^{-1} s^{-1}), respectively, indicating better affinity of enzyme for feruloyl CoA over other substrates and preference of reduction reaction over oxidation. Activation energy, E_a for various substrates was found to be in the range of 20-50 kJ/mol. Involvement of probable carboxylate ion, histidine, lysine or tyrosine at the active site of enzyme was predicted by pH activity profile. SAXS studies of protein showed radius 3.04 nm and volume 49.25 nm^3 with oblate ellipsoid shape. Metal ion inhibition studies revealed that Ll-CCRH1 is a metal independent enzyme.

Chapter 3A: Probing the active site of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *L. leucocephala*

This chapter deals with the active site characterization of Ll-CCRH1 by means of molecular, computational and biochemical methods. Putative active site residues involved in the substrate/NADPH binding and catalysis for *L. leucocephala* CCR (Ll-CCRH1; GenBank: DQ986907) were identified by amino acid sequence alignment and homology modeling. Putative active site residues and proximal H215 were subjected for site directed mutagenesis, and mutated enzymes were expressed, purified and assayed to confirm their functional roles. Mutagenesis of S136, Y170 and K174 showed complete loss of activity, indicating their pivotal roles in catalysis. Mutant S212G exhibited the catalytic efficiencies less than 10% of wild type, showing its indirect involvement in substrate binding or catalysis. R51G, D77G, F30V and I31N double mutants showed significant changes in K_m values, specifying their roles in substrate binding. Finally, chemical modification and substrate protection studies corroborated the presence Ser, Tyr, Lys, Arg and carboxylate group at the active site of Ll-CCRH1.

Chapter 3 B: *In silico* mutagenesis and docking studies of active site residues of cinnamoyl Co reductase 1 (Ll-CCRH1)

This chapter deals with the *in silico* mutagenesis and docking studies of active site residues of cinnamoyl Co reductase 1 (Ll-CCRH1). CCR shows multiple substrate specificity towards various cinnamoyl CoA esters. Here, *in silico* mutagenesis studies of active site residues of Ll-CCRH1 were carried out. Homology modeling based modeled 3D structure of Ll-CCRH1 was used as template for *in silico* mutant preparations. Docking simulations of Ll-CCRH1 mutants with CoA esters by

AutoDock Vina tools showed altered substrate specificity as compared to wild type. The study evidences that conformational changes and changes in geometry or architecture of active site pocket occurred following mutations. The altered substrate specificity for active site mutants suggests the possible physiological role of CCR either in lignin formation or in defense system in plants.

Chapter 4: Conformational transitions of cinnamoyl CoA reductase 1(Ll-CCRH1) from *L. leucocephala*

Lack of folding- unfolding studies on CCR limits their structural characterization. As a step towards structural characterization, Ll-CCRH1 was subjected to denaturation with GdnHCl, pH and temperature. Conformational transitions of cinnamoyl CoA reductase; a key regulatory enzyme in lignin biosynthesis, from *L. leucocephala* (Ll-CCRH1) were studied using fluorescence and circular dichroism spectroscopy. The native protein possesses four trp residues exposed on the surface and 66% of helical structure, undergoes rapid structural transitions at and above 45 °C and starts forming aggregates at 55 °C. Ll-CCRH1 was transformed into acid induced (pH 2.0) molten globule like structure, exhibiting altered yet compact secondary structure, diminished tertiary structure and exposed hydrophobic residues. The molten globule like structure was examined for the thermal and chemical stability. The altered secondary structure of L1-CCRH1 at pH 2.0 was stable upto 90 °C and also in the vicinity of 2 M Guanidine hydrochloride (GdnHCl) (as compared to drastic loss of native structure in 2 M GdnHCl) as seen in far UV-CD spectra. The structural transition of Ll-CCRH1 at pH 2.0 was reversible, as all the characteristics of molten globule had diminished after readjusting the pH to 8.0, that of native protein.

CHAPTER 5: Steady state fluorescence studies of wild type recombinant cinnamoyl CoA reductase (Ll-CCRH1) and its active site mutants

Fluorescence quenching and time resolved fluorescence of wild type recombinant Ll-CCRH1, a multitryptophan protein from *L. leucocephala* and ten different active site mutants were carried out to investigate tryptophan environment. The enzyme showed highest affinity for feruloyl CoA (K_a= 3.72 x 10^5 M^{-1}) over other CoA esters and cinnamaldehydes, as determined by fluorescence spectroscopy. Quenching of the fluorescence by acrylamide for wild type and active site mutants was collisional with almost 100% of the tryptophan fluorescence accessible under native condition and remained same after denaturation of protein with 6 M GdnHCl. In wild type Ll-CCRH1, the extent of quenching achieved with iodide (f_a=1.0) was significantly higher than cesium ions (f_a=0.33) suggesting more density of positive charge around surface of trp conformers under native conditions. Denaturation of wild type protein with 6 M GdnHCl led to significant increase in the quenching with cesium (f_a=0.54), whereas quenching with iodide ion was decreased (f_a=0.78), indicating reorientation of charge density around trp from positive to negative and heterogeneity in trp environment. The Stern-Volmer plots for wild type and mutants Ll-CCRH1 under native and denatured conditions with cesium ion yielded biphasic quenching profiles, indicating that the trp residues in the protein fall into at least two groups that differ considerably in their accessibility and/or environment. The extent of quenching for cesium and iodide ions under native and denatured conditions observed in active site mutants is significantly different from wild type Ll-CCRH1 under same conditions. Thus, single substitution type mutations of active site residues showed heterogeneity in tryptophan microenvironment and differential degree of conformation of protein under native or denatured conditions. The native enzyme showed two different lifetimes, τ1 (2.27 ns)

and τ2 (7.92 ns) with average lifetime, (τ) 5.28 ns, which decreased τ1 (1.38 ns) and τ2 (3.59 ns) with (τ) 2.16 ns after denaturation with 6 M GdnHCl.

List of Publications:

- **Prashant Sonawane**, K. Patel, R.K. Vishwakarma, S. Srivastava, S. Singh, S.M. Gaikwad, Bashir M. Khan, **(2013)** Probing the active site of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, International Journal of Biological Macromolecules 60, 33-38.
- **Prashant Sonawane,** R.K. Vishwakarma, Bashir M. Khan, **(2013)** Biochemical characterization of recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, International Journal of Biological Macromolecules 58, 154-159.
- **Prashant Sonawane**, K. Patel, R.K. Vishwakarma, S. Singh, Bashir M. Khan, **(2013)** In silico mutagenesis and docking studies of active site residues suggest altered substrate specificity and possible physiological role of Cinnamoyl CoA Reductase 1 (Ll-CCRH1), Bioinformation 9, 224-232.
- **Prashant Sonawane,** B.M. Khan, Sushama M. Gaikwad, Cinnamoyl CoA reductase 1 from *Leucaena leucocephala*: Acid induced, thermostable molten globule. (Under review).
- R.J. Santoshkumar, Ruby, S. Singh, **Prashant Sonawane**, R.K. Vishwakarma, Bashir M. Khan, **(2013)** Functional Characterization of a Glucosyltransferase Specific to Flavonoid 7-OGlucosides from *Withania somnifera*, Plant Molecular Biology Reporter (Accepted, http: //dx. doi. org/ 10.1007/s11105-013-0573-4).
- R.K. Vishwakarma, Ruby, S. Singh, **Prashant Sonawane**, S. Srivastava, U. Kumari, R.J. Santoshkumar, Bashir M. Khan, **(2013)** Molecular Cloning, Biochemical Characterization, and differential expression of an Acetyl-CoA C-Acetyltransferase

gene (AACT) of Brahmi (*Bacopa monniera*), Plant Molecular Biology Reporter 31, 547-557.

- S. Singh, R.K. Vishwakarma, R.J. Santoshkumar, **Prashant Sonawane**, Ruby, Bashir M. Khan, **(2013)** Functional characterization of a flavonoid glycosyltransferase gene from *Withania somnifera* (Ashwagandha), Applied Biochemistry and Biotechnology (Accepted, http: //dx. doi. org/ 10.1007/s12010-013-0230-2).
- R.K. Vishwakarma, K. Patel, **Prashant Sonawane**, S. Singh, Ruby, U. Kumari, Bashir M. Khan, **(2012)** Molecular Characterization of Farnesyl pyrophosphate Synthase from *Bacopa monniera* by comparative modeling and docking studies, Bioinformation 22, 1075-1081.
- R.K. Vishwakarma, **Prashant Sonawane**, S. Singh, U. Kumari, Ruby, Bashir M. Khan, **(2013)** Molecular characterization and differential expression studies of an oxidosqualene cyclase (OSC) gene of Brahmi (*Bacopa monniera*), Physiology and Molecular Biology of Plants (Accepted, http: //dx. doi. org/10.1007/s12298-013-0195-1).
- B.M. Khan, S.K. Rawal, M. Arha, S.K. Gupta, S. Srivastava, N.M. Shaikh, A.K. Yadav, P.S. Kulkarni, O.U. Abhilash, S. Gupta, S. Omer, R.K. Vishwakarma, S. Singh, R.J. Santoshkumar, **Prashant Sonawane**, P. Patel, K. Chinnathambi, S. Abbassi, **(2012)** Genetic Engineering of Phenylpropanoid Pathway in *Leucaena leucocephala*, In: Genetic Engineering- Basics, new applications and responsibilities InTech Croatia, 93-120, ISBN 978-953-307-671-3

CHAPTER 1

Introduction

1.1 Paper and pulp industry

Trees are reservoirs of many economically and biotechnologically significant products. Wood, one such gifts of nature, has diverse application for mankind. One of the most well known applications of wood is pulp and paper industry. Wood is the major raw material for the production of pulp and paper (Food and Agriculture Organization, 2001a). Nowadays, paper has become the integral part of our life, as its significance is evident in numerous ways we use it. Worldwide annual production of paper has increased more than three to four folds in the past forty years, amounting to a total production of around 350 millions tones (FAO, Forest report 2003). The growth of paper and paper product consumption in world is expected to be the highest during the upcoming decades.

1.2 Papermaking processes

Two major processes are used for the production of paper pulp: chemical and mechanical. The chemical process mainly uses chemicals to remove lignin from fiber cell walls to get long and flexible fibers that consist of polysaccharide only. Chemical pulping yields individual intact fibers that can interact with other fibers via hydrogen bonds due to removal of hydrophobic lignin and thus, making a very strong paper. The mechanical pulping process focuses on mechanical separation of fibers without removal of lignin [1]. Although mechanical pulping gives the highest pulp yield, but the obtained pulp has limited bleachability and may revert in brightness upon exposure to light, that is, paper becomes yellow due to presence of lignin. Previously, kraft process was the most widely used chemical procedure for the production of pulp and paper [1,2]. Nowadays, this process has been gradually replaced by thermomechanical (TMP) and chemithermomechanical (CTMP) pulping methods that gives higher pulp yields

and consume less water. The details of different pulping methods and pulping yields are mentioned in Table 1.1.

Table 1.1 Main wood pulping processes*

Pulping process	Chemical treatment	Mechanical treatment	Pulp Yield (%)
Mechanical process			
Stone groundwood (SGW)	-	Grinder	93-95
Pressure groundwood (PGW)	Steam	Grinder	93-95
Refiner mechanical (RMP)	-	Disc refiner	93-95
Thernomechanical (TMP)	Steam	Disc refiner	80-90
Chemithermomechanical (CTMP)	Steam+Na_2SO_3, NaOH	Disc refiner	80-90
Chemimechanical (CMP)	Na_2SO_3, NaOH	Disc refiner	80-90
Semimechanical process			
Neutral sulfite (NSSC)	Na_2SO_3 and Na_2CO_3 or $NaHCO_3$	Disc refiner	70-85
Green liquor (GLSC)	Na_2SO_3 and Na_2CO_3	Disc refiner	70-85
Non-sulfer	NaOH and Na_2CO_3	Disc refiner	70-85
Chemical process			
Kraft	NaOH and Na_2S	-	45-55
Soda	NaOH	-	40-50
Soda-anthraquinone	NaOH and anthraquinone	-	45-55
Soda-oxygen	NaOH and oxygen	-	45-55
Sulfite or bisulfate	$Ca(HSO_3)_2$, $NaHSO_3$, NH_4HSO_3 or $Mg(HSO_3)_2$ and H_2SO_4	-	45-55

* Data has been obtained from Baucher *et al*., 2003.

1.3 *Leucaena leucocephala* as source of pulp

Paper industry mainly uses bamboos, poplar, *Eucalyptus* sp. and *Casuarina* sp. as a source of raw material for paper and pulp [3]. Selection of the species depends upon availability, price and acceptability by any given industrial unit. Although all these plant species are of importance to paper industry, *Leucaena* sp. is extensively used in India and nearly a 25-30% of raw materials for paper industry are contributed by this plant (Fig. 1.1). *L. leucocephala*, a multipurpose nitrogen fixing tropical tree legume, also has a tremendous prospective as a raw material for nutritious forage, timber, organic fertilizer, firewood, industrial fuel and depilatory agent due to its adaptability to thrive under farthest agro-climatic conditions [3-6].

Fig. 1.1 (a) *Leucaena leucocephala* cv. K-636 growing at NCL premises, (b) A transgenic plant cultivated in our green house. (Photographs courtesy of Shakeel Abbassi, NCL)

It has been estimated that dried leaves of *Leucaena* contain nearly 28-35% of protein content of high nutritional quality. A native to Central America, *Leucaena* has been naturalized pan-tropically, with members of its genera being vigorous, drought tolerant, highly palatable, high yielding and grow in wide range of soils [3,5]. However, these attributes are limited by the occurrence of anti-nutritive factors in the fodder, such as tannins and mimosine [5,7,8].

1.4 Lignin: Occurrence, structure and function

Raw materials for several industries, viz. wood, agricultural residues, etc. consist largely of lignin and cellulose. As the second most abundant natural (terrestrial) organic polymer, eclipsed only by cellulose, lignin is a major constituent of wood and accounts, on an average, 25% of the terrestrial plant biomass [9-11]. The evolution of the ability to synthesize lignin is thought to be an important step in evolution of the land plants [12].

Lignins are complex, heterogeneous, three dimensional aromatic polymers synthesized from dehydrogenative polymerization of monolignols, namely p-coumaryl, coniferyl and sinapyl alcohol monomers differing in their degree of methoxylation [1,13,14]. These monolignols produce respectively, *p*-hydroxyphenyl (H), guaiacyl (G, more compact) and syringyl (S, less compact) phenylpropanoid units when incorporated into the lignin polymer (Fig. 1.2). Lignin is primarily synthesized and deposited in the secondary cell wall of specialized cells such as xylem vessels, tracheids and fibers. It is also deposited in minor amounts in the periderm where in association with suberin provides a protective role against pathogens. Both the lignin content and composition are known to depend on plant species, tissue type, subcellular location and also influenced by developmental and environmental cues [15]. For

example, while lignin from gymnosperm consists mainly of G units, angiosperm lignin is predominantly made up of both S and G units along with traces of H units [1]. Lignin content is higher in softwoods (27-33%) than in hardwood (18-25%) and grasses (17-24%). The highest amount of lignin (35-40%) occurs in compression wood on the lower part of branches and leaning stems [9].

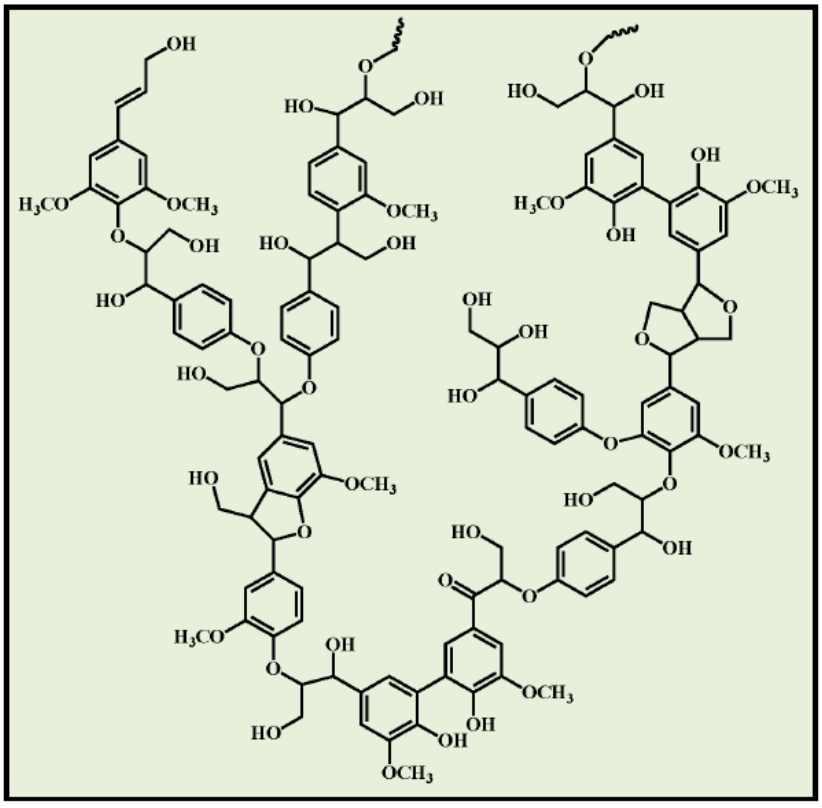


Fig. 1.2 Structure of lignin oligomer. These structures consist of H, G and S phenyl propanoid units. Number of such oligomers cross-polymerize to form a complex structure of lignin.

Lignin plays a major role in growth and development of plants such as rigidity and strength to cell wall, water/nutrient transport in conducting cells and also provides a physico-chemical barrier against pathogen attack [16,17,18]. However, an agro-economical opinion considers lignin as an obstacle for utilization of plant biomass

because it affects pulping and paper manufacture, forage digestibility and biofuel production processes [19-21]. The manufacture of high quality paper involves complete separation of lignin from polysaccharide component of wood during pulping and bleaching reactions [1]. These processes not only damage cellulose component but also consumes large quantities of energy and hazardous chemicals resulting in reduced pulp quality and paper strength as well as poor environmental image [19, 22-26]. Thus, reduced lignin content or altered quality of lignin in pulp wood species without compromising mechanical strength of plant is desirable. The proportion of S and G units in lignin determines the efficacy of paper production; that is, higher S/G ratio is enviable for paper as removal of S lignin (less compact) from pulp is more favorable than G lignin [27].

1.5 Lignin biosynthesis

Although lignin has been studied for a century, many aspects of its biosynthesis still remain unresolved. The monolignol biosynthetic pathway has been redrawn many times, yet still remains a matter of debate. During the last two decades, significant headway has been made in isolating and characterizing a number of genes pertaining monolignol biosynthesis from different plants. Several reviews on the advancements of monolignol biosynthesis pathways are available [10,11,16,28].

In plants, biosynthesis of monolignols is specialized branch of phenylpropanoid metabolism, a complex series of branching biochemical reactions responsible for synthesis of variety of products like lignin, flavonoids and hydoxycinnamic acid conjugates [10,29,30]. Many intermediates and end products of this pathway function in the production of aroma, fruit flavor, and as molecular signals, antimicrobial pigments, antioxidants and UV protectants [28, 31-33]. Finally, the diverse functions of

lignin and related products in resistance to biotic and abiotic stresses make this pathway vital to the health and survival of plants.

The synthesis of lignin represents one of the most energy demanding biosynthetic pathways in plants, requiring large quantities of carbon skeletons. Deposition of lignin in plants proceeds via the following steps:

1. The biosynthesis of monolignols
2. Transport of monolignols from the site of synthesis to the site of polymerization.
3. Dehydrogenation and polymerization of monolignols

1.5.1 Biosynthesis of monolignols

The biosynthesis of lignin precursors proceed through the phenylpropanoid pathway starting with deamination of phenylalanine to produce cinnamic acid and involves successive hydroxylation reactions of the aromatic ring, followed by phenolic o-methylation and conversion of the side chain carboxyl to an alcohol group. Several enzymes involved in lignin biosynthesis have been isolated and characterized. Monolignol biosynthesis is a complex pathway comprising of enzymes with functions like methyltransferase, hydroxylase, reductase and dehydrogenase. Highlights of some of these enzymes are given below (Fig. 1.3):

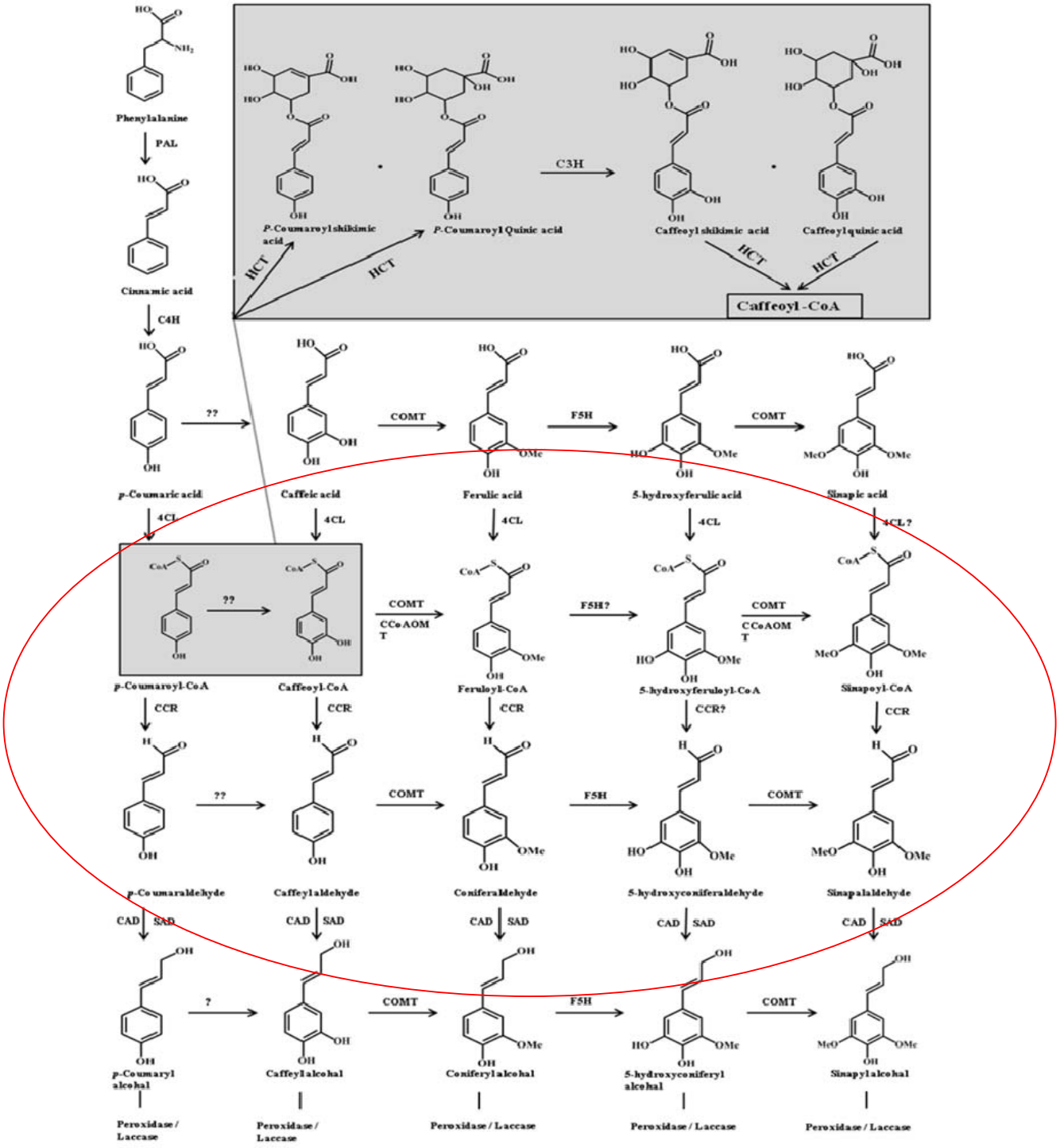


Fig. 1.3 An overview of monolignol biosynthesis pathway: PAL, Phenylalanine ammonia-lyase; C4H, Cinnamic acid 4-hydroxylase; C3H, *p*-Coumarate 3-hydroxylase;

HCT, Hydroxycinnamoyltransferase; 4CL, 4-Coumarate-CoA ligase; CCoAOMT, Caffeoyl-CoA O-methyltransferase; CCR, Cinnamoyl CoA reductase; CAld5H, Coniferyl aldehyde 5-hydroxylase; AldOMT, 5- hydroxyconiferyl aldehyde O-methyltransferase; CAD, Cinnamyl alcohol dehydrogenase; SAD, Sinapyl alcohol dehydrogenase.

1.5.1.1 Phenylalanine ammonia-lyase (PAL)

The enzyme phenylalanine ammonia-lyase (PAL; EC: 4.3.1.5) that catalyzes the conversion of phenylalanine to transcinnamic acid, is the initial step towards monolignol biosynthesis and other phenolic secondary plant metabolites. Genes encoding *PAL* have been studied in *Populus* species [34,35], loblolly pine and other plant species [36-41]. *PAL* exists as a multiple member gene family and the individual members can be involved in different metabolic pathways suggested by their expression patterns in association with certain secondary compounds accumulated in specific tissue or developmental stage. The biochemical activity of all known PALs is verified as a specific deamination of phenylalanine, but genetic and physiological function may vary among different PAL members. The expression of *PAL* genetic function is controlled by various genetic circuits and signaling pathways.

1.5.1.2 Cinnamate 4-Hydroxylase (C4H)

Hydroxylation of cinnamic acid to *p*-coumaric acid is catalyzed by cinnamate hydroxylase (C4H; EC: 1.14.13.11), a cytochrome P-450-linked monooxygenase belonging to the CYP73 family. Molecular oxygen is cleaved during this reaction, with one oxygen atom added to the aromatic ring and the other reduced to water. Similar to PAL, C4H is thought to be involved in a number of secondary metabolism pathways in

addition to monolignol biosynthesis, as *p*-coumarate is an intermediate for biosynthesis of many secondary compounds [42].

1.5.1.3 Coumarate 3-Hydroxylase (C3H)

Early biochemical evidence suggested that conversion of coumarate to caffeate is catalyzed by a nonspecific phenolase, but that suggestion did not receive much support in other studies [43-46]. The gene encoding *p*-coumarate 3-hydroxylase (C3H) was cloned and an alternative pathway was proposed based on the enzyme activity of CYP98A3 gene from *Arabidopsis* [47-49].

1.5.1.4 *p*-hydroxycinnamoyl-CoA: quinate shikimate *p*-hydroxycinnamoyl-transferase (HCT)

The enzyme *p*-coumarate 3-hydroxylase (C3H) converts *p*-coumaric acid into caffeic acid and has shown to be a cytochrome p450-dependent monooxygenase. It is interesting to note that enzymatic assays have demonstrated that the shikimate and quinate esters of *p*-coumaric acid are the preferred substrates for C3H over *p*-coumaric acid, *p*-coumaroyl CoA, *p*-coumaraldehyde, *p*-coumaryl alcohol, and not the 1-O-glucose ester or the 4-O-glucoside of *p*-coumaric acid [47-49]. *p*-coumarate is first converted to *p*-coumaroyl CoA by 4CL, with subsequent conversion to p-coumaroyl-shikimate and *p*-coumaroyl-quinate, the substrates for C3H, by *p*-hydroxycinnamoyl-CoA:quinate-(CQT) or *p*-hydroxycinnamoyl-CoA:shikimate *p*-hydroxycinnamoyltransferase (CST) [49]. These enzymes, described as reversible enzymes, can convert caffeoyl-shikimate or caffeoyl-quinate (chlorogenic acid) into caffeoyl CoA, the substrate for CCoAOMT. A reversible acyltransferase with both CQT and CST activity, designated HCT, has been purified and the corresponding gene cloned from tobacco [50]. Silencing of HCT through RNA interference (RNAi) led to

reduction in lignin, hyper accumulation of flavonoids and growth inhibition in *Arabidopsis* [51-53].

1.5.1.5 Coumarate Coenzyme-A ligase (4CL)

Genetic and biochemical functions of 4-Coumarate Coenzyme A ligase (4CL; EC: 6.2.1.12) genes have been clearly demonstrated in association with monolignol biosynthesis [14, 54-56]. 4CL catalyzes the formation of CoA thioesters of cinnamic acids in the biosynthesis of a wide variety of phenolic derivatives, including benzoic acid, condensed tannins, flavonoids and cinnamyl alcohols [57]. *4CL* genes usually exist as a multi-gene family. Four *4CL* genes were detected in the *Arabidopsis* genome and expression of each member was regulated differentially in tissues and developmental stages [58]. The enzymatic activities of 4CL members from aspen, loblolly pine, tobacco, soyabean, *Arabidopsis*, and many other species were found to have distinct substrate specificities [55,59,60,61]. As the 4CL catalytic kinetics vary among species, it is also likely that the mainstream pathway mediated by 4CL may not be the same in all plant species or tissues.

1.5.1.6 O-methyltransferases (O-MT)

S-adenosyl-L-methionine methyltransferases are key enzymes in the phenylpropanoid, flavanoid and many other metabolic pathways in plants. The enzymes Caffeate 3-O methyltransferase (COMT; EC: 2.1.1.68) and Caffeoyl CoA 3-O methyltransferase (CCoAOMT; EC: 2.1.1.104) control the degree of methoxylation in lignin precursors. The 3-carbon methylation leads to guaiacyl (G) unit formation and methylations on 3- and 5- position result in a syringyl (S) unit [11,62,63]. The methylation reactions at the C3 and C5 hydroxyl functions of the lignin precursors were thought to occur mainly at the cinnamic acid level via a bifunctional COMT. However, an enzyme distinct from

COMT, CCoAOMT has been identified in connection with the defense response in several dicot species [64-69]. Ye *et al*., (1994) proposed that CCoAOMT plays a role in methylation of both caffeoyl CoA and 5-hydroxyferuloyl CoA during monolignol biosynthesis.

1.5.1.7 Cinnamoyl CoA Reductase (CCR)

As mentioned earlier, the biosynthesis of lignin precursors proceed through common phenylpropanoid pathway, starting with conversion of phenylalanine to cinnamate and subsequent formation of hydroxycinnamoyl CoA esters via hydroxylation and methylation at different positions in aromatic ring. These esters are the end products of common phenylpropanoid metabolism and common precursors of a wide array of end-products such as flavonoids, coumarins and many small phenolic molecules [11]. The reduction of cinnamoyl CoA esters to cinnamaldehydes is the first metabolic step committed to monolignol formation [70]. In monolignol biosynthesis, cinnamoyl CoA reductase (CCR; EC: 1.2.1.44) catalyzes the NADPH dependent reduction of CoA esters to corresponding hydroxycinnamaldehydes and it controls the overall carbon flux towards lignin. Several *CCR* gene sequences from various plants have been deposited in the GenBank database, but their functions have still not been clearly demonstrated. It is proposed that all CCR enzymes have a similar catalysis mechanism for converting CoA esters to aldehyde in monolignol biosynthesis. A thorough literature survey of work done on CCR in the area of genetic engineering with lignin biosynthesis genes, current status of research in this area and need of more research on CCR has been discussed in further sections.

1.5.1.8 Coniferaldehyde 5-hydroxylase (CAld5H)

CAld5H enzyme like C4H belongs to cytochrome P450 monooxygenase family. The hydroxylation reaction in the biosynthesis of S-unit (syringyl) was first considered to occur at ferulate level [71], and hence, the enzyme was called ferulate 5-hydroxylase. However, later studies have revealed that F5H can also function at later steps in the pathway, mainly at the coniferyl aldehyde or coniferyl alcohol level [72,73]. This enzyme was therefore alternatively renamed as coniferaldehyde 5-hydroxylase [74]. F5H/CAld5H is a multifunctional plant P450 with three physiologically relevant substrates (coniferaldehyde, coniferyl alcohol and ferulate). Biochemical studies showed that the coniferaldehyde is the most preferred substrate for the enzyme [72].

1.5.1.9 Cinnamyl/sinapyl alcohol dehydrogenase (CAD/SAD)

In gymnosperm wood, coniferyl alcohol is the major monolignol units, while both coniferyl and sinapyl alcohols are present in angiosperm wood. Cinnamyl alcohol dehydrogenase (CAD; EC: 1.1.1.195) depicts a class of NADPH dependent oxidoreductase, suggested to catalyze multiple cinnamyl alcohol formations from their corresponding cinnamaldehydes [10,14,16]. This reduction of aldehydes to corresponding alcohols has been considered to be an indicator of lignin biosynthesis pathway [75,76]. When the *populus* tree was studied for monolignol biosynthesis in wood forming tissue, in addition to *CAD*; another gene whose sequence similar to *CAD*, but distinct from *CAD* was also found to be associated with lignin formation [77]. The biochemical characterization of the recombinant protein encoded by this gene indicated that the enzyme has specific affinity toward sinapaldehyde, therefore it was named as SAD. Compared to SAD, CAD showed a catalytic specificity towards coniferaldehyde instead. The catalytic specificities of the two enzymes have been

further verified in protein structure analysis [78]. Furthermore, it was demonstrated that the expression of CAD was associated with G-lignin accumulation while SAD was associated with S-lignin formation during xylem differentiation [77].

1.5.2. Transport of monolignols

After the synthesis, the lignin precursors or monolignols are transported to the cell wall where they are oxidized and polymerized. The monolignols formed are insoluble and toxic to the plant cell, and hence are converted to their respective glucosides by the action of UDP-glycosyltransferases (UDP-GT) [1,11]. This conversion renders the monolignols, soluble and less toxic to plant cells, which can be stored in plant vacuoles, and transported to the cell wall as need arises. It has been hypothesized that these monolignol glucosides are storage or transport forms of the monolignols [79].

1.5.3 Dehydrogenation and polymerization

After transport of the monolignols to the cell wall, lignin is formed through dehydrogenative polymerization of the monolignols [25]. The dehydrogenation of monolignol radicals has been attributed to different class of enzymes, such as peroxidases (POX), laccases (LAC), polyphenol oxidases, and coniferyl alcohol oxidase. Lignin is a hydrophobic, optically inactive, highly complex and heterogeneous polymer in nature. Lignin polymerization is a radical coupling reaction, where the monolignols are first activated into phenoxy radicals in an enzyme catalyzed dehydrogenation reaction. These radicals couple to form dimers, oligomers and eventually the lignin polymer [13]. Peroxidases are heme-containing oxidoreductases that use H_2O_2 as the ultimate electron acceptor. The natural electron donor molecules in a peroxidase catalyzed reaction vary and include, monolignols, hydroxycinnamic acids [80], tyrosine residues in extensions [81] and auxin [82]. Several reports on peroxidase

activity or gene expression in lignin-forming tissues have appeared, but only a few isoenzymes or genes have been specifically associated with lignifications [83-86].

1.6 CCR as a candidate gene

Among several genes involved in lignin biosynthesis, CCR plays a key regulatory role in lignin formation. Hydroxycinnamoyl CoA esters of general phenylpropanoid pathway when acted upon by CCR become destined to form respective monolignols. As the first committed step in monolignol biosynthesis, CCR diverts phenylpropanoid derived metabolites towards lignin.

1.6.1 Cloning and characterization of CCR genes from various plant species

CCR activity is found to be generally low in plants, so it is hypothesized that it may play a crucial role as a rate limiting step in regulation of lignin biosynthesis [87]. First cloning of a cDNA encoding CCR was achieved by screening a differentiating xylem cDNA library with degenerate oligonucleotides designed from internal peptide sequences, obtained from purified *Eucalyptus* CCR. The identity of the clone has been unambiguously proven through expression of a functional recombinant enzyme in *E. coli* [70]. CCR has been purified and partially characterized from *Forsythia* and soyabean cultures [88-90], spruce cambial sap [89], poplar xylem [91] and *Eucalyptus gunnii* xylem [92].

CCR genes have also been cloned and characterized from ryegrass [93], *Hordeum vulgare* and *Solanum tuberosum* [94], *Arabidopsis* [95], *Isatis indigotica* Fort [96], aspen [97], maize [98], *Medicago trancatula* [99], switchgrass [100] and Wheat [101]. CCR is apparently encoded by a single gene per haploid genome in *Eucalyptus* [70], poplar [102], ryegrass [93,103], *Triticum* [101] and tobacco [104], and by two genes in maize [98], *Arabidopsis* [105]. The CCR genes in various species also appear

as a multiple member family. In the *Populus* genome, there exist 8 CCR-homolog or CCR-like gene sequences [97]. *Triticum* [87,101], maize [98], switchgrass [100], *Medicago* [99] and *Arabidopsis* [105] have been shown to possess two or more than two isoforms (CCR1 and CCR2) which are involved in mutually exclusive or redundant functions like, constitutive lignifications or defense cascades.

1.6.2 Downregulation studies on CCR

Being entry point enzyme in lignin biosynthesis, CCR has proven to be a good target for down regulation to reduce the lignin content, to improve saccharification efficiency in bioenergy crops without compromising yields, to enhance the forage quality and produce optimal feedstock plants for biofuel production. The reduction in lignin content through down regulation of CCR has been observed in tobacco [106,107,108], *Arabidopsis* [109] and Norway spruce [110]. In poplar, CCR down regulation led to reduced levels of lignin and hemicelluloses with corresponding increase in cellulose level [102]; in alfalfa, the reduction in lignin levels in CCR down regulated transgenic plants was shown to be associated with improved cell wall saccharification efficiency [111]. In tomato, the lignin content was significantly reduced by down regulation of CCR gene along with enhanced availability of hydroxycinnamoyl CoA ester precursors, thereby stimulating the production of soluble, potentially health promoting phenolic compounds [112]. Recent studies in CCR deficient perennial ryegrass showed enhanced forage digestibility without detrimental effects on either plant fitness or biomass production [113]. All above studies suggest the importance of CCR as a candidate gene in lignin biosynthesis pathway.

1.6.3 Biochemical studies on CCR

CCR catalyzes the NADPH dependent reduction of cinnamoyl CoA esters to corresponding cinnamaldehydes. CCR exhibits the substrate specificity for various hydroxycinnamoyl CoA esters such as *p*-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA. CCR also carries out oxidation of hydroxycinnamaldehydes to their corresponding CoA esters (reverse reaction) [89,91,92] (Fig. 1.4). Differential substrate specificity of CCR suggests that there is possibility of differential activity of CCR in spatio-temporal manner in order to facilitate the formation of particular type of lignin in a particular tissue. This differential substrate specificity also suggests that there could be presence of more than one isoforms performing exclusive or redundant function inside the cell [97,99,100]. For example, PvCCR1 from switchgrass has a strong preference for feruloyl CoA indicating its role in lignin formation during plant development; while PvCCR2 shows coumaroyl CoA as favored substrate over other CoA esters suggesting that PvCCR2 could have a functional role in defense [100].

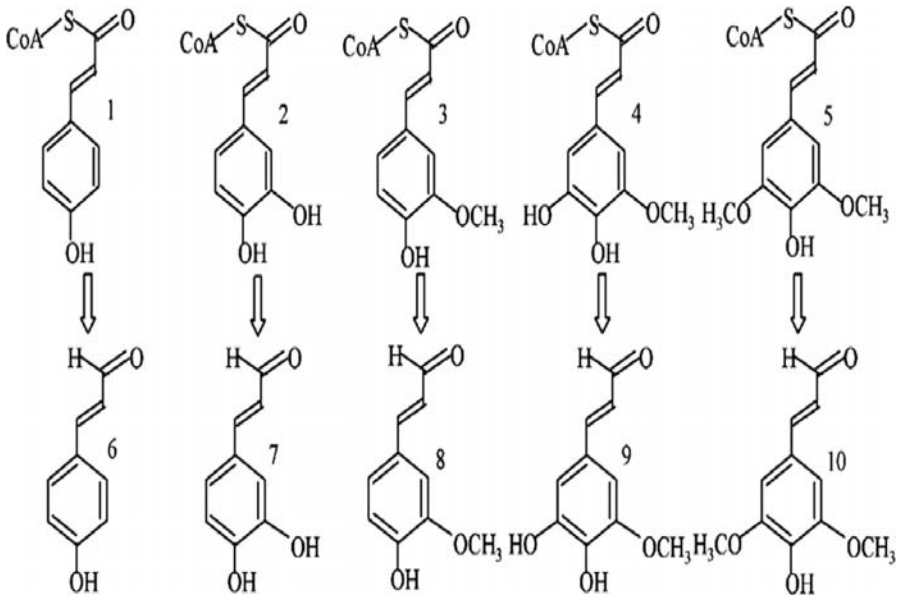


Fig. 1.4 Structures of hydroxycinnamoyl CoA esters and their corresponding cinnamldehydes. (1) *p*-coumaroyl-CoA, (2) caffeoyl-CoA, (3) feruloyl-CoA, (4) 5-hydroxyferuloyl-CoA, (5) sinapoyl-CoA, (6) *p*-coumaraldehyde, (7) caffealdehyde, (8) coniferaldehyde, (9) 5-hydroxyconiferaldehyde, (10) sinapaldehyde

Majority of biochemical studies on CCR from various plants showed the preference of enzyme for feruloyl CoA over other CoA esters. CCR from *Glycine max* exhibited high affinity for feruloyl CoA, 5-hydroxyferuloyl CoA and *p*-coumaroyl CoA with apparent K_m values in the order of 50 to 100 µM and the highest reaction rate was obtained with feruloyl CoA as substrate [90]. Also, feruloyl CoA was found to be preferred substrate under physiological conditions for the CCRs isolated from spruce (*Picea abies* L.) and soyabean (*Glycine max* L.) [89]. In poplar stems, cinnamoyl CoA reductase showed a decreasing affinity towards feruloyl CoA, sinapoyl CoA and *p*-coumaroyl CoA [91]. The substrate affinity data of CCR from *E. gunni* suggested that the enzyme did not prefer one cinnamoyl CoA ester over another. Although nearly identical K_m values were obtained for all CoA esters tested, approximately three times higher conversion rate was observed for feruloyl CoA than that for other substrates [92]. Kinetic parameters obtained for CCR from *Arabidopsis*, AtCCR1 were in the same order of magnitude and the enzyme showed almost equal affinity for feruloyl CoA and sinapoyl CoA over coumaroyl CoA [95]. In aspen, enzyme kinetics demonstrated that CCR selectively catalyzed the reduction of feruloyl CoA from a mixture of five cinnamoyl CoA esters. Furthermore, feruloyl CoA showed a strong competitive inhibition of the CCR catalysis of other cinnamoyl CoA esters [97]. The two CCR enzymes, LeCCR1 and LeCCR2 from *Solanum lycopersicum* (tomato) exhibited contrasting activities on the four cinnamoyl CoA substrates tested; while

LeCCR1 could use all substrates, LeCCR2 could only reduce feruloyl CoA [112]. Recombinant Ta-CCR1 protein from wheat converted feruloyl CoA, 5-OH-feruloyl CoA, sinapoyl CoA and caffeoyl CoA, but feruloyl CoA was the best substrate as evident from calculated catalytic efficiency [101]. These results are quite different with TaCCR2 from same plant, which can convert above substrates with almost similar efficiency [87]. Two cDNAs, PvCCR1 and PvCCR2, encoded enzymes with CCR activity in switchgrass. They are phylogenetically distinct, differentially expressed and the corresponding enzymes showed different biochemical properties with respect to substrate preference. PvCCR1 has higher specific activity and prefers feruloyl CoA as substrate, whereas PvCCR2 showed better affinity for caffeoyl and coumaroyl CoAs [100]

1.6.4 Structural studies on CCR

Three dimensional structures of enzymes provide valuable information to understand substrate specificity and to engineer the enzymes with novel specificity and/or improved catalytic function. Homology modeling and docking studies constitutes an alternative approach for the study of tertiary structures when enough data on crystal structures are not available. Although CCR is one of the most investigated enzyme of lignin biosynthesis pathway, its three dimensional crystal structure remains to be determined. The lack of information on three dimensional structure of CCR has precluded a clarification of functional active site residues involved in substrate binding and catalysis.

Surprisingly, structural studies on CCR, reported in literature are very few; in fact, only one molecular modeling and docking report is available on CCR from *L. leucocephala* [114]. In this report, the three dimensional model of cinnamoyl CoA

reductase was constructed based on the crystal structure of dihydroflavanol reductase from *Vitis vinifera* (PDB ID: 2c29) [115]. Substrate binding site pocket was found to be made up of 15 residues, that is, Asn52, Pro53, Asp54, Asp55, Ser56, Lys57, Asn58, Ser59, His60, Lys62, Glu63, Leu73, Gln206, Ser207 and Thr208; while NADPH binding site pocket composed with residues Asp77, Leu78, Leu79, Ser136, Tyr180, Pro197, Val198, Leu199 and Ser212. Furthermore docking studies were performed to understand the substrate interactions with active site of CCR. Docking simulations showed that residues Arg51, Asn52, Asp54 and Asn58 were involved in substrate binding. Arg51 in CCR was supposed to be the determinant residue in substrate specificity. Despite of all these observations, it is quite surprising that both substrate and cofactor binding site of CCR did not contain any of the amino acid residues in catalytic motif, NWYCYGK. Also little has been described about the significance of cofactor (NADPH) and its interaction with CCR enzyme. Furthermore, no information was provided about CCR catalysis mechanism and confirmatory biochemical evidences for active site residues. Thus, further detailed computational and biochemical investigations warrants a clear view of active site of CCR.

A very preliminary three dimensional structure of Ta-CCR2 from wheat was predicted using *Sporobolomyces salmonicolor* aldehyde reductase (PDB ID: 1UJM) as a template and shared only 20% similarity with the template [87]. In general, sequence similarity of 30% is enough to construct 3D model of target proteins through homology modeling. Here, very low similarity/identity of Ta-CCR2 with template urged the question on reliability of generated Ta-CCR2 model.

1.7 Protein folding

Protein molecules represent a remarkable relationship between structure and function at the molecular level. The surface pattern of proteins regarding their shape, charge and hydrophobicity determines their diverse and highly specific function. This surface pattern is determined by the unique three dimensional structure of the polypeptide chain. The process by which a linear polypeptide chain transforms into three dimensional structure is known as protein folding.

The mechanism governing protein folding remains a central point in biophysics and molecular biology. Each amino acid in the chain can be thought of having certain 'gross' chemical structures, which may be hydrophobic, hydrophilic and electrically charged. These amino acids interact with each other and their surroundings to produce a well-defined, three dimensional shape of the folded protein, known as native state. Protein folding is a topic of fundamental interest since it concerns the mechanism by which linear information of genetic message is transformed into three dimensional and functional structure of protein [116].

To be a biologically active, all proteins must adopt specific folded three dimensional structures. Till date, the genetic information for the protein specifies only the primary structure, the linear sequence of amino acids in the polypeptide chain. Most purified proteins can spontaneously refold in vitro after being completely unfolded, so the three dimensional structure must be determined by the primary structure. How this occurs has been known as 'protein folding problem' [117-120]. Protein folding problem involves a number of related questions such as (a) How does the given sequence find its specific native structure in a finite time among the astronomical number of possible conformations that a polypeptide could adopt? (b) How is the folding process initiated and what is (are) the pathway (s) of folding? (c)

What is the physical basis of the stability of the folded conformations? (d) Are the main rules of protein folding deduced from in vitro studies valid for folding in vivo? [121,122].

The process of protein folding begins from the unfolded state and proceeds to the native or misfolded states via various kinds of folding intermediates. Revealing the structural and dynamic properties of the unfolded states and folding intermediates at atomic level is crucial for understanding protein folding pathways [123].

1.7.1 Protein folding intermediates

A number of equilibrium and kinetic studies have led to structural characterization of folding/unfolding intermediates, a prerequisite to solve the folding problem. Partially folded states are characterized at equilibrium under mild denaturing conditions, such as by altering pH, addition of salts, chemical denaturants (urea and guanidine hydrochloride) or by changing temperature and pressure. Equilibrium intermediates characterized in different proteins were found to be related to kinetic folding intermediate transiently populated in early phase of folding reaction. This partially folded state was termed as ‘molten globule’ since it had shape with loosely collapsed hydrophobic core [124].

Molten globule is a compact intermediate with high content of native-like secondary structure but altered tertiary structure [125-127]. It contains accessible hydrophobic surfaces which bind to a hydrophobic dye, 1-anilinonapthalene sulfonate (ANS). The absence of near UV circular dichroism spectrum shows that aromatic residues in intermediate can rotate in a symmetrical environment. Sub-millisecond kinetic methods have improved resolution time of kinetics studies allowing detection of early events of folding.

Table 1.2 experimental techniques used to monitor protein folding

Techniques	Information about folding process
Protein engineering	Role of individual amino acids in stability
Laser Scattering	Radius of gyration
Gel filtration chromatography	Radius of gyration
DSC	Thermodynamics of folding process
Fluorescence spectroscopy	
1. Intrinsic	Environment and orientation of Trp
2. Polarization and anisotropy	Dynamics of fluorophore
3. FRET	Distance between two points in a protein
4. Quenching	Accessibility and environment of fluorophore
5. REES	Difference between environment fluorophore
6. Stopped flow	Time scale of fluorescence changes
7. Substrate binding	Formation of the active site
8. ANS binding	Exposure of hydrophobic area
Circular Dichroism (CD)	
1. Far UV	Secondary structure
2. Near UV	Tertiary structure
3. Stopped flow	Time for secondary and tertiary structure formation
Small angle X-ray scattering	Dimension and shape of polypeptide
FTIR	Secondary structure formation
Hydrogen exchange (HX)	
1. Native state	Detection of metastable state
2. Pulsed HX NMR	Rate of formation of backbone hydrogen bonds in specific amino acids
3. Pulsed HX ESI MS	Folding population determination
NMR	
1. Real time	Environment of protein side chain
2. Dynamic NMR	Detect equilibrium species
Laser temperature jump	Trigger folding/unfolding at nanoseconds
Force spectroscopy (AFM or optical tweezers)	Unfolding forces and unfolding rate constants of single molecules.

More recently, techniques, such as hydrogen exchange NMR, solution X-ray scattering, protein engineering and site directed mutagenesis have provided detailed picture of molten globule of many proteins. In the last decade, a wide range of physical and chemical methods complemented the fundamental techniques to study protein folding. A brief summary of experimental techniques used in characterization of protein folding intermediates have been described in Table 1.2.

The ultimate objective of such studies is to define complete energy landscape for the folding reaction and to understand in detail how this is defined by sequence. Various studies have shown that molten globule (MG) state has heterogeneous structure in which one portion of molecule is more organized and native like; while other portions being less organized [128]. How much native-like structure is present depends on protein species and solution conditions, and a remarkable diversity in molten-globule structure of different proteins has been observed [129-133]. Thus it will be useful to describe MG state of different globular proteins to investigate the common as well as different features among the different proteins.

1.7.2 Protein folding and Biotechnology

An intense research is focused to understand the structural basis of protein folding and stability, and mechanistic role of early folding intermediates. Finding a solution to protein folding problem has many practical applications.

1.7.2.1 Predicting structure of proteins

A thorough understanding of the physical principles that govern protein folding is required for rational efforts to predict three dimensional structures of proteins. Our understanding of protein folding has been greatly advanced through computational models and simulations. The methodologies used can be divided into three general

categories: comparative modelling, Fold recognition or threading and ab-initio methods. These methods involve the development of an energy function capable of identifying the most stable conformation of a protein and a scoring function for the evaluation of protein models. These methods also provide information complementary to that obtained from experiments. In particular, molecular dynamics and simulation studies can help to identify or predict the transition or intermediate states along the folding pathway, provide predictions of the rate of folding and in some cases, predict the final, folded structures.

1.7.2.2 Solving the protein aggregation problem

Native state of proteins always represents thermodynamically the most stable conformation under physiological conditions [134]. All the information regarding the native structure is hidden in the amino acid sequence. However, correct folding is a challenge for proteins in a living cell and only a part of the protein can assume their native structure spontaneously. Efficient protein folding and transport depend on the presence of a complex machinery of chaperones, chaperonins and cofactors. The primary mission of this machinery is to prevent the aggregation of nascent polypeptide chains and proteins that unfold upon environment stress [135].

Proteins and peptides carry out a multitude of functions required for survival and sustenance of life with the help of complex and usually transient networks of intermolecular interactions. Establishment of protein intermolecular contacts is not always beneficial for the fitness of organism, as can be seen from the increasing evidences of human pathologies, like Alzheimer's and Parkinson's diseases, which are associated with the formation of abnormal interactions between adjacent protein molecules. Such aberrant interactions then result in misfolding or misbinding events in

proteins or protein complexes, thus exposing previously hidden regions that can now establish specific, but unwanted contacts. These contacts in most cases cause self-assembly and formation of large and insoluble structures. The abnormal association of misfolded proteins usually leads to the formation of aggregates. Small aggregates can remain soluble, but large protein aggregates precipitate out of solution under physiological conditions.

Protein aggregation in vivo is a widespread phenomenon that arises from early folding intermediates through kinetic competition between proper folding and misfolding. A wide range of human diseases resulting from incorrect folding or altered conformation of proteins have been found (Table 1.3). Possible treatment of these diseases could exploit detailed knowledge of protein folding and the prevention of abnormal folding.

1.7.2.3 De novo designing

An important field of research dependent upon our understanding of protein folding is de novo protein designing for constructing completely new positions with determined function. Grado *et al*., have reviewed the different methods for de novo protein design, which usually consist of choosing a function of protein followed by searching a protein's scaffold capable of supporting the reactive groups in desired geometry [136]. It is then necessary to determine an amino acid sequence capable of folding into an adequate and stable three dimensional structure. Exploring genetic methods or combinatorial or computational algorithms may help in this. It is now evident that de novo protein design represents a growing field of research that will be useful both in testing the principles of protein folding and in offering the perspectives to design new proteins with practical applications for pharmaceuticals and diagnostics.

Table 1.3 Some protein misfolding diseases

Diseases	Protein involved	Cause
Crueltzfeldt-jacob	Prion protein	Toxic folding/ aggregation
Alzheimer's	Beta-amyloid	Toxic folding/ aggregation
Cystic fibrosis	CFTR	Misfolding
Cancer	Protein 53 (transcription factor)	Misfolding
Type II diabetes	Islet amyloid polypeptide	Amyloid fibril formation
Familial amyloid polyneuropathy	Apolipoprotein	Aggregation
Familial visceral amyloidosis	Lysozyme	Aggregation
Parkinson's	Alpha-synuclein	Amyloid fibril formation
Huntington's	Huntington	Amyloid fibril formation
Marfan syndrome	Fibrillin	Misfolding
Cataract	Crystallins	Aggregation
Scurvey	Collagen	Misfolding
Osteigenesis imperfect	Type I procollagen	Misassembly
Kelanddic cerebral angiopathy	Cystatin C	Amyloid fibril formation
Amyotropic lateral sclerosis	Superoxide dismutase	misfolding
Finish type familial amyloidosis	Gesolin	Amyloid fibril formation

Rationale of the thesis

This chapter gives brief information on lignin, its biosynthesis and role of *L. leucocephala* in paper and pulp industry in India. Background of research done on Cinnamoyl CoA reductase (CCR), a key regulatory enzyme in lignin biosynthesis has been dealt in detail. A thorough literature survey of work done on CCR in the area of genetic engineering with lignin biosynthesis pathway genes with regards to the current status of research in this area has been presented. Structure-Functional studies on CCR from various plants have been discussed in detail and need of more research on CCR

has been elaborated. Finally the importance of protein folding and bioinformatics also has been discussed. Our group at the CSIR-National Chemical Laboratory (Pune, India) has been working on genes involved in lignin biosynthesis pathway in *L. leucocephala*, for the past 10 years [137]. *L. leucocephala*; being one of the important resources for paper and pulp production in India, yet so far not much study has been done on structure-functional aspects of enzymes taking part in lignin formation in this species. Although CCR is one of the most investigated enzyme of lignin biosynthesis pathway, its structure-function relationship remains to be determined. In this context, the present study was aimed at understanding the structure-functional aspects of cinnamoyl CoA reductase from *L. leucocephala* (Ll-CCRH1).

As a first step towards understanding the structure-function relationship of Ll-CCRH1 (GenBank: DQ986907), we report here the detailed biochemical characterization (expression, purification, pH/temperature stability, substrate kinetics, activation energies determination, pH activity profile, effect of metal ions, small angle X-ray scattering studies and dynamic light scattering studies) of recombinant Ll-CCRH1 by various biophysical and biochemical techniques. Homology modelling was carried out to predict three dimensional structure and catalytic active site residues. Furthermore, docking simulations were performed to study interactions of active site residues with different hydroxycinnamoyl CoA esters to shed light on multiple substrate specificity to understand mechanism of catalysis. Active site residues and their roles in binding or catalysis obtained from *in silico* studies were validated and further investigated by site directed mutagenesis and chemical modification studies. Also, *in silico* mutagenesis and docking studies of active site residues of Ll-CCRH1 were performed to study altered substrate specificity and possible physiological role of Ll-CCRH1 either in lignifications or in defence system in plants. Conformational

transitions in Ll-CCRH1 under the effect of pH, temperature and guanidine hydrochloride were studied by biophysical methods like circular dichroism and fluorescence spectroscopy. The tryptophan environment of wild type recombinant Ll-CCRH1 and active site mutant enzymes was investigated with solute quenching studies.

CHAPTER 2

Biochemical characterization of recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*

Prashant Sonawane, Rishi Kishore Vishwakarma, Bashir M. Khan (2013) Biochemical characterization of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, ***Int J Biol Macromol***, 58, 154-159.

Summary

Recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) protein from *Leucaena leucocephala* was overexpressed in *E. coli* BL21 (DE3) strain and purified to apparent homogeneity. Optimum pH for forward and reverse reaction was found to be 6.5 and 7.8 respectively. The enzyme was most stable around pH 6.5 at 25 °C for 90 min. The enzyme showed K_{cat}/K_m for feruloyl, caffeoyl, sinapoyl, coumaroyl CoA, coniferaldehyde and sinapaldehyde as 4.6, 2.4, 2.3, 1.7, 1.9 and 1.2 (x 10^6 M^{-1} s^{-1}), respectively, indicating affinity of enzyme for feruloyl CoA over other substrates and preference of reduction reaction over oxidation. Activation energy, E_a for various substrates was found to be in the range of 20-50 kJ/mol. Involvement of probable carboxylate ion, histidine, lysine or tyrosine at the active site of enzyme was predicted by pH activity profile. SAXS studies of protein showed 3.04 nm radius and 49.25 nm^3 volume with oblate ellipsoid shape. Metal ion inhibition studies revealed that Ll-CCRH1 is a metal independent enzyme.

2.1 Introduction

Lignins are complex, heterogeneous, three dimensional aromatic polymers derived from oxidative polymerization of three monolignols namely, *p*-coumaryl alcohol (H unit), coniferyl alcohol (G unit, more compact) and sinapyl alcohol (S unit, less compact), differing in their degree of methoxylation [11]. Besides its critical role in normal plant growth and development, an agro-economical opinion considers lignin as an obstacle for utilization of plant biomass because it affects paper manufacture and limits digestibility of forage crops [19]. In plants, biosynthesis of lignin precursors proceeds through common phenylpropanoid pathway [30]. In monolignol biosynthesis,

cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the first committed step and is considered as a first regulatory point in lignin formation [70].

CCR catalyzes the NADPH dependent reduction of cinnamoyl CoA esters to corresponding hydroxycinnamaldehydes. CCR exhibits the substrate specificity for various hydroxycinnamoyl CoA esters such as *p*-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA [89-91]. CCR also carries out oxidation of hydroxycinnamaldehydes to their corresponding CoA esters [89,92].

Leucaena leucocephala, a multipurpose nitrogen fixing tropical legume, has a tremendous prospective as a raw material (30-40%) for paper and pulp industry in India [27]. Surprisingly, structural studies on CCR reported in literature are very few; in fact only one molecular modeling and docking report is available on CCR from *L. leucocephala* till date [114]. Biochemical characterization of CCR from wheat, switchgrass and *Arabidopsis* provides the better understanding about its key role in lignin biosynthesis. [87,95,100]. A three dimensional crystal structure of CCR remains to be determined. As a first step towards understanding the structure-function relationship, we report biochemical characterization of recombinant Ll-CCRH1 protein by various biophysical and biochemical techniques. To the present author's knowledge, this is the first detailed report on the biochemical characterization of recombinant Ll-CCRH1 from *L. leucocephala*.

2.2 Materials and Methods

Materials

Various cinnamoyl CoA esters (feruloyl CoA, caffeoyl CoA, *p*-coumaroyl CoA and sinapoyl CoA) were obtained from TransMIT GmbH, Giessen, Germany.

2.2.1 Heterologous expression and purification of recombinant Ll-CCRH1 from *L. leucocephala*

The full length Ll-CCRH1 gene was directionally cloned into expression vector pET30b (+) by digesting the vector and Ll-CCRH1 gene (cloned in pGEMT Easy Vector) with restriction enzymes *Nde* I and *Xho* I, and ligated together. The recombinant plasmids were transformed into *E. coli* strain BL21 (DE3) competent cells. Colonies for recombinant plasmids were screened by colony PCR method. *E. coli* cells harboring recombinant pET30b (+) vector were grown for 12 h in 5 ml LB-kanamycin (50 µg/ml) at 37 °C and 200 rpm. This culture was used to inoculate 200 ml of LB-kanamycin (50 µg/ml), which was incubated until A_{600} reached 0.6-0.8. Production of recombinant Ll-CCRH1 was then induced by addition of 0.5 mM IPTG followed by incubation at 20 °C for 14 h. Induced *E. coli* cells were harvested by centrifugation at 7000 g at 4 °C for 20 min and stored at -20 °C.

Pellets of induced bacteria were resuspended in 50 mM Tris-HCl, pH 8.0 containing 500 mM NaCl, 10% glycerol, 10 mM β-Mercaptoethanol, 20 mM imidazole, and 1% (v/v) Triton X-100. This suspension was sonnicated for 3 min with pulses of 15 s on/off at 70% amplitude (Misonix, USA). Lysozyme (0.5 mg/ml) was added after sonication and kept at 4 °C for 30 min. Finally cell debris was harvested by centrifugation at 10,000 g for 20 min, and the supernatant (lysate) was filtered through 0.22µm Millex filters (Millipore, USA). Purification of recombinant Ll-CCRH1 was performed at 4 °C by FPLC (AKTA Explorer, GE Healthcare, Sweden) as follows: 1. The lysate was loaded on Ni Sepharose 6 Fast Flow column previously equilibrated with 50 mM Tris-HCl, 500 mM NaCl and 30 mM imidazole. Non retained proteins were washed with 5-10 column volumes (CV) with wash buffer containing 40 mM imidazole. Proteins specifically retained in resin were eluted with elution buffer having

150 mM imidazole (flow rate 1 ml/min). Fractions collected from Ni-Sepharose were pooled, concentrated and desalted on Sephadex G-25 column with 20 mM Tris-HCl, pH 8.0 buffer. 2. The desalted fractions of Ll-CCRH1 were then applied to DEAE Sepharose Fast Flow column, equilibrated with same buffer. The proteins were eluted with linear gradient (10 CV) from 0-0.5 M NaCl at a flow rate of 0.5 ml/min. Eluted fractions were then concentrated through Amicon Ultra-15 with 10 kDa molecular weight cut off. 3. The preparation from step 2 was finally applied to Hi Load 16/60 Superdex 75 Pg and protein was eluted with 20 mM Tris-HCl, pH 8.0 buffer containing 150 mM NaCl (flow rate 0.3 ml/min). Eluted fractions were checked for Ll-CCRH1 activity and those showing activity were pooled and analyzed by SDS-PAGE followed by silver staining. Protein concentration was determined by Bradford assay (Bradford reagent, Bio-Rad, USA).

2.2.2 Ll-CCRH1 identification by ESI-MS/MS and Western blot analysis

Ll-CCRH1 protein spots were excised from SDS-PAGE gels and subjected to in situ trypsin digestion [138]. The digested peptides with a final concentration of 100 ng/µl were analyzed by using nanoACQUITY UPLC online coupled to a SYNAPT HDMS system (MS^E) (Waters Corporation, Milford, MA) as described by Cheng *et al*. (2009) [139]. After MS^E analysis, data were analyzed with Protein Lynx Global Server software (PLGS version 2.4 Waters Corporation, Milford, MA). For Western blot, the 6X-His tagged Ll-CCRH1 protein was run on 10% SDS-PAGE, transferred onto PVDF membrane and confirmed by detection with Anti-His antibodies. Molecular mass of native Ll-CCRH1 protein was determined by MALDI-TOF/TOF mass spectrometer.

2.2.3 Enzyme assay

The reaction mixture contained 100 mM phosphate buffer pH 6.5, 100 µM Cinnamoyl CoA ester, 70 µM NADPH and Ll-CCRH1 protein (0.131 µM) in a total volume of 1000 µl at 30 °C. Reduction was determined by the decrease in absorbance at 366 nm. Ll-CCRH1 activity was calculated by linear decrease in A_{366} and correction factors given by Luderitz and Grisebach [89,90].

For the reverse reaction, the oxidation of cinnamaldehyde was estimated by an increase in absorbance at 366 nm using similar types of calculations given above. Here, the incubation mixture consisted 100 mM Tris-HCl pH 7.8, 100 µM cinnamaldehyde, 150 µM $NADP^+$, 150 µM CoA and Ll-CCRH1 protein (0.131 µM) in 1000 µl reaction.

2.2.4 Optimum pH and pH stability

Optimum pH for both forward and reverse reaction was determined by assaying Ll-CCRH1 (5-10 µg) at different pH, ranging from 2-12. For pH stability studies, 150-200 µg enzyme was incubated in buffers of pH, ranging from 2-12. Aliquots were taken out every 15 min and assayed for Ll-CCRH1 activity.

2.2.5 Temperature optimum and stability

For determination of optimum temperature, 10 µg enzyme was assayed at different temperature (10-80 ^oC) for Ll-CCRH1 activity. For temperature stability determination, 150-200 µg enzyme was incubated at different temperatures (10-80 °C). Aliquots were taken at suitable time intervals and residual activity was determined.

2.2.6 Substrate kinetics studies

K_m and V_{max} of Ll-CCRH1 for various substrates (feruloyl CoA, caffeoyl CoA, coumaroyl CoA and sinapoyl CoA for forward reaction; coniferaldehyde and sinapaldehyde for reverse assay) were determined by Michaelis-Menten method by

varying the substrate concentrations. The activity of the enzyme was expressed as units/mg/min.

2.2.7 Activation energies determination

K_m and V_{max} of Ll-CCRH1 for various substrates were determined at different temperatures ranging from 10 to 40 °C. The activation energy E_a, with each substrate was calculated from the slope of plot of the ln V_{max} Vs 1000/T, as E_a = -slope X R (R, gas constant= 8.314 X10^{-3} kJ/mol). ΔG was calculated by using the equation ΔG= RT ln K_d (here, the values of K_m were used for dissociation constant K_d) [140,141].

2.2.8 pH activity profile

K_m and V_{max} of Ll-CCRH1 at different pH (4-10) were determined under standard assay conditions. Ln (V_{max} / K_m) Vs pH was plotted to determine pK_a values.

2.2.9 Effect of metal ions and detergents

Ll-CCRH1 enzyme (0.131 µM) was incubated with different metal ions (1 mM, 2 mM concentration) and detergents (1%) in 100 mM phosphate buffer pH 6.5 for 5-10 min at 25 °C, and activity was determined as described above. Reversibility of metal ion inhibition was checked with various concentrations of EDTA.

2.2.10 Dynamic light scattering (DLS) studies

The protein solution (250 µg/ml) incubated at 28 °C was subjected to particle size analysis at 28 °C on a Brookhaven 90 plus particle size analyzer.

2.2.11 Small angle X-ray scattering (SAXS) studies

SAXS measurements were performed on a Bruker Nanostar equipped with a rotating anode generator. The X-rays are collimated through a three-pinhole system, and data is acquired using a 2D gas filled Hi-Star detector over a *q*-range of 0.011-0.2 Å^{-1}. The detector was calibrated using a silver behenate sample. Samples were sealed in quartz

capillaries having an outer diameter of 2 mm and wall thickness of ≈10 μm. 0.28 % Ll-CCRH1 protein in Tris-HCl, pH 8.0 was used for study. Data was corrected for background scattering (including solvent scattering and air scattering). The 2D data were reduced to 1D by circularly averaging, using the software provided with the instrument [142].

2.3 Results and discussion

2.3.1 Purification of recombinant Ll-CCRH1 protein

The recombinant Ll-CCRH1 cloned in pET 30b (+) plasmid was overexpressed in *E. coli* BL21 (DE3). The SDS-PAGE analysis showed that the soluble and pellet fraction of recombinant bacteria contained an overexpressed protein with an apparent molecular mass of 38 kDa (Fig. 2.1).

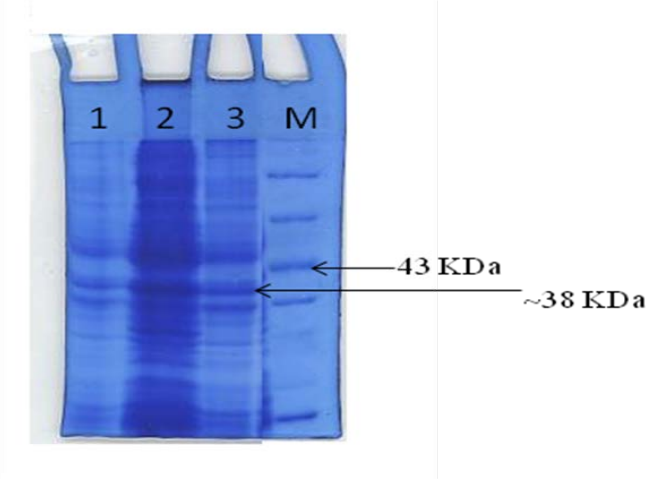


Fig. 2.1 Over expression of Ll-CCRH1 protein; M: Protein Molecular weight marker, Lane 1: pET 30b (+) vector without insert; Lane 2: Induced Ll-CCRH1 protein; Lane 3: Uninduced Ll-CCRH1. Lane 1-3 is the total bacterial protein obtained under each growing condition respectively.

An array of purification techniques involving ammonium sulfate fractionation, different chromatographic techniques such as affinity, ion exchange, hydrophobic interaction and gel filtration were applied to purify recombinant Ll-CCRH1. However, a simple and efficient three step protocol comprising of an affinity chromatography followed by anion-exchange chromatography and final gel filtration was found to be best suited for Ll-CCRH1, hence this protocol was optimized for purification of the protein from the cell lysate (Fig. 2.2 to Fig. 2.5).

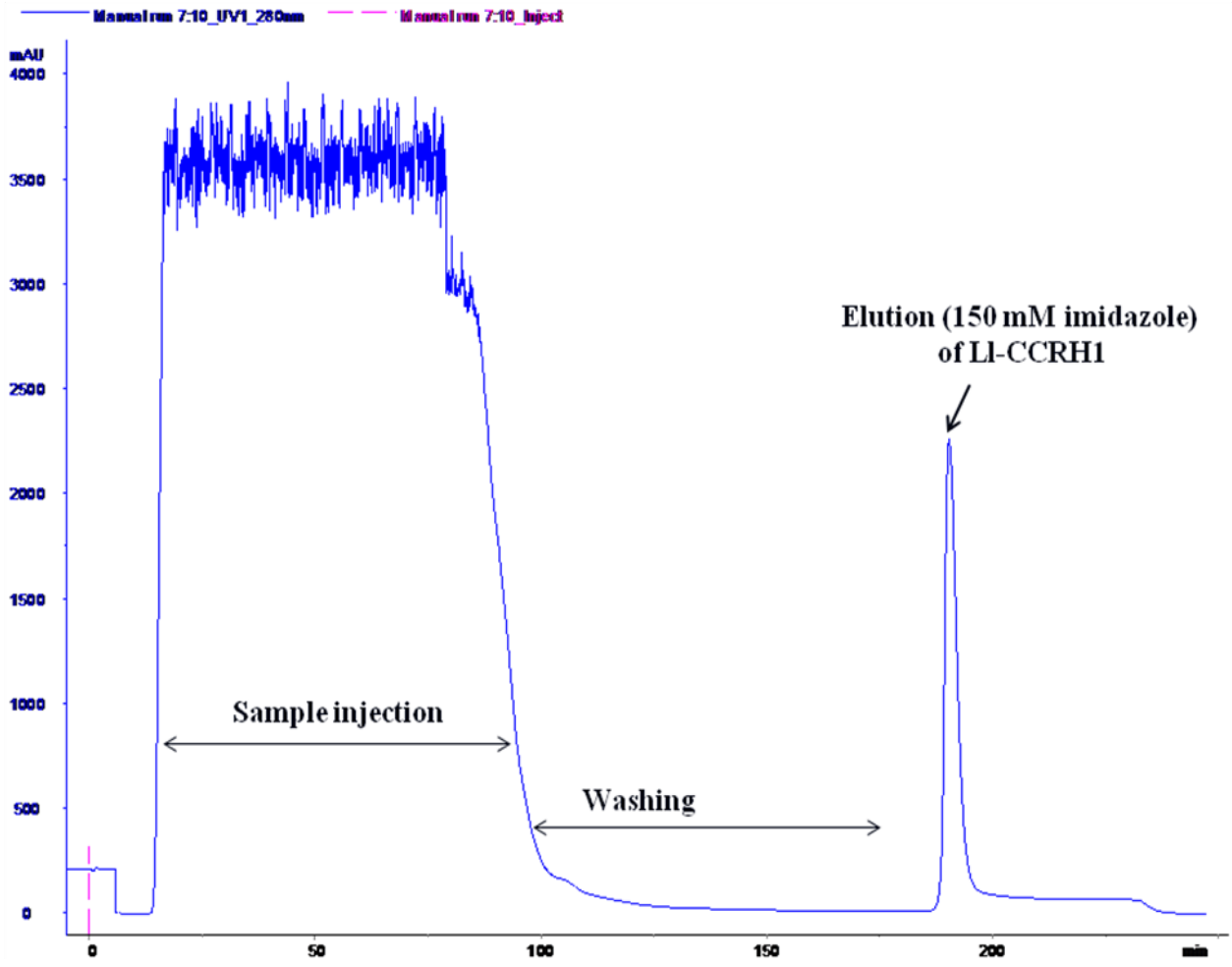


Fig. 2.2 Affinity chromatogram showing first step purification of Ll-CCRH1.

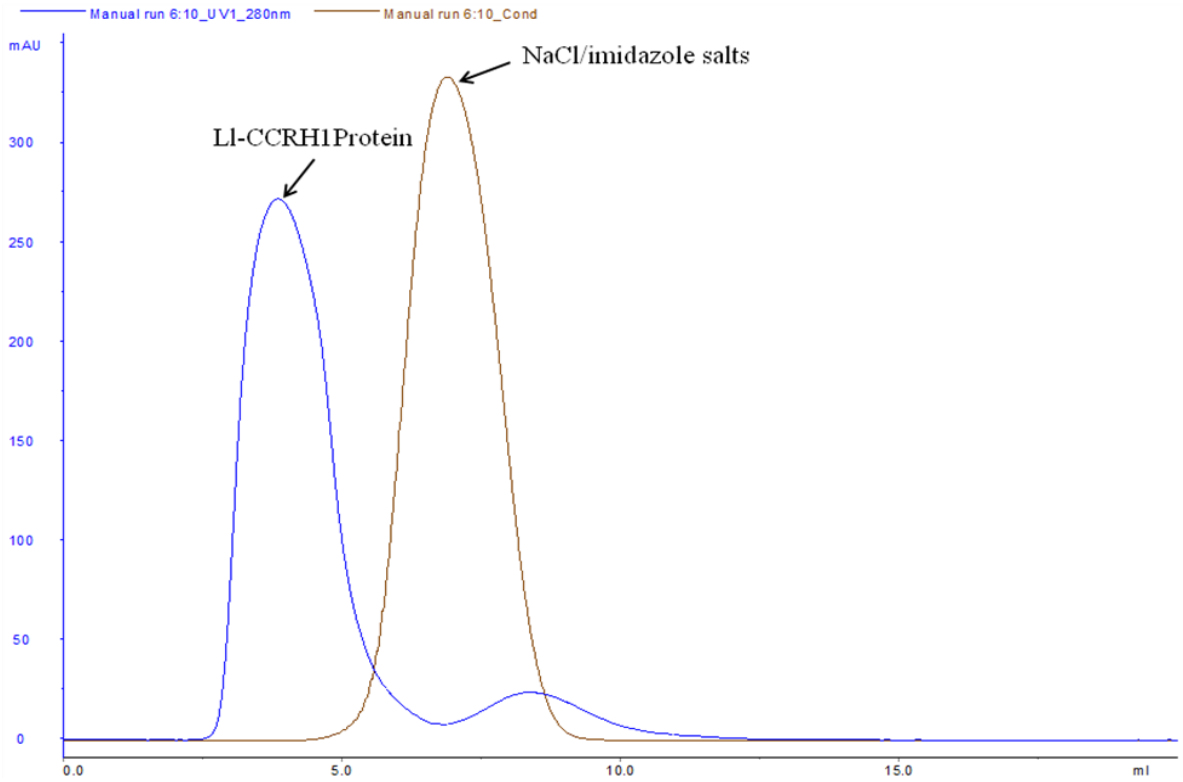


Fig. 2.3 Chromatogram for desalting of Ll-CCRH1 (HiPrep Sephadex G-25).

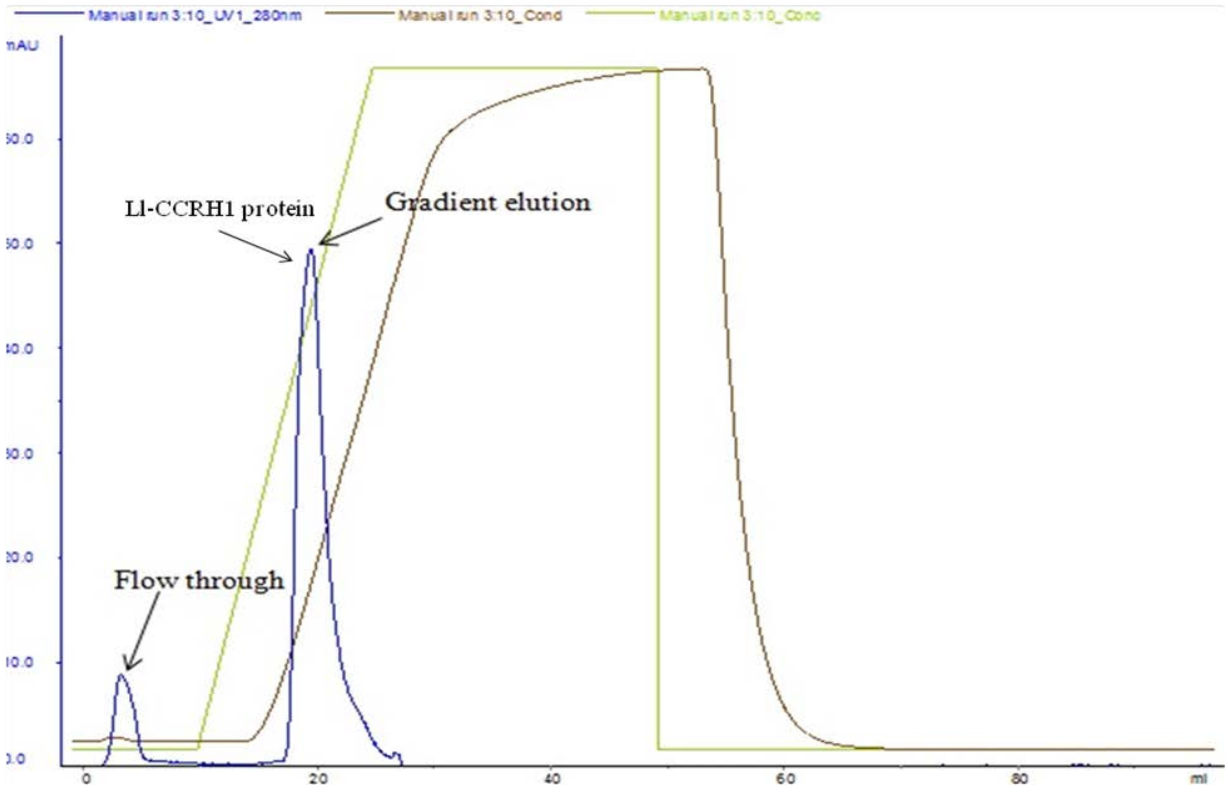


Fig. 2.4 Anion exchange chromatogram of Ll-CCRH1 using DEAE-Sepharose.

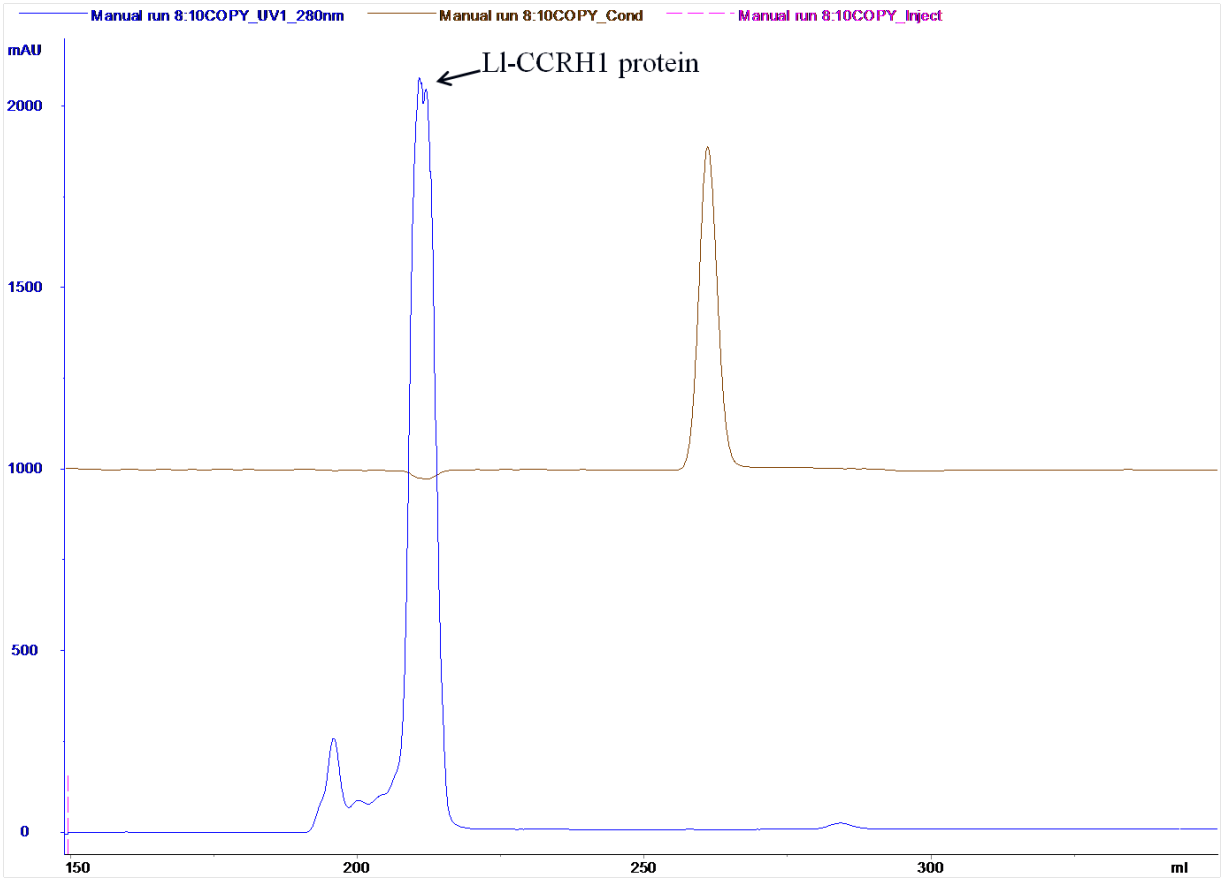


Fig. 2.5 Gel filtration chromatogram of Ll-CCRH1.

The results of the purification procedure for the recombinant Ll-CCRH1 are summarized (Table 2.1). The final purification step resulted in 53 fold purification with specific activity of 1.2 U/mg and relatively good recovery of 66%. Fig. 2.6 shows SDS–PAGE analysis of fractions obtained at each purification step and showed a single purified band of Ll-CCRH1 after gel filtration chromatography.

Table 2.1: Purification procedure for recombinant cinnamoyl CoA reductase 1 (Ll-CCRH1) expressed in *E.coli* BL21 (DE3). The starting soluble crude lysate was obtained from1L culture broth.

Purification step	Total volume (ml)	Total protein (mg)	Total activity (U)	Specific activity (U/mg)	Fold Purification	Recovery (%)
Crude lysate	80	1600	36	0.023	-	100
Ni Sepharose 6 FF	20	40	30.6	0.765	34	85
DEAE Sepharose FF	12	27	27	1.0	44	75
Superdex 75 Pg	8	20	24	1.2	53	66

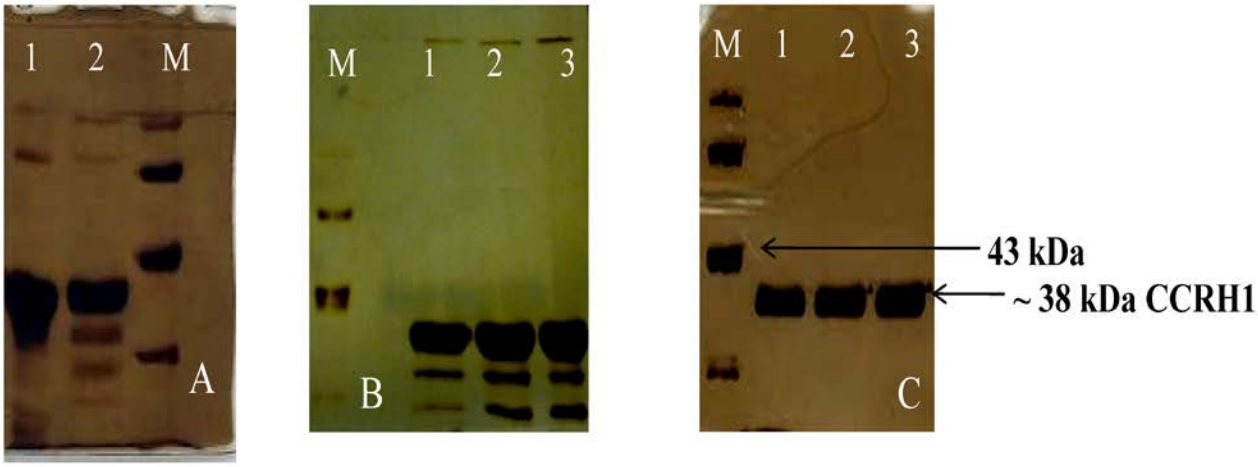


Fig. 2.6 SDS–PAGE analysis of fractions obtained at each purification step of overexpressed recombinant Ll-CCRH1 protein. (A), (B), (C) represent affinity, ion exchange and gel filtration purified Ll-CCRH1 protein respectively; Lane M: Protein Molecular weight marker (97, 66, 43, 29 kDa respectively starting from top), Lane 1-2 (A) and Lane 1-3 (B,C): fractions collected from respective chromatography purification.

2.3.2 ESI-MS/MS, Western blot, MALDI-TOF-MS and DLS analysis of Ll-CCRH1

Detection and purification of Ll-CCRH1 protein was monitored by ESI-MS/MS and western blot analysis (Fig. 2.7 and 2.8). ESI-MS analysis of Ll-CCRH1 showed 97% coverage only with our reported CCR sequence from *L. leucocephala* (GenBank protein: ABL01801], indicating that the expressed protein is Ll-CCRH1.

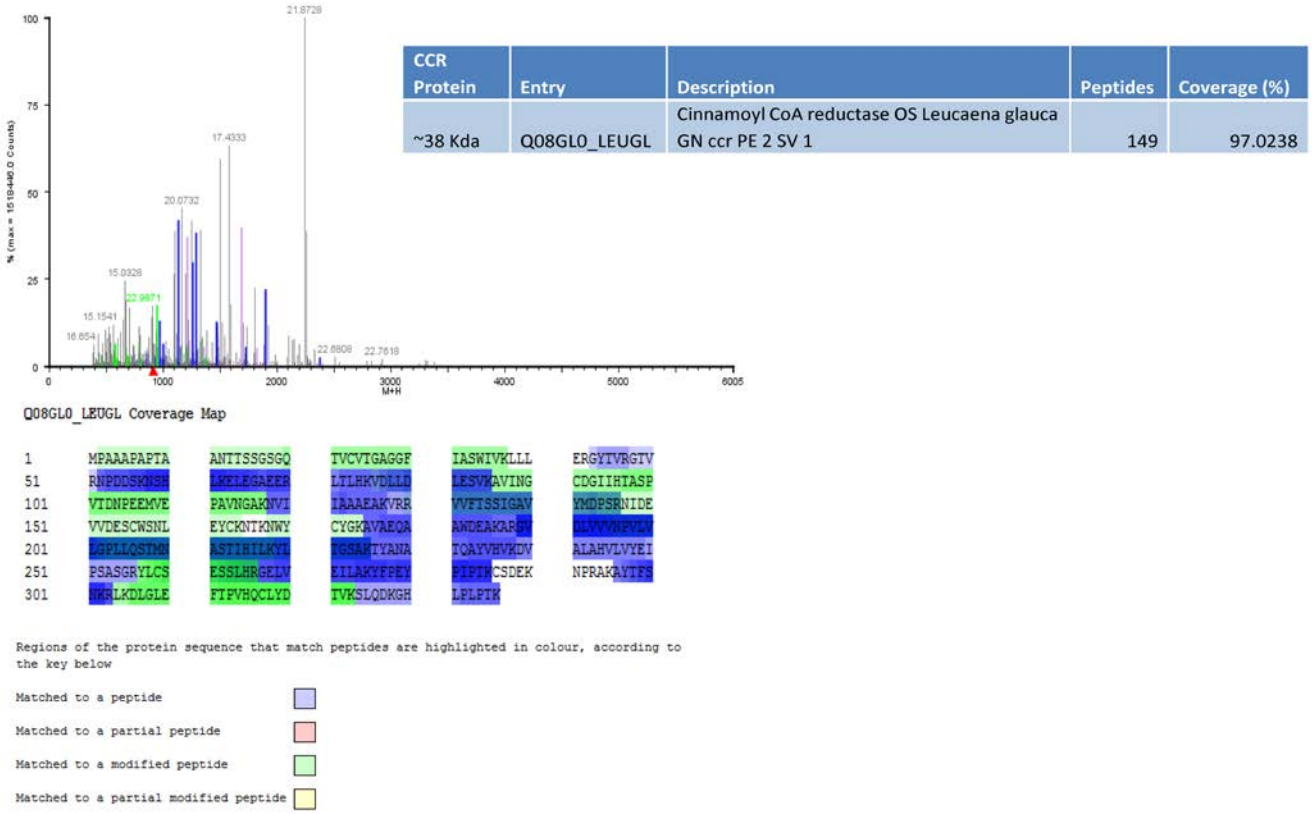


Fig. 2.7 ESI-MS/MS analysis of trypsin digested Ll-CCRH1.

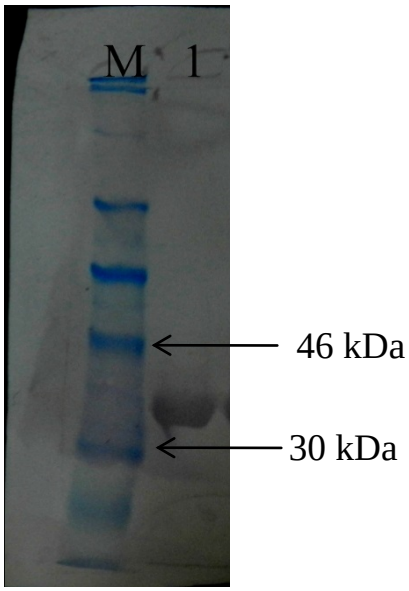


Fig. 2.8 Western blot analysis of recombinant Ll-CCRH1 protein. M: Prestained molecular weight marker (NEB), Lane 1: recombinant Ll-CCRH1, an apparent purified single 38 kDa band was observed.

The purified 6X His tagged Ll-CCRH1 was subjected to MALDI-TOF-MS analysis and observed mass value for $[M + H]^+$ was m/z 37477.9414 (Fig. 2.9).

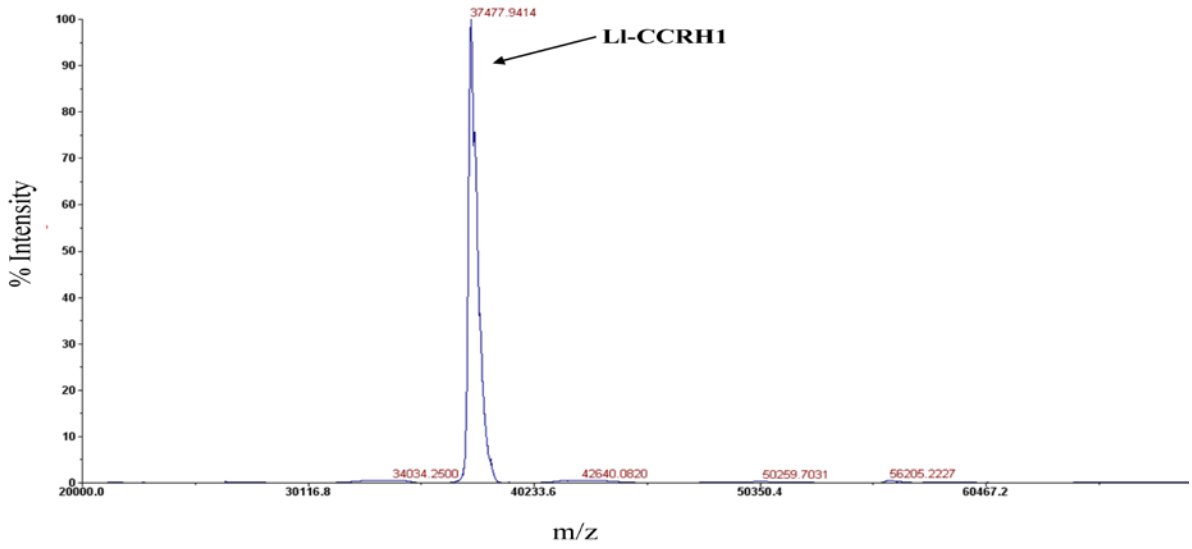


Fig. 2.9 Molecular mass determination of Ll-CCRH1 by MALDI-TOF/ MS analysis

Native recombinant Ll-CCRH1 enzyme showed mean hydrodynamic diameter of 6.4 nm by DLS analysis. Also, Ll-CCRH1 treated with 2-mercaptoethanol and SDS showed mean hydrodynamic diameter of 6 nm. The slight reduction in diameter suggests that native Ll-CCRH1 exists as monomer and show uniform particle size distribution (Fig. 2.10).

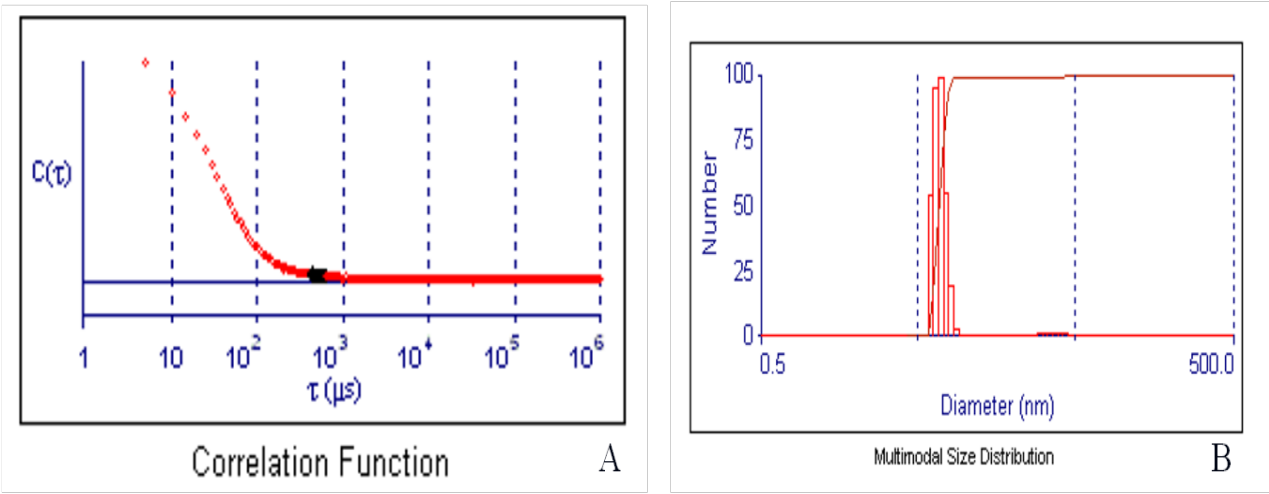


Fig. 2.10 Dynamic light scattering studies of Ll-CCRH1 protein. A, B: Correlation and multimodal size distribution plots for Ll-CCRH1 respectively.

2.3.3 Characterization of recombinant Ll-CCRH1

CCR carries out both reactions; the reduction of the hydroxycinnamoyl CoA esters to hydroxycinnamaldehydes (forward) and oxidation of cinnamaldehydes to corresponding CoA esters (reverse). The enzyme showed maximum activity at pH 6.5 in the direction of reduction. For oxidation reaction, Tris-HCl (pH 7.5-9) buffer was used and enzyme activity was optimal at pH 7.8 (Fig. 2.11 A). The optimum temperature for Ll-CCRH1 was found to be 30 °C (Fig. 2.11 B).

pH stability studies showed that Ll-CCRH1 was most stable around pH 6.5 at 25 °C for 90 min (Fig. 2.11 C); while at pH 6 and 7, Ll-CCRH1 was stable for 45 and

60 min, respectively. Below pH 6.0 and above pH 8.0, significant reduction in enzyme activity was recorded indicating the stability of Ll-CCRH1 over a specific pH range for

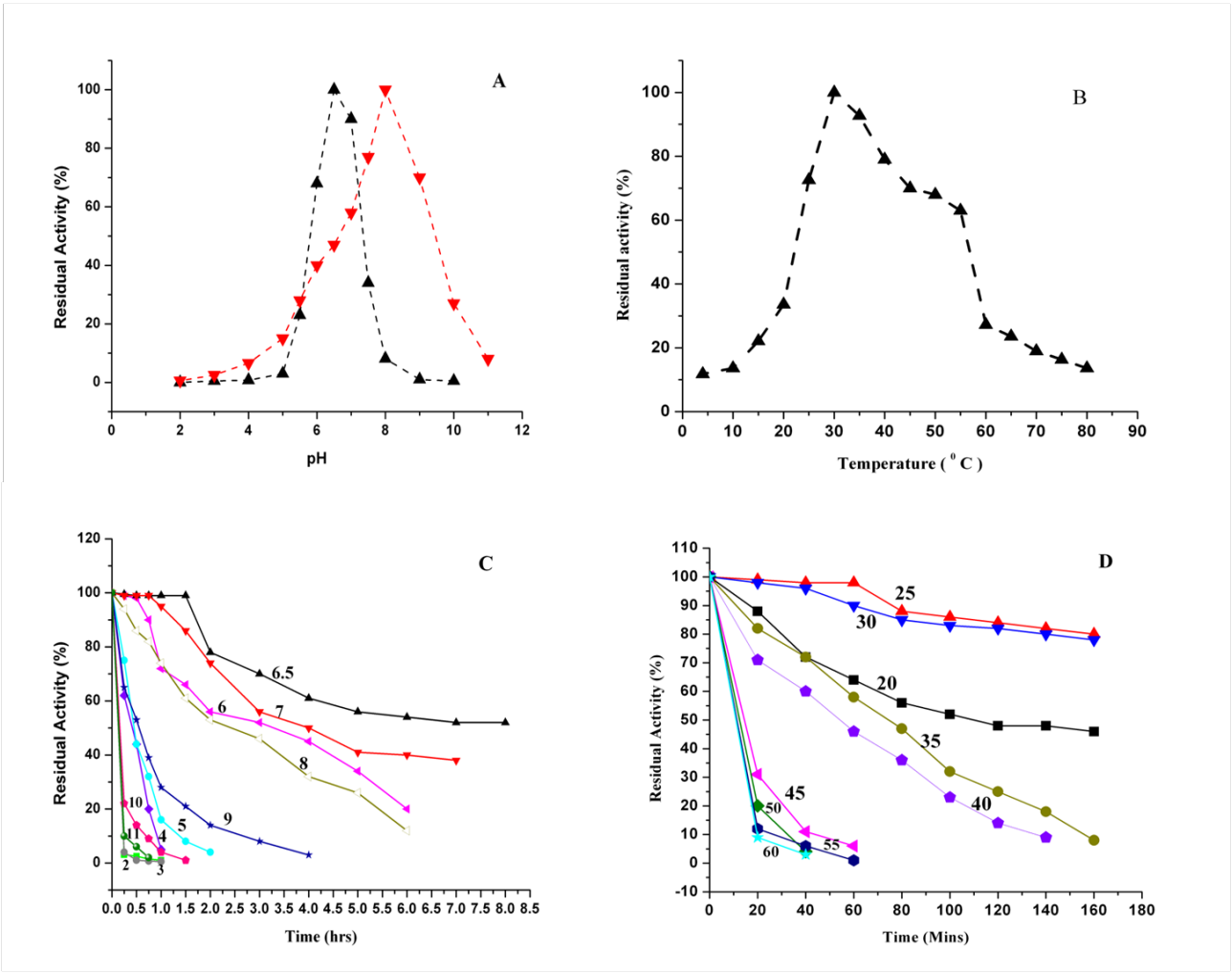


Fig. 2.11 Effect of pH and temperature on activity/stability of recombinant Ll-CCRH1. (A) Optimum pH determination of Ll-CCRH1 for forward and reverse reaction. ▲– –▲ represents forward reaction, while ▼– –▼ shows reverse reaction. (B) Temperature optima determination for Ll-CCRH1. (C), (D) pH and temperature stability of Ll-CCRH1; here aliquots of Ll-CCRH1 enzyme were incubated in pH range 2-11 and temperature range 10-80 °C respectively, and enzyme activity was measured using feruloyl CoA as a substrate at 30 °C. The numbers on the lines in graph indicate pH and temperature in °C respectively.

short duration of time. Temperature stability shows that Ll-CCRH1 was also found to be stable for 60 min at 25 °C and at optimum temperature 30 ^{o}C, stability was observed for 40 min. However, at 20 °C and 35 °C, 36% and 42% loss in activity was observed after 60 min. Above 40 °C, significant loss in Ll-CCRH1 activity was observed within 30 min (Fig. 2.11 D). The circular dichroism (CD) spectra of Ll-CCRH1 at temperature range from 20-40 °C almost remain constant. Also, secondary structure elements (helix, sheets, turns and unordered structures) calculated from CD pro software were also in similar range, indicating the unaltered structure of Ll-CCRH1 in this temperature range (discussed in chapter 4). Ll-CCRH1 shows optimum temperature of 25 °C for its activity and stability. Below and above 25 °C, gradual loss in the enzyme activity was recorded with respect to change in temperature. Thus, there may be change in architecture or geometry of the active site of Ll-CCRH1 at these temperatures irrespective of unchanged overall structure.

2.3.4 Substrate specificity of Ll-CCRH1

The activity of Ll-CCRH1 in presence of various concentrations of different cinnamoyl CoA esters and cinnamaldehydes was estimated separately. The K_m and V_{max} values were determined from Michaelis-Menten plots obtained by GraphPad prism and Sigma 12.0 softwares for each substrate. K_m values for forward and reverse substrates were in the same order of magnitude, ranging from 30-80 µM [92]. The k_{cat} values of Ll-CCRH1 for feruloyl, caffeoyl, sinapoyl, coumaroyl CoA, coniferaldehyde and sinapaldehyde were 164, 113, 131, 105, 94, 81 s^{-1} respectively (Table 2.2). These values are 10-1000 times higher than those reported for the CCR from various plant sources [87,95,100]. The specificity constant k_{cat} / K_m of feruloyl CoA, caffeoyl CoA, sinapoyl CoA and coumaroyl CoA at 30 °C were 4.6, 2.4, 2.3 and 1.7 (x 10^6 M^{-1} s^{-1}),

respectively, indicating preference of the enzyme for substrates and the rate of the reaction in order of feruloyl > caffeoyl > sinapoyl > coumaroyl CoA (Table 2.2).

Apparent kinetic parameters for reverse reaction were also determined using coniferaldehyde and sinapaldehyde as substrates at pH 7.8 as well as carrying out the reaction at pH 6.5. The K_m and V_{max} values were determined as described in previous section. Ll-CCRH1 showed comparatively higher affinity for coniferaldehyde than sinapaldehyde as reverse reaction substrate. At pH 7.8, the specificity constants for coniferaldehyde and sinapaldehyde were 7.5 and 5.3 (x 10^6 M^{-1} s^{-1}) respectively, and around 1.5-2.5 times higher than forward substrates (feruloyl CoA and sinapoyl CoA). On the other hand, reverse reaction at pH 6.5 (favorable for forward reaction) gives specificity constants 1.9 and 1.20 (x 10^6 M^{-1} s^{-1}), respectively (Table 2.2).

In plants, CCR catalyzed reduction reaction occurs in the pH range of 6-6.5 [87,89,91,95,100]. Comparison of kinetic parameters of the reduction (forward) and oxidation (reverse) reactions for Ll-CCRH1 at pH 6.5 showed that reduction of cinnamoyl CoA esters is favored over oxidation reaction in vitro too. Here, we report reverse reaction favored at pH 7.8 with higher specificity constants. Such reaction would not be possible in plants due to controlled and coordinated regulation of metabolic flux of lignin biosynthesis pathway genes.

Differential substrate specificity of CCR has been correlated to its exclusive or redundant function inside the cell either in lignification (feruloyl CoA/ sinapoyl CoA as most preferred substrate) during plant development or in defense mechanism (Coumaroyl CoA as favored substrate) [87,95,100]. Biochemical studies of Ll-CCRH1 showed better affinity for feruloyl CoA and caffeoyl CoA over other CoA esters; providing hint about its role in lignification rather than defense mechanism in plants.

Table 2.2 Steady state kinetic parameters, energy of activation and free energy of binding for Ll-CCRH1 with forward and reverse reaction substrates

Substrate	K_m **(µM)**	V_{max} **(Umg^{-1} min^{-1})**	k_{cat} **(s^{-1})**	k_{cat}/K_m **(M^{-1} s^{-1})**	E_a **(kJ/mol)**	ΔG **(kJ/ mol)**
Feruloyl CoA	36±1.12	258	164	4.6 x 10^6	30.26	-25.78
Caffeoyl CoA	48±1.08	177	113	2.4 x 10^6	31.08	-25.05
Sinapoyl CoA	57±1.44	206	131	2.3 x 10^6	32.67	-24.62
Coumaroyl CoA	60±1.69	164	105	1.7 x 10^6	35.98	-24.48
Coniferaldehyde (pH 7.8)	31±0.90	371	236	7.5 x 10^6	25.77	-26.12
Sinapaldehyde (pH 7.8)	39±1.65	330	210	5.3 x 10^6	28.26	-25.54
Coniferaldehyde (pH 6.5)	50±1.15	147	94	1.9 x 10^6	40.32	-24.95
Sinapaldehyde (pH 6.5)	67±1.21	127	81	1.20 x 10^6	46.72	-24.19

Reverse reaction (oxidation) of Ll-CCRH1 at pH 7.8 is an important feature for further study and holds significant biotechnological application. Hydroxycinnamoyl CoA esters are generally synthesized by chemical methods described earlier [143]. The reported method is tedious, time consuming and not eco-friendly; further, yields obtained are comparatively lower with presence of non-specific products. Using reverse reaction parameters, we can synthesize CoA esters biochemically within shorter period of time and with higher yields. Thus, recombinant Ll-CCRH1 can be utilized as a biocatalyst for production of cinnamoyl CoA esters through biotransformation.

2.3.5 Activation energies of CoA esters

From the slope of the plot of ln V_{max} versus 1000/T (R^2 =0.98 and 0.99), energy of activation, E_a for all the substrates was calculated (Table 2.2). For forward substrates, activation energies are almost more or less similar and in the range of 30-35 kJ/mol (Fig. 2.12 A). For reverse reaction at pH 6.5, E_a for coniferaldehyde and sinapaldehyde was 40.32 kJ/mol and 46.72 kJ/mol respectively (Fig. 2.12 B) indicating that there may be more constraints during binding and catalysis due to which oxidation reaction is slow compared to reduction reaction. Thus, the preference of enzyme for reduction over oxidation reaction seems to be reflected by *Ea*. The free energy of binding (ΔG) of forward and reverse substrates ranges from -24 to -26 kJ/mol.

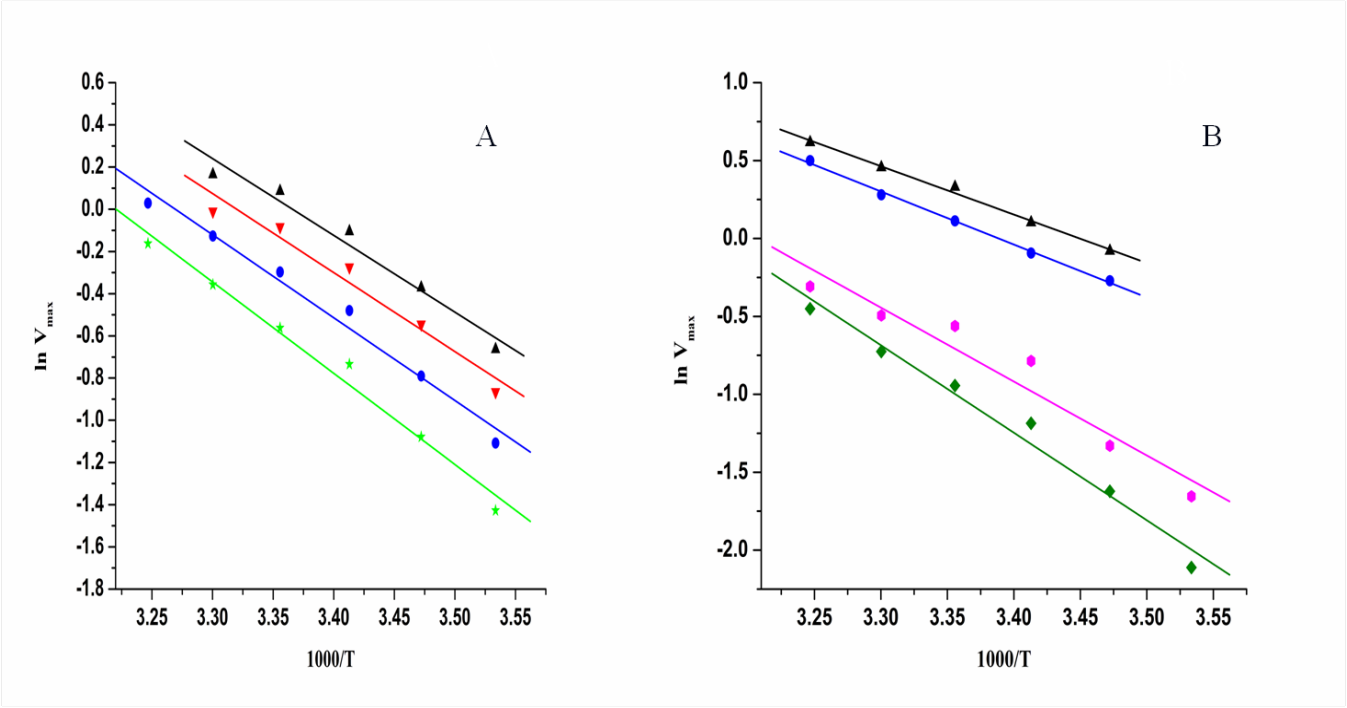


Fig. 2.12 Energy of activation of recombinant Ll-CCRH1 with various cinnamoyl CoA esters and cinnamaldehydes. 0.131 µM enzyme was taken for the forward and reverse assay. The reaction mixtures containing different concentrations of substrates were incubated at different temperatures (10-40 °C) and, K_m and V_{max} values were

determined. ln V_{max} was plotted against temperature. The Values were subjected to linear fitting procedure to obtain the E_a values. (A) feruloyl CoA (▲), caffeoyl CoA (▼), sinapoyl CoA (●) and coumaroyl CoA (★). (B) coniferaldehyde pH 8 (▲), sinapaldehyde pH 8 (●), coniferaldehyde pH 6.5 (♦) and sinapaldehyde pH 6.5 (■).

2.3.6 pH activity profile

The pH activity profile obtained by plotting ln (V_{max}/ K_m) against pH of purified Ll-CCRH1 revealed the participation of two ionizable groups with apparent pK_a values of 6.6 and 8.2 at the active site (Fig. 2.13).

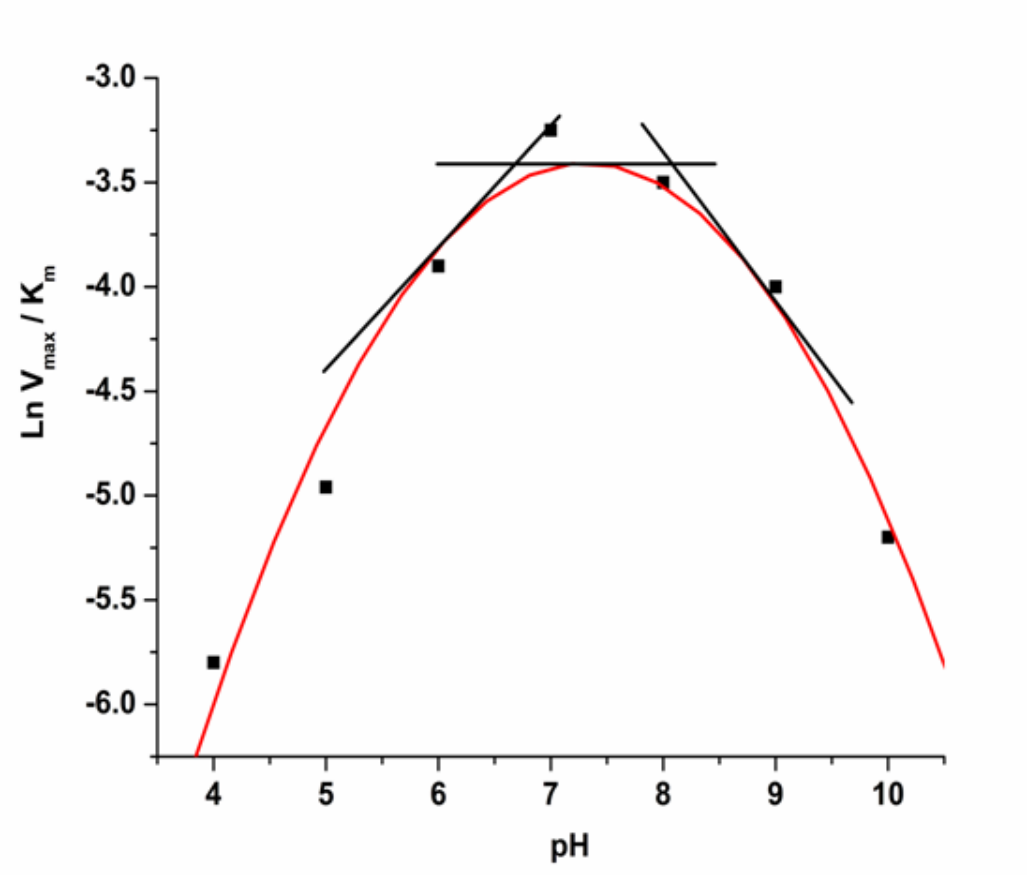


Fig. 2.13 pH activity profile of recombinant Ll-CCRH1. K_m and V_{max} values of the enzyme were determined in the pH range of 3-10. ln (V_{max} /K_m) of Ll-CCRH1 was

plotted against pH. The values were subjected to polynomial fitting procedure to yield the pK_a values.

The pK_a value of 6.6 indicates the probable involvement of carboxylate ion (COO^-) (pK_a value may be changed due to microenvironment) or histidine at active site of enzyme. The other pK_a of 8.2 suggests the possible participation of lysine, arginine or tyrosine at the active site.

This is the first preliminary biochemical evidence for the probable active site residues of Ll-CCRH1. Further, site directed mutagenesis and X-ray crystallographic studies will clarify the active site architecture. Multiple sequence alignment of Ll-CCRH1 with CCRs from other plant species by ClustalX shows the presence of highly conserved region, that is, a common signatural catalytic site ^{168}NWYCYGK^{174}. Here, tyrosine and lysine residues in catalytic motif seem to be reflected in pH activity profile also. In Ll-CCRH1, presence of conserved domain homologous with flavanol reductase from 22-309 amino acids is observed. The domain mainly consists of three motifs: catalytic active site residues, NADP binding site and substrate binding pocket (Fig. 2.14).

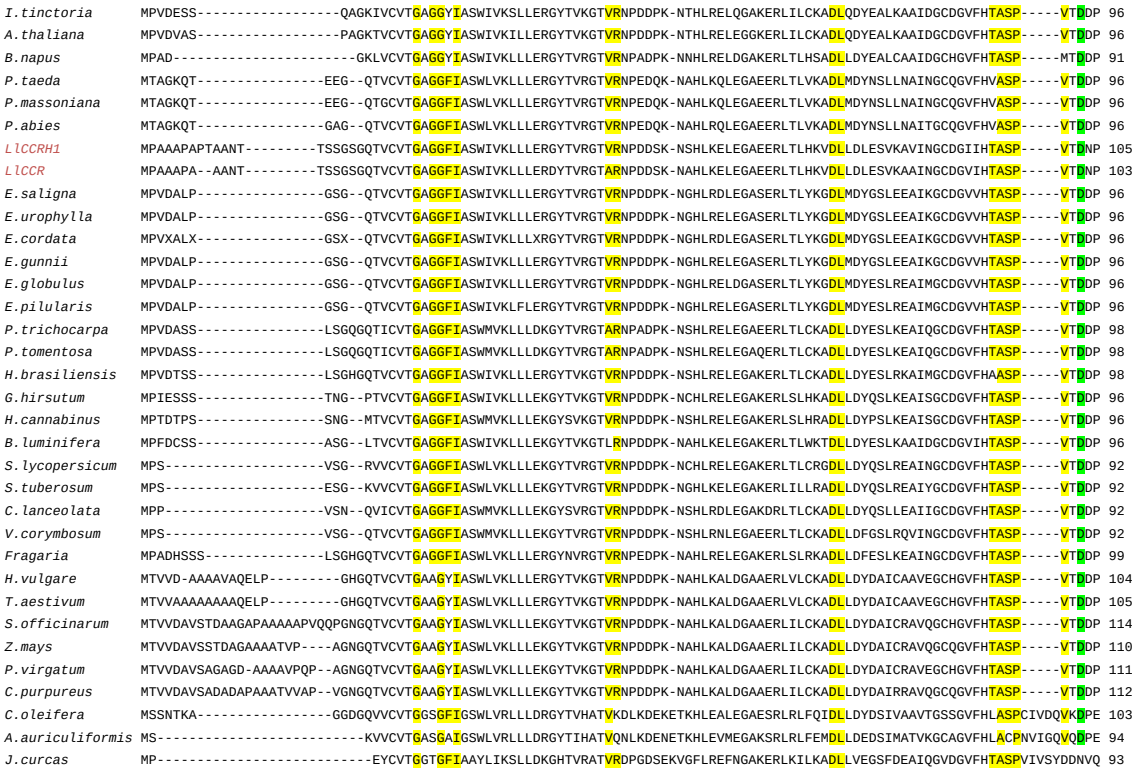

I.tinctoria MPVDESS----------QAGKIVCVTGAGGYIASWIVKSLLERGYTVKGTVRNPDDPK-NTHLRELQGAKERLILCKADLQDYEALKAAIDGCDGVFHTASP-----VTDDP 96
A.thaliana MPVDVAS----------PAGKTVCVTGAGGYIASWIVKILLERGYTVKGTVRNPDDPK-NTHLRELEGGKERLILCKADLQDYEALKAAIDGCDGVFHTASP-----VTDDP 96
B.napus MPAD----------GKLVCVTGAGGYIASWIVKLLLERGYTVRGTVRNPADPK-NNHLRELDGAKERLTLHSADLLDYEALCAAIDGCHGVFHTASP-----MTDDP 91
P.taeda MTAGKQT----------EEG--QTVCVTGAGGFIASWLVKLLLERGYTVRGTVRNPEDQK-NAHLKQLEGAEERLTLVKADLMDYNSLLNAINGCQGVFHVASP-----VTDDP 96
P.massoniana MTAGKQT----------EEG--QTGCVTGAGGFIASWLVKLLLERGYTVRGTVRNPEDQK-NAHLKQLEGAEERLTLVKADLMDYNSLLNAINGCQGVFHVASP-----VTDDP 96
P.abies MTAGKQT----------GAG--QTVCVTGAGGFIASWLVKLLLERGYTVRGTVRNPEDQK-NAHLRQLEGAEERLTLVKADLMDYNSLLNAITGCQGVFHVASP-----VTDDP 96
LlCCRH1 MPAAAPAPTAANT---------TSSGSGQTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDSK-NSHLKELEGAEERLTLHKVDLLDLESVKAVINGCDGIIHTASP-----VTDNP 105
LlCCR MPAAAPA--AANT---------TSSGSGQTVCVTGAGGFIASWIVKLLLERDYTVRGTARNPDDSK-NAHLKELEGAEERLTLHKVDLLDLESVKAAINGCDGVIHTASP-----VTDNP 103
E.saligna MPVDALP----------GSG--QTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDPK-NGHLRDLEGASERLTLYKGDLMDYGSLEEAIKGCDGVVHTASP-----VTDDP 96
E.urophylla MPVDALP----------GSG--QTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDPK-NGHLRELEGASERLTLYKGDLMDYGSLEEAIKGCDGVVHTASP-----VTDDP 96
E.cordata MPVXALX----------GSX--QTVCVTGAGGFIASWIVKLLLXRGYTVRGTVRNPDDPK-NGHLRDLEGASERLTLYKGDLMDYGSLEEAIKGCDGVVHTASP-----VTDDP 96
E.gunnii MPVDALP----------GSG--QTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDPK-NGHLRELEGASERLTLYKGDLMDYGSLEEAIKGCDGVVHTASP-----VTDDP 96
E.globulus MPVDALP----------GSG--QTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDPK-NGHLRELDGASERLTLYKGDLMDYESLREAIMGCDGVVHTASP-----VTDDP 96
E.pilularis MPVDALP----------GSG--QTVCVTGAGGFIASWIVKLFLERGYTVRGTVRNPDDPK-NGHLRELEGASERLTLYKGDLMDYESLREAIMGCDGVVHTASP-----VTDDP 96
P.trichocarpa MPVDASS----------LSGQGQTICVTGAGGFIASWMVKLLLDKGYTVRGTARNPADPK-NSHLRELEGAEERLTLCKADLLDYESLKEAIQGCDGVFHTASP-----VTDDP 98
P.tomentosa MPVDASS----------LSGQGQTICVTGAGGFIASWMVKLLLDKGYTVRGTARNPADPK-NSHLRELEGAQERLTLCKADLLDYESLKEAIQGCDGVFHTASP-----VTDDP 98
H.brasiliensis MPVDTSS----------LSGHGQTVCVTGAGGFIASWIVKLLLERGYTVKGTVRNPDDPK-NSHLRELEGAKERLTLCKADLLDYESLRKAIMGCDGVFHAASP-----VTDDP 98
G.hirsutum MPIESSS----------TNG--PTVCVTGAGGFIASWIVKLLLEKGYTVKGTVRNPDDPK-NCHLRELEGAKERLSLHKADLLDYQSLKEAISGCDGVFHTASP-----VTDDP 96
H.cannabinus MPTDTPS----------SNG--MTVCVTGAGGFIASWMVKLLLEKGYSVKGTVRNPDDPK-NSHLRELEGAKERLSLHRADLLDYPSLKEAISGCDGVFHTASP-----VTDDP 96
B.luminifera MPFDCSS----------ASG--LTVCVTGAGGFIASWIVKLLLEKGYTVKGTLRNPDDPK-NAHLKELEGAKERLTLWKTDLLDYESLKAAIDGCDGVIHTASP-----VTDDP 96
S.lycopersicum MPS----------VSG--RVVCVTGAGGFIASWLVKLLLEKGYTVRGTVRNPDDPK-NCHLRELEGAKERLTLCRGDLLDYQSLREAINGCDGVFHTASP-----VTDDP 92
S.tuberosum MPS----------ESG--KVVCVTGAGGFIASWLVKLLLEKGYTVRGTVRNPDDPK-NGHLKELEGAKERLILLRADLLDYQSLREAIYGCDGVFHTASP-----VTDDP 92
C.lanceolata MPP----------VSN--QVICVTGAGGFIASWMVKLLLEKGYSVRGTVRNPDDPK-NSHLRDLEGAKDRLTLCKADLLDYQSLLEAIIGCDGVFHTASP-----VTDDP 92
V.corymbosum MPS----------VSG--QTVCVTGAGGFIASWMVKLLLEKGYTVRGTVRNPDDPK-NSHLRNLEGAEERLTLCKADLLDFGSLRQVINGCDGVFHTASP-----VTDDP 92
Fragaria MPADHSSS----------LSGHGQTVCVTGAGGFIASWLVKLLLERGYNVRGTVRNPEDPK-NAHLRELEGAKERLSLRKADLLDFESLKEAINGCDGVFHTASP-----VTDDP 99
H.vulgare MTVVD-AAAAVAQELP---------GHGQTVCVTGAAGYIASWLVKLLLERGYTVKGTVRNPDDPK-NAHLKALDGAAERLVLCKADLLDYDAICAAVEGCHGVFHTASP-----VTDDP 104
T.aestivum MTVVAAAAAAAAQELP---------GHGQTVCVTGAAGYIASWLVKLLLERGYTVKGTVRNPDDPK-NAHLKALDGAAERLVLCKADLLDYDAICAAVEGCHGVFHTASP-----VTDDP 105
S.officinarum MTVVDAVSTDAAGAPAAAAAPVQQPGNGQTVCVTGAAGYIASWLVKLLLEKGYTVKGTVRNPDDPK-NAHLKALDGAAERLILCKADLLDYDAICRAVQGCHGVFHTASP-----VTDDP 114
Z.mays MTVVDAVSSTDAGAAAATVP----AGNGQTVCVTGAAGYIASWLVKLLLEKGYTVKGTVRNPDDPK-NAHLKALDGAAERLILCKADLLDYDAICRAVQGCQGVFHTASP-----VTDDP 110
P.virgatum MTVVDAVSAGAGD-AAAAVPQP--AGNGQTVCVTGAAGYIASWLVKLLLEKGYTVKGTVRNPDDPK-NAHLKALDGAAERLILCKADLLDYDAICRAVEGCHGVFHTASP-----VTDDP 111
C.purpureus MTVVDAVSADADAPAAATVVAP--VGNGQTVCVTGAAGYIASWLVKLLLEKGYTVKGTVRNPDDPK-NAHLKALDGAAERLILCKADLLDYDAIRRAVQGCQGVFHTASP-----VTDDP 112
C.oleifera MSSNTKA----------GGDGQVVCVTGGSGFIGSWLVRLLLDRGYTVHATVKDLKDEKETKHLEALEGAESRLRLFQIDLLDYDSIVAAVTGSSGVFHLASPCIVDQVKDPE 103
A.auriculiformis MS----------KVVCVTGASGAIGSWLVRLLLDRGYTIHATVQNLKDENETKHLEVMEGAKSRLRLFEMDLLDEDSIMATVKGCAGVFHLACPNVIGQVQDPE 94
J.curcas MP----------EYCVTGGTGFIAAYLIKSLLDKGHTVRATVRDPGDSEKVGFLRERNGAKERLKILKADLLVEGSFDEAIQGVDGVFHTASPVIVSYDDNVQ 93
I.tinctoria EQMVEPAVNGAKFVINAAAET-KVKRVVITSSIGAIYMDPNRDPEVVVDESCWSDLEFCKDTKNWYCYGKMVAEQAAWETAKEKGVDLVVLNPVLVLGPPLQPTINASLFHVLKYLTGSA 215
A.thaliana EQMVEPAVNGAKFVINAAAEA-KVKRVVITSSIGAVYMDPNRDPEAVVDESCWSDLDFCKNTKNWYCYGKMVAEQAAWETAKEKGVDLVVLNPVLVLGPPLQPTINASLYHVLKYLTGSA 215
B.napus ETMLEPAVNGAKFVIDAAAKA-KVKRVVFTSSIGAVYMNPNRDPQTIVNEDCWSDLDFCKNTKNWYCYGKMVAEQSAWETAKAKGVDLVVLNPVLVLGPPLQSAVNASLVHILKYLTGSA 210
P.taeda EEMVEPAVNGTKNVLDACAVA-GVRRVVFTSSIGAVYMDPSRDYDALVDENCWSNLDYCKETKNWYCYGKTVAEKAAWERAKDKGLDLVVVNPCVVLGPVLQSSINSSIIHILKYLTGSA 215
P.massoniana EEMVEPAVNGTKNVLDACAVA-GVRRVVFTSSIGAVYMDPSRDYDALVDENCWSNLDYCKETKNWYCYGKTVAEKAAWERAKDKGLDLVVVNPCVVLGPVLQSSINASIIHILKYLTGSA 215
P.abies VQMVEPAVNGTKNVLDACAEA-AVRRVVFTSSIGAVYMDPTRDYDALVDESCWSNLDFCKDTKNWYCYGKAVAEKAAWDRAKEKGLDLVVVNPCVVLGPVLQSSINASILHILKYLTGSA 215
LlCCRH1 EEMVEPAVNGAKNVIIAAAEA-KVRRVVFTSSIGAVYMDPSRNIDEVVDESCWSNLEYCKNTKNWYCYGKAVAEQAAWDEAKARGVDLVVVNPVLVLGPLLQSTMNASTIHILKYLTGSA 224
LlCCR EEMVEPAVNGAKNVIIAAAEA-KVRRVVFTSSIGAVYMDPSRNIDEVVDESCWSNLEYCKTTKNWYCYGKAVAEQAAWDEAKARGVDLVVVNPVLVLGPLLQTTMNASTIHILKYLTGSA 222
E.saligna EQMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKSTKNWYCYGKAVAEKARCAEAKERGVDLVVINPVLVLGPLLQSTINASIIHILKYLTGSA 215
E.urophylla EQMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKSTKNWYCYGKAVAEKAACAEAKERGVDLVVINPVLVLGPLLQSTINASIIHILKYLTGSA 215
E.cordata EQMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKSTKNWYCYGKAVAEKAACAEAKERGVDLVVINPVLVLGPLLQSTINASIIHILKYLTGSA 215
E.gunnii EQMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKSTKNWYCYGKAVAEKAAWPEAKERGVDLVVINPVLVLGPLLQSTINASIIHILKYLTGSA 215
E.globulus EQMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKSTKNWYCYGKAVAEKAACAXAKERGVDLVVINPVLVLGPLLQSTINASIIHILKYLTGSA 215
E.pilularis EQMVEPAVIGTKNVIVAAAEA-KVQRVVFTSSVGAITMDPNRGLDVVVDESCWSDLDFCKSTKNWYCYGKAVAEKSACAEAKERGVDLVVINPVLVLGPLLQSTVNASIIHILKYLTGSA 215
P.trichocarpa EEMVEPAVNGTKNVIIAAAEA-KVRRVVFTSSIGAVYMDPNKGPDVVIDESCWSDLEFCKNTKNWYCYGKAVAEQAAWDMAKEKGVDLVVVNPVLVLGPLLQPTVNASITHILKYLTGSA 217
P.tomentosa EEMVEPAVNGTKNVIIAAAEA-KVRRVVFTSSIGAVYMDPNKGPDVVIDESCWSDLEFCKNTKNWYCYGKAVAEQAAWDMAKEKGVDLVVVNPVLVLGPLLQPTVNASITHILKYLTGSA 217
H.brasiliensis EQMVEPAVNGTKNVVIAAAEA-KVRRVVFTSSIGAVYMDPNRNPDVVVDESCWSDLDFCKNTKNWYCYGKAVAEQAAWEVAKEKGVDLVAVNPVLVLGPLLQSTVNASIIHILKYLTGSA 217
G.hirsutum EQMVEPAVIGTKNVIMAAAEA-KVRRVVFTSSIGAVYMDPNRSPDVVVDESCWSDLEFCKNTKNWYCYGKAVAEQAAWETAKEKGVDLVAITPVLVLGPLLQPTVNASIVHILKYLTGSA 215
H.cannabinus EQMVEPAVNGTKNVIMAAAEA-KVRRVVFTSSIGAVYMDPNRSPDVVVDESCWSDLEFCKNTKNWYCYGKAVAEQAAWETAKEKGVDLVVVAPVLVLGPLLQSTVNASTVHILKYLTGSA 215
B.luminifera ELMVEPAVDGTKNVIIAAAET-KVRRVVFTSSIGAVYMDPNRGPDVVVDESCWSDLEFCKNTKNWYCYGKAVAEQAAWEVAEEKGVDLVVVNPVLVLGPLLQPNVNASVVHVLKYLTGSA 215
S.lycopersicum EQMVEPAVIGTKNVIVAAAEA-NVRRVVFTSSIGAVYMDPSRDPEKVVDETCWSDPDFCKNTKNWYCYGKAVAEQAAWDEARERGVDLVAINPVLVLGPLLQNTVNASVLHILKYLTGSA 211
S.tuberosum EQMVEPAVIGTKNVITAAAEA-KVGRVVFTSSIGTVYMDPNRAPDKVVDETCWSDLGFCKNTKNWYCYGKTVAEKTAWDEAREKGVDLVVINPVLVLGPLLQPTVNASVLHILKYLTGSA 211
C.lanceolata EQMVEPAVIGTKNVIVAAAEA-KCRRVVFTSSIGAVYMDPNRSPDAVVDETCWSDLEFCKNTKNWYCYGKAVAEQAAWDEAKVRGVDLVVVNPVLVLGPLLQHTVNASIVHVQKYLTGSA 211
V.corymbosum EEMVEPAVIGTKNVIVAAAEA-KVRRVVFTSSIGAVTMDPNRGPDTVVDESCWSDLEFCKNTKNWYCYGKAVAEQAAWDEAKDKGVDLVVVTPVLVMGPLLQPTLNASIIHVLKKYLNGSA 211
Fragaria EQMVEPAVNGTKNVIVAAAEA-KVKRVVFTSSIGAVYMDPARGPDVVVDESCWSDLEFCKNTKNWYCYGKAVAEQAAWEEAKERGVDLVVVNPVLVLGPLLQPTINASIIHILKYLTGSA 218
H.vulgare EQMVEPAVRGTEYVIDAAADAGTVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKKTKNWYCYGKAVAEQAAWEKARARGVDLVVVNPVLVVGPLLQPTVNASAAHILKYLDGSA 224
T.aestivum EQMVEPAVRGTEYVINAAADAGTVRRVGVTSSIGAVTMDPNRGPDVVVDESCWSDLEFCKKTKNWYCYGKAVAEQAAWEKAAARGVDLVVVNPVLVVGPLLQPTVNASAAHILKYLDGSA 225
S.officinarum EQMVEPAVRGTEYVINAAAEAGTVRRVVFTSSIGAVTMDPSRGPDVVVDESCWSDLEFCKKTKNWYCYGKAVAEQAAWDAARQRGVDLVVVNPVLVVGPLLQPTVNASIAHVVKYLDGSA 234
Z.mays EQMVEPAVRGTEYVINAAAEAGTVRRVVFTSSIGAVTMDPKRGPDVVVDESCWSDLEFCEKTRNWYCYGKAVAEQAAWETARRRGVDLVVVNPVLVVGPLLQATVNASIAHILKYLDGSA 230
P.virgatum EQMVEPAVRGTEYVIRAAAEAGTVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLDFCKKTRNWYCYGKAVAEQAAWDAARQRGVDLVVVNPVLVVGPLLQPTVNASIAHILKYLDGSA 231
C.purpureus EQMVEPAVRGTEYVLSAAAEAGTVRRVVFTSSIGAVTMDPNRGPDVVVDESCWSDLDFCKKTRNWYCYGKAVAEQSAWDAARQRGVDLVVVNPVLVVGPLLQPTVNASIAHILKYLDGSA 232
C.oleifera RELLEPAIKGTLNVLTAAKEL-GVRRVVVTSSNTAITPSPNWPADKVKNEDCWTDVEYCKQNGLWYPLSKTLAEKAAWEFAKEKGLDVVVVNPGTVMGPIIPPALNASMLMLLRFLQGCT 222
A.auriculiformis KEIVEPAVKGTVNVLKAAREA-GVERVVATSSISAIIPSPNWPSDRIKNEDCWCDLDYCKRKGLWYPIAKTLAEKAGWEFAKETGYDVVMINPGTALGPLIPPRLNSSMAVLLGVLKGDT 213
J.curcas ATLIDPCIKGTLNVLSSCTKATSVKRVVLTSSCSSIRYRYDVQQVCPLNESHWSDTDYCKRYNLWYAYAKTIGETEAWRIAKESGIDLVVVNPSFVVGPLLAPQPTSTLHLILSIVKGSL 213

Fig. 2.14 Alignment of Ll-CCRH1 (presented in grey color) with homolog CCR sequences from different plants illustrates a functional conserved domain FR_SDR_e. Signatural CCR sequence (NWYCYGK) is marked with pink color. Active site residues are shown in red color; NADP binding domain residues are distinguished by yellow color; substrate binding pocket is indicated by green color. Gaps are introduced to maximize homology and are shown by dashes. Highly conserved) residues; Tyr (170), Lys (174) and Ser (136) are present in all conserved motifs of both Ll-CCRs and are supposed to play critical role in catalysis. GenBank accession numbers of all CCRs used in alignment are as follows (starting from top): ADC40029, NP_173047, AEK27166, AAL47684, ACE76870, CAK18610, ABL01801.3, EU195224, AF297877_1, CBG37721, AAT74875, CAA56103, AAT74876, ACZ59064, CAC07424, ACE95172, ADU64758, ACQ59094, ADK24219, ACJ38670, AAY41879, AAN71761, BAE48787, ACI14382, AAP46143, AAN71760, ABE01883, CAA13176, ACG33996, ACZ74584, ADY39751, ACQ41893, ADQ53455, ACS32301.

2.3.7 Effect of metal ions and detergents

Various metal ions were checked for effect on Ll-CCRH1 enzyme activity. The enzyme was strongly inhibited by Ni^{2+}, Hg^{2+}, Mn^{2+} and partially inhibited by Zn^{2+}, Cu^{2+} and Co^{2+}. Ll-CCRH1 enzyme when treated with different concentrations of EDTA showed no inhibition in activity; thus showed that Ll-CCRH1 is neither a metal ion dependant nor gets inactivated by EDTA (Table 2.3). Incubation of metal ion inhibited Ll-CCRH1 enzyme with EDTA resulted in recovery of enzyme activity. Table 2.3 also shows that enzyme is inhibited by anionic detergent, SDS; but activated by Triton X-100 and Tween 20, because they are non-ionic detergents.

Table 2.3 Effect of metal ions and other reagents on Ll-CCRH1 activity.

Reagent	Residual activity (%)
EDTA (2mM, 5mM, 10 mM)	92, 88, 89
Metal ions (2mM)	
Control	100
Ni ($NiSO_4$)	25
Zn ($ZnSO_4$) / ($ZnCl_2$)	45, 55
Mg ($MgSO_4$)	74
Cu ($CuSO_4$)	49
Mn ($MnCl_2$)	21
K (KCl)	106
Ca ($CaCl_2$)	86
Cs ($CsCl_2$)	83
Co ($CoCl_2$)	42
Hg ($HgCl_2$)	No activity
Fe($FeSO_4$)	92
Enzyme (0.171 μM) + **Metal ions** (Ni, Zn, Cu, Mn, Co) 2mM + 2mM **EDTA**	100, 85, 95, 91, 80
Detergents (1%)	
SDS	6
Triton X-100	106
Tween 20	113

2.3.8 Small Angle X-ray Scattering (SAXS) studies of Ll-CCRH1

As three dimensional crystal structure of Ll-CCRH1 is unknown, its structure-function relationship is still not established. SAXS studies were performed to determine the structure of a protein molecule in terms of average particle size and shape. Scattering data were fitted using the SASFIT package (version 0.93.3) developed by Joachim Kohlbrecher at the Paul Scherrer Institute (http://kur.web.psi.ch/sans1/SANSSoft/sasfit).

Attempts to fit the data to models for polydisperse spherical scatterers were unsuccessful. Therefore, we fitted the scattering data, assuming that the protein can be modeled as monodisperse oblate ellipsoids. This model fitted the data reasonably well (Fig. 2.15).

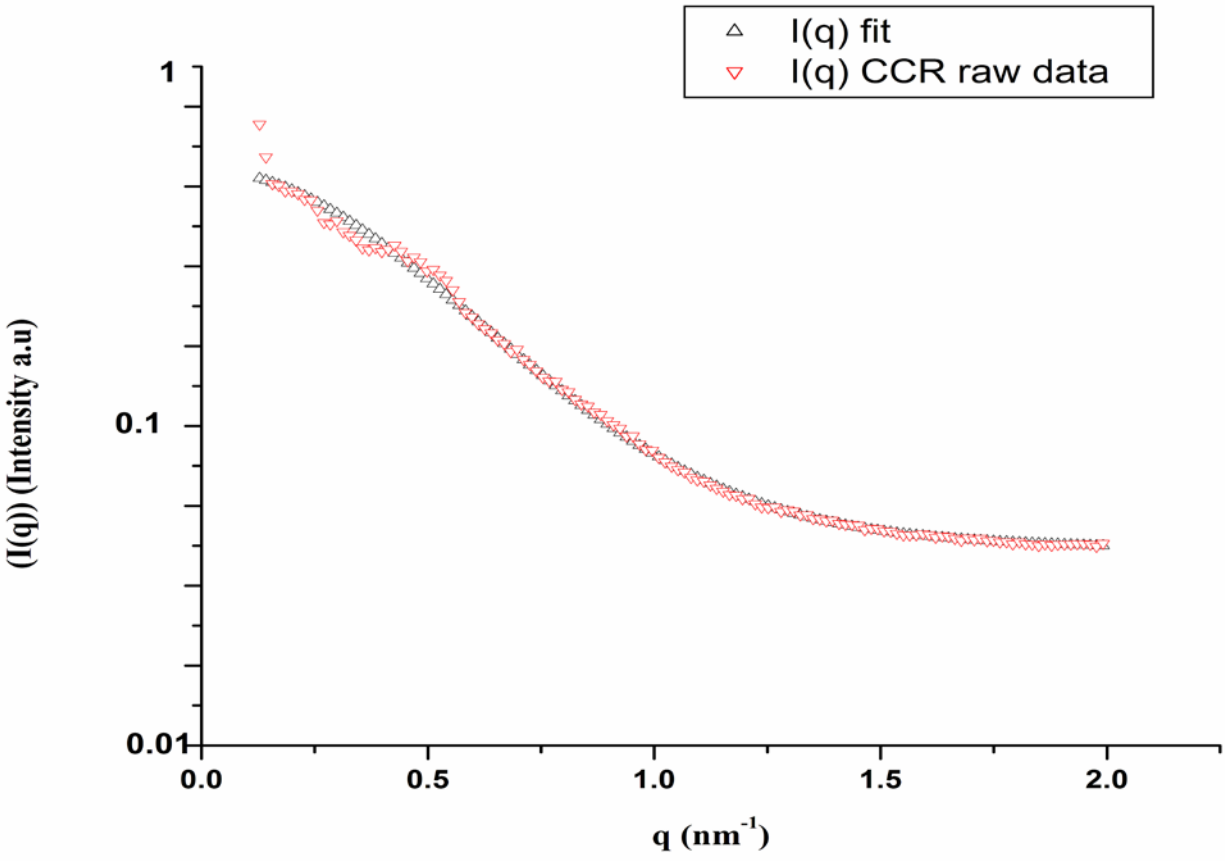


Fig. 2.15 SAXS data of Ll-CCRH1 in Tris-HCl, pH 8.0 buffer at 25 ^{o}C. The data was fitted using SASFIT program.

The fitted parameters were obtained as: radius 3.03 nm and volume 49.25 mm^3. We note that the size obtained from this fit matches well with the DLS size (radius 3.2 nm).

CHAPTER 3A

Probing the active site of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*

Prashant Sonawane et al. (2013) Biochemical characterization of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, ***Int J Biol Macromol***, 60, 33-38.

Summary

Lack of three dimensional crystal structure of cinnamoyl CoA reductase (CCR) limits its detailed active site characterization studies. Putative active site residues involved in the substrate/NADPH binding and catalysis for *Leucaena leucocephala* CCR (Ll-CCRH1; GenBank: DQ986907) were identified by amino acid sequence alignment and homology modeling. Putative active site residues and proximal H215 were subjected for site directed mutagenesis, and mutated enzymes were expressed, purified and assayed to confirm their functional roles. Mutagenesis of S136, Y170 and K174 showed complete loss of activity, indicating their pivotal roles in catalysis. Mutant S212G exhibited the catalytic efficiencies less than 10% of wild type, showing its indirect involvement in substrate binding or catalysis. R51G, D77G, F30V and I31N double mutants showed significant changes in K_m values, specifying their roles in substrate binding. Finally, chemical modification and substrate protection studies corroborated the presence of Ser, Tyr, Lys, Arg and carboxylate group at the active site of Ll-CCRH1.

3.1 Introduction

Lignin plays a major role in normal growth and development of plants; however, it is considered as an obstacle in number of agro-industrial processes, such as pulping and paper manufacture, forage digestibility and biofuel production. In plants, biosynthesis of monolignols is a specialized branch of phenylpropanoid metabolism, a complex series of branching biochemical reactions responsible for synthesis of variety of products like lignin, flavonoids and hydroxycinnamic acid conjugates [11]. Cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the first committed step in monolignol biosynthesis and considered as a first regulatory point in lignin formation [70]. CCR

carries out the NADPH dependent reduction of various hydroxycinnamoyl CoA esters to corresponding hydroxycinnamaldehydes and vice-versa [89,91]. Previous studies on CCR from various plants were basically focused on isolation, cloning, molecular characterization and downregulation aspects [87,95,99,101,104,106,108,109].

Functional attributes of any protein or enzyme are always correlated and dependent on 3D structures. Three dimensional structures of enzymes often provide significant information to understand catalysis mechanisms, substrate specificity and to modify the enzymes with altered specificity and/or advance catalytic function. Although CCR is one of the most investigated enzyme of lignin biosynthesis pathway, its three dimensional structure remains to be determined. The lack of the three dimensional crystal structure of CCR has precluded a clarification of functional active site residues involved in substrate binding and catalysis.

Leucaena leucocephala, a multipurpose nitrogen fixing tropical legume, has a tremendous prospective as a raw material (30-40%) for paper and pulp industry in India. As a step towards the active site characterization of Ll-CCRH1 (GenBank: DQ986907), homology modeling was carried out to predict 3D structure and catalytic active site residues. To confirm the roles of putative active site residues, fourteen different Ll-CCRH1 mutants were prepared by site directed mutagenesis and assayed for catalytic activity. To corroborate our mutational studies, chemical modification of Ll-CCRH1 using various amino acid group specific chemical reagents has been carried out. Finally, substrate protection studies of Ll-CCRH1 confirmed presence of modified amino acid residues at the active site.

3.2 Materials and Methods

3.2.1 Homology Modeling and substrate docking

Ll-CCRH1 protein sequence from *L. leucocephala* (NCBI protein accession No: ABL01801) was selected and BLASTP search was carried out against PDB (Protein Data Bank) for selecting template structures of closest homologues. Multiple sequence alignments of the target and template sequences were carried out using Clustal W 2.0 program with default parameters. The coordinates of crystal structure of template were used to build initial models of Ll-CCRH1. Finally 3D models of Ll-CCRH1 were generated using MODELLER 9v9 software. The quality of the selected models was assessed by DOPE (Discrete Optimized Protein Energy) method. The models generated were improved by molecular dynamics and equilibration methods using Chemistry at HARvard Macromolecular Mechanics (CHARM27) force field for lipids, proteins and nucleotides to remove bad contacts in the side chains.

The least energy homology models were evaluated for stereochemical quality by Ramchandran's plot using PROCHECK and environmental profile using ERRAT graph (Structure analysis and verification Server). ProSA server (protein structure analysis) was used to obtain Z-scores of Ramchandran's plot. Finally, the resultant energy minimized protein model was used for the active site identification and for docking analysis with the substrate.

Active site prediction studies were carried out with site finder in MOE (Molecular operating environment, Chemical computing group Inc. Canada), DALI server, SCF Bio Depth server, CASTp, Qsite finder, pocket finder and observing the residues by aligning the template sequences using Clustal W and also by 3D structural alignment of models.

Docking studies were carried out with five different hydroxycinnamoyl CoA esters; that is, 4-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5- hydroxyferuloyl CoA, sinapoyl CoA along with cofactor NADPH. Substrate molecules and NADPH were downloaded from Pubchem database on NCBI, and converted to 3D molecules using LigPrep module in the schrodinger suite (LigPrep, version 2.4; Schrodinger: New York, 2010). Protein-ligand complexes were minimized within an rmsd of 0.30 Å with force field OPLS2005 using MacroModel package (MacroModel, version 9.8; Schrodinger: New York, 2010). Protein- ligand docking simulations were conducted using AutoDock Vina tool to prepare the systems for calculations. For each ligand, around 100 docking runs with default parameters were performed treating protein as rigid and the ligand as flexible. The results were visualized using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.), wherein all the conformations of each of the ligand were found to be within the cavity of protein indicating that the docking run was free from errors. The conformational clusters with lowest binding energy (Ea) for each ligand were considered for further studies.

3.2.2 Heterologous expression and purification of wild type recombinant CCRH1 from *L .leucocephala*

Heterologous expression, purification and assay of the enzyme with cinnamoyl coA esters were carried out as described in chapter 2, section 2.2.1 and 2.2.3 [145].

3.2.3 Site directed mutagenesis of Ll-CCRH1 active site residues

Site directed mutagenesis was performed using the Quickchange Lightening Site Directed Mutagenesis Kit (Stratagene, USA) according to manufacturer's instructions. The wild type recombinant Ll-CCRH1 cDNA in pET30b (+) vector was used as template for PCR. Primers used for mutagenesis studies are shown in Table 3.1.

Putative positive clones were picked, and plasmids were isolated and sequenced. For double substitution mutant construction, F30V and R51G mutant plasmids were used as template.

Table 3.1 Primers used in mutagenesis studies

primer name	Length (base pair)	T_m	Primer Sequence (5' to 3')
R51G F	35	78.10°C	5'-CTGTTAGAGGCACCGTCGGAAATCCAGATGATTCT-3'
R51G R	35	78.10°C	5'-AGAATCATCTGGATTTCCGACGGTGCCTCTAACAG-3'
Y170H F	37	78.28°C	5'-AAGAACACAAAGAACTGGCATTGCTATGGGAAGGCAG-3'
Y170H R	37	78.28°C	5'-CTGCCTTCCCATAGCAATGCCAGTTCTTTGTGTTCTT-3'
K174M F	33	79.14°C	5'-AGAACTGGTATTGCTATGGGATGGCAGTGGCAG-3'
K174M R	33	79.14°C	5'-CTGCCACTGCCATCCCATAGCAATACCAGTTCT-3'
D77G F	43	78.73°C	5'-GGCTAACTCTTCATAAGGTTGGTCTTCTTGATCTTGAATCTGT-3'
D77G R	43	78.73°C	5'-ACAGATTCAAGATCAAGAAGACCAACCTTATGAAGAGTTAGCC-3'
V200E F	35	79.27°C	5'-GGTGAATCCAGTTTTGGAGTTGGGACCATTGCTTC-3'
V200E R	35	79.27°C	5'-GAAGCAATGGTCCCAACTCCAAAACTGGATTCACC-3'
S212G F	38	78.37°C	5'-CTTCAATCCACCATGAATGCAGGCACAATTCACATCCT-3'
S212G R	38	78.37°C	5'-AGGATGTGAATTGTGCCTGCATTCATGGTGGATTGAAG-3'
L64W F	39	79.50°C	5'-TCTAAGAACTCTCACTTGAAAGAGTGGGAAGGAGCAGAG-3'
L64W R	39	79.50°C	5'-CTCTGCTCCTTCCCACTCTTTCAAGTGAGAGTTCTTAGA-3'
S136A F	33	79.14°C	5'-AGAGTAGTGTTCACGTCAGCCATTGGAGCCGTC-3'
S136A R	33	79.14°C	5'-GACGGCTCCAATGGCTGACGTGAACACTACTCT-3'
H215L F	41	79.60°C	5'-CATGAATGCAAGCACAATTCTCATCCTCAAGTATCTCACTG-3'
H215L R	41	79.60°C	5'-CAGTGAGATACTTGAGGATGAGAATTGTGCTTGCATTCATG-3'
F30V F	25	78.38°C	5'-GGCGCCGGTGGCGTCATCGCCTCTT-3'
F30V R	25	78.38°C	5'-AAGAGGCGATGACGCCACCGGCGCC-3'
I31N F	26	80.08°C	5'-GGCGCCGGTGGCTTCAACGCCTCTTG-3'
I31N R	26	80.08°C	5'-CAAGAGGCGTTGAAGCCACCGGCGCC-3'

3.2.4 Expression, purification and biochemical characterization of Ll-CCRH1 mutants

After sequencing, mutant plasmids were transformed into *E. coli* strain BL21 (DE3) cells for protein expression. The generated Ll-CCRH1 mutant proteins were purified

and, the standard enzyme assay and substrate kinetics studies were also performed similarly as that for wild type recombinant Ll-CCRH1.

3.2.5 Chemical modifications and substrate protection studies of Ll-CCRH1

Modification of arginine, tyrosine, lysine, serine, histidine, tryptophan, cysteine residues and carboxylate groups by phenylglyoxal (Pg), N-acetyl imidazole (NAI), citraconic acid (CA), phenylmethylsulfonyl fluoride (PMSF), diethylpyrocarbonate (DEPC), N-Bromosuccinimide (NBS), N-ethylmaleimide (NEM)/p-hydroxymercurybenzoate (pHMB) and Woodward's reagent K (WRK) respectively, were carried out using methods described previously [146-148].

Substrate protection studies were carried out by incubating enzyme with various concentrations of feruloyl CoA (0.01-0.2 mM) for 5-10 min prior to treatment with modifying agents and then assayed modified enzyme with proper reaction controls.

3.3 Results and discussion

3.3.1 Construction of Ll-CCRH1 homology model

Full length cinnamoyl CoA Reductase gene (*Ll-CCRH1*, 336 amino acids, GenBank: DQ986907) had previously been isolated from *L. leucocephala* by similar methods described earlier for other CCR (*Ll-CCR*, 334 amino acids, GenBank: EU195224) from same plant [27]. Protein-protein BLAST search of Ll-CCRH1 against PDB showed a hit against Dihydroflavanol Reductase from *Vitis vinifera* (PDB ID: 2c29) and it demonstrated 43% identity with Ll-CCRH1. Thus 2c29 was chosen as template for modeling of the Ll-CCRH1 protein on the basis of sequence similarity, residues completeness, crystal resolution and functional similarities. Out of 50 models generated by MODELLER for Ll-CCRH1; only 5 models were selected based on DOPE score.

The final 3D model structure of Ll-CCRH1 satisfying evaluation criteria is displayed in Fig. 3.1A. Multiple sequence alignment of Ll-CCRH1 with template (PDB ID: 2c29) is shown in Fig. 3.1B.

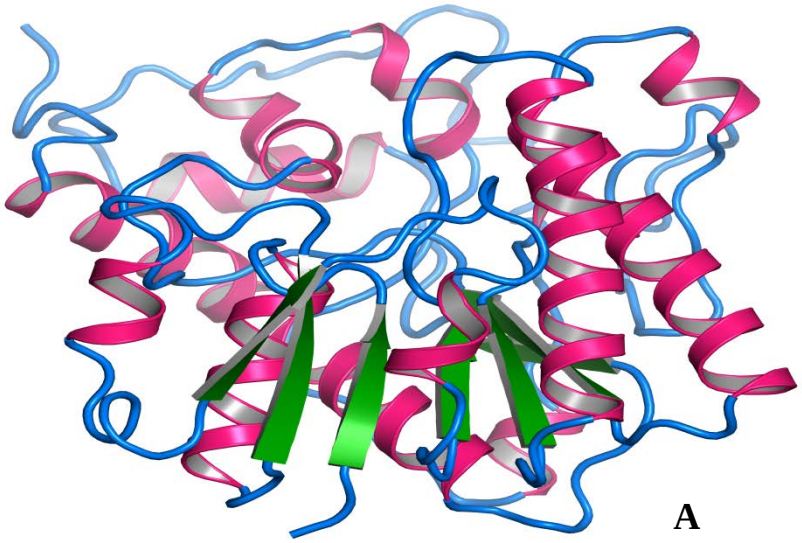


Fig. 3.1 A: Predicted three dimensional structure of Ll-CCRH1 represented in ribbon diagram.

```
2C29            ------------MGSQS-ETVCVTGASGFIGSWLVMRLLERGYTVRATVRDPTNVKKVK 46
ABL01801        MPAAAPAPTAANTTSSGSGQTVCVTGAGGFIASWIVKLLLERGYTVRGTVRNPDDSKNS- 59
                             .* * :*******.***.**:*  *********.***:* : *:
2C29            HLLDLPKAETHLTLWKADLADEGSFDEAIKGCTGVFHVATPMDFESKDPENEVIKPTIEG 106
ABL01801        HLKELEGAEERLTLHKVDLLDLESVKAVINGCDGIIHTASPV---TDNPE-EMVEPAVNG 115
                ** :*  ** :*** *.** *  *.. .*:** *::*.*:*:   :.:** *:::*:::*
2C29            MLGIMKSCAAAKTVRRLVFTSSAGTVNIQEHQLP--VYDESCWSDMEFCRAKKMTAWMYF 164
ABL01801        AKNVIIAAAEAK-VRRVVFTSSIGAVYMDPSRNIDEVVDESCWSNLEYCKNTKN---WYC 171
                  .:: :.* ** ***:***** *:* ::  :    * ******::*:*: .*     *
2C29            VSKTLAEQAAWKYAKENNIDFITIIPTLVVGPFIMSSMPPSLITALSPITGNEAHYSIIR 224
ABL01801        YGKAVAEQAAWDEAKARGVDLVVVNPVLVLGPLLQSTMNASTIHILKYLTGSAKTYANAT 231
                 .*::******. ** ..:*::.: *.**:**:: *:* .* *  *. :**.   *:
2C29            QGQFVHLDDLCNAHIYLFENPKAEGRYICSSHDCIILDLAKMLREKYPEYNIPTEFKG-V 283
ABL01801        QA-YVHVKDVALAHVLVYEIPSASGRYLCSESSLHRGELVEILAKYFPEYPIPTKCSDEK 290
                *. :**:.*:. **: ::* *.*.***:**. .    :*.::* : :*** ***: ..
2C29            DENLKSVCFSSKKLTDLGFEFKYSLEDMFTGAVDTCRAKGLLPPSHEKPVDGKT 337
ABL01801        NPRAKAYTFSNKRLKDLGLEFTPVHQCLYD-TVKSLQDKGHLPLPTK------- 336
                : . *:  **.*:*.***:**.   : ::  :*.: : ** ** . :
```

B

Fig. 3.1 B: Multiple sequence alignment of Ll-CCRH1 (GenBank protein ID: ABL01801) with template (PDB ID: 2c29). Highly conserved residues which correspond to active site of Ll-CCRH1 are marked in red color. The ' ***** ' indicates identical or conserved residues; ' : ' indicates a conserved substitution and ' **.** ' represent a semiconserved substitution.

The three dimensional structure of Ta-CCR2 from wheat was predicted using *Sporobolomyces salmonicolor* aldehyde reductase (PDB ID: 1UJM) as template and shares only 20% similarity with the template. The NADPH binding motif (TGAGGFI) and the reaction active site (NWYCY) were located in Ta-CCR2 [87]. Homologous alignment of amino acid sequence of Ll-CCRH1 with CCRs from other plant species shows presence of two highly conserved regions; a common signatural catalytic site NWYCYGK and a putative NADPH binding domain. In Ll-CCRH1, catalytic motif (NWYCYGK) lies from 168-174 amino acids. Escamilla Trevino et al. (2010) showed that the NWYCY motif is indispensable for CCR activity by analysing two recombinant CCR enzymes; one having NWYCY motif showed CCR activity, while another enzyme deficient in motif did not show any CCR activity [100].

Our investigations (computational and biochemical) found out that the entire motif is not required for CCR activity; Y170 and K174 are absolutely essential for catalysis of Ll-CCRH1 mediated reactions. Homology modeling and docking studies of CCR has also been previously reported from *L. leucocephala* [114]. In this report, it is quite surprising that both substrate and cofactor binding site of CCR did not contain any of the amino acid residue in catalytic motif, NWYCYGK. Also, no information was provided about CCR reaction mechanism and confirmatory biochemical evidences for active site residues.

3.3.2 Validation of Ll-CCRH1 model

The geometry of the final refined model was evaluated with Ramachandran's plot calculations computed with PROCHECK program and shows 92.6 % of Ll-CCRH1 residues are located in most favored regions, 7.4 % residues are in additional allowed region (Fig. 3.2). Further evaluation of Ll-CCRH1 model was carried out using PROSA server and the Z-score -9.98 for Ll-CCRH1, revealed it to be good quality model. Also overall PROCHECK G-factor for Ll-CCRH1 was -0.11 which is indicative of good quality model; meanwhile Ll-CCRH1 showed satisfactory overall ERRAT quality factor 81.967 (Fig. 3.3).

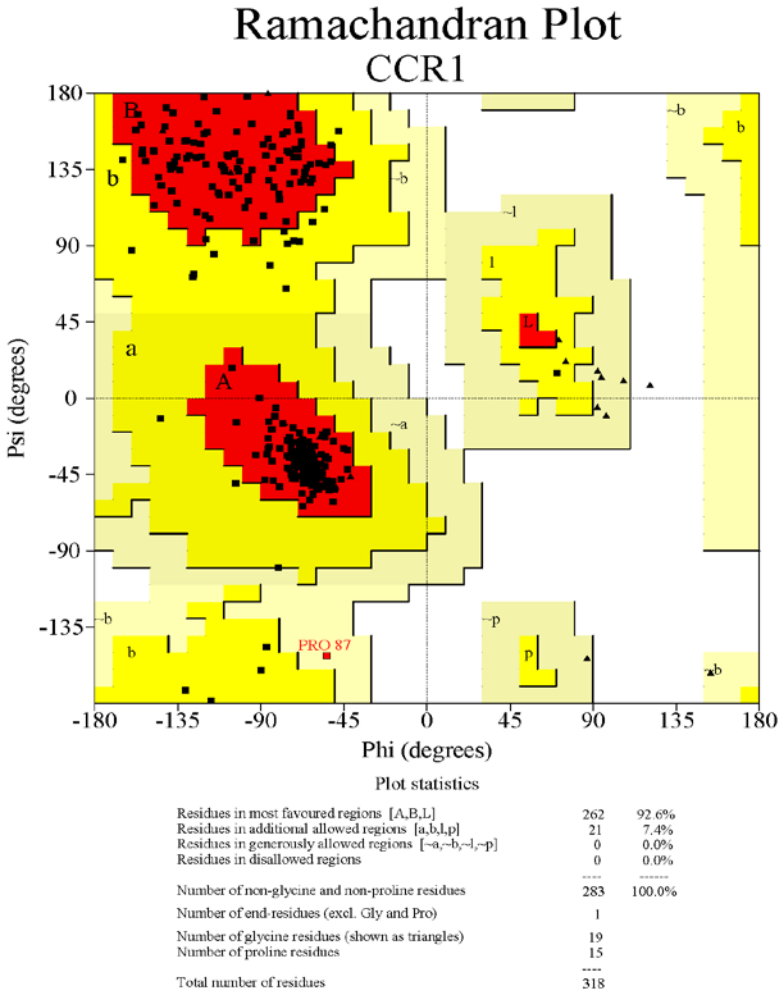


Fig. 3.2 Residue profile of Ll-CCRH1 in Ramachandran's plot

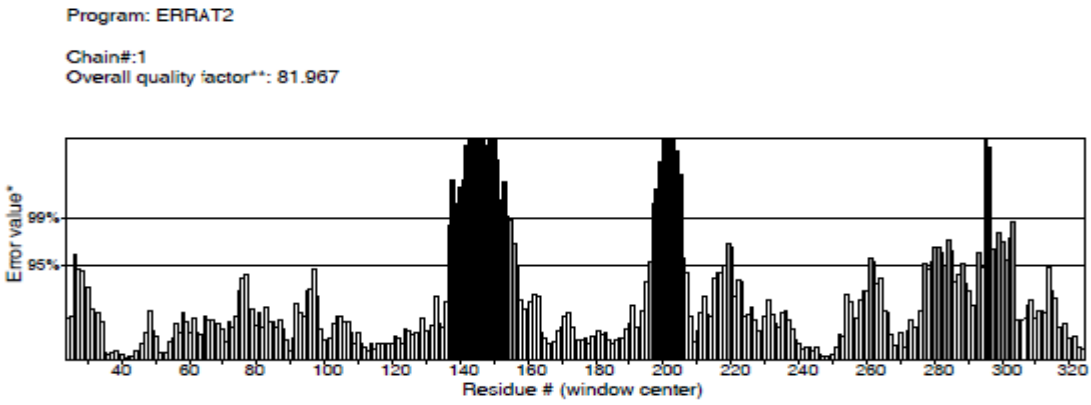


Fig. 3.3 3D profiles of Ll-CCRH1 model verified using ERRAT server. Overall quality factor indicates reasonably folded residues.

3.3.3 *In silico* active site identification of Ll-CCRH1

3.3.3.1 NADPH / Ll-CCRH1 interaction

Multiple sequence alignment of Ll-CCRH1 with other homologous CCRs (chapter 2, Fig. 2.14) and also with template 2c29 (Fig. 3.1B) was performed using ClustalW 2.0 to locate the conserved residues. The coenzyme NADPH binding site of Ll-CCRH1 is held by the structural motif known as Rossmann fold domain. NADPH binding site pocket consists of G26, G28, G29, T49, L78, S99, S135, Y170, K174, P197, V198, Q206 and A211 residues (Fig. 3.4). All the residues except S99, Q206, and A211 are conserved with NADPH binding site of the template 2c29 [115]. The diphosphate group of NADPH molecule binds to atoms of Gln206 and Ala211. On the other hand, P197 and V198 form electrostatic interactions with adenine moiety of AMP. Ribose sugar and phosphate of AMP are surrounded by side chains of S135, T97, Y170 and K174, which may play important role in stabilization of NADPH molecule. The nicotinamide ring forms interactions with G26, T49 and V76. Finally, S99 interacts with both hydroxyl groups of nicotinamide ribose (2’ and 3’ OH) (Fig. 3.4). In template

2c29, K167 forms H-bonds with nicotinamide ribose and reduces pKa of hydroxyl group of Y163 to promote proton transfer [115]. Here, S99 interacts with both hydroxyl groups of nicotinamide ribose (2' and 3' OH) (Fig. 3.4). Thus, similar types of interactions were observed with S99, indicating the residue might play important catalytic role in hydride transfer from the cofactor to the substrate in CCR mediated reduction reaction.

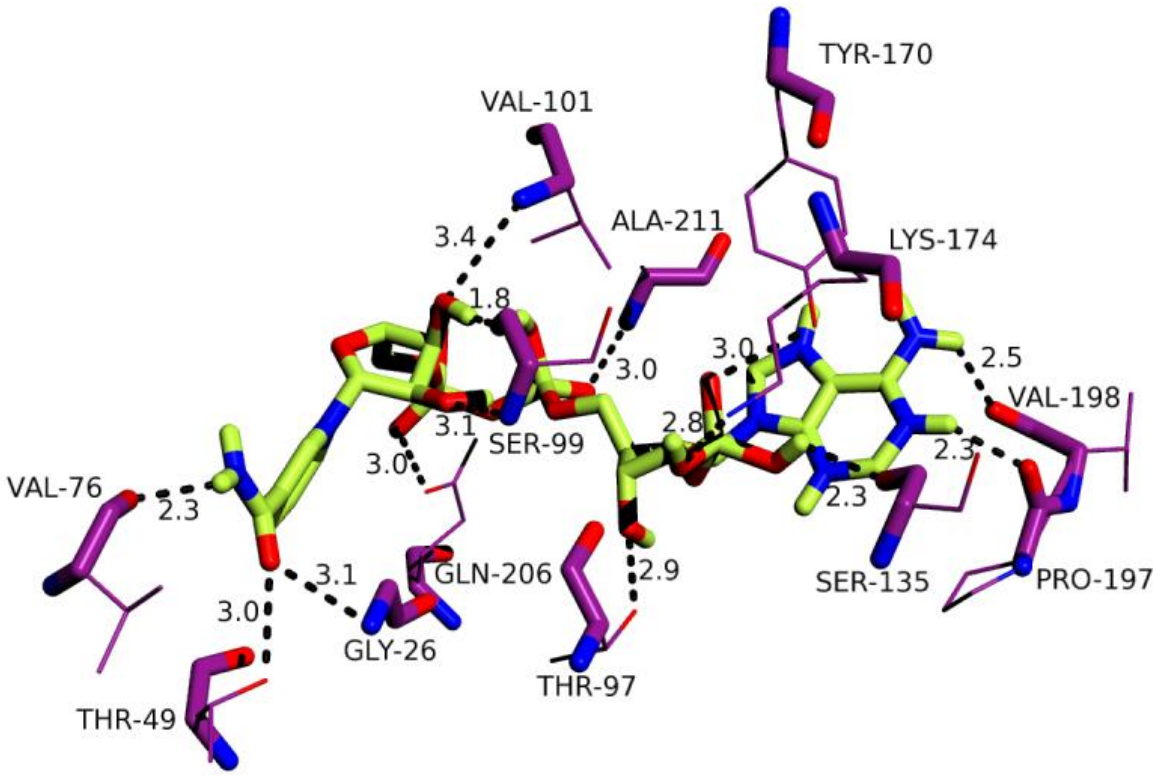


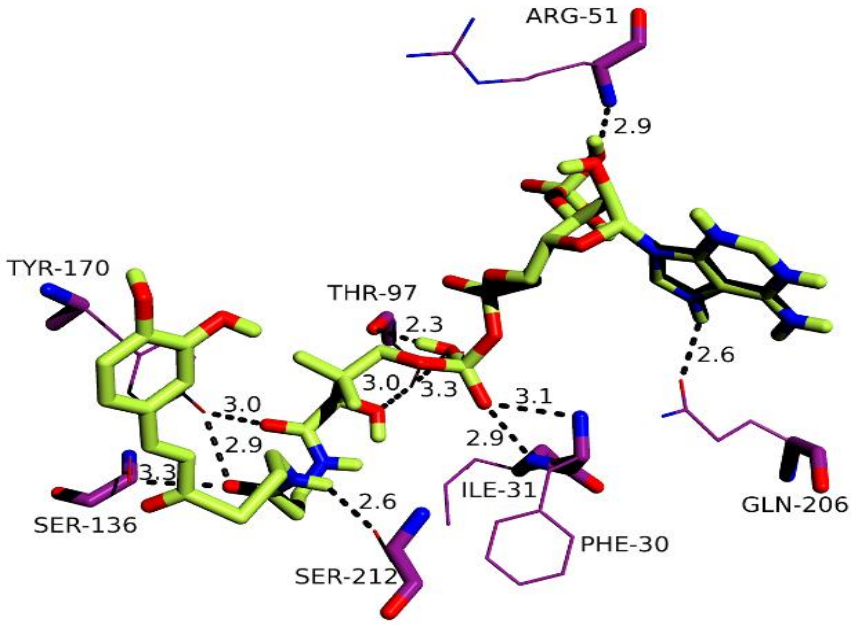
Fig. 3.4 Stick line diagram showing NADPH interactions with Ll-CCRH1. Dotted dark black lines shows bond distances (Å) between interacting groups of NADPH and amino acid residues of protein. Carbon atoms of the NADPH are shown in lemon green, whereas those of Ll-CCRH1 are displayed in deep purple. Oxygen and nitrogen atoms of the substrates are presented in red and blue colors respectively.

3.3.3.2 Putative substrate binding site of Ll-CCRH1

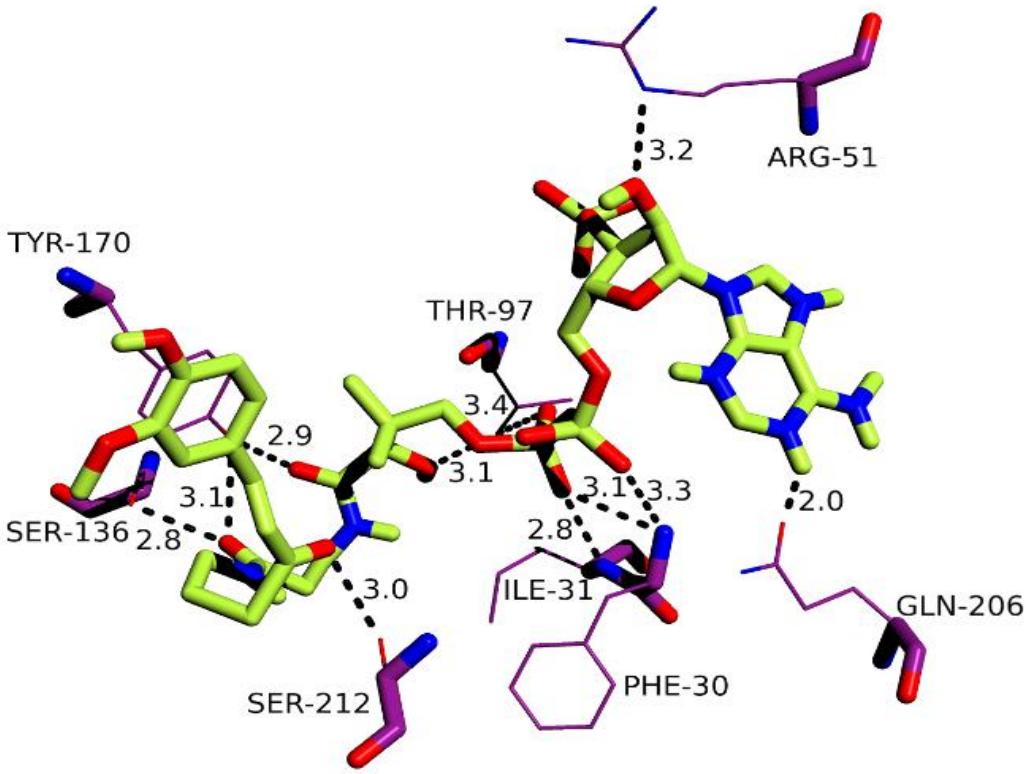
The substrate binding site of Ll-CCRH1 is made up of 10 residues, that is, F30, I31, R51, L64, D77, S136, Y170, K174, V200 and S212 as found by various bioinformatics tools. Residues corresponding to the active site in Ll-CCRH1 are highly conserved among the CCRs from various plant species and also with template 2c29 (Fig. 3.1B).

3.3.4 Docking simulations

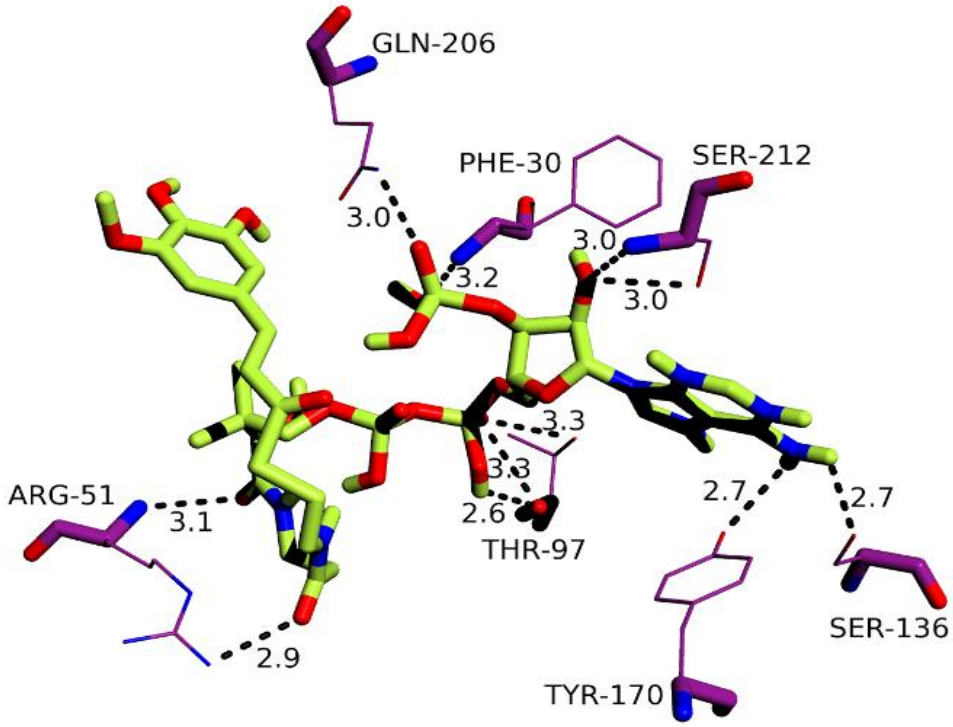
Docking conformations of CoA esters with Ll-CCRH1, obtained from AutoDock Vina [149] showed binding energies -9.9, -9.7, -9.4, -9.3 and -8.8 Kcal mole^{-1} for feruloyl CoA, 5-hydroxyferuloyl CoA, caffeoyl CoA, sinapoyl CoA and *p*-coumaroyl CoA respectively (Fig. 3.5 A-E, Table 3.2). The key intermolecular interactions between Ll-CCRH1 and CoA esters were identified from docking studies and are discussed in further section. The combined docking interactions of Ll-CCRH1 with NADPH and feruloyl CoA (most favored substrate) is shown in Fig. 3.6



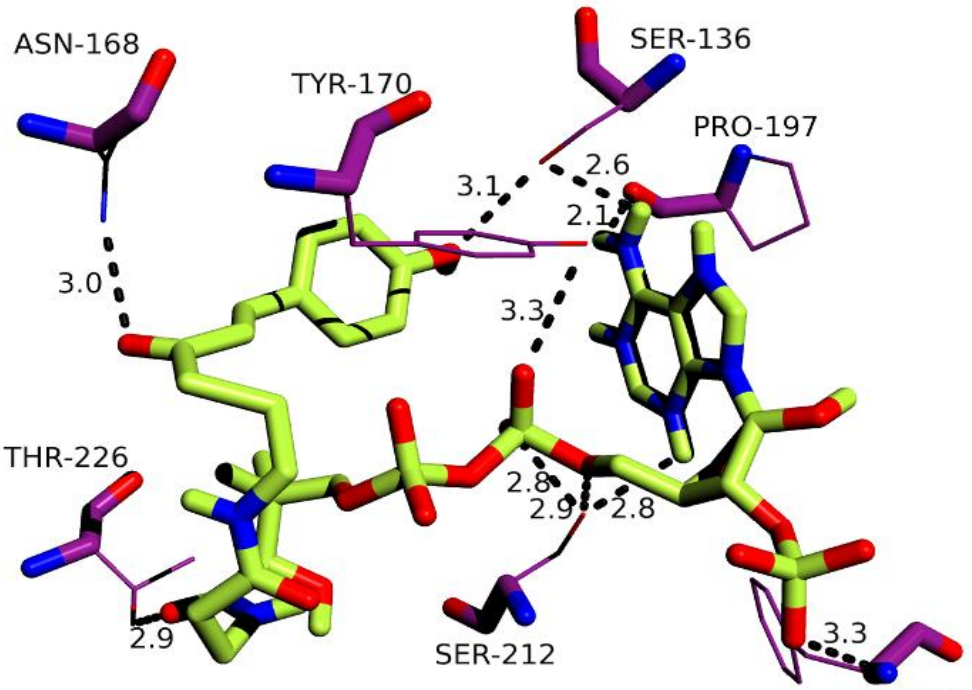
A: Caffeoyl CoA



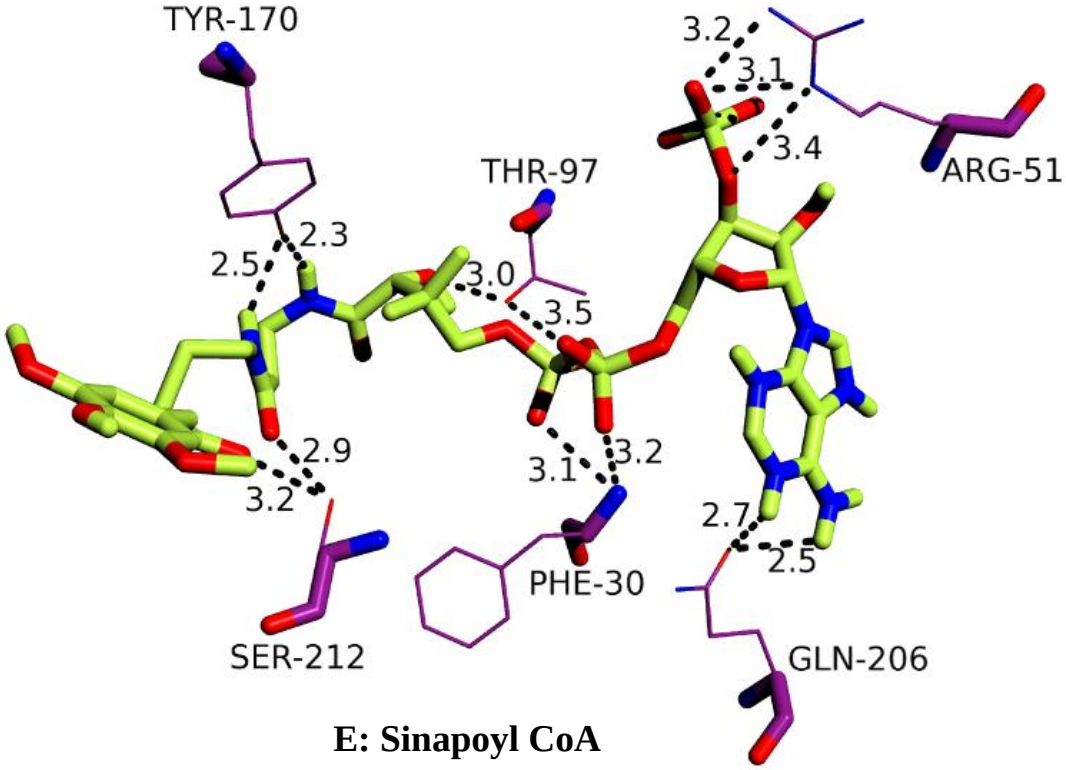
B: Feruloyl CoA



C: Hydroxyferuloyl CoA



D: Coumaroyl CoA



E: Sinapoyl CoA

Fig. 3.5 Sticks presentation showing interactions: A- caffeoyl CoA, B- feruloyl CoA, C- hydroxyferuloyl CoA, D- coumaroyl CoA and E- sinapoyl CoA substrates respectively with Ll-CCRH1. Interactions are expressed in terms of bond distances (Å) as dotted black lines. The color code is the same as in the legend for Figure 3.4.

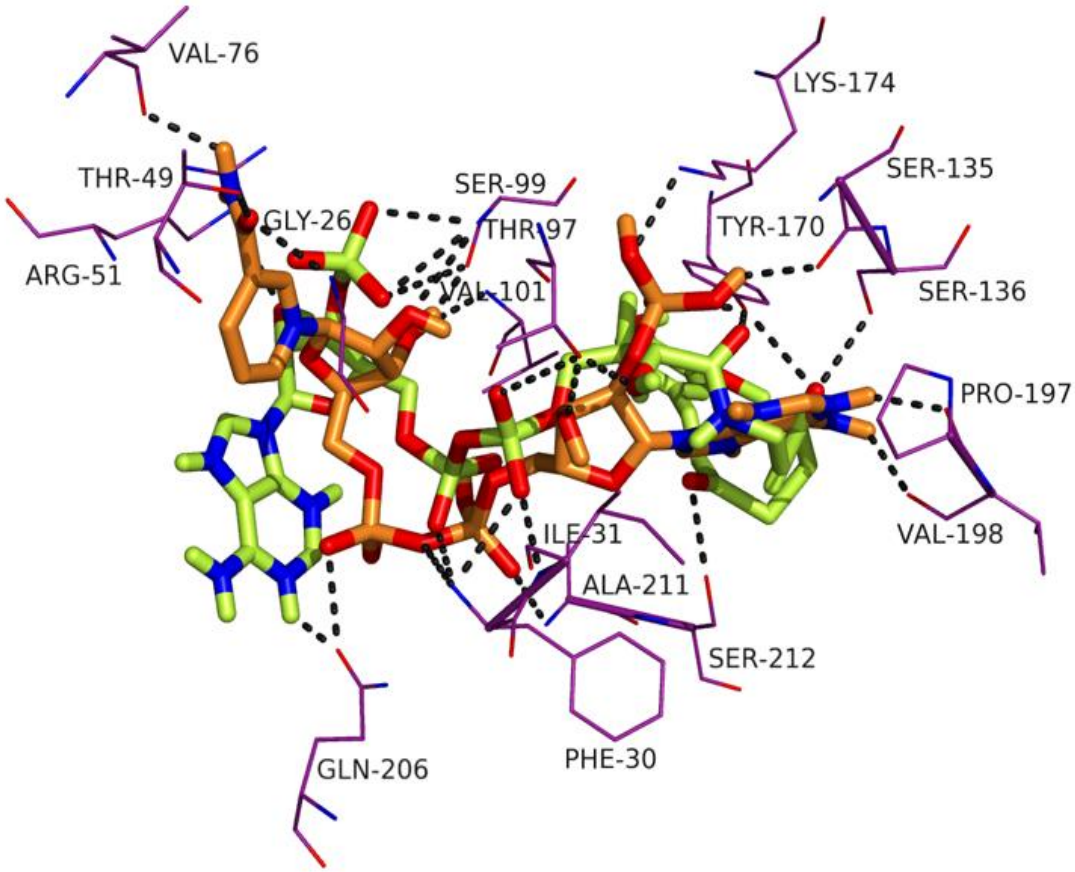


Fig. 3.6 Interaction of Ll-CCRH1 with cofactor NADPH and feruloyl CoA. Carbon atoms of the NADPH are shown in tv orange, whereas those of feruloyl CoA are displayed in lemon green color. Oxygen and nitrogen atoms of the cofactor and substrate are presented in red and blue colors respectively. Amino acid residues of Ll-CCRH1 are shown in deep purple color.

Table 3.2 Summary of results from docking simulations for Ll-CCRH1 and different CoA esters.

Substrate	Amino acid residues involved in binding	Substrate groups involved in interactions	Hydrogen bond distance (Å)	Docking energy (Kcal/mole)
Caffeoyl CoA	Phe30, Ile31, Thr97	Pantothenate of CoA	2.3-3.3	-9.4
	Arg51	3’ phosphate of ribose	2.9	
	Gln206	Adenine base	2.6	
	Ser136, Tyr170, Ser212	Cysteine of CoA group	2.6-3	
Feruloyl CoA	Phe30, Ile31, Thr97	Pantothenate of CoA	2.8-3.4	-9.9
	Arg51	2’ OH of ribose	3.2	
	Gln206	Adenine base	2.0	
	Ser136, Tyr170	Cysteine of CoA group	2.8-3.1	
	Ser212	Carbonyl carbon in aromatic ring	3.0	
Hydroxy-feruloyl CoA	Thr97	Pantothenate of CoA	2.6-3.3	-9.7
	Arg51	Cysteine and pantothenate of CoA group	2.9-3.1	
	Phe30, Gln206	3’ Phospahate of ribose	3-3.2	
	Ser136, Tyr170	Adenine base	2.7	
	Ser212	2’OH of ribose	3.0	
Coumaroyl CoA	Ser136	4’ OH on aromatic ring	2.6-3.1	-8.8
	Asn 168	Carbonyl carbon	3.0	
	Thr 226	Pantothenate group	2.9	
	Tyr170, Ser212	Phosphate in pantothenate	2.8-3.3	
	Phe30	3’ Phosphate of ribose	3.3	
	Pro197	Adenine base	2.1	
Sinapoyl CoA	Ser212	Carbonyl carbon and cysteine group	2.9-3.2	-9.3
	Tyr170	Cysteine and pantothenate group	2.3-2.5	
	Phe30, Thr97, Ile31	Pantothenate group	2.9-3.5	
	Arg51	3’ Phosphate of ribose	3.1-3.4	
	Gln206	Adenine base	2.5-2.7	

3.3.5 Predicted role of active site residues in Ll-CCRH1 mediated reactions

In template 2c29, catalysis mechanism was based on formation of canonical triad formed by lysine (proton transfer), serine (substrate stabilization) and tyrosine (catalytic base) [115]. In CCR catalyzed reaction, reduction is carried out with the hydride transfer mechanism from NADPH to Carbonyl carbon in aromatic ring of CoA ester. Docking simulations with modeled Ll-CCRH1 showed that the residues S136,

Y170 and S212 are present in vicinity of the reaction centre and form interactions with carbonyl group, cysteine moiety and pantothenate group of CoA (Fig. 3.5, Table 3.2). On the other hand, S99, a probable catalytic residue promoting hydride transfer is also in proximity of reaction site, and with residues S136, Y170 and S212. Thus, all these four residues are supposed to form catalytic reaction centre and may play important role in catalysis. Further, F30 and I31 interact with CoA esters via their main chain polar function; while R51 binds to esters via main chain as well as side chain. Certain residues, T97, Q206, H215, and T226, are located in nearby proximity (2-4 Å) of active site and may be involved in binding with substrate. Thus, these residues could participate in binding and stabilizing CoA esters. The role of binding and catalytic residues was further confirmed by site direct mutagenesis and chemical modification studies.

3.3.6 Characterization of Ll-CCRH1 mutant proteins

Based on the active site identification studies and interactions observed in modeling/ docking simulations, site directed mutant proteins were generated to confirm the functional role of putative residues involved in substrate binding or catalysis. Substitutions type mutations were designed to have minimal effect on secondary structure and 3D conformation of Ll-CCRH1 protein. Fourteen different Ll-CCRH1 mutants were constructed, cloned, over-expressed and purified to homogeneity before assaying enzyme activity.

3.3.6.1 Catalytic residues identification for Ll-CCRH1

Complete loss of activity was observed in mutants Y170H, K174M and S136A with feruloyl CoA as substrate (Table 3.3); indicating its critical role in catalysis. Surprisingly, in modeling and docking, no interactions were observed for putative

K174 residue with substrate; but mutational analysis and chemical modification studies of wild type recombinant Ll-CCRH1 proved essentiality of lysine in CCR reactions. Also an apparent 2.5 fold increase in $K_{m,}$ 7 fold reduction in k_{cat} and 15 fold decrease in k_{cat} / K_m for S212G suggests its indirect role in binding (holding reactive groups in proper orientation) or in catalysis (holding the catalytic conformation) (Table 3.3).

3.3.6.2 Substrate binding residues in Ll-CCRH1

Mutants, R51G, V200E, D77G, F30V and I31N displayed higher K_m values (up to 4 fold), indicating lower affinity towards substrate, while catalytic efficiencies for theses mutants were notably decreased to 2-3.5 fold (Table 3.3), suggesting their role in substrate binding and its stabilization in CCR reactions. To strengthen our findings, double substitution mutants F30V-I31N and R51G-D77G were generated from F30V, I31N, R51G and D77G mutants. F30V-I31N mutant showed approximately 7 fold increase in K_m and 8 fold decrease in k_{cat} / K_m values; while R51G-D77G displayed 5 fold increase in K_m and around 9 fold reduction in specificity constant suggesting essential role of F30, I31, R51 and D77, amino acids in substrate binding at or near the active site. Lastly, H215 residue (present in proximity of active site pocket) mutant also showed a reduction about 4-4.5 fold in k_{cat} / K_m and K_m (Table 3.3). H215 residue has not shown any interactions in docking simulations like K174, and also no significant changes in kinetic parameters were observed. But the probable involvement of histidine was confirmed by chemical modification studies (Section 3.3.7). Thus, H215 may be involved in substrate binding and stabilization, but not in catalysis.

Table 3.3 Kinetic parameters for wild type and mutant Ll-CCRH1 enzymes with feruloyl CoA as substrate (NA: No activity detectable).

Enzyme	K_m **(µM)**	V_{max} **($Umg^{-1}min^{-1}$)**	k_{cat} **(s^{-1})**	k_{cat}/K_m **($M^{-1}s^{-1}$)**
Wild type	36±1.12	258	164	4.6×10^6
R51G	47±0.95	158	100	2.1×10^6
Y170H	NA	NA	NA	NA
K174M	NA	NA	NA	NA
S212G	79±1.86	38	24	3.10×10^5
V200E	62±2.1	192	122	2.0×10^6
D77G	57±1.08	170	108	1.9×10^6
F30V	96±1.68	216	137	1.4×10^6
I31N	102±0.98	244	155	1.5×10^6
L64W	39±1.39	252	160	4.1×10^6
S136A	NA	NA	NA	NA
H215L	42±2.16	69	44	1.0×10^6
F30V- I31N	244±2.90	224	142	5.9×10^5

R51G- D77G	166±3.12	132	84	5.1×10^5
S99G	32±1.51	226	144	4.5×10^6

3.3.6.3 Possible role of K174 in hydride transfer

The S99 residue, which was supposed to be most important in transfer of hydride ion from NADPH to substrate, showed no significant loss of Ll-CCRH1 activity with any of the substrates after mutation (Table 3.3). Kinetic parameters for the S99G mutants were in the same order of magnitude as compared to wild type. Thus, role of S99 remains doubtful in CCR mediated reactions. The critical residue involved in hydride transfer still remains a matter of question. Here, K174 in Ll-CCRH1 might be performing the same function of hydride transfer as in template 2C29 [115]. K174 interactions might be reducing the pKa of Y170; thus facilitating proton transfer from Y170. Except L64W and S99G, all putative residues viz. F30, I31, R51, D77, S136, Y170, K174, S212 and H215 demonstrated their role in binding and catalysis based on kinetic parameters like $K_{m,}$ $V_{max,}$ k_{cat} and $k_{cat} / K_{m.}$ (Table 3.3).

3.3.7 Active site characterization by chemical modification

It is evident that even single mutation at or near the active site residues results in change of catalytic activity or completely altered substrate specificity. We have performed site directed mutagenesis studies based on modeling and docking simulations. Chemical modification of Ll-CCRH1 using various amino acid group specific chemical reagents was carried out to corroborate our mutational studies.

Treatment of enzyme with NEM/pHMB and NBS (1-10 mM) resulted in no loss of activity. Lack of inhibition by NEM and NBS suggested non involvement of

cysteine and tryptophan in enzyme catalysis respectively. Ll-CCRH1 activity was strongly inhibited by phenylglyoxal (1 mM), NAI (3 mM), DEPC (3 mM), CA (1.5 mM), PMSF (1 mM) and Woodward's reagent K (0.02 mM), indicating probable presence of arginine, tyrosine, histidine, lysine, serine and carboxylate at or near the active site (Table 3.4).

Table 3.4 Effect of group specific chemical modifiers on Ll-CCRH1 activity.

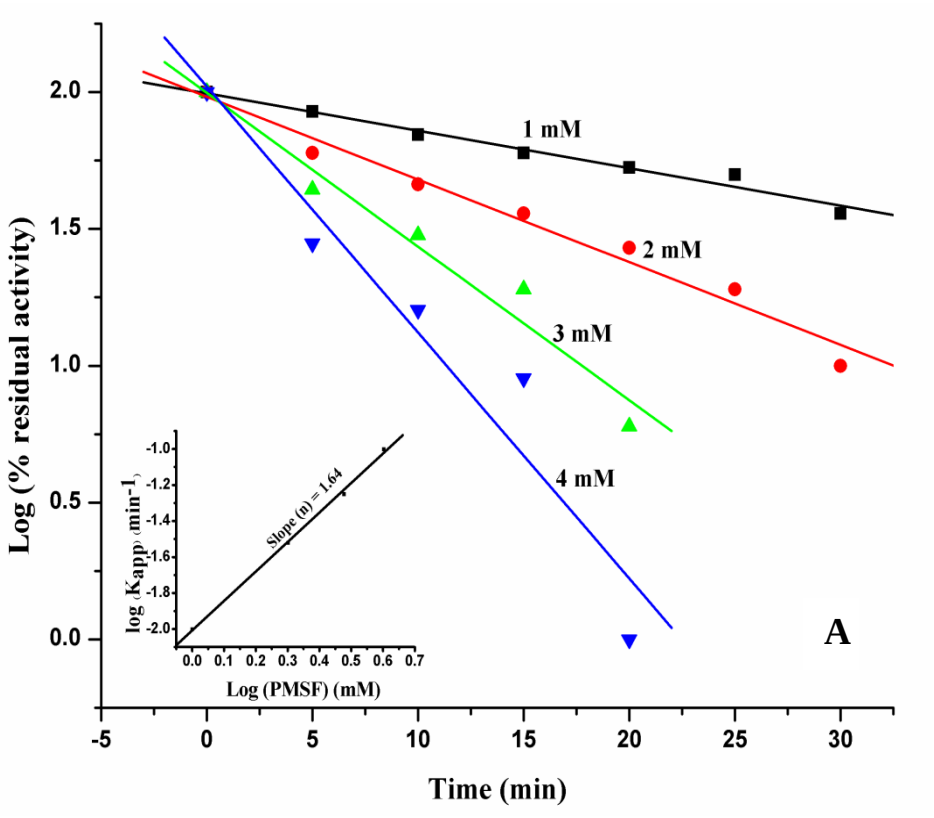
Chemical reagents	**Concentration**	**Reaction site**	**% Residual activity**	
			Feruloyl CoA	**Sinapoyl CoA**
Control	-	-	100	100
NEM	1, 5, 10 mM	Cysteine	98, 98, 98	97, 97, 97
pHMB	1, 5, 10 mM	Cysteine	98, 97, 96	100, 98, 98
NBS	1, 5, 10 mM	Tryptophan	100, 100, 98	98, 95, 95
Phenylglyoxal	1 mM	Arginine	47	35
NAI	3 mM	Tyrosine	50	43
DEPC	3 mM	Histidine	44	30
Citraconic anhydride	1.5 mM	Lysine	31	28
PMSF	1 mM	Serine	60	40
Woodward's reagent K	0.02 mM	Carboxylate	54	38

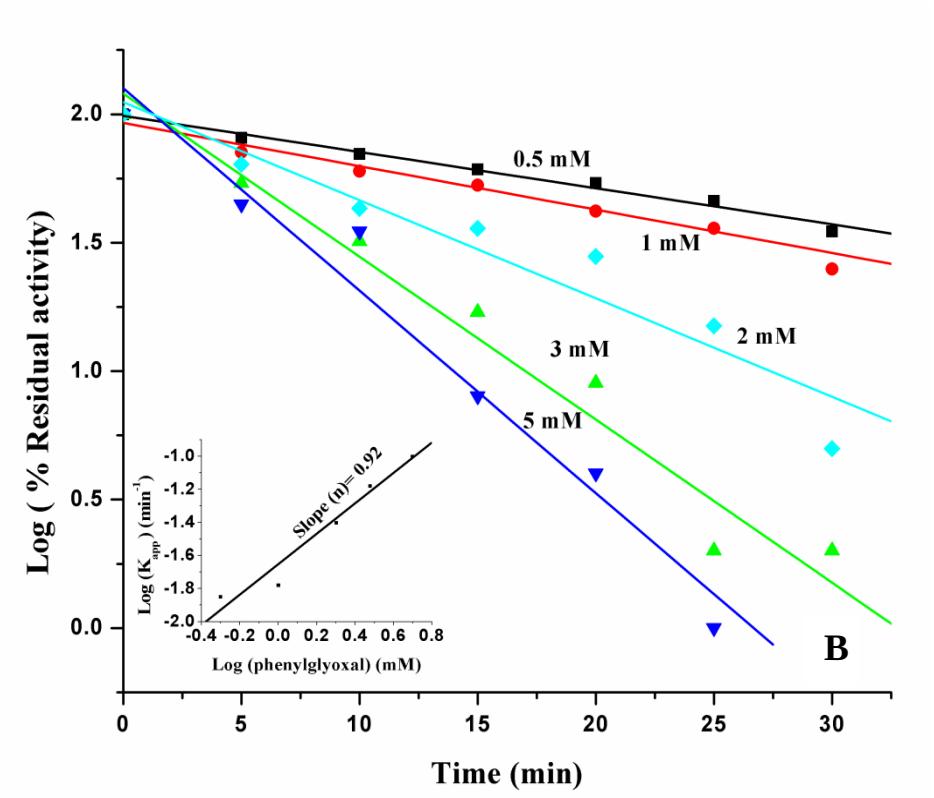
Except for carboxylate group residue, modification and subsequent inactivation of the enzyme by above mentioned chemical reagents followed pseudo first order kinetics at any fixed concentration of the reagent (Fig. 3.7). The individual slopes of plots (Log % residual activity Vs Time) were used to determine the respective rate

constant (K_{app}). Double logarithmic plots of rate constants (K_{app}) Vs reagent concentration led us to derive the number of respective residues at the active site as one residue each of arginine, tyrosine, histidine and lysine, and two of serine residues (Fig. 3.7). From the results of modeling/docking and mutational analysis, the two serine residues are most likely to be S136 and S212.

Table 3.5 Protection of Ll-CCRH1 against inactivation by various chemical modifiers (E: Enzyme, S: Substrate, feruloyl CoA).

Amino acids modified	Conditions	Residual activity (%)
Control	E	100
Arginine	E + Phenylglyoxal (1, 5 mM)	47, 21
	E + S (100, 140 μM) + Phenylglyoxal (1, 5 mM)	84, 89
Tyrosine	E + NAI (3, 10 mM)	50, 16
	E + S (150, 200 μM) + NAI (3, 10 mM)	88,70
Histidine	E + DEPC (3, 5 mM)	44, 32
	E + S (100, 125 μM) + DEPC (3, 5 mM)	84, 76
	E + DEPC (3, 5 mM) + Hydroxylamine (10 mM)	79, 74
Lysine	E + CA (2, 6 mM)	31, 20
	E + S (120, 180 μM) + CA (2, 6 mM)	81,68
Serine	E + PMSF (1, 10 mM)	60, 14
	E + S (75, 150 μM) + PMSF (1, 10 mM)	89, 93
Carboxylate group	E + WRK (20, 40 μM)	54, 30
	E + S (50, 100 μM) + WRK (20, 40 μM)	78, 76


Log (% residual activity)
Time (min)
1 mM
2 mM
3 mM
4 mM
A
log (Kapp) (min-1)
Slope (n) = 1.64
Log (PMSF) (mM)


Log (% Residual activity)
Time (min)
0.5 mM
1 mM
2 mM
3 mM
5 mM
B
Log (Kapp) (min-1)
Slope (n)= 0.92
Log (phenylglyoxal) (mM)

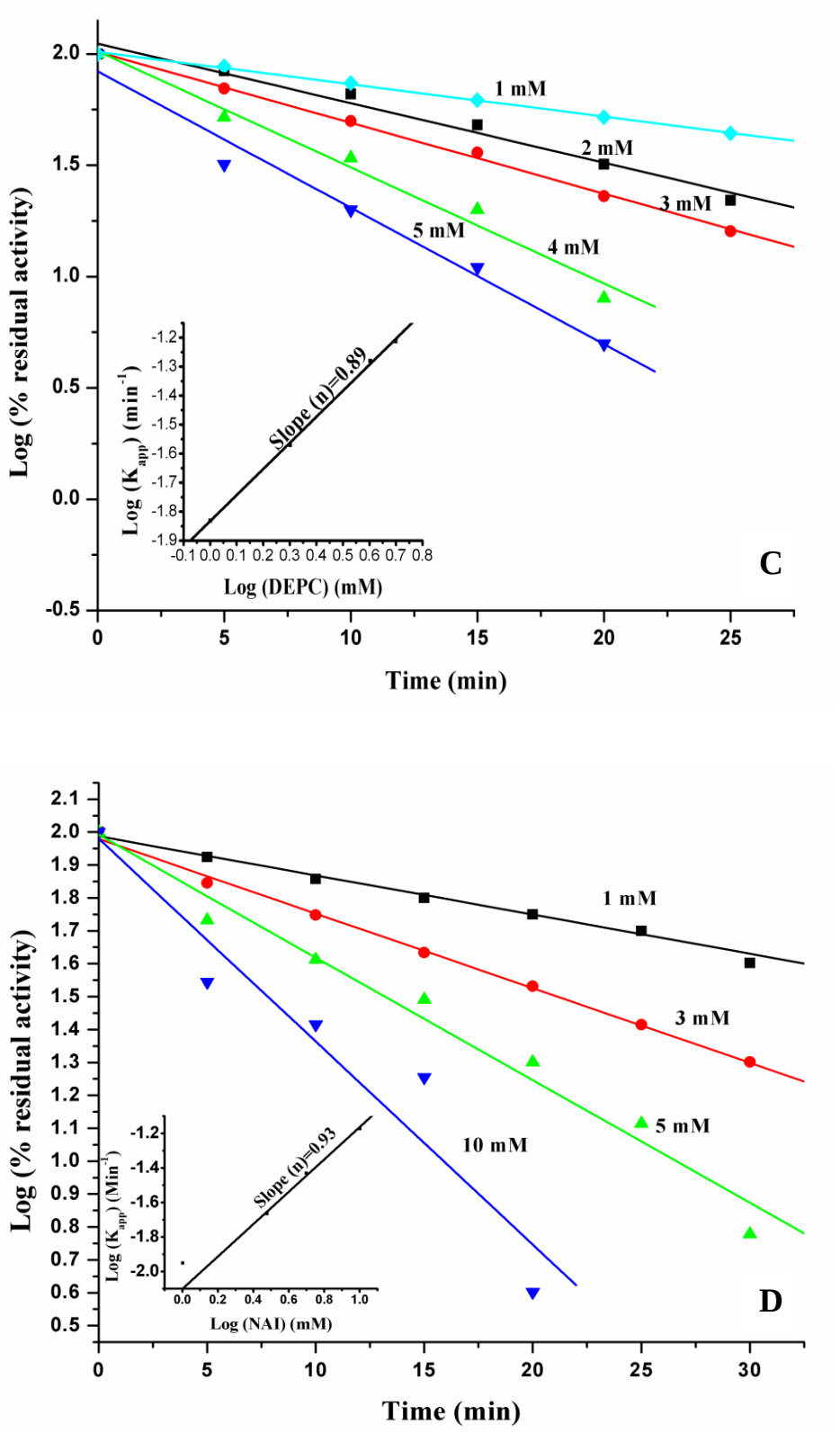


Fig. 3.7 Pseudo first order plots for inactivation of Ll-CCRH1 by (A) PMSF (B) Phenylglyoxal (C) DEPC (D) NAI. Number above lines indicates the respective

concentration of reagent. Inset: Double logarithmic plot of the first order rate constants (K_{app}) as a function of modifier concentrations.

The presence of the above amino acid residues at or near the active site was further confirmed by carrying out the chemical modification reactions in presence of substrate. The inactivation of enzyme was prevented to substantial extent by the substrate (Table 3.5). Based on substrate protection results, it has been concluded that the modified residues are present at or near the active site of Ll-CCRH1. Finally, results for active site residues of Ll-CCRH1 obtained by modeling/docking, site directed mutagenesis and chemical modification are summarized in Table 3.6.

Table 3.6 Summary for active site residues of Ll-CCRH1 obtained by modeling/docking, site directed mutagenesis and chemical modification studies. Active site residues identified by all above mentioned methods were highlighted in bold.

Methods	Active site residues
1. Modeling/docking simulations and various online softwares (CASTp, DALI server, MOE etc.)	Phe30, Ile31, **Arg51**, Leu64, **Asp77**, Val200 **Ser136**, **Tyr170**, **Lys174**, and **Ser212**
2. Site directed mutagenesis	Phe30, Ile31, **Arg51**, **Asp77** (substrate binding or stabilization) **Ser136**, **Tyr170**, **Lys174** (Catalysis) **Ser212**, His215 (Binding /catalysis)
3. Chemical modification and substrate protection	**Tyr**, **Lys**, His, **Carboxylate group**, **Arg** (one each) and **Ser** residues (two) in binding and catalysis

CHAPTER 3B

In silico mutagenesis and docking studies of active site residues of cinnamoyl CoA reductase 1 (Ll-CCRH1)

Prashant Sonawane et al. (2013) In silico mutagenesis and docking studies of active active site residues suggest altered substrate specificity and possible physiological role of cinnamoyl CoA reductase 1 (Ll-CCRH1), ***Bioinformation***, 9, 224-232.

Summary

Cinnamoyl CoA reductase (CCR) carries out the first committed step in monolignol biosynthesis and acts as a first regulatory point in lignin formation. CCR shows multiple substrate specificity towards various cinnamoyl CoA esters. Here, *in silico* mutagenesis studies of active site residues of Ll-CCRH1 were carried out. Homology modeling based modeled 3D structure of Ll-CCRH1 was used as template for *in silico* mutant preparations. Docking simulations of Ll-CCRH1 mutants with CoA esters by AutoDock Vina tools showed altered substrate specificity as compared to wild type. The study suggests that conformational changes, and change in geometry or architecture of active site pocket occurred following mutations. The altered substrate specificity for active site mutants suggests the possible physiological role of CCR either in lignin formation or in defense system in plants.

3.1 Introduction

Lignin, an integral cell wall component of plants, is a phenolic heteropolymer of monolignols namely, *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol [11]. Among several genes involved in lignin biosynthesis, cinnamoyl CoA reductase (CCR, EC 1.2.1.44) plays a key regulatory role in lignin formation [70]. Hydroxycinnamoyl CoA esters of general phenylpropanoid pathway become destined to form respective monolignols after action of CCR. Being entry point enzyme in monolignol biosynthesis, CCR diverts phenylpropanoid derived metabolites towards lignin.

CCR exhibits substrate specificity for different hydroxycinnamoyl CoA esters (p-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA); and reduces them to corresponding aldehydes. Cinnamoyl CoA esters are the common precursors of wide range of phenolic compounds and flavonoids. For

example, coumaroyl CoA esters are the substrates for chalcone synthase enzyme, the first catalytic step towards secondary metabolites synthesis. Secondary metabolites are considered as the first line of defense against pathogens and diseases. Differential substrate specificity of CCR has been correlated to its exclusive or redundant function inside the cell either in lignification (feruloyl CoA/ sinapoyl CoA as most preferred substrate) during plant development or in defense mechanism (coumaroyl CoA as favored substrate) [97,99,100].

The major limitation in understanding structure-function relationship of CCR is lack of its three dimensional structure till date by any experimental means. Homology modeling provides an alternative approach for constructing three dimensional structures if X-ray crystal structure data is not available. We constructed 3D model of Ll-CCRH1 using dihydroflavanol reductase from *Vitis vinifera* as template (PDB ID: 2c29). Putative active site residues involved in substrate binding, stabilization and catalysis were identified based on amino acid sequence analysis and docking simulations. These residues were further investigated and confirmed by site directed mutagenesis and chemical modification studies [145,150].

Here, the present study was aimed to investigate the effects of various substitution mutations (*in silico*) of active site residues on substrate specificity of Ll-CCRH1. Five different *in silico* mutants were prepared for each amino acid residue and subjected to docking simulations with different hydroxycinnamoyl CoA esters. Based on docking energies obtained, substrate preferences for each mutant were determined. These substrate specificities were used to predict the possible role of *in silico* Ll-CCRH1 mutants either in lignin formation or in defense mechanisms.

3.2 Materials and Methods

3.2.1 Starting molecule

Three dimensional model of Ll-CCRH1, generated using MODELLER 9v9, was used as template for *in silico* mutagenesis studies (Protein Model Database ID: PM0078699) [150].

3.2.2 Preparation of the *in silico* mutants

Active site of Ll-CCRH1 is made up of ten residues, Phe30, Ile31, Arg51, Asp77, Ser136, Tyr170, Lys174, Val200, Ser212 and His215. Five one-point substitution mutants were prepared for each active site residue. The three dimensional structures of *in silico* mutants were constructed by homology modeling, using the program TRITON interfaced with MODELLER [151-153]. Each mutant homology model was evaluated for its stereo chemical quality using PROCHECK [154] and also checked for environmental profile using ERRAT (Structure analysis and verification servers).

3.2.3 Substrate docking

Docking studies were carried out with five different hydroxycinnamoyl CoA esters; 4-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5- hydroxyferuloyl CoA, sinapoyl CoA similar to methods as described in chapter 3A.

3.3 Results and discussion

3.3.1 Wild type Ll-CCRH1

Docking simulations of different hydroxycinnamoyl CoA esters with Ll-CCRH1, obtained from AutoDock Vina [149] showed feruloyl CoA (-9.9 Kcal/mole) as most favored substrate over others (-8.8 to -9.7 Kcal/mole). Better affinity of Ll-CCRH1 towards feruloyl CoA indicates its possible role in lignification during growth and

development of plants [145]. Docking of various cinnamoyl CoA esters in substrate binding pocket of Ll-CCRH1 is shown in Fig. 3.1

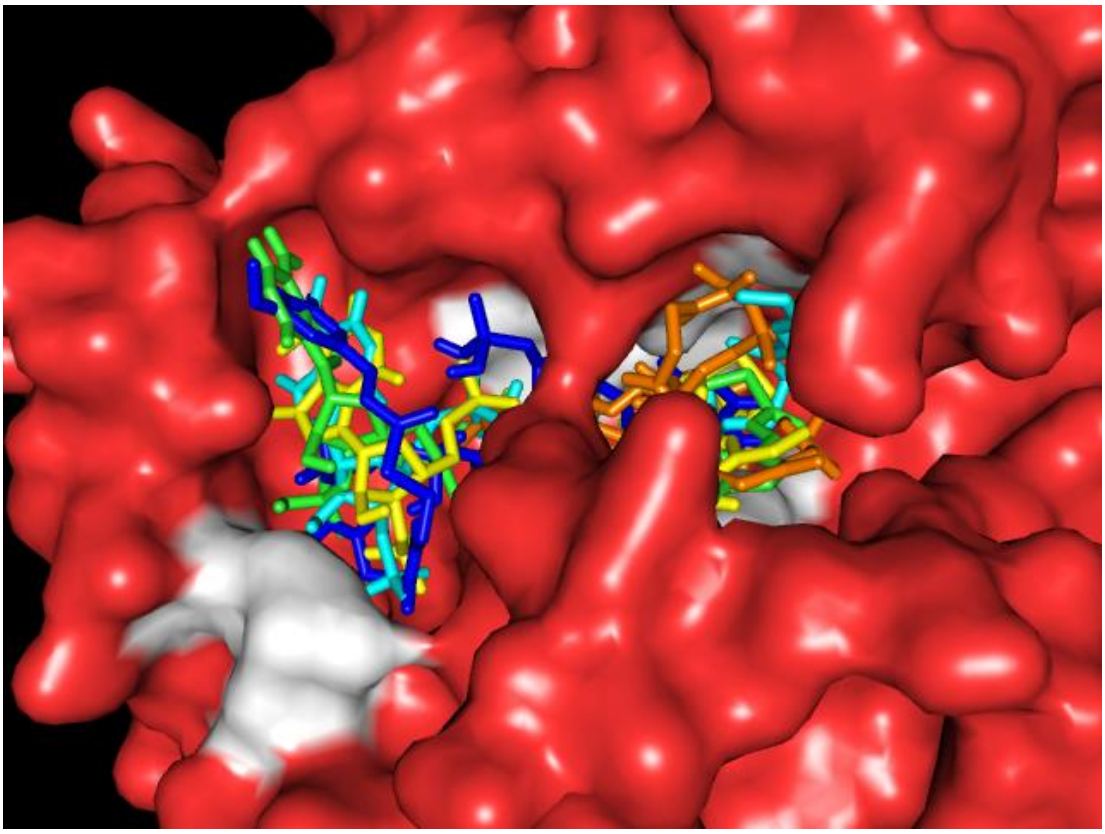


Fig. 3.1 Surface representation of the catalytic active site in Ll-CCRH1 model (close up view) and docking of cinnamoyl CoA esters in substrate binding pocket of Ll-CCRH1; caffeoyl CoA (tv green), feruloyl CoA (yellow), hydroxyferuloyl CoA (blue), coumaroyl CoA (orange) and sinapoyl CoA (cyan). Active site residues are shown in white color.

3.3.2 Phenylalanine30

Phe30 residue interacts with CoA esters via its main chain function and takes part in substrate binding or stabilization. Five mutants, namely F30C, F30L, F30S, F30V and F30Y were generated by homology modeling using TRITON software. Table 3.1 shows a comparisons of the MODELLER produced mutants with respect to RMSD.

Table 3.1: Root Mean Square Deviations (RMSD) of TRITON/MODELLER produced active site mutants of Ll-CCRH1.

Residues	Mutants	RMSD (Å)
Phe30	F30C	0.125
	F30L	0.119
	F30S	0.107
	F30V	0.122
	F30Y	0.120
Ile31	I31F	0.107
	I31M	0.144
	I31N	0.115
	I31S	0.110
	I31T	0.104
Arg51	R51G	0.153
	R51I	0.115
	R51K	0.138
	R51S	0.116
	R51T	0.099
Asp77	D77A	0.112
	D77G	0.122
	D77H	0.111
	D77N	0.104
	D77Y	0.104
Ser136	S136A	0.144
	S136C	0.120
	S136P	0.134
	S136T	0.131
	S136Y	0.134
Tyr170	Y170C	0.139
	Y170D	0.124
	Y170F	0.114
	Y170H	0.115
	Y170N	0.140
Lys174	K174E	0.153
	K174M	0.119
	K174N	0.108
	K174R	0.097
	K174T	0.123
Val200	V200E	0.128
	V200G	0.116
	V200L	0.132
	V200M	0.111
	V200A	0.116

	S212G	0.169
	S212I	0.114
Ser212	S212N	0.139
	S212R	0.130
	S212T	0.121
	H215D	0.117
	H215L	0.108
His215	H215Q	0.139
	H215R	0.125
	H215Y	0.110

The RMSD values for phe30 mutants lay between 0.107-0.125 Å. Docking studies showed that F30S and F30V showed feruloyl CoA as favored substrate; while mutants F30L, F30S and F30Y showed preference for coumaroyl CoA (Table 3.2, Fig. 3.2). The main chain function of all mutants was expected to remain same even after mutation. Thus, there should be conformational alterations or change in geometry of active site following mutations resulting in differential substrate specificity (Fig. 3.3). Mutants F30C and F30V may play a role in lignin biosynthesis; while F30S, F30Y and F30L could be important for defense cascades [97,99,100].

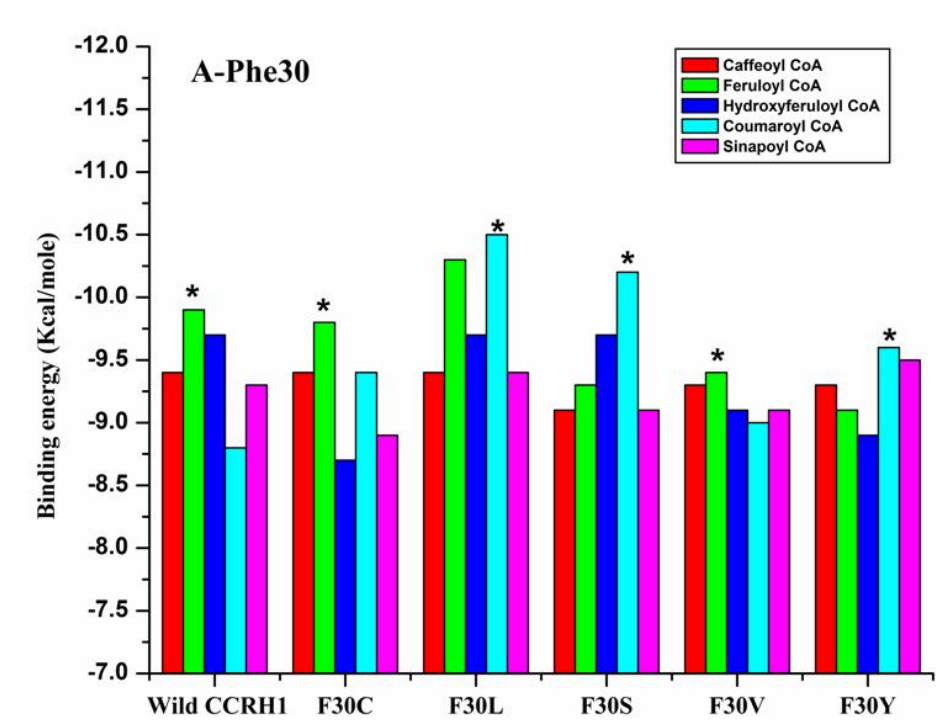


Fig. 3.2 Cinnamoyl CoA esters binding energy change for *in silico* phe30 active site mutants. X-axis represents Ll-CCRH1 amino acid mutants generated by homology modeling, using TRITON/MODELLER software, and Y-axis is respective binding energies for CoA esters docked with each individual mutant. The '*' shows best docked substrate with most negative binding energy among particular mutant.

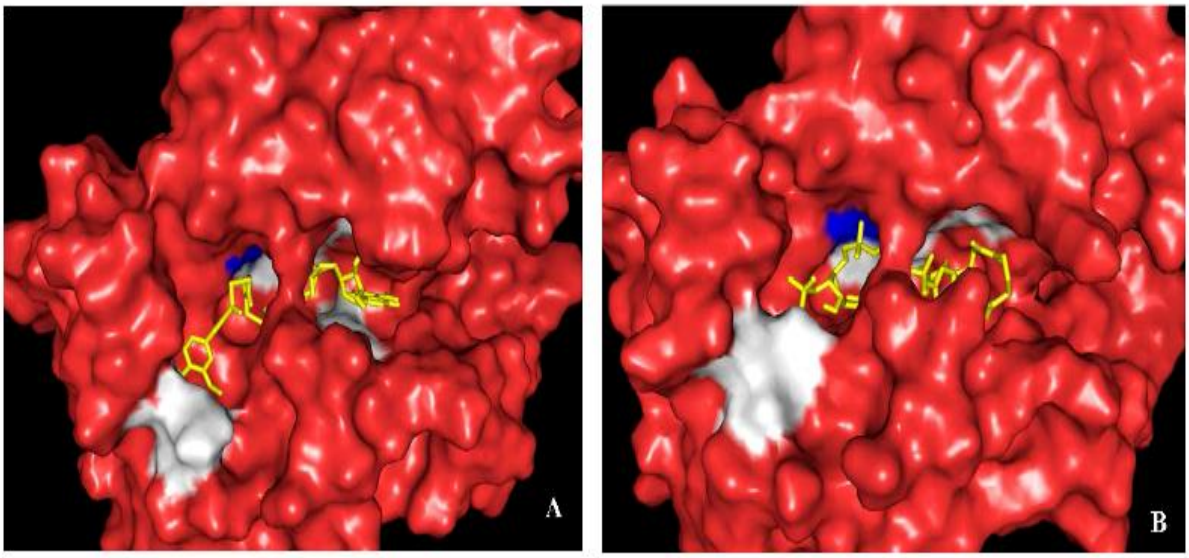


Fig. 3.3 The best docked structures of cinnamoyl CoA esters into the active site of *in silico* Ll-CCRH1 mutants at various positions. Mutated Ll-CCRH1 molecule (red) is represented as solid surface, whereas docked structure is shown as sticks. Mutated amino acid is displayed in blue color and remaining active site residues are shown in white. (A) F30C mutant, feruloyl CoA (B) F30Y mutant, coumaroyl CoA**.**

3.3.3 Isoleucine31

Ile31 residue also plays role in substrate binding through its main chain function. Mutant I31F showed more negative binding energy for hydroxyferuloyl CoA; while I31M exhibited equal affinity for coumaroyl and hydroxyferulol CoA. Mutant I31N demonstrated better affinity for sinapoyl CoA over others. Substitution with serine and threonine (I31S and I31T) resulted in feruloyl CoA as preferred substrate with equal

binding energy (-10 Kcal/mole) (Fig. 3.4, Table 3.2). *In silico* mutations of Ile31 lead to overall change in architecture of active site pocket (Fig. 3.5). Slightly higher RMSD value was observed for mutant I31M compared to other Ile31 mutants (Table 3.1).

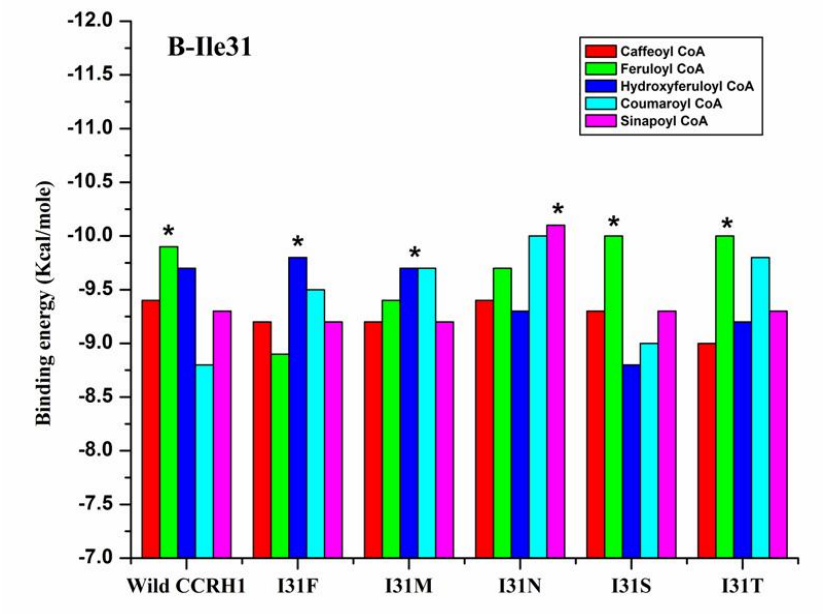


Fig. 3.4 Cinnamoyl CoA esters binding energy change for *in silico* Ile31 active site mutants. The '*' indicates best docked substrate with most negative binding energy among particular mutant.

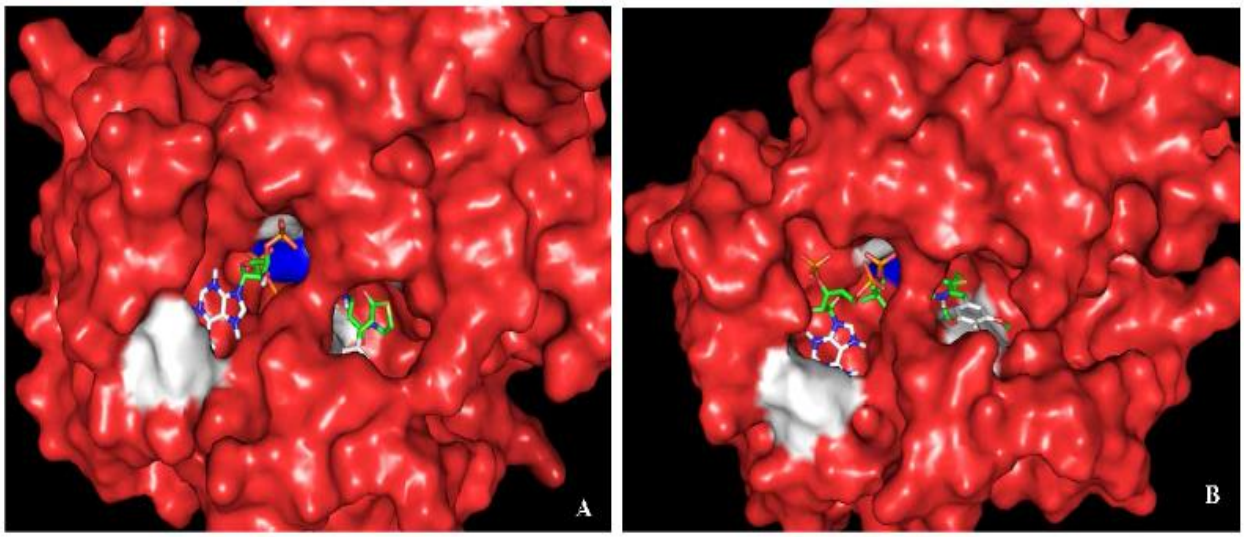


Fig. 3.5 The best docked structures of cinnamoyl CoA esters with Ile31 mutants. (A) I31F mutant, coumaroyl CoA (B) I31N mutant, sinapoyl CoA.

3.3.4 Arginine51

In wild type Ll-CCRH1, Arg51 interacts with CoA esters via its main chain as well as side chain and plays important role in substrate binding or stabilization. In mutant R51G, coumaroyl CoA showed most favorable binding energy (-9.7 Kcal/mole) (Table 3.2, Fig. 3.6). Substitution of Arg51 by glycine altered side chain from polar charged to small compact neutral residue, resulting in marked decrease in accessible surface area and leads to loss of interactions with substrate. Similar decrease in surface area was observed with mutants R51I, R51S and R51T except for R51K, which showed similar area as that of wild type (Fig. 3.7). Except R51K, all three remaining mutants have shown better preference for coumaroyl CoA; and thus could play role in defense. Mutant R51K showed affinity towards caffeoyl CoA (Table 3.2) (Fig. 3.6 and 3.7). Thus, all mutants R51G, R51I, R51S, R51K and R51T may function in defense [97,99].

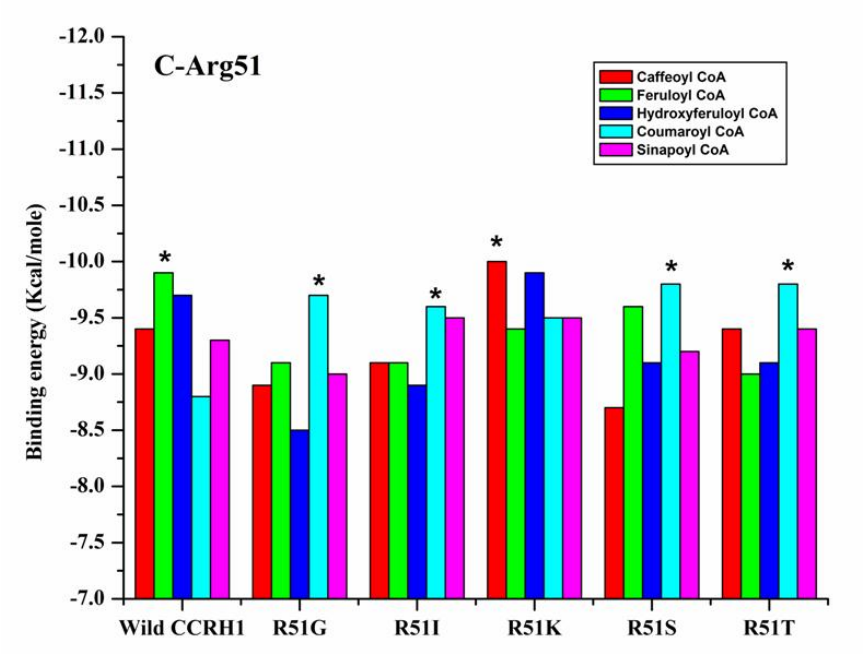


Fig. 3.6 Docking energies obtained for *in silico* Arg51 active site residue mutants. The '*' shows best docked substrate with most negative binding energy among particular mutant.

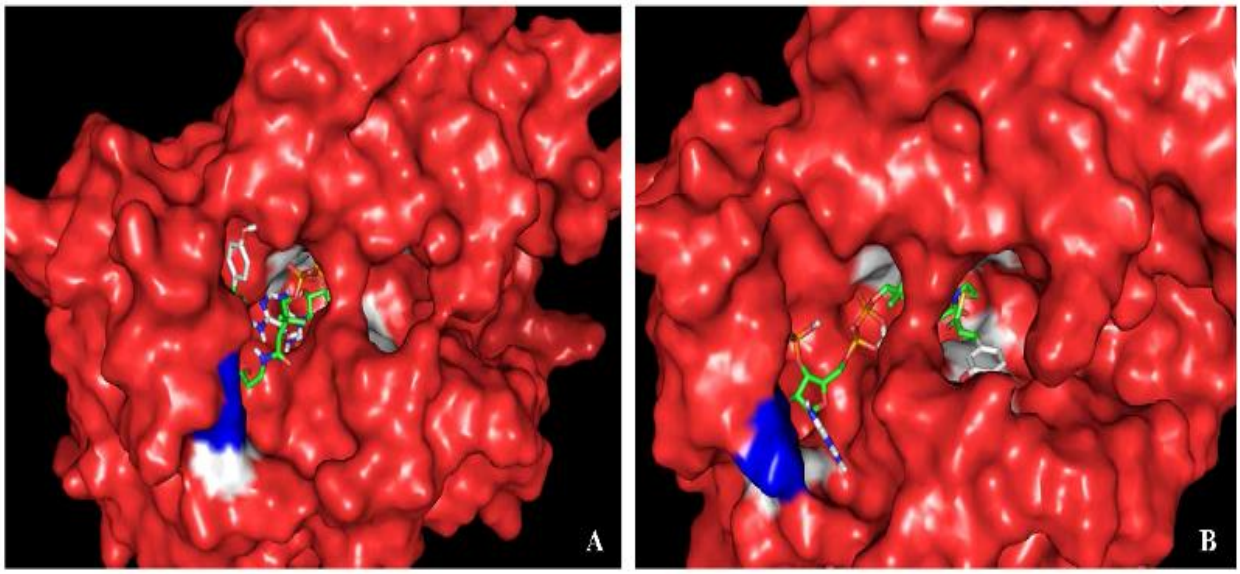


Fig. 3.7 Docking of cinnamoyl CoA esters with Arg51 mutants. Mutated Ll-CCRH1 molecule (red) is represented as solid surface, whereas docked structure is shown as sticks. Mutated amino acid is displayed in blue color and remaining active site residues are shown in white (A) R51I mutant, coumaroyl CoA (B) R51K mutant, caffeoyl CoA.

3.3.5 Aspartate77

No Asp77 interactions were observed during docking of Ll-CCRH1 with different CoA esters; but site directed mutagenesis and chemical modification confirmed its role in CCR catalyzed reaction [150]. D77 is present partially on surface and in proximity of R51. Substitution with alanine and glycine resulted in slight decrease in surface areas, but substrate affinities differ for both these mutants; D77A showed specificity towards coumaroyl CoA while D77G displayed equal affinity for Caffeoyl CoA and sinapoyl CoA (Fig. 3.8, Table 3.2). This may be due to change in structure of binding pocket and allowing better conformations to interact with other amino acids. Mutant D77H showed more negative binding energy for sinapoyl CoA (-9.7 Kcal/mol). D77N has same substrate affinity (feruloyl CoA) as that of wild type (Fig. 3.8, Table 3.2). In mutation

D77Y, change from D to Y altered small polar side chain to bulky hydrophobic aromatic ring, induced a significant increase in accessibility for surface area (Fig. 3.9). RMSD values for all mutants are very comparable to each other (0.104-0.122 Å) (Table 3.1).

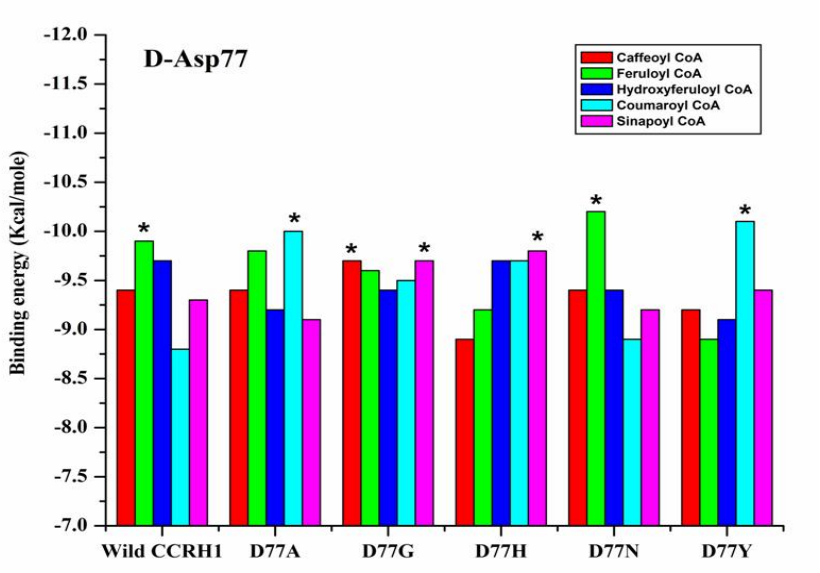


Fig. 3.8 Summary of results for docking simulations of Asp77 mutants and different CoA esters. The '*' shows best docked substrate with most negative binding energy among particular mutant.

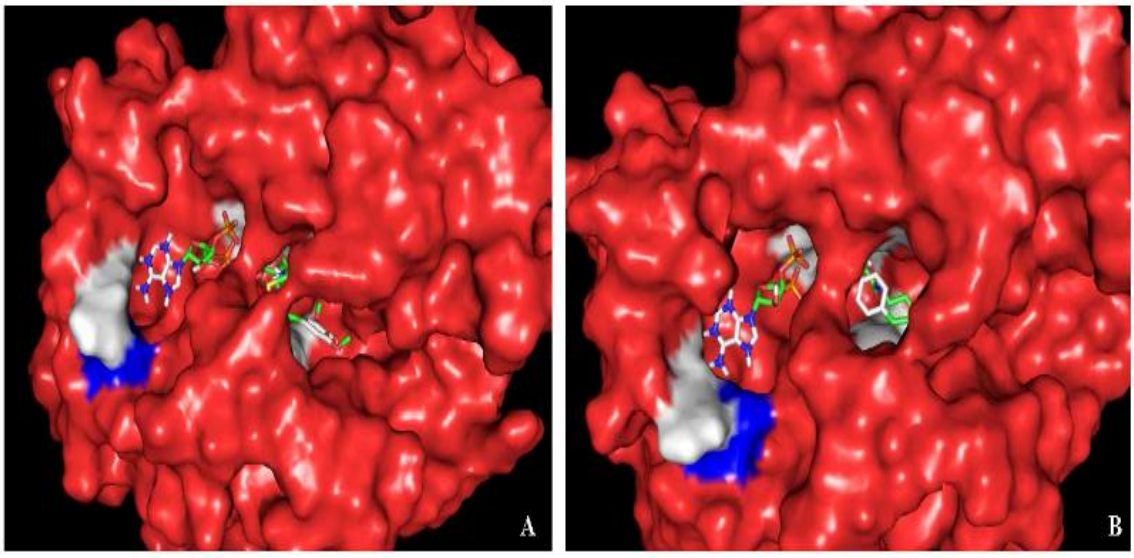


Fig. 3.9 Cinnamoyl CoA esters conformations (stucks) in the active site pocket of

Asp77 mutants (surface, red). Color code for mutated and remaining active site residue is same as mentioned in Fig. 3.7. (A) D77H mutant, sinapoyl CoA (B) D77Y mutant, coumaroyl CoA.

3.3.6 Serine136

In wild type Ll-CCRH1, Ser136 plays a key role in catalysis and is a part of reaction centre. Mutations in ser136 caused change in the geometry of active site and non specific interactions with substrates have been increased. Ser136 is buried in the substrate binding pocket and mutations resulted in partial or complete exposure of mutant residue. Coumaroyl CoA was found to be better substrate for mutant S136A and S136C. S136Y showed preference for caffeoyl CoA and, S136P along with S136T showed favored specificity for feruloyl CoA (Table 3.2) (Fig. 3.10, 3.11). Mutants S136P and S136T could have functional role in lignin formation; while mutants S136Y, S136A and S136C might be involved in defense [97,99,100].

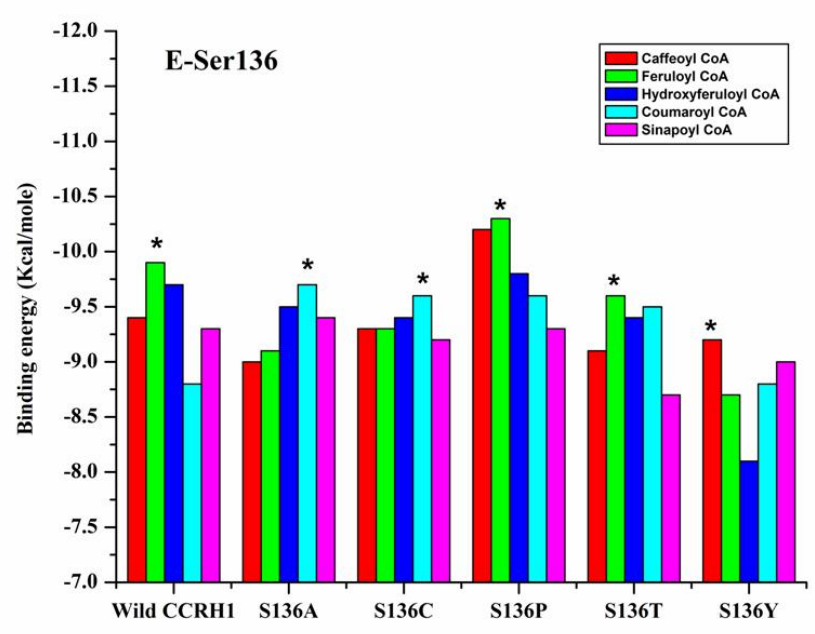


Fig. 3.10 Cinnamoyl CoA esters binding energy change for *in silico* Ser136 mutants.

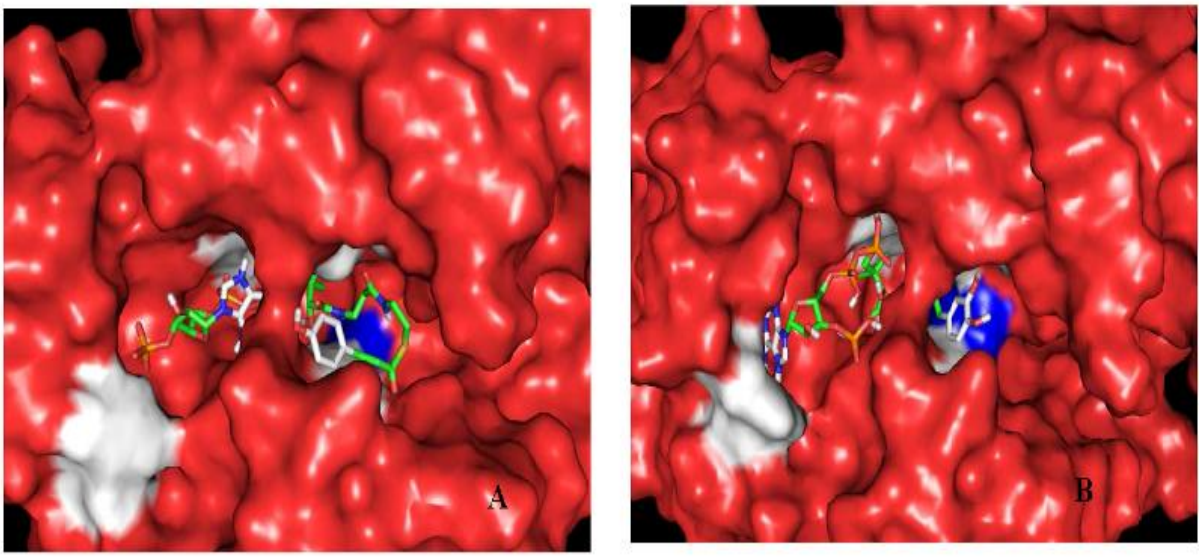


Fig. 3.11 Docking of Cinnamoyl CoA esters in the active site of Ser136 mutants (surface, red). (A) S136C mutant, coumaroyl CoA (B) S136Y mutant, caffeoyl CoA.

3.3.7 Tyrosine170

Tyr170 acts as catalytic base and accepts hydride from NADPH and transfers to Serine residues in catalysis reaction. Tyr170 is completely buried in active site pocket and is surrounded by Lys174, Ser212, His215 and Ser136 catalytic residues. Mutant Y170C displayed less number of interactions compared to wild type; still sufficient for displaying affinity towards coumaroyl CoA. Mutant Y170D exhibited better binding energy for feruloyl CoA (-9.7 Kcal/mol) (Fig. 3.12). In case of mutant Y170D, increase in number of hydrogen bond interactions with Arg51 was observed. Arg51 is distantly present from Y170D. This indicates the drastic change in architecture of binding pocket and significant changes in pKa values of pocket. Y170F and Y170N showed preference for coumaroyl CoA; while Y170H demonstrated affinity for feruloyl CoA (Table 3.2, Fig. 3.13).

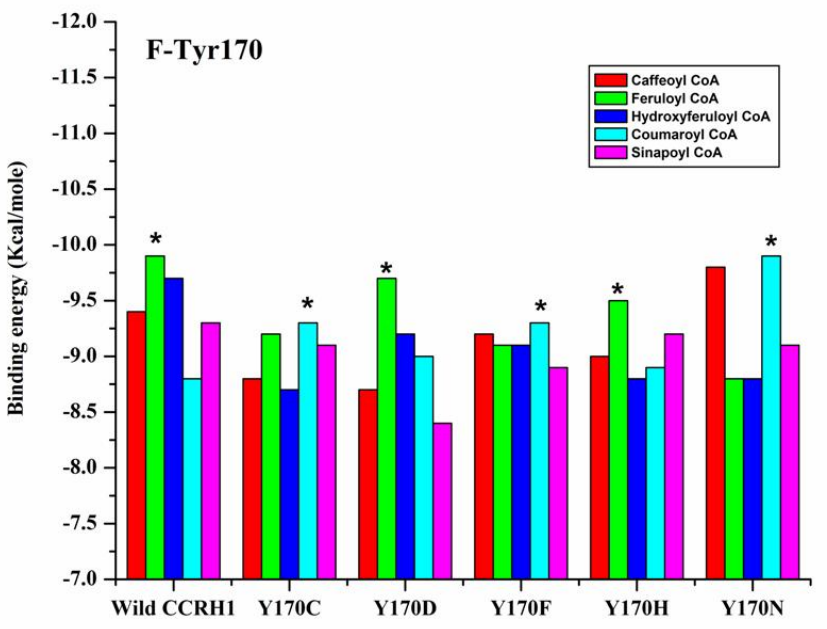


Fig. 3.12 Docking summary of Tyr170 mutants and cinnamoyl CoA esters.

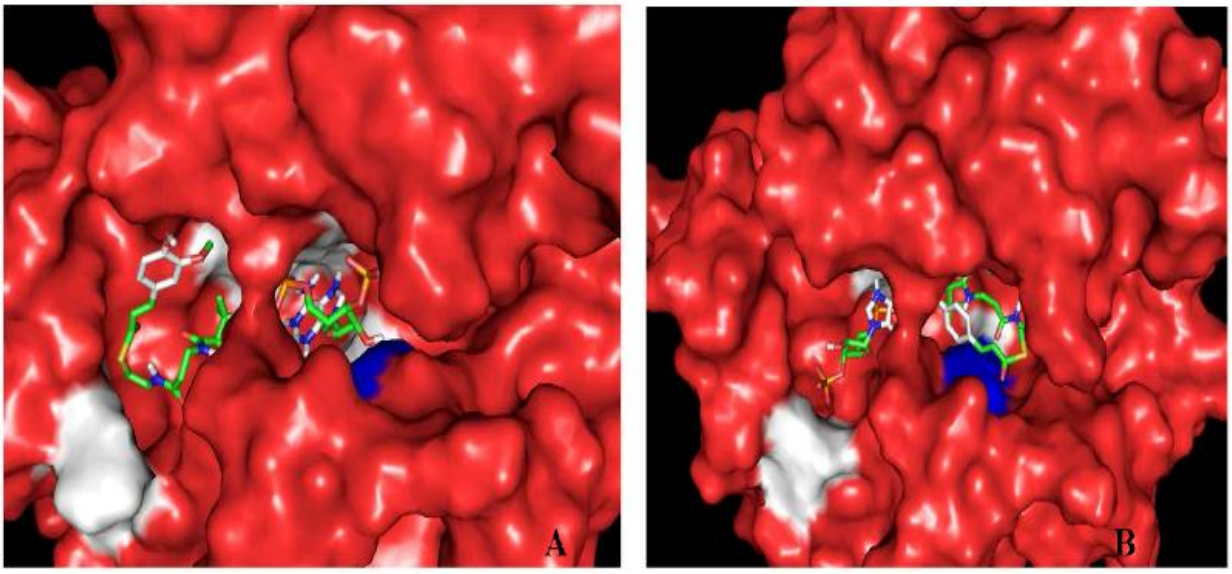


Fig. 3.13 Cinnamoyl CoA esters conformations (sticks) in the active site of mutants Ll-CCRH1 (surface, red). (A) Y170F mutant, coumaroyl CoA (B) Y170H mutant, coumaroyl CoA.

3.3.8 Lysine174

Lys174 residue promotes hydride transfer in CCR mediated reduction reactions. Docking simulations of K174E and K174T mutants showed coumaroyl CoA as preferred substrate (-10 Kcal/mole) (Fig. 3.14 and Table 3.2). On the other hand, K174M mutant is specific for feruloyl CoA; while K174N and K174R have favorable binding energy for hydroxyferuloyl CoA (Fig. 3.14). Thus, mutant K174R may play role in both, either in lignin biosynthesis or defense system [97,99]. Lys174 is deeply buried in active site pocket and same conformational profile was also observed in all mutants (Fig. 3.15). Mutant K174R shows the lowest RMSD (0.097 Å) among the all mutants generated (Table 3.1).

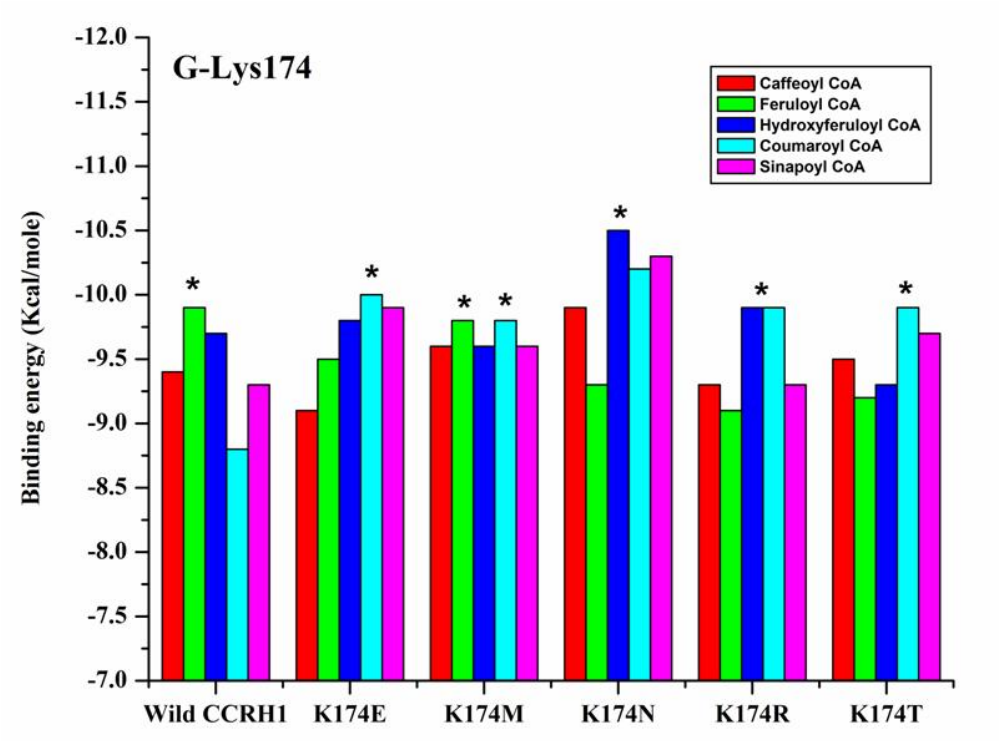


Fig. 3.14 Summary of results for docking simulations of Lys174 mutants and different CoA esters.

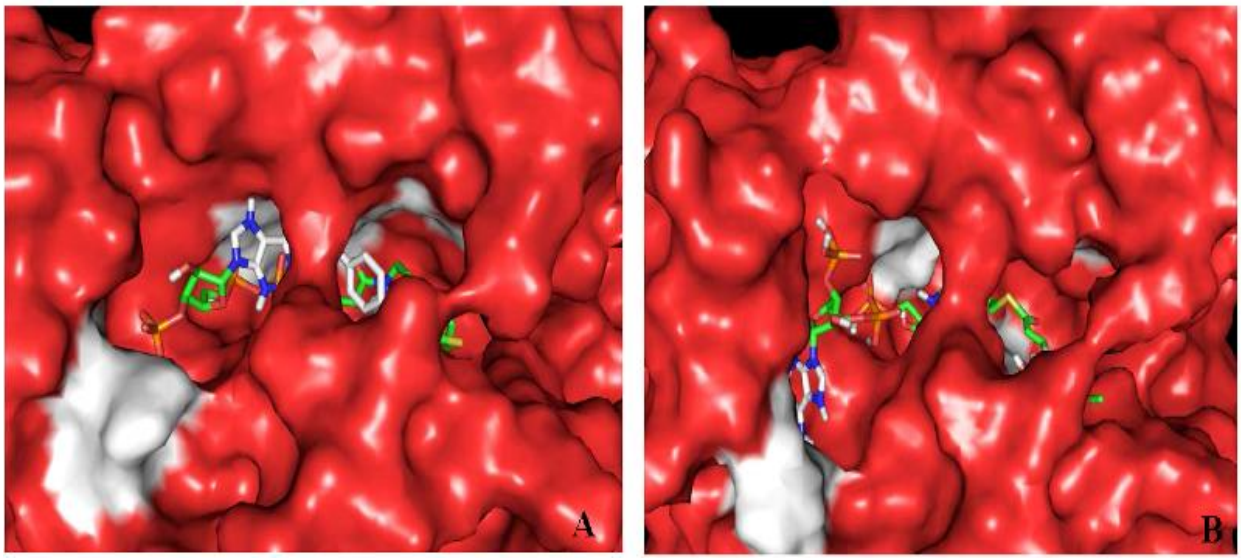


Fig. 3.15 The best docked structures of cinnamoyl CoA esters into the active site of *in silico* Lys174 mutants. (A) K174E mutant, coumaroyl CoA (B) K174N mutant, hydroxyferuloyl CoA. Lysine174 residue is deeply buried inside the binding pocket, thus it is not visible in A, B surface diagrams.

3.3.9 Valine200

Replacement of Val200 with glycine and alanine displayed substrate specificity towards coumaroyl CoA. These substitutions have not altered aliphatic side chain profile of mutant residues. On the other hand, substitution of valine by positively charged aspargine resulted in charge redistribution along the active site pocket and allowed favorable conformational changes for substrate binding (feruloyl CoA). V200M mutant exhibited increased affinity for coumaroyl CoA (Table 3.2, Fig. 3.16 and 3.17). V200E mutant could possibly be involved in monolignol biosynthesis. Mutants V200G, V200M, V200A may take part in secondary metabolite synthesis [97,99].

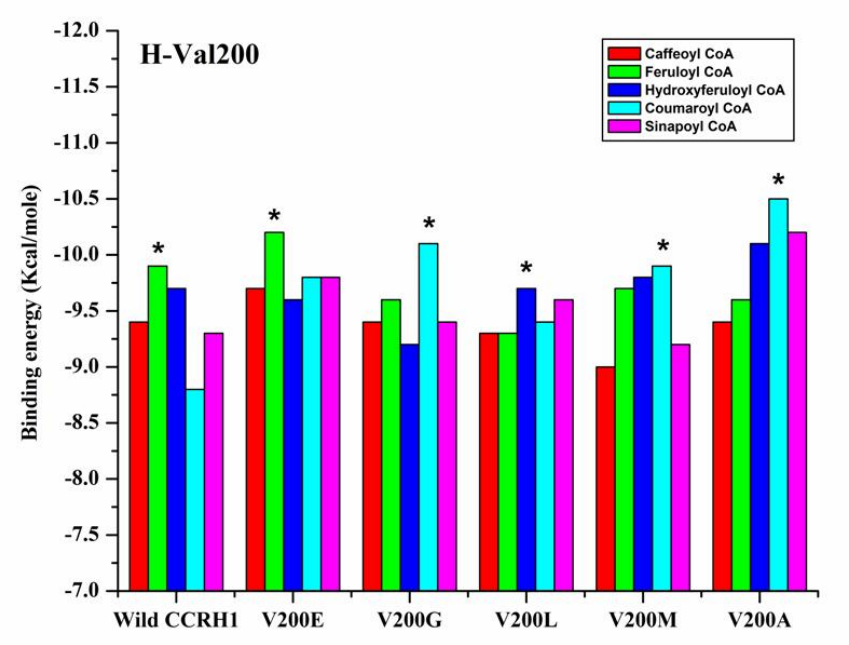


Fig. 3.16 Docking energies obtained for *in silico* Val200 active site residue mutants.

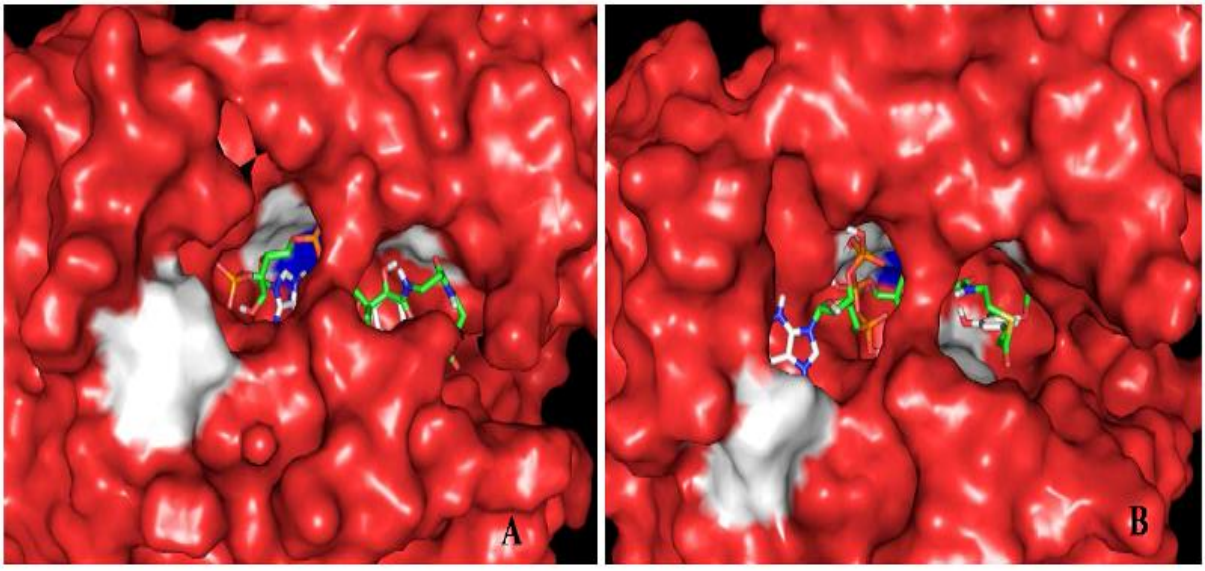


Fig. 3.17 Cinnamoyl CoA esters conformations (sticks) in the active site of Val200 mutants Ll-CCRH1 (surface, red). (A) V200L mutant, hydroxyferuloyl CoA (B) V200M mutant, coumaroyl CoA.

3.3.10 Serine212

Ser212 residue is involved in proton shuttle mechanism and thus participates in CCR catalysis. Mutations in Ser residue (deeply buried) by Gly, Ile, Glu and Arg showed partial or complete exposure of mutated residues. These exposed mutated residues altered substrate binding conformation and favored coumaroyl CoA as substrate (Fig. 3.18 and 3.19 A, B). Thus these structural changes altered shape of active site pocket and assisted more number of interacting residues with maximum interactions. S212T has shown feruloyl CoA as promising substrate (Fig. 3.18 and 3.19 C) (Table 3.2). All mutants of ser212 except S212T may function in secondary metabolism, ultimately in providing defense [97,99,100].

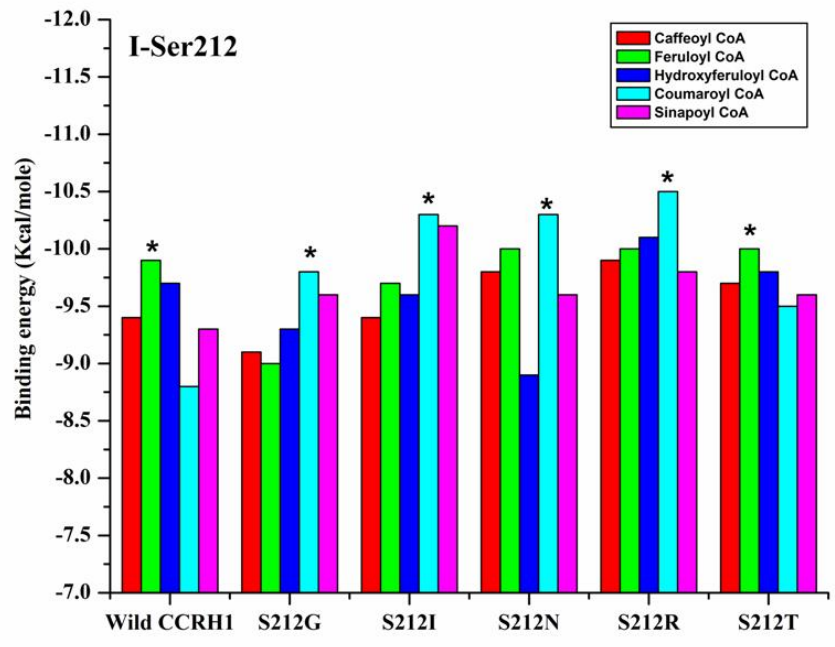


Fig. 3.18 Docking summary of Ser212 mutants and CoA esters.

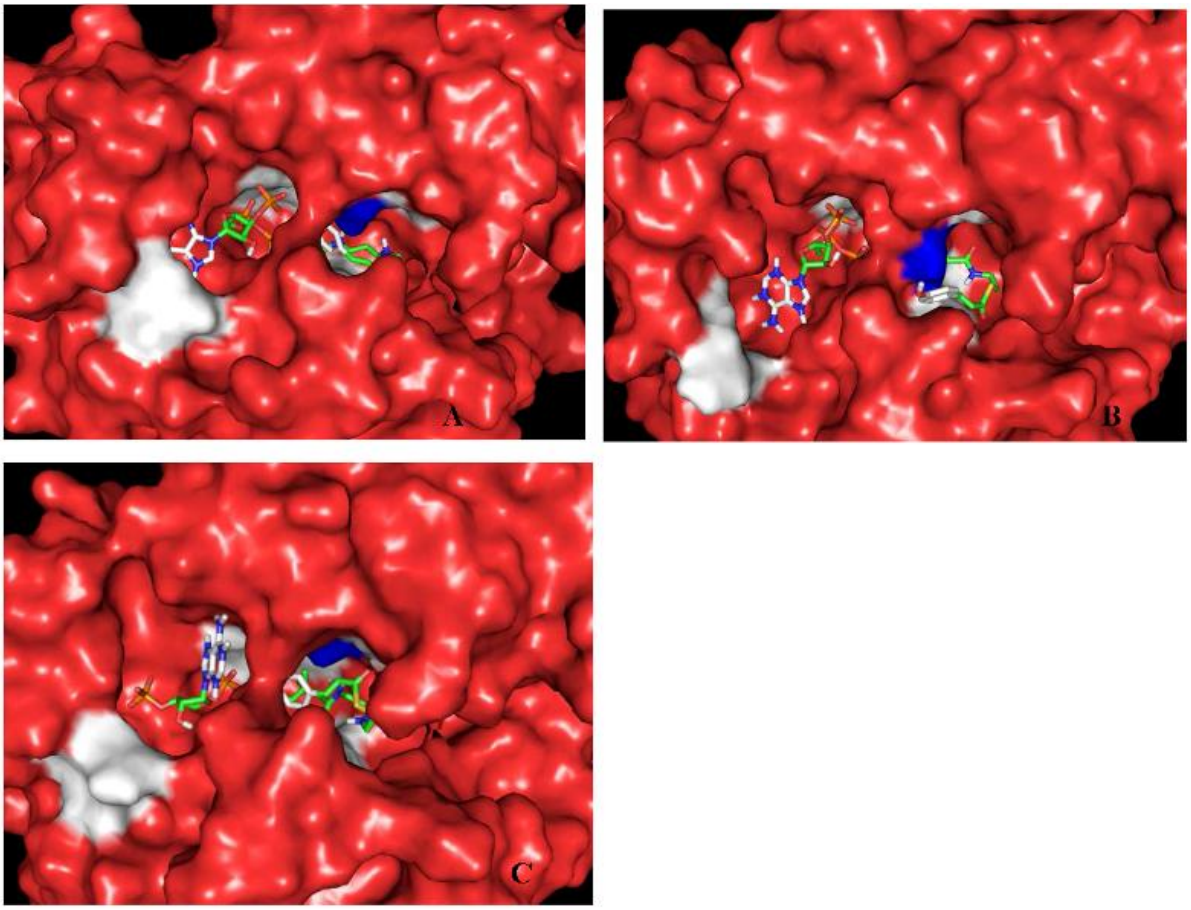


Fig. 3.19 Docking studies of cinnamoyl CoA esters in active site of Ser212 mutants. A) S212I mutant, coumaroyl CoA (B) S212R mutant, coumaroyl CoA (C) S212T mutant, feruloyl CoA.

3.3.11 Histidine215

His215 also takes part in CCR mediated reduction reactions in substrate catalysis. H215R and H215Y mutants showed significant increase in surface area due to bulky and long side chains. Both mutants showed same binding energy (-10.5 Kcal/mol), but for different substrates; that is, coumaroyl CoA is specific for H215R and hydroxyferuloyl CoA for H215Y (Fig. 3.20 and 3.21 B,C). Replacement of His215 by Asp and Gln demonstrated affinity for coumaroyl CoA (Fig. 3.20). Lastly, replacement of His with Leu displayed preference for sinapoyl CoA (Table 3.2) (Fig. 3.20 and 3.21

A). Thus only H215L mutant may possibly take part in lignification; while remaining mutants might prefer secondary metabolite pathway [97,99].

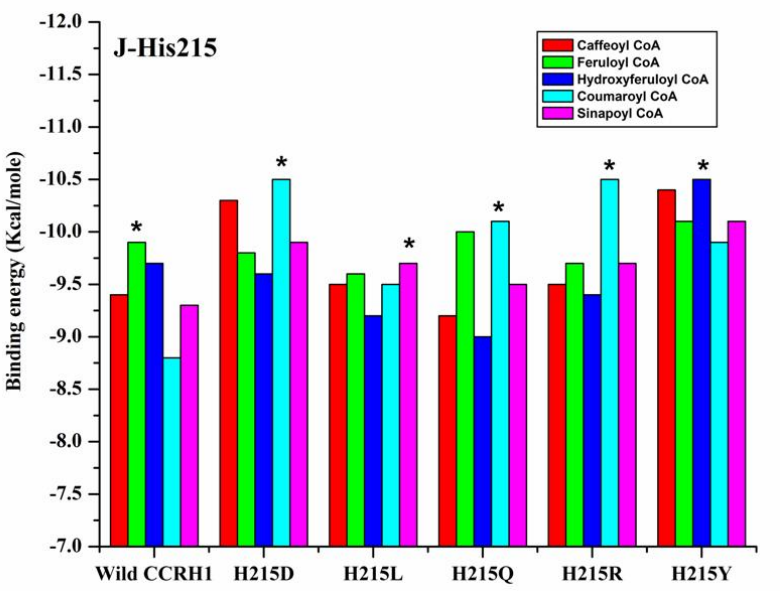


Fig. 3.20 Cinnamoyl CoA esters binding energy change for *in silico* His215 mutants.

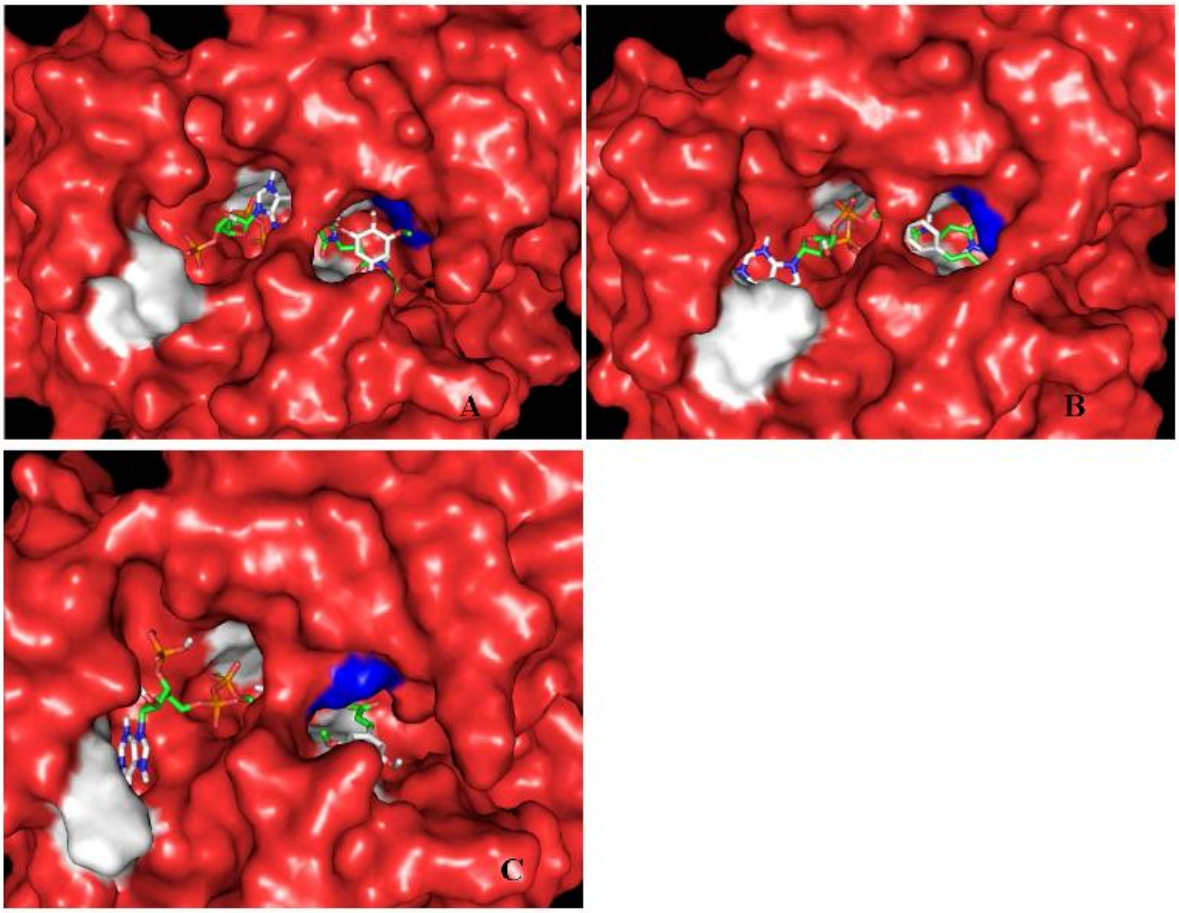


Fig. 3.21 Cinnamoyl CoA esters conformations (sticks) in the active site of His215 mutants Ll-CCRH1 (surface, red). (A) H215L mutant, sinapoyl CoA (B) H215R mutant, coumaroyl CoA (C) H215Y mutant, hydroxyferuloyl CoA.

Table 3.2 Summary from docking simulations of Ll-CCRH1 mutants with five different cinnamoyl CoA esters. For each mutant, more negative binding energy (preferred substrate) is indicated in red color.

Active site residues in Ll-CCRH1	Mutants Ll-CCRH1	Binding Energies (Kcal/mole)				
		Caffeoyl CoA	Feruloyl CoA	Hydroxy-Feuloyl CoA	Coumaroyl CoA	Sinapoyl CoA
Phe30	F30C	-9.4	-9.8	-8.7	-9.4	-8.9
	F30L	-9.4	-10.3	-9.7	-10.5	-9.4
	F30S	-9.1	-9.3	-9.7	-10.2	-9.1
	F30V	-9.3	-9.4	-9.1	-9.0	-9.1
	F30Y	-9.3	-9.1	-8.9	-9.6	-9.5
Ile31	I31F	-9.2	-8.9	-9.8	-9.5	-9.2
	I31M	-9.2	-9.4	-9.7	-9.7	-9.2
	I31N	-9.4	-9.7	-9.3	-10.0	-10.1
	I31S	-9.3	-10.0	-8.8	-9.0	-9.3
	I31T	-9.0	-10.0	-9.2	-9.8	-9.3
Arg51	R51G	-8.9	-9.1	-8.5	-9.7	-9.0
	R51I	-9.1	-9.1	-8.9	-9.6	-9.5
	R51K	-10.0	-9.4	-9.9	-9.5	-9.5
	R51S	-8.7	-9.6	-9.1	-9.8	-9.2
	R51T	-9.4	-9.0	-9.1	-9.8	-9.4
Asp77	D77A	-9.4	-9.8	-9.2	-10.0	-9.1
	D77G	-9.7	-9.6	-9.4	-9.5	-9.7
	D77H	-8.9	-9.2	-9.7	-9.7	-9.8
	D77N	-9.4	-10.2	-9.4	-8.9	-9.2
	D77Y	-9.2	-8.9	-9.1	-10.1	-9.4
Ser136	S136A	-9.0	-9.1	-9.5	-9.7	-9.4
	S136C	-9.3	-9.3	-9.4	-9.6	-9.2
	S136P	-10.2	-10.3	-9.8	-9.6	-9.3
	S136T	-9.1	-9.6	-9.4	-9.5	-8.7
	S136Y	-9.2	-8.7	-8.1	-8.8	-9.0
Tyr170	Y170C	-8.8	-9.2	-8.7	-9.3	-9.1
	Y170D	-8.7	-9.7	-9.2	-9.0	-8.4
	Y170F	-9.2	-9.1	-9.1	-9.3	-8.9
	Y170H	-9.0	-9.5	-8.8	-8.9	-9.2
	Y170N	-9.8	-8.8	-8.8	-9.9	-9.1
Lys174	K174E	-9.1	-9.5	-9.8	-10.0	-9.9

	K174M	-9.6	-9.8	-9.6	-9.8	-9.6
	K174N	-9.9	-9.3	-10.5	-10.2	-10.3
	K174R	-9.3	-9.1	-9.9	-9.9	-9.3
	K174T	-9.5	-9.2	-9.3	-9.9	-9.7
Val200	V200E	-9.7	-10.2	-9.6	-9.8	-9.8
	V200G	-9.4	-9.6	-9.2	-10.1	-9.4
	V200L	-9.3	-9.3	-9.7	-9.4	-9.6
	V200M	-9.0	-9.7	-9.8	-9.9	-9.2
	V200A	-9.4	-9.6	-10.1	-10.5	-10.2
Ser212	S212G	-9.1	-9.0	-9.3	-9.8	-9.6
	S212I	-9.4	-9.7	-9.6	-10.3	-10.2
	S212N	-9.8	-10.0	-8.9	-10.3	-9.6
	S212R	-9.9	-10.0	-10.1	-10.5	-9.8
	S212T	-9.7	-10.0	-9.8	-9.5	-9.6
His215	H215D	-10.3	-9.8	-9.6	-10.5	-9.9
	H215L	-9.5	-9.6	-9.2	-9.5	-9.7
	H215Q	-9.2	-10.0	-9.0	-10.1	-9.5
	H215R	-9.5	-9.7	-9.4	-10.5	-9.7
	H215Y	-10.4	-10.1	-10.5	-9.9	-10.1

CHAPTER 4

Conformational transitions of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*

Prashant Sonawane, Bashir Khan, Sushama Gaikwad (2013) Cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*: Acid induced, thermostable molten globule, (Under review).

Summary

Conformational transitions of cinnamoyl CoA reductase; a key regulatory enzyme in lignin biosynthesis, from *Leucaena leucocephala* (Ll-CCRH1) were studied using fluorescence and circular dichroism spectroscopy. The native protein, with four trp residues exposed on the surface and 66% helical structure, undergoes rapid structural transitions at and above 45 °C and starts forming aggregates at 55 °C. Ll-CCRH1 was transformed into acid induced (pH 2.0) molten globule like structure, exhibiting altered yet compact secondary structure, diminished tertiary structure and exposed hydrophobic residues. The molten globule like structure was examined for the thermal and chemical stability. The altered secondary structure of L1-CCRH1 at pH 2.0 was stable upto 90 °C and also in the vicinity of 2 M Guanidine hydrochloride (GdnHCl) (as compared to drastic loss of native structure in 2 M GdnHCl) as seen in far UV-CD spectra. The structural transition of Ll-CCRH1 at pH 2.0 was reversible, as all the characteristics of molten globule had diminished after readjusting the pH to 8.0, that of native protein.

4.1 Introduction

Leucaena leucocephala, a multipurpose nitrogen fixing tropical legume tree, has a tremendous prospective as a raw material for paper and pulp industry, nutritious forage, timber, organic fertilizer, firewood, industrial fuel and depilatory agent due to its adaptability to thrive under farthest agro-climatic conditions [3-5]. Although lignin plays a major role in growth and development of plants, an agro-economical opinion considers lignin as an obstacle for utilization of plant biomass as it affects many industrial processes [1]. The biosynthesis of lignin precursors proceeds through common phenylpropanoid pathway, starting with conversion of phenylalanine to

cinnamate and subsequent formation of hydroxycinnamoyl CoA esters [11]. These esters are the end products of common phenylpropanoid metabolism. In monolignol biosynthesis, cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the NADPH dependent reduction of cinnamoyl CoA esters to corresponding cinnamaldehydes. As the first committed step in monolignol biosynthesis, CCR plays a key regulatory role by controlling the overall carbon flux of metabolites towards lignin [70].

The characterization of unfolded and partly folded conformations is vital to understand the principles governing protein folding and its stability [124,155]. Many studies have shown that protein folding is a discrete pathway with distinct intermediate states between native and denatured one. These intermediate states have been observed under different conditions from native or completely unfolded states and analysis of these intermediates provides an insight into protein folding pathway [126,130,133].

Although CCR is one of the most investigated enzymes in lignin biosynthesis pathway, its three dimensional crystal structure remains to be determined [27,87,89,91,95,99,108]. Also, hardly any reports on CCR conformation and structure-function relationship are available. The recombinant Ll-CCRH1 from *L. leucocephala* has already been purified and characterized by us [145,150]. The enzyme was able to catalyze reduction of CoA ester moieties which had OH, OCH_3 substituents at 3 and 5 positions with different rates. Also, the enzyme showed stringent conditions of pH optimum for forward (pH 6.5) and reverse (pH 7.8) catalysis. In the present study, as a step towards structural characterization, the Ll-CCRH1 was subjected to thermal, chemical as well as acid and alkali induced denaturation. Structural transitions in Ll-CCRH1 were monitored using biophysical techniques to understand its conformational stability. Here, we also report the existence and characterization of an acid induced molten globule like intermediate at pH 2.0.

4.2 Materials and Methods

Materials

Guanidine hydrochloride (GdnHCl) and 1-anilino-8-napthalenesulfonate (ANS) were obtained from Sigma, USA. All other reagents were of analytical grade. Solutions prepared for spectroscopic measurements were in sterile MilliQ water.

4.2.1 Heterologous expression and purification of recombinant CCRH1 from *L .leucocephala*

Heterologous expression, purification and assay of the Ll-CCRH1 enzyme with cinnamoyl coA esters were carried out as described in chapter 2, section 2.2.1 and 2.2.3 [145].

4.2.2 Fluorescence measurements

Intrinsic fluorescence of the Ll-CCRH1 enzyme was measured at 30 °C using a Perkin-Elmer LS 50B spectrofluorimeter connected to a Julabo F20 circulating water bath. The protein solution (0.47 µM) was excited at 295 nm and emission was recorded in the range of 310-400 nm. Both the excitation and emission spectra were obtained by setting the slit width at 7 nm, and a scan speed of 100 nm/min. To eliminate the background emission, the signal produced by either buffer solution, or buffer containing the desired quantity of denaturants was subtracted.

4.2.3 Decomposition analysis of fluorescence spectra

The decomposition analysis of trp fluorescence spectra was carried out using PFAST program (http://pfast.phys.uri.edu/pfast/) based on the SIMS and PHREQ algorithm as described elsewhere [156].

4.2.4 Circular Dichroism (CD) measurements

The CD spectra of the enzyme were recorded on a Jasco 815-1505 (Jasco, Tokyo, Japan) spectropolarimeter connected to a Peltier Type CD/FL cell circulating water bath (Jasco, Tokyo, Japan) at 25 °C. Far UV CD measurements of protein (54 µg/ml) were recorded in the wavelength range of 190-250 nm with a 1 mm path length cell. Each CD spectrum was accumulated from five scans of 50 nm/min with a 1 nm slit width and a time constant of 1 sec for a nominal resolution of 0.5 nm. Near UV CD measurements were recorded at 0.5 mg/ml protein concentration in the wavelength range of 250-300 nm with a 1 cm path length cell.

All spectra were corrected by subtracting buffer contributions and observed values were converted to mean residue ellipticity (MRE) in deg cm^2 $dmol^{-1}$ defined as

$MRE = M \theta_\lambda / 10\ d\ c\ r$

Where M is the molecular weight of protein, θ_λ is CD in millidegree, d is the path length in cm, c is the protein concentration in mg/ml and r is the average number of amino acid residues in the protein. Secondary structure elements were calculated by using CD pro software available online.

4.2.5 Effect of pH

Ll-CCRH1 samples (0.47 µM) were incubated in an appropriate buffer over the pH range (2-12) at 25 °C. The following buffers (20 mM) were used for these studies: Glycine-HCl (pH 2-3), acetate (pH 4-5), phosphate (pH 6-7), Tris-HCl (pH 8-9) and glycine-NaOH (pH 11-12). Fluorescence spectra were recorded as described in section 4.2.2. For refolding experiments, the pH of each sample was adjusted back to pH 8.0 and incubated for 25 °C for 1-2 h before recording the spectral measurement. Effect of pH on the secondary structure was studied by incubating the protein samples in 20 mM

buffers in pH range 2.0-12.0 and taking the far UV and/or near UV CD scans as described above.

4.2.6 Thermal denaturation of Ll-CCRH1

Effect of temperature on Ll-CCRH1 was monitored using a thermostatic cuvette holder connected to an external constant temperature circulation water bath in the range of 25-90 °C. The protein sample (0.47 µM) was incubated for 10 min at specified temperature before taking fluorescence scan. For renaturation, the samples were cooled to 25 °C in two ways (90 °C to 25 °C and in other, 50 °C to 25 °C) and left for 1-2 h before recording the spectra. Fluorescence measurements were performed as described in section 4.2.2. The effect of temperature on the protein was studied by two methods. In first method, the protein samples were incubated for 5 min at different temperatures ranging from 25 to 90 °C and CD scans were recorded independently. In other method, the temperature of the protein sample was increased at the rate of 1 °C /min for the temperature ranging from 25 to 90 °C and ellipticity was recorded at 208 nm. From the plot of ellipticity values against temperature, the T_m of the protein sample was determined.

4.2.7 Hydrophobic dye binding studies

The intermediate states of Ll-CCRH1 during unfolding and refolding under different denaturing conditions (pH, temperature and GdnHCl) were analyzed by performing hydrophobic dye (ANS) binding studies. ANS emission spectra were recorded in the range of 410-550 nm with excitation at 375 nm using slit widths of 7 nm each for excitation and emission monochromators. The final ANS concentration used was 50 µM. The spectrum of ANS in the buffer of respective pH or in desired concentration of GdnHCl was subtracted from the combined protein-ANS spectrum.

4.2.8 Guanidine hydrochloride (GdnHCl) mediated unfolding

Protein samples (0.47 µM) were incubated in 0-6 M GdnHCl solution at pH 8.0 for 4 h to attain the equilibrium. For refolding experiments, the protein samples in 10-fold excess concentration was first denatured in 6 M GdnHCl at 25 °C for 16 h and subsequently diluted ten times in buffer containing 0-5.5 M GdnHCl and incubated at 25 °C for 2 h. Fluorescence were recorded as described above. Effect of GdnHCl on the secondary structure was studied by incubating the protein samples (54µg/ml) in 20 mM buffers at pH 8.0 in the presence of 0.25 -6.0 M GdnHCl and taking the far UV CD scans as described above.

4.2.9 Light scattering

Protein aggregation upon thermal or pH denaturation was monitored using Ray-leigh light scattering with fluorescence spectroscopy. Both excitation and emission wavelengths, and slit widths were set at 400 nm, and 5 nm and 2.5 nm, respectively. Scattering was recorded for 60 sec.

4.3 Results and discussion

Biochemical properties of recombinant Ll-CCRH1 were studied in detail and have been reported by Sonawane *et al.* [145]. Conformational transition studies of the enzyme are presented here.

4.3.1 Monitoring conformation of Ll-CCRH1 at different pH

4.3.1.1 Fluorescence measurements

The native Ll-CCRH1, a multi tryptophan protein (Trp 34, 169, 156 and 182; GenBank ID: DQ986907) showed maximum intensity of intrinsic fluorescence, λ_{max} at 352 nm (Fig. 4.1) indicating the trp residues to be exposed to solvent.

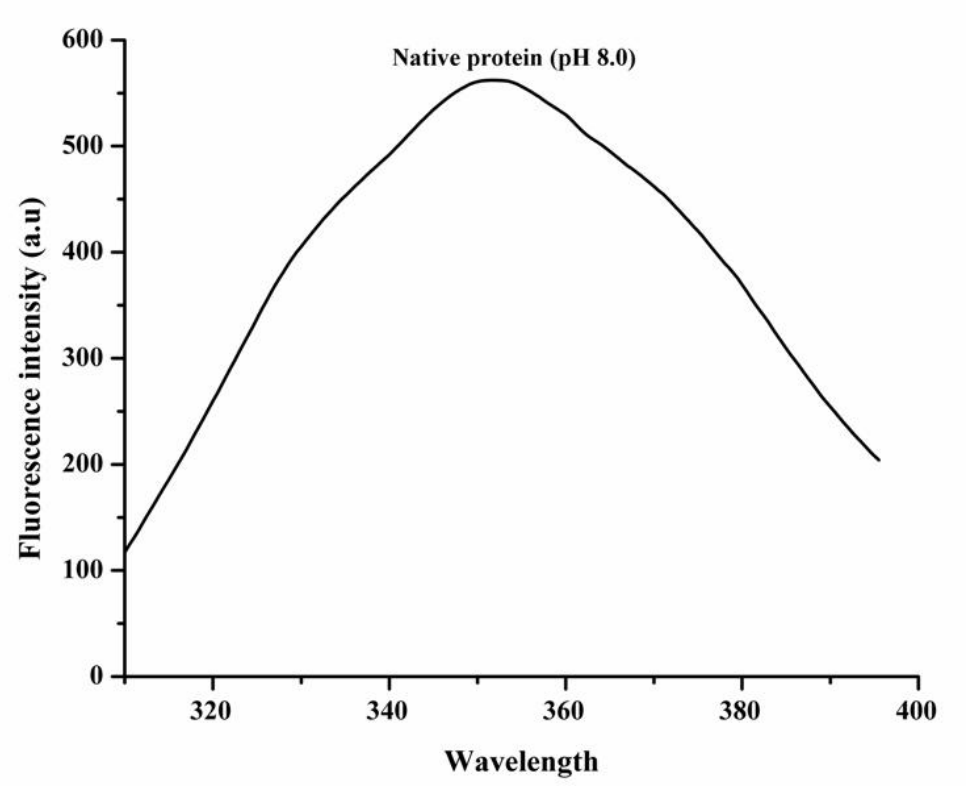


Fig. 4.1 Fluorescence spectra of native Ll-CCRH1 (0.018 mg/ml) at pH 8.0.

Decomposition analysis of the intrinsic fluorescence profile revealed two populations/conformers of the trp, 1) class A or S (33%) existing in the hydrophobic environment at any given time and 2) class III (67%), existing in completely polar environment. Fluorescence intensities of the enzyme incubated in the ranges of pH 2-5 and pH 11-12 were significantly reduced as compared to that at pH 8.0 due to protonation and deprotonation of amino acids in the vicinity of tryptophan, respectively (Fig. 4.2). The intrinsic fluorescence λ_{max} of protein incubated at different pH were observed in the range 350-352 nm suggesting apparently unchanged overall conformation.

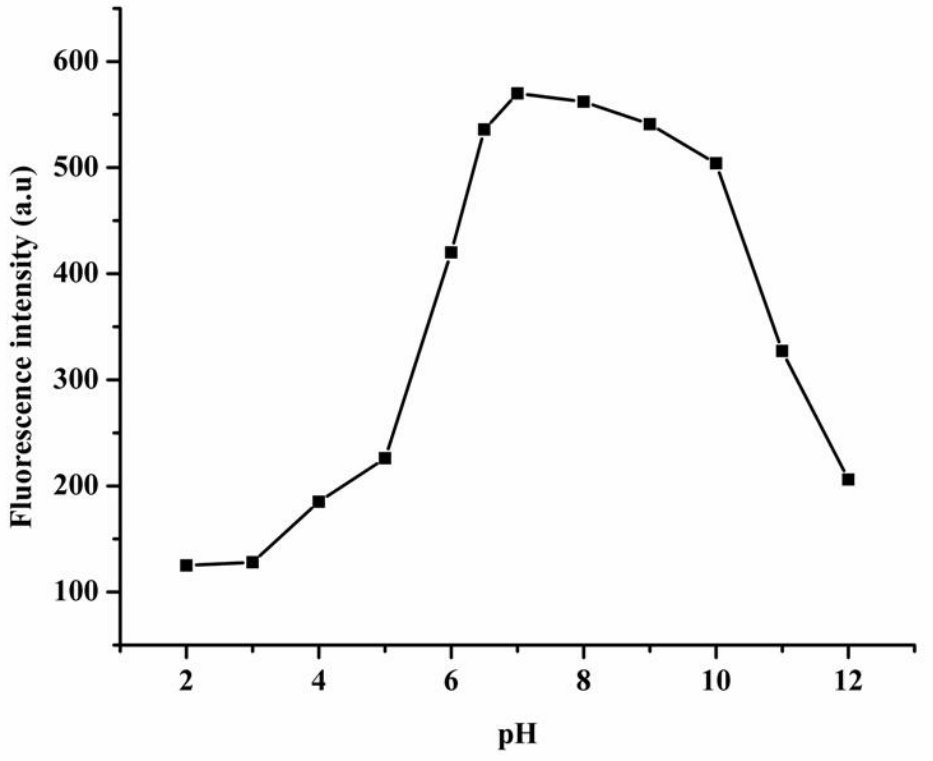


Fig. 4.2 Effect of pH on fluorescence intensity of Ll-CCRH1 protein.

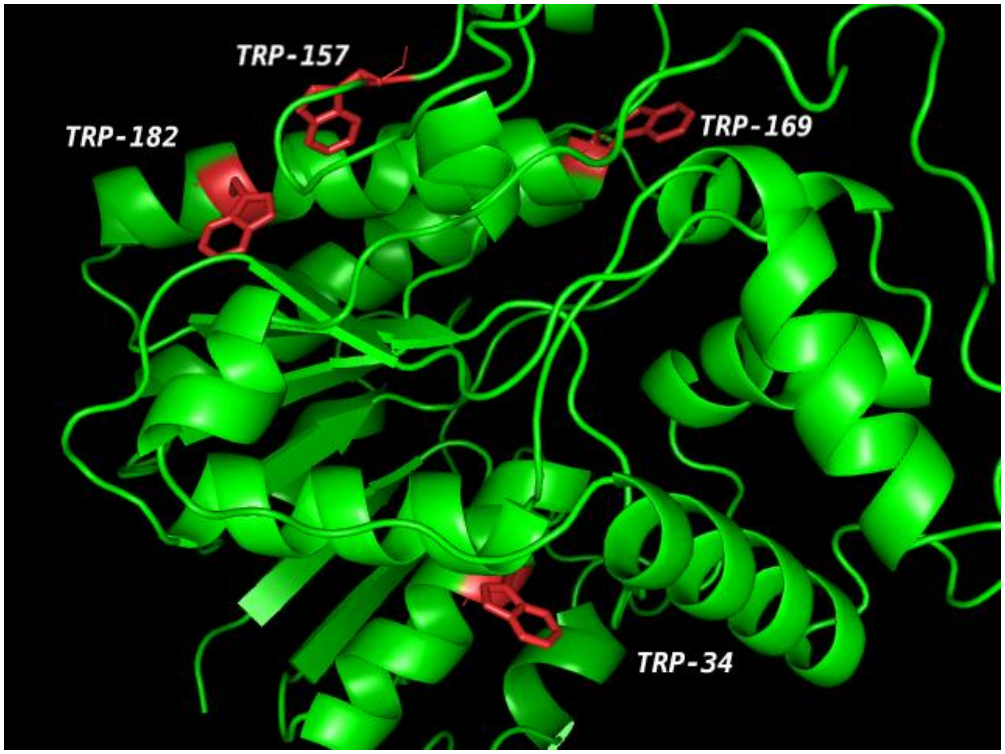


Fig. 4.3 Ribbon diagram of Ll-CCRH1 showing location of tryptophan residues (red) in protein. The diagram is prepared in PyMol (The PyMOL Molecular Graphics System,

Version 1.5.0.4Schrodinger, LLC.). The homology model of L1-CCRH1 has been submitted to protein model database (PMDB ID PM0078699).

The polar environment of exposed tryptophan residues observed in the three dimensional model of Ll-CCRH1, where all four trp residues are seen on the surface of protein is confirmed from the intrinsic fluorescence studies (Fig. 4.3) [157]. The conformers observed existing in the hydrophobic environment as deduced from decomposition analysis, could be due to hydrophobic amino acid residues around them.

4.3.1.2 Far UV CD

The far-UV CD spectra of the Ll-CCRH1 samples incubated at different pH (2-12) are shown in Figure 4.4A. The spectrum of native protein (at pH 8.0) showed minimum at 208 nm and broad shoulder (216-220 nm) indicating predominant α-helix content and very few β-sheets. The spectra were analyzed using CDSSTR program (CDPro) available online. Composition of the secondary structure elements in native protein obtained was: α-helix-65.7%, β-sheet-7.3%, turns-8.7% and unordered structures-18.3% **(**NRMSD 0.057**)** (Table 4.1).

The α-helical content of the enzyme reduced sharply when incubated at pH below 7, significantly at acidic pH; while no major alterations were observed between pH 7-11, indicating the structural stability of the protein at alkaline pH. However; at pH 12, significant change in the spectrum was observed as the α-helix and β-sheet contents were altered. A single minimum at 204 nm for Ll-CCRH1 at pH 2.0 (Fig. 4.4A) and the composition of the altered structure deduced from CDSSTR was α-helix-4.5%, β-sheet-40.1%, 22% turns and 33.4 % unordered (NRMSD 0.046) revealed major loss of helical conformation and formation of β-sheet structure in the protein. Thus,

protonation of the amino acid side chains on the surface of the native protein leads to change in the conformation.

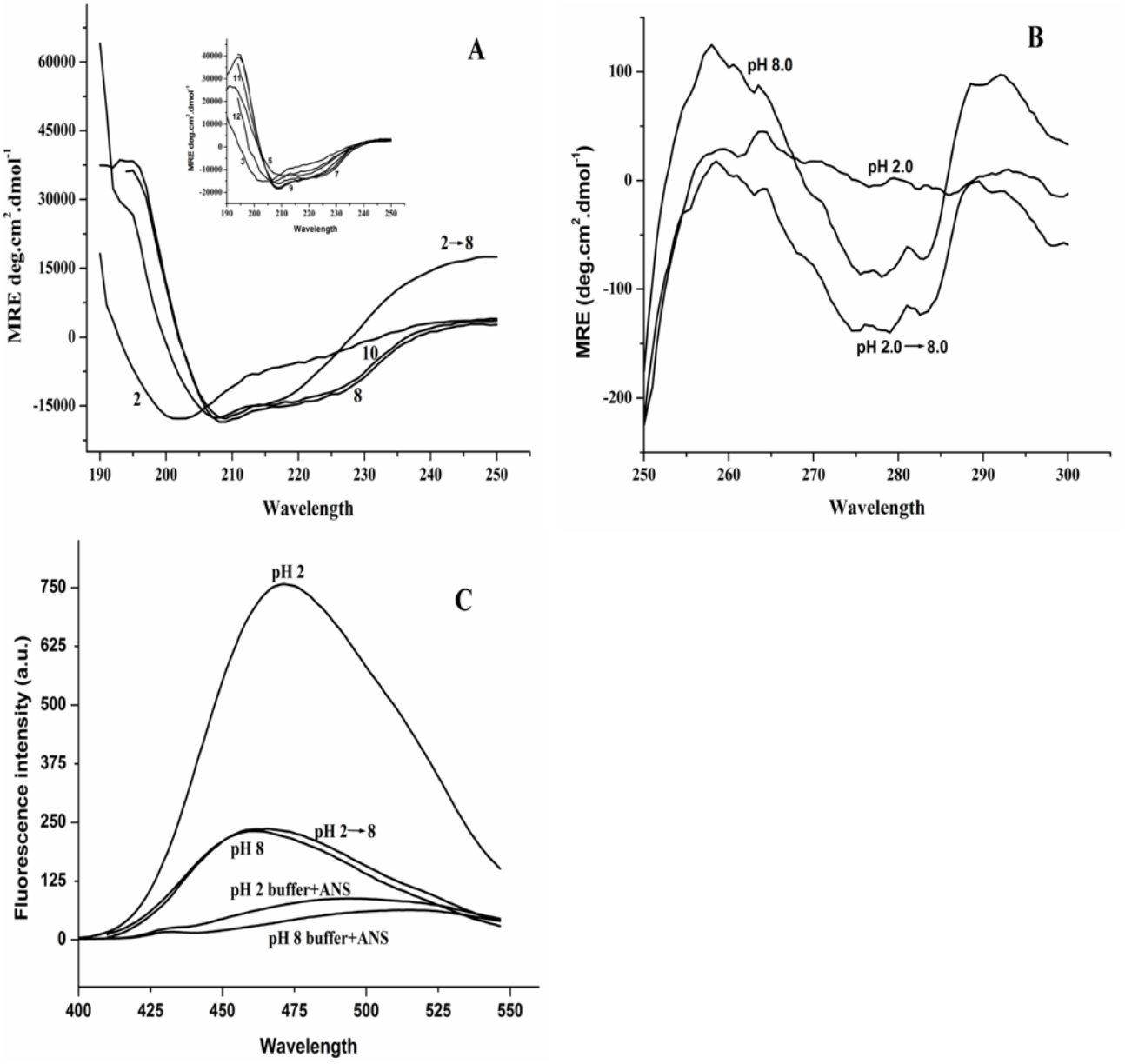


Fig. 4.4 pH induced conformational changes in Ll-CCRH1. (A) Far UV CD spectra of native Ll-CCRH1 (0.054 mg/ml) at pH 8.0, acid induced state at pH 2.0, protein incubated at pH 10.0 and the sample of which pH was readjusted from pH 2.0 to pH 8.0. Inset shows Far UV CD spectra of Ll-CCRH1 incubated at different pH (3, 5, 7, 9, 11, and 12). (B) Near UV CD spectra of Ll-CCRH1 at native, acid induced and

readjusted pH (2.0 →8.0) state. (C) ANS fluorescence emission spectra of Ll-CCRH1 (0.018 mg/ml for pH 8.0 and 10.0; 0.006 mg/ml for pH 2.0) incubated at the respective pH for 1.5 hr. The samples were excited at 375 nm. The numbers on or near the spectra indicate the respective pH value.

4.3.1.3 Near UV CD

The near UV-CD spectrum of the native Ll-CCRH1 showed two prominent maxima, one at 258 nm and the other at 292 nm, and a minima of negative ellipticity around 275-279 nm indicating ordered tertiary structure of Ll-CCRH1 (Fig. 4.4B). Significant decrease in the overall tertiary structure in the acid induced state (pH 2.0) suggested change in the environment of aromatic amino acid residues.

4.3.1.4 Hydrophobic dye binding

ANS binding to Ll-CCRH1 was checked at different pH. Native Ll-CCRH1 (at pH 8.0) did bind to ANS (five times increase in the fluorescence intensity with blue shift in the λ_{max} to 460 nm) showing native enzyme already has few hydrophobic amino acid side chains exposed to the surface (Fig. 4.4C). This confirms the class A of trp conformers, supposed to be in the hydrophobic environment as deduced from decomposition analysis (section 4.3.1.1). ANS binding to the protein to similar extent was observed below and above pH 7.0, up to pH 5.0 and pH 10, respectively. Absolutely no ANS binding was observed at extreme alkaline pH 11-12 (Fig. 4.5). However, almost sixteen fold increase in the intensity of the dye was observed with protein incubated at pH 2.0, indicating much enhanced exposure of hydrophobic amino acid residues (Fig. 4.4C).

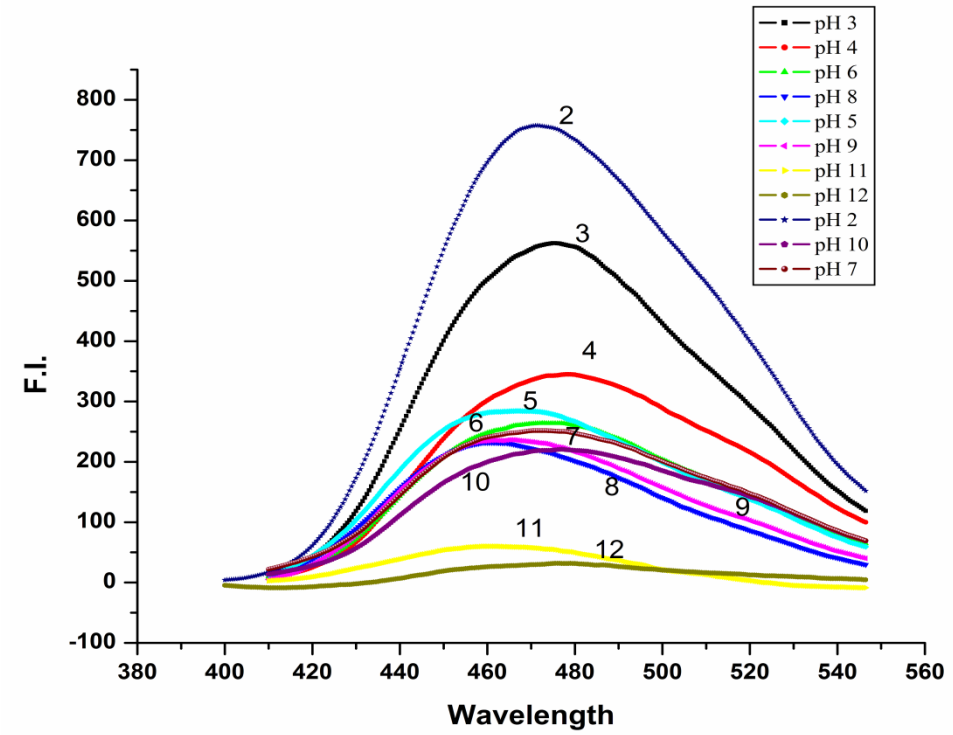


Fig. 4.5 Fluorescence spectra showing ANS binding of Ll-CCRH1 (0.47 µM) at different pH (2-12)

Thus, a rearranged yet compact secondary structure, reduced tertiary structure and exposed hydrophobic patches on the surface, indicate the existence of an acid induced molten globule like structure of Ll-CCRH1 at pH 2.0. Characterization of the molten globule like structure for its properties like reversibility, thermal and chemical stability was taken up further.

4.3.2 Characterization of molten globule

4.3.2.1 Reversibility of the acid induced state

Adjustment of pH of the protein incubated at pH 2 back to pH 8 resulted in significant reorganization of the secondary structure of the protein (Fig. 4.4A). The composition of the renatured structure was estimated to be: α-helix-54.8%, β-sheet-28.6%, turns-5.6% and unordered structures-11.0%, showing reformation of the helical structure. Also, partial reorganization of the tertiary structure of Ll-CCRH1 was

observed in near UV CD spectra (Fig. 4.4B). The enhanced exposure of hydrophobic amino acids was reversed to the extent of native state of protein (Fig. 4.4C). Thus, acid induced formation of molten globule appeared to be a reversible process.

4.3.2.2 Thermo stability of molten globule

Both native and acid treated (pH 2.0) Ll-CCRH1 samples were examined for the thermal transitions using fluorescence and CD spectroscopy and compared thereafter.

4.3.2.2.1 Steady state fluorescence

Gradual decrease in the fluorescence intensity of the native protein was observed with increasing temperature, which could be due to the deactivation of the single excited state by non radiative process (Fig. 4.6).

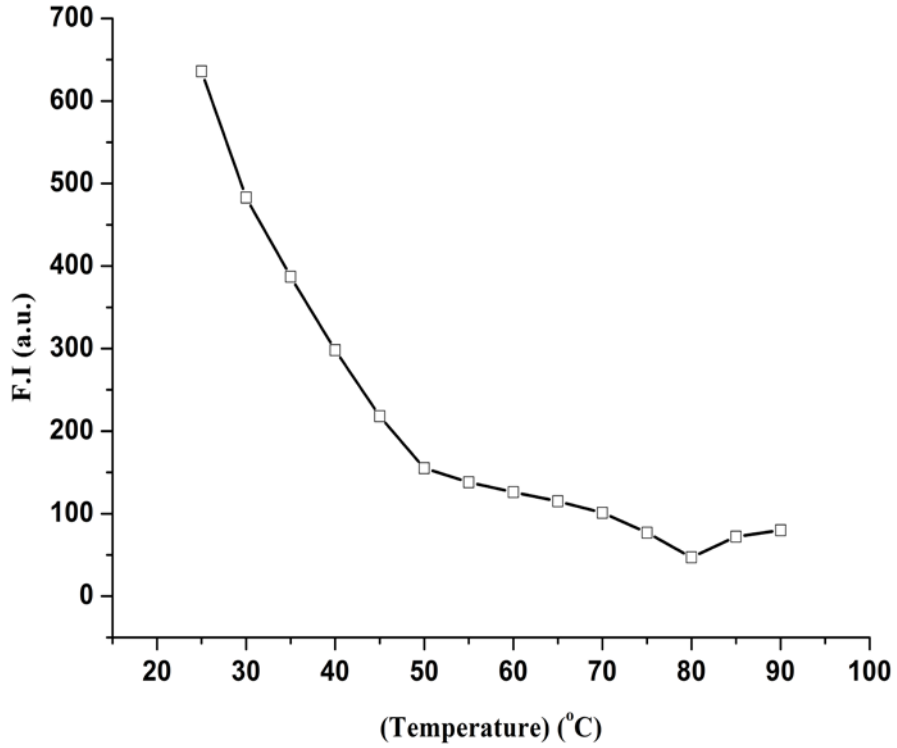


Fig. 4.6 Effect of the temperature on the fluorescence intensity of protein at pH 8.0

The observed λ_{max} values were in the range 349-352 nm upto 70 °C while above 70 °C, fluorescence spectrum showed a red shift to 356 nm due to complete exposure of trp residues of unfolded protein (Fig. 4.7A). However, at pH 2.0, Ll-CCRH1 showed red shift in the λ_{max} at and above 50 °C, indicating more polar environment of trp residues (Fig. 4.7A).

4.3.2.2.2 Far UV CD

As mentioned earlier, native protein at 25 °C contains majority of α-helical element (65.7%). Far UV CD spectra of Ll-CCRH1 exposed to increasing temperature from 25 to 90 °C are shown in Fig. 4.7B. No change in secondary structures was observed up to 40 °C. With slight change at 45 °C, sudden loss in the helical content and simultaneous increase in the beta sheet structure was observed at 50 °C. The estimated values are shown in Table 4.1. A rapid transition in the secondary structure was observed between 45 to 55 °C. At 55 °C, intermolecular beta sheet structure formation could be taking place as seen in the increased content of the respective element, similar observation was made in α-synuclein [158]. Above 65 °C, total loss in the structure was observed (Fig. 4.7B). All these conformational transitions are not reflected in the fluorescence profile of thermal denaturation of Ll-CCRH1 (section 4.3.2.2.1). As evident from the sigmoidal fit, temperature induced structural changes follow a cooperative single step two state transition (Fig. 4.7C). The T_m of the protein was calculated by monitoring changes in MRE at 208 nm with increase in temperature from 25-90 °C at the rate of 1 °C /min and was obtained as 57.77 °C (Fig. 4.7C).

. Interestingly, the altered secondary structure of molten globule formed at pH 2.0 was stable upto 90 °C indicating the thermostable nature of the species (Fig. 4.7D).

The newly formed interactions in the altered conformation due to protonation of surface amino acid residues could be playing the role in increasing the stability of Ll-CCRH1.

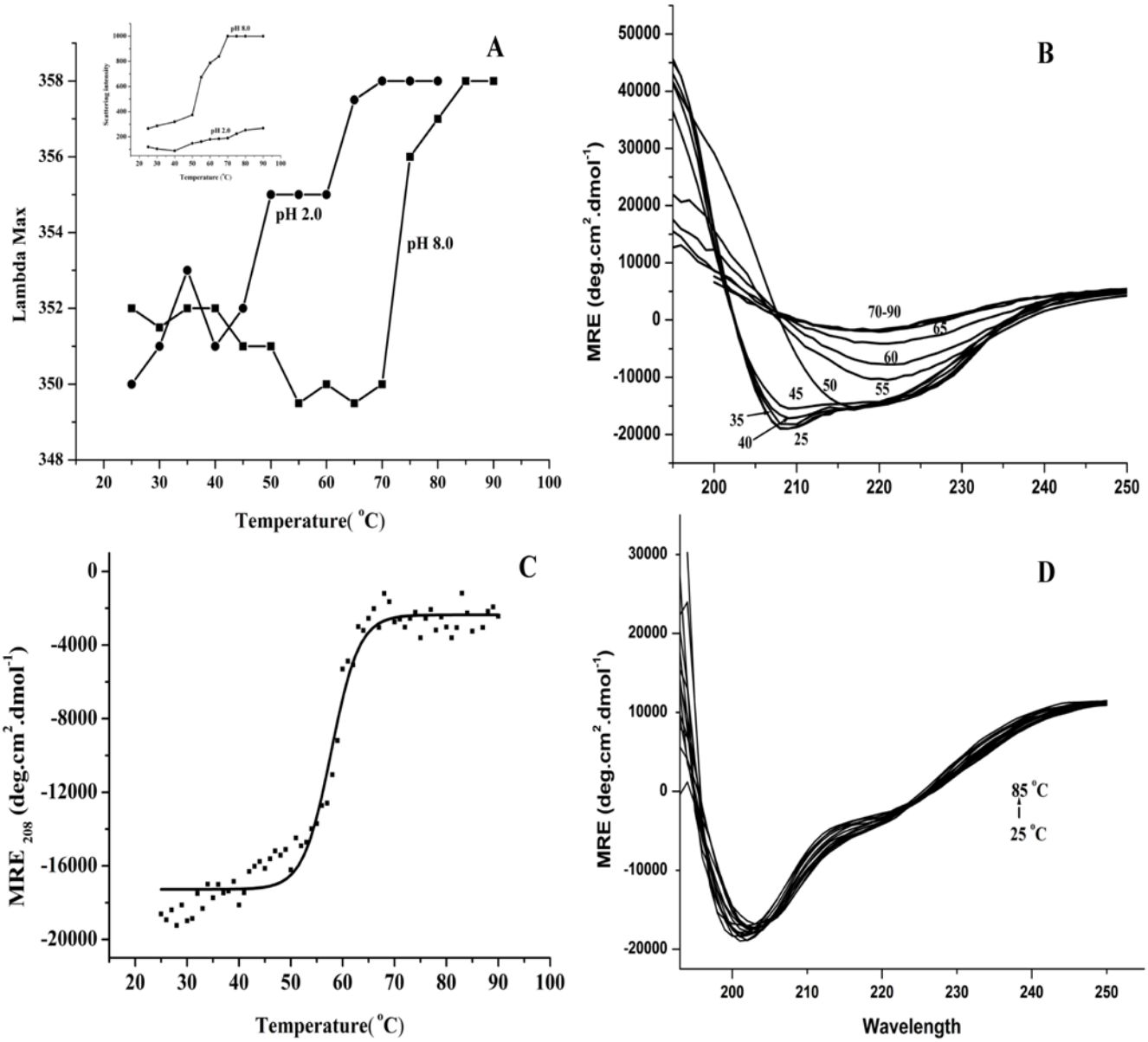


Fig. 4.7 Characterization of the acid induced molten globule state of Ll-CCRH1. (A) Effect of temperature on λ_{max} of intrinsic fluorescence of the enzyme; native (filled square) and incubated at pH 2.0 (filled circle), at increasing temperature from 25 °C to 90 °C at the interval of 5 °C for 10 min. Inset: Ray-leigh light scattering of Ll-CCRH1 at pH 8.0 (filled squares) and pH 2.0 (filled circles) (B) Thermal denaturation of Ll-

CCRH1, far-UV CD spectra of the protein incubated at different temperatures starting from 25 to 90 ^{o}C with an interval of 5 ^{o}C. The protein was incubated for 10 min at respective temperature. (C) Thermal unfolding profile showing changes in ellipticity at 208 nm when the temperature of the sample was increased at the rate of 1 ^{o}C / min (D) Thermal stability studies of the acid induced molten globule state (pH 2.0) of Ll-CCRH1: far-UV CD spectra of protein incubated at different temperatures from 25 ^{o}C to 85 ^{o}C.

4.3.3 Ray-leigh light scattering

The protein showed sudden increase in the scattering intensity at 55 °C which rapidly increased with increase in temperature (Fig. 4.7A inset) upto 70 °C indicating instant formation of aggregates before total unfolding of the protein on thermal denaturation. However, the scattering due to aggregation of protein was not observed at pH 2.0 which could be due the charge repulsion at acidic pH.

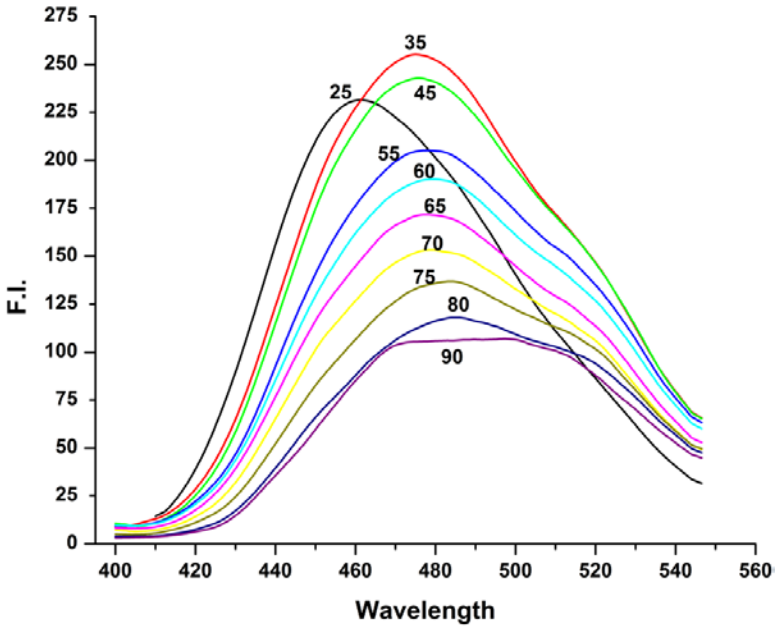


Fig. 4.8 ANS fluorescence emission spectra of Ll-CCRH1 (0.018mg/ml) incubated at respective pH for 5-10 min. The samples were excited at 375 nm and spectra were

recorded in the range of 410-550 nm. The numbers on or near the spectra indicate the respective temperature value.

Generally, due to thermal denaturation, hydrophobic patches get exposed and protein tends to aggregate. However, insignificant change in ANS binding capacity of the protein was observed after heating of the sample at 55 °C (Fig. 4.8). The already exposed hydrophobic patches on Ll-CCRH1 might be causing the aggregation at this temperature.

Table 4.1 Secondary structure elements of Ll-CCRH1 under various denaturing conditions*.

Condition	α-helix (%)	β-sheet (%)	Turns (%)	Unordered structures (%)	NRMSD
Native	65.7	7.3	8.7	18.3	0.057
pH 2.0	4.5	40.1	22.0	33.4	0.087
pH 10.0	58.3	6.9	11.9	22.9	0.074
pH 12.0	34.5	20.6	13.7	31.2	0.067
pH 2.0 → 8.0	54.8	28.6	5.6	11.0	0.046
45 °C	49.9	16.0	12.6	21.5	0.059
50 °C	38.5	31.4	12.1	18.0	0.074
55 °C	6.1	47.8	19.5	26.6	0.069
60 °C	3.1	50.2	21.8	24.9	0.070

* The calculations were carried out by CDSSTR program available online, by feeding the far UV CD spectra of Ll-CCRH1.

4.3.4 Chemical denaturation of Ll-CCRH1 molten globule

Both native and acid treated (pH 2.0) Ll-CCRH1 were subjected to chemical denaturation and conformational transitions were monitored.

4.3.4.1 Fluorescence studies

The gradual red shift in the λ_{max} (Fig. 4.9) and pronounced decrease in the fluorescence intensity of the enzyme were recorded with the increasing concentration of GdnHCl (Fig. 4.10).

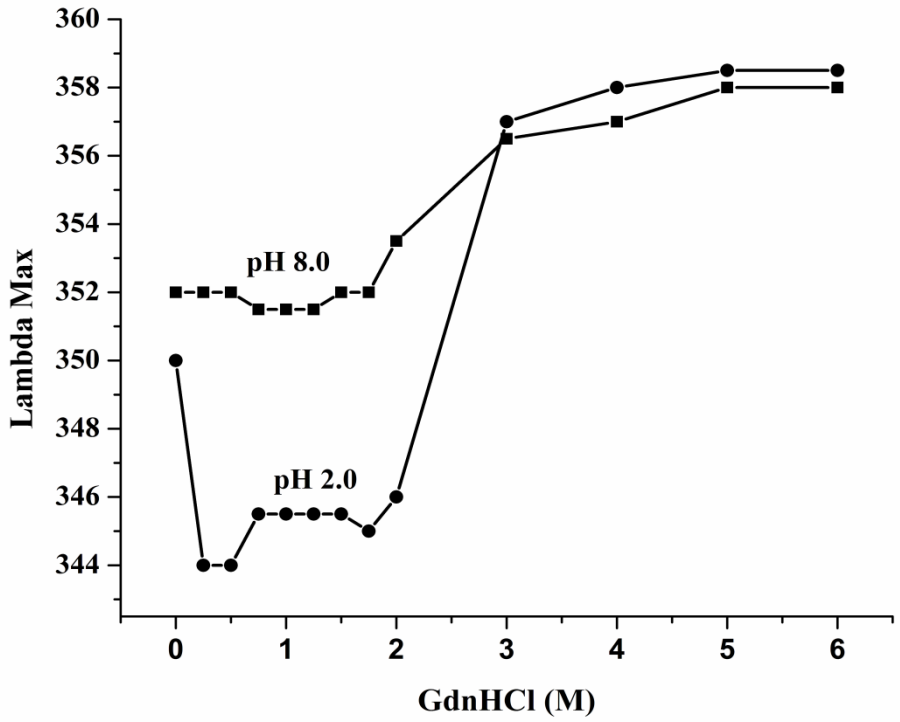


Fig. 4.9 Effect of GdnHCl on fluorescence λ_{max} of the protein incubated at pH 8.0 (filled squares) and at pH 2.0 (filled circles).

In the vicinity of 2.25 M to 6 M GdnHCl, the fluorescence λ_{max} of Ll-CCRH1 consequently shifted to 355 nm and ultimately to 358 nm indicating increase in the polarity of trp residues due to gradual unfolding of the protein. On the other hand, the molten globule like species of Ll-CCRH1 showed sudden blue shift in the λ_{max} in GdnHCl at concentrations 0.5 to 2.0 M indicating more of non polar environment of trp. At and above 3.0 M concentration, the red shift in the λ_{max} was observed similar to that after denaturation of the native protein.

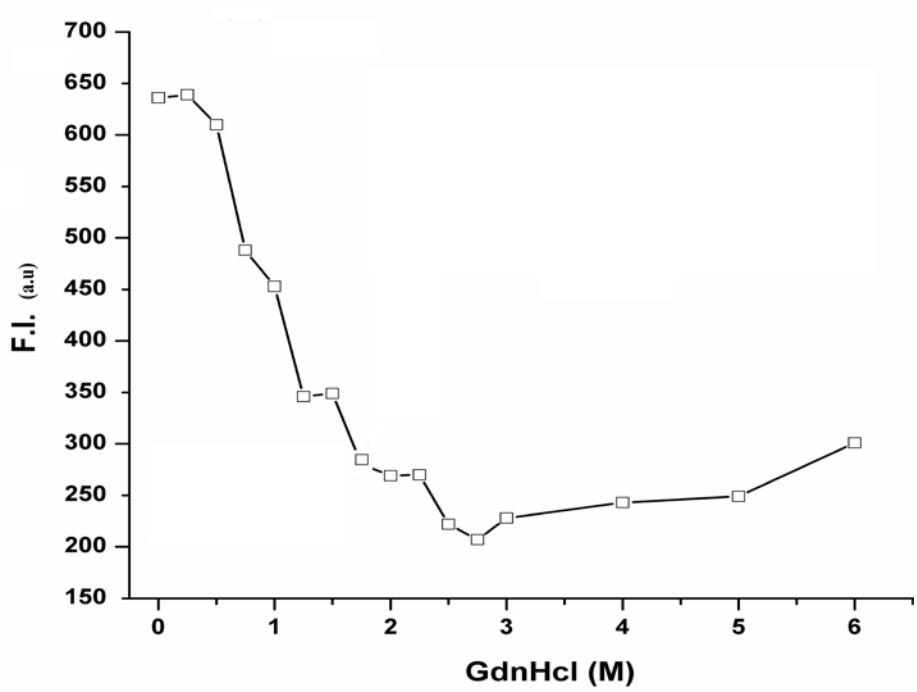


Fig. 4.10 Effect of GdnHCl on fluorescence intensity of Ll-CCRH1 protein.

4.3.4.2 Far UV CD

Major retention of the secondary structure of native Ll-CCRH1 in the vicinity of GdnHCl up to 1.5 M was observed (Fig. 4.12A). Negative MRE at 219 nm had increased, implying little pronounced secondary structure. The stability or slight increase in the compactness of the structure in presence of low concentration of GdnHCl has been observed in number of proteins [159-162]. However, significant unfolding of the protein occurred in presence of 2 M GdnHCl, and was drastic in 3 M denaturant (Fig. 4.12A). Both CD and fluorescence data analysis revealed the multi state unfolding of Ll-CCRH1 in presence of increasing GdnHCl concentrations.

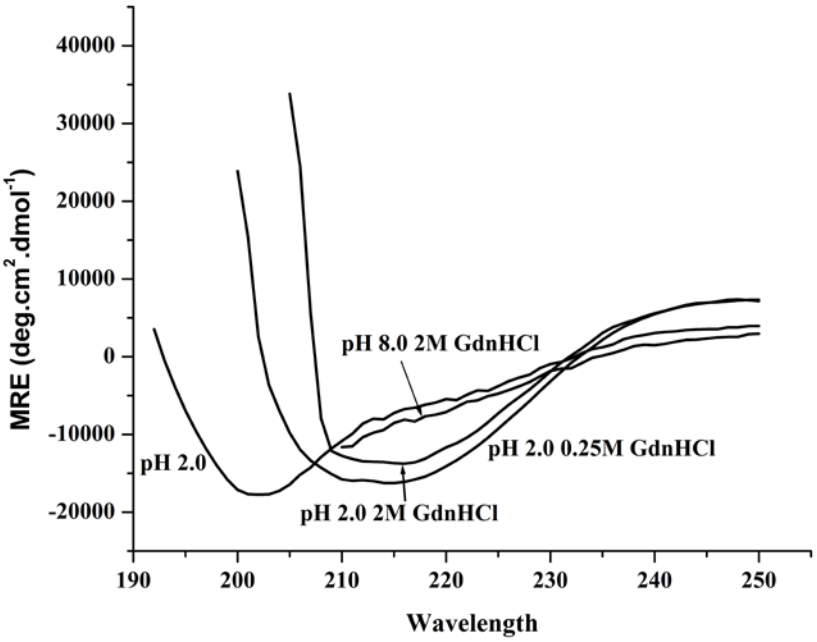


Fig. 4.11 Far UV CD spectra of native and acid induced state at pH 2.0 incubated in presence of 0.25 M and 2 M GdnHCl respectively (4 h incubation). The numbers on or near the spectra indicate respective GdnHCl concentration.

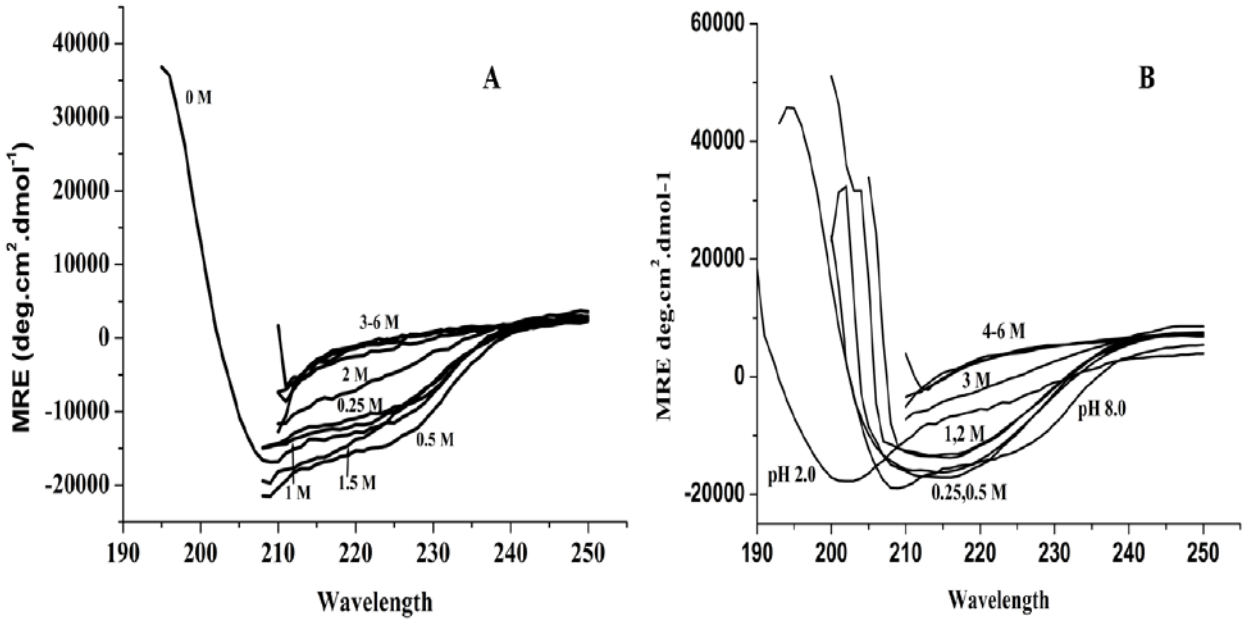


Fig. 4.12 GdnHCl induced unfolding of Ll-CCRH1. Far UV CD spectra of Ll-CCRH1 incubated in pH 8.0 (A) and pH 2.0 (B) at different concentrations of GdnHCl (0-6 M) respectively.

Although the structural change induced in Ll-CCRH1 at pH 2.0 remained stable even after heating upto 90 °C, when subjected to chemical denaturation, the protein showed distinct change in the shape of the spectrum even in the vicinity of 0.25 M GdnHCl (Fig. 4.11 and 4.12B). This altered structure was, however stable in Gdn-HCl upto 2 M concentration, whereas native protein (at pH 8.0) lost significant structure under similar conditions (Fig. 4.11). This can be correlated with the observation of blue shift in the fluorescence λ_{max} of the acid induced molten globule in the vicinity of 2M GdnHCl indicating burial of trp residues, ultimately causing tightening of the structure. Complete loss of structure was observed in the samples treated with 3-6 M GdnHCl, indicating similar extent of sensitivity of the molten globule as that of native protein (Fig. 4.12B). Thus, the molten globule structure of Ll-CCRH1 formed at pH 2.0 showed significantly increased thermostability and slightly increased stability in chemical denaturant as compared to the native protein. While unfolding of the protein in GdnHCl was observed to be reversible based on MRE 219 nm that of acid induced molten globule like structure was irreversible (Fig. 4.13).

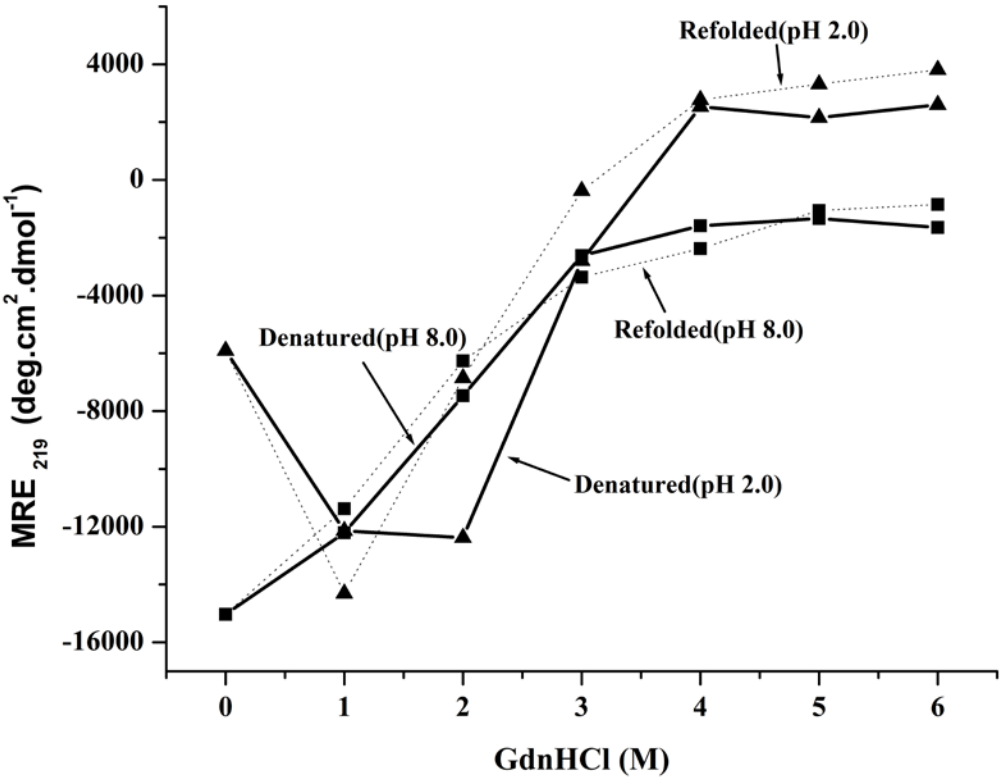


Fig. 4.13 Refolding of GdnHCl induced unfolded Ll-CCRH1: Plot of MRE at 219 nm obtained from far UV CD spectra, as a function of increasing GdnHCl concentrations representing unfolding of protein at pH 8.0 (solid line, filled squares) and pH 2.0 (solid line, filled triangles), and refolding at pH 8.0 (dotted line, filled squares) and pH 2.0 (dotted line, filled triangles), respectively.

CHAPTER 5

Steady state fluorescence studies of wild type recombinant cinnamoyl CoA reductase (Ll-CCRH1) and its active site mutants

Prashant Sonawane, Bashir Khan, Sushama Gaikwad (2013) Steady state and time resolved fluorescence studies of cinnamoyl CoA reductase (Ll-CCRH1) and its active site mutants, (Under review).

Summary

The first reductive committed step in monolignol biosynthesis is catalyzed by cinnamoyl CoA reductase. Fluorescence quenching and time resolved fluorescence studies of wild type recombinant Ll-CCRH1, a multitryptophan protein from *Leucaena leucocephala* and ten different active site mutants were carried out to investigate tryptophan environment. The enzyme showed highest affinity for feruloyl CoA (K_a= 3.72 x 10^5 M^{-1}) over other CoA esters and cinnamaldehydes, as determined by fluorescence spectroscopy. Quenching of the fluorescence by acrylamide for wild type and active site mutants was collisional with almost 100% of the tryptophan fluorescence accessible under native condition and remained same after denaturation of protein with 6 M GdnHCl. In wild type Ll-CCRH1, the extent of quenching achieved with iodide (f_a=1.0) was significantly higher than cesium ions (f_a=0.33) suggesting more density of positive charge around surface of trp conformers under native conditions. Denaturation of wild type protein with 6 M GdnHCl led to significant increase in the quenching with cesium (f_a=0.54), whereas quenching with iodide ion was decreased (f_a=0.78), indicating reorientation of charge density around trp from positive to negative and heterogeneity in trp environment. The Stern-Volmer plots for wild type and mutants Ll-CCRH1 under native and denatured conditions with cesium ion yielded biphasic quenching profiles, indicating that the trp residues in the protein fall into at least two groups that differ considerably in their accessibility and/or environment. The extent of quenching for cesium and iodide ions under native and denatured conditions observed in active site mutants is significantly different from wild type Ll-CCRH1 under the same conditions. Thus, single substitution type mutations of active site residues showed heterogeneity in tryptophan microenvironment and differential degree of conformation of protein under native or denatured conditions. The

native enzyme showed two different lifetimes, τ1 (2.27 ns) and τ2 (7.92 ns) with average lifetime, (τ) 5.28 ns, which decreased τ1 (1.38 ns) and τ2 (3.59 ns) with (τ) 2.16 ns after denaturation with 6 M GdnHCl.

5.1 Introduction

Lignin biosynthesis is a significant part of phenylpropanoid pathway which leads to formation of monolignols and subsequently lignin [11]. Among several enzymes involved in lignin biosynthesis, cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the first committed step and plays a key regulatory role in lignin formation [70]. Hydroxycinnamoyl CoA esters of general phenylpropanoid pathway when acted upon by CCR become destined to form respective monolignols. During the last two decades, significant headway has been made in characterizing CCR genes from variety of plants like *Arabidopsis*, poplar, *Medicago* and wheat [87,95,99,101,102,109]. Being entry point enzyme in lignin biosynthesis, CCR has proven to be good target for down regulation to reduce the lignin content, to improve saccharification efficiency in bioenergy crops without compromising yields, to enhance the forage quality and produce optimal feedstock plants for biofuel production [100,108]. CCR exhibits substrate specificity for different hydroxycinnamoyl CoA esters and hydroxycinnamaldehydes [89,91,145].

Studies on the intrinsic fluorescence properties have been widely used to obtain the information about the protein structure and the conformational changes induced by alteration of the environment and/or ligand binding [163-166]. Among several intrinsic fluorescent probes, tryptophan is the most popular probe. The fluorescence of the indole chromophore is highly sensitive to environment, making it an ideal choice for reporting protein conformation changes and interaction with other molecules. Also,

indole fluorescence quenching by added solutes have provided valuable information regarding structure and dynamics of protein in solution [164, 167-169].

Structure-functional studies on CCR reported in literature are very few; in fact no studies regarding intrinsic fluorescence and characterization of trp microenvironment have been reported so far. As a first step towards understanding structure-function relationship, we isolated, cloned, overexpressed, purified and characterized cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala* in detail [27,145,157]. Active site characterization of Ll-CCRH1 was carried out using modeling/docking, site directed mutagenesis and chemical modification studies [150]. Conformational transitions of Ll-CCRH1 were studied using fluorescence and circular dichroism (CD) spectroscopy. The present paper describes the exposure and differential environment of tryptophan residues in wild type recombinant Ll-CCRH1 and active site mutants on the basis of steady state fluorescence and solute quenching studies. To the present author's knowledge, this is the first report on fluorescence studies of Ll-CCRH1 from plant species.

5.2 Materials and Methods

Materials

Acrylamide, potassium iodide, cesium chloride and sodium thiosulphate, were purchased from Sigma-Aldrich, USA. All other reagents used were of high purity and analytical grade. Solutions prepared for spectroscopic measurements were in sterile MilliQ water.

5.2.1 Heterologous expression and purification of recombinant Ll-CCRH1 from *L .leucocephala*

Heterologous expression, purification and assay of the Ll-CCRH1 enzyme with cinnamoyl coA esters were carried out as described by Sonawane *et al*. [145, also in chapter 2, section 2.2.1 and 2.2.3].

5.2.2 Steady state fluorescence

Intrinsic fluorescence of the protein was measured at 30 °C using a Perkin-Elmer LS 50B spectrofluorimeter connected to a Julabo F20 circulating water bath. The protein solution (0.47 µM) was excited at 295 nm and emission was recorded in the range of 310-400 nm. Both the excitation and emission spectra were obtained by setting the slit width at 7 nm, and a scan speed of 100 nm/min. To eliminate the background emission, the signal produced by either buffer solution, or buffer containing the desired quantity of denaturants was subtracted.

5.2.3 Ligand binding analysis

Fluorescence measurements were carried out as described in section 5.2.2. The binding of CoA esters to Ll-CCRH1 was studied by intrinsic fluorescence titrations. The CoA ester/ligand solution was added in 10-12 aliquots (1-10 µl). The ligands were used in the concentration of 1 mM forward substrates (feruloyl CoA and sinapoyl CoA), reverse substrates (coniferaldehyde and sinapaldehyde), and cofactor (NADPH and $NADP^+$). The fluorescence intensity at 352 nm (λ_{max} of the protein) was considered for further analysis. Corrections were also made to compensate the dilution effect upon addition of CoA ester/ligand to the Ll-CCRH1.

The association constants were calculated according to the method described by Chipman *et al*. [170]. The abscissa intercept of the plot of log $[C]_f$ against log $\{(\Delta F)/(Fc- F_\infty)\}$, where $[C]_f$ is the free ligand concentration, yielded pK_a value for protein-ligand interaction according to the relationship

$$\log [F_o - F_c / F_c - F_\infty] = \log K_a + \log \{[C]_t - [P]_t (\Delta F/\Delta F_\infty)\} \quad (1)$$

where F_c is the fluorescence intensity of the protein at any point during the titration, $[P]_t$ is the total protein concentration, ΔF_∞ is the change in fluorescence intensity at saturation binding, $[C]_t$ is the total ligand concentration, and $[C]_f$ is the free ligand concentration, given by,

$$[C]_f = \{[C]_t - [P]_t (\Delta F/\Delta F_\infty)\} \quad (2)$$

Free energy changes of association (ΔG) were determined by the equation,

$$\Delta G = -RT \ln K_a \quad (3)$$

5.2.4 Solute quenching studies

Fluorescence measurements were performed for native and denatured protein with different quenchers like acrylamide (5 M) (neutral quencher), iodide (5 M) and cesium ions (5 M) (charged quenchers), as described above. Small aliquots of quencher stock solutions were added to protein samples and fluorescence spectra were recorded after each addition. Sodium thiosulphate (0.2 M) was added to the iodide stock solution to prevent the formation of tri-iodide ions (I^{-3}). For quenching studies with denatured Ll-CCRH1, the protein was incubated in 6 M GdnHCl overnight at room temperature. Volume correction was done for fluorescence intensities before analyzing the quenching data.

The steady state fluorescence quenching data obtained with different quenchers were analyzed by Stern-Volmer (Eq. 1) and modified Stern-Volmer (Eq. 2) equations in order to obtain quantitative quenching parameters [168,171].

$$F_o/F_c = 1 + K_{sv} [Q] \quad (1)$$

$$F_o/\Delta F = f_a^{-1} + 1/[K_a f_a (Q)] \quad (2)$$

Where F_o and F_c are the relative fluorescence intensities in the absence and presence of quencher, respectively, [Q] is the quencher concentration, K_{sv} is Stern-Volmer quenching constant, $\Delta F=F_o-F_c$ is the change in fluorescence intensity at any point in the quenching titration, K_a is the corresponding Stern-Volmer constant for accessible fraction of the total fluorophores and f_a is the fraction of the total fluorophore accessible to quencher. Equation 2 shows that the slope of the plot of $F_o/\Delta F$ versus $[Q]^{-1}$ (modified Stern-Volmer plot) gives the value of $(K_a f_a)^{-1}$ and its Y-intercept gives the value of f_a^{-1}.

5.2.5 Fluorescence measurements and quenching studies of active site mutants

Active site characterization of Ll-CCRH1 was previously carried out by means of molecular, computational and biochemical methods. Putative active site residues identified in homology modeling and docking studies were further confirmed by site directed mutagenesis and chemical modification. Active site of Ll-CCRH1 is made up of 10 residues, that is, Phe30, Ile31, Arg51, Asp77, Ser136, Tyr170, Lys174, Val200, Ser212 and His215 [150]. Respective ten site directed mutant proteins were generated for each of active site residue. These active site mutants were constructed, cloned, over-expressed and purified similarly as that of wild type recombinant Ll-CCRH1 [150]. Fluorescence measurements and quenching studies of active site mutants were performed as described in section 5.2.2 and 5.2.4.

5.2.6 Lifetime measurement of fluorescence decay

Lifetime fluorescence measurements were carried out on an FLS920 single photon counting spectrometer supplied by Edinburgh instruments. A xenon flash lamp of pulse

width 1 ns was used for excitation and a synchronization photomultiplier was used to detect fluorescence. The diluted Ludox solution was used for measuring Instrument Response Function (IRF). Protein sample (0.2 mg/ml) was excited at 295 nm and emission was recorded at 352 nm. Slit widths of 10 nm each were used on the excitation and emission monochromators. The resultant decay curves were analyzed by a multiexponential iterative fitting program provided by instrument.

5.3 Results and discussion

5.3.1 Steady state fluorescence

The native Ll-CCRH1, a multi tryptophan protein (GenBank: DQ986907) showed fluorescence maximum, λ_{max} at 352 nm indicating the trp residues to be exposed to solvent. However, decomposition analysis of the intrinsic fluorescence profile revealed two populations/conformers of the trp, 1) class A or S (33%), hydrophobic environment and 2) class III (67%), polar environment. The denatured protein showed red shift in λ_{max} to 358 nm, indicating more enhanced exposure of trp to completely polar environment due to unfolding of protein (Fig. 5.1)

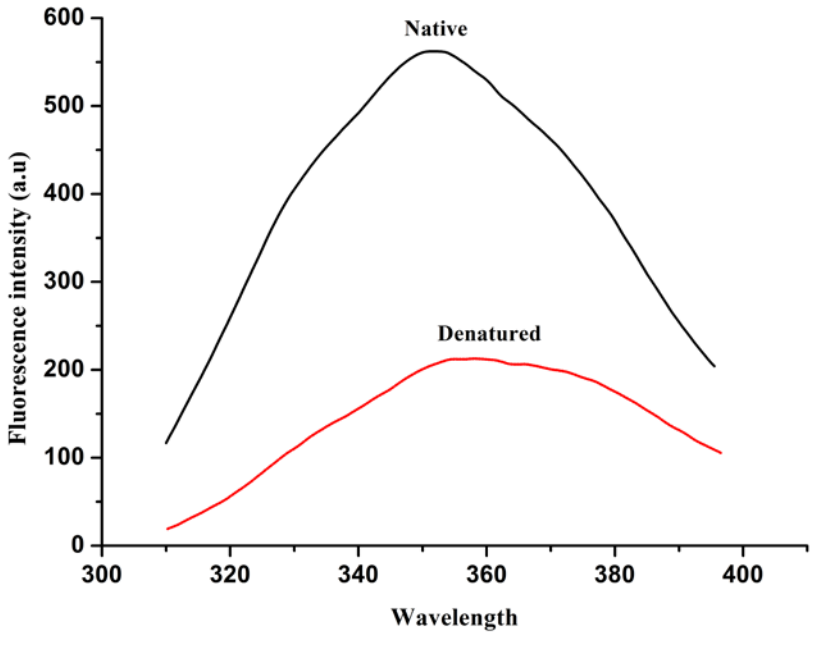


Fig. 5.1 Fluorescence emission spectra of the native (0.018 mg/ml) and denatured Ll-CCRH1 (in 6 M GdnHCl).

Tryptophan residues appear to be uniquely sensitive to quenching by variety of solutes as a result of a propensity of the excited indole nucleus to donate electrons while in the excited state. Fluorescence spectra of the native and denatured Ll-CCRH1 recorded in the absence and in presence of increasing concentrations of acrylamide (Fig. 5.2 A and B) displayed higher extent of quenching in the presence of 6 M GdnHCl, clearly revealing that unfolding results in significant increase in the accessibility of the tryptophan residues to the quencher. Also, denaturation led to significant increase in the extent of quenching with other quencher, namely cesium (Table 5.1). The percentage quenching was calculated from raw data.

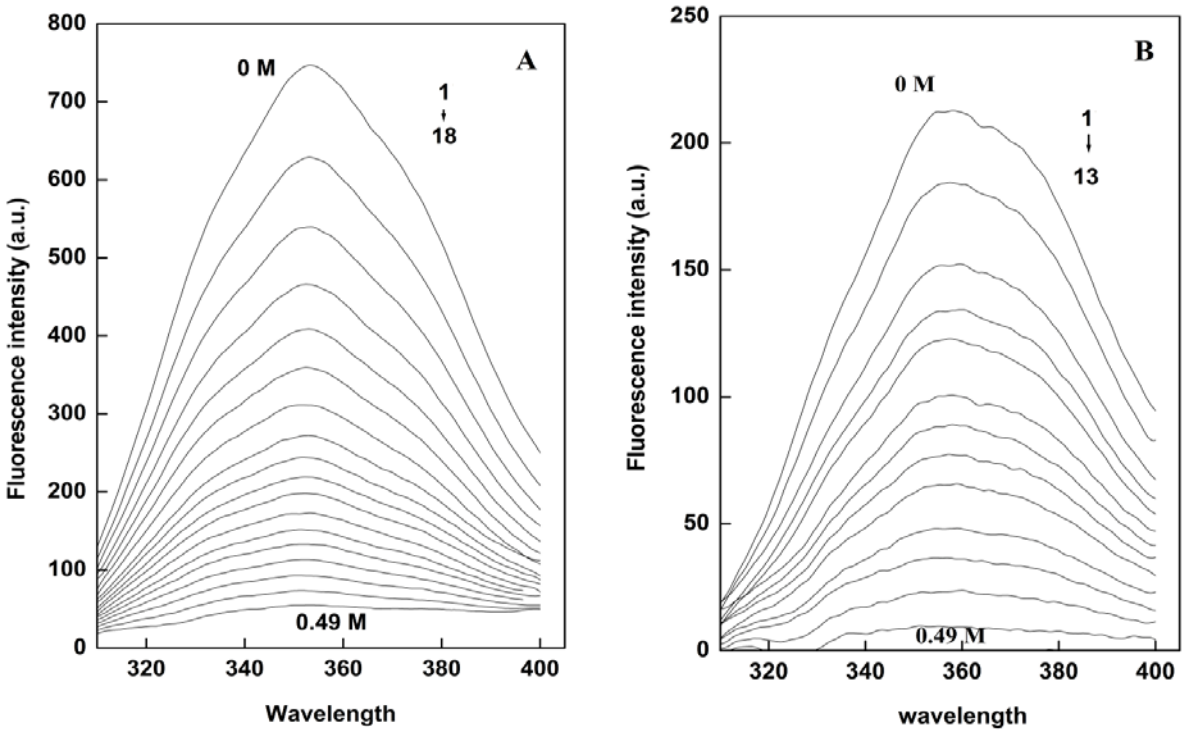


Fig. 5.2 Fluorescence spectra of Ll-CCRH1 in the absence and presence of acrylamide. (A) Under native conditions, (B) Under denatured conditions (6 M GdnHCl). Spectrum

1 corresponds to protein alone and spectra 2-18/2-13 correspond to protein in the presence of increasing concentrations of acrylamide. The final concentration of acrylamide in both A and B is 0.49 M.

Table 5.1 Extent of fluorescence quenching of Ll-CCRH1 with different quenchers.

Quenchers	Quenching (%)	
	Native	In 6 M GdnHCl
Acrylamide (0.49 M)	93	100
CsCl (0.5 M)	39.5	53
KI (0.56 M)	77	72

Of the three quenchers used, acrylamide was the most effective, quenching 93% of the total intrinsic fluorescence of the protein (at 0.49 M). Amongst ionic quenchers, iodide and cesium ions, which cannot penetrate into the protein matrix and can access only surface exposed tryptophans, were found to quench only 77% and 39.5% respectively, of the total intrinsic fluorescence of Ll-CCRH1. With charged quenchers, higher quenching was observed by iodide ions compared to cesium, with native Ll-CCRH1 indicated surface tryptophan in the protein to have more positively charged amino acids around them. Also, the inherently low quenching efficiency of cesium ions may be responsible for lower quenching observed. On the other hand, the iodide ions could get concentrated in a positively charged environment in the vicinity of surface tryptophans and increased probability of iodide ions colliding with them and quenched the fluorescence.

Denaturation of Ll-CCRH1 resulted in a significant increase in quenching by acrylamide and cesium ions, with extent of quenching observed being 100% and 53% respectively. In case of cesium ions, the extent of quenching increased almost by 40%

after denaturation (Table 5.1). Denaturation of Ll-CCRH1 led to increase in the accessibility of tryptophan fluorescence by cesium ions indicating increase in the density of negative charge around the tryptophan residues. Thus, in native Ll-CCRH1 condition, the environment of tryptophan residues is electropositive; while in denatured state, charge reorientation resulted in electronegative environment around tryptophan residues. This may be the reason for decrease in extent of quenching by iodide ions (72%) in denatured condition. The presence of negatively charged residues in proximity of trp residues repel the negatively charged iodide ions, resulting in decreased quenching efficiency (Table 5.1).

5.3.2 Lifetime measurements

The fluorescence decay of the native Ll-CCRH1 on a nanosecond time scale obtained from time resolved measurements is presented in Fig. 5.3. When fitted into a biexponential curve ($\chi 2 = 1.076$), two decay times $\tau 1$ (2.27 ns) and $\tau 2$ (7.92 ns) with their relative contributions to the overall fluorescence being 20% and 80% respectively, were obtained indicating the presence of two conformers of the tryptophans in Ll-CCRH1. The shorter life time component is supposed to be on the surface of the protein and its fluorescence decays faster, while the longer conformer lies in the interior and decays slowly. The λ_{max} at 352 nm is hence the cumulative intrinsic fluorescence of Ll-CCRH1. The two life times of present protein can be correlated with decomposition analysis of intrinsic fluorescence profile showing polar as well as non polar environment of trp. The average life time was calculated using the following equation [25]:

$$\tau = \sum \alpha_i \tau_i / \sum \alpha_i \quad i=1, 2, \ldots\ldots$$

Where τ is the average life time and α is the weighting factor. The average τ was found to be 5.28 ns for native Ll-CCRH1.

For the denatured enzyme, again two decay times τ1 (1.3822 ns) and τ2 (3.5810 ns) with 41% and 59% contribution to the total fluorescence, respectively were observed.

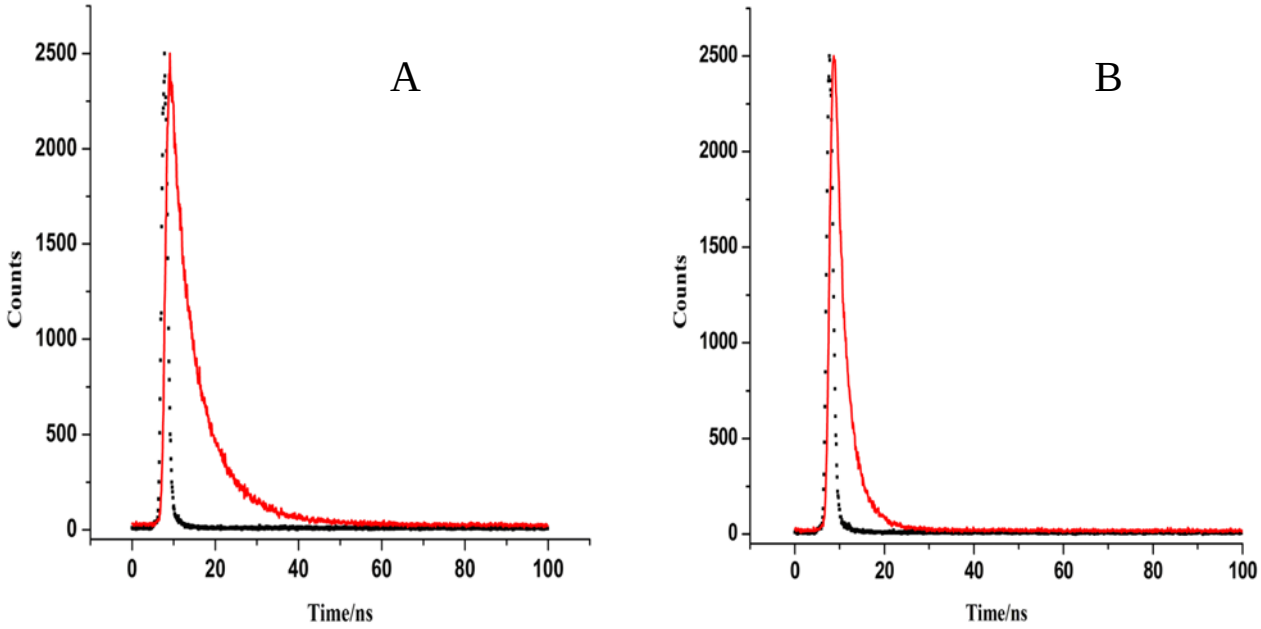
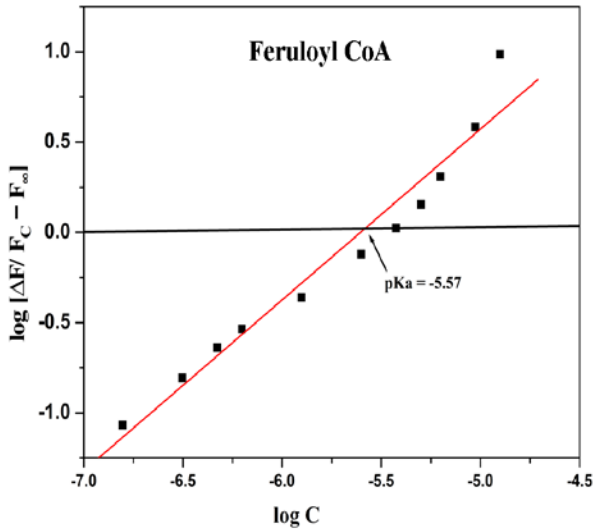
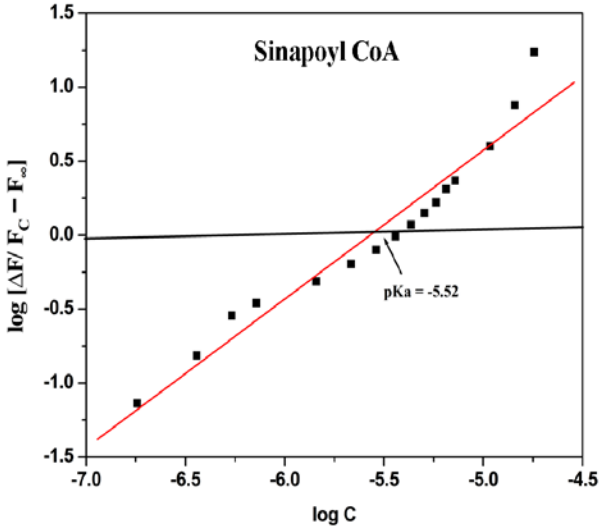


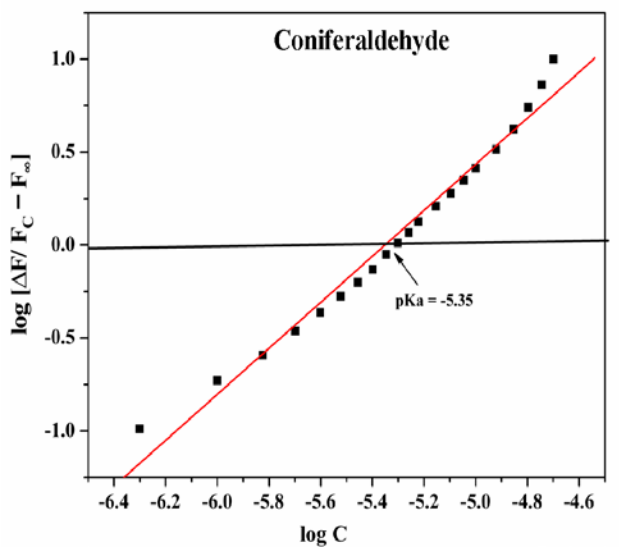
Fig. 5.3 Life time measurement of Ll-CCRH1. Time resolved fluorescence decay profile of (A) native Ll-CCRH1 (B) 6 M GdnHCl denatured Ll-CCRH1. The solid line corresponds to the non linear least square fit of the exponential data.

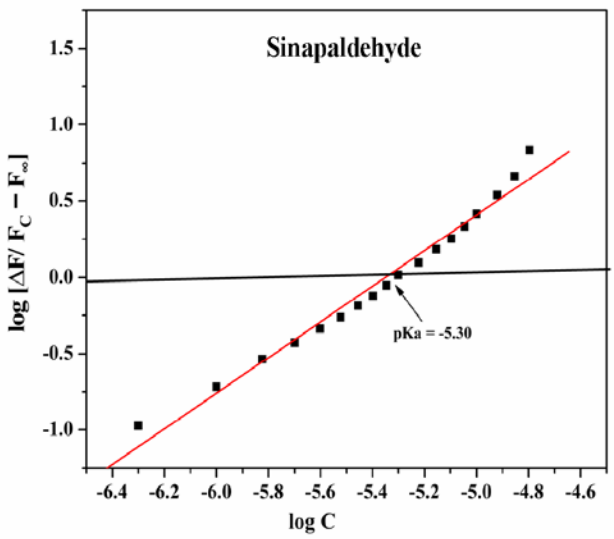
5.3.3 Fluorimetric analysis of CoA esters binding

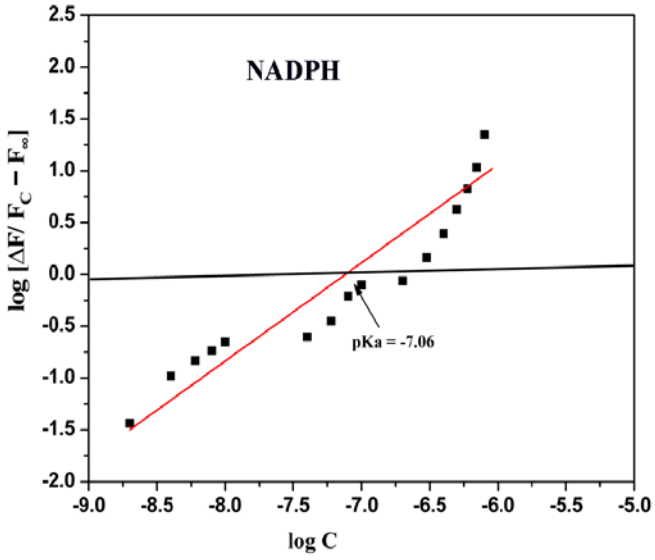
Cinnamoyl CoA esters binding studies were carried out using fluorescence spectroscopy. Specific binding of CoA ester may change the microenvironment of tryptophan either by enhancing [172] or by quenching of the fluorescence [173]. Titration of Ll-CCRH1 with substrates (feruloyl CoA, sinapoyl CoA, coniferaldehyde and sinapaldehyde) and cofactors (NADPH and NADP) resulted in quenching of fluorescence. Biochemical studies of Ll-CCRH1 showed better affinity of enzyme for feruloyl CoA over other substrates and preference of reduction reaction over oxidation [12]. Here, the enzyme showed association constant, K_a for feruloyl CoA, sinapoyl CoA, coniferaldehyde and sinapaldehyde as 3.72, 3.3, 2.2., 1.99 (x 10^5 M^{-1}) respectively, confirming the differential binding affinity of substrates towards protein (Fig. 5.4, Table 5.2). The binding constants for cofactors, NADPH and $NADP^+$ to protein are 11.49 x 10^6 M^{-1} and 7.46 x 10^5 M^{-1} respectively showing preference for NADPH over NADP in Ll-CCRH1 mediated reaction. ΔG values for different ligands are in the range of -31 to -41 $kJmol^{-1}$, indicating spontaneous nature of binding.


Feruloyl CoA
pKa = -5.57
log [ΔF/ F_C − F_∞]
log C


Sinapoyl CoA
pKa = -5.52
log [ΔF/ F_C − F_∞]
log C


Coniferaldehyde
pKa = -5.35
log [ΔF/ F_C − F_∞]
log C


Sinapaldehyde
pKa = -5.30
log [ΔF/ F_C − F_∞]
log C


NADPH
pKa = -7.06
log [ΔF/ F_C − F_∞]
log C

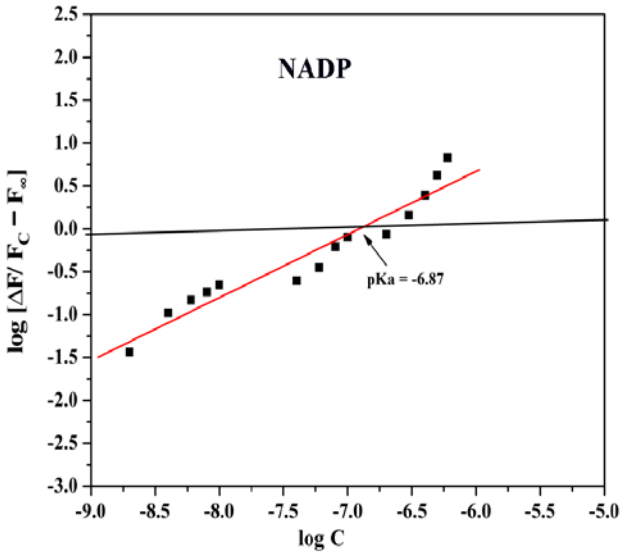

NADP
pKa = -6.87
log [ΔF/ F_C − F_∞]
log C

Fig. 5.4 Determination of the association constants (K_a) for the interaction of Ll-CCRH1 with CoA esters, cinnamaldehydes and cofactors by fluorescence titration using double log plot of log $\{\Delta F/ (F_c - F_\infty)\}$ versus log $[C]_f$. The X-intercept of the plot gives pKa value for the interaction between Ll-CCRH1 and substrates.

Table 5.2 Summary of ligand binding to Ll-CCRH1 with fluorescence analysis.

Substrates	K_a **(M^{-1})**	ΔG **($kJmol^{-1}$)**
Feruloyl CoA	3.72×10^5	- 31.77
Sinapoyl CoA	3.3×10^5	- 31.40
Coniferaldehyde	2.23×10^5	-30.51
Sinapaldehyde	1.99×10^5	-30.15
NADPH	11.49×10^6	-40.28
$NADP^+$	7.46×10^5	-33.50

5.3.4 Analysis of the fluorescence quenching data of Ll-CCRH1

5.3.4.1 Quenching with acrylamide

Acrylamide, being neutral by nature, was found to be the most efficient quencher for intrinsic fluorescence of Ll-CCRH1 as it could penetrate into the interior of protein. Quenching of the Ll-CCRH1 with acrylamide gave a linear Stern-Volmer plot with K_{sv} value as 23.15 M^{-1} for native protein, indicating exceptionally higher rate of quenching (Fig. 5.5 A). 100% of the tryptophan fluorescence is accessible to protein under native condition (Fig. 5.5 B, Table 5.3).

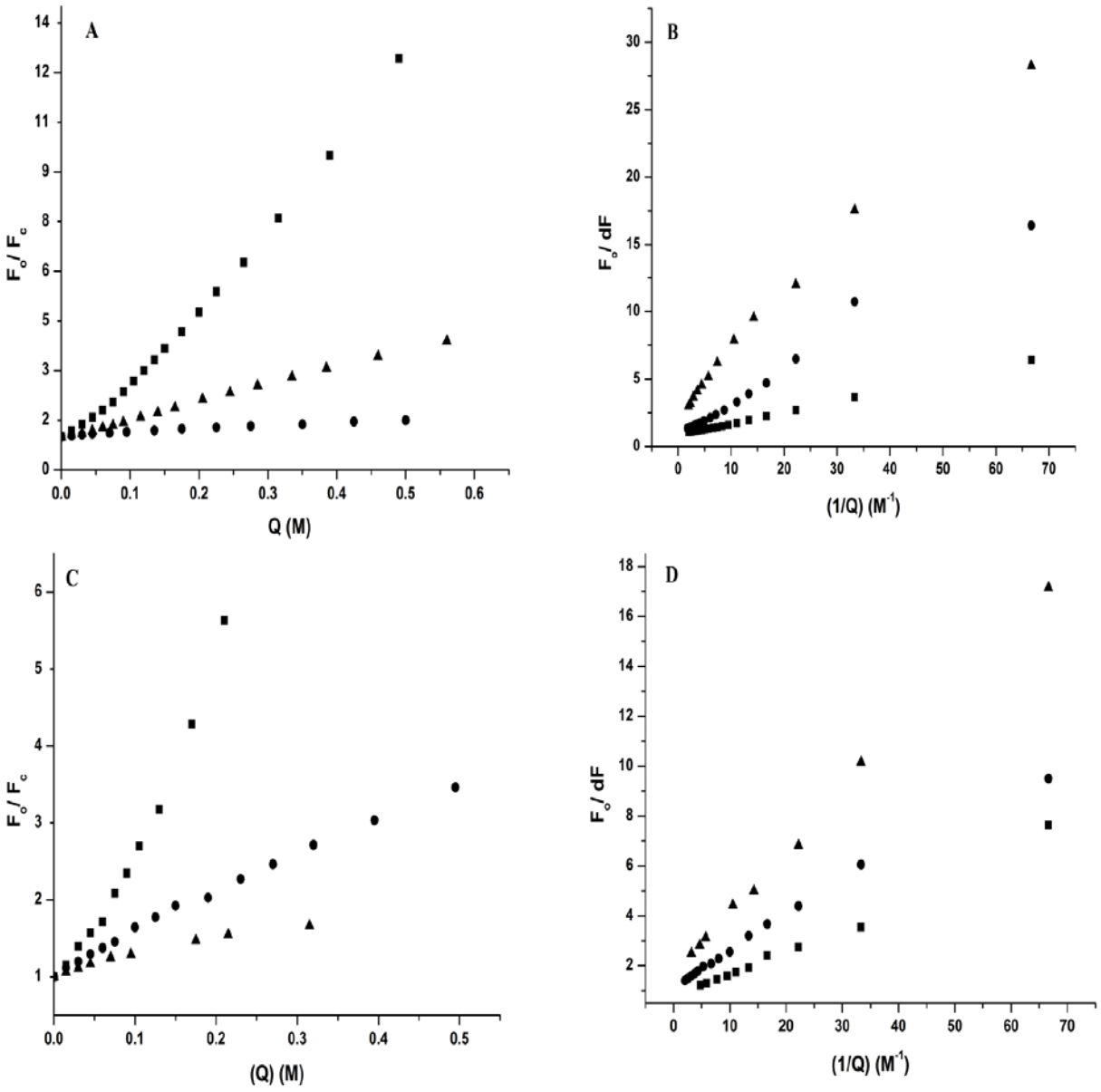


Fig. 5.5 Stern-Volmer plots (A,C) and modified Stern-Volmer plots (B,D) for the quenching of native (A,B) and 6 M GdnHCl denatured (C,D) Ll-CCRH1 (0.018 mg/ml). Acrylamide (filled squares), cesium (filled triangles) and iodide (filled circles) was used as quenchers. After fitting the data, the R^2 value in each case was 0.99. The downward curves seen in the quenching profiles of iodide and cesium under native and denatured conditions respectively (A,C) were split into two linear components.

Accessibility of acrylamide to fluorescence remained same, 100% upon denaturation of the protein with 6 M GdnHCl as determined from modified Stern-Volmer plot (Fig. 5.5 D, Table 5.4). An upward curvature was obtained in the Stern-Volmer's plot, indicating the involvement of both collisional and static components (Fig. 5.5 C). The static mechanism is a consequence of complex formation, while the dynamic mechanism involves collisions with acrylamide during the lifetime of tryptophan in excited state [167]. Almost similar K_{sv} value, 21.51 M^{-1} for denatured protein was also observed.

5.3.4.2 Quenching with cesium and iodide ions

The Stern-Volmer plots for the quenching of Ll-CCRH1 with ionic quenchers (CsCl and KI) are shown in Fig. 5.5. The quenching profile obtained for iodide ion under native and denatured conditions showed linear dependence. The Stern-Volmer plot obtained for cesium ion under same conditions displayed negative curvature (Fig. 5.5 A,C). Downward/negative curvature of the plots indicated that certain tryptophans are selectively quenched before others. At low concentration of quencher, the slop of the Stern-Volmer plots reflects largely the quenching of the more accessible residues. At higher concentrations, the easily quenched fluorescence has been depleted, and those tryptophans having lower quenching constants become dominant. Similar quenching patterns have been observed for several single or multi-tryptophan proteins [174-180]. Downward curvatures of quenching profiles with cesium ions shown for native and denatured Ll-CCRH1 suggest the heterogeneity in the microenvironment of surface exposed tryptophans. These results also indicate the presence of tryptophans in Ll-CCRH1 in different environments where some tryptophans are partially or fully exposed to solvent while others are present inside hydrophobic environment. These

observations are in good agreement with the decomposition analysis profile of native enzyme [23].

The biphasic curvature obtained for CsCl quenching in native state was split into two linear components and the values of K_{sv1} and K_{sv2} obtained were 1.31 M^{-1} and 0.82 M^{-1} respectively; while K_{sv} of 5.35 M^{-1} was observed for KI (Fig. 5.5 A, Table 5.3). The higher K_{sv} value of iodide as compared to cesium indicates higher efficiency of quenching by iodide due to presence of an electropositive environment around trp residues in the enzyme.

Modified Stern-Volmer plots obtained with ionic quenchers are shown in Fig. 5.5 B, D, from which f_a or fractional accessibility of the total fluorescence and K_a, quenching constant were obtained and listed in Table 5.4. From table 5.4, it is seen that 33% and 100% of the total fluorescence is accessible to cesium and iodide ions, respectively, under native condition. These results also confirm higher extent of quenching by iodide ions and suggest more density of positive charge around surface of trp conformers.

The ionic solute quenching profile was different for denatured protein. The denatured protein gave a linear Stern-Volmer plot (Fig. 5.5 C) and the K_{sv} value observed was 6.86 M^{-1} for iodide ion. From the slopes of the two linear components of the Stern-Volmer constants, K_{sv1} of 3.56 M^{-1} and K_{sv2} of 1.69 M^{-1} were obtained for cesium in denatured conditions. K_{sv1} is considerably greater than K_{sv2}, giving substantial evidence for selective tryptophan quenching. Upon denaturation of protein with 6 M GdnHCl, fraction accessible has been increased to 54% for cesium ions (from 33%); while same has been decreased to 78% (from 100%) for iodide ions. This may be due to change in conformation of the protein upon unfolding. The extent of quenching achieved with cesium is significantly higher than iodide ions, as evident from Table 5.3

and 5.4. Thus, denaturation of protein with 6 M GdnHCl resulted in the reorientation of charge density around trp from positive to negative, that is, it is quite likely that a larger fraction of tryptophan residues have positively charged residues in their closer proximity.

5.3.5 Solute quenching studies for active site mutants

The exposure and environment of the tryptophan residues in active site mutants was investigated by solute quenching technique using steady state fluorescence studies. Ten different mutant proteins, namely F30V, I31N, R51G, D77G, S136A, Y170H, K174M, V200E, S212G and H215L were used in this study. Substitution type mutations were designed in order to have minimal effect on secondary structure and 3D conformation of protein [150]. Tryptophan fluorescence spectra of wild type Ll-CCRH1 and mutants were similar, suggesting that mutant proteins are properly folded. λ_{max} for active site mutants were observed in the range of 350-352.5 nm, suggesting apparently overall unchanged conformation (Fig. 5.6).

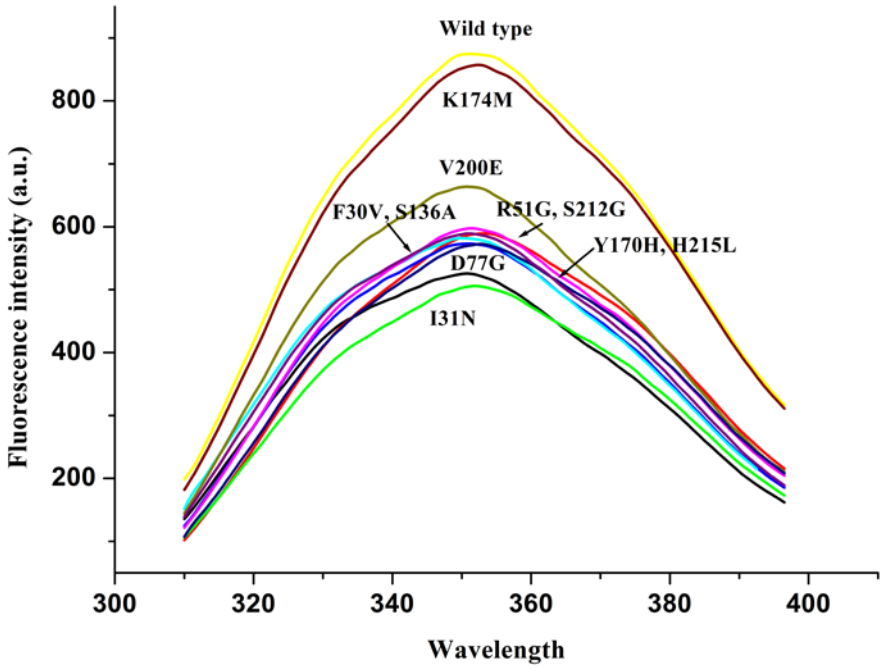


Fig. 5.6 Fluorescence emission spectra of wild type and active site mutants Ll-CCRH1.

5.3.5.1 Quenching of mutants fluorescence with acrylamide

Quenching of the intrinsic fluorescence of all active site mutants gave linear Stern-Volmer plots with K_{sv} values as 6-20 M^{-1} and 8-14 M^{-1} for native and denatured proteins, respectively, indicating differential degree of quenching (Table 5.3 and 5.4). Among the three quenchers used, acrylamide was the most efficient quencher for intrinsic fluorescence of Ll-CCRH1 mutants. Almost 100% of the trp fluorescence of mutants was accessible to acrylamide except V200E mutant, where around 80% of the total fluorescence is accessible under native condition. Accessibility of acrylamide to fluorescence increased to 100% upon denaturation of mutant proteins with 6 M GdnHCl (Table 5.3 and 5.4). An upward curvature was obtained for all mutants in the Stern-Volmer plots, indicating the involvement of both collisional and static components.

5.3.5.2 Quenching with cesium ions

The quenching profile obtained for native and denatured Ll-CCRH1 mutants with cesium ion showed downward curvature. The biphasic curvature obtained for CsCl quenching indicated heterogeneous ionic environment around the trp with presence of two conformers- one getting quenched earlier than other. Downwards curves were divided into two linear components and the values of K_{sv1} and K_{sv2} obtained were 0.7-2.4 M^{-1} and 0.2-0.9 M^{-1}, respectively for native state and, 0.9-1.4 M^{-1} and 0.4-1.0 M^{-1}, respectively for denatured state. K_{sv1} is considerably greater than the K_{sv2} in each case, suggesting selective tryptophan quenching. Based on the f_a values, 33% and 54% of the total fluorescence was found to be accessible to cesium ions in native and denatured wild type protein respectively. 20-40% accessibility was obtained for CsCl in F30V, I31N, R51G, D77G, S136A, K174M, S212G and H215L mutants, showing apparently

similar quenching profile as that of native wild type protein (Table 5.3). Mutants Y170H and V200E showed only 13% accessibility of fluorescence suggesting lower extent of quenching with cesium ions. For cesium ions, the fraction accessible increased to 26%, 32%, 52% and 33% for mutants (upto 2 fold) Y170H, V200E, S212G and H215L respectively after denaturation. However, analysis of the quenching data shows that there is slight increase in the accessibility of fluorescence (25-40%) compared to wild type Ll-CCRH1, indicating there may be partial unfolding or minor change in the conformation after denaturation (Table 5.4).

5.3.5.3 Quenching with iodide

The Stern-Volmer plots for the quenching of Ll-CCRH1 with iodide under native and denatured conditions showed linear dependence. Quenching of the intrinsic fluorescence of active site mutants with K_{sv} values, 2.3-4.3 M^{-1} and 0.8-4.1 M^{-1} for native and denatured mutant proteins respectively, indicating lower rate of quenching as compared to wild type Ll-CCRH1 (Table 5.3 and 5.4). 100% and 78% accessibility was observed for KI in native and denatured wild Ll-CCRH1 respectively, showing reorientation of charge density around trp from positive (native state) to negative (denatured state). From table 5.3, it is seen that 100%, 100% and 89% of the total fluorescence of the mutants S212G, K174M and R51G is accessible to iodide ions respectively, similar to wild type Ll-CCRH1; and it is decreased to 50-70% approximately in mutants F30V, I31N, D77G, S136A, Y170H, V200E and H215L. Denaturation of protein with 6 M GdnHCl led to decrease in accessibility (60-70%) with respect to all mutants, similar to denatured wild Ll-CCRH1 (78%) (Table 5.4).

Table 5.3 Summary of parameters obtained from the intrinsic fluorescence quenching of wild Ll-CCRH1 and different active site mutants under native conditions with different quenchers.

	Acrylamide			**Cscl**			**KI**		
	K_{sv}	***f_a***	**K_a**	**K_{sv1}/K_{sv2}**	***f_a***	**K_a**	**K_{sv}**	***f_a***	**K_a**
Native	23.15	1.17	12.78	1.31/0.82	0.33	5.69	5.35	1.23	5.47
F30V	13.05	1.02	10.08	1.31/0.48	0.27	3.20	3.78	0.69	7.68
I31N	15.09	0.93	15.15	1.26/0.72	0.23	26.02	3.45	0.64	10.68
R51G	20.41	1.05	12.78	1.08/0.85	0.39	11.21	4.17	0.89	5.47
D77G	9.86	0.91	10.62	2.34/0.38	0.20	27.11	2.47	0.66	7.26
S136A	11.37	0.89	13.37	1.31/0.57	0.20	15.37	2.84	0.58	10.29
Y170H	10.37	0.90	11.58	0.40/0.27	0.13	18.9	2.31	0.58	10.60
K174M	14.64	1.01	12.22	0.73/0.47	0.29	4.50	3.42	2.08	1.30
V200E	6.07	0.82	12.1	0.79/0.23	0.13	5.98	2.54	0.59	11.44
S212G	19.17	1.0	13.90	0.94/0.64	0.32	4.77	4.26	1.33	1.71
H215L	8.80	0.91	9.12	1.18/0.42	0.18	17.02	2.40	0.52	13.96

Table 5.4 Summary of parameters from Stern-Volmer and modified Stern-Volmer analysis of the intrinsic fluorescence quenching of wild Ll-CCRH1 and different active site mutants under denatured conditions with different quenchers.

	Acrylamide			**Cscl**			**KI**		
	K_{sv}	***f_a***	**K_a**	**K_{sv1}/K_{sv2}**	***f_a***	**K_a**	**K_{sv}**	***f_a***	**K_a**
Denatured	21.51	1.55	4.47	3.56/1.69	0.54	11.71	6.86	0.78	10.74
F30V	12.59	0.97	9.27	1.33/0.77	0.32	7.72	4.02	0.68	8.33
I31N	12.84	1.02	8.44	1.25/0.83	0.37	7.11	3.75	0.70	9.03
R51G	11.58	1.48	4.47	1.39/0.90	0.31	11.71	3.27	0.59	10.74
D77G	12.65	1.02	8.01	0.97/0.66	0.28	9.83	3.60	0.61	11.21
S136A	11.34	1.37	4.23	1.08/0.41	0.27	10.11	0.96	0.68	9.34
Y170H	8.86	1.21	4.75	0.92/0.73	0.26	11.45	3.24	0.60	9.35
K174M	11.64	0.91	10.70	1.11/0.68	0.36	5.24	3.24	0.60	8.75
V200E	10.04	1.14	6.25	1.20/0.76	0.32	8.30	3.16	0.66	8.91
S212G	12.15	0.94	10.79	1.08/0.85	0.52	2.91	3.83	0.66	10.42
H215L	11.33	1.18	5.43	1.25/0.92	0.33	7.43	3.87	0.64	10.84

The extent of quenching for cesium and iodide ions under native and denatured conditions observed in active site mutants is significantly different from wild type Ll-CCRH1 under same conditions. Thus, substitution type mutations of active site residues showed heterogeneity in tryptophan microenvironment and differential degree of conformation of protein under native or denatured conditions.

Summary

Leucaena leucocephala, a multipurpose nitrogen fixing tropical tree legume, has a tremendous prospective as a raw material for paper & pulp industry, nutritious forage, timber, organic fertilizer, firewood, industrial fuel and depilatory agent due to its adaptability to thrive under farthest agro-climatic conditions [3-5]. As the second most abundant natural (terrestrial) organic polymer, eclipsed only by cellulose, lignin is a major constituent of wood and accounts, on an average, 25% of the terrestrial plant biomass [9,10]. Lignin plays a major role in growth and development of plants such as rigidity and strength to cell wall, water/ nutrient transport in conducting cells and also provides physico-chemical barrier against pathogen attack [17,106]. However, an agro-economical opinion considers lignin as an obstacle for utilization of plant biomass because it affects paper manufacture and limits digestibility of forage crops [19].

Lignins are complex, heterogeneous, three dimensional aromatic polymers derived from oxidative polymerization of three monolignols namely, *p*-coumaryl alcohol (H unit), coniferyl alcohol (G unit, more compact) and sinapyl alcohol (S unit, less compact) differing in their degree of methoxylation [11,13,14]. In plants, biosynthesis of monolignols is a specialized branch of phenylpropanoid metabolism, a complex series of branching biochemical reactions responsible for synthesis of variety of products like lignin, flavonoids and hydroxycinnamic acid conjugates [11,29-31]. The biosynthesis of lignin precursors proceed through common phenylpropanoid pathway, starting with conversion of phenylalanine to cinnamate and subsequent formation of hydroxycinnamoyl CoA esters via hydroxylation and methylation at different positions on the aromatic ring. These esters are end products of common phenylpropanoid metabolism. In monolignol biosynthesis, cinnamoyl CoA reductase (CCR, EC 1.2.1.44) catalyzes the NADPH dependent reduction of cinnamoyl esters to corresponding hydroxycinnamaldehydes. These aldehydes are further channeled by

cinnamyl alcohol dehydrogenase (CAD) to form monolignol units (S, G, H) and finally polymerized into cross linked plant polymer lignin.

CCR activity is generally low in plants, indicating that it may play a crucial role as a rate limiting step in regulation of lignin biosynthesis [87]. As the first committed step in monolignol biosynthesis, CCR plays a key regulatory role by controlling overall carbon flux of metabolites towards lignin [70]. During the last two decades, significant headway has been made in characterizing CCR genes from variety of plants like *Arabidopsis*, poplar, *Medicago*, wheat etc. As CCR catalyzes a step devoted specifically to monolignol biosynthesis, its downregulation studies could be a potential avenue to improve pulping and saccharification efficiency in bioenergy crops. The reduction in lignin content through downregulation of CCR has been observed in tobacco [104,106,107], *Arabidopsis* [109], Norway spruce [110], poplar [102], alfalfa [111], tomato [112] and ryegrss [113]. All above studies show significance of CCR as a candidate gene in lignin biosynthesis, not only to reduce lignin content for paper and pulp industry but also to enhance the forage quality and produce optimal feedstock plants for biofuel production.

CCR exhibits the substrate specificity for different hydroxycinnamoyl CoA esters like *p*-coumaryl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA and reduce them into corresponding aldehydes respectively. CCR also carries out oxidation of hydroxycinnamldehydes to their corresponding CoA esters [89,91].

L. leucocephala; being one of the important resources for paper and pulp production in India, yet so far not much work has been done on structural-functional aspects of enzymes taking part in lignin formation in this species. Although CCR is one of the most investigated enzyme of lignin biosynthesis pathway, its three dimensional

structure remain to be determined. The lack of the three dimensional crystal structure of CCR has precluded a clarification of functional active site residues involved in substrate binding and catalysis. On the other hand, one of the molecular modeling and docking report on CCR has not provided any significant information as it has completely ignored the role of serine residues and signatural catalytic sequence (NWYCYGK) in CCR, required for catalysis. Also, little has been described about the significance of cofactor (NADPH) and its interaction with CCR enzyme. Thus, the functional features of CCR enzyme like mechanism of catalysis and multiple substrate specificity are still unresolved. Structure-functional studies on CCR reported in literature are very few; in fact no studies regarding conformation, intrinsic fluorescence and characterization of trp microenvironment have been reported so far. In this context, the present study is aimed at understanding the structural and functional aspects of cinnamoyl CoA reductase from *L. leucocephala* (Ll-CCRH1).

As a first step towards understanding the structure-function relationship of *Ll-CCRH1* (GenBank: DQ986907), we report the detailed biochemical characterization of recombinant Ll-CCRH1 by various biophysical and biochemical techniques. The recombinant Ll-CCRH1 cloned in pET 30b (+) plasmid was overexpressed in *E. coli* BL21 (DE3). The SDS-PAGE analysis showed that the soluble and pellet fraction of recombinant bacteria contained an overexpressed protein with an apparent molecular mass of 38 kDa. A simple and efficient three step protocol comprising of an affinity chromatography followed by anion-exchange chromatography and final gel filtration was found to be best suited for Ll-CCRH1, hence this protocol was optimized for purification of the protein from the cell lysate. The final purification step resulted in 53 fold purification with specific activity of 1.2 U/mg and relatively good recovery of 66%.

Optimum pH for forward and reverse reaction was found to be 6.5 and 7.8 respectively. The enzyme was most stable around pH 6.5 at 25 ^{o}C for 90 min. Activation energy, E_a for various substrates was found to be in the range of 20-50 kJ/mol. Involvement of probable carboxylate ion, histidine, lysine or tyrosine at the active site of enzyme was predicted by pH activity profile. SAXS studies of protein showed radius 3.04 nm and volume 49.25 nm^3 with oblate ellipsoid shape. Metal ion inhibition studies revealed that Ll-CCRH1 is a metal independent enzyme.

Lack of three dimensional structure of cinnamoyl CoA reductase (CCR) limits its detailed active site characterization studies. Active site residues involved in substrate/NADPH binding, stabilization and catalysis for Ll-CCRH1 were identified by amino acid sequence analysis and homology modeling. Protein-protein BLAST search of Ll-CCRH1 against PDB showed a hit against dihydroflavanol reductase from *Vitis vinifera* (PDB ID: 2c29) and it demonstrated 43% identity with Ll-CCRH1. Thus 2c29 was chosen as template for modeling of the Ll-CCRH1 protein on the basis of sequence similarity, residues completeness, crystal resolution and functional similarities. Out of 50 models generated by MODELLER for Ll-CCRH1; only 5 models were selected based on DOPE score. Multiple sequence alignment of Ll-CCRH1 with other homologous CCRs and also with template 2c29 was performed using ClustalW 2.0 to locate the conserved residues. The coenzyme NADPH binding site of Ll-CCRH1 is held by the structural motif known as Rossmann fold domain. NADPH binding site pocket consists of Gly26, Gly28, Gly29, Thr49, Leu78, Ser99, Ser135, Tyr170, Lys174, Pro197, Val198, Gln206 and Ala211 residues. All the residues except Ser99, Gln206, and Ala211 are conserved with NADPH binding site of the template 2c29. The substrate binding site of Ll-CCRH1 is made up of 10 residues, that is, Phe30, Ile31, Arg51, Leu64, Asp77, Ser136, Tyr170, Lys174, Val200 and Ser212. Residues

corresponding to the active site in Ll-CCRH1 are highly conserved among the CCRs from various plant species and also with template 2c29.

Docking simulations with five different hydroxycinnamoyl CoA esters showed feruloyl CoA as most preferred substrate and possible roles of Ser136, Tyr170, and Ser212 in catalysis, and Phe30, Ile31 and Arg51 in substrate binding were pointed out. Wild type recombinant Ll-CCRH1 enzyme showed highest K_{cat}/K_m for feruloy CoA (4.57 x 10^6 M^{-1} s^{-1}) indicating better affinity over other CoA esters. Apparent kinetic parameters for reverse reaction were also determined using coniferaldehyde and sinapaldehyde as substrates at pH 7.8 as well as carrying out the reaction at pH 6.5. Ll-CCRH1 showed comparatively higher affinity for coniferaldehyde than sinapaldehyde as reverse reaction substrate. At pH 7.8, the specificity constants for coniferaldehyde and sinapaldehyde were 7.5 and 5.3 (x 10^6 M^{-1} s^{-1}) respectively, and around 1.5-2.5 times higher than forward substrates (feruloyl CoA and sinapoyl CoA). On the other hand, reverse reaction at pH 6.5 (favorable for forward reaction) gives specificity constants 1.9 and 1.20 (x 10^6 M^{-1} s^{-1}) for both aldehydes, respectively. In plants, CCR catalyzed reduction reaction occurs in the pH range of 6-6.5. Comparison of kinetic parameters of the reduction (forward) and oxidation (reverse) reactions for Ll-CCRH1 at pH 6.5 showed that reduction of cinnamoyl CoA esters is favored over oxidation reaction in vitro too. Here, we report reverse reaction favored at pH 7.8 with higher specificity constants. Such reaction would not be possible in plants due to controlled and coordinated regulation of metabolic flux of lignin biosynthesis pathway genes.

Putative active site residues and proximal H215 were subjected for site directed mutagenesis, and mutated enzymes were expressed, purified and assayed to confirm their functional roles. Mutagenesis of S136, Y170, K174 and S212 exhibited catalytic efficiencies less than 10-15% of wild type, indicating their pivotal roles in catalysis.

1.2-4 fold increase in K_m values were observed with R51, V200, D77, F30 and I31 mutants; while double mutants F30V-I31N and R51G-D77G demonstrated 5-7 fold increase in K_m and 8-9 fold decrease in K_{cat} / K_m values specifying their role in substrate binding or stabilization. H215 mutant displayed about 4.5 fold reduction in K_{cat} / K_m. Further, treatment of Ll-CCRH1 with amino acid group specific chemical modifiers verified involvement of Tyr, His, Lys, Ser, Arg and carboxylate group in reaction mechanisms. Finally, substrate protection studies of Ll-CCRH1 confirmed presence of modified amino acid residues at the active site.

As mentioned earlier, CCR exhibits substrate specificity for different hydroxycinnamoyl CoA esters (*p*-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA); and reduce them to corresponding aldehydes. Cinnamoyl CoA esters are the common precursors of wide range of phenolic compounds and flavonoids. For example, coumaroyl CoA esters are the substrates for chalcone synthase enzyme, the first catalytic step towards secondary metabolites synthesis. Secondary metabolites are considered as the first line of defense against pathogens and diseases. Differential substrate specificity of CCR has been correlated to its exclusive or redundant function inside the cell either in lignification (feruloyl CoA/ sinapoyl CoA as most preferred substrate) during plant development or in defense mechanism (coumaroyl CoA as favored substrate) [97,99,100]. *In silico* mutagenesis studies of active site residues of Ll-CCRH1 were carried out. Homology modeling based modeled 3D structure of Ll-CCRH1 was used as template for *in silico* mutant preparations. Docking simulations of Ll-CCRH1 mutants with CoA esters by AutoDock Vina tools showed altered substrate specificity as compared to wild type. This differential substrate specificity was mainly due to conformational changes in substrate binding pocket or change in geometry/architecture/shape of active site or

increase/decrease in number of interactions following mutations or charge redistribution along active site and altered physicochemical properties of mutated residues. The altered substrate specificity for active site mutants suggests the possible physiological role of CCR either in lignin formation or in defense system in plants.

Conformational transitions of cinnamoyl CoA reductase; a key regulatory enzyme in lignin biosynthesis, from *L. leucocephala* (Ll-CCRH1) were studied using fluorescence and circular dichroism spectroscopy. The native Ll-CCRH1, a multi tryptophan protein (Trp 34, 169, 156 and 182; GenBank ID: DQ986907) showed maximum intensity of intrinsic fluorescence, λ_{max} at 352 nm indicating the trp residues to be exposed to solvent. Decomposition analysis of the intrinsic fluorescence profile revealed two populations/conformers of the trp, 1) class A or S (33%) existing in the hydrophobic environment at any given time and 2) class III (67%), existing in completely polar environment. Fluorescence intensities of the enzyme incubated in the ranges of pH 2-5 and pH 11-12 were significantly reduced as compared to that at pH 8.0 due to protonation and deprotonation of amino acids in the vicinity of tryptophan, respectively. λ_{max} of protein incubated at different pH were observed in the range 350-352 nm suggesting apparently unchanged overall conformation.

The native protein possesses 66% of helical structure, undergoes rapid structural transitions at and above 45 °C and starts forming aggregates at 55 °C. Ll-CCRH1 was transformed into acid induced (pH 2.0) molten globule like structure, exhibiting altered yet compact secondary structure, diminished tertiary structure and exposed hydrophobic residues. The molten globule like structure was examined for the thermal and chemical stability. The altered secondary structure of L1-CCRH1 at pH 2.0 was stable upto 90 °C and also in the vicinity of 2 M Guanidine hydrochloride (GdnHCl) (as compared to drastic loss of native structure in 2 M GdnHCl) as seen in far UV-CD

spectra. The structural transition of Ll-CCRH1 at pH 2.0 was reversible, as all the characteristics of molten globule had diminished after readjusting the pH to 8.0, that of native protein. While unfolding of the protein in GdnHCl was observed to be reversible based on MRE 219 nm that of acid induced molten globule like structure was irreversible

Fluorescence quenching and time resolved fluorescence of wild type recombinant Ll-CCRH1, a multitryptophan protein from *L. leucocephala* and ten different active site mutants were carried out to investigate tryptophan environment. Tryptophan residues appear to be uniquely sensitive to quenching by a variety of solutes as a result of a propensity of the excited indole nucleus to donate electrons while in the excited state. Fluorescence spectra of the native, denatured Ll-CCRH1 and active site mutants were recorded in the presence of different quenchers like acrylamide (neutral), cesium chloride (positively charged) and potassium iodide (negatively charged). The enzyme showed highest affinity for feruloyl CoA (K_a= 3.72 x 10^5 M^{-1}) over other CoA esters and cinnamaldehydes, as determined by fluorescence spectroscopy. Quenching of the fluorescence by acrylamide for wild type and active site mutants was collisional with almost 100% of the tryptophan fluorescence is accessible under native condition and remained same after denaturation of protein with 6 M GdnHCl. In wild type Ll-CCRH1, the extent of quenching achieved with iodide (f_a=1.0) was significantly higher than cesium ions (f_a=0.33) suggesting more density of positive charge around surface of trp conformers under native conditions. Denaturation of wild type protein with 6 M GdnHCl led to significant increase in the quenching with cesium (f_a=0.54), whereas quenching with iodide ion was decreased (f_a=0.78), indicating reorientation of charge density around trp from positive to negative and heterogeneity in trp environment. The Stern-Volmer plots for wild type and mutants Ll-

CCRH1 under native and denatured conditions with cesium ion yielded biphasic quenching profiles, indicating that the trp residues in the protein fall into at least two groups that differ considerably in their accessibility and/or environment. The extent of quenching for cesium and iodide ions under native and denatured conditions observed in active site mutants is significantly different from wild type Ll-CCRH1 under same conditions. Thus, single substitution type mutations of active site residues showed heterogeneity in tryptophan microenvironment and differential degree of conformation of protein under native or denatured conditions. The native enzyme showed two different lifetimes, $\tau 1$ (2.27 ns) and $\tau 2$ (7.92 ns) with average lifetime, (τ) 5.28 ns, which decreased $\tau 1$ (1.38 ns) and $\tau 2$ (3.59 ns) with (τ) 2.16 ns after denaturation with 6 M GdnHCl.

.

References

[1] M. Baucher, C. Halpin, M. Petit-Conil, W. Boerjan, Lignin: genetic engineering and impact on pulping, Crit. Rev. Biochem. Mol. 38 (2003) 305-350.

[2] C.J. Biermann, Handbook of pulping and papermaking, (Ed.), Academic press, San Diego, CA. 1996.

[3] C.E. Hughes, *Leucaena*: A genetic resources handbook. Tropical Forestry Papers No.37. Oxford Forestry Institute, Department of Plant science. Oxford university, UK., 1998.

[4] N. Vietmeyer, B. Cotton, F.R. Ruskin, *Leucaena* promising forage and tree crop for the tropics, first ed., National Academy of Science, Washington, D.C., 1977.

[5] R.J. Jones, The value of *Leucaena leucocephala* as feed for ruminants in the tropics, World Animal review 32 (1979) 13-23.

[6] S. Rastogi, U.N. Dwivedi, *Agrobacterium tumefaciens* mediated transformation of *Leucaena leucocephala* multipurpose tree legume, Physiol. Mol. Biol. Plants 9 (2003a) 207-216.

[7] A.C. Hammond, M.J. Allison, M.J. Williams, Persistence of DHP-degrading bacteria between growing seasons in subtropical Florida, *Leucaena* Research Report, 10, 1989.

[8] M.P. Hegarty, R.D. Court, P.M. Throne, The determination of mimosine and 3,4-DHP in biological material, Australia Journal of Agricultural Research, 15 (1964) 168.

[9] K.V. Sarkanen, C.H. Ludwig, Lignins: Occurrence, Formation, Structure, and Reactions, New York: Wiley Interscience, pp 916, 1971.

[10] R.W. Whetten, J.J. MacKay, R.R. Sederoff, Recent advances in understanding lignin biosynthesis, Annu. Rev. Plant Physiol. Mol. Biol. 49 (1998) 585-609.

[11] W. Boerjan, J. Ralph, M. Baucher, Lignin Biosynthesis, Annu. Rev. Plant Biol. 54 (2003) 519-546.

[12] P.H. Raven, R.F. Evert, S.E. Elchhorn, Biology of plants, 5th Ed., New York: Worth Publishers, 1992.

[13] K. Freudenberg, A.C. Neish, Constitution and biosynthesis of lignin, Springer-Verlag, Berlin, pp 129, 1968.

[14] N.G. Lewis, E. Yamamoto, Lignin: occurrence, biogenesis and biodegradation, Annu. Rev. Plant Physiol. Plant Mol. Biol. 41 (1990) 455-496.

[15] M.M Campbell, R.R. Sederoff, Variation in lignin content and composition. Mechanisms of control and implications for the genetic improvement of plants, Plant Physiol. 110 (1996) 3-13.

[16] R.W. Whetten, R.R. Sederoff, Lignin Biosynthesis, Plant cell 7 (1995) 1001-1013.

[17] L. Jones, A.R. Enos, S.R. Turner, cloning and characterization of irregular xylem (irx4): a severely lignin deficient mutant of *Arabidopsis,* Plant J. 26 (2001) 205-216.

[18] A.M. Boudet, Lignins and lignifications: Selected issues, Plant Physiol. Biochem. 38 (2000) 81-96.

[19] V.L. Chiang, R.J Puumala, H. Takeuchi, R.E. Eckert, Comparison of softwood and hardwood kraft pulping, Tappi J. 71 (1998) 173-176.

[20] L. Chen, C.K. Auh, P. Dowling, J. Bell, F. Chen, A. Hopkins, R.A. Dixon, Z.Y. Wang, Improved forage digestibility of tall fescue (*Festuca arundinacea*) by

transgenic down-regulation of cinnamyl alcohol dehydrogenase, Plant Biotechnol. J. 1 (2003) 437-449.

[21] F. Chen, R.A. Dixon, Lignin modification improves fermentable sugar yields for biofuel production, Nat. Biotechnol. 25 (2007) 759-761.

[22] T. Higuchi, Biosynthesis of Lignin, In: Higuchi, T. (ed.), Biosynthesis and biodegradation of wood components, Academic Press, New York, 141-160.

[23] C.J. Biermann, Essentials of Pulping and Papermaking, Academic Press, San Diego, CA. 1993.

[24] P. Axegard, B. Jacobson, S. Ljunggren, N.O. Nilvebrant, Bleaching of kraft pulps-A research perspective, Papier 46 (1992) V16-V25.

[25] J.H. Christensen, M. Baucher, A.P. O'Connell, M. Van Montagu, W. Boerjan, Control of lignin biosynthesis, Molecular Biology of Woody Plants 64 (2000) 227-267.

[26] S. Odendahl, Environmental protection and consumer demands: a review of trends and impacts, Pulp and Paper Canada 95 (1994) 144-148.

[27] S. Srivastava, R.K. Gupta, M. Arha, R.K. Vishwakarma, S.K. Rawal, P.B. Kavi Kishor, B.M. Khan, Expression analysis of cinnamoyl-CoA reductase (CCR) gene in developing seedlings of *Leucaena leucocephala*: A pulp yielding tree species, Plant Physiol. Biochem. 49 (2011) 138-145.

[28] J.M. Humphreys, C. Chapple, Rewriting the lignin roadmap, Curr. Opin. Plant Biol. 5 (2002) 224-229.

[29] G. Gross, Biosynthesis and metabolism of phenolic acids and monolignols. In Biosynthesis and degradation of wood components, Higuchi, T. (ed.), Academic Press, New York, 229-271.

[30] K. Hahlbrock, D. Scheel, Physiology and molecular biology of phenylpropanoid metabolism, Annu. Rev. Plant Physiol. Mol. Biol. 40 (1989) 347-369.

[31] R.A. Dixon, Natural products and disease resistance, Nature 411 (2001) 843-847.

[32] L. Davin, D.L. Bedgar, T. Katayama, N. Lewis, On the stereoselective synthesis of (+)-pinoresinol in *Forsythia suspensa* from its achiral precursor, coniferyl alcohol, Phytochemistry 31 (1992) 3869-3874.

[33] L. Davin, N. Lewis, An historical perspective of lignan biosynthesis: monolignol, allylphenol, and hydroxycinnamic acid coupling and downstream mechanisms, Phytochem. Rev. 2 (2003) 257-288.

[34] Y.Y. Kao, S.A. Harding, C.J. Tsai, Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen, Plant physiol. 130 (2002) 796-807.

[35] Y. Osakabe, Y. Ohtsubo, S. Kawai, Y. Katayama, N. Morohoshi, Structures and tissue specific expression of genes for phenylalanine ammonia-lyase from a hybrid aspen, Plant Sci. 105 (1995) 217-226.

[36] N.J. Bate, J. Orr, W. Ni, A. Meromi, T. Nadler-Hassar, P.W. Doerner, R.A. Dixon, C.J. Lamb, Y. Elkind, Quantitative relationship between phenylalalnine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate determining step in natural product synthesis, Proc. Natl. Acad. Sci. USA. 91 (1994) 7608-7612.

[37] D. Hatton, R. Sablowski, M.H. Yung, C. Smith, W. Schuch, M. Bevan, Two classes of cis sequences contribute to tissue-specific expression of a *PAL2* promoter in transgenic tobacco, Plant J. 7 (1995) 859-876.

[38] H.D. Jones, Phenylalanine ammonia-lyase: Regulation of its induction, and its role in plant development, Phytochemistry 23 (1984) 1349-1359.

[39] A. Kumar, B.E. Ellis, The phenylalanine ammonia-lyase gene family in raspberry. Structure, expression and evolution, Plant physiol. 127 (2001) 230-239.

[40] A. Levya, X. Liang, J.A. Pintor-Toro, R.A. Dixon, C.J. Lamb, cis-Element combinations determine phenylalanine ammonia-lyase gene tissue specific expression patterns, Plant Cell 4 (1992) 263-271.

[41] S. Ohl, S.A. Hedrick, J. Chory, C.J. Lamb, Functional properties of a phenylalanine ammonia lyase promoter from *Arabidopsis*, Plant Cell 2 (1990) 837-848.

[42] R. Croteau, T.M. Kutchan, N.G. Lewis, Natural Products, Biochemistry and Molecular Biology of Plants (2000) 1250-1318.

[43] J.M. Boniwell, V.S. Butt, Flavin nucleotide-dependent 3-hydroxyaltion of 4-hydroxyphenylpropanoid carboxylic acids by particulate preparations from potato tubers, Z Natureforsch 41 (1986) 56-60.

[44] M. Kojima, W. Takeuchi, Detection and characterization of p-coumaric acid hydroxylase in mung bean, *Vigna mungo*, seedlings, Biochem J. 105 (1989) 265-270.

[45] M. Petersen, D. Strack, U. Matern, Biosynthesis of phenylpropanoids and related compounds, In: Biochemistry of plant secondary metabolism, M. Wink (Ed.), 2 pp 151-222 Sheffield Academic Press, Sheffield, UK, 1999.

[46] H.A. Stafford, S. Dresler, 4-Hydroxycinnamic acid hydroxylase and polyphenolase activities in *Sorghum vulgare*, Plant Physiol. 49 (1972) 590-595.

[47] R. Franke, J.M. Humphreys, M.R. Hemm, J.W. Denault, M.O. Ruegger, J.C. Cusumano, C. Chapple, The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism, Plant J. 30 (2002) 33-45.

[48] R.B. Nair, Q. Xia, C.J. Kartha, E. Kurylo, R.N. Hirji, R. Datla, G. Selvaraj, *Arabidopsis CYP98A3* mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast, Plant Physiol. 130 (2002) 210-220.

[49] G. Schoch, S. Goepfert, M. Morant, A. Hehn, D. Meyer, P. Ullmann, D. Werck-Reichhart, CYP98A3 from *Arabidopsis thaliana* is a 3-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway, J. Biol. Chem. 276 (2001) 36566-36574.

[50] L. Hoffmann, S. Maur, F. Martz, P. Geoffre, M. Legrand, Purification, cloning and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism, J. Biol. Chem. 278 (2002) 95-103.

[51] X. Li, J.K. Weng, C. Chapple, The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids, Plant Cell 22 (2010) 1620-1632.

[52] L. Hoffmann, S. Besseau, P. Geoffroy, C. Ritzenthaler, D. Meyer, C. Lapierre, B. Pollet, M. Legrand, Silencing of hydroxycinnamoyl-Coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis, Plant Cell, 16 (2004) 1446-1465.

[53] L. Hoffmann, S. Besseau, P. Geoffroy, D. Meyer, C. Lapierre, B. Pollet, M. Legrand, Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth, Plant Cell, 19 (2007) 148-162.

[54] D. Lee, K. Meyer, C. Chapple, C.J. Douglas, Antisense suppression of 4-coumarate coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition, Plant Cell, 9 (1997) 985-998.

[55] W.J. Hu, A. Kawaoka, C.J. Tsai, J. Lung, K. Osakabe, H. Ebinuma, V.L. Chiang, Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase gene in aspen (*Populus tremuloides*), Proc. Natl. Acad. Sci. USA. 95 (1998) 5407-5412.

[56] W.J. Hu, S.A. Harding, J. Lung, J.L. Popko, J. Ralph, D.D. Stokke, C.J. Tsai, V.L. Chiang, Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees, Nat. Biotechnol. 17 (1999) 808-812.

[57] S.A. Harding, J. Leshkevich, V.L. Chiang, C.J. Tsai, Differential substrate inhibition couples kinetically distinct 4-coumarate: coenzyme A ligase with spatially distinct metabolic roles in quaking aspen, Plant physiol. 128 (2002) 428-438.

[58] J. Raes, A. Rohde, J.H. Christensen, Y.V. Peer, W. Boerjan, Genome-wide characterization of the lignifications toolbox in *Arabidopsis*, Plant Physiol. 133 (2003) 1051-1071.

[59] C. Lindermayr, J. Fliegmann, J. Ebel, Deletion of a single amino acid residue from different 4-coumarate-CoA ligase from soyabean results in the generation of new substrate specificities, J. Biol. Chem. 278 (2003) 2781-2786.

[60] K.S. Woo, R.W. Whetten, D.M. O'Malley, R.R. Sederoff, 4-coumarate: coenzyme A ligase from loblolly pine xylem. Isolation, characterization and complementary DNA cloning, Plant Physiol. 108 (1995) 85-97.

[61] C. Grand, A. Boudet, A.M. Boudet, Isoenzymes of hydroxycinnamte: CoA ligase from poplar stems-properties and tissue distribution, Planta 158 (1983) 225-229.

[62] T. Higuchi, Lignin biochemistry: biosynthesis and biodegradation, Wood Sci. Technol. 24 (1990) 23-63.

[63] J. Ralph, R.D. Hatfield, J. Piquemal, N. Yahiaoui, M. Pean, C. Lapierre, A.M. Boudet, NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignifications enzymes cinnamyl alcohol dehydrogenase and cinnamoyl CoA reductase, Proc. Natl. Acad. Sci. USA. 95 (1998) 12803-12808.

[64] A.E. Pakusch, R.E. Kneusel, U. Matern, S-Adenosyl- L-methionine: trans-caffeoyl coenzyme A 3-O-methyltransferase from elicitor treated parsley cell suspension cultures, Arch. Biochem. Biophy. 271 (1989) 488-494.

[65] Z.H. Ye, R.E. Kneusel, U. Matern, J.E. Varner, An alternative methylation pathway in lignin biosynthesis in *Zinnia,* Plant Cell, 6 (1994) 1427-1439.

[66] D. Schmitt, A.E. Pakusch, U. Matern, Molecular cloning, induction, and taxonomic distribution of Caffeoyl-CoA 3-O-methyltransferase, an enzyme involved in disease resistance, J. Biol. Chem. 266 (1991) 17416-17423.

[67] F. Martz, S. Maury, G. Pincon, M. Legrand, cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme, Plant Mol. Biol. 36 (1998) 427-437.

[68] Z.H. Ye, Association of caffeoyl Coenzyme A 3-O-methyltransferase expression with lignifying tissues in several dicot plants, Plant Physiol. 115 (1997) 1341-1350.

[69] Z.H. Ye, J.E. Varner, Differential expression of two O-methyltransferases in lignin biosynthesis in *Zinnia elegans*, Plant Physiol. 1084 (1995) 459-467.

[70] E. Lacombe, S. Hawkins, J. Piquemal, J. Van Doorsselaere, D. Goffner, O. Poeydomenge, A. Boudet, J. Grima- Pettenati, Cinnamoyl-CoA reductase, the first committed enzymes of lignin branch biosynthetic pathway: cloning, expression, and phylogenetic relationships, Plant J. 11 (1997) 429-441.

[71] C. Grand, Ferulic acid 5-hydroxylase: A new cytochrome P-450-dependent enzyme from higher plant microsomes involved in lignin biosynthesis, FEBS Lett. 169 (1984) 7-11.

[72] J.M. Humphreys, M.R. Hemm, C. Chapple, New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase, Proc. Natl. Acad. Sci. USA. 96 (1999) 10045-10050.

[73] L. Li, J.L. Popko, T. Umezawa, V.L. Chiang, 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms, J. Biol. Chem. 275 (2000) 6537-6545.

[74] K. Osakabe, C. Tsao, L. Li, J.L. Popko, T. Umezawa, D.T. Carraway, R.H. Smeltzer, C.P. Joshi, V.L. Chiang, Coniferyl aldehyde 5-hydroxyaltion and methylation direct syringyl lignin biosynthesis in angiosperms, Proc. Natl. Acad. Sci. USA. 96 (1999) 8955-8960.

[75] M. Baucher, B. Chabbert, G. Pilate, J. Van Doorsselaere, M. Tollier, M. Petit-Conil, B. Monties, M. Van Montagu, D. Inze, L. Jouanin, W. Boerjan, Red xylem and higher lignin extractability by down regulating a cinnamyl alcohol dehydrogenase in poplar, Plant Physiol. 112 (1996) 1479-1490.

[76] M. Baucher, M.A. Bernard-Vailhe, B. Chabbert, J.M. Besle, C. Opsomer, M. Van Montagu, J. Botterman, Downregulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility, Plant Mol. Biol. 39 (1999) 437-447.

[77] L. Li, X.F. Cheng, J. Leshkevich, T. Umezawa, S.A. Harding, V.L. Chiang, The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase, Plant Cell, 13 (2001) 1567-1585.

[78] E.K. Bomati, J.P. Noel, Structural and kinetic basis for substrate selectivity in *Populus tremuloides* sinapyl alcohol dehydrogenase, Plant Cell 17 (2005) 1598-1611.

[79] C. Steeves, H. Forster, U. Pommer, R. Savidge, Coniferyl alcohol metabolism in conifers. Glucosidic turnover of cinnamyl aldehydes by UDPG: Coniferyl alcohol glucosyltransferase from pine cambium, Phytochem. 57 (2001) 1085-1093.

[80] A. Zimmerlin, P. Wojtaszek, G.P. Bolwell, Synthesis of dehydrogenation polymers of ferulic acid with high specificity by a purified cell wall peroxidase from French bean (*Phaseolus vulgaris* L.) Biochem J. 299 (1994) 747-753.

[81] M.D. Brownleader, N. Ahmed, M. Trevan, M.F. Chaplin, P.M. Dey, Purification and partial characterization of tomato extension peroxidase, Plant Physiol. 109 (1995) 1115-1123.

[82] R.L. Hinman, J. Lang, Peroxidase catalyzed oxidation of indole-3-acetic acid, Biochemistry 4 (1965) 144-158.

[83] J.H. Christensen, M. Van Montagu, G. Bauw, W. Boerjan, Xylem peroxidases: purification and altered expression, Molecular Breeding of Woody Plants 18 (2001) 171-176.

[84] K. Marjamaa, K. Hilden, E. Kukkola, M. Lehtonen, H. Holkeri, P. Haapaniemi, S. Koutaniemi, T.H. Teeri, k. Fagerstedt, T. Lundell, Cloning, characterization and localization of three novels class III peroxidase in lignifying xylem of Norway spruce (*Picea abies*), Plant Mol. Biol. 61 (2006) 719-732.

[85] M. Quiroga, C. Guerrero, M.A. Botella, A. Barcelo, I. Amaya, M.I. Medina, F.J. Alonso, S.M. Forchetti, H. Tigier, V. Valpuesta, A tomato peroxidase involved in the synthesis of lignin and suberin, Plant Physiol. 122 (2000) 1119-1127.

[86] Y. Sato, M. Sugiyama, R.J. Gorecki, H. Fukuda, A. Komamine, Interrelationship between lignin deposition and the activities of peroxidase isoenzymes in differentiating tracheary elements of Zinnia. Analysis using Lα-aminoxy-β-phenylpropionic acid and 2-aminoindan-2-phosphonic acid, Planta 189 (1993) 584-589.

[87] Q. Ma, Tian B, Biochemical characterization of a cinnamoyl- CoA reductase from wheat, Biol. Chem. 386 (2005) 553–560.

[88] G.G. Gross, W. Kreiten, Reduction of Coenzyme A thioesters of cinnamic acids with an enzyme preparation from lignifying tissue of Forsythia, FEBS Lett. 54 (1975) 259-262.

[89] T. Luderitz, H. Grisebach, Enzymic synthesis of lignin precursors Comparison of Cinnamoyl-CoA reductase and Cinnamyl Alcohol:NADP dehydrogenase from spruce (*Picea abies* L) and soybean (*Glycine max* L), Eur. J Biochem. 119 (1981) 115–124.

[90] H. Wengenmayer, J. Ebel, H. Griesebach, Enzymic synthesis of lignin precursors: purification and properties of a cinnamoyl- CoA:NADPH reductase from cell suspension cultures of soybean *(Glycine max)*, Eur. J. Biochem. 65 (1976) 529-536.

[91] F. Sarni, C. Grand, A. Boudet, Purification and properties of cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase from poplar stems (*Populus X euramerican*), Eur. J Biochem. 139 (1984) 259–265.

[92] D. Goffner, M.M. Campbell, A.M. Boudet, Purification and characterization of cinnamoyl-Coenzyme A: NADP oxidoreductase in *Eucalyptus gunnii*, Plant Physiol. 106 (1994) 625-632.

[93] K. Larsen, Cloning and characterization of a ryegrass (Lolium perenne) gene encoding cinnamoyl CoA reductase (CCR), Plant Sci. 166 (2004) 569-581.

[94] K. Larsen, Molecular cloning and characterization of cDNAs encoding cinnamoyl CoA reductase (CCR) from barley (*Hordeum vulgare*) and potato (*Solanum tuberosum*), J. Plant Physiol. 161 (2004) 105-112.

[95] M. Baltas, C. Lapeyre, F. Bedos-Belval, M. Maturano, P. Saint-Aguet, L. Roussel, H. Duran, J. Grima-Pettenati, Kinetic and inhibition studies of cinnamoyl-CoA reductase 1 from *Arabidopsis thaliana*, Plant Physiol. Biochem. 43 (2005) 746–753.

[96] Y. Hu, P. Di, J. Chen, Y. Xiao, L. Zhang, W. Chen, Isolation and characterization of a gene encoding cinnamoyl CoA reductase from *Isatis indigotica* Fort, Mol. Biol. Rep. 38 (2011) 2075-2083.

[97] L. Li, X. Cheng, S. Lu, T. Nakatsubo, T. Umezawa, V.L. Chiang, Clarification of cinnamoyl co-enzyme A reductase catalysis in monolignol biosynthesis of aspen, Plant Cell Physiol. 46 (2005) 1073-1082.

[98] M. Pichon, I. Courbon, M. Bechert, A.M. Boudet, J. Grima-Pettenati, Cloning and characterization of two maize cDNAs encoding cinnamoyl CoA reductase (CCR) and differential expression of the corresponding genes, Plant Mol. Biol. 38 (1998) 671-676.

[99] R. Zhou, L. Jackson, G. Shadle, J. Nakashima, S. Temple, F. Chen, R.A. Dixon, Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*, Proc. Natl. Acad. Sci. USA. 1071 (2010) 17803-17808.

[100] L. Escamilla-Trevino, H. Shen, S. Uppalapati, T. Ray, Y. Tang, T. Hernandez, Y. Yin, Y. Xu, R.A. Dixon, Switchgrass (*Panicum virgatum*) possesses a divergent family of cinnamoyl CoA reductases with distinct biochemical properties, New Phytol. 185 (2010) 143–155.

[101] Q.H. Ma, Characterization of cinnamoyl CoA reductase that is associated with stem development in wheat, J. Exp. Bot. 58 (2007) 2011-2021.

[102] J.C. Leple, R. Dauwe, K. Morrel, V. Storme, C. Lapierre, B. Pollet, A. Naumann, W. Boerjan, Downregulation of Cinnamoyl-Coenzyme A Reductase in Poplar: Multiple-Level Phenotyping Reveals Effects on Cell Wall Polymer Metabolism and Structure, The Plant Cell 19 (2007) 3669-3691.

[103] R. McInnes, A. Lidgett, D. Lynch, H. Huxley, E. Jones, N. Mahoney, G. Spanganberg, Isolation and characterization of cinamoyl CoA reductase gene from perennial ryegrass (*Lolium perenne* L.), J. Plant Physiol. 159 (2002) 415-422.

[104] J. Piquemal, C. Lapierre, K. Myton, A. O'Conell, W. Schuch, J. Grima-Pettenati, A.M. Boudet, Down-regulation in cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants, Plant J. 13 (1998) 71–83.

[105] V. Lauvergeat, C. Lacommer, E. Lacombe, E. Lasserre, D. Roby, J. Grima-Pettenati, Two cinamoyl CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria, Phytochemistry 57 (2001) 1187-1193.

[106] M. Chabannes, A. Barakate, C. Lapierre, J.M. Marita, J. Ralph, M. Pean, S. Danoun, C. Halpin, J. Grima-Pettenati, A.M. Boudet, Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants, Plant J. 28 (2001a) 257-270.

[107] J.C. Abbott, A. Barakate, G. Pincon, M. Legrand, C. Lapierre, I. Mila, W. Schusch, C. Halpin, Simultaneous suppression of multiple genes by single transgenes. Down-regulation of three unrelated lignin biosynthetic genes in Tobacco, Plant Physiol. 128 (2002) 844-853.

[108] R. Dauwe, K. Morreel, G. Goeminne, B. Gielen, A. Rohde, J. Van Beeumen, J. Ralph, A.M. Boudet, J. Kopka, S. Rochange, C. Halpin, E. Messens, W. Boerjan, Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration, Plant J. 52 (2007) 263–285.

[109] T. Goujon, V. Ferret, I. Mila, B. Pollet, K. Ruel, V. Burlat, J.P. Joseleau, Y. Barriere, C. Lapierre, L. Jouanin, Downregulation of the AtCCR1 gene in *Arabidopsis*

thaliana: effects on phenotype, lignins and cell wall degrability, Planta 217 (2003) 218-228.

[110] J. Wadenback, S. Von Arnold, U. Egertsdotter, M.H. Walter, J. Grima-Pettenati, D. Goffner, G. Gellerstedt, T. Gullion, D. Clapham, Lignin biosynthesis in transgenic Norway spruce plants harboring an antisense construct for cinnamoyl CoA reductase, Transgenic Res. 17 (2008) 379-392.

[111] L.A. Jackson, G.L. Shadle, R. Zhou, J. Nakashima, F. Chen, R.A. Dixon, Improving saccharification efficiency of alfalfa stems through modification of terminal stages of monolignol biosynthesis, Bioenergy Res. 1 (2008) 180-192.

[112] B. Van der Rest, S. Danoun, A.M. Boudet, S.F. Rochange, Downregulation of cinnamoyl CoA reductase in tomato (*Solanum lycopersicum* L.) induces dramatic change in soluble phenolic pools, J. Exp. Bot. 57 (2006) 1399-1411.

[113] Y. Tu, S. Rochfort, Z. Liu, Y. Ran, M. Griffith, P. Badenhorst, G. Louie, M. Bowman, K. Smith, J.P. Noel, A. Mouradov, G. Spangenberg, Functional analysis of caffeic acid O-methyltransferase and cinnamoyl CoA reductase genes from perennial ryegrass (*Lolium perenne*), Plant Cell, 33 (2010) 3357-3373.

[114] N. Prasad, V. Vindal, V. Kumar, A. Kabra, N. Phogat, M. Kumar, Structural and docking studies *Leucaena leucocephala* Cinnamoyl CoA reductase, J. Mol. Model. 17 (2011) 533-541.

[115] P. Petit, T. Granier, B. d'Estaintot, C. Manigand, K. Bathany, J.M. Schmitter, V. Lauvergeat, S. Hamdi, B. Gallois, Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis, J. Mol. Biol. 368 (2007) 1345-1357.

[116] T.G. Oas, P.S. Kim, A peptide model of a protein folding intermediate, Nature 336 (1988) 42-48.

[117] T.E. Creighton, Protein folding, Biochem J. 270 (1990) 1-16.

[118] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133-7155.

[119] K.A. Dill, Theory for the folding and stability of globular proteins, Biochemistry 24 (1985) 1501-1509.

[120] K.A. Dill, S.B. Ozakan, M.S. Shell, T.R. Weikl, The protein folding problem, Annu. Rev. Biophys. 37 (2008) 289-316.

[121] H. Feng, Z. Zhou, Y. Bai, A protein folding pathway with multiple folding intermediates at atomic resolution, Proc. Natl. Acad. Sci. USA. 102 (2005) 5026-5031.

[122] D.A. Simmons, L. Konermann, Characterization of transient protein folding intermediates during myoglobin reconstitution by time-resolved electrospray mass spectrometry with on-line isotopic pulse labeling, Biochemistry 41 (2002) 1906-1914.

[123] C.M. Dobson, Protein folding and misfolding, Nature 426 (2003) 18-25.

[124] K. Kuwajima, The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure, Proteins: Struct. Funct. Genet. 6 (1989) 87-103.

[125] O.B. Ptitsyn, Protein folding: Hypotheses and experiments, J. Protein Chem. 6 (1987) 273-293.

[126] O.B. Ptitsyn, Molten Globule and protein folding, Adv. Protein Chem. 47 (1995) 83-229.

[127] D.A. Dolgikh, R.I. Glimanshin, E.V. Brazhnikov, V.E. Bychkova, G.V. Semisotnov, S.Y. Venyaminov, O.B. Ptitsyn, Alpha-Lactalbumin: compact state with fluctuating tertiary structure?, FEBS Lett. 136 (1981) 311-315.

[128] M. Arai, K. Kuwajima, Role of the molten globule state in protein folding, Adv. Protein Chem. 53 (2000) 209-271.

[129] S.K. Haq, S. Rasheedi, R.H. Khan, Characterization of a partially folded intermediate of stem bromelain at low pH, Eur. J. Biochem. 269 (2002) 47-52.

[130] P.L. Privalov, Intermediate states in protein folding, J. Mol. Biol. 258 (1996) 707-725.

[131] P.L. Privalov, Stability of proteins: small globular proteins, Adv. Protein Chem. 33 (1979) 167-241.

[132] P.L.Privalov, E.I. Tiktopulo, S. Venyaminov, Y. Griko, G. Makhatadze, N. Khechinashvili, Heat capacity and conformation of proteins in the denatured state, J. Mol. Biol. 205 (1989) 737-750.

[133] P.S. Kim, R.L. Baldwin, Intermediates in the folding pathway of small proteins, Annu. Rev. Biochem. 59 (1990) 631-660.

[134] M. Vendruscolo, J. Zurdo, C.E. MacPhee, C.M. Dobson, Protein folding and misfolding: A paradigm of self assembly and regulation in complex biological systems, Philos Transact A Math Phy Eng Sci 361 (2003) 1205-1222.

[135] J. Frydmann, Folding of newly translated proteins in vivo: The role of molecular chaperones, Annu. Rev. Biochem, 70 (2001) 603-647.

[136] W.F. DeGrado, Z.R. Wasserman, J.D. Lear, Protein design, a minimalist approach, Science 243 (1989) 622-628.

[137] B.M. Khan, S.K. Rawal, M. Arha, S.K. Gupta, S. Srivastava, N.M. Shaikh, A.K. Yadav, P.S. Kulkarni, O.U. Abhilash, S. Gupta, S. Omer, R.K. Vishwakarma, S. Singh, R.J. Santoshkumar, Prashant Sonawane, P. Patel, K. Chinnathambi, S. Abbassi, (2012) Genetic Engineering of Phenylpropanoid Pathway in *Leucaena leucocephala*, In: Genetic Engineering- Basics, new applications and responsibilities InTech Croatia, 93-120, ISBN 978-953-307-671-3

[138] H.S. Bhonsle, A.M. Korwar, S.S. Kote, S.B. Golegaonkar, A.D. Chougale, M.L. Shaik, N.L. Dhande, A.P. Giri, K.M. Shelgikar, R. Boppana, M.J. Kulkarni, Low plasma albumin levels are associated with increased plasma protein glycation and HbA1c in diabetes, J. Proteome Res. 11 (2012) 1391-1396.

[139] F.Y. Cheng, K. Blackburn, Y.M. Lin, M.B. Goshe, J.D. Williamson, Absolute protein quantification by LC/MS (E) for global analysis of salicylic acid-induced plant protein secretion responses, J. Proteome Res. 8 (2009) 82−93.

[140] K.S. Shashidhara, S.M. Gaikwad, Class II α-mannosidase from *Aspergillus fischeri*: Energetics of catalysis and inhibition, Int. J. Biol. Macromol. 44 (2009) 112–115.

[141] A. Kumar, S.M. Gaikwad, Jack bean α-mannosidase (Jbα-man): Tolerance to alkali, chelating and reducing agents and energetics of catalysis and inhibition, Int. J. Biol. Macromol. 49 (2011) 1066–1071.

[142] K. Sharma, V. Aswal, G. Kumarwamy, Adsorption of nonionic surfactant on silica nanoparticles: structure and resultant inter particle interactions, J. Phy. Chem. 114 (2010) 10984-10996.

[143] J. Stockigt, M. Zenk, Chemical synthesis and properties of hydroxycinnamoyl Coenzyme A derivatives, Z Naturfosch. 30 (1975) 352-358.

[144] D. Laskar, M. Jourdes, A. Patten, G. Helms, L. Davin, N.G. Lewis, The Arabidopsis cinnamoyl CoA reductase irx4 mutant has a delayed but coherent (normal) program of lignifications, Plant J. 48 (2006) 674–686.

[145] P. Sonawane, R.K. Vishwakarma, B.M. Khan, Biochemical characterization of recombinant cinnamoyl coA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, Int. J. Biol. Macromol. 58 (2013) 154-159.

[146] L. Wen, Z. Miao, W. Qing, Chemical modification of Xylanase from *Trichosporon cutaneum* shows the presence of carboxyl groups and cysteine residues essential for enzyme activity, Protein J. 18 (1999) 677-686.

[147] H. Balkrishnan, L. Satyanarayana, S.M. Gaikwad, C.G. Suresh, Structural and active site modification studies implicate Glu, Trp and Arg in the activity of xylanase from alkalophilic *Bacillus* sp. (NCL 87-6-10), Enzyme Microb. Tech. 39 (2005) 67-73.

[148] M. Gote, M. Khan, J. Khire, Active site directed chemical modification of α-galactosidase from *Bacillus stearothermophilus* (NCIM 5146): Involvement of lysine, tryptophan and carboxylate residues in catalytic site, Enzyme Microb. Tech. 40 (2007) 1312-1320.

[149] O. Trott, A. Olson, Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455-461.

[150] P. Sonawane, K. Patel, R.K. Vishwakarma, S. Srivastava, S. Singh, S.M. Gaikwad, B.M. Khan, Probing the active site of cinnamoyl CoA reductase 1 (Ll-CCRH1) from *Leucaena leucocephala*, Int. J. Biol. Macromol. 60 (2013) 33-38.

[151] N. Eswar, M.A. Marti-Renom, B. Webb, M.S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative protein structure modeling with MODELLER, Current protocols in bioinformatics, Wiley suppl 15, 5.6.1-5.6.30.

[152] A. Sali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779-815.

[153] M. Prokop, J. Adam, Z. Kriz, M. Wimmerova, J. Koca, Triton: A graphical tool for ligand-binding protein engineering, Bioinformatics 24 (2008) 1955-1956.

[154] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: A program to check the stereochemical quality of protein structures, J. Appl.Crystallogr. 26 (1993) 283-291.

[155] A. Kobata, Glycobiology: an expanding research area in carbohydrate chemistry, Acc. Chem. Res. 26 (1993) 319-324.

[156] E.A. Burstein, S.M. Abornev, Y.K. Reshetnyak, Decomposition of protein tryptophan fluorescence spectra into log-normal components. I. Decomposition algorithms, Biophys. J. 81 (2001) 1699-1709.

[157] P. Sonawane, K. Patel, R.K. Vishwakarma, S. Singh, B.M. Khan, *In silico* mutagenesis and docking studies of active site residues suggest altered substrate specificity and possible physiological role of Cinnamoyl CoA Reductase (Ll-CCRH1), Bioinformation 9 (2013) 224-232.

[158] V.N. Uvesrsky, J. Li, A.L. Fink, Evidence for a partially folded intermediate in α-synuclein fibril formation, J Biol. Chem. 276 (2001) 10737-10744.

[159] P.R. Louzada, A. Sebollela, M.E. Scaramello, S.T. Ferreira, Predissociated dimers and molten globule monomers in the equilibrium unfolding of yeast glutathione reductase, Biophys. J. 85 (2003) 3255-3261.

[160] B. Reddy, V.R. Srinivas, N. Ahmad, A. Surolia, Molten globule-like state of peanut lectin monomer retains its carbohydrate specificity-implications in protein folding and legume lectin oligomerization, J. Biol. Chem. 274 (1999) 4500-4503.

[161] D. Sarkar, C. Dasgupta, Characterization of a molten globule intermediate during GdnHCl- induced unfolding of RTEM beta-lactamase from *Escherichia coli*, Biochim. Biophys. Acta 1296 (1996) 85-94.

[162] A. Kumar, S.M. Gaikwad, Multistate unfolding of α-mannosidase from *Canavalia ensiformis* (Jack Bean): Evidence for the thermostable molten globule, Biochem. Biophys. Res. Commun. 403 (2010) 391-397.

[163] A. Kumar, N. Gowda, S. Gaikwad, A. Pundle, *Rhodotorula aurantiaca* penicillin V acylase: Active site characterization and fluorometric studies, J Photochem. Photobiol. B Biol. 97 (2009) 109-116.

[164] J.R. Lakawiez (Ed.), Principles of fluorescence spectroscopy, Second ed., Kluwer, Academic/Plenum, New York, 1999, 4445-486.

[165] M.R. Eflink, C.A. Ghiron, Fluorescence quenching studies with proteins, Anal. Biochem. 114 (1981) 199-227.

[166] A. Grinvald, I.Z. Steinberg, The fluorescence decay of tryptophan residue in native and denatured proteins, Biochem. Biophys. Acta 427 (1976) 663-678.

[167] S.S. Lehrer, Solute perturbation of protein fluorescence. The quenching of tryptophyl fluorescence of model compounds and of lysozyme by iodide ion, Biochemistry 10 (1971) 3254-3263.

[168] S.S. Lehrer, P.C. Leavis, Solute quenching of protein fluorescence, Methods Enzymol. 49 (1978) 222-236.

[169] E.M. Lakowiez, G. Weber, Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins on the nanosecond time scale, Biochemistry 12 (1973) 4171-4179.

[170] D.M. Chipman, V. Grisrao, N. Sharon, The binding of oligosaccharides containing N-acetylglucosamine and N-acetylmuramic acid to lysozyme, J Biol. Chem. 242 (1967) 4388-4394.

[171] R. Kenoth, M.J. Swamy, Steady state and time resolved fluorescence studies on *Trichosanthes cucumerina* seed lectin, J. Photochem. Photobiol. B Biol. 69 (2003) 193-201.

[172] S.M. Gaikwad, M.I. Khan, Binding of T-antigen disaccharides to *Artocarpus hirsuta* lectin and jacalin are energetically different, Photochem. Photobiol. 82 (2006) 1315-1318.

[173] F. Khan, A. Ahmad, M.I. Khan, Purification and characterization of a lectin from endophytic fungus *Fusarium solani* having complex sugar specificity, Arch. Biochem. Biophys. 457 (2007) 243-251.

[174] U. Katre, C.G. Suresh, M.I. Khan, S.M. Gaikwad, Steady state and Time-Resolved fluorescence studies of a Hemagglutinin from Moringa oleifera, J. Fluoresc. 18 (2008) 479-485.

[175] R.K. Sreejith, V. Yadav, N. Varshney, S. Berwal, C.G. Suresh, S.M. Gaikwad, J.K. Pal, Conformational characterization of human eukaryotic initiation factor 2α: A single tryptophan protein, Biochem. Biophys. Res. Comm. 390 (2009) 273-279.

[176] K.S. Shashidhara, S.M. Gaikwad, Conforamtional and Functional Transitions in class II α-mannosidase from *Aspergillus fischeri*, 20 (2010) 827-836.

[177] K.S. Shashidhara, S.M. Gaikwad, Fluorescence quenching and time resolved fluorescence studies of α-mannosidase from *Aspergillus fischeri* (NCIM 508), J. Fluoresc. 7 (2007) 599-605.

[178] R.K. Sreejith, C.G. Suresh, S.H. Bhosale, V. Bhavnani, A. Kumar, S.M. Gaikwad, J.K Pal, Conformational Transitions of the catalytic domain of Heme-regulated eukaryotic initiation factor 2α kinase, a key translational regulatory molecule, J. Fluoresc. 22 (2012) 431-441.

[179] P.N. Dharker, S.M. Gaikwad, C.G. Suresh, V. Dhuna, M.I. Khan, J. Singh, S.S. Kamboj, Comparative studies of two araceous lectins by steady state and time resolved fluorescence and CD spectroscopy, J. Fluoresc. 19 (2009) 239-248.

[180] U. Sharma, S.M. Gaikwad, C.G. Suresh, V. Dhuna, J. Singh, S.S. Kamboj, Conformational transitions in *Ariesaema curvatum* lectin: Characterization of an acid induced active molten globule, J. Fluoresc. 21 (2011) 753-763.