



P-3090



**CRYSTAL STRUCTURE DETERMINATION OF (A)
MATURED cyp38 FROM *Arabidopsis thaliana* AND (B)
SUPEROXIDE DISMUTASE (SOD) FROM *Spirulina
platensis***

PROJECT REPORT

Submitted by

S.MANIVANNAN



Register No: 0820203006

P-3090

in partial fulfillment for the award of the degree

Of

MASTER OF TECHNOLOGY

in

BIOTECHNOLOGY

KUMARAGURU COLLEGE OF TECHNOLOGY, COIMBATORE-06.

(An Autonomous Institution affiliated to Anna University, Coimbatore)

MAY 2010

ANNA UNIVERSITY:COIMBATORE
BONAFIDE CERTIFICATE
KUMARAGURU COLLEGE OF TECHNOLOGY

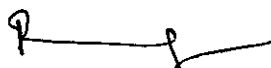
COIMBATORE-641 006

Department of Biotechnology

PROJECT WORK -PHASE II

MAY 2010

This is to certify that the project entitled **CRYSTAL STRUCTURE DETERMINATION OF (A) MATURED cyp38 FROM *Arabidopsis thaliana* AND (B) SUPEROXIDE DISMUTASE (SOD) FROM *Spirulina platensis*** is the bonafide record of project work done by **S.MANIVANNAN Register No: 0820203006** of M.Tech during the year 2009-2010.



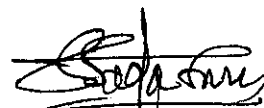
Dr. P.RAMALINGAM

Assistant Professor

Department of Biotechnology

Kumaraguru College of Technology

Coimbatore - 641 006



Dr. S.SADASIVAM

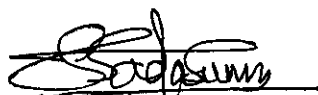
Dean - Biotechnology

Department of Biotechnology

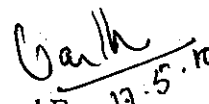
Kumaraguru College of Technology

Coimbatore - 641 006

Submitted for the Project Viva-Voce examination held on 17.05.2010




Internal Examiner



External Examiner

DECLARATION

I affirm that the project work titled **Crystal Structure Determination Of (A) Matured Cyp38 From *Arabidopsis thaliana* And (B) Superoxide Dismutase (SOD) From *Spirulina platensis*** being submitted in partial fulfilment for the award of **M.Tech(Biotechnology)** is the original work carried out by me. It has not formed the part of any other project work submitted for award of any degree or diploma, either in this or any other University.


S.MANIVANNAN
0820203006

I certify that the declaration made by the above candidate is true.

Signature of the Guide,


Dr.P.Ramalingam

Assistant Professor

Department of Biotechnology

Kumaraguru College of Technology

Coimbatore

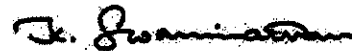
April 24, 2010

To whom it may concern

Mr. Manivannan Sethurajan of Department of Biotechnology, Kumaraguru College of Technology, Coimbatore worked in my laboratory as a research trainee during December 1, 2009 to March 31, 2010 for his Master's project. His work was on **Crystal structure determination of (a) matured Cyp38 from *Arabidopsis thaliana* and (b) Superoxide dismutase (SOD) from *Spirulina platensis*.**

I am glad to attest that Mr. Sethurajan was very focused, committed and sincere. He realized his level of responsibility in his undertaking and proved to be an excellent researcher. Also, he has very pleasing personality and moves well with others.

Feel free to contact me (dbksks@nus.edu.sg) should you need any further information and reference about him.



(Electronic signature)

Kunchithapadam Swaminathan
Associate Professor

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

First of all, I express my deep sense of gratitude to **Correspondent, Joint correspondent, Principal and Dr. S. Sadasivam, Dean**, Department of Biotechnology, Kumaraguru College of Technology, Coimbatore for allowing me to undertake my Master thesis at NUS, Singapore.

My heartfelt gratitude to my internal thesis supervisor, **Dr.P.Ramalingam**, Assistant Professor, Department of Biotechnology, Kumaraguru college of Technology.

I would like to express my sincere thanks and gratitude to my external thesis supervisor, **Dr. K. Swaminathan**, for giving me an opportunity to work on these high end research projects. His experience, inspiration, enthusiasm, and patience helped me get trained in the most prominent fields of molecular and structural biology and also I am very much thankful for his unfailing guidance and support throughout the work and thereafter.

I record my thanks to all faculty members and also lab assistants of Department of Biotechnology, Kumaraguru college of Technology.

I also remember the kind help of my friends and classmates in KCT, Coimbatore and also the co workers and lab mates in Department of biological sciences, NUS, Singapore.

Last but not the least I am thankful to all my family members, who have contributed significantly to bring this day in my life. My parents and my brother remain as a constant source of strength throughout my educational career and to achieve my goals.


S.Manivannan

CHAPTER NO	TITLE	PAGE
	Abstract	ix
	List of Tables	x
	List of figures	xi
	List of abbreviations	xiii
1	Macromolecular X-Ray Crystallography	1
	1.1 Protein Structure Determination	1
	1.2 Protein Crystallization	1
	1.3 Basic Concepts In Protein Crystallography	5
	1.3.1 Lattices, point groups and space groups	6
	1.3.2 hkl plane	7
	1.3.3 Principle of X-ray diffraction and Bragg's law	8
	1.3.4 Reciprocal space	9
	1.3.5 The Ewald sphere	10
	1.3.6 Fourier transformation and structure factor equation	11
	1.3.7 Phase problem	11
	1.4 Structure Determination	12
	1.4.1 Solution to phase problem	12
	1.4.2 Direct methods	12
	1.4.3 Molecular replacement (MR)	13
	1.4.4 Multiple isomorphous replacement (MIR)	13
	1.4.5 Anomalous scattering	13
	1.4.5.1 MAD	14
	1.4.5.2 SAD	14
	1.4.6 Phase improvement	14
	1.4.7 Model building and refinement	15
	1.4.8 Validation and structure deposition	16
2	Biological Background	18

2.1	Organ Transplantation & Immunosuppressive Drugs	18
2.1.1	Cyclosporin	18
2.1.2	FK506	19
2.1.3	Rapamycin	20
2.2	Immunophilins	21
2.2.1	PPIase activity of immunophilins and protein folding	22
2.2.2	Immunosuppressive activity of immunophilins	24
2.3	Diversity Of Immunophilins	24
2.3.1	Archaeal cyclophilins	25
2.3.2	Bacterial cyclophilins	25
2.3.3	Fungal cyclophilins	25
2.3.4	Animal cyclophilins	26
2.3.5	Plant cyclophilins	28
	2.3.5.1 Domain organization & phylogenetic analysis in Arabidopsis cyclophilins	30
	2.3.5.2 Evolutionary dynamics of the luminal cyclophilins	36
2.4	Cylophilin Structures	37
2.4.1	Ligand-free cyclophilin structures	37
2.4.2	Cyclophilin-ligand complex structures	42
	2.4.2.1 Structural features of cyclosporin A binding	44
	2.4.2.2 Structural features of prolyl-peptide	

	binding	45
2.5	Homologs Of Atcyp38	46
2.6	Superoxide Dismuatse	48
	2.6.1 Human SOD	49
	2.6.2 Biochemistry	51
	2.6.3 Physiology	51
	2.6.4 Role in disease	51
2.7	Aim & Objectives	53
3	Materials & Methods	54
	3.1 Cloning Of Sod Gene	54
	3.1.1 Primer design for PCR	54
	3.1.2 Gradient PCR	54
	3.1.3 PCR products purification	54
	3.1.4 Cloning of SOD in pGEM-T-Easy cloning vector	54
	3.1.5 Preparation of E. coli DH5 α competent cells	55
	3.1.6 Transformation into DH5 α Competent cells and blue-white screening	55
	3.1.7 Colony PCR	56
	3.1.8 Agarose gel electrophoresis	56
	3.1.9 Sequencing PCR	57
	3.1.10 Ethanol precipitation	57
	3.2 Expression & Purification of Matured Cyp38	57
	3.2.1 Affinity chromatographic purification of matured Cyp38	57
	3.2.2 Size exclusion chromatography	58
	3.3 Crystallization	58
	3.3.1 Hanging drop vapor diffusion method	58
4	Results & Discussion	59

	4.1	Cloning Of Sod Gene	59
		4.1.1 Gradient PCR	59
		4.1.3 Transformation of Ligated SOD-pGEM-T-Easy vector into DH5 α competent cells and Colony PCR	59
		4.1.3 DNA Sequencing	60
	4.2	Expression & Purification Of Matured cyp38	61
		4.2.1 Affinity chromatographic purification	61
		4.2.1 Size exclusion chromatography	61
	4.3	Crystallization	64
	4.4	Conclusions	65
	4.5	Future Work	65
5		Appendix	66
	5.1	SOD sequencing results	66
	5.2	pGEM:SOD plasmid DNA sequencing	67
6		References	68

Abstract

Abstract

Cyclophilin 38 (CyP38) is one of the highly divergent multi-domain cyclophilins from *Arabidopsis thaliana*. A recombinant form of matured AtCyP38 (residues 100-437) was expressed as a fusion protein with GST tag in *Escherichia coli*. Purification of this matured GST - CyP38 was first carried out by Affinity chromatography with Sepharose GST beads. Then it was purified to homogeneity by size exclusion chromatography. Polydispersity index of this protein was found to be 0.10 and SOS error was 4.48 by Dynamic light scattering. Crystallization trials were performed with commercially available kits.

Superoxide dismutase (SOD, EC 1.15.1.1) is one of the antioxidants from *Spirulina platensis*. Genes coding 170 aminoacids of Fe SOD were isolated from *Spirulina* and were cloned in pGEM-T-Easy vector. The template was found to be mutated by DNA sequencing.

LIST OF TABLES

Table 1.1.	The crystal systems and their related unit-cells and lattices.	7
-------------------	---	----------

LIST OF FIGURES

		Page
Figure 1.	Pictorial representation of Structure determination process	4
Figure 2.	Schematic representation of key stages in the structure determination process.	5
Figure 3.	The unit-cell	6
Figure 4.	(a) Constructive interference (b) destructive interference	8
Figure 5.	Bragg's law.	8
Figure 6.	Reciprocal lattice.	9
Figure 7.	The Ewald sphere.	10
Figure 8.	The chemical structure of Cyclosporin A	19
Figure 9.	The chemical structure of FK506	20
Figure 10.	The chemical structure of Rapamycin	21
Figure 11.	Peptidyl-prolyl isomerization reaction in the tri-peptide Ser-Pro-Ala	23
Figure 12.	Domain organization in <i>Arabidopsis</i> cyclophilins	32
Figure 13.	The phylogenetic relationships of <i>Arabidopsis</i> cyclophilins	33
Figure 14.	<i>Arabidopsis</i> cyclophilin isoforms and their sub-cellular Localization	35
Figure 15.	Structure overlap of various cyclophilin domains: bovine CyP40 (cyan), human CyPA (orange), human CyPB (pale yellow) and <i>E. coli</i> CyPB (violet). The overlap was generated with the MULTIPROT server (Shatsky <i>et al.</i> , 2004) and the figure was prepared using PYMOL.	42
Figure 16.	The structure of human CyPA. α -helices are shown in red, loops in green and β -strands as yellow arrows. The residues of the active site that are involved in cyclosporine A (CsA) binding and PPIase activity are highlighted in blue as stick model. This figure was prepared with the PYMOL program.	45

Figure 17.	Sequence alignment (by CLUSTALW) of AtCyp38 with TLP40 from spinach, pea and rice. The predicted helical domain is given in blue and the cyclophilin domain in red. The identical sequences are highlighted in yellow.	47
Figure 18.	Biological function of superoxide dismutase (SOD).	48
Figure 19.	Crystallographic structure of the human SOD1 enzyme	50
Figure 20.	The active site of the active site of human superoxide dismutase 2.	50
Figure 21.	Agarose gel electrophoresis showing Gradient PCR results. Lane 1 – 1 kb DNA ladder. Lanes 2, 3, 4, 5, 6 correspond to temperatures 59, 58.1, 57, 56.1 and 55 respectively.	59
Figure 22.	Agarose Gel electrophoresis showing colony PCR results. Lane 1 – 1 kb DNA ladder, Lanes 2 – 8 correspond to colonies 1 – 7.	60
Figure 23.	SDS-PAGE showing affinity chromatographic purification of the matured Cyp38-GST fusion protein. Lane 1 – Low molecular weight marker, Lane 2 – Pellet, Lane 3 – Flow through, Lane 4 – Wash 1, Lane 5 – Wash 2, Lanes 6, 7, 8, 9 correspond to Elutions 1, 2, 3, 4 respectively, Lane 10 – Wash 3, Lanes 11,12 correspond to Elutions 5 and 6	61
Figure 24.	Profile of size-exclusion chromatographic purification of matured Cyp38-GST fusion protein. The higher peak around 70ml corresponds to the protein	62
Figure 25.	SDS-PAGE of size-exclusion chromatographic fractions of Cyp38-GST fusion protein. Lane 1 – Low molecular weight marker, Lanes 2-15 correspond to fractions E7 to F11.	63
Figure 26.	DLS profile of size-exclusion chromatographic fractions of matured Cyp38-GST fusion protein.	64

LIST OF ABBREVIATIONS

ADSC	Area Detector Systems Corporation
BnP	Buffalo and Pittsburg Software
CaM	calcium-loaded calmodulin
CCD	charge coupled device
Cn	calcineurin
CNS	Crystallography and Nuclear magnetic resonance System
CPR	cyclosporin-sensitive proline rotamase
CsA	cyclosporin A
CyP	cyclophilin
DLS	dynamic light scattering
DMSO	dimethylsulfoxide
DTT	1,4-dithio-DL-threitol
EDTA	ethylene diamine tetra-acetic acid
ER	endoplasmic reticulum
FKBP	FK506 binding protein
GRASP	graphical representation and analysis of structural properties
GST	glutathione sepharose transferase
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethane sulfonic acid
HIV	human immuno-deficiency virus
hnRNP	heterogeneous ribonucleoprotein particle
HSP	heat shock protein
IgE	Immunoglobulin E

IPTG	isopropyl-beta-D-thiogalactopyranoside
xiMAD	multi-wavelength anomalous dispersion
MALDI-TOF	matrix assisted laser desorption/ionization – time of flight
MCSG	midwest center for structural genomics
MIR	multiple isomorphous replacement
MPTP	mitochondrial permeability transition pores
MS	mass spectrometry
MWPC	multi-wire proportional chamber
NMR	nuclear magnetic resonance
PAGE	poly acrylamide gel electrophoresis
PDB	Protein Data Bank
PEG	poly ethylene glycol
PMSF	phenyl methyl sulfonic fluoride
pNA	para-nitroaniline
Ppi	Peptidyl-prolyl isomerase
PPIase	peptidyl-prolyl cis-trans isomerase
PS II	photosystem II
RRM	RNA recognition motif
SDS	sodium dodecyl sulfate
snRNP	small nuclear ribonucleoprotein
TLP	thylakoid lumen protein
TM1367	<i>Thermotoga maritima</i> protein 1367
TPR	tetratricopeptide repeat

*DEDICATED TO MY BELOVED
PARENTS*

*MACROMOLECULAR X-RAY
CRYSTALLOGRAPHY*

CHAPTER 1 MACROMOLECULAR X-RAY CRYSTALLOGRAPHY

1.1 PROTEIN STRUCTURE DETERMINATION

Proteins perform a variety of functions in living systems. Enzymes, which are proteins, catalyze and control the rate of chemical reactions in an organism. Other proteins function as structural, transport, signaling and defense entities of living systems. Even if we know the amino acid sequence of a protein, we still cannot appreciate its function and properties if we do not know what exactly the protein looks like or how the active site functions. Thus, knowing the 3-dimensional structure of a protein is of great significance.

Several methods are available for studying the 3-dimensional structure of proteins such as Nuclear Magnetic Resonance (NMR) spectroscopy, X-ray crystallography and cryo-electron microscopy. At present, X-ray crystallography is the most effective of these techniques. Certainly, the other two techniques complement crystallography and are regarded for their respective value in macromolecular structure determination. X-ray crystallographic techniques were first developed in 1912 and initially applied to small molecules. Based on its extraordinary success, the recording of the first diffraction pattern from a protein crystal, that of pepsin, was successfully attempted in the 1930s (Bernal & Crowfoot, 1934). However, the complete application of X-ray diffraction to protein structure determination happened much later. The first protein structures, those of myoglobin (Kendrew *et al.* 1960) and hemoglobin (Perutz *et al.* 1960), were solved only in the late 1950's. After that macromolecular crystallography has progressed tremendously and in the last 20 years nearly 35,000 protein structures have been determined. However, this collection represents only a very small fraction of the hundreds of thousands of various protein molecules that actively define the process of 'life'.

1.2 PROTEIN CRYSTALLIZATION

There are several bottlenecks in the determination of a crystal structure, of which obtaining a useful crystal is the most serious one. If one cannot collect diffraction data of suitable quality, protein structure determination will not be possible. Proteins have

irregularly shaped surfaces, which result in the formation of large channels within a crystal. Therefore, the non-covalent bonds that hold together the lattice must often be hydrogen bonds, through intervening water molecules (Rhodes, 2000). Furthermore, the successful production of crystals for diffraction studies depends on a number of environmental factors like pH, temperature, ionic strength etc. with each protein requiring a unique condition for successful crystallization. Therefore, protein crystallization is almost always a very tedious process.

Crystallization of proteins is achieved by driving a protein solution to supersaturation in several ways such as: cooling, evaporation of water, addition of an ionic solute, e.g. salt, and variation of pH. Once the protein solution becomes super-saturated, protein crystals tend to form themselves upon appearance of a microscopic nucleus, and crystallization can recur very rapidly.

Macromolecules always require some precipitants to invoke nucleation and crystal growth. The most popular precipitants are inorganic salts (ammonium sulfate, sodium chloride etc.) and organic polyols, polyethylene glycol of different molecular weights, methylpentanediol and similar additives. A multitude of other parameters is varied, e.g. the buffer type and pH, temperature, purity and concentration of protein and other components. In order to obtain sufficient homogeneity, the protein should usually be at least 97% pure. Also, pH conditions, regulated by different buffers, are very important, as different pHs can result in different packing orientations within a crystal. Some proteins require small amounts of special additives such as phenol, isopropanol or various co-factors or effectors to produce good-quality crystals.

Many screening methods, such as full factorial, incomplete factorial, random and sparse matrix screens, are in use for the efficient identification of crystallization conditions of protein molecules. Of these, the most common one is the sparse matrix method, which utilizes conditions based on known successful crystallization conditions of macromolecules. Various sets of crystallization screening conditions, duly selected by sparse matrix sampling, have been proposed and many ready-made solutions are commercially available. These screens make use of several conditions that have a wide range of pH, salts and precipitants. Additional information regarding a protein's affinity for metal ions, ligands, etc. also helps in the crystallization process. Results from

preliminary trials of crystallization screening can provide considerable information regarding the solubility of a protein, the choice of a precipitant, pH and salts and at times may even give crystals. Crystals obtained initially may not always diffract well. They need further optimization to give better diffraction quality.

There are several crystallization techniques, including sitting drop vapor diffusion, hanging drop vapor diffusion, sandwich drop, batch, micro-batch, under oil, vapor batch, micro-dialysis and free interface diffusion. The hanging drop vapor diffusion technique is the most popular method for the crystallization of macromolecules. The principle of vapor diffusion is quite straightforward. A drop of a mixture of a protein sample and a reagent is placed in vapor equilibration with a liquid reservoir of the reagent. Typically, the drop contains a lower reagent concentration than that in the reservoir. To achieve equilibrium, water vapor leaves the drop and eventually ends up in the reservoir. As water leaves the drop, the protein undergoes an increase in relative super-saturation. Equilibration is reached when the reagent concentration in the drop is approximately the same as that in the reservoir, favoring the protein to crystallize.

For decades, crystals were mounted in sealed glass (or quartz) capillaries, in presence with a drop of mother liquor. A breakthrough in protein crystal handling is the cryo-cooling technique. Using this method, now commonly referred to as cryo-crystallography, the crystal and the mother liquor surrounding and inside the crystal in solvent channels can be solidified to a state of amorphous glass with the addition of a suitable cryoprotectant. The crystal is then scooped from the liquid with a small fiber loop and quickly plunged into liquid nitrogen or placed in a very cold nitrogen gas stream at temperatures around -173 °C (Teng, 1990). The crystal is preserved in its crystalline order to retain its diffraction properties (Hope, 1988). This method reduces mechanical stress and diminishes the amount of radiation damage induced by exposure to X-rays (Garman & Schneider, 1997). Various modern modifications in crystal handling protocols have made the whole process of getting good quality data-sets from a given protein crystal a lot easier than how it used to be two decades back.

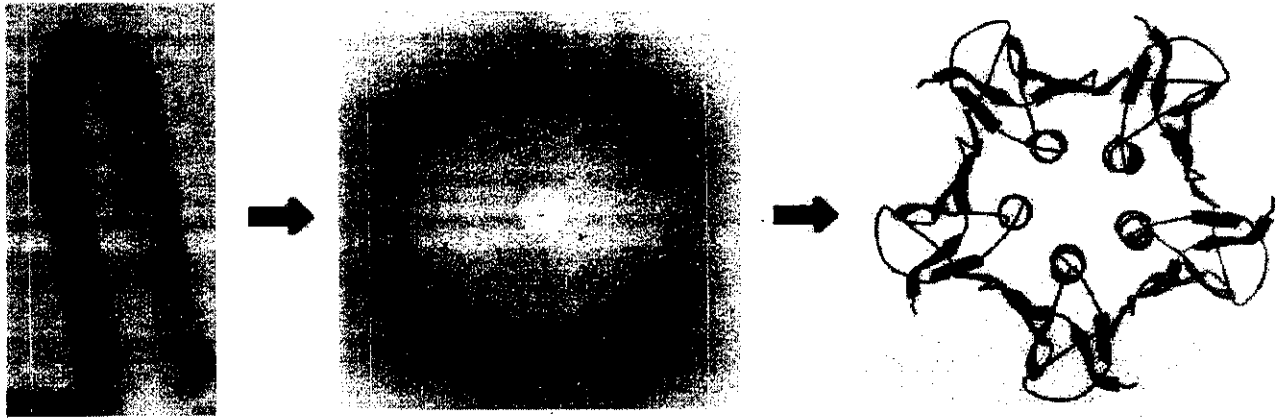


Figure 1. Pictorial representation of Structure determination process

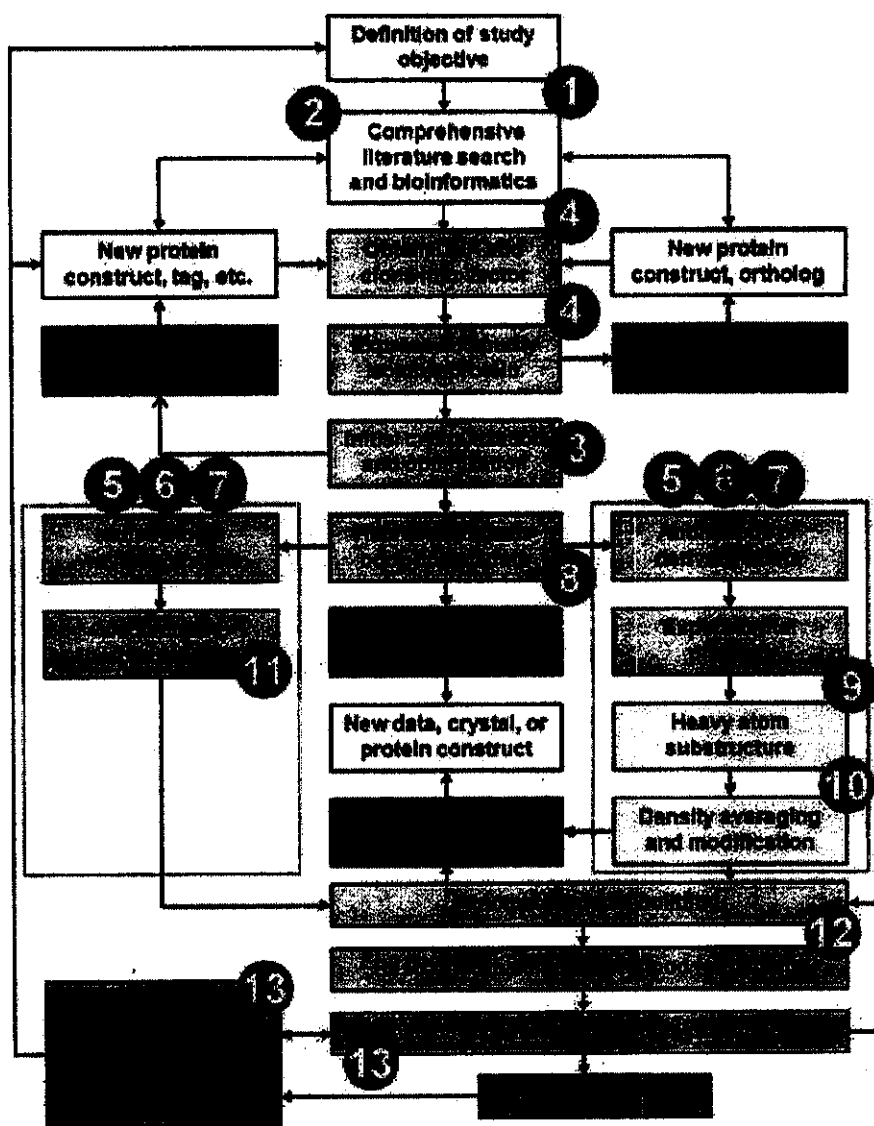


Figure 2. Schematic representation of key stages in the structure determination process.

1.3 BASIC CONCEPTS IN PROTEIN CRYSTALLOGRAPHY

Crystals are defined as solids with a periodic three dimensional arrangement of an atomic structure. The simplest repeating unit in a crystal is called a unit-cell with three basis vectors a , b and c , and inter-axial angles α , β and γ between them (Fig. 3). The smallest and unique volume within the unit-cell that can be rotated and translated to generate one unit-cell is called the asymmetric unit. Unit-cells are classified into seven different systems based on their dimensions.

The arrangement of molecules in a unit-cell is governed by symmetry (Section 1.3.1) and this symmetrical arrangement defines the system of that crystal. Table 1.1 shows the 7 different crystal systems:

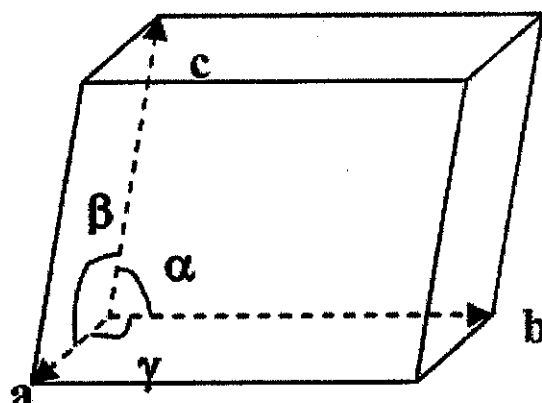


Figure 3. The unit-cell

1.3.1 Lattices, point groups and space groups

In a crystal, unit-cells are arranged in a contiguous way to fill space. If a point is assumed to represent a whole unit-cell, then the array of all points will form a lattice. Table 1.1 summarises the collection of 14 such lattices, known as Bravais lattices. The Bravais lattices are classified as primitive, P (simple unit-cell with one point for each unit-cell), face centered, F (an additional lattice point at the center of each face), body centered, I (an additional point at the center of the cell), end centered, C (an additional point at the center of one face).

Molecules are arranged in the unit-cell with certain symmetry operations when packed into a crystal. A symmetry operation gives an identical or similar image of an object. Besides unit translations along the three unit-cell axes, called three-dimensional translation symmetry, the three crystallographic symmetry elements are rotation, reflection, and inversion. The combination of these symmetry elements that acts on a unit-cell is commonly called a point group. There are totally 32 crystallographic point groups (11 with proper rotations and 21 with improper rotations).

Table 1.1. The crystal systems and their related unit-cells and lattices.

<u>System</u>	<u># of independent cell parameters</u>	<u>Cell restrictions</u>		<u>Types of Centering</u>
1. Triclinic	6	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$	P
2. Monoclinic	4	$a \neq b \neq c$	$\alpha \neq \gamma \neq 90; \beta = 90; \gamma = 90$	P,C
3. Orthorhombic	3	$a \neq b \neq c$	$\alpha = \gamma = \beta = 90; \gamma = 90$	P,C,I,F
4. Tetragonal	2	$a = b \neq c$	$\alpha = \gamma = \beta = 90; \gamma = 90$	P,I
5. Hexagonal (6)	2	$a = b \neq c$	$\alpha = \beta = 90; \gamma = 120; \gamma = 120$	P
6. Trigonal (3)*	2	$a = b \neq c$	$\alpha = \beta = 90; \gamma = 120; \gamma = 120$	P
7. Cubic	1	$a = b = c$	$\alpha = \gamma = \beta = 90; \gamma = 90$	P,I,F

7 = Crystal systems

Bravais Lattices = 14

Rotation or reflection, when combined with translation, will generate screw or glide symmetry, respectively. The combination of lattices and point groups leads to 230 different possible ways of molecular packing in a crystal, known as space groups, out of which, only 65 space groups are applicable to protein crystals (McRee, 1999).

1.3.2 hkl plane

The diffraction effect of a crystal (Section 1.3.3), is well explained using the concept of hkl planes. When X-rays are diffracted by a crystal, each resulting spot is created by an imaginary set of parallel 'hkl' planes that slice the entire crystal in a particular direction. The index h is the number of integral parts into which the set of planes cuts the X direction (or a-axis) of each unit-cell. Similarly, the indices k and l specify how many such planes exist per unit-cell in the Y and Z directions. The family of planes having indices hkl (h, k, l must be integers) is the hkl family of planes.

1.3.3 Principle of X-ray diffraction and Bragg's law

X-rays are electromagnetic waves which interact with matter like electrons. X-rays, scattered from different electrons, will travel different distances, differing in their relative phases. Thus, there will be interference when they are in phase (constructive interference) or when out of phase (destructive interference), Fig. 4.

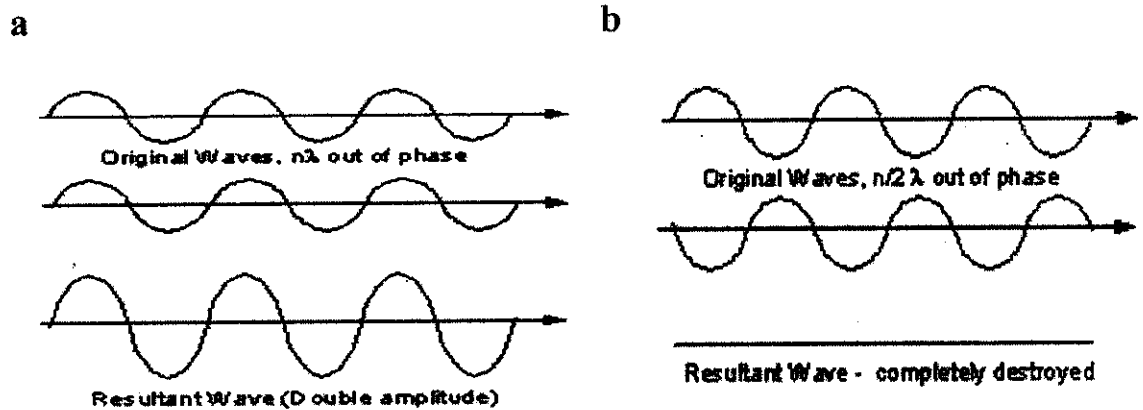


Figure 4. (a) Constructive interference (b) destructive interference

The geometric requirements needed for diffraction to occur were first explained by Bragg. Bragg showed that a set of parallel planes, n , with indices, hkl and interplanar spacing, d_{hkl} produces a diffracted beam when X-rays of wavelength λ impinge on the planes at an angle θ and are reflected at the same angle, only if θ meets the condition

$$2 d_{hkl} \sin\theta = n\lambda \quad (\text{Eq. 1.1})$$

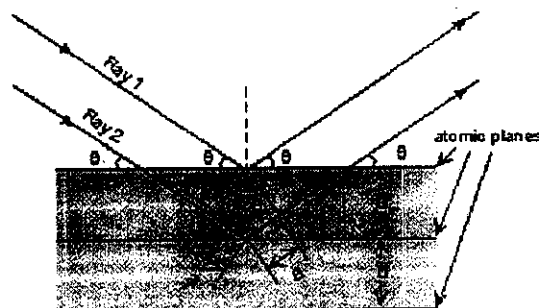


Figure 5. Bragg's law.

This is known as Bragg's law for X-ray diffraction, Fig. 5.

1.3.4 Reciprocal space

A reciprocal lattice is defined as a discrete set of diffracted rays whose vectors are perpendicular to the real lattice planes from which they are derived.

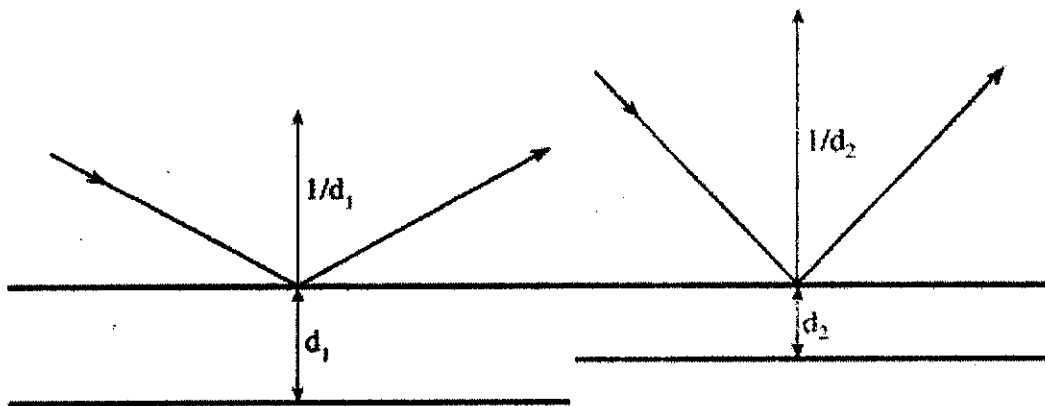


Figure 6. Reciprocal lattice.

In Fig. 6, the Bragg planes and the incoming and reflected rays are shown, as before, for two diffraction angles, but now a vector perpendicular to the Bragg plane is added for each. The phase of a wave diffracting from an object lying in between two Bragg's planes depends on the fraction of the distance of the object from one Bragg plane to the next. If the position of the object is considered as a vector, then the distance of that object from one of the Bragg planes can be obtained by projecting that vector on the plane normal. The phase shift can then be obtained by dividing the projected distance by the Bragg spacing between the planes. Mathematically, we can carry out the projection and division by giving the plane normal a length equal to the reciprocal of the Bragg spacing, and then computing the dot product between the position vector and the plane normal. Because the plane normal is a vector with a length reciprocal to the spacing in the object, we call it a vector in reciprocal space.

1.3.5 The Ewald sphere

Ewald's geometrical construction, Fig. 1.16, helps us visualise which Bragg planes are in the correct orientation to diffract.

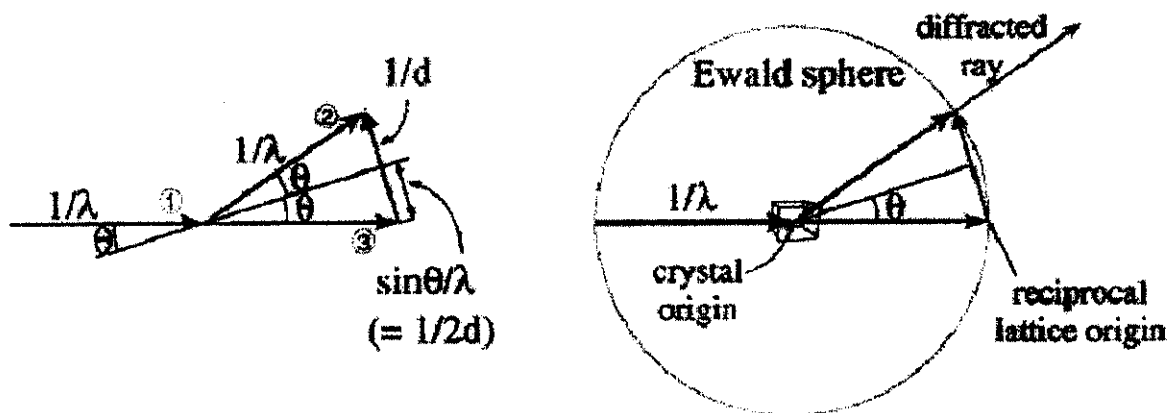


Figure 7. The Ewald sphere.

Rearranging Bragg's law in reciprocal space,

$$\sin \theta = \frac{\lambda}{2d_{hkl}} = \frac{\lambda}{2} \times \frac{1}{d_{hkl}} = \frac{1}{\frac{2}{\lambda}}$$

(Eq. 1.2)

Fig. 7 demonstrates how each reciprocal lattice point must be arranged with respect to the X-ray beam in order to satisfy Bragg's law and produce a reflection from the crystal. Since the Ewald sphere is three dimensional, reflections by hkl planes when exposed to X-rays form a three dimensional network of diffraction spots. If the Ewald sphere has a diameter of $2/\lambda$, then any reciprocal-lattice point within a distance $2/\lambda$ from the origin can be rotated into contact with the sphere to form a diffraction spot.

1.3.6 Fourier transformation and structure factor equation

The atomic arrangement (or electron density) in a crystal is related to all the diffraction spots that are obtained from a crystal through the Fourier transformation principle. The electron density at any point can be calculated using Eq. 1.3.

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx + ky + lz)}$$

(Eq. 1.3)

In the above equation, we are transforming the reciprocal space of lattice planes to real or direct space electron density at the point x, y, z . Thus, if we know F_{hkl} , the structure factor (inverse space entity from diffraction by electrons), we can calculate the actual real structure (the density of electrons in real space).

The Structure factor F_{hkl} for a reflection h, k, l is a complex number derived quite straightforward as follows:

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

(Eq. 1.4)

This is a simple summation, which extends over all atoms j at fractional coordinates x_j, y_j and z_j and $f(j)$ is the scattering factor of each atom which depends on the atomic number of that atom and the diffraction angle of the corresponding reflection (h, k, l) . Ironically, Eq. 1.4 shows that if we know the positions of all atoms (the structure) then we can easily calculate all structure factors and with the help of Eq. 1.3, we can calculate the electron density at any point (again, the structure).

1.3.7 Phase problem

In order to compute Eq. 1.4, we need the amplitudes of all diffracted waves (involving the scattering factor component), which are obtained from the intensities of



P-3090

the diffraction spots and also their relative phase shifts with respect to the origin of the unit-cell (involving the positions of all atoms), which cannot be measured directly. The immeasurability of the phase angle is known as phase problem. The structure factor equation, Eq. 1.4, can be rearranged as shown in Eq. 1.5 to explicitly show the phase angle.

$$F(hkl) = \sum_{j=1}^N |f_j| [\cos 2\pi(hx + ky + lz)_j + i \sin 2\pi(hx + ky + lz)_j]$$

(Eq. 1.5)

The methods used for phase angle estimation are discussed in the next section. Solving the phase problem is the most crucial step in crystal structure determination

1.4 STRUCTURE DETERMINATION

1.4.1 Solution to phase problem

Four methods are used to solve the phase problem in macromolecular structure determination. They are: direct methods, heavy-atom method (or isomorphous replacement method), anomalous scattering method (also called anomalous dispersion) and molecular replacement method. All these methods only yield phase estimates for a limited set of reflections. To improve the accuracy of the phase and to get an interpretable electron density map, refinement at both reciprocal and real space is carried out with the help of Fourier transformation.

1.4.2 Direct methods

This method relies on the possibility of development of useful statistical relationships between sets of structure factors to deduce their phases. However, we must assume a crystal to be made up of similarly-shaped atoms with positive electron density everywhere. The direct methods estimate the initial phases for a selected set of reflections using a triple relation and extend phases to more reflections. A triple relation is one in which there are trio of reflections in which the intensity and phase of one reflection can be explained by the other two.

The prime requirement for the direct methods to be successful in protein crystallography is very high resolution data ($> 1.2 \text{ \AA}$). This has limited its usefulness as only structures of small molecules can be solved easily, although the direct methods have been used to phase proteins up to 1000 atoms.

1.4.3 Molecular replacement (MR)

The molecular replacement method is used when a protein molecule has high sequence and hence structural similarity to an already solved protein structure (referred to as the search model). Usually, the Patterson function (a function derived using structure factors) of the search model is first correctly orientated in the new crystal unit-cell by means of rotation functions and then translated to achieve the best fit that is supported by a convincing correlation factor and a residual factor (details of the residual factor in section 1.4.7).

1.4.4 Multiple isomorphous replacement (MIR)

In this technique, first an X-ray data set for a protein crystal (native crystal) is collected. Following this, the native crystal is soaked in a specific heavy atom, such as mercury, platinum or gold. The goal is to obtain derivative crystals in which heavy atoms bind specifically and consistently to each protein molecule in the unit-cell. Another data set is collected for the derivative and the positions of the heavy atoms are determined using difference Patterson maps. Once the initial heavy atom locations have been determined, the coordinates, occupancy and temperature factors of each heavy atom are refined. At least two isomorphous derivatives (which are essentially isomorphous to the native crystal) are needed for successful structure determination by MIR.

1.4.5 Anomalous scattering

The atomic scattering factor of an atom has three components as shown in Eq. 1.6.

$$f = f_0 + f' + if'' \quad (\text{Eq. 1.6})$$

f_0 is a scattering term that is dependent on the Bragg angle and the two terms f' and f'' are independent of the scattering angle, but dependent on wavelengths. These latter

two terms represent the anomalous scattering that occurs at the absorption edge when the X-ray photon energy is sufficient to promote an electron from an inner shell to the next shell. The dispersive term f' reduces f_0 whereas the absorption term f'' is 90° advanced in phase with respect to f' . This leads to a breakdown in Friedel's law (which in a normal case gives $F_{hkl} = F_{-h-k-l}$), giving rise to anomalous differences that can be used to locate anomalous scatterers in a crystal, if any.

1.4.5.1 MAD

Isomorphous replacement has several problems: non-isomorphism between crystals (unit-cell changes, reorientation of the protein, conformational changes, changes in salt and solvent ions), problems in locating all the heavy atoms, problems in refining heavy-atom positions, occupancies and thermal parameters and errors in intensity measurements. The use of the multiwavelength anomalous dispersion (MAD) method overcomes the non-isomorphism problems. The method used is similar to MIR in the calculation of difference Patterson maps to locate anomalous scatterers but now, we calculate difference anomalous Patterson maps. Data are collected at several, typically three, wavelengths in order to maximise the absorption and dispersive effects. The method demands excellent optics for accurate wavelength setting with minimum wavelength dispersion. Generally, all data are collected from a single frozen crystal with high redundancy in order to increase the statistical significance of the measurements and data are collected with as high a completeness as possible.

1.4.5.2 SAD

Single anomalous dispersion (SAD) is a sub-set of MAD. It is becoming increasingly practical to collect data at the absorption peak and use density-modification protocols to break the phase ambiguity and provide interpretable maps.

1.4.6 Phase improvement

The experimentally determined phases are usually not accurate enough to give a completely interpretable electron-density map. Thus, a variety of phase improvement techniques are available, like density modification, solvent flattening, histogram

matching and non-crystallographic averaging to modify electron density and improve phases. Solvent flattening is a powerful noise suppression technique that removes negative electron density and sets the value of electron density in the solvent regions to a value lower than that of protein electron density. Histogram matching alters the values of electron-density points to concur with an expected distribution of electron-density values. Non-crystallographic symmetry averaging imposes equivalence on electron density values when more than one copy of a molecule is present in the asymmetric unit. Density modification is a cyclic procedure, involving back-transformation of the modified electron-density map to give modified phases, recombination of these phases with the experimental phases and calculation of a new map which is then modified iteratively until convergence.

1.4.7 Model building and refinement

A model of the subject protein is produced by fitting the components of the structure into the experimentally derived electron-density map, followed by refinement. Generation of an atomic model of the molecule(s) is a crucial step in the structure determination process. The model-building task may be far from straightforward because of poor phase information and the resolution of the diffraction data may be limited. However, automation in the recent years has enormously reduced the amount of time involved in manual model building. When an initial atomic model is available, a vast amount of geometrical restraints of the structure can be applied in the structure refinement and rebuilding process, in order to generate better phases and in turn, a better / accurate atomic model. These steps are carried out in an iterative way to achieve gradual improvement of the model. Refinement of a model is the optimisation of a function of a set of observations by changing selected associated parameters so that the atomic model agrees well with the diffraction data. Refinement programs refine the geometry and temperature factors of the model atoms to improve the fitness of observed and calculated structure-factor amplitudes. Low-resolution data must be collected for proper evaluation of the structure because this is used in bulk solvent averaging. The highest possible resolution limit should be used because this maximises the accuracy and precision of the structure. This is determined by the signal-to-noise ratio $[I/\sigma(I)]$, completeness and

redundancy of data within the highest resolution shell. After several rounds of refinement and map fitting, the model is considered as an acceptable final model. The refinement of the structural model against the experimentally observed X-ray diffraction data is measured by a 'residual' or 'reliability' factor (R-factor), Eq. 1.7. The progress in iterative real and reciprocal space refinement is monitored by computing the difference between the measured structure factor $|F_{obs}|$ and the calculated structure factor $|F_{calc}|$ from the current model.

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

(Eq. 1.7)

When the model converges to a correct structure, the difference between measured F's and calculated F's will also converge. A desirable target R factor for a protein model should be less than 0.25.

1.4.8 Validation and structure deposition

The correctness and precision of atomic parameters in a structure will need to be assessed thoroughly, both during and after refinement. Validation of the correctness of the model of a macromolecule obtained from crystal structure analysis is necessary for several reasons. The number of atoms in the model and therefore the number of refined parameters may not be significantly higher than the number of measured unique reflections, especially at low resolution. Moreover, a large percentage of reflection intensities may be rather weak in relation to their estimated uncertainties. In addition, some parts of the model are usually more flexible or disordered and their conformation cannot be confidently defined in an electron-density map. When the model is refined without restraints, even at atomic resolution, these weakly defined regions of the model become very disordered, in contrast to well defined fragments (Dauter et al. 1992; Kleywegt, 2000). As long as both the R and R_{free} (calculated with about 10% of reflections that are not included for refinement) values continue to decrease with each

cycle of refinement, the process is repeated and an improved protein structure model can be built. Furthermore, as torsion angles are usually not restrained during refinement, their agreement with expected values in the Ramachandran plot (Ramachandran et al. 1963; Ramakrishnan & Ramachandran, 1965) and proper side chain rotamers are extremely useful in model validation.

Once an X-ray structure for a macromolecule is determined and it has successfully passed a proper validation process, the resulting model is deposited in the Protein Data Bank (PDB) (<http://www.rcsb.org/pdb>) and is available to the world.

BIOLOGICAL BACKGROUND

CHAPTER 2 BIOLOGICAL BACKGROUND

2.1 ORGAN TRANSPLANTATION AND IMMUNOSUPPRESSIVE DRUGS

When any organ from an individual is transplanted to another, the immune system of the recipient recognizes the organ as foreign in a short while and starts to reject and at times destroy it. The success of organ transplantation, which is the treatment of choice in a variety of end-stage diseases, is limited primarily by this type of allograft rejection. In its early days organ transplantation had severe setbacks due to such rejections. Direct manipulation of the host's immune response provides practical means of achieving success even when tissue matching of the donor and recipient is less than optimal.

The discovery of cyclosporin A, FK-506 and rapamycin, the family of highly effective and specific fungal and bacterial metabolites, ranks among one of the most important advancements in the field of organ transplantation. These compounds have been shown to act as potent immunosuppressants and their administration to transplant recipients has proved to significantly reduce organ rejection rates. Although structurally different, cyclosporin A, FK506 and rapamycin belong to a class of immunosuppressive natural products whose major pharmacological attributes stem from their ability to disrupt certain intracellular signal transduction process of T-lymphocyte activation.

2.1.1 Cyclosporin

Cyclosporin A (CsA) is a cyclic peptide made up of 11 amino-acids which was originally isolated from the cultures of the soil microorganism *Tolypocladium inflatum*. The drug's remarkable immunosuppressive activity, characterized by a lack of myelotoxicity, was identified by Borel and co-workers (1976). A series of experimental studies demonstrated the extraordinary effectiveness of cyclosporin A in preventing organ allograft rejection in animal models (Borel, 1990). It has been studied in transplants of skin, heart, kidney, lung, pancreas, bone marrow and small intestine. Currently, cyclosporin is probably the most powerful of the available drugs used to prevent tissue graft rejection, with the best risk-to-benefit ratio.

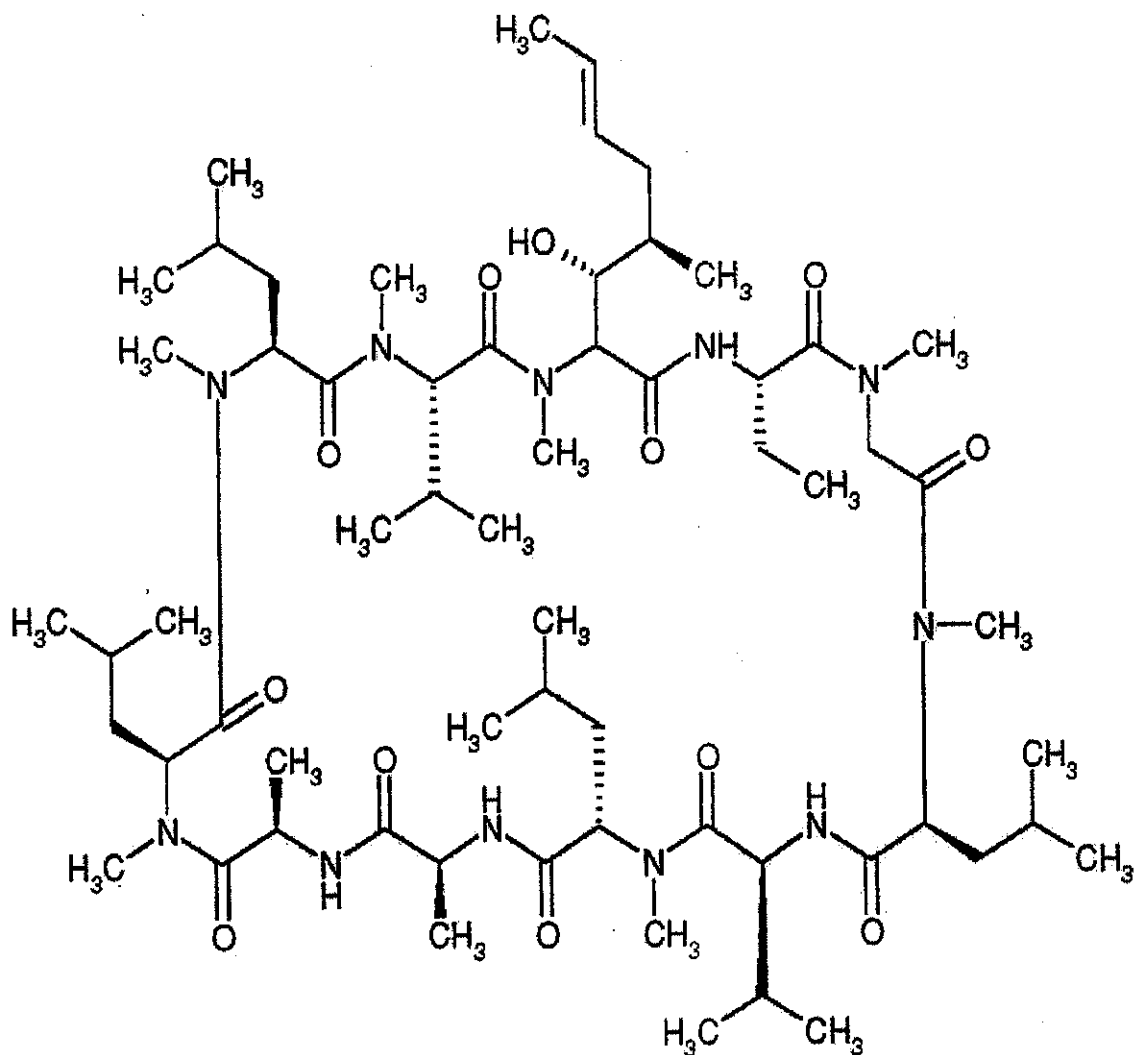


Figure 8. The chemical structure of Cyclosporin A

2.1.2 FK506

FK506 or Tacrolimus (Tsukuba macrolide immunosuppressant) is another FDA-approved immunosuppressant used for the prevention of allograft rejection, especially during liver and kidney transplantations. It is a macrolide lactone with a 23-member ring and is totally different from cyclosporin A in structure. FK506 was isolated in 1984 from the fermentation broth of *Streptomyces tsukubaensis* and was shown to be capable of suppressing a mixed-lymphocyte response *in vitro* (Kino *et al.* 1987).

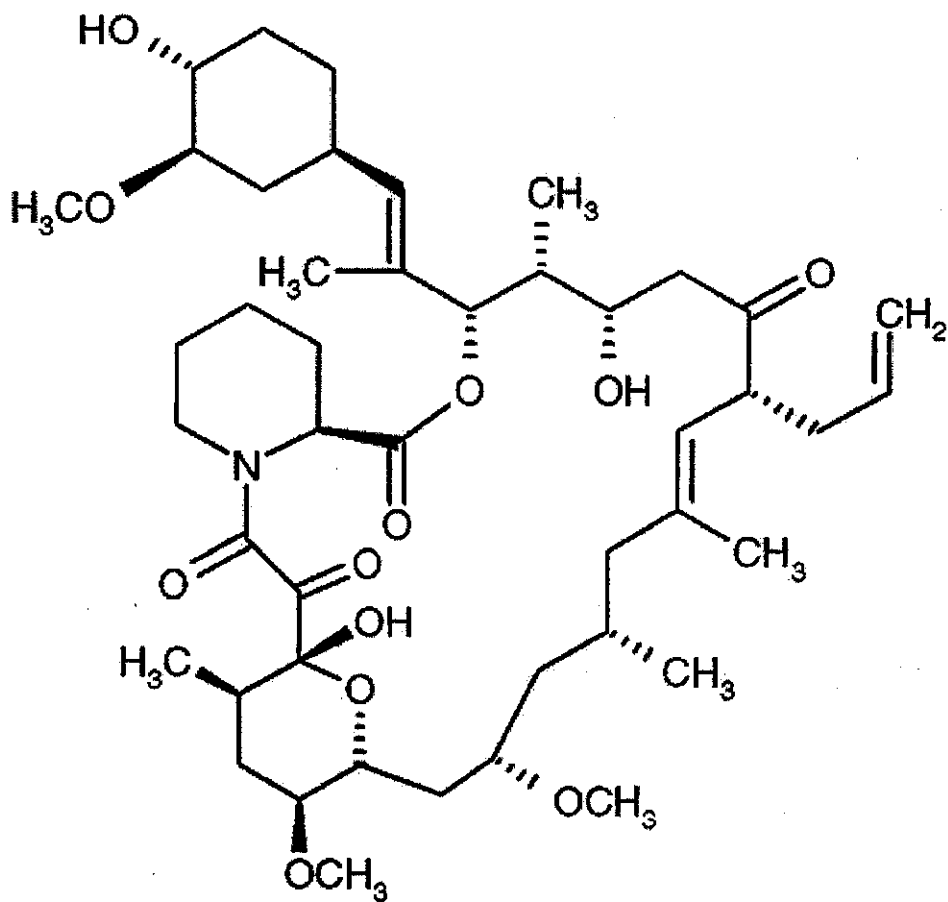


Figure 9. The chemical structure of FK506

2.1.3 Rapamycin

Rapamycin is a structural analogue of FK506. Also known as Sirolimus, it was discovered as a product of the bacterium *Streptomyces hygroscopicus* in a soil sample from Easter Island (or Rapa Nui, as it is known by locals) on a Canadian scientific expedition and isolated by the Ayerst researchers in Montreal. Rapamycin was later found to be an efficient antifungal agent (Sehgal *et al.* 1975). The first demonstration of rapamycin's immunosuppressive activity was from the studies that showed its inhibitory effects upon production of humoral IgE as well as its preventative effects in two animal models of human autoimmune diseases, experimental autoimmune encephalitis and adjuvant arthritis (Martel *et al.* 1977). Currently, rapamycin is being used as a successful immunosuppressive drug during kidney transplantations.

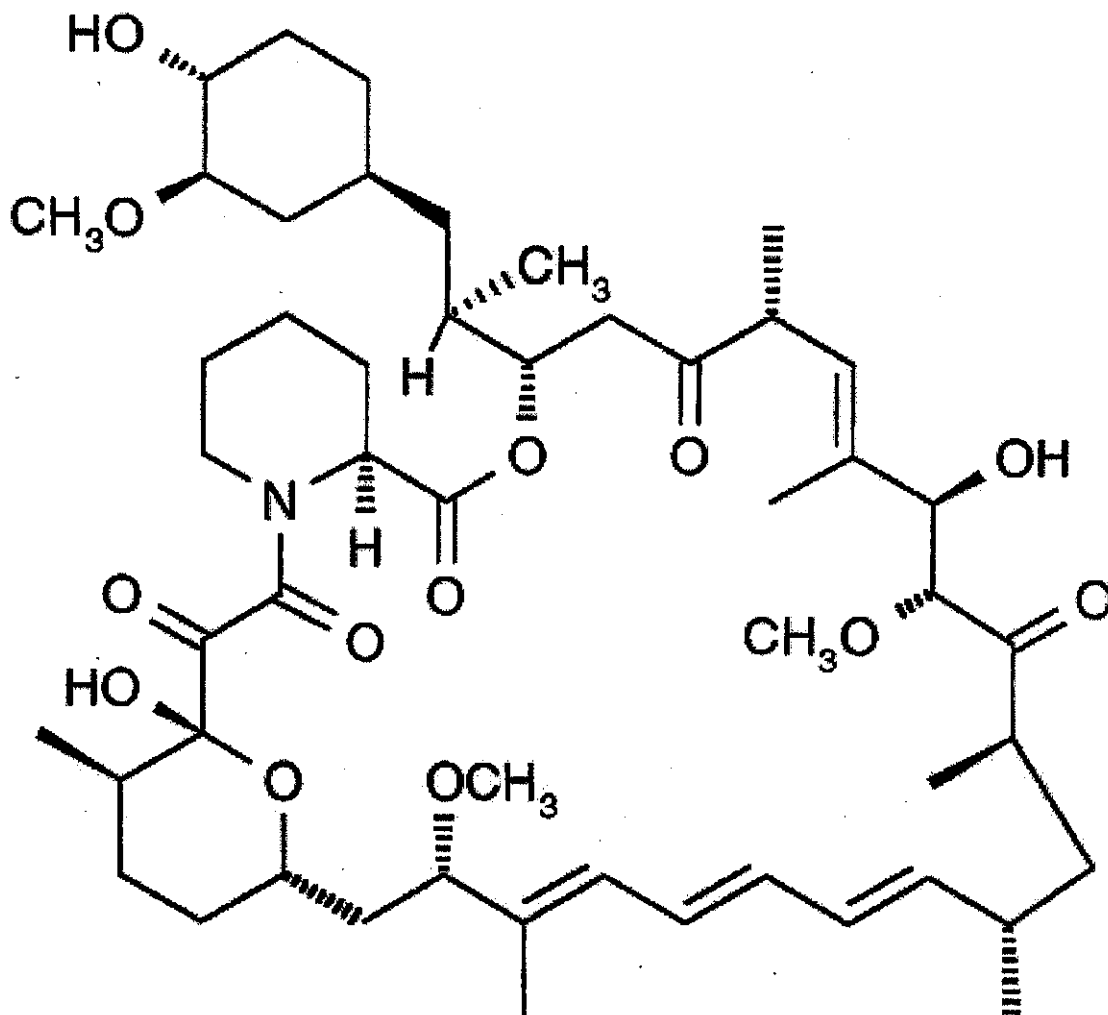


Figure 10. The chemical structure of Rapamycin

2.2 IMMUNOPHILINS

In order to understand the molecular mechanisms of immunosuppression by cyclosporin A, FK506 and rapamycin, the cellular receptors of these drugs have been purified and characterized (Schreiber, 1991; Fruman *et al.* 1994). These immunosuppressive drugs do not elicit immunosuppressive activity by themselves. Their complexes formed with their respective receptor proteins form the functional modules for immunosuppression.

An 18 kDa protein isolated from the porcine kidney was found capable of catalyzing the inter-conversion of *cis* and *trans* rotamers of amide bonds adjacent to

proline residues in peptidic substrates and the enzymatic activity was termed as peptidyl-prolyl *cis-trans* isomerization. Due to its enzymatic activity, the protein was called peptidyl-prolyl *cis-trans* isomerases or PPIase in short (Fischer *et al.* 1984). Proteins with PPIase activity are also known as rotamases. The *cis-trans* isomerization of bonds located between proline and its preceding residue is a rate-limiting step in the folding of newly synthesized proteins (Brandts *et al.* 1975). Furthermore, while investigating the cellular actions of cyclosporin A, a protein from the calf thymus was isolated as the principal binding protein for cyclosporin and was termed cyclophilin (Handschumacher *et al.* 1984). Later it became clear that PPIase and cyclophilin (CyP) are the same protein (Fischer *et al.* 1989). Also, a target protein for FK506 was identified and the protein was found to bind rapamycin as well. This new 12 kDa protein, called FK506 binding protein (FKBP) has no sequence homology to CyP, and is shown to possess PPIase activity (Harding *et al.* 1989).

Immunophilins are classified into two major families according to their binding partners: the FK-506 and rapamycin-binding proteins or FKBP; the cyclosporin-binding proteins or cyclophilins (CyP). Whilst immunophilins are defined by their ability to bind immunosuppressant ligands, they share their PPIase activity with a third family of proteins known as the parvulins (Fischer and Aumüller, 2003), which show a higher degree of specificity for their substrates. Based on domain organization, the immunophilins; both FKBP and CyP, can be classified as single-domain or multi-domain proteins. The single-domain members contain a FKBP or CyP catalytic domain and other regions that do not show identifiable domain structures except for targeting sequences for different compartments. The multi-domain members have other functional domains in addition to a single or duplicated CyP or FKBP domain(s).

2.2.1 PPIase activity of immunophilins and protein folding

In order to perform their function within the cell, newly synthesized proteins are rapidly and efficiently converted from their primary structure into their well-defined functionally competent tertiary (or) quaternary structures. While folding of globular single-domain polypeptides occurs on a second-millisecond timescale, the isomerization of the imidic peptide bond that precedes proline residues in a protein, which is also

known as the peptidyl-prolyl bond, constitutes a slower and rate-limiting step in the folding process (Kiefhaber *et al.* 1990). This is due to electron distribution and steric constraints intrinsic to prolyl bonds (Odefey *et al.* 1995; Scherer *et al.* 1998). Prolyl bonds can occur in both the *cis* and *trans* conformations. *Cis*-prolyl bonds develop unfavorable contacts with adjacent amino acid residues in this isomeric form. The *trans* form is preferred overwhelmingly in most peptide bonds; roughly 1000:1 ratio in *trans:cis* populations (MacArthur and Thornton, 1991; Reimer and Fischer, 2002). However, peptidyl-prolyl bonds in newly synthesized polypeptides tend to have a roughly 3:1 ratio, presumably because the symmetry between the C α and C δ atoms of proline makes the *cis* and *trans* isomers nearly equal in energy. Peptidyl-prolyl *cis-trans* isomerases catalyze the rapid isomerization of prolyl bonds from the *cis* to the *trans* configuration. Whilst few *in vivo* substrates of immunophilin PPIase activity have been identified, several *in vitro* protocols have been developed which allow the determination of the catalytic properties and substrate specificity of these enzymes. Data for short proline containing peptides indicate that prolyl-isomerization proceeds through a single step mechanism. In a classic PPIase assay, a synthetic peptide substrate containing a proline residue and a C-terminal para-nitroaniline (pNA) group (Kofron *et al.* 1991) is incubated with the PPIase of interest and chymotrypsin. The latter cleaves the chromogenic pNA moiety upon PPIase-catalyzed *cis-trans* conversion of the substrate, resulting in emission at 390 nm. Thus the folding kinetics of PPIases can be ascertained with small peptides.

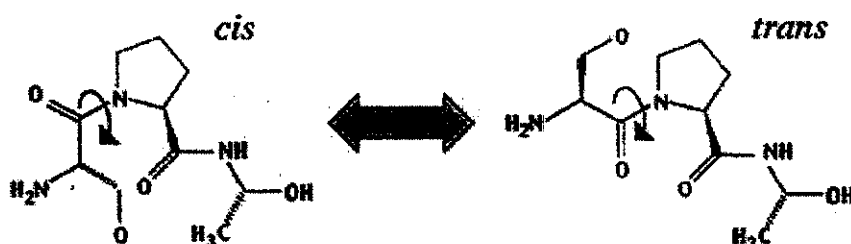


Figure 11. Peptidyl-prolyl isomerization reaction in the tri-peptide Ser-Pro-Ala

2.2.2 Immunosuppressive activity of immunophilins

Cyclosporin and FK506 are incapable of producing the immunosuppressive effect in the absence of their respective binding partners, the immunophilins. Studies so far suggest that low molecular weight, single domain immunophilins such as FKBP12 (in complex with FK506) and CyPs A and B (in complex with cyclosporin A) inhibit the activity of calcineurin, a Ca^{2+} / calmodulin-dependent protein phosphatase, thereby blocking the signaling pathway for T-cell activation (Liu *et al.* 1991b; Clipstone and Crabtree, 1992). When bound to their respective binding partners, immunophilins lose their PPIase activity as well. However, the PPIase activity of these proteins is not essential for their immunosuppressive effects. Instead, the specific surfaces of the cyclophilin-CsA and FKBP-FK506 complexes mediate the immunosuppressive action. As such, the immunosuppressive action of immunophilins is not due to the inhibition of PPIase activity (Durette *et al.* 1988).

2.3 DIVERSITY OF IMMUNOPHILINS

Immunophilins vary considerably in both form and biological function. Over three hundred sequences of cyclophilins have been identified from a wide variety of organisms, ranging from Archaea to humans (Andreeva *et al.* 1999; Ivery, 2000; Galat, 2003). Multiple immunophilins within a single organism are quite common. Recently, PPIase activity has been suggested as playing a role in the regulation of transcription and differentiation as well. Even though the initial function assigned to this group of proteins is the PPIase activity, which is responsible for proper protein folding, immunophilins are found to have several other functions as well (Romano *et al.* 2005). Not all immunophilin functions are explained by the PPIase activity. The abundance and diversity of single and multi-domain immunophilins identified to date underline the functional versatility of this family, which is further exemplified by the wide range of cellular processes in which they are involved. The diversity of immunophilins has been reviewed in a number of publications (Galat and Metcalfe, 1995; Hunter, 1998; Ivery, 2000; Galat, 2003; He *et al.* 2004; Romano *et al.* 2005; Wang and Heitman, 2005).

2.3.1 Archaeal cyclophilins

Among 2 cyclophilins found in Archaea, only the *Halobacterium cutirubrum* cyclophilin (HcCyP19) has been characterized. It is a cyclosporin A (CsA) sensitive CyP with a MW of 19.4 kDa. The PPIase activity and CsA sensitivity of this CyP is higher at higher salt concentration in the medium. Though the archaeal protein TM1367 (from *Thermotoga maritima*) has a typical cyclophilin fold, its functional features are not known yet (Jin *et al.* 2006).

2.3.2 Bacterial cyclophilins

Pathogenic bacteria possess a diverse complement of immunophilins and immunophilin-like proteins, which have been shown to play varied roles in the biology of these organisms. *Escherichia coli* possess at least six immunophilin-like proteins, including two cyclophilins such as CyPA and CyPB (Liu *et al.* 1991a). *Mycobacterium tuberculosis* is known to have two cyclophilins, termed as PpiA and PpiB. The structure of PpiA is known and a comparison of the sequence of PpiA with PpiB shows that the structural and binding features of PpiB will be much different from that of PpiA (Henriksson, *et al.* 2004).

2.3.3 Fungal cyclophilins

Immunophilins have been isolated as IgE binding proteins from several fungi (Flückiger *et al.* 2002; Horner *et al.* 1995; Lindborg *et al.* 1999), including CyP from the mold *Aspergillus fumigatus* (Asp f 11) (Cramer, 1999). Prolyl isomerases of the budding yeast, *Saccharomyces cerevisiae* have been reviewed by Arevalo-Rodriguez and his coworkers (2004). The *S. cerevisiae* proteome contains 12 immunophilin isoforms which are localized throughout the cell and have been shown, individually and collectively, to be indispensable for viability. These are encoded by eight Cyclosporin-sensitive Proline Rotamase (CPR) genes and four FK506-sensitive Proline Rotamase (FPR) genes (Dolinski *et al.* 1997). Apart from metabolic control, CPR1 of *S. cerevisiae* has been shown to mediate the import of fructose-1,6-bisphosphatase into intermediate transport vesicles (Brown *et al.* 2001). By far, the best-characterized yeast cyclophilin is CyP40. It

possesses both PPIase and three-unit tetratricopeptide repeat (TPR) domains and it constitutes an integral component of the Hsp90 chaperone machinery (Duina *et al.* 1996).

The cyclophilin repertoire of the fission yeast *Schizosaccharomyces pombe* comprises of nine members that range from small single-domain to large multi-domain proteins (Pemberton and Kay, 2005). Each *S. pombe* cyclophilin possesses only a single prolyl-isomerase domain, and these proteins vary in their degree of consensus, including the positions that are likely to affect their drug-binding ability and catalytic activity. Additionally identified motifs are involved in putative protein or RNA interactions, while a novel domain of SpCyP7 and its orthologs may have functions that include interaction with heterogeneous ribonucleoprotein (hnRNP) particles. The *S. pombe* cyclophilins are found throughout the cell but appear to be absent from the mitochondrion which is unique among the characterized eukaryotic repertoires. SpCyP5, SpCyP6 and SpCyP8 exhibit significant upregulation of their expression during the meiotic cycle. Additionally, SpCyP5 exhibit significant upregulation of its expression during heat stress. All nine *S. pombe* cyclophilins have their respective orthologs in *H. sapiens*, *D. melanogaster* and *A. thaliana*. However *S. cerevisiae* has only three cyclophilins which have orthologs in higher eukaryotes.

2.3.4 Animal cyclophilins

The soil nematode *Caenorhabditis elegans* and the fruit-fly *Drosophila melanogaster* both possess in excess of 20 immunophilin isoforms, including the ubiquitous single-domain members and more complex isoforms consisting of multiple domains. The *C. elegans* cyclophilin CyP3 is highly expressed during embryonic development and is characterized by a 'divergent loop' motif with two neighboring, unoxidized cysteine residues and two unique histidine and glutamic acid residues (Dornan *et al.* 1999). Based on the presence of the putative regulatory pair of cysteine residues and its expression pattern during development, CyP3 may function as both a redox-regulated stress responsive protein and as a foldase that is required to fold newly synthesized structural proteins during larval development. One of the earliest characterized immunophilins, the *D. melanogaster* cyclophilin NinaA forms a stable, specific complex with newly synthesized rhodopsin, thus acting as a putative molecular

chaperone in the trafficking of rhodopsin from the endoplasmic reticulum to the photosensitive membranes (Colley *et al.* 1991; Baker *et al.* 1994).

At least eight different forms of human cyclophilins have been found, ranging from 18 kDa to 150 kDa in molecular mass. Cyclophilin A (CyPA) is a cytosolic protein with a molecular mass of 17.7 kDa. CyPA binds CsA in nanomolar concentrations and appears abundantly expressed in all tissues (Liu *et al.* 1991a). CyPA functionally interacts with the major virion-associated HIV accessory protein, viral protein R and also modulates the catalytic activity of interleukin-2 Tyr kinase (Brazin *et al.* 2002; Zander *et al.* 2003). The assembly of HIV viral core requires an interaction between hCyPA and the HIV-1 *gag* protein (Colgan *et al.* 1996). Nucleocapsid protein of SARS corona virus (SARS-CoV) can bind to hCyPA and this interaction is expected to play a functional role in facilitating invasion of host cells by SARS-CoV (Luo *et al.* 2004). Several other mammalian cyclophilin isoforms have been cloned and characterized biochemically. All these cyclophilins share more than 50% sequence homology with CyPA. Cyclophilin isoforms differ in their sub-cellular localization and binding affinity to CsA. Many of them contain an N-terminal leader sequence delegating the cyclophilin to their specific sub-cellular locations.

CyPB is a 21 kDa protein containing an N-terminal signal sequence, which is thought to mediate translocation into the endoplasmic reticulum (Price *et al.* 1991). Immunofluorescent staining revealed that the avian homologue of CyPB, s-cyclophilin, is co-localized with the calcium storage protein calreticulin, a 55-kDa Ca²⁺- and calmodulin-dependent serine/threonine phosphatase, suggesting that CyPB could play a role in Ca²⁺-dependent signaling pathways (Arber *et al.* 1992). CyPB is known to exert its function in the nucleus as part of the prolactin transcriptional control machinery (Rycyzyn and Clevenger, 2002) and also triggers the integrin-mediated adhesion of peripheral blood T lymphocytes to the extra-cellular matrix (Allain *et al.* 2002).

CyPC protein is also localized in the ER. It is 23 kDa in size and contains an N-terminal signal sequence (Friedman and Weissman, 1991). The CyPC-CsA complex binds specifically to calcineurin (Liu *et al.* 1991b).

CyPD is an integral part of the mitochondrial permeability transition complex, which is regarded to have a crucial role in the mechanism of cell death. In support of a

role in stress response, the expression of certain cyclophilins, including CyPD, has recently been shown to be up-regulated under various stressful conditions (Halestrap *et al.* 2002).

Human CyP40 is a multi-domain cyclophilin and it forms a chaperone-like complex with the heat shock protein HSP90, which is required for proper protein folding of steroid receptors, similar to its bovine and yeast counterparts (Duina *et al.* 1996). CyPG or SRCyP is another multi-domain cyclophilin. In addition to the N-terminal CyP domain, human CyPG contains a serine/arginine-rich (SR) domain, similar to that found in the SR protein family of pre-mRNA splicing factors, which is required for interaction with phosphorylated C-terminal domain of RNA polymerase II. CyPG is a nuclear protein with a characteristic distribution in large and irregularly shaped nuclear speckles and co-localizes perfectly with the SR domain-containing splicing factor (Bourquin *et al.* 1997).

A novel spliceosomal-specific cyclophilin, SnuCyP20 or CyPH, has been identified (Teigelkamp *et al.* 1998). The spliceosome, which catalyzes the removal of introns from nuclear pre-messenger RNA (mRNA) molecules, is a complex consisting of the small nuclear ribonucleoprotein (snRNP) particles U1, U2, U4:U6 and U5, as well as an undefined number of non-snRNP splicing factors. CyPH assembles into the spliceosomal complex where it assists in mRNA splicing (Reidt *et al.* 2003; Ingelfinger *et al.* 2003). CyPJ is another novel member of the CyP family and its encoding gene was first identified from the human fetal brain cDNA library (Zhou *et al.* 2001).

2.3.5 Plant cyclophilins

The discovery cyclophilins in plants dates to the early 1990s during the cloning era that yielded the first plant cyclophilin-like cDNA sequences from tomato (*Lycopersicon esculentum*), maize (*Zea mays*) and oilseed rape (*Brassica napus*) (Gasser *et al.* 1990). The *Arabidopsis* genome-sequencing project has allowed the identification of all the members of the immunophilin superfamily for a single plant species. Comprising one of the largest such families identified to date, it consists of b). Extensive phylogenetic analysis has allowed thorough classification of all *Arabidopsis* isoforms and constituted a sound foundation for the ongoing functional characterization of these

proteins in plants. The discovery of plant cyclophilins has not only demonstrated conservation of these proteins in a full spectrum of biological systems, but has also provided clues to their potential functions in plants. For example, some plant cyclophilin genes have been shown to be induced by a variety of biotic and abiotic stresses, suggesting that they may play a role in environmental response processes (Chou and Gasser, 1997; Kurek *et al.* 1999). Early works that indicated the distribution of cyclophilins throughout the plant cell (Breiman *et al.* 1992; Luan *et al.* 1994) have recently been confirmed and expanded by genomic and proteomic approaches, which have provided detailed sub-cellular localization data for these large gene families. Most recently, the sequence data obtained for the unicellular green alga *Chlamydomonas reinhardtii* suggest that this sharp increase in the number and diversity of immunophilin isoforms has occurred following the divergence of red and green algae. Remarkably, the *Chlamydomonas reinhardtii* genome has been found to encode a total of 52 putative isoforms, including a high number of distinct orthologs of *Arabidopsis* immunophilins (). 29 CyP isoforms and 23 FKBP isoforms (He *et al.* 2004; Romano *et al.* 2004 Vallon, 2005

The plant cyclophilin family consists of a wide variety of isoforms varying in form, function and cellular location. While information regarding the function of this large and diverse protein family remains sporadic in plants, evidence suggests that they are involved in functions including, but not limited to, development and photosynthesis. Genetic analysis of developmental mutants has identified three multi-domain immunophilins containing TPR repeats that regulate plant development. One among these, *Arabidopsis* CyP40 has been shown to be required for the vegetative maturation of the shoot and regulates the development of leaf shape (Berardini *et al.* 2001). *Arabidopsis* CyP40 is only 48 % identical and 64 % similar to the previously studied bovine CyP40. Point mutations in the catalytic domain of the tomato cyclophilin LeCyP1 are shown to be responsible for the pleiotropic and auxin-resistant *diageotropica* phenotype (Oh *et al.* 2003). Several plant CyPs interact with an endonuclease involved in T-DNA transfer from agrobacterium to host plant cells (Deng *et al.* 1998).

Innate immunity in higher plants invokes a sophisticated surveillance system capable of recognizing bacterial effector proteins. In *Arabidopsis*, resistance to infection

by the strains of *Pseudomonas syringae* that express an effector protein requires a plant resistance protein. This bacterial effector protein was identified as a putative cysteine protease that results in the elimination of the *Arabidopsis* protein RIN4. RIN4 cleavage serves as an activation signal to the resistance protein. The effector protein gets delivered into the plant cell, where it gets activated by a eukaryotic factor identified as a cyclophilin. This activation of effector protein is necessary for protease activity. Active effector protein can then directly cleave RIN4. The discovery that eukaryotic cyclophilin can activate an effector protein not only provides a mechanism for effector activation upon delivery into the plant cell but also implicates *Arabidopsis* single domain cyclophilins as general protein folding catalysts (Coaker *et al.* 2005).

2.3.5.1 Domain organization and phylogenetic analysis in Arabidopsis cyclophilins

Romano and his coworkers (2004b) performed a comprehensive phylogenetic analysis of the *Arabidopsis* cyclophilin protein family, the largest cyclophilin family identified in any organism to date with 29 members. Their studies on both groups of immunophilins (CyPs and FKBP) revealed the structural and potentially functional diversity of these protein foldases, the broad distribution of PPIases in the sub-cellular compartments, and the large number of PPIases localized to the chloroplast.

Among the 29 cyclophilins from *Arabidopsis*, 21 are single domain forms and the remaining 8 have multi-domain organization. These eight cyclophilins have other functional domains in addition to the characteristic CyP domain. The multi-domain cyclophilin isoforms encompass a set of proteins possessing unique domain arrangements. Unlike the FKBP, none of the CyPs contains multiple catalytic domains. In addition, the functional domains in CyPs are more divergent. Besides the TPR domain that is also found in some of the AtFKBPs, some CyP members contain other functional domains such as the WD40 repeat, U-box domain, zinc finger domain and leucine zipper domain, each of which is involved in protein-protein or protein-DNA interactions. AtCyP38, the subject protein of this thesis is the smallest and the thylakoid luminal multi-domain isoforms. It is composed of a C-terminal putative leucine zipper domain, a central acidic region and the N-terminal catalytic PPIase domain. The details of domain organization among the various *Arabidopsis* cyclophilins can be seen in Fig.

An RNA recognition motif (RRM) that may interact with RNA is found in AtCyP59. It is noteworthy that most of the multiple domain CyPs have Arg- or Lys-rich domains with unknown functions. *Arabidopsis* cyclophilin AtCyP20-2 was found to associate with the periphery of Photo System II super-complexes (Romano *et al.* 2004a). Together with earlier reports (Fulgosi *et al.* 1998; Gupta *et al.* 2002), these studies suggest that a subset of immunophilins might have specifically evolved to fulfill functions that are unique to photosynthetic organisms.

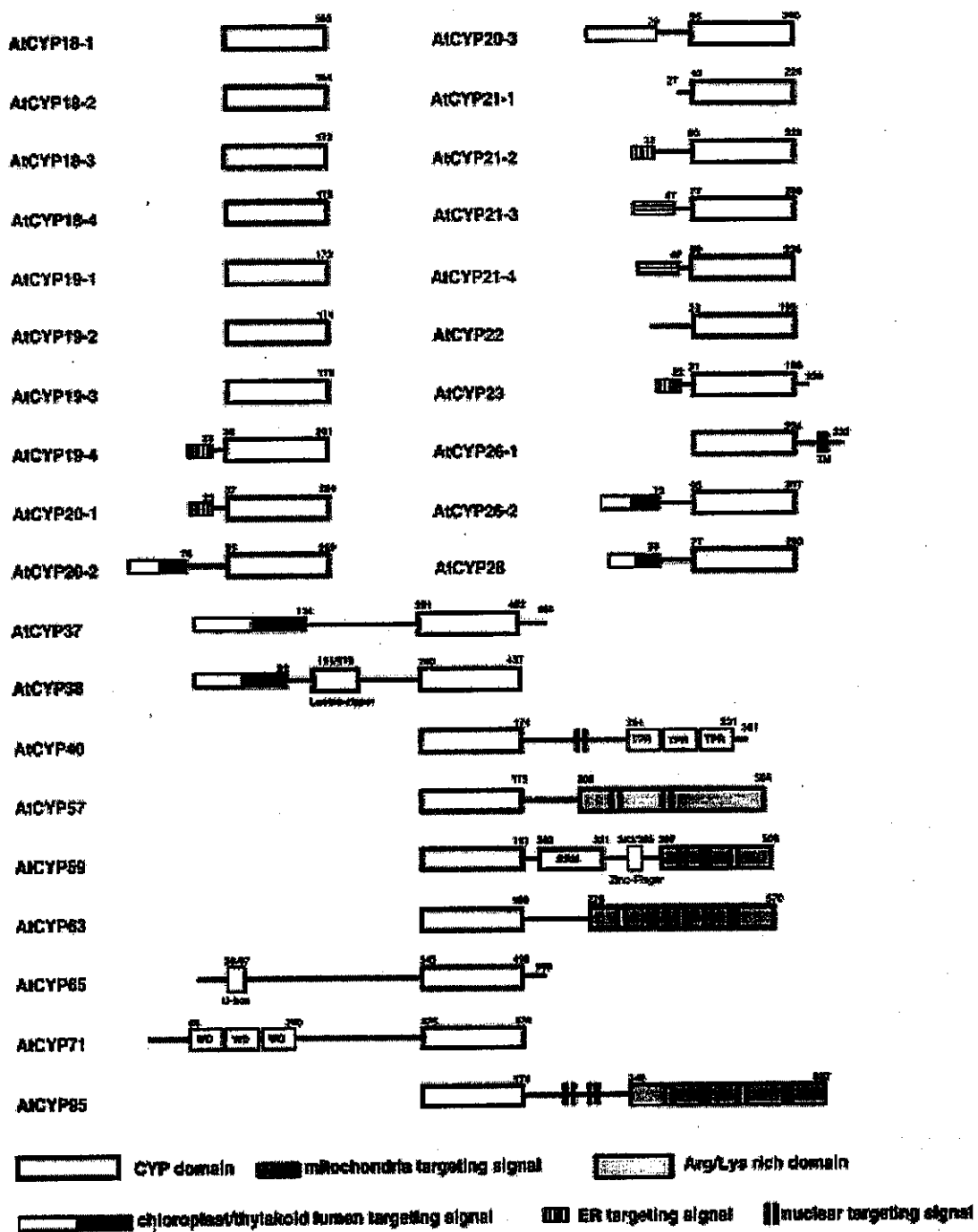


Figure 12. Domain organization in *Arabidopsis* cyclophilins

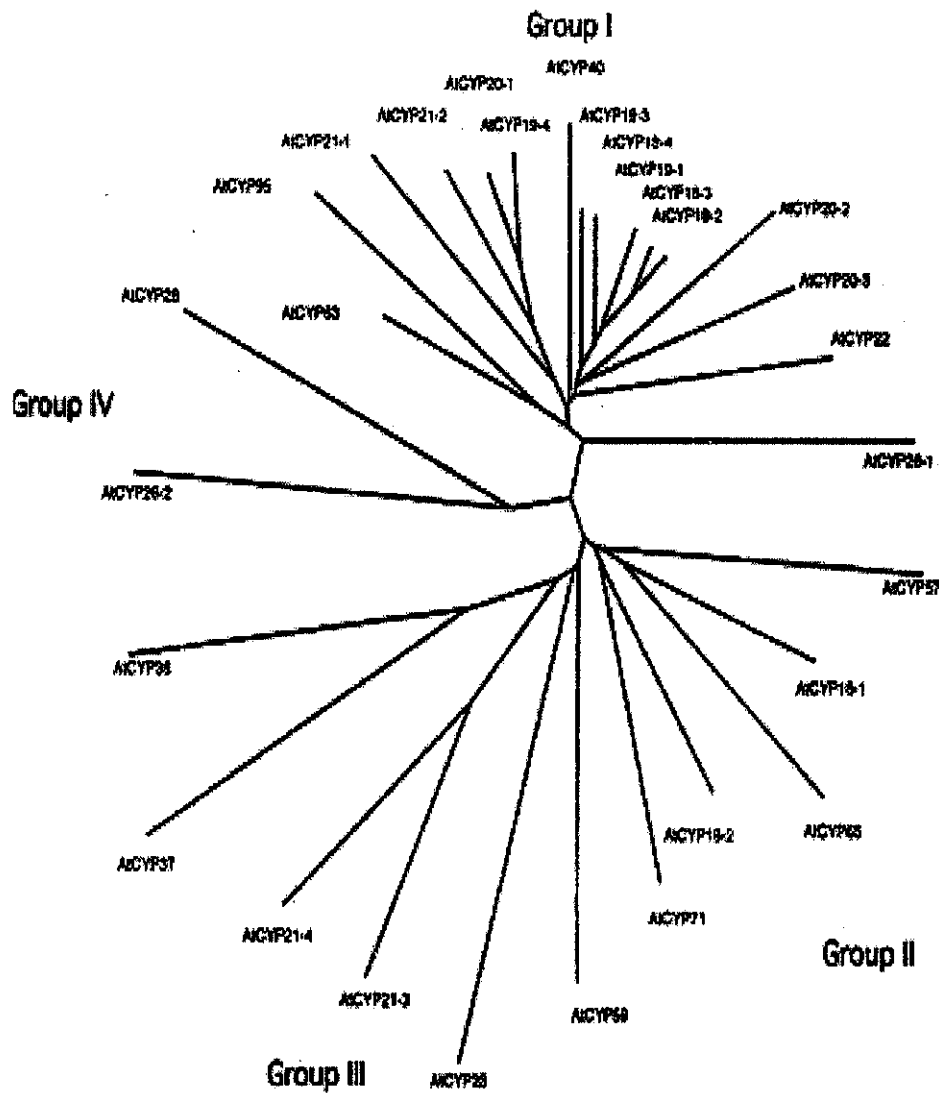


Figure 13. The phylogenetic relationships of *Arabidopsis cyclophilins*

AtCyPs could be clustered into four groups. Group I is the most complex and includes 16 members. This group could be further divided into several subgroups. Subgroup I contains five cytosolic single domain CyPs without any significant N- or C-terminal extensions. Furthermore, these five members have similar genomic structures that do not contain introns. The high similarity in both protein sequences and genomic structures indicates that they might have originated from gene duplication. The only other

Arabidopsis immunophilin gene without any intron is AtCyP26-1 (an isoform of CyP26) that also belongs to Group I but contains a unique C-terminal transmembrane domain. Subgroup II contains two chloroplast CyPs, one located in stroma (AtCyP20-3) and the other in the thylakoid lumen (AtCyP20-2). Subgroup III contains four secretory pathway CyPs, including AtCyP19-4, 20-1, 21-1, and 21-2. Three multiple domain CyPs, AtCyP40, 63, and 95, are predicted to be localized in the nucleus and form subgroup IV. The remaining two members of Group I are AtCyP22 (with an N-terminal extension) and AtCyP26-1 (with a C-terminal extension) that might be located in the cytosol.

Group II of the AtCyP family includes five cytosolic cyclophilins: three multiple-domain cyclophilins, AtCyP57, AtCyP65, and AtCyP71, and two single domain cyclophilins, AtCyP18-1 and AtCyP18-2. The genes that encode the latter two single domain cyclophilins contain five and six introns, respectively, in contrast to those in subgroup I of group I (with no intron).

Group III contains six members which have different sub-cellular localization patterns. AtCyP21-3 and AtCyP21-4 are the only two *Arabidopsis* immunophilins that are predicted to be located in the mitochondria. The high similarity between both the precursors and the mature proteins of these two CyPs indicates that they might be generated by gene duplication. AtCyP37 and AtCyP38 are thylakoid lumen cyclophilins which are highly divergent AtCyPs. AtCyP23 (a secretory cyclophilin) and AtCyP59 (a nuclear multi-domain cyclophilin) also belong to this group.

Group IV contains two thylakoid lumen cyclophilins, AtCyP26-2 and AtCyP28. They are the most divergent AtCyPs as only a maximum of 3 out of the 13 important PPIase and drug binding amino acids are conserved.

The single-domain cyclophilins have just the cyclophilin domain and some of them have additional signal peptides for delivery to different organelles. Interestingly, a number of CyPs are predicted to get translocated to chloroplast. All these CyPs (total of six) contain the typical N-terminal transit peptide for chloroplast import. Five of them, including AtCyP20-2, 26-3, 28, 37, and 38, may be translocated into the thylakoid lumen as they contain the bipartite signal sequences for crossing both the envelope and thylakoid membrane.

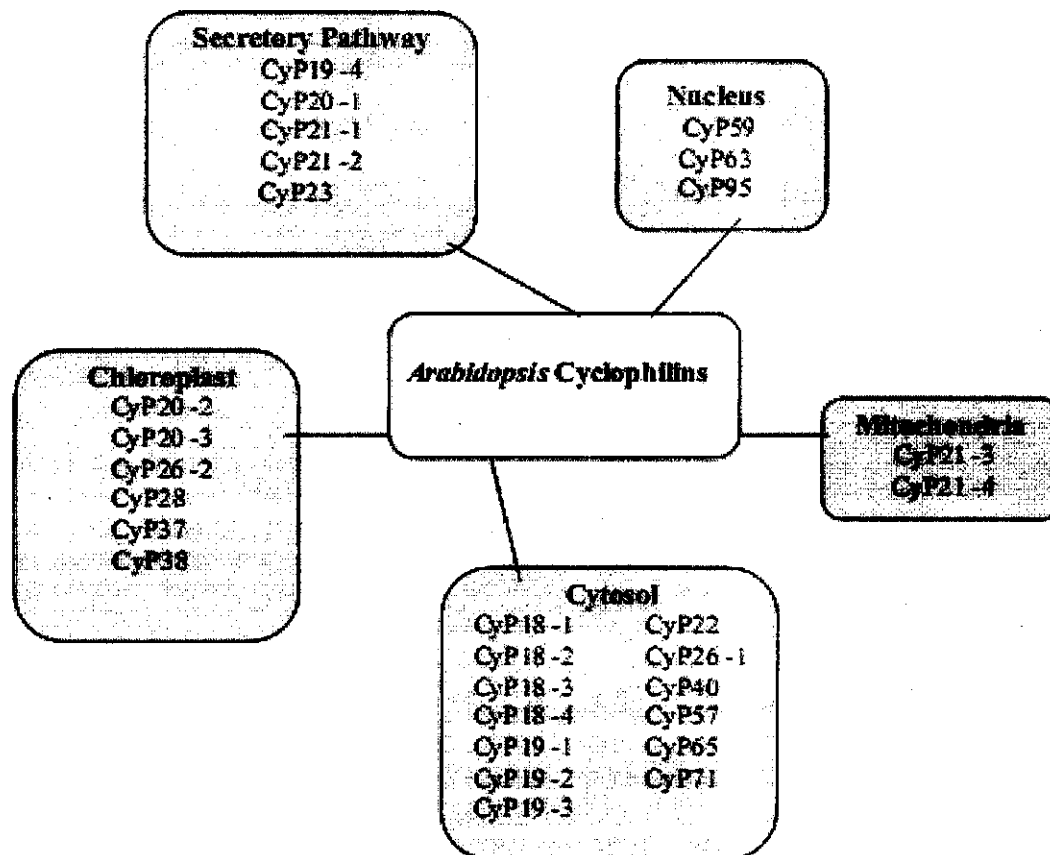


Figure 14. *Arabidopsis* cyclophilin isoforms and their sub-cellular localization

The homolog of AtCyP38 from pea (*Pisum sativum*), TLP40, has been shown to be located in the thylakoid lumen and the partition between thylakoid membrane and soluble luminal fraction (Fulgosi *et al.* 1998). The association and/or dissociation of TLP40 from the thylakoid membrane regulates phosphorylation of the photosystem II components (Vener *et al.* 1999). The chloroplast (luminal or stromal) localization *Arabidopsis* Cyclophilins Chloroplast CyP20 -2 CyP20 -3 CyP26 -2 CyP28 of some of the predicted CyPs have also been confirmed by chloroplast import, immunodetection and peptide sequencing efforts.

Two of the CyPs, AtCyP21-3 and AtCyP21-4, are predicted to be targeted to the mitochondrion. They contain very similar signal peptides of 47 amino acids and have highly homologous C-terminal cyclophilin domains. A mitochondrial cyclophilin (CYPD) in animal cells modulates the mitochondrial permeability transition pore (MPTP)

and plays a critical role in apoptotic and necrotic cell death (Connern and Halestrap, 1994; Crompton, 1999). A similar process may exist in plants and plant mitochondrial cyclophilins may be involved in similar functions (Arpagaus *et al.* 2002). The presence of mitochondrial cyclophilins has also been indicated by the identification of CsA-sensitive rotamase activity and purification using a CsA affinity column (Breiman *et al.* 1992; Luan *et al.* 1994).

There are five cyclophilins predicted to go through the cellular secretory pathway. These include AtCyP19-4, AtCyP20-1, AtCyP21-1, AtCyP21-2, and AtCyP23. Unlike the ER-located AtFKBPs, these cyclophilins do not have a C-terminal ER-retention signal. Their exact location needs to be determined by further experiments. AtCyP19-4 (CYP5) has been shown to be located in the ER by N-terminal green fluorescent protein fusion (Saito *et al.* 1999). This protein displays PPIase and protein refolding activities that are sensitive to CsA. Furthermore, AtCyP19-4 is suggested to modulate the function of an *Arabidopsis* protein, GNOM, a guanine nucleotide exchange factor, which is essential for vesicle trafficking in many organisms (Grebe *et al.* 2000). Another secretory pathway cyclophilin is AtROC7 (Jackson and Soll, 1999), which may regulate the protein phosphatase 2A activity.

Several single domain cyclophilins are predicted to be cytosolic proteins. These include AtCyP18-1, AtCyP18-2, AtCyP18-3, AtCyP18-4, AtCyP19-1, AtCyP19-2, and AtCyP19-3. These proteins are highly conserved and may have similar functions. AtCyP18-3 is shown to interact with *Agrobacterium* VirD2, an endonuclease that is covalently bound to the 5' end of T-DNA. Because T-DNA transfer is inhibited by CsA, CYP interaction with VirD2 may be important for agrobacterial infection (Deng *et al.* 1998). AtCyP26-1 is another CyP that is predicted to be cytosolic. However, there is a potential trans-membrane domain at the C-terminus which may serve as a membrane anchor.

2.3.5.2 Evolutionary dynamics of the luminal cyclophilins

Even though the sub-cellular localization patterns of all the chloroplast cyclophilins are known, their *in vivo* functions are not known. It is interesting to note that, with the exception of AtCyP20-2, the luminal cyclophilin-like proteins possess an

atypical and poorly conserved primary structure. AtCyP26-2, AtCyP28, AtCyP37 and AtCyP38 show only 10-40% similarity to human cyclophilin A (Romano *et al.* 2004b). The reduced pH tolerance that is observed for AtCyP20-2 PPIase activity may explain the reduced number and conservation of this cyclophilin-type immunophilin. AtCyP26-2, AtCyP28, AtCyP37 and AtCyP38 may have forfeited their PPIase function in order to carry out novel roles within this catalytically suboptimal cellular environment. Given the localization of AtCyP20-2 and AtCyP38, one may surmise that these luminal cyclophilins may have evolved functions related to membrane proteins.

Studies in both animal and plant systems have revealed a diverse array of functions for individual cyclophilin members. Such functions can be a result of their protein foldase activity, chaperone activity, scaffolding activity, and other unknown activities. A very clear distinction between cyclophilins and other types of protein foldases and molecular chaperones is that each member of the cyclophilin family appears to have specific targets and functions in the cell. This is consistent with the fact that the sequence and structure of cyclophilins are rather divergent although a conserved drug-binding core is present in all members. The specific sequence motif in each member may present the structural basis for interaction with specific targets. Another general rule for the immunophilin function is their association with super-molecular complexes in both animal and plant cells. It is speculated that each immunophilin member may function in the maintenance of these complexes.

2.4 CYLOPHILIN STRUCTURES

2.4.1 Ligand-free cyclophilin structures

Even though no three-dimensional structures are available for the members of the plant cyclophilin family, structures of several cyclophilin isoforms from other sources are known. The available cyclophilin structures have thoroughly been reviewed (Taylor *et al.* 1997; Dorman *et al.* 2003; Ke and Huai, 2004).

Several cyclophilin structures have been studied in detail and all of them share a common fold. All cyclophilins have an eight stranded β -barrel, capped by two α -helices. The topology of this fold is rather different from other β -barrel structures which transport hydrophobic ligands in a pocket in the middle of the barrel. In CyPA, the center of the

sandwich is filled with closely packed aromatic groups. Whereas, in the case of most other β -barrel structures, the hydrophobic core of the barrel is almost always open for the substrate to enter. The N- and C- termini lie close together on the same side of the protein, distal to the active site face. The X-ray structures of numerous complexes define the PPIase active site that is mainly formed by residues located on one face of the β -sheets. Cyclophilins are neither functionally nor evolutionally related to other eight stranded β -barrel proteins. The anti-parallel β -strands of a typical cyclophilin form a right-handed β -barrel. The tilted angle of the strands against the barrel axis is about 49°. Even though all known eight-stranded β -barrels are right-handed, their tilt angles are considerably different from that of a cyclophilin and hence it is difficult to superimpose them properly.

The hCyPA protein is the first and most widely studied cyclophilin (Kallen *et al.* 1991; Ke *et al.* 1991). The β -barrel of hCyPA has dimensions of about 14 Å height, and 15 and 17 Å for the minor and major axes of the elliptical cross section of the barrel. In this structure, helices, strands and turns are 21%, 36% and 10% of the total amino acid residues, respectively. The hydrophobic core of the barrel is closed by Helix 2 and Helix 4 on both ends so that ligands cannot enter the core but bind only to the surfaces of the barrel. There are four cysteine residues in CyPA, but no disulfide bond is formed among them (Ke, 1992). The X-ray (Edwards *et al.* 1997) and NMR (Clubb *et al.* 1994) structures of *E. coli* CypA have also been determined and show a similar fold (with an r.m.s.d. fit of 1.4 Å for the core residues) but with significant differences in the surface loops.

Cyclophilin B, the second member of the cyclophilin family is 208 amino acids in length and contains an N-terminal extension, characteristic of endoplasmic reticulum directed signal sequences. The core cyclophilin domain of human CyPB shares 65% amino acid identity with human CyPA, but is flanked by variable N and C terminal regions. The structure of human CyPB, in complex with CsA, was solved in 1994 (Mikol *et al.* 1994). The overall structure of CyPB is very similar to that of CypA. The main differences are at the N and C-terminal extensions and in two loop regions. *Escherichia coli* CyPB, one of its two cyclophilins, is also shown to have a fold similar to that of human CyPB (Konno *et al.* 2004). Two of the cyclophilin isoforms from *Caenorhabditis*

elegans, CyP5 and CyP6, are closely related and the crystal structure of recombinant CyP5 most strongly resembles that of human CypB with changes in loop regions and the N and C-terminal extensions (Picken *et al.* 2002).

Unlike CyPA and CyPB, the expression pattern of Cyclophilin C (CyPC), the third member of the cyclophilin family was found to be highly restricted to specific kidney tissues. The sequence identity between murine CyPC and human CyPA in their core cyclophilin domain is about 56%. The structure of murine CyPC differs from that of human CypA mainly in the conformations of three surface loop regions (Ke *et al.* 1993). Another study showed that the binding affinity of CyPC for CsA is much lower than that of CyPA (Schneider *et al.* 1994) as was the case for CyPB (Price *et al.* 1991). The reduced affinity for CsA is unexpected, as the active site residues are well conserved in these two cyclophilins as compared to CyPA.

The structure of a point mutant of human CyPD shows that it has very much the same fold as that of CyPA and there are only minor differences in its active site as compared to that of CyPA (Schlatter *et al.* 2004).

Divergent cyclophilins form a separate class of cyclophilins wherein, an additional loop of 5 to several amino acids is present at about residue 50 (hCyPA numbering). Also, another common feature of this group of proteins is the presence of two reduced cysteine residues, which come together in close proximity and hence might be involved in a redox signaling process. The presence of a conserved histidine in the divergent loop and close to the two cysteine residues is suggestive of a co-ordination site for zinc or other metals. The overall fold of the protein however remains the same as that of hCyPA. Most of the members of this family have only minor differences in their active site as compared to hCyPA and hence show similar CsA binding characteristics. However, the loop conformations adopted by all known divergent-loop structures are significantly different from each other.

The structures of a few divergent cyclophilins such as *C. elegans* CyP3 (Dornan *et al.* 1999), *B. malayi* CyP1 (Taylor *et al.* 1998), bovine CyP40 (Taylor *et al.* 2001), *P. falciparum* CyP (Peterson *et al.* 2000) and human CyPH (Reidt *et al.* 2000; 2003) are known. In all these structures, the cyclophilin fold brings the two cysteine side chains into close proximity with the sulfur atoms less than 6Å apart. This leads to a speculation

that these proteins may be involved in a redox signaling process that might be triggered by the formation of a disulphide bond (Dornan *et al.* 1999).

Mycobacterium tuberculosis PpiA also belongs to one among the divergent cyclophilins with an extra insert and a different N-terminal segment. The X-ray structure of this cyclophilin reveals an overall fold similar to that of human CyPA (Henriksson *et al.* 2004). The catalytic arginine and the residues in the active site are highly conserved and the CsA binding characteristics are expected to be similar to those of human CyPA.

Human CyPJ is another divergent cyclophilin that shows about 45-51% identity to human CyPA, CyPC, CyPH and murine CyPC. Even though the overall three-dimensional structure of CyPJ is similar to that of CyPA, it has some segments with striking differences and an additional disulfide bridge (Huang *et al.* 2005). Barnstein and co-workers (unpublished) have solved the crystal structure of the PPIase domain of the multi-domain human cyclophilin CyPG (PDB id: 2GW2).

Even though several multi-domain cyclophilins are known, the only multi-domain cyclophilin structure known to date is that of bovine cyclophilin 40 (CyP40), which has a cyclophilin domain at the N-terminus and a tetratricopeptide repeat (TPR) domain at the C-terminus (Taylor *et al.* 2001). An acidic linker region connects the N-terminal cyclophilin domain with the helical TPR domain. The TPR domain consists of seven helices and consists of 11 aspartate and glutamate residues. This domain forms a well-defined hydrogen bonded structure with two β -turns and a larger loop comprising 194PKDGS GD200. There is no direct contact between the helices of the TPR domain and the cyclophilin domain, although the first two helices of the TPR domain form an intimate interface with the linker region. Generally, each TPR repeat is made up of about 34 amino acids and forms a helix-turn helix motif. However, in CyP40, the first three helices are longer than expected lengths for TPR sequences. The surface of this domain shows a pronounced charge distribution in which the inner groove forms a positively charged surface while the outer groove forms a negatively charged surface. The positively charged concave groove provides a recognition and binding site for the conserved C-terminal 'MEEVD' sequence of Hsp90 (Carrello *et al.* 1999). Being a multi-domain cyclophilin, AtCyP38 was predicted to have a domain organization similar to that of CyP40, except for the fact that unlike CyP40, the cyclophilin domain of AtCyP38 is

located at the C-terminus. The structure overlap of the cyclophilin domains of hCyPA, bovine CyP40 and some other cyclophilins is shown in Fig. 16.

The crystal structure of *Aspergillus fumigatus* cyclophilin (Asp f 11) reveals dimerization by 3D domain swapping and represents one of the first proteins with a swapped central domain (Limacher *et al.* 2006). The domain-swapped element consists of two β -strands and a subsequent loop carrying a conserved tryptophan. The tryptophan binds into the active site, inactivating *cis-trans* isomerization. Therefore, dimerization of Asp f 11 *in vivo* could be a means of regulating its own biological function. The *Thermotoga maritima* protein TM1367 exhibits a remarkable structural similarity to human CyPA. It has an eight-stranded, anti-parallel, closed β -barrel structure. But it does not have even a single residue of the PPIase active site conserved. Also, significant structural differences are observed in the loop regions that surround the active site. Thus, it represents a novel member of the cyclophilin (peptidyl-prolyl isomerase) fold. TM1367 also exhibits a strong structural similarity to the C-terminal domain of a protein of unknown function from *B. cereus*. However, this protein forms a seven-stranded β -barrel rather than an eight-stranded β -barrel and lacks the corresponding N-terminal β -strand in the TM1367 structure. Despite lacking one β -strand of the β -barrel, the MCSG hypothetical protein superimposes remarkably well with CyPA and TM1367 (Jin *et al.* 2006).

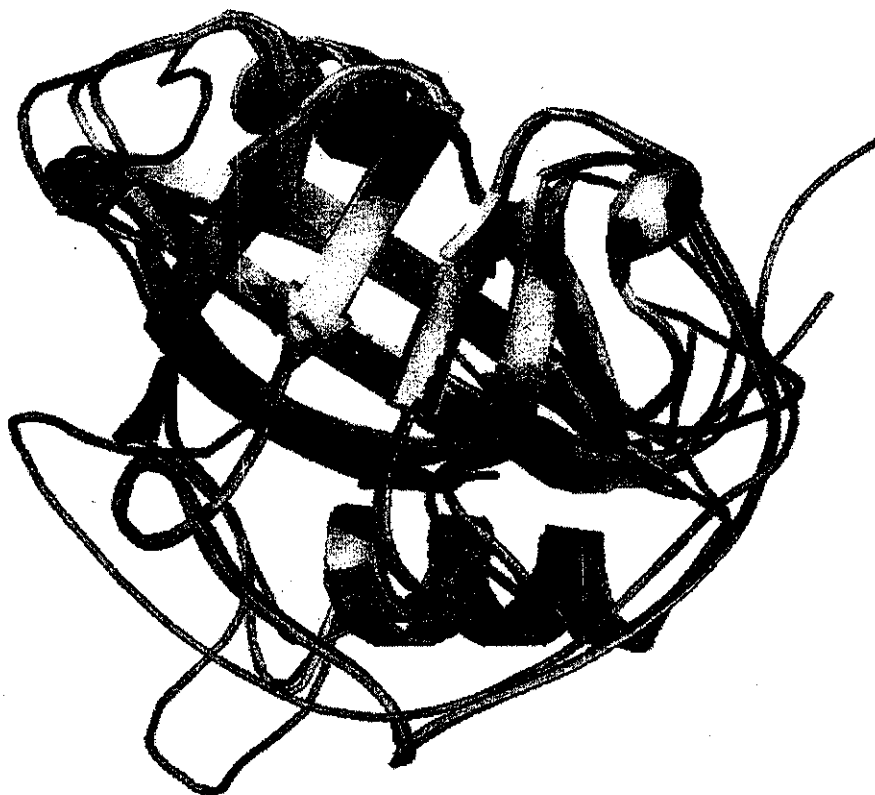


Figure 15. Structure overlap of various cyclophilin domains: bovine CyP40 (cyan), human CyPA (orange), human CyPB (pale yellow) and *E. coli* CyPB (violet). The overlap was generated with the MULTIPROT server (Shatsky *et al.* 2004) and the figure was prepared using PYMOL.

2.4.2 Cyclophilin-ligand complex structures

Several cyclophilin structures are available in complex with a ligand. More than 20 different X-ray structures of native and CsA analogue structures have been determined. All structures of native and CsA analogues show very similar conformations and share the same binding mode. The binding is enhanced by 6 conserved hydrogen bonds between CsA and CyPA. The different CsA analogues can have a significant effect on binding and on calcineurin inhibition (Kallen *et al.* 1998), even though there are few significant changes in their 3D structures. Even chemically, different inhibitors like peptolide 214103 adopt the same bound conformation as CsA (Mikol *et al.* 1998). A series of non-immunosuppressant CsA analogues have also been developed which can bind strongly to CyPA but have small chemical differences in the effector loops.

However, the binding of the complex to calcineurin is dramatically reduced and there is no measurable immunosuppressant effect. Such modified cyclosporins are of interest in anti-HIV therapy, where inhibition of CyPA alone prevents replication of the HIV virion (Thali *et al.* 1994; Franke and Luban, 1996). Another non-immunosuppressive CsA derivative is the drug Valspodar (PSC833), which exhibits the capacity to chemosensitise tumor cells that show multi-drug resistance associated with MDR1-glycoprotein over-expression. The changes in Valspodar, as compared to CsA, significantly reduce the binding to CyPA and there is a very significant drop in immunosuppressive activity (Keller *et al.* 1992).

Three families of small molecular ligands that can bind to the active site of cyclophilins have also been discovered. The piperidine family has been designed, based on the chemical similarity with dipeptide structures. The X-ray structures of complexes with 1-acetyl-3-methyl piperidine and ethyl-piperidine glyoxylate have been characterised (Kontopidis *et al.* 2004). The structure of a complex of hCyPA with dimethylsulfoxide (DMSO) bound in the active site has led to the design of a series of related inhibitors. When crystallized in the presence of DMSO, two DMSO molecules were found in close proximity to the active site of human CyPD (Schlatter *et al.* 2005). A number of reports have appeared in the literature describing the isolation, characterization and immunosuppressive activity of a novel cyclophilin binding compound, Sanglifehrin A (Zenke *et al.* 2001).

A number of anti-arthritic gold (I) drugs are in clinical use. Their critical target sites are thought to be thiolate sulphurs (cysteine residues). However, there is little structural data on the adducts of anti-arthritic Au (I) complexes with proteins. The crystal structure of a medically relevant gold compound [Au (I) phosphine complex (AuPEt₃Cl)] in complex with CyP3 from *C. elegans* has been reported (Zou *et al.* 2000). The gold is bound to the active site histidine and it is shown to be inhibitory to PPIase activity.

Calcineurin (protein phosphatase 2B, Cn) is activated by calcium-loaded calmodulin (CaM) and is inhibited by both the complexes, FKBP12-FK506 and CyPA-CsA. It is a heterodimer of chains CnA (60 kDa) and CnB (19 kDa). The crystal structures of the ternary complex between CyPA, CsA, truncated CnA and full length CnB have been reported (Huai *et al.* 2002; Jin and Harrison, 2002). The protein-drug

complex binds at the interface between the catalytic and regulatory subunits of Cn. The conformation of both CyPA and CsA are largely unaltered from those seen in the binary complex alone.

Cyclophilin complexes with proline containing dipeptides, tripeptides and tetrapeptides have been determined (Kallen *et al.* 1991; Kallen and Walkinshaw, 1992; Ke *et al.* 1993; Konno *et al.* 1996; Zhao and Ke, 1996). In all these cases, binding of the peptide to a cyclophilin is similar and always the proline residue in the peptide adopts a *cis*-conformation.

Cyclophilins are potential drug targets for anti-HIV therapy as the interaction between human CyPA and the HIV-1 *gag* protein is required to promote the assembly of the viral core (Colgan *et al.* 1996). A number of structures of CyPA in complex with the various fragment lengths of the amino terminal domain of HIV-1 capsid (Gamble *et al.* 1996; Zhao *et al.* 1997) as well as with a series of hexapeptides (Vajdos *et al.* 1997) have been characterized. In contrast to all other peptide complexes, the proline residue of the HIV-1 *gag* peptide binds in the active site of the cyclophilin in the *trans* conformation. With longer peptides, one of the CyP loops moves in order to accommodate the peptide, whereas binding of smaller peptides leaves the CyP structures unaltered.

2.4.2.1 Structural features of cyclosporin A binding

The crystal structures of CyPA-CsA complex reveal the formation of pentameric or decameric complexes, which however do not seem to have any biological relevance, as monomers in complex with CsA are capable of causing immunosuppression (Pflugl *et al.* 1993; 1994). In a monomer of CyPA-CsA, six residues of CsA form close contacts with the CyPA active site, which is made up of the 13 residues: Arg55, Phe60, Met61, Gln63, Gly72, Ala101, Asn102, Ala103, Gln111, Phe113, Trp121, Leu122 and His126 (Ke and Huai, 2004).

The binding is enhanced by 5 conserved hydrogen bonds between CsA and CyPA. All peptide bonds of the bound CsA have the *trans* conformation in the complex structures. Most other cyclophilins have the active site residues well conserved and hence the binding of CsA basically remains the same. In cases where two or more of the active site residues are different, the level of CsA binding appears to be significantly less.



Figure 16. The structure of human CyPA. α -helices are shown in red, loops in green and β -strands as yellow arrows. The residues of the active site that are involved in cyclosporine A (CsA) binding and PPIase activity are highlighted in blue as stick model. This figure was prepared with the PYMOL program.

2.4.2.2 Structural features of prolyl-peptide binding

As discussed before, cyclophilins perform PPIase activity by catalyzing the *cis-trans* isomerization of a peptidyl-prolyl amide bond. The active site can bind to a peptide of varying lengths having a proline residue with at least one residue preceding the proline. The active site can bind to CsA as well. As CsA and the peptide essentially compete for the same active site, CsA binding leads to the complete loss of PPIase activity. Twelve active site residues of hCyPA that are located on the surface of its hydrophobic pocket and interact with a peptide substrate are more or less the same as that for CsA binding. The subtle difference is in the fact that Ile57 also takes part in binding

the peptide whereas Gly72 and Ala103 do not. Four hydrogen bonds between the peptide and the active site residues are known to stabilize the interaction. In addition, the side chain of Arg148 of CyPA also forms a hydrogen bond with the peptide. Most crystal structures reveal that the *cis* form of the proline peptide binds to CyP (Kallen *et al.* 1991; Zhao and Ke, 1996; Konno *et al.* 1996). But the N-terminal domain of HIV-1 capsid p24 is known to bind to CyPA in the *trans* peptidyl form (Zhao *et al.* 1997).

The cyclophilin domain of CyP40 is quite similar to that of hCyPA and the active site residues are identical to CyPA, apart from His141 of CyP40 that corresponds to Trp121 in CyPA. Site directed mutagenesis of this residue in CyPA has shown that it plays a significant role in CsA binding, but not very important in the PPIase enzymatic activity (Liu *et al.* 1991). This observation is in line with the catalytic and CsA binding properties of bovine CyP40.

2.5 HOMOLOGS OF ATCYP38

Thylakoid lumenal proteins (TLP) are present in higher plants as well as lower forms of life such as the blue green algae. Thylakoid Lumen Protein 40 (TLP40) from spinach, rice and pea has a very similar sequence as that of AtCyP38 with 82% identity, Fig. 18. Such a high degree of identity suggests that they might perform very similar functions. Furthermore, AtCyP38 is known to have homologs in other plants, all of which are localized in the thylakoid lumen of chloroplasts. However, cyclophilins with the same domain organization of AtCyP38 have not been reported in any higher organism and hence these are thought to have evolved to perform functions that are specific to plants.

AtCyP38	VANFVIPDVSVLISGPPKIDPEALLRYALPIDNKAIREVOKBEEDITDSEKIAGVKALDS
Spinach TLP40	LTSPVLPDLAVLISGPPKIDPEALLRYALPIDNKAIREVOKBEEDITESEKISGKALDS
Pea TLP40	AANSALSDELVLISGPPKIDPGALLRYALPIDNKAIREVOKBEEDITDSEKISGKALDS
Rice TLP40	PLEFVIPDVSVLISGPPKIDPGALLRYALPIDNKAVREVOKPLEEITDSEKISGKALDS
AtCyP38	VERNVRQASRTLQQGKSIIVAGFAESKDHGEMIEKLEAGMQLKIVEDRKRDAVAPK
Spinach TLP40	VERNLRQASRTLQQGKSLIIAGLAESKDHGVELLDKLEAGMQLQIVENRNREGVAPK
Pea TLP40	VERNVRQASRTLQQGKTLVSLAESKDHGIELIDKLEAGIDEFELTLDGIEALLDQN
Rice TLP40	VERNVRQASRALNNGRNLLLGGLAESKRANGEELLDKLAVALDELQRIVEDRNRDAVAPK
AtCyP38	QKEILKYVGGIEEDMVDGFPYEVPEEYRNMPLLKGRASVDMKVKIKDNPN-IEDCVIRIV
Spinach TLP40	QRELLQYVGSVEEDMVDGFPYEVPEEYQTMPLKGRAVVMKVKVDNPN-VDNCRIRIV
Pea TLP40	RKNFLQYVGGIEEDMVDGFPYELPEEYRNMPLLKGRAAVDMKIKIKDNPKRVDECVFHLV
Rice TLP40	QKELLQYVGTVEEDMVDGFPYEVPEEYSSMPLKGRATVDMKVKIKDNPN-LEDCVIRIV
AtCyP38	LDGYNAPVTAGNFVDLVERHEYDGMETIQSDGFVVQTDPEGPAEGTIDESTEKTRIVEL
Spinach TLP40	LDGYNAPVTAGNFVDLVERHEYDGMETIQSDGFVVQTDPEGPAEGTIDESTEKTRIVEL
Pea TLP40	LDGYNAPVTAGNFVDLVERHEYDGMETIQSDGFVVQTDPEGPAEGTIDESTEKTRIVEL
Rice TLP40	LDGYNAPVTAGNFVDLVERKPYDGMETIQSDGFVVQTDPEGPAEGTIDESTEKTRIVEL
AtCyP38	EIMVTGERTFFYGSTLEELGLYKAQVVFNFNAGFTMAMARETEEDNSGSSQVEWLLKESE
Spinach TLP40	EIMVEGEKVFYVYGSTLEELGLYKAQTKLFFNAGFTMAMARETEEDNSGSSQVEWLLKESE
Pea TLP40	EIMVEGEKAPVYGETLEELGLYKAQTKLFFNAGFTMAMARETEEDNSGSSQVEWLLKESE
Rice TLP40	ELMVDGDKAPVYGETLEELGLYKAQTKLFFNAGFTMAMARETEEDNSASSQVEWLLKESE
AtCyP38	LTPSNANILDGRYAVFGYVTDNEDFLADLKVGDVIESIQVVSGLNLANPSYKLAG
Spinach TLP40	LTPSNANILDGRYAVFGYVTDNEDFLADLKVGDVIESIQVAVSGVDNLVNPYKLAG
Pea TLP40	LTPSNANILHGRYAVFGYVTDNEDFLADLKVGDVIESIQVVSGLNLANPSYKLAG
Rice TLP40	LTPSNANILDGRYAVFGYVTDNEDFLADLKVGDVIESIQVVSGLNLANPSYKRIAG

Figure 17. Sequence alignment (by CLUSTALW) of AtCyP38 with TLP40 from spinach, pea and rice. The predicted helical domain is given in blue and the cyclophilin domain in red. The identical sequences are highlighted in yellow.

2.6 SUPEROXIDE DISMUTASE:

Enzymes occupy a primary position in our life. They are present in all organisms and carry out vital functions. Enzymes are proteins that act as catalysts, speeding the rate at which biochemical reactions precede but not altering the direction or nature of the reactions.

Superoxide dismutase (SOD, EC 1.15.1.1) catalyzes the dismutation of superoxide into oxygen and hydrogen peroxide, Fig. 11. It functions as an antioxidant. Irwin Fridovich and Joe McCord discovered superoxide dismutases. It is an anti-oxidative enzyme that catalyzes the dismutation of two superoxide anions to hydrogen peroxide and molecular oxygen. The toxic hydrogen peroxide is further rapidly reduced by catalase or glutathione peroxidase (GPX) into water and molecular oxygen. SOD is the rate limiting enzyme in this pathway (Sun and Chen, 1998; Allen and Tresini, 2000) and its upregulation is of high therapeutic interest.

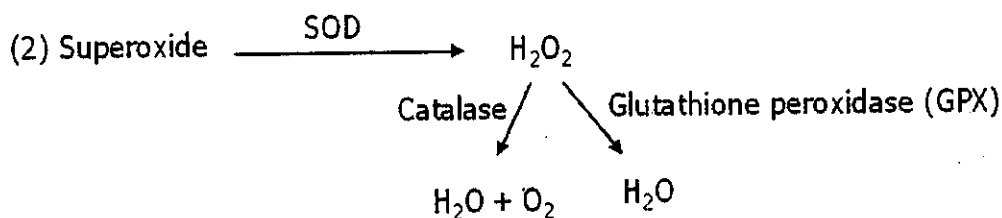


Figure 18. Biological function of superoxide dismutase (SOD).

Superoxide radicals play a detrimental role in a variety of neurodegenerative diseases (Venarucci *et al.* 1999), including stroke, and their accumulation ultimately leads to oxidative stress, causing cellular damage and death. The rapid breakdown of superoxide anions is needed to minimize overall tissue damage. Overexpression of SOD in transgenic mice decreases infarct volume following ischemia (Sheng *et al.* 1999; Sampei *et al.* 2000) and increases resistance to neurotoxin MPTP (Przedborski *et al.* 1992; Klivenyi *et al.* 1998). Over-expression of SOD in transgenic rabbits provides cardioprotection from myocardial infarction (Li *et al.* 2001). When designing therapeutics to over-express SOD, it is important to determine whether SOD's over-expression might have deleterious effects. It has been shown in transgenic models that over-expression of SOD, by as much as 30-fold in all tissues, not only resists ischemia

but the animals remains healthy (Przedborski *et al.* 1992; Chan *et al.* 1993; Chan and Kawase, 1998; Sheng *et al.* 1999b; Li *et al.* 2001). In addition, SOD has been found to have a potential role in deterring the aging process, thus over-expression of SOD in transgenic models shows multiple positive effects (Parkes *et al.* 1998).

They exist in different forms. They are mostly cofactored with copper, zinc, nickel, iron or manganese. Accordingly they can be classified into three major families.

Copper and zinc: Mostly eukaryotic SOD exists as Cu-Zn-enzyme, homodimer (32500 Da), linked by hydrogen and electrostatic interactions (Tainer *et al.* 1983).

Iron or manganese: Most prokaryotes and protists uses SOD cofactored with either iron or manganese. The active site of Mn and Fe superoxide dismutases contains the same type of amino acid side chains.

Nickel: This is also used by prokaryotes. Ni-SOD has a unique hexameric structure which provides critical interactions for metal binding and catalytic activity (Barondeau *et al.* 2004).

2.6.1 Human SOD

Human SOD can also be classified into three forms based on their location. SOD1, Fig. 12, is located in cytoplasm and exists mostly as a dimer. SOD2, Fig. 13, is found in mitochondria. SOD3 is extracellular. Both SOD2 and SOD3 always exist as tetramers. SOD1 and SOD3 are cofactored with copper and zinc while SOD2 found with manganese in their reaction centers.

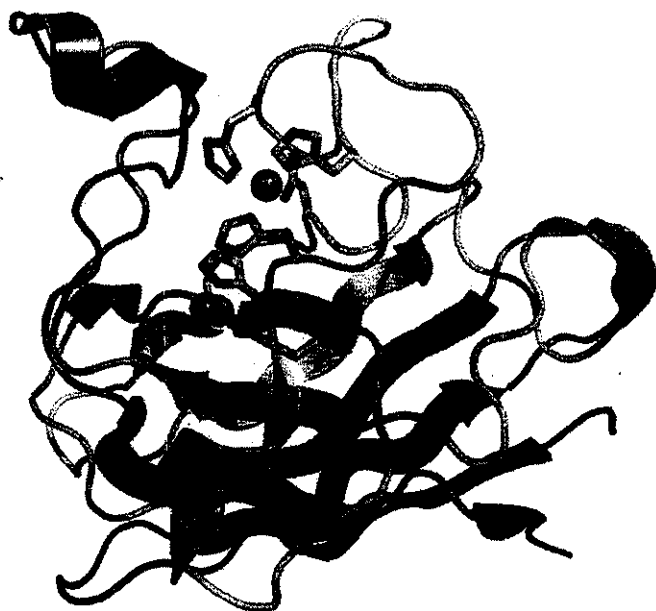


Figure 19. Crystallographic structure of the human SOD1 enzyme

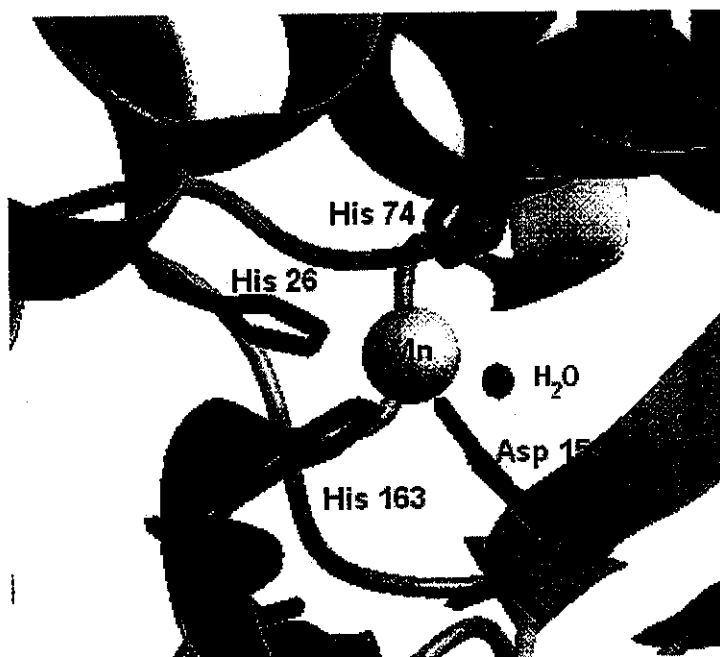
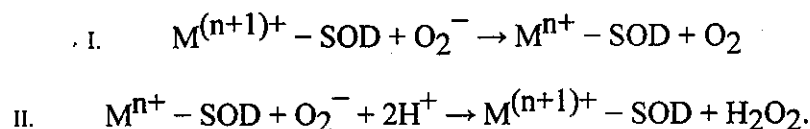


Figure 20. The active site of the active site of human superoxide dismutase 2.

2.6.2 Biochemistry

SOD protects a cell from damaging reactions of superoxide. The SOD-catalysed dismutation of superoxide may be written with the following half-reactions :



where M = Cu (n=1) ; Mn (n=2) ; Fe (n=2) ; Ni (n=2).

In this reaction the oxidation state of the metal cation oscillates between n and n+1.

In biological systems, this means its main reactions are with itself (dismutation) or with another biological radical such as nitric oxide (NO) or a metal. The superoxide anion radical (O_2^-) spontaneously dismutates to O_2 and hydrogen peroxide (H_2O_2) quite rapidly. SOD is essential because it can react with superoxide (which reacts with critical cellular targets). Superoxide reacts with NO and forms peroxynitrite. The half life of superoxide is very short for low concentrations but quite long for very low concentrations. SOD has the largest turnover number of any known enzyme.

2.6.3 Physiology

Each type of SOD plays a different role in keeping cells healthy. Physiological importance of SOD can be studied by genetically engineered mice which carry these enzymes. Deficiency of SOD2 results in massive oxidative stress in mice (Li *et al.* 1995). Mice lacking SOD1 result in severe pathological conditions, including hepatocellular carcinoma (Elchuri, 2005) and also in age related muscle loss (Muller *et al.* 2006). Loss of SOD3 in mice does not show any obvious effects (Sentman *et al.* 2006). Superoxide inducing drugs, like paraquat and diquat, can produce lethal effects in any SOD knockout mice. Similar studies in *Drosophila* and *C. elegans* show their lifespan is reduced if they lack SOD.

2.6.4 Role in disease

ALS (amyotrophic lateral sclerosis) can be caused by mutations in the SOD1 enzyme (AL-chalabi *et al.* 2000). Overexpression of SOD1 can also lead to the Down syndrome (Groner *et al.*, 1994). In human, the deficiency of SOD2 and SOD3 does not

show any lethal effects, but in mice absence of SOD2 results in perinatal lethality. SOD is also used for clinical purposes. Reactive oxygen species generation and oxidative stress can be reduced, which in turn inhibits endothelial activation. It can also be used as a new tool for the treatment of inflammatory bowel disease (Segui *et al*, 2004).

2.7 AIM AND OBJECTIVES

The aim of this study is to determine the crystal structures of

a) Cyp38 from *Arabidopsis thaliana* and

b) Superoxide dismutase (SOD) from *Spirulina platensis*

MATERIALS AND METHODS

CHAPTER 3. MATERIALS AND METHODS

3.1 CLONING OF SOD GENE

3.1.1 Primer design for PCR

The forward and reverse primers were designed for PCR of the SOD gene (coding residues 1-170) by incorporating the flanking restriction sites (EcoRI and XhoI) that was present in the multiple cloning site (MCS) of the vector that was used for cloning. Note that these restriction sites are not present in the insert.

3.1.2 Gradient PCR

Gradient PCR was first performed using an MJ Mini Gradient Thermal Cycler machine to optimize the annealing temperature across a broad range (55 °C - 65 °C) using a set of upstream and downstream oligonucleotide primers with respective restriction endonuclease sites flanking the primer ends.

A 50 µl PCR reaction mixture was prepared for amplifying the SOD gene containing 5 µl of 10X Thermopol buffer, 3 µl of 25 mM MgCl₂, 1 µl each of 10 µM upstream and downstream oligonucleotide primers, 0.5 µl each of dNTPs (10 mM) and Taq DNA polymerase (5000 U/ml) enzyme and 2 µl of cDNA as DNA template.

A gradient PCR program was designed to amplify SOD gene with a 30 cycles involved initial denaturation at 96 °C for 3 min, followed by denaturation at 96 °C for 30 secs, annealing (55 °C - 65 °C) for 30 secs, elongation at 72 °C for 45 secs, followed by a final elongation at 72 °C for 10 min.

3.1.3 PCR products purification

The PCR products were purified by using the QIAquick PCR Purification kit according to the manufacturer's protocol. DNA was eluted in 40-50 µl of elution buffer.

3.1.4 Cloning of SOD in pGEM-T-Easy cloning vector

Ligation of the target DNA after PCR and purification was first carried out into the pGEM®-T Easy vector system (Promega) in a 10 µl reaction. The DNA concentration of the PCR product was checked using a Nanovue (GE Healthcare)

spectrophotometer. The following equation was used to calculate the amount of PCR product to be used for the ligation reaction. An insert:vector molar ratio of 3:1 was taken for the ligation reaction according to the manufacturer's recommendations.

$$\frac{\text{ng of vector} \times \text{kb size of insert} \times \text{insert : vector molar ratio}}{\text{kb size of vector}} = \text{ng of insert}$$

1 μ l of the vector, 5 μ l of 2X ligation buffer and 1 μ l T4 DNA ligase (Promega) were used to set up a ligation reaction. The ligation mixtures were incubated at room temperature for 1 hour before being transformed into DH5 α competent cells.

3.1.5 Preparation of E. coli DH5 α competent cells

A frozen glycerol stock of DH5 α cells was used to streak onto a Luria Bertani (LB) agar plate which does not contain ampicillin (Sigma) and incubated overnight at 37 $^{\circ}$ C. A fresh single colony was picked and inoculated into 5 ml of LB broth and incubated in a shaker overnight at 37 $^{\circ}$ C. The overnight culture was transferred into 100 ml of LB broth and incubated in a shaker at 37 $^{\circ}$ C with revolution of 200 rpm. Cells were continued shaking until O.D. at 600 nm wavelength reached 0.4-0.6 and then transferred into 50 ml tubes and again centrifuged at 1,520 g at 4 $^{\circ}$ C for 10 min. The supernatant was discarded and the cell pellet was dissolved gently with 40 ml of cold glycerol buffer (0.1 M CaCl₂, 15% glycerol) and then placed on ice for 30 min. Cells were centrifuged at 1,520 g for another 10 min at 4 $^{\circ}$ C and the supernatant was removed. The pellet was dissolved completely in 4 ml of cold glycerol buffer, distributed at 50 μ l per tube, frozen in liquid nitrogen and stored at -80 $^{\circ}$ C.

3.1.6 Transformation into DH5 α Competent cells and blue-white screening

A 50 μ l aliquot of the previously prepared competent E. coli DH5 α cells was thawed on ice for 5 min. 2 μ l (concentration above 20ng) of the plasmid was transferred, mixed gently and placed on ice for 30 min. The cells were then heat shocked at 42 $^{\circ}$ C for 1.5min and immediately kept back on ice for 2 min. 750 μ l of LB medium (without any antibiotic) was added into the tube, the cells were incubated with shaking at 37 $^{\circ}$ C for an hour. Cells were harvested in a microfuge at 13000 rpm for 1 min and 600 μ l of the

supernatant was discarded. The cell pellet was resuspended with the remaining medium and plated on LB agar plate added with a suitable antibiotic for selection. Plates were then incubated overnight at 37 °C. Additionally, for blue / white colony screening of colonies with pGEM-T-Easy vector system, 10 µl of 1 M IPTG and 20 µl of X-gal solution (Biorad) were added to the LB Plates (containing appropriate antibiotic) before plating the cells on it.

3.1.7 Colony PCR

Colony PCR screening of transformants was used for immediate screening of ligated transformants. In this method, 7 well grown single colonies were selected from the transformant plate and mixed gently by pipeting up and down to resuspend them in 40 µl of sterilized water taken in 7 different tubes and this cell suspension was used as a template for doing PCR. A PCR mix was prepared in the same way to the mix prepared for target gene amplification, mentioned in section 2.2 and a PCR protocol unique to each gene target for amplification was used. The PCR products were analyzed on agarose gel to confirm the presence of amplified target inserts at the right gene fragment size.

3.1.8 Agarose gel electrophoresis

Agarose gel electrophoresis was done using 1% UltraPure™ agarose (Invitrogen) in 60-100 ml 1X Tris-acetate (TAE) running buffer, added with 2 µl of ethidium bromide (10 mg/ml) solution. Each DNA sample was first mixed with 6X loading dye (6:1 concentration) (Promega). 3 µl of 1 kb DNA ladder (Promega) was also similarly added with the dye and loaded alongside the sample wells, serving as a marker for estimation of molecular weight. After loading, the gel was run at 120 V for about an hour until the loading dye reached the bottom of the gel. DNA bands were then examined using the GeneSnap machine (SynGene BioImaging System) under an ultraviolet (UV) transilluminator.

3.1.9 Sequencing PCR

1 μ l of plasmid DNA containing the insert DNA was mixed with 7 μ l of water, 1 μ l of the respective vector specific sequencing primer (3.2 μ M), and 1 μ l of Big Dye Terminator V3.1 (Applied Biosystems). 25 cycles of PCR reaction were carried out with the denaturation step at 96 °C for 30 sec, annealing at 50 °C for 15 sec and extension at 60 °C for 4 min.

3.1.10 Ethanol precipitation

DNA was then precipitated by adding 1/10th volume of 3 M sodium acetate of pH 5.2 and mixed by vortexing briefly. Then, 50 μ l ice cold 100% ethanol was added to this. The mixture was mixed completely well and placed in a -80 °C freezer for 20 min followed by the tubes were spun at 13000 rpm for 20 min at 4 °C. The supernatant was removed and the pellet was washed with 500 μ l of 70% ethanol and spun at 13000 rpm for 20 min at 4 °C. The supernatant was carefully aspirated and the sample was dried before using for sequencing.

3.2 EXPRESSION AND PURIFICATION OF MATURED Cyp38

3.2.1 Affinity chromatographic purification of matured Cyp38

Already expressed matured GST: Cyp38 (residues 100-437) fusion protein cell pellet was suspended in ice cold lysis buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF and subjected to sonication. The crude lysate was centrifuged at 42,400 g for 45 min at 4 °C and the supernatant was collected. SDS-PAGE analysis was carried out to confirm the presence of the soluble protein. The crude supernatant from the previous step was applied to a GST-affinity column (Glutathione 4 Fast Flow; Amersham Biosciences) and the fusion protein was allowed to bind to the resin overnight at 4 °C. The contaminant proteins that were loosely bound to the affinity column were removed by giving washes with high salt wash buffer [20 mM Tris-HCl (pH 7.5), 900 mM NaCl, 1 mM DTT, 1 mM EDTA] and low salt wash buffer [20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM DTT, and 1 mM EDTA]. SDS-PAGE of the various samples from the process was carried out to confirm the success of the purification step.

3.2.2 Size exclusion chromatography

The GST fusion protein was further purified by size exclusion chromatography on a pre-equilibrated HiLoad 16/60 Superdex-200 column (Amersham Biosciences) using the low salt wash buffer with 20mM TRIS, 150 mM NaCl, and 5% Glycerol. Fractions corresponding to the protein peak were analyzed by SDS-PAGE and pure fractions were pooled together and concentrated.

3.3 CRYSTALLIZATION

3.3.1 Hanging drop vapor diffusion method

Crystallization trials are being setup using the hanging drop vapor diffusion method at 20 °C with Hampton and other commercial screens.

RESULTS AND DISCUSSION

CHAPTER 4. RESULTS AND DISCUSSION

4.1 CLONING OF SOD GENE

4.1.1 Gradient PCR:

PCR conditions were optimized for the upstream and downstream primers in the temperature range 55-65 °C. We found good amplification at 56.1 °C, Fig. 14.

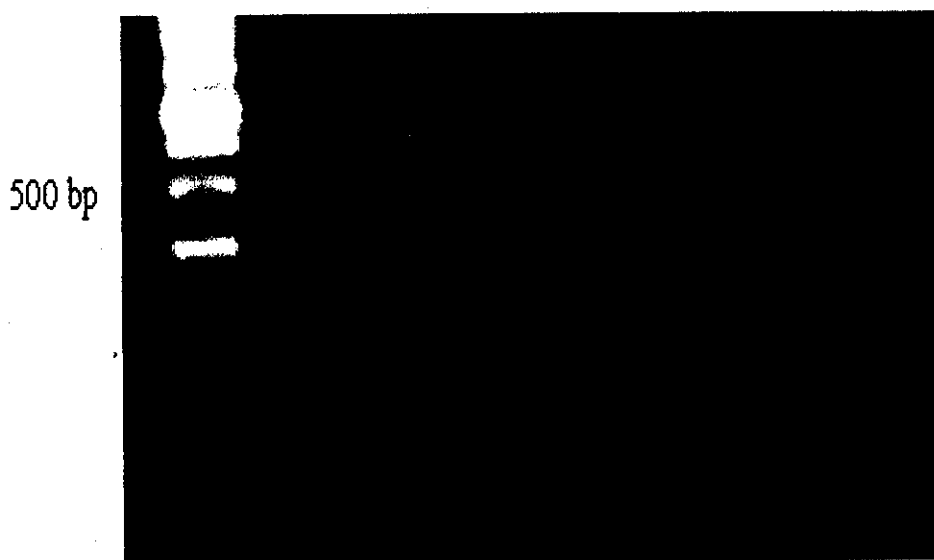


Figure 21. Agarose gel electrophoresis showing Gradient PCR results. Lane 1 – 1 kb DNA ladder. Lanes 2, 3, 4, 5, 6 correspond to temperatures 59, 58.1, 57, 56.1 and 55 respectively.

4.1.2 Transformation of Ligated SOD-pGEM-T-Easy vector into DH5 α competent cells and Colony PCR

The amplified PCR products were purified with using the QIAquick PCR Purification kit and the concentration was 106 ng, when measured using a Nanovue (GE Healthcare) spectrophotometer. Then this purified DNA was ligated overnight in to pGEM®-T Easy vector system. The ligated SOD-pGEM-T-Easy vector was then transformed to DH5 α competent cells. Seven well separated distinct colonies (Blue/White selection) were selected for colony PCR and among them, six were found as transformed, Fig. 15.

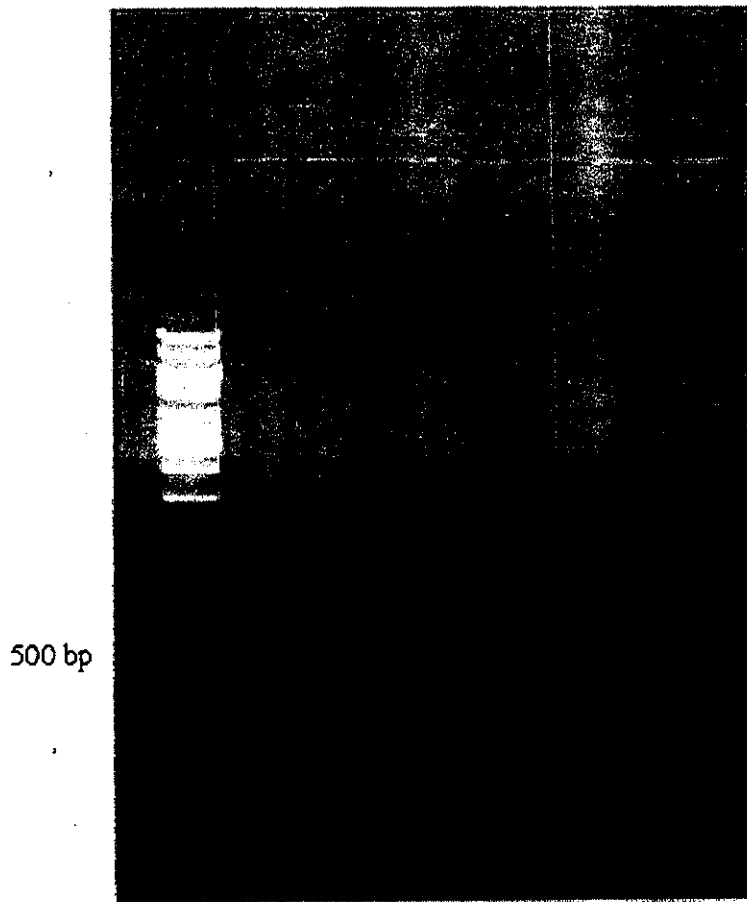


Figure 22. Agarose Gel electrophoresis showing colony PCR results. Lane 1 – 1 kb DNA ladder, Lanes 2 – 8 correspond to colonies 1 – 7.

4.1.3 DNA Sequencing

Samples 5 and 6 were selected for plasmid DNA sequencing. These samples were subjected to sequencing PCR followed by ethanol precipitation. Then the samples were given for DNA sequencing. The results from DNA sequencing is blasted (using Nucleotide BLAST, NCBI online software) against the Fe-SOD sequence (*Spirulina platensis*) retrieved from UNIPROT. The gene sequence was found be mutated, Appendix 1. So we are not able to proceed further with this gene.

4.2 EXPRESSION AND PURIFICATION OF MATURED CYP38

4.2.1 Affinity chromatographic purification

The overexpressed truncated GST:Cyp38 (100-437) fusion protein was purified using glutathione-sepharose beads (Amersham Biosciences), and results shown in Fig. 16. Due to the slow kinetics of GST binding, considerable amount of the fusion protein escaped in the flow through. Hence the flow through from the first step was subjected to a second round of purification process to retrieve the unbound fusion protein.

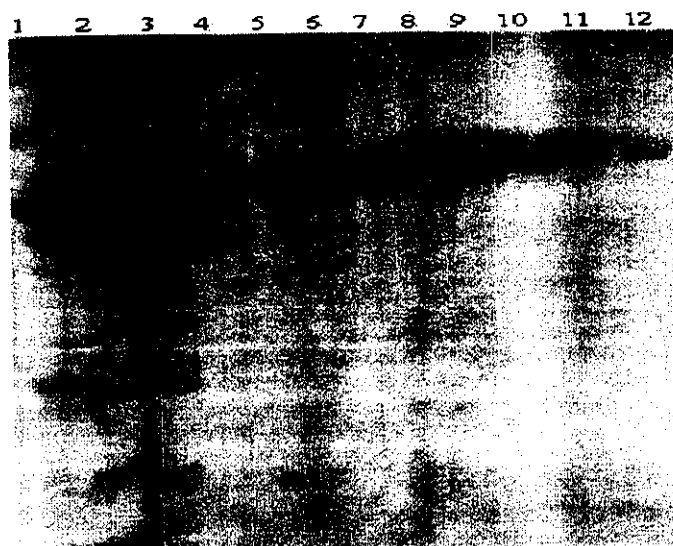


Figure 23. SDS-PAGE showing affinity chromatographic purification of the matured Cyp38-GST fusion protein. Lane 1 - Low molecular weight marker, Lane 2 - Pellet, Lane 3 - Flow through, Lane 4 - Wash 1, Lane 5 - Wash 2, Lanes 6, 7, 8, 9 correspond to Elutions 1, 2, 3, 4 respectively, Lane 10 - Wash 3, Lanes 11, 12 correspond to Elutions 5 and 6

4.2.2 Size exclusion chromatography

HiLoad 16/60 Superdex-200 column (Amersham Biosciences) was used for the size exclusion chromatographic experiment. Pure truncated Cyp38 protein was eluted out as a single peak, Fig.17, and the corresponding fractions were pooled together. SDS-PAGE of the pooled fractions, Fig.18, had no other contaminant proteins. The integrity of the protein was verified by dynamic light scattering (DLS), Fig.19. The polydispersity

index is 0.10 and the SOS error is 4.48. This shows the protein is monodisperse and suitable for crystallization.

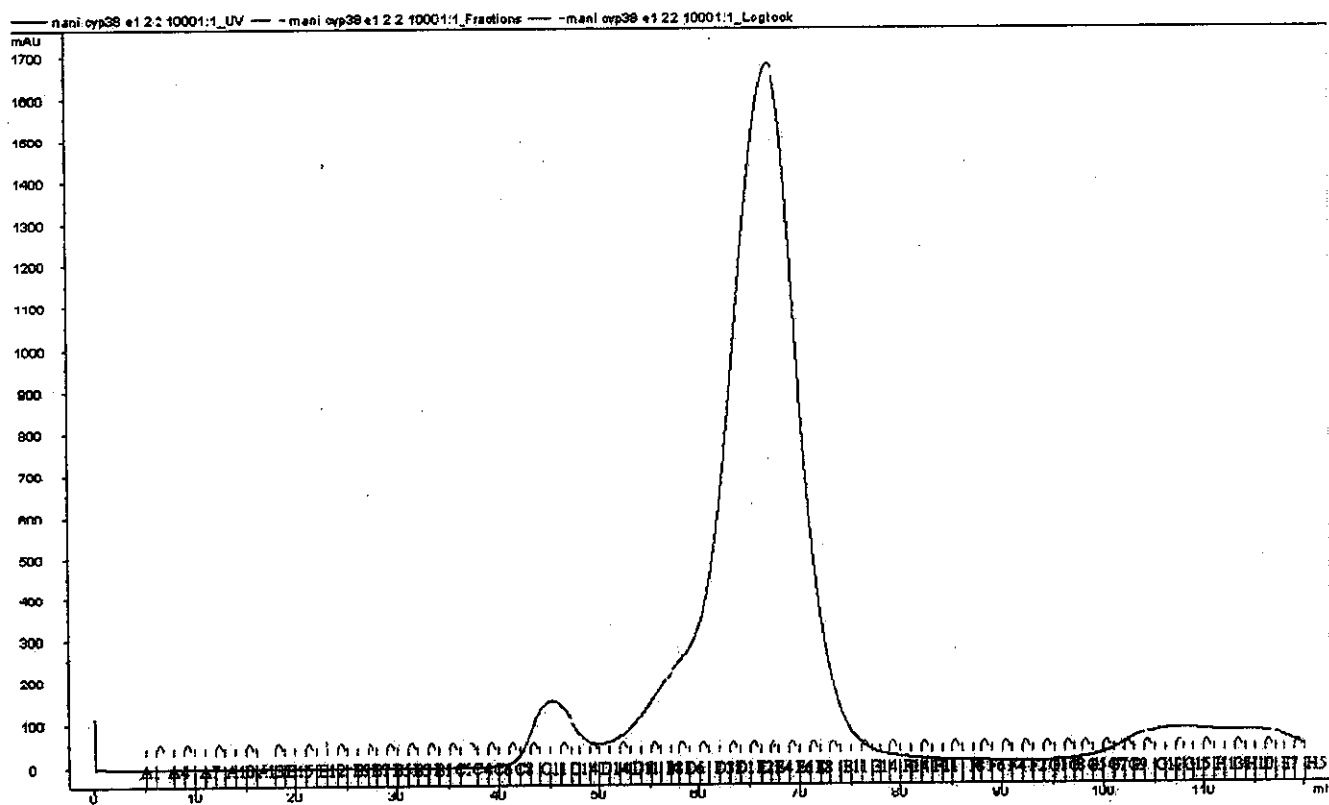


Figure 24. Profile of size-exclusion chromatographic purification of matured Cyp38-GST fusion protein. The higher peak around 70ml corresponds to the protein

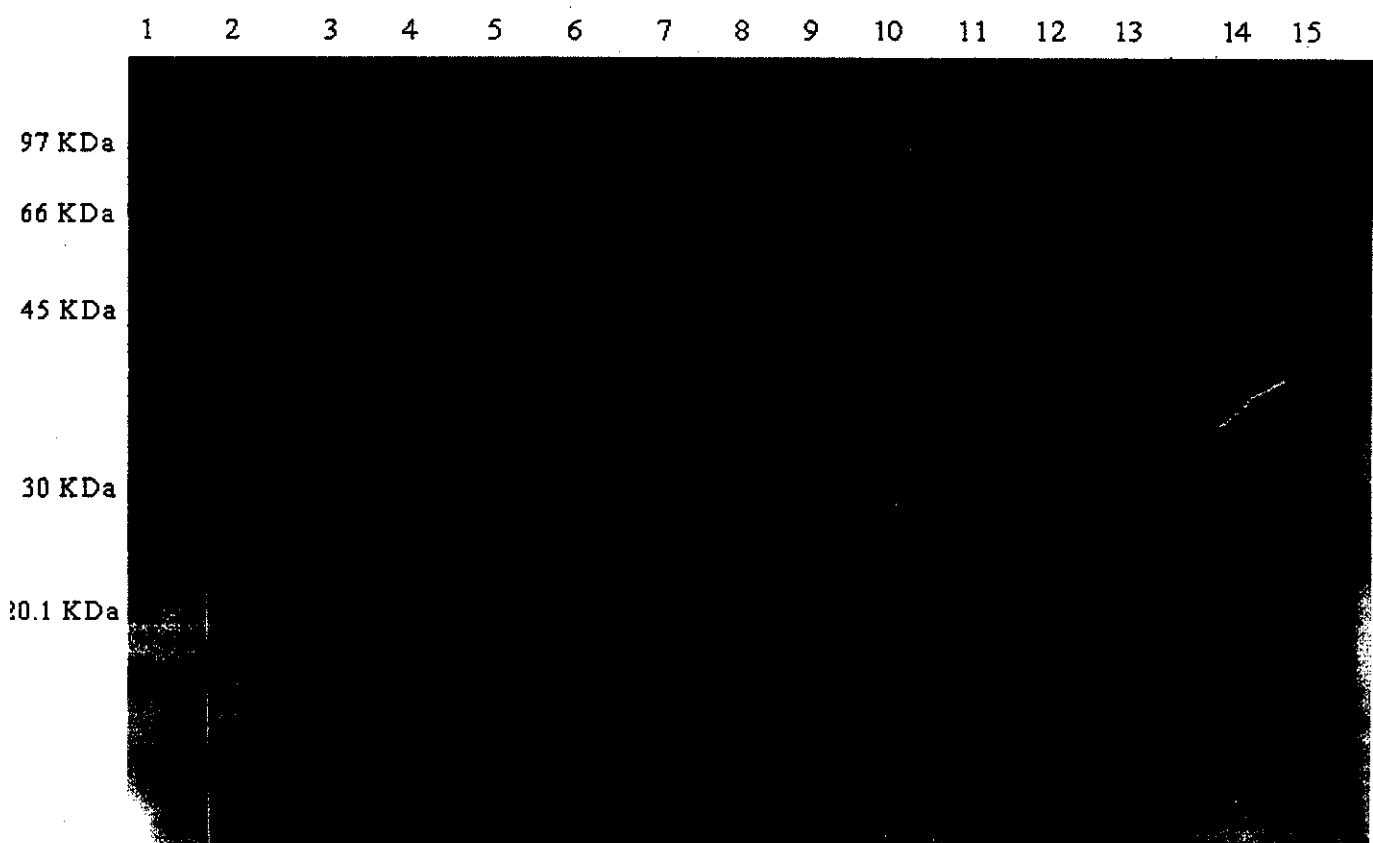


Figure 25. SDS-PAGE of size-exclusion chromatographic fractions of Cyp38-GST fusion protein. Lane 1 – Low molecular weight marker, Lanes 2-15 correspond to fractions E7 to F11.

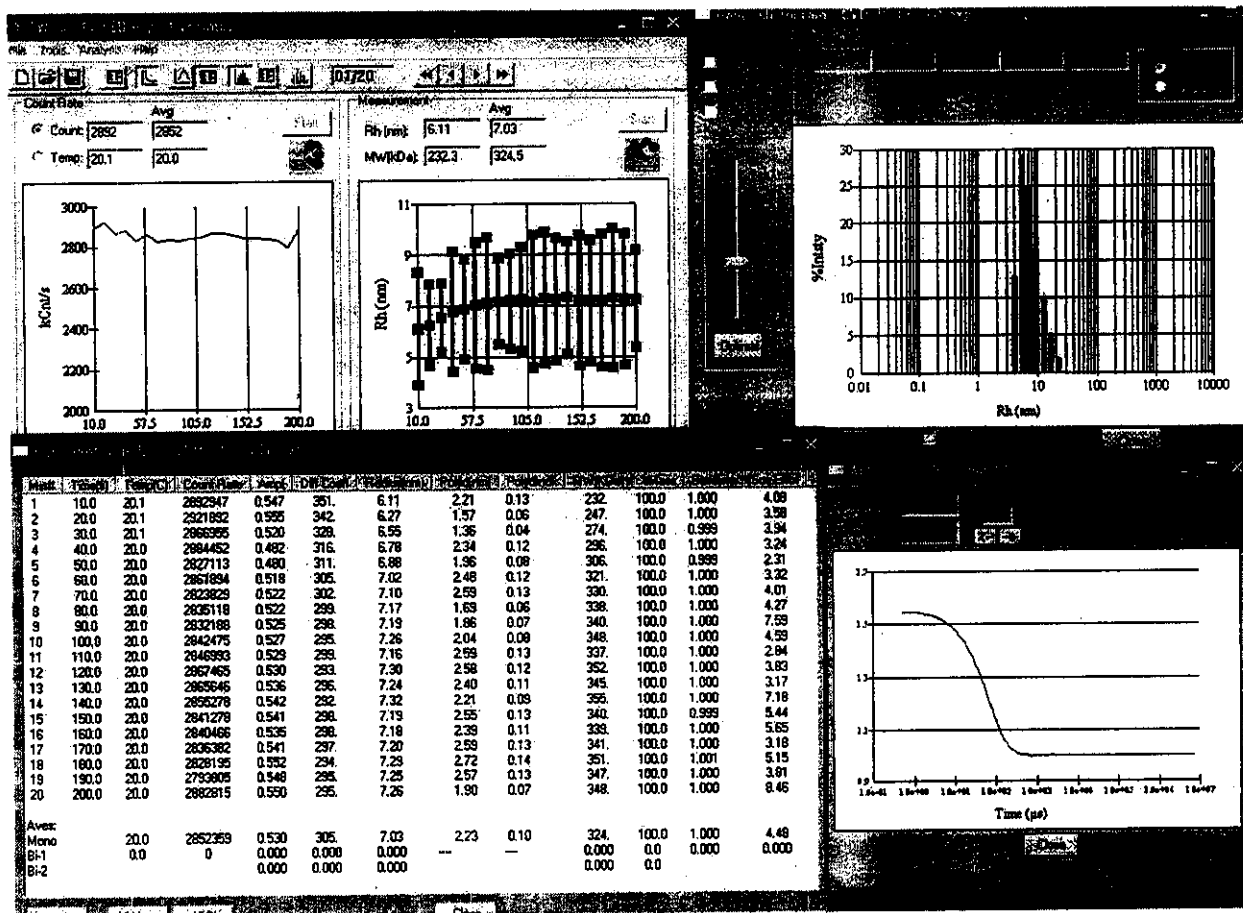


Figure 26. DLS profile of size-exclusion chromatographic fractions of matured Cyp38-GST fusion protein.

4.3 CRYSTALLIZATION

The matured Cyp38 (100:437) protein aggregates when the GST tag is cleaved. So we have decided to crystallize the protein with the tag. The matured GST: Cyp38 (100-437) protein is being setup for crystallization with Hampton research crystallization commercial kits and other crystallization screens like flexible sparse matrix screens (Zeelen, 1999) using the hanging drop method.

4.4 CONCLUSIONS

From the present study,

The matured Cyp38 (100:437) protein aggregates when the GST tag is cleaved. But the expression of GST tagged Cyp38 fusion protein is good. Affinity chromatography followed by Size exclusion chromatography purifies the fusion protein to homogenous. SDS PAGE analysis reveals the molecular weight of this matured fusion protein, is 66 KDa. 0.10 monodispersity index and 4.48 SOS error shows that the protein homogenous and monodisperse and is suitable for crystallization.

Genes coding 1 – 170 aminoacids of Fe SOD isolated from *Spirulina platensis* is found to be mutated.

4.5 FUTURE WORK

When a crystal of suitable size is obtained, we will determine the crystal structure using the molecular replacement method. As our lab had already solved the crystal structure of non-functional Cyp38 (83-437), the present work will shed light on the conformational changes of the Cyp38 protein upon maturation.

APPENDIX

APPENDIX

1. SOD sequencing results

>5R_F03_011.ab1 (with Reverse sequencing primers)

NNNNNNNNNNNNNNNNNNNNNTGGGAGCTCTCCCATATGGTCGNNNNNNNNGCGGCCGCGA
ATTCACTAGTGNNNNNNNNNNNNNNNGNNNTTNGAACTTCCCAGTTTACCTTTTGANNNNN
NNNANNANAGTCTAGCAAATGTCCGCTAATACTCTATCCTACCATCATGGTAAGCACC
ATGCAGCTTATGTCAAAAACCTTAATGCTGCTATTGAAGGTACAGACATGGCTAATATG
TCCCTGGAAGAGATCATCAAAGCCACTTACAATGATCCTTCTAAGTCTGGTATATTTAA
TAACGCTGCCAAGTCTGGAATCACAGTTTCTTCTGGAAGTGCCCTCAAGCCCAATGGTG
ATGGTCAACCTACAGGAGCTTTGGCTGACAAGATCCAAGCAGATTCGGTAGCTTTGAT
GCATTCATTCAAGAATTTAAGAACGCAGCGGCTACCCAGTTCGGTAGTGGTTGGGNTTG
GNNAGTCCTGGATAATGGCACACTGAAAGTTACCAAGACTGCTAATGCTGTTAACCCCA
TGGTTGAGGGTAAAACTCCCCTGCTAACTCTCGAGGTTTGGGAAGACGCCTACTACAAT
TGACTCGAGAAGAAAATCGAATTCGCGCGGCCGCGCATGGCGGCCGGGAGCATGCGACGT
CGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCCTGGCCGTGTTTTACAAC
GTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCT
TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCG
CAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGNGTGGT
GGTTACGCGCAGCGTGANCGCTACACTTGNCAGCGCCNTANCGCCCGCTCCTTTTCGCTT
TCTTCCCNTCCTTTCTCGNCACNNCGNCGGCTTTCCCGNCAAGCTCTAATCGGGGGGC
TCCNTTNNGGGTNCNATTTNNNGNTTTNNNGCNCCTCNACCNNAAAAANNTGANNANG
GNNANGGNNCNCGNANNGGNNANCNNNCNNANANANNNNNTCNNNNNNTNGNNNNNNNNN
ANNNCNCNNTNNNNNNNNNNNNNNNNNNNCNNNNNNNNNNNNNNNANNANNNNNN

2. pGEM:SOD plasmid DNA sequencing

BLASTN 2.2.23+

Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb Miller (2000), "A greedy algorithm for aligning DNA sequences", J Comput Biol 2000; 7(1-2):203-14.

RID: UKWDW6BU11R

Query=Length=1172

Sequences producing significant alignments: (Bits) Score E Value
815 0.0

ALIGNMENTS

>lcl|11821

Length=511

Score = .815 bits (441), Expect = 0.0
Identities = 480/505 (95%), Gaps = 6/505 (1%)
Strand=Plus/Plus

```
Query 91   GAACTTCCCAGTTTACCTTTTga-nnnnnnannanaGTCTAGCBAATGTCGGCTAAT 149
          |||||                               | | |||||
Sbjct 10   GAACTTCCCAGTTTACCTTTGATCAGSATGCCTAGAGTCTAGCBAATGTCGGCTAAT 69

Query 150  ACTCTATCCTACCATCATGGTAAGCACCATGCAGCTTATGTCAAAAACCTTAATGCTGCT 209
          |||||
Sbjct 70   ACTCTATCCTACCATCATGGTAAGCACCATGCAGCTTATGTCAAAAACCTTAATGCTGCT 129

Query 210  ATTGAAGGTACAGCATGGCTAATATGTCCTGGAAAGAGATCATCAAAGCCACTTACAAT 269
          |||||
Sbjct 130  ATTGAAGGTACAGCATGGCTGATATGTCCTGGAAAGAGATCATCAAAGCCACTTACAAT 189

Query 270  GATCCTTCTAAGTCTGGTATATTTAATAACCGCTGCCAAGTCTGGAATCACAGTTTCTTC 329
          |||||
Sbjct 190  GATCCTTCTAAGTCTGGTATATTTAATAACCGCTGCCAAGTCTGGAATCACAGTTTCTTC 249

Query 330  TGGAAAGTGCCTCAAGCCCAATGGTGGTCAACCTACAGGAGCTTTGGCTGACAAGATC 389
          |||||
Sbjct 250  TGGAAAGTGCCTCAAGCCCAATGGTGGTGGTCAACCTACAGGAGCTTTGGCTGACAAGATC 309

Query 390  CAAGCAGATTTGGTAGCTTTGATGCATTATTCAAGAAATTAAGAAACGACGGCTACC 449
          |||||
Sbjct 310  CAAGCAGATTTGGTAGCTTTGATGCATTATTCAAGAAATTAAGAAACGACGGCTACC 369

Query 450  CAGTTCGGTAGTGGTTGGGNTTGGNNACTCCTGGATAATGGCACACTGAAAGTTACCAAG 509
          |||||
Sbjct 370  CAGTTCGGTAGTGGTTGGGCTTGGCTAGTCTGGATAATGGCACACTGAAAGTTACCAAG 429

Query 510  ACTGCTAATGCTGTAAACCCCATGG-TTGAGGGTAAACTCCCCTGCTAACTCTCGA-GG 567
          |||||
Sbjct 430  ACTGCTAATGCTGTAAACCC-ATGGGTTGAGGGTAAACTTCCCTGCTAAC-CTCGATGG 487

Query 568  TTTGGGAAGACGCCTACTACAATTG 592
          |||||
Sbjct 488  TTTGGGAAGACGCCTACTACA-TTG 511
```

REFERENCES

REFERENCES

1. Vasudevan, D. Gopalan, G. He, Z. Luan, S. Swaminathan, K. (2005). Expression, purification, crystallization and preliminary X-ray diffraction analysis of *Arabidopsis thaliana* cyclophilin 38 (AtCyp38). *ACTA Cryst. F* 61, 1087-1089.
2. Vasudevan, D. Radhakrishnan, A. Luan, S. Swaminathan, K. (2007). Crystal structure of an intermediate form of *Arabidopsis thaliana* cyclophilin 38 (AtCyp38) (to be submitted).
3. Andreeva, L. Heads, R. Green, C.J. (1999). Cyclophilins and their possible role in the stress response. *Int. J. Exp. Pathol.* 80, 305-315.
4. Borel, JF. Feurer, C. Gubler, HU. Stahelin, H. (1976). Biological effects of cyclosporin A: a new antilymphocytic agent. *Agents Actions.* 6, 468-475.
5. Borel, JF. (1990). Pharmacology of cyclosporine (sandimmune). IV. Pharmacological properties in vivo. *Pharmacol. Rev.* 41, 259-371.
6. Dornan, J. Taylor, P. Walkinshaw, MD. (2003). Structures of immunophilins and their ligand complexes. *Curr. Top. Med. Chem.* 3, 1392-1409.
7. Fischer, G. Bang, H. Mech, C. (1984). Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides. *Biomed Biochim. Acta.* 43, 1101-1111.
8. Galat A. (2003). Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity-targets-functions. *Curr. Top. Med. Chem.* 55, 423-436.
9. Gasser, CS. Gunning, DA. Budelier, KA. Brown, SM. (1990). Structure and expression of cytosolic cyclophilin peptidyl-prolyl cis-trans isomerase of higher-plants and production of active tomato cyclophilin in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 87, 9519-9523.

10. He, Z. Li, L. Luan, S. (2004). Immunophilins and parvulins. Superfamily of peptidyl prolyl isomerases in Arabidopsis. *Plant Physiol.* 134, 1248-1267.
11. Huai, Q. Kim, HY. Liu, Y. Zhao, Y. Mondragon, A. Liu, JO. (2002). Crystal structure of calcineurin-cyclophilin-cyclosporin shows common but distinct recognition of immunophilin-drug complexes. *Proc. Natl. Acad. Sci. USA.* 99, 12037-12042.
12. Ivery, MT. (2000). Immunophilins: switched on protein binding domains? *Med. Res. Rev.* 230, 452-484.
13. Ke, HM. Huai, Q. (2004). Crystal structures of cyclophilins and its partners. *Front. Biosci.* 9, 2285-2296.
14. Liu, J. Chen, CM. Walsh, CT. (1991). Human and Escherichia coli cyclophilins: sensitivity to inhibition by the immunosuppressant cyclosporin A correlates with a specific tryptophan residue. *Biochemistry.* 30, 2306-2310.
15. Liu, J. Farmer, JD. Lane, WS. Friedman, J. Weissman, I. Schreiber, SL. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell.* 66, 807-815.
16. Luan, S. Albers, MW. Schreiber, SL. (1994). Light-regulated, tissue-specific immunophilins in a higher plant. *Proc. Natl. Acad. Sci. USA.* 91, 984-988.
17. Kino, T. Hatanaka, H. Hashimoto, M. Nishiyama, M. Goto, T. Okuhara, M. (1987). FK-506, a novel immunosuppressant isolated from a Streptomyces. I. Fermentation, isolation, and physico-chemical and biological characteristics. *J. Antibiot. (Tokyo).* 40, 1249-1255.
18. Martel, RR. Klicius, J. Galet, S. (1977). Inhibition of the immune response by rapamycin, a new antifungal antibiotic. *Can. J. Physiol. Pharmacol.* 55, 48-51.
19. Romano, PGN. Edvardsson, A. Ruban, AV. Andersson, B. Vener, AV. Gray, JE. (2004). Arabidopsis AtCyp20 2 is a light-regulated cyclophilin-type peptidyl- prolyl

- cis-trans isomerase associated with the photosynthetic membranes. *Plant Physiol.* 134, 1244-1247.
20. Romano, PGN. Horton, P. Gray, JE. (2004). The Arabidopsis cyclophilin gene family. *Plant Physiol.* 134, 1268-1282.
 21. Sehgal, SN. Baker, H. Vezina, C. (1975). Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. *J. Antibiot. (Tokyo)*. 28, 727-732.
 22. Schreiber, SL. (1991). Chemistry and biology of the immunophilins and their immunosuppressive ligands. *Science*. 251, 283-287.
 23. Schlatter, D. Thoma, R. Kung, E. Stihle, M. Muller, F. Borroni, E. (2005). Crystal engineering yields crystals of cyclophilin D diffracting to 1.7 Å resolution. *Acta Crystallogr. D* 61, 513-519.
 24. Taylor, P. Husi, H. Kontopidis, G. Walkinshaw, MD. (1997). Structures of cyclophilin ligand complexes. *Prog. Biophys. Mol. Biol.* 67, 155-181.
 25. Breiman, A. Fawcett, TW. Ghirardi, ML. Mattoo, AK. (1992). Plant organelles contain distinct peptidylprolyl cis,trans-isomerases. *J. Biol. Chem.* 267, 21293-21296.
 26. McCord, JM. Fridovich, I. (1988). Superoxide dismutase: the first twenty years (1968-1988). *Free Radic. Biol. Med.* 5, 363-9.
 27. Muller, FL. Song, W. Liu, Y. Chaudhuri, A. Pieke-Dahl, S. Strong, R. Huang, TT. Epstein, CJ. Roberts, LJ. Csete, M. Faulkner, JA. Van Remmen, H. (2006). Absence of CuZn superoxide dismutase leads to elevated oxidative stress and acceleration of age-dependent skeletal muscle atrophy. *Free Radic. Biol. Med.* 40, 1993-2004
 28. Elchuri, S. Oberley, TD. Qi, W. Eisenstein, RS. Jackson Roberts, L. Van Remmen, H. Epstein, CJ. Huang, TT. (2005). CuZnSOD deficiency leads to persistent and widespread oxidative damage and hepatocarcinogenesis later in life. *Oncogene* 24 ,

29. Li, Y. Huang, TT. Carlson, EJ. Melov, S. Ursell, PC. Olson, JL. Noble, LJ. Yoshimura, MP. Berger, C. Chan, PH. Wallace, DC. Epstein, CJ. (1995). Dilated cardiomyopathy and neonatal lethality in mutant mice lacking manganese superoxide dismutase. *Nat. Genet.* 11, 376–81
30. Biomolecular crystallography: principles, practice, and application to structural biology .Bernhard Rupp (2010).
31. Al-Chalabi, A. Leigh, PN. (2000). Recent advances in amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* 13, 397–405.
32. Sentman, ML. Granström, M. Jakobson, H. Reaume, A. Basu, S. Marklund, SL, (2006). Phenotypes of mice lacking extracellular superoxide dismutase and copper- and zinc-containing superoxide dismutase. *J. Biol. Chem.* 281, 6904–9
33. Seguí, J. Gironella, M. Sans, M. Granell, S. Gil, F. Gimeno, M. Coronel, P. Piqué, JM. Panés, J. (2004). Superoxide dismutase ameliorates TNBS-induced colitis by reducing oxidative stress, adhesion molecule expression, and leukocyte recruitment into the inflamed intestine". *J. Leukoc. Biol.* 76, 537–44.
34. Groner, Y. Elroy-Stein, O. Avraham, KB. Schickler, M. Knobler, H. Minc-Golomb, D. Bar-Peled, O. Yarom, R. Rotshenker, S. (1994). Cell damage by excess CuZnSOD and Down's syndrome". *Biomed. Pharmacother.* 48, 231–40.
35. Allen, RG. Tresini, M. (2000). Oxidative stress and gene regulation. *Free Radical Biology and Medicine.* 28, 463-499.
36. Chan, PH. Kinouchi, C. Epstein, E. Carlson, S. Chen, S. (1993). Role of Superoxide dismutase in ischemic brain injury: reduction of edema and infarction in transgenic mice following focal cerebral ischemia. *Progress in Brain Research.* 96, 97-104.
37. Chan, PH. and Kawase, M. (1998). Overexpression of SOD1 in transgenic rats

- protects vulnerable neurons against ischemic damage after cerebral ischemia and reperfusion. *Journal of Neuroscience*. 20, 8292-8299.
38. Li, Q. Bolli, R. Qiu, Y. Tang, XL. Guo, Y. French, BA. (2001). Gene therapy with extracellular superoxide dismutase protects conscious rabbits against myocardial infarction. *Circulation*. 103, 1893-1898.
39. Klivenyi, P. St Clair, D. Wermer, M. Yen, HC. Oberley, T. Yang, L. Flint Beal, M. (1998). Manganese superoxide dismutase overexpression attenuates MPTP toxicity. *Neurobiology Disorders*. 4, 253-258.
40. Parkes, TL. Elia, AJ. Dickinson, D. Hilliker, AJ. Phillips, JP. Boulianne, GL. (1998). Extension of *Drosophila* lifespan by overexpression of human SOD1 in motoneurons. *Nature Genetics*. 19, 171-174.
41. Przedborski, S. Kostic, V. Jackson-Lewis, V. Naini, AB. Simonetti, S. Fahn, S. Carlson, E. Epstein, CJ. Cadet, JL. (1992). Transgenic mice with increased Cu/Zn superoxide dismutase activity are resistant to N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced neurotoxicity. *Journal of Neuroscience*. 12, 1658-1667.
42. Sampei, K. Mandir, AS. Asano, Y. Wong, PC. Traystman, RJ. Dawson, VL. Dawson, TM. Hum, PD. (2000). Stroke outcome in double-mutant antioxidant transgenic mice. *Stroke*. 31, 2685-2691.
43. Sun, AY. and Chen, Y. (1998). Oxidative stress and neurodegenerative disorders. *Journal of Biomedical Science*. 5, 401-414.
44. Sheng, H. Bart, RD. Oury, TD. Pearlstein, RD. Crapo, JD. Warner, DS. (1999a). Mice overexpressing extracellular superoxide dismutase have increased resistance to focal cerebral ischemia. *Neuroscience*. 88, 185-191.
45. Sheng, H. Brady, T. Pearlstein, R. Crapo, J. Warner, D. (1999b). Extracellular superoxide dismutase deficiency worsens outcome from fetal focal cerebral ischemia in the mouse. *Neuroscience Letters*. 267, 13-16.

46. Venarucci, D. Venarucci, V. Vallese, A. Battila, L. Casado, A. De la Torre, R. Lopez-fernandez, ME. (1999). Free radicals: important cause of pathologies related to ageing. *Panminerva Medicine*. 4, 335-339.
47. Barondeau, DP. Kassmann, CJ. Bruns, CK. Tainer, JA. Getzoff, ED. (2004). Ni-superoxide dismutase structure and mechanism. *Biochemistry*. 25, 8038-8047.