Crystallographic structure analysis of selected pharmaceutically important plant compounds and protein complexes

Dissertation

Zur Erlangung des Doktorgrades
des Department Chemie
der Universität Hamburg

Vorgelegt von
Ahmed Akrem
aus Lahore, Pakistan

Hamburg, 2010

Die vorliegende Arbeit wurde im Zeitraum von Oktober 2006 bis September 2010 in der Arbeitsgruppe von Prof. Ch. Betzel am Institut für Biochemie und Molekularbiologie am Department Chemie der Universität Hamburg durchgeführt.

1. Gutachter Prof. Ch. Betzel
2. Gutachter Prof. U. Hahn
3. Tag der Disputation: 11-02-2011

To my wife and daughters

Abstract

Protective plant proteins specifically induced in pathological or related situations have been intensively studied from a biotechnological perspective. Pathogenesis related proteins can act as a constitutive defense tool against microbial pathogens and invertebrate pests. Structural and Functional Analysis of seed or bulb proteins have exploited their scope towards industrial and pharmaceutical applications. In this context, a pharmaceutically important ribosome-inhibiting Lectin protein and biotechnological important pesticide chitinase were thoroughly studied. The X-ray structure of a tetrameric lectin (CVL) from the spring *Crocus* plant (*Crocus vernus*) was determined and refined to an final *R*-factor of 17.3% and R_{free} of 22.5%, respectively. The protein was crystallized by the hanging drop method using 4 M Sodium formate as a precipitant. Crystals diffracted to 1.94 Å resolution having unit-cell parameters of a = 48.2 b = 98.0 and c = 105.9 Å and the space group $P2_12_12$. There were two dimers in the asymmetric unit. The overall folding of the lectin is similar to that observed for snowdrop lectin. The tetramer is composed of two homo dimers with chains A and B as one dimer formed by chains C and D. Stable dimers have been particular formed by monomers through the exchange of C-terminal strands interactions. The two chains of one dimer share a primary amino acid sequence homology of approx. 38 %. The protein is rich in Val, Asn, Gly and Leu and contains eight Cys residues. A single intra-chain disulfide bridge has been found in each of the individual chains of the tetramer. A total of 339 bound water molecules, three glycerol molecules and eight formate ($HCOO^-$) molecules, which could be located in the refined 2 F_o-F_c electron density map, were included in the final model. The amino acid sequence showed marginal sequence homology of approx. around 40% with already deposited mannose specific lectin structures in the protein data bank. The coordinate information for CVL has already been deposited in the protein data bank (PDB) with code 3mez. A superposition of with those mannose specific lectins with CVL showed particular highly conserved structural motifs of anti-parallel β-sheets. The structure of CVL has been solved by Multiple Isomorphous Replacement (MIR), using Platinum and Mercury metal derivatives.

Further, the *C. vernus* chitinase (CVC) was purified and crystallized. Crystals belonged to monoclinic, space group C2, with unit cell parameters a = 171.81, b = 36.55, c = 125.28 Å and β = 126.9°. The crystals diffracted to 2.1 Å resolution applying synchrotron radiation and the X-ray structure was refined to initial *R*-factor of 30% and R_{free} of 37%, respectively.

Inhaltsangabe

Die Struktur-Funktions-Beziehungen ausgewählter pflanzlicher Abwehrproteine die unter pathologischen oder pathologisch verwandten Bedingungen induziert und aktiviert werden, wurden im Rahmen der hier zusammengefassten Promotionsarbeit im Hinblick auf biotechnologische und pharmazeutische Anwendungen untersucht. Die ausgewählten Proteine werden von Pflanzen als konstitutives Verteidigungsmittel gegen mikrobielle Krankheitserreger und wirbellose Schädlinge eingesetzt.. In diesem Kontext wurde ein pharmazeutisch bedeutsames ribosominhibierendes Lektin sowie ein biotecnologisch bedeutschames, als Pestizid wirkendes Enyzm aus der Familie der Chitinasen besonders intensiv analysiert. Die Röntgenstruktur des tetrameren Lektins (CVL) aus dem Frühlingscrocus (*Crocus vernus*) wurde aufgeklärt und zu einem finalen R-Faktor von 17,3% und einem R_{free}-Wert von 22,5% verfeinert. Vorausgehend wurde das Protein mittels der "hanging drop" Methode kristallisiert, wobei 4 M Natriumformiat als Präzipitant diente. Die Kristalle diffraktierten zu einer Auflösung von 1,94 A. Die Raumgruppe der Kristalle ergab sich zu $P2_12_12_1$ und Zellparametern von a = 48.2 b = 98.0 and c = 105.9 Å mit jeweils zwei Dimeren in der asymmetrischen Einheit. Die Faltung dieses Lektins ist derjenigen des sogenannten Schneetropfen-Lektins ähnlich. Ein Tetramer setzt sich aus jeweils zwei Homodimeren mit den Ketten A und B und den Ketten C und D zusammen. Die Dimere werden insbesondere über Wechselwirkungen zwischen den C-terminalen Strängen stabilisiert. Die beiden Ketten eines Dimers teilen eine primäre Aminosäuresequenzhomologie von ca. 38%. Das Protein ist reich an Val, Asn, Gly und Leu und enthält acht Cys-Reste. Eine intramolekulare Disulfidbrücke findet sich in jeder Kette eines Tetramers. Zusätzlich wurden 339 gebundene Wassermoleküle, drei Glycerinmoleküle und acht Formiationen in der verfeinerten 2 *Fo-Fc* Elektronendichtekarte gefunden und in das endgültige Modell aufgenommen. Die Aminosäuresequenz des CVL zeigt eine geringe Sequenzhomologie von ca. 40% mit bereits in der Protein Data Bank hinterlegten Strukturen von Mannose-spezifischem Lektin. Die Koordinierung von Informationen für CVL bereits hinterlegt in der Protein Data Bank (PDB) mit Code 3mez. Allerdings zeigte und bestätigte die Überlagerung mit diesen Mannose-spezifischen Lektinen hochkonservierte Strukturmotive von

antiparallelen β-Faltblättern. Die Struktur des CVL konnte durch Anwendung der Methode des multiisomorphen Ersatzes (Multiple Isomorphous Replacement, MIR) mit Platin-und Quecksilber-Derivaten gelöst worden.

Weiterhin wurde eine Chitinase (CVC) aus *Crocus vernus* isoliert, aufgereinigt und kristallisiert. Die Kristalle zeigten Diffraktion bis 2,1 Å und gehörten zur monoklinen Raumgruppe C2 mit den Zellkonstanten a = 171.81, b = 36.55, c = 125.28 Å und β= 126,9 ° . Die initiale Verfeinerung der Struktur resultierte zu einem *R*-faktor von 30% und R_{free} von 37%.

Contents

List of Figures

List of Tables

List of Abbreviations

2W-MAD	Two wavelength-Multi-wavelengths Anomalous Dispersion
4MU	4-methylumbelliferone
4MUC	4-Methylumbelliferyl N, N´-diacetyl-b-D-chitobioside
4MUG	4-Methylumbelliferyl N-acetyl-b-D-glucosaminide
4MUT	4-Methylumbelliferyl b-D-N, N´, N´´-triacetylchitotriose
ACF	Auto-Correlation Function
C2	C Centered Monoclinic Space group, Number 5
$C222_1$	C Centered Orthorhombic Space group, Number 20
CCD	Charge-coupled Device
CCP4i	Collaborative Computational Project Number 4
cDNA	Complementary Deoxyribonucleic acid
CRD	Carbohydrate Recognition Domain
CSA	*Crocus sativus* agglutinin
CV	Column Volumn
CVA	*Crocus vernus* Agglutinin
CVC	*Crocus vernus* Chitinase
CVL	*Crocus vernus* Lectin
DESY	Deutsches Elektronen-Synchrotron
DLS	Dynamic Light Scattering
DNA	Deoxyribonucleic acid
EMBL	European Molecular Biology Laboratory
F_c	Calculated Structure-factor
F_o	Measured Structure-factor
FPLC	Fast Protein Liquid Chromatography
GNA	*Galanthus nivalis* Agglutinin
HAD	Heavy Atom Derivatives
HCl	Hydrochloric acid
HIV	Human Immunodeficiency Virus
HKL	Miller Indices

MBL	Mannose-binding Lectin
MIR	Multiple Isomorphous Replacement
MR	Molecular Replacement
MWCO	Molecular Weight Cut Off
NCBI	National Center for Biotechnology Information
$P2_12_12$	Primitive Orthorhombic Space group, Number 18
PAGE	Polyacrylamide Gel Electrophoresis
PCS	Photon Correlation Spectroscopy
PCT	Pre-Crystallization Test
PDB	Protein Data Bank
PDB-ID	Identification Code for Protein Data Bank
PEG	Polyethylene Glycol
PR	Pathogenesis Related
R_{free}	*R*-facotr, based on selection of reflections not considered for structure solution
RIPs	Ribosome-inactivating Proteins
R_{merge}	Reliability factor of all symmetry-equivalent reflexes
RNA	Ribonucleic acid
SAD	Single wavelength Anomalous Dispersion
SBA	Soybean Agglutinin
SBL	Soybean Lectin
SCA	*Scilla campanulata* agglutinin
SDS	Sodium Dodecyl Sulfate
SIRAS	Single Isomorphous Replacement with Anomalous Signal
T	Temperature
Tris	tris (hydroxymethyl) aminomethane
UEA	*Ulex europaeus* Agglutinin
UK	United Kingdom
USA	United States of America
UV	Ultraviolet
WGA	Wheat Germ Agglutinin

Physical Units

°	Degree
°C	Degree Celsius
Å	Angstrom
eV	Electronvolt
g	Gram
k	Boltzmann's Constant
K	Kelvin
kDa	Kilo Dalton
M	Molar
mg	Milligram
min	Minute
ml	Milliliter
mm	Millimeter
mM	Milli-molar
nl	Nano liter
nm	Nanometer
R_H	Hydrodynamic Radius
rpm	Revolutions per minute
V_M	Matthews Coefficient
η	Viscosity
λ	Wavelength
µl	Micro liter
µs	Micro-second
χ^2	Chi square

Symbols for Amino Acids

A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartate
C	Cys	Cysteine
E	Glu	Glutamate
Q	Gln	Glutamine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine

1 INTRODUCTION

Iridaceae is a plant family of worldwide distribution having great morphological diversity [1] and is a relatively large family of petaloid monocots [2]. The family has a marked concentration on the southern continents and the major center of radiation in Africa south of the Sahara [3]. With over 2030 species divided among 65–75 genera, Iridaceae are among the largest families of the order Asparagales [4] and they have perennial evergreen or seasonal herbs, with rhizomes, bulbs or corms, rarely annuals or shrubs [5]. The family has taken its name from the genus *Iris* [6].

Several cultures have used species of Iridaceae as food, ornamental, condiment or as medicinal plants. Iridaceae have a long history of traditional medicinal use, recognized as a rich source of secondary metabolites predominantly Flavonoids [7]. The corms of a few common species of *Lapeirousia* and *Moraea* formed part of the traditional diet of some African tribes. Pieces of the *Iris missouriensis* rhizome were used to relieve toothaches [8]. *Belamcanda chinensis* has a long history of use as medicinal plant in China. Apparently, plant extracts were very effective in controlling bacterial, viral and fungal diseases, and in reducing fever and inflammations [9]. Several other genera (e.g., *Crocus*, *Dietes*, *Sparaxis*, *Tritonia*, *Watsonia*) are cultivated in gardens in both tropical and temperate areas. Another popular member of the family is *Crocus sativus*; the spice saffron – obtained from the stigma – has been used for centuries in folk medicine as an antispasmodic, aphrodisiac, expectorant, narcotic and sedative [10].

To date there are about 85 species of *Crocus*, of which approximately 30 species are cultivated. *Crocus* is a genus of perennial flowering plants, native to a large area from coastal and subalpine areas of central and southern Europe (including the islands of the Aegean), North Africa and the Middle East, across Central Asia to western China. *Crocus vernus* (Spring *Crocus*, Giant Dutch *Crocus*) is a plant of the *Crocus* genus in the Iridaceae family. So far two subspecies are recognized: *C. vernus* Hill ssp. albiflorus (Schultes) Asch. & Graeb., a small white or purple-striped or purple type growing mostly

at high altitude in the Alps and Pyrenees and in Sicily and *C. vernus* ssp. *vernus* a somewhat larger and more richly colored type of the Carpeters, the former Yugoslavia and Central Italy [11].

Figure 1: Purple and White colored flowers of *Crocus vernus*

1.1 *Crocus vernus* Defense Proteins

By changing their physiological conditions, higher plants protect themselves from harsh growing environment. These protective reactions are known as "defense responses" of higher plants, and the proteins actively synthesized in accordance with this reaction are called "**defense-related proteins**" [12]. In particular, protective plant proteins specifically induced in pathological or related situations have been intensively studied from an agricultural perspective and are called "**pathogenesis-related proteins**" (**PR proteins**). Recently, protective proteins of higher plants have drawn much attention from plant breeders [13].

1.2 Chitinase

Studies on defense responses have mainly focussed on the endohydrolases that exhibit 1, 3-glucanase and chitinase activities. **Chitinases** are digestive enzymes that break down glycosidic bonds in chitin [14]. Chitinases occur in a wide range of organisms including plants, animals, viruses, bacteria, fungi and insects and play a variety of roles in these organisms [15, 16]. Chitinases are generally found in organisms that either needs to reshape their own chitin or to dissolve and digest the chitin of fungi or animals.

1.3 Classification of Chitinases

Based on sequence similarity, plant chitinases have been classified into seven classes (I–VII) and further grouped into two families of glycosyl hydrolases: family 18 and family 19 [17, 18]. Family 19 is comprised of class I, II, IV and VII chitinases; their structures have high α-helical contents similar to that of lysozyme. Chitinases of class III, V and VI belong to family 18 and their catalytic domain possess a common α, β-TIM barrel. Chitinases of family 18 use a substrate-assisted double-displacement mechanism where as that of family 19 uses a single-displacement mechanism. This difference in mechanism leads to retention for family 18 and inversion of the configuration of the anomeric carbon in case of family 19 chitinases [19, 20, 21, 22, 23].

1.4 Chitin

Excluding cellulose, chitin is one of earth's most abundant natural sources of complex carbohydrates (polysaccharides). Chitin is used as a structural component of the fungal cell wall [24], and insect exoskeletons [25] and is also found in nematode egg shells [26]. Chitin is an abundant product of microbial walls. Chitin is an unbranched homopolymer of 1, 4 - linked N-acetyl-d-glucosamine (GlcNac), a derivative of glucose, and it is found in many places throughout the natural world [28]. It is the main component of the cell walls of fungi, the exoskeletons of arthropods, such as crustaceans (like the crab, lobster and shrimp) and insects, including ants, beetles and butterflies, the radula of mollusks and the beaks of the cephalopods, including squid and octopi. Chitin is a biological substance which may be compared to the polysaccharide cellulose and to the protein keratin.

1.5 Importance of Chitinases

Chitinases are capable of converting the chitin exoskeleton of insects and crustaceans to monomeric sugars. These are digestive enzymes that break down glycosidic bonds in chitin [14]. There is good evidence that the action of the endohydrolases leads to detrimental effects, such as the inhibition of hyphal growth as well as the probable release signaling molecules (1,3-glucans and chitin/chitosan oligomers) that activate defense genes [27]. The proteins have been found in monocots

and dicots and are known to accumulate under developmental regulation, as well as in response to defense-related stimuli. Waste from processing crustaceans, such as shrimp and crabs, forms a major source of inexpensive raw material for the enzymatic conversion of chitin to a more easily fermentable sugar. This, in turn, may be used for alcohol fuel and sugar production.

In plants, chitinases play a major role in their defense by attacking chitin of insects. Therefore, chitinase inhibitors have chemotherapeutic potential against insects [29], fungi [30] and malaria transmission [31, 32]. Chitinase inhibitors were also suggested to have anti-inflammatory potential against asthma and allergic diseases, including atopic dermatitis and allergic rhinitis [33].

1.6 Ribosome-inactivating Proteins (RIPs)

Plants can also contain one or more so-called ribosome-inactivating proteins (RIPs). RIPs are *N*-glycosidases that catalytically inactivate eukaryotic ribosomes [34] by removing a single adenine residue from the large rRNA. RIPs are divided into three groups on the basis of their molecular structure. Type 1 RIP is a single chain protein consisting of an enzymatically active polypeptide chain of approx. 30 kDa. They are widespread among higher plants and occur in monocots as well as in dicots. Type 2 RIPs are built up of one or more protomers consisting of two different disulfide linked A and B chains. The A chain of the [A-s-s-B] pair shows sequence similarity to type 1 RIPs and has RNA *N*-glycosidase activity, whereas the B chain has no enzymatic activity, but contains carbohydrate-binding sites. Due to the presence of these carbohydrates-binding sites, type 2 RIP is also regarded as **Lectins** [35].

The concept of type 3 RIPs was introduced only recently after the discovery of RIPs that are synthesized as single chain zymogens of about 30 kDa and are converted into the enzymatically active form through a posttranslational processing. Hitherto, type 3 RIPs have been identified only in maize and barley [36].

1.7 Lectins

Lectins are sugar-binding proteins that are highly specific for their sugar moieties. These proteins reversibly and non-enzymatically bind specific carbohydrates [37]. Lectins are found in cells, membranes, and secretomes of organisms from all kingdoms of life, and tolerate a degree of binding-pocket sequence variation for ligand specificity rivaling that of immunoglobins [38]. It was presumed that in addition to carbohydrate binding sites lectins may contain one or more sites that interact with non-carbohydrate ligands.

1.8 Classification of Lectins

Because of their diversity, classifying lectins into families is still a challenge. Lectin classification is still evolving, and general agreement has not been achieved yet. Looking at the lectins as a whole, [39] proposed six families based on comparisons of the carbohydrate recognition domains (CRD): (1) legume lectins, (2) cereal lectins, (3) P-, (4) S-, and (5) C-type lectins, and (6) the pentraxins. The first two are found in plants whereas the others are more typical for animals. The plant lectins were previously classified into seven families based on the CRDs: amaranthins, Cucurbitaceae phloem lectins, lectins with hevein domain(s), jacalin related lectins, legume lectins, mannose-binding lectins from monocots, and type-2 ribosome-inactivating proteins [40]. For example, the legume (*Canavalia ensiformis*) lectin Concanavalin A (ConA) binds glucose/mannose residues; soybean (*Glycine max*) agglutinin (SBA), also known as soybean lectin (SBL), binds N-acetyl-D-galactosamine/galactose; gorse (*Ulex europaeus*) lectin (UEA1) binds L-fucose; and the hevein-domain cereal lectin wheat germ agglutinin (WGA) binds N-acetyl D-glucosamine. More recently Damme *et al*., 2005 [41] updated their CRD-based lectin classification scheme from 7 to 12 families that encompass nearly all known and certainly all well characterized plant lectins. This new system incorporates sequence and structural homology as well as evolutionary relatedness to classify the plant lectins. Moreover, these recent studies show that previously reported carbohydrate-binding affinities are sometimes misleading for lectin classification, because lectins, as determined by glycoarray analyses, generally show higher binding affinities to complex oligosaccharides than to simple oligosaccharides [41, 42]. For this reason, the traditional

methods of carbohydrate monomer-based lectin classification are being supplanted by the more informative homology-based systems [41].

1.9 Monocot Mannose-binding Lectins

Over the last two decades, abundant constitutively expressed lectins have been identified in many plant species and these show very diverse molecular structures and sugar specificities [43]. Higher plants developed different structural motifs to bind mannose or oligomannosides [44]. Mannose-binding lectin (MBL; also referred to as mannan-binding lectin and mannose-binding protein) is an important constituent of the innate immune system and 1 of the 30 or so proteins of the complement system [45]. The original discovery of MBL may be traced back to findings made in approx. 1950, when it was discovered that factors in non-immune bovine and murine serums could inhibit influenza virus-induced haemagglutination and lead to influenza virus neutralization [46]. MBL has been shown to bind to a wide range of micro-organisms [47] and in some cases there is a good correlation with known structural features. The monocot mannose-binding lectin family comprises only lectins with an exclusive specificity towards mannose [44]. Numerous members of this super family of lectins have been characterized and cloned from species of the families Alliaceae, Amaryllidaceae, Araceae, Bromeliaceae, Liliaceae and Orchidaceae.

1.1.1 Importance of Mannose-binding Lectins

In the last two decades, plant lectins have been used as recognizing tools to differentiate malignant tumors from benign and the degree of glycosylation associated with metastasis [48]. Recently, many studies have further presented the anti-tumour activities of plant lectins on a variety of malignant cells [49, 50]. Much excitement was created by the findings that lectins, such as those of wheat germ [51, 52], jack bean (*Canavalia ensiformis*) (concanavalin A) [53], and soybean [54] agglutinated malignantly transformed cells but not their normal parental cells. Monocot mannose-binding lectins recently obtained importance because of their potential to recognize anti retrovial molecules [55]. They can recognize the mannosylated region of the envelope

glycoproteins of these viruses e.g., human immunodeficiency virus, simian immunodeficiency virus and feline immunodeficiency virus [56].

2 MATERIALS AND METHODS

A plant has different biologically important organs e.g., Roots, Branches, Leaves, Seeds, Fruits and Underground storage organs like bulbs, corms, tubers etc. Based on plant specie, different organs have different protein profiles. Therefore, initially an extended plant screening was performed to identify and characterize protein profiles. Protein profiles of the following plant species have been screened and analyzed:
Allium, Anemone, Begonia, Chionodoxa, Crocus, Dahlia, Dichelostemma, Eranthis, Fritillaria, Gladiolus, Hyacinth, Hymenocallis, Ipheion, Leucojum, Lilium, Moringa, Muscari, Nectaroscordum, Ornithogalum, Puschkinia, Ranunculus, Salvadora, Snow ball

Roots of *Salvadora*, seeds of *Moringa* and bulbs or corms for the rest of the species were finally selected as target candidates.

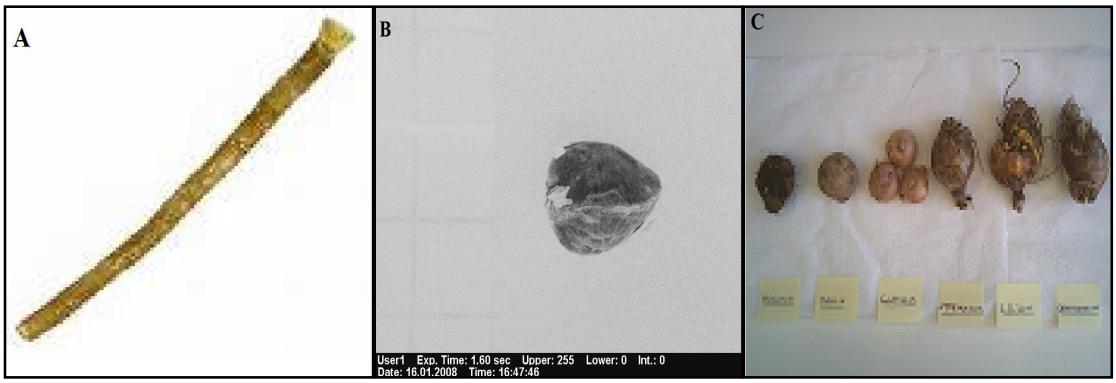

Figure 2: Root of *Salvadora*, B: Seed of *Moringa*, C: Bulbs and Corms

2.1 Crude Extract Preparation

Plant material from different sources was thoroughly washed with distilled water to remove dust or mud. The related plant material was de-shelled or peeled off to remove the outer unnecessary layers. Every individual material was weighted equally to 10 grams using a balance and was powdered by using a blender or pestle and mortar after freezing in liquid nitrogen. Each and every powder material was dissolved in 20mM Tris pH 7.5 buffer and stirred constantly for four hours at 4^{o}C. After stirring, the whole solution was

centrifuged at 14000 rpm for 15 minutes to remove the plant debris. The supernatant was collected in a separate vessel.

2.2 Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide Gel Electrophoresis (PAGE) was performed under both reduced and un-reduced conditions according to standard protocols using the 15% gel [57]. The protein bands were stained with Coomassie Brilliant Blue G250 and destained with a mixture of methanol and acetic acid solution. Protein weight markers (Fermantas, SM0431) were used for the estimation of the molecular mass.

2.3 Spring *Crocus*: *Crocus vernus*

Based on the physical characteristics of crude extracts, protein profiles obtained from the gels, literature studied and detailed search inthe protein data bank (PDB), spring *Crocus* proved to be the ultimate choice for distinct lectin purification. Bulbs (corms) of spring *Crocus* (*Crocus vernus*) were obtained from a local supplier. The bulbs were then immediately stored at – 80 ^{o}C to prevent their germination and as well as the attack of pathogenic organisms.

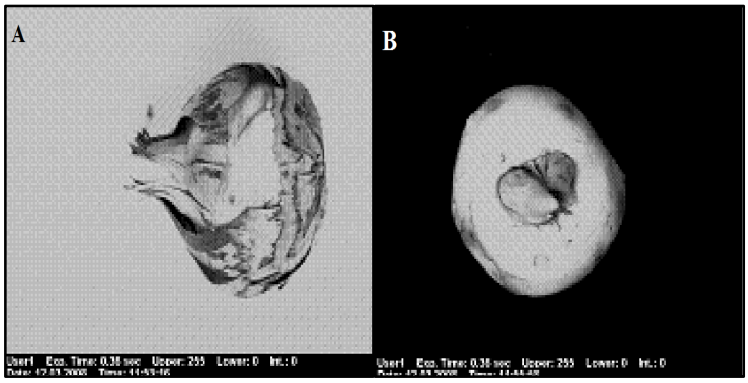

Figure 3: Side view of corm, B: Vertical view of corm

2.3.1 Purification of *Crocus vernus* Proteins

Bulbs of *C. vernus* (83 g) were powdered and dissolved in 500 ml of the sodium acetate buffer A (20 mM, pH 5.0) and stirred for 4-6 hours at 4 ^{o}C. The crude extract obtained was centrifuged at 1000 ×g for 15 min. The supernatant was collected and passed through a series of filters with different range of pore sizes.

2.3.2 Fast Protein Liquid Chromatography (FPLC)

The crude extract was applied to an ÄKTApurifier^{TM} machine containing cationic exchange column (Mono S® HR 10/10), which was previously equilibrated with buffer A. The column was washed with many bed volumes of the equilibrated buffer and then with buffer B (buffer A with 1M NaCl) to elute the desired proteins. The peaks were analyzed by PAGE and purified fractions were pooled to check the final purity. The pooled sample was concentrated until 5ml and loaded on a Gel filtration column Superdex 200 10/300 GL column (GE Healthcare, UK) to remove minor impurities and to obtain crystallization grade purity.

The unbound fraction of Mono S column was dialyzed against buffer C (20mM Tris, pH 8.5) and further applied on anionic exchange column (Mono Q® HR 10/10) already equilibrated with Buffer C (20mM Tris pH 8.5). The column was washed with many bed volumes of the equilibrated buffer and then Buffer D (buffer C with 1M NaCl) was applied to elute the proteins. The purity of the peaks was analyzed with PAGE and purified fractions were pooled together.

2.3.3 N-Terminal Sequences of Proteins

Amino acid sequences were determined using the peptide and protein sequencing system (Applied Biosystems Edman sequencer 476, Germany) using Edman degradation methodology [58]. A Search for homologous sequences and alignments of related sequences was carried out using Protein Blast (NCBI, USA) [59].

2.3.4 UltrafleXtreme™ MALDI-TOF/TOF Mass Spectrometer

The UltrafleXtreme™ MALDI-TOF/TOF mass spectrometer (Bruker, Germany) is using 1 kHz smart beam II laser technology in both TOF and TOF/TOF modes. The

MALDI top-down sequencing (TDS) data were obtained by MALDI in-source decay (ISD) measurements. Matrix substances were sDHB (2, 5-dihydoxybenzoic acid plus 5-methoxysalicylic acid) and 1, 5-DAN (1, 5-diaminonaphthalene). The determination of the mass of the intact protein was done with 2, 5-DHAP (2, 5-dihydroxyacetophenone) as matrix.

2.3.5 Chitinase Assay Kit, Fluorimetric

An enzymatic assay (Chitinase Assay Kit, Fluorimetric, CS1030 Sigma) was applied and confirmed the identification of the purified protein as a Chitinase enzyme. The kit assay is based on the enzymatic hydrolysis of chitinase substrates. This enzymatic hydrolysis releases 4-methylumbelliferone (4MU), which upon ionization in basic pH, can be measured fluorimetrically at an excitation wavelength of 360 nm and an emission wavelength of 450 nm [90]. The use of fluorimetric substrates provides a very sensitive detection system. The fluorimetric absorbance was recorded applying the GENios Instrument XFlour4, version 4.51 (Tecan, Austria).

The chitinolytic enzymes are also categorized based on their enzymatic action on chitin substrates. Endochitinases are defined as the enzymes catalyzing the random cleavage at internal points in the chitin chain. Exochitinases catalyze the progressive release of acetylchitobiose or N-acetyl glucosamine from the non-reducing end of chitin, and are referred to as chitobiosidase and β-N-acetylglucosaminidase, respectively. Kit provides three different substrates for the detection of the various types of the chitinolytic activity [91, 92].

1. 4-Methylumbelliferyl N, N´-diacetyl-b-D-chitobioside (4MUC) – a substrate suitable for exochitinase activity detection (chitobiosidase activity)
2. 4-Methylumbelliferyl N-acetyl-b-D-glucosaminide (4MUG) – a substrate suitable for exochitinase activity detection (b-N-acetylglucosaminidase activity)
3. 4-Methylumbelliferyl b-D-N, N´ ,N´´-triacetylchitotriose (4MUT) – a substrate suitable for endochitinase activity detection

For this purpose, a Chitinase Fluorimetric Assay Kit (Sigma) was used according to the following plan.

No.	Assay	Substrate Working Solution	Sample or Standard Solution	Assay Buffer
1	Blank*	100µl	-------	--------
2	Positive control**	90 µl	10 µl of chitinase control Enzyme	---------
3	**Sample**	**90 µl (3 substrates)**	**10 µl of sample**	---------
4	Standard Blank***	--------	----------	100 µl
5	10ng/assay standard	--------	2 µl of 5µg/ml	98 µl
6	100ng/assay standard	--------	2 µl of 50µg/ml	98 µl
7	500ng/assay standard	--------	10 µl of 50µg/ml	90 µl
8	1000ng/assay standard	--------	2 µl of 500µg/ml	98 µl

*A blank reaction (Substrate Solution without enzyme) should be run to account for the spontaneous hydrolysis of the substrate during the incubation time.

** The volume of the enzyme can range between 1-10 µl, depending on the reaction duration (i.e., for shorter time a higher enzyme concentration is required).

*** Standards should be run when activity calculations are required (samples No. 4-8). A standard curve may be determined with the five standard samples indicated in the table. It is also possible to use only one standard concentration with in the range of 10-1000ng and use the equation in the calculation section.

GENios; Serial number: 12900400599; Firmware: V 4.62 - 07/01 GENios; XFLUOR4 Version: V 4.51

Date:	2.7.09
Time:	14:37
Measurement mode:	Fluorescence Top
Excitation wavelength:	360 nm
Emission wavelength:	465 nm
Gain (Manual):	40
Number of flashes:	10
Lag time:	0 µs
Integration time:	40 µs
Plate definition file:	GRE96ft.pdf
Part of the plate:	A1 - H3

Figure 4: Chitinase 4assay specifications

2.4 Crystallization of the Proteins

After the completion of the first step of protein purification the proteins were immediately subjected to the different stages of crystallization protocols.

2.4.1 Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) Spectroscatter 201 (Molecular Dimensions, UK) is an instrument which can be used to determine the size distribution profile of small particles in suspension or polymers in solution. Analysis of the autocorrelation function in terms of particle size distribution is done by program CONTIN [88]. A truly monodisperse sample would give rise to a single exponential decay to which fitting of a calculated particle size distribution is relatively straightforward. Dynamic light scattering (also known as Photon Correlation Spectroscopy (PCS)) is the general designation for a method to determine the size of small particles in the submicron range. These particles are found in Brownian motion in suspension or emulsion. The diffusion speeds for this are inversely proportional to the radius **rp** (hydrodynamic radius) of the particles.

$$D = \frac{k \times T}{3 \times \pi \times \eta \times r_p}$$

(k = Boltzmann's constant)

Temperature **T** and viscosity **η** of the solution are important parameters, which must be accurately known. Since solutions have Brownian particles motion, the back

scattered light will have a frequency shift (Doppler shift). The resulting intensity variation is recorded by a highly sensitive detector.

2.4.2 Pre-Crystallization Test

The PCT Pre-Crystallization Test was used to determine the optimized protein concentrations for crystallization screening. Highly concentrated samples resulted in amorphous precipitate, while diluted samples had produced transparent drops. Amorphous precipitate and clear drop production was avoided by changing the protein concentration accordingly.

The four reagents of PCT kit, used to evaluate protein concentration for crystallization screening, are:

1) Reagent A1: 0.1M Tris Hydrochloride pH 8.5, 2.0M Ammonium Sulfate

2) Reagent B1: 0.1M Tris Hydrochloride pH 8.5, 1.0M Ammonium Sulfate

3) Reagent A2: 0.1M Tris Hydrochloride pH 8.5, 0.2M Magnesium Chloride Hexahydrate, 30% w/v Polyethylene Glycol 4,000

4) Reagent B2: 0.1M Tris Hydrochloride pH 8.5, 0.2M Magnesium Chloride Hexahydrate, 15% w/v Polyethylene Glycol 4,000

2.4.3 Robotic Screening for Crystallization

In structural genomics/proteomics, major advances in protein crystallization have been made using robotics, which has automated the crystallization experiment and reduced the amount of protein required by an order of magnitude, improving the reproducibility of the experiments and allowing a large number of set ups.

Figure 5: Zinsser Pipetting Robot (Digilab Genomic Solution, Germany)

2.4.4 Optimization of Crystals

The hanging drop vapor diffusion method has been applied to improve initially obtained crystals using the Linbro plates. A 2 µl protein solution was mixed with 2 µl reservoir solution and equilibrated against 1.0 ml reservoir solution. X-ray diffracting crystals possessing approx. dimensions of 0.4 × 0.2 × 0.05 mm were obtained by incubating plates at 20 ^{o}C for 3-5 days. Optimal size crystals were produced using 12% PEG4000, 200mM $MgCl_2.6H_2O$ and 100mM Tris-HCl buffer.

2.5 Data Collection

2.5.1 Native Diffraction Data

Native X-ray diffraction data from single crystals were collected by exposing them at the synchrotron Consortium-Beamline X13 DESY, Hamburg. A crystal was mounted on a nylon loop and flash-cooled in cold nitrogen-gas stream at 100 K. All intensity data were indexed, integrated and scaled with the HKL-2000 package [60]. Phase information was generated using the Molecular Replacement (MR) strategy. Therefore, the program Molrep from the CCP4i suite [61] was used.

2.5.2 Matthews Coefficient (V_M)

One of the most important parameter for most density modification methods is the fraction of the unit cell (or asymmetric unit) that is considered as solvent. The most often used way to calculate the *solvent content* is via the Matthews parameter [62]. The

Matthews coefficient provides this number. The Matthews coefficient (Vm) is calculated as:

$$\frac{\text{volume of unit cell}}{\text{the molecular weight of macromolecule} \times Z \times X}$$

Z is the number of asymmetric units in the unit cell (i.e. the number of symmetry operators in space group). The unknown variable, X, is the number of molecules in the asymmetric unit.

2.5.3 Rigaku RU-200B: X-ray Generator

A Rigaku Rotating anode generator, Rotaflex RU-200B series with a PSPC (position sensitive proportional counter) type microdiffractometer with auxiliary cabinet, H.V. transformer, water circulator and an IP scanner, was used for the first diffraction data collection.

2.5.4 Cryogenic Crystal Recovery Techniques

After successful data collection, a small Dewar was filled with liquid nitrogen to keep the internal environment in freezing temperature while crystal was still on the goniometer. The head of the cryo-tongs was placed into the nitrogen along with cryo-vial and forceps until they stopped boiling.

Figure 6: Cryogenic technique accessories

The tong was clamped around the crystal and quickly returned with the crystal to the liquid nitrogen in the Dewar. With another pair of forceps clamped the base of the cryo-pin, and the crystal was released from the cryo-tongs.

The crystal was transferred to the cryo-vial. During the whole procedure of crystal preservation, the crystal was kept strictly under the liquid nitrogen surface as little change in temperature can damage the crystal order and ultimately the data quality. The cryo-pin was screwed onto the cryo-vial using the magnetic wand and was transferred into the liquid nitrogen filled Dewar.

2.5.5 Macromolecular Crystallography Beamline X13

Beamline X13 is a bending magnet macromolecular crystallography beamline located in the HASYLAB Hall 5. It is a monochromatic fixed wavelength beamline operating at a wavelength of 0.81 Å (15.3 keV). The beamline is equipped with a MARCCD detector (165 mm). The station is optimized for rapid data collection, high resolution studies and Single Wavelength Anomalous Diffraction (SAD) experiments and for halide (Br) soaked crystals.

2.5.6 Heavy Metal Derivative Diffraction Data

Derivative X-ray diffraction data from metal soaked crystals were collected also by exposing them at the synchrotron EMBL-Beamline X12, DESY, Hamburg. Crystals were mounted on a nylon loop and flash-cooled in a cold nitrogen-gas stream at 100 K. All intensity data were indexed, integrated and scaled with the HKL-2000 package. Phase information was derived by Multiple Isomorphous Replacement (MIR) strategy using the program SOLVE/RESOLVE [63].

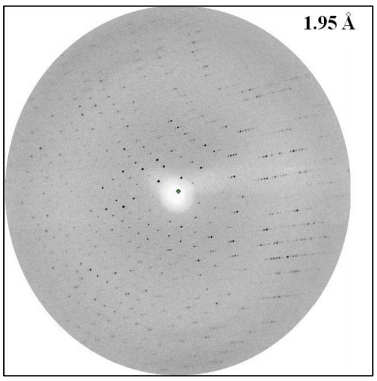

Figure 7: A diffraction image

2.5.7 Macromolecular Crystallography Beamline X12

Beamline X12 at EMBL Hamburg is tunable in the energy range of 5500 - 18000 eV (wavelengths of 2.1 - 0.7 Å respectively). The beamline is provided with a MAR desktop beamline, a large MARMosaic-CCD detector (225 mm) and a Bruker/AXS energy-sensitive X-ray fluorescence detector. The station is optimized for MAD or SAD data collection.

Table 1: **Heavy metal derivatives used for soaking experiments**

HAD Name	Nr.	Mol. Wt. g/mol	Stock Conc.
Ethyl mercury chloride	A	265.1	12mM
Ethyl mercuric thiosalicylic acid	B	404.8	13mM
Mercurric acetate	C	318.7	25mM
Mercurric chloride	D	271.5	30mM
Methyl mercuric acetate	E	275.0	17mM
Methyl mercuric chloride	F	251.1	17mM
Mersalyl acid	G	483.9	8mM
P-chloromercuricbezoic acid	H	357.1	11mM
Potassium dicyanoaurate I	I	288.1	49mM
Potassium hexachloride Iridate	J	483.1	4mM
Potassium hexachloride osmate	K	481.1	9mM
Potassium hexachloro platinateII	L	486.0	8mM
Potassium tetracholoraurate	M	377.8	19mM
Potassium tetrachloro platinateII	N	415.1	13mM
Samarium acetate hydrate	O	327.5	15mM
Samarium III chloride hexahydrate	P	256.8	20mM
Samarium III nitrate hexahydrate	Q	444.4	15mM
Silver Nitrate	R	265.1	32mM

2.5.8 Multiple Isomorphous Replacement (MIR)

Multiple Isomorphous Replacement, or MIR, is an approach to solve the phase problem in X-ray crystallography. This method involves soaking of crystals into a heavy atom solution. It is obligatory that heavy atom should not affect the crystal lattice or unit cell dimensions in comparison to its native form, hence, they should be isomorphic.

Native and heavy-atom derivative data sets of the samples were collected. Then the interpretation of the Patterson difference map revealed the heavy atom's position in the unit cell. This allowed determining of both the amplitude and the phase of the atom. Since the structure factor of the heavy atom derivative (F_{ph}) of the crystal is the vector sum of the heavy atom (F_h) and the native crystal (F_p), the phase of the native F_p and F_{ph} vectors was solved by geometrically.

$$\mathbf{F}_{ph} = \mathbf{F}_p + \mathbf{F}_h$$

2.5.9 SOLVE/RESOLVE Program Package

SOLVE/RESOLVE are the crystallographic soft wares that can carry out macromolecular structure determination and electron density maps improvements automatically. SOLVE scales data, solves Patterson functions, and calculates difference Fourier to distinct solvent and protein regions, and can score partial MAD and MIR solutions to build up a complete solution.

RESOLVE uses a statistical approach to combine experimental X-ray diffraction information about an electron density map of a macromolecule [63].

2.5.10 Auto-Rickshaw Web Server

The EMBL-Hamburg automated crystal structure determination platform can executes many macromolecular crystallographic software programs to produce a software pipeline for automated and efficient crystal structure determination. The corresponding processes have been optimized so that the pipeline can be used efficiently for validating the X-ray experiment at a synchrotron beamline [64, 65].

2.6 Model Building and Refinement

After solving the phase problem, the three dimensional model of the proteins were build using Coot [66] and refined by Refmac5 [67].

2.6.1 Coot: Crystallographic Object-Oriented Toolkit

Coot is designed for macromolecular model building, model completion and validation, particularly suitable for protein modeling using X-ray data. Coot displays maps and models and allows model manipulations such as idealization, real space refinement, manual rotation/translation, rigid-body fitting, ligand search, solvation, mutations, rotamers, Ramachandran plots, skeletonization, non-crystallographic symmetry and much more.

2.6.2 Refmac5

Refmac5 is a refinement program for macromolecular structures. The Refmac5 program can carry out rigid body, TLS, restrained or unrestrained refinement against X-ray data. Refmac5 will refine an atomic model by adjusting the model parameters (coordinates, B-factors, TLS etc.) in order to obtain the model which best explains the experimental data (i.e. maximizes the likelihood). Progress is monitored by the R-factor and Free R-factor [89], as well as by the likelihood scores themselves [67].

2.7 Ligand Binding Experiments

The main use of protein structures in medicine and biotechnology is of course guiding medicinal chemists in their effort to synthesize better compounds to bind with the target proteins. To do this, it is necessary to analyze the structures of protein-ligand complexes. Many of drug discovery projects are supported by X-ray crystallographic efforts and two common ways to obtain these ligand protein complexes are co-crystallization and crystal soaking.

2.7.1 D - (+) - Mannose

Mannose is a sugar monomer of the aldohexose series of carbohydrates. Mannose is present in numerous glyco-conjugates including N-linked glycosylation of proteins. Both co-crystallization and crystal soaking experiments with different concentrations of Mannose were performed.

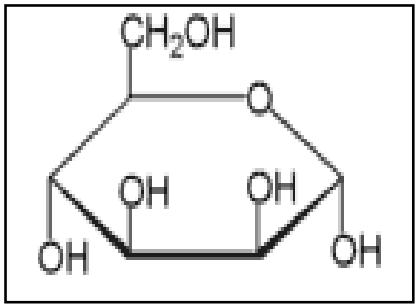

Figure 8: D-Mannose

The following different concentrations of Mannose were used in both co-crystallization and soaking experiments.

Table 2: Different molar concentrations of Mannose sugar

Co-crystallization solutions	Crystal soaking solutions
10mM Mannose	10mM Mannose
25mM Mannose	25mM Mannose
50mM Mannose	50mM Mannose
75mM Mannose	75mM Mannose
100mM Mannose	100mM Mannose
150mM Mannose	150mM Mannose

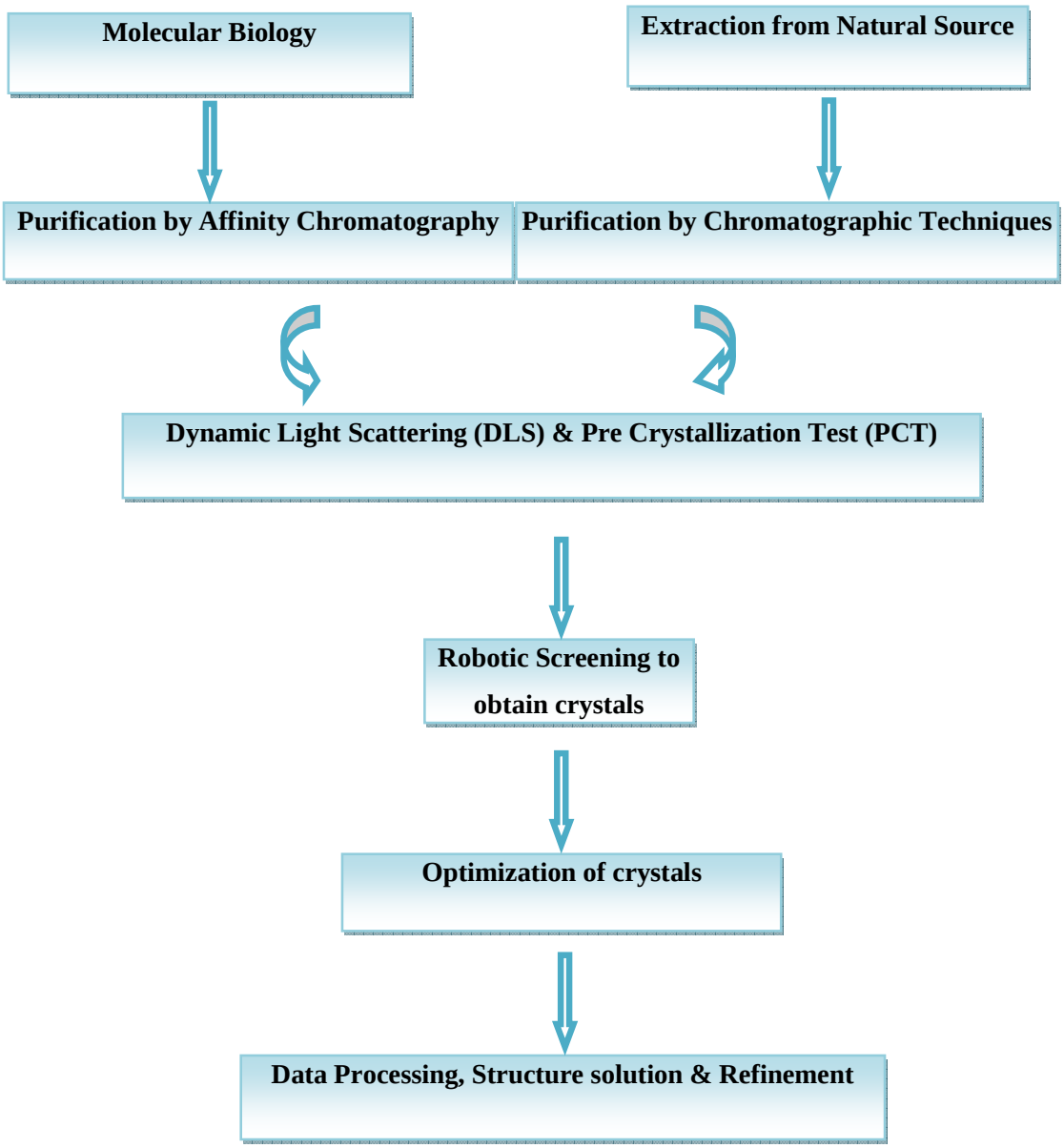

Figure 9: A brief Flow Diagram showing different steps involved in the determination of 3-dimensional protein structure

3 RESULTS AND DISCUSSION

3.1 Initial Screening

Around 23 species were screened to obtain the ultimate choice, based on the physical characters of the crude extract and the gel analysis of the protein profiles. Many of the crude extracts showed good protein profiles on the gel, but their crude extracts were unfortunately either highly pigmented or highly viscous because of the presence of abundant polysaccharides in the solutions.

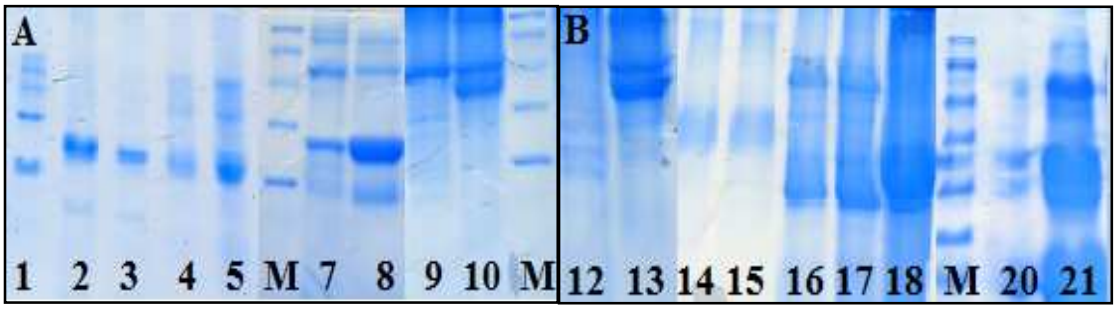

Figure 10: SDS-polyacrylamide gel electrophoresis of crude extracts

Well 1: 10µl *Viburnum* crude	+	10µl sample buffer
Well 2: 10µl *Muscari* crude	+	10µl sample buffer
Well 3: 10µl *Fritillaria* crude	+	10µl sample buffer
Well 4: 10µl *Ipheion* crude	+	10µl sample buffer
Well 5: 10µl *Ornithogalum* crude	+	10µl sample buffer
Well 6: 10µl Marker		
Well 7: 10µl *Crocus* Paris d Claus crude	+	10µl sample buffer
Well 8: 10µl *Crocus* Advance crude	+	10µl sample buffer
Well 9: 10µl *Eranthis* crude	+	10µl sample buffer
Well 10: 10µl *Hymenocallis* crude	+	10µl sample buffer
Well 11: 10µl Marker		
Well 12: 10µl *Salvadora* crude	+	10µl sample buffer
Well 13: 10µl *Chionodoxa* crude	+	10µl sample buffer
Well 14: 10µl *Nectaroscordum*	+	10µl sample buffer
Well 15: 10µl *Leucojum* crude	+	10µl sample buffer
Well 16: 10µl *Puschkinia* crude	+	10µl sample buffer
Well 17: 10µl *Ranunculus* crude	+	10µl sample buffer
Well 18: 10µl *Dichelostemma* crude	+	10µl sample buffer

Well 19: 10µl Marker

Well 20: 10µl *Anemone* + 10µl sample buffer

Well 21: 10µl *Hyacinth* crude + 10µl sample buffer

Spring *Crocus* or *Crocus vernus* was the ultimate choice based on the physical character of the crude extract solution and the better protein profile obtained from the results of the gels shown before.

3.2 Purification of *Crocus vernus* Proteins

Crocus vernus (Spring *Crocus*, Giant Dutch *Crocus*) is a plant of the *Crocus* genus in the Iridaceae family. Its cultivars are used as ornamental plants. The Dutch *Crocuses* are larger than the other cultivated species (Snow *Crocuses*), and tend to bloom two weeks later.

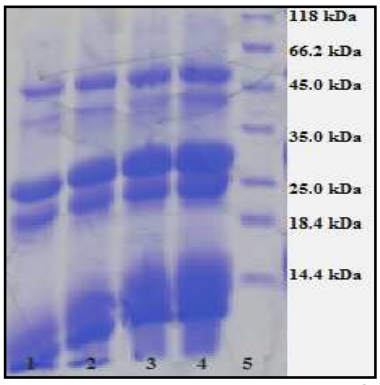

Figure 11: SDS-polyacrylamide gel electrophoresis of *Crocus* crude extract

Well 1: 2µl *Crocus* crude + 2µl sample buffer

Well 2: 4µl *Crocus* crude + 4µl sample buffer

Well 3: 6µl *Crocus* crude + 6µl sample buffer

Well 4: 8µl *Crocus* crude + 8µl sample buffer

Well 5: 10µl Marker

3.2.1 Fast Protein Liquid Chromatography (FPLC)

There have already been two separate reports about the isolation of two important proteins from the genus *Crocus*. A 50 kDa tetramer monocot mannose-binding Lectin [68] which appears as 12-14 kDa protein bands on the gels due to the presence of only hydrogen bonds between the chains. Similarly another enzymatic Chitinase protein has been reported from the same genus *Crocus* [69]. However, no 3-dimensional molecular structure had been determined for the proteins. 50 ml of the filtered sample was loaded on

a MonoS column with 0-100% buffer B (20mM Sodium acetae, pH 5.0, 1.0M NaCl) in 230 column volumes (CV) and the active fractions were ran on the 15% gel to see the protein profile.

Unbound parts were further used in the purification of the second protein through anion exchange chromatography.

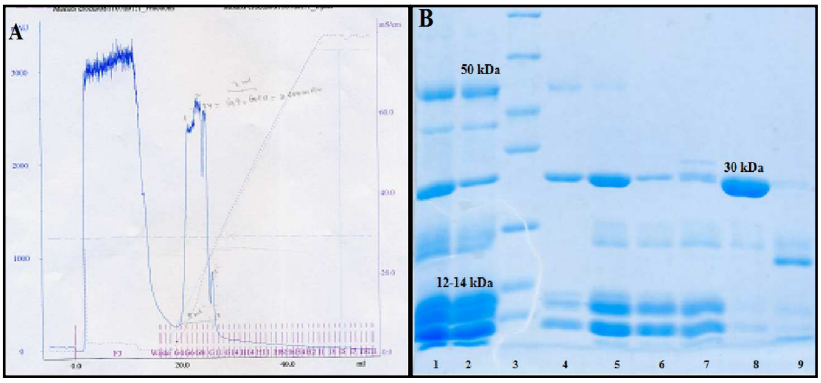

Figure 12 A: Elution profile of *Crocus* crude extract B: SDS-polyacrylamide gel electrophoresis

Well 1: 10µl *Crocus* crude + 10µl sample buffer
Well 2: 10µl Unbound + 10µl sample buffer
Well 3: 10µl Marker
Well 4: 10µl G6 + 10µl sample buffer
Well 5: 10µl G7 + 10µl sample buffer
Well 6: 10µl G8 + 10µl sample buffer
Well 7: 10µl G9 + 10µl sample buffer
Well 8: 10µl G10 + 10µl sample buffer
Well 9: 10µl G611 + 10µl sample buffer

3.2.2 Optimized Chromatogram for Chitinase

Successful elution of the protein was obtained between 0-20% buffer B in 200 column volumes (CV).

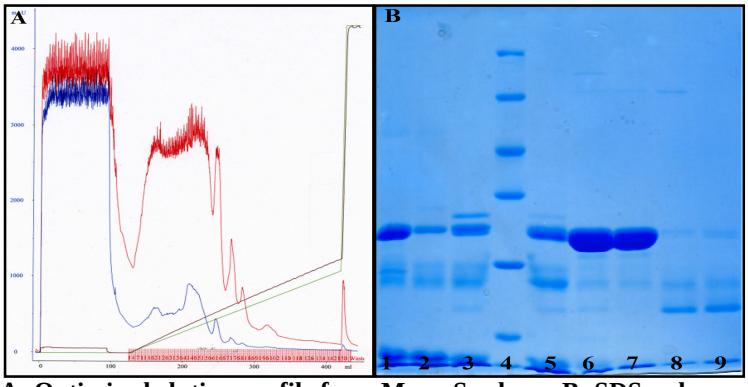

Figure 13 A: Optimized elution profile from Mono S column B: SDS-polyacrylamide gel electrophoresis

Well 1: 10µl fraction 74	+	10µl sample buffer
Well 2: 10µl fraction 78	+	10µl sample buffer
Well 3: 10µl fraction 81	+	10µl sample buffer
Well 4: 10µl Marker		
Well 5: 10µl fraction 83	+	10µl sample buffer
Well 6: 10µl fraction 85	+	10µl sample buffer
Well 7: 10µl fraction 87	+	10µl sample buffer
Well 8: 10µl fraction 92	+	10µl sample buffer
Well 9: 10µl fraction 94	+	10µl sample buffer

3.2.3 Superdex 200 Gel Filtration for Chitinase

The most purified fractions, after the analysis of the gel, were pooled together and were dialyzed using a dialysis membrane with a MWCO (Molecular Weight Cut Off) 3.5 kDa against the Buffer A (20mM sodium acetate, pH 5.0) to get rid of any salt traces from it.

For the gel filtration run, a Superdex 200 10/300 GL column (GE Healthcare) was equilibrated with 2 CV of 20mM acetate pH 5.0 and 100mM NaCl inside to prevent any non-ionic interactions between the gel matrix and the protein. It took almost 1 CV to elute all proteins from the column with a considerable separation of the contaminants

from the target protein, as indicated from the gel. Again the most purified fractions were pooled and dialyzed to get rid of all salts.

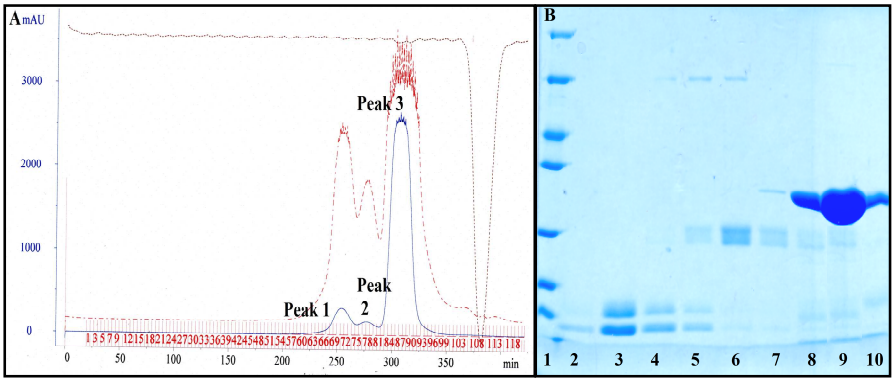

Figure 14 A: Gel filtration of chitinase on Superdex 200 B: SDS-polyacrylamide gel electrophoresis

Well 1: 10µl Marker

Well 2: 10µl fraction 69 peak 1	+	10µl sample buffer
Well 3: 10µl fraction 72 peak 1	+	10µl sample buffer
Well 4: 10µl fraction 74 peak 1	+	10µl sample buffer
Well 5: 10µl fraction 76 peak 2	+	10µl sample buffer
Well 6: 10µl fraction 78 peak 2	+	10µl sample buffer
Well 7: 10µl fraction 80 peak 2	+	10µl sample buffer
Well 8: 10µl fraction 84 peak 3	+	10µl sample buffer
Well 9: 10µl fraction 89 peak 3	+	10µl sample buffer
Well 10: 10µl fraction 92 peak 3	+	10µl sample buffer

3.2.4 N-Terminal Amino Acid Sequence Analysis of Chitinase

The partial N-terminal amino acid sequence analysis of the 30 kDa protein A using Edman sequencer 476 (Applied Biosystems, USA) provides the following sequence.

TLFVEYIGYP LFSGVKFSDV PINPHITKFQ FVLSFAVDYT

Feeding this information into the protein blast (ExPasy Proteomics Server) [70] indicated that it correspond to a Chitinase, indicating 76% sequence identity with ´´Chitinase a`` from *Gladiolus x gandavensis* [71].

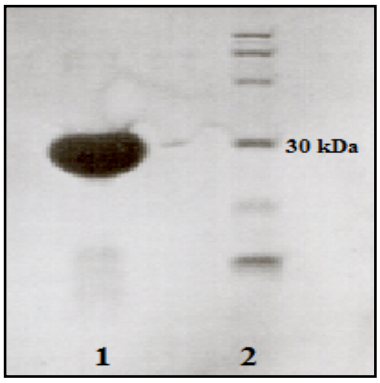

Figure 15: Western Blot for N-Terminal sequencing

Well 1: 10ul protein A sample + 10ul reduced buffer

Well 2: 10µl Marker

3.2.5 UltrafleXtreme™ MALDI-TOF/TOF Mass Spectrometer

MALDI-TOF/TOF mass spectrometer produced the following amino acid fragmented sequence of the chitinase protein.

TLFVEYIGYP LFSGVKFSDV PINPHITKFQ FVLSFAVDYT ASSPHTSTNG

KFNVFWDSAL LGPDQLGVDK (gap of about 15 amino acids)

VGSNTVQFQA ASVDSWVSNA VTSLTR (gap of 3-4 amino acids)

YNLDGIDIDY EHFQNTDKNT FAECIGR (gap of about 60 amino acids)

FLGYYNNDNK (gap of 20-25 amino acids)

GFFDAATSLK NKGKLHGIAV WTADTSKKSD FRYEEEAQAF LVS

Both protein blast of the N-terminal sequence and mass spectrometry sequences identified protein as Chitinase protein.

3.2.6 Enzyme Assay

For further identification of the protein, an enzymatic assay was performed. The assay was performed in a micro titre plate and wells were filled respectively. Micro titre plate was incubated for half an hour at 37 ^{o}C. The measurement was recorded fluorimetrically at an excitation wavelength of 360 nm and an emission wavelength of 450 nm.

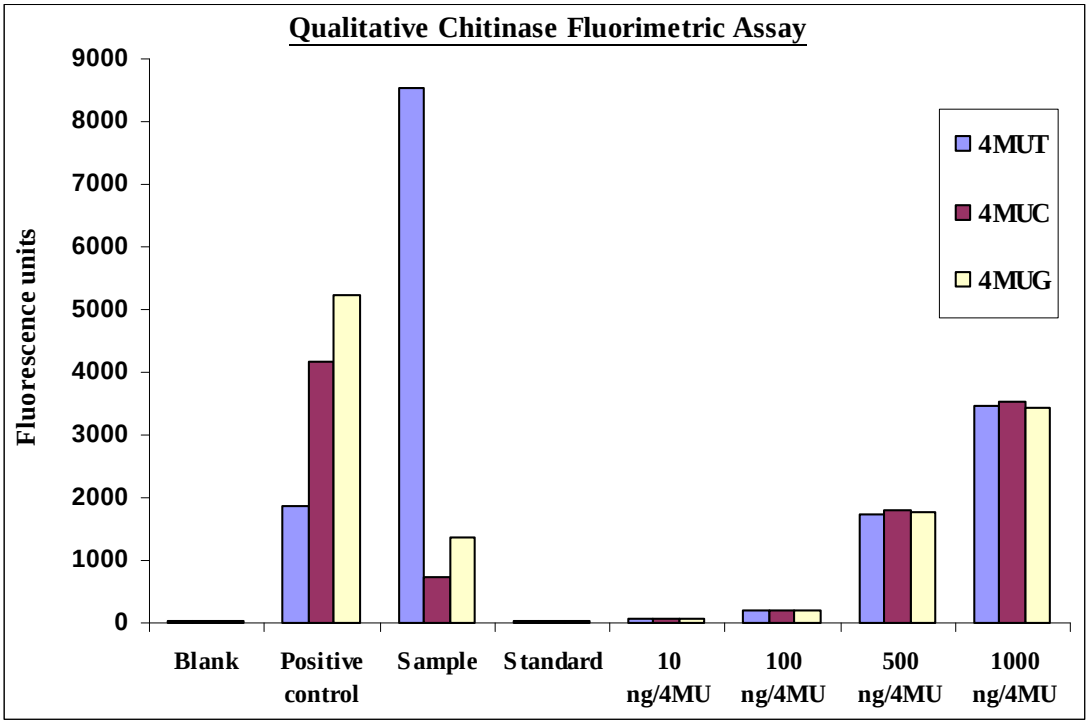

Figure 16: Endochitinase activity for the purified chitinase

The crude enzyme preparation hydrolyzed 4MUT at significant higher rates than 4-MUC and 4-MUG. Hence, the different activities towards these substrates confirmed CVC to be an endochitinase.

3.2.7 Purification of *Crocus* Lectin

The unbound part coming from the Mono S column was collected, dialyzed against buffer C (20mM Tris pH 8.0) and was directly applied on the anion exchange Mono Q column, which was washed and equilibrated with many column volumes of the same buffer C. Under optimized conditions, purified protein B was eluted between 0-30% of buffer D (buffer C with 1M NaCl) in 50 CV. It appeared as 12-14 kDa bands on the gel.

But we have already observed during the gel filtration of chitinase that in the solution it possesses the molecular weight of 50 kDa and is eluting before the 30 kDa Chitinase.

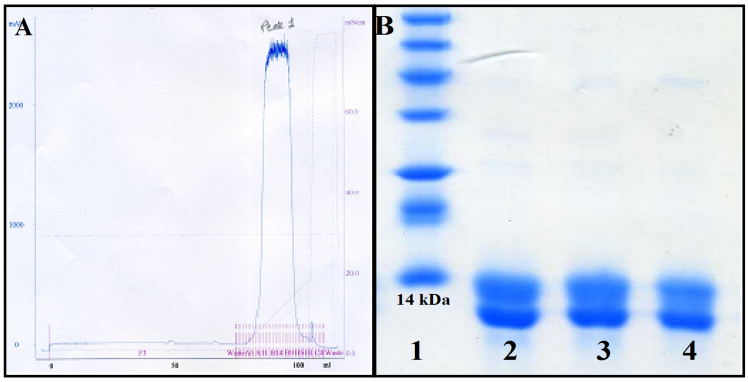

Figure 17 A: Optimized elution profile from Mono Q column B: SDS-polyacrylamide gel electrophoresis

Well 1: 10µl Marker

Well 2: 10µl fraction 12 + 10µl sample buffer

Well 3: 10µl fraction 15 + 10µl sample buffer

Well 4: 10µl fraction 17 + 10µl sample buffer

3.2.8 N-Terminal Amino Acid Sequence Analysis of Lectin

The partial N-Terminal amino acid sequence analysis of the 50 kDa lectin using Edman sequencer 476 (Applied Biosystems, USA) provides the following sequence.

DNNVLLTGDV

Feeding this residual stretch into protein blast (ExPasy Proteomics Server) [70] provided a 100% match with a **Lectin** protein from *Crocus vernus* [68].

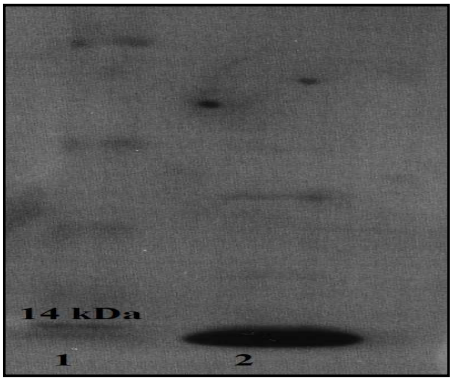

Figure 18: Western Blot for N-Terminal sequencing

Well 1: 10µl Marker

Well 2: 10ul protein B sample + 10ul reduced buffer

3.3 Crystallization of the Chitinase

To start crystallization experiments with the chitinase protein, the purified protein was first dialyzed against acetate buffer (20mM pH 5.0) and then subjected to DLS measurements.

3.3.1 Dynamic Light Scattering (DLS) Measurement

To check the monodispersive nature of the protein solution, purified protein was subjected to DLS measurements and a strong single signal confirmed that the protein was monodisperse under the given buffer conditions. Moreover the molecular weight calculated from the obtained hydrodynamic radius indicated a monomeric protein solution.

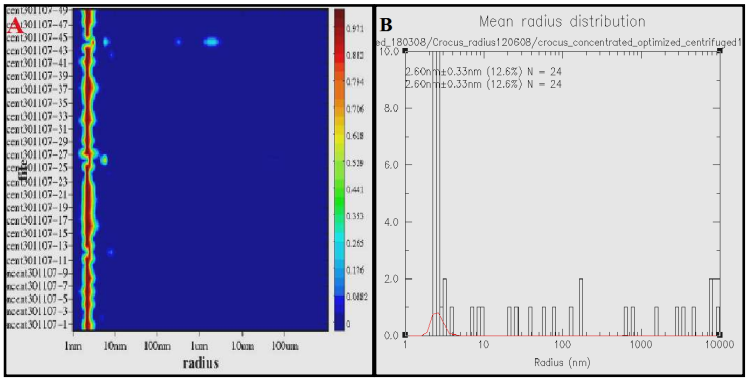

Figure 19 A: DLS measurement showing monodisperse protein solution B: Hydrodynamic radius (R_H) of 2.6nm confirmed a monomeric form in solution

3.3.2 Pre-Crystallization Test

To optimize the concentration of the protein a PCT™ Pre-Crystallization Test (Hampton Research) was performed to align most suitable protein concentration for crystallization experiments.

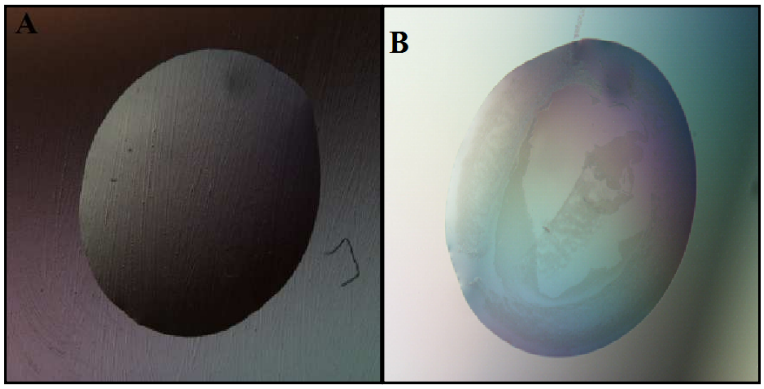

Figure 20 A: Solution A1 showing a transparent drop B: Solution B1 showing homogenous granulate precipitation

3.3.3 Robotic Screening for Crystallization

Robotic screening applying the sitting drop method was performed with 480 different potential conditions in the NeXtal 96 wells plates (Qiagen) using 300 nl spotted in each well. Five screening plates were placed at 20 ^{o}C and analyzed after one week. A few potent conditions were identified via UV microscope through the fluorescence due to tryptophan presence in the primary sequence. The most suitable conditions are listed below:

1. Condition: Hampton Research Crystal Screen 2 Number 33: 16% w/v PEG 4000, 0.1M Tris pH 8.5, 0.2M Magnesium Chloride (Figure 21 A)
2. Condition: Hampton Research Crystal Screen 3 Number 23: 17% w/v PEG 20000, 0.1M Tris pH 8.5, 0.1M Magnesium Chloride (Figure 21 B)

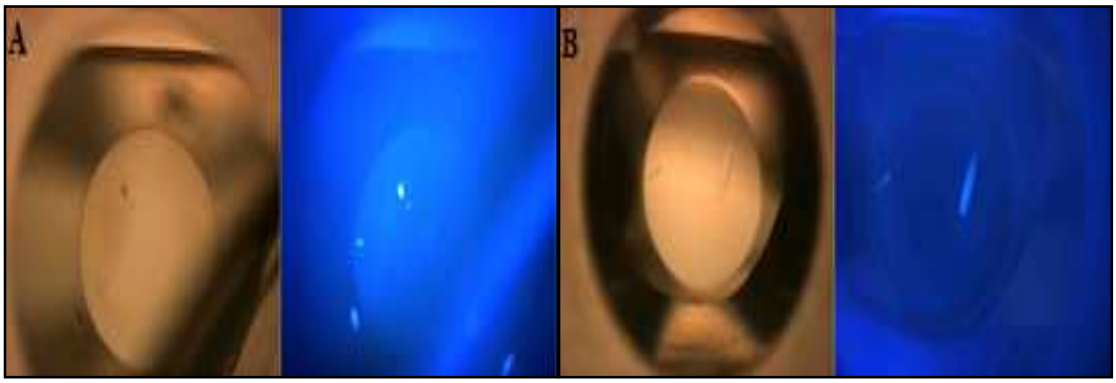

Figure 21 A & B: Visible and UV microscopic photos for drops showing initial crystals

3.3.4 Optimization of Chitinase Crystallization Experiments

Above two described conditions were optimized using the Hanging drop strategy with linbro plates and 2 + 2 µl drop size. The linbro plates were put at 20 ^{o}C temperature. Good shaped crystals with suitable dimensions were grown using 14% w/v PEG 4000, 0.2M Magnesium Chloride and 0.1M Tris pH 8.5 buffer (Figure 22).

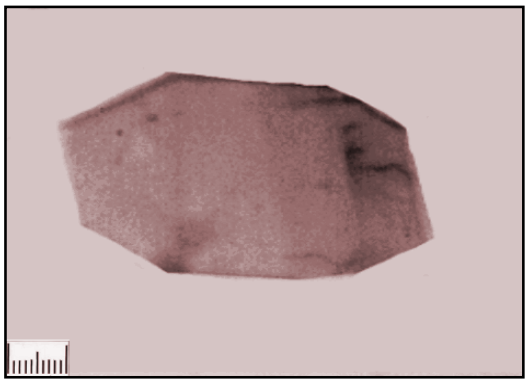

Figure 22: Chitinase crystal dimensions of 0.625 × 0.370 × 0.01 mm. Scale bar, 0.1 mm

3.4 Crystallization of the Lectin

To start crystallization of the lectin protein, purified protein was dialyzed first against Tris buffer (20mM pH 8.0) and subjected to DLS instrument.

3.4.1 Dynamic Light Scattering (DLS) Measurement

To check the monodispersive nature of the protein solution, purified protein was subjected to DLS measurement and a strong single signal confirmed that the protein was monodisperse under the given buffer conditions. Moreover the molecular weight calculated from the obtained hydrodynamic radius indicated a monomeric protein solution.

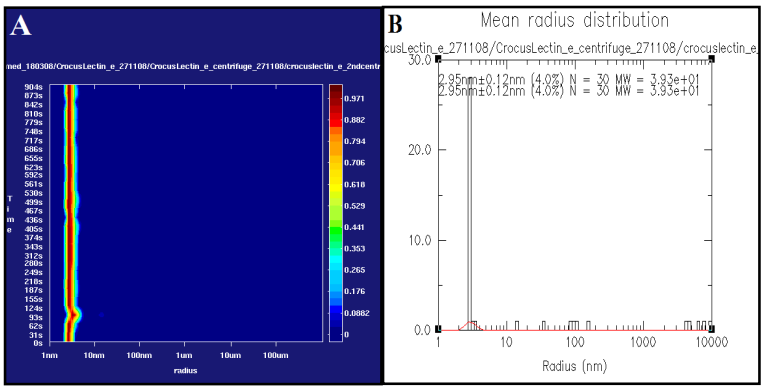

Figure 23 A: DLS measurement showing a monodispersive protein solution B: Hydrodynamic radius (R_H) of 2.95nm confirmed the monomeric form

3.4.2 Pre-Crystallization Test

To optimize the concentration of the protein a PCT™ Pre-Crystallization Test (HAMPTON RESEARCH) was performed to align most suitable protein concentration for crystallization experiments.

Formulation B2 of PCT showed crystal growth in a drop straight away. Reagent B2: 15% w/v Polyethylene Glycol 4000, 0.2M Magnesium Chloride hexahydrate, 0.1M Tris Hydrochloride pH 8.5.

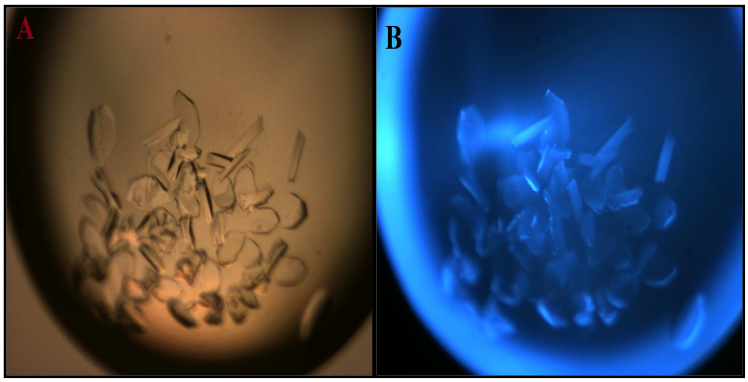

Figure 24 A & B: Visible and UV microscopic photos showing the crystals

3.4.3 Optimization of Lectin Crystallization

The above described condition was optimized using the hanging drop method applying linbro plates with 2 + 2 µl drop size. A linbro was put at 20 ^{o}C temperature

which after three days provided good shaped crystals with suitable dimensions of around 0.50 – 0.70 mm using the following crystallization recipe.

14% w/v PEG 4000, 0.2M Magnesium Chloride, 0.1M Tris pH 8.5 (Figure 25)

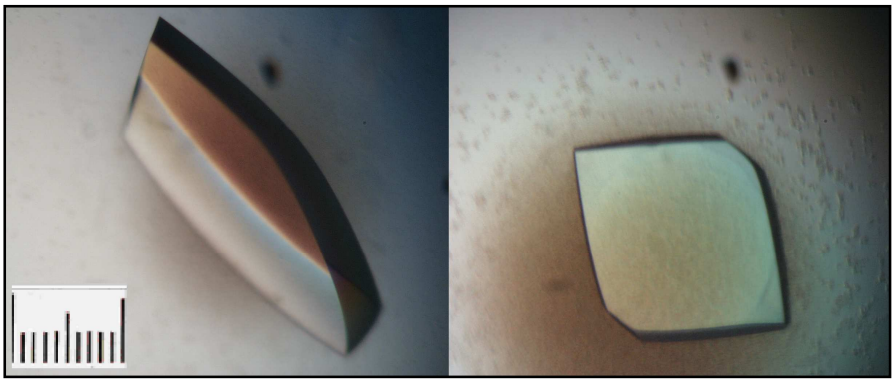

Figure 25: Crystal with dimensions of 0.5 × 0.370 × 0.1 mm. Scale bar, 0.2 mm.

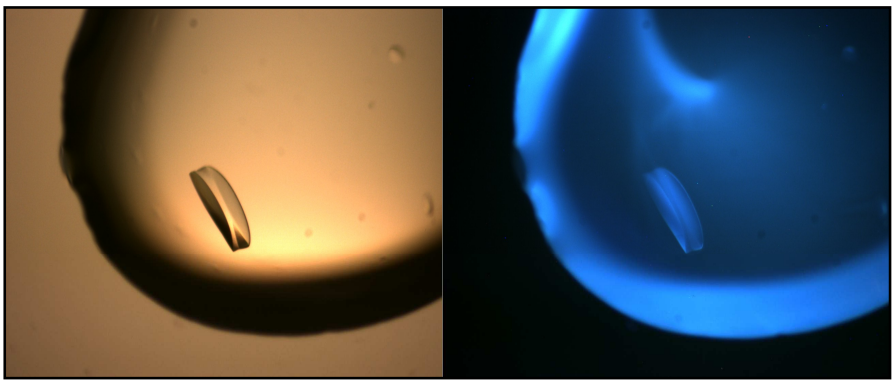

Figure 26 A & B: Visible and UV microscopic photos showing the crystals

3.5 Data Collection

First native diffraction data were collected at the Consortium-Beamline X13, DESY Hamburg. The detector type used was a MarCCD 165 mm and the wavelength of the X-ray used was 0.8123 Å. Data collection was performed at 100 K and the ice production was prevented by soaking the crystal in the same mother liquor with 20% Glycerol inside just before the crystal was mounted. Data with maximum resolution of 2.5 Å and Phi of 1.0^{o} per image were collected. The data were processed using the HKL-2000 program package [60]. The unit cell was orthorhombic and the space group was

identified to be $C222_1$ with cell dimensions a = 88.09, b = 145.22, c = 218.96 Å. The Matthews Coefficient V_M calculation [62] indicated the presence of three molecules in the asymmetric unit.

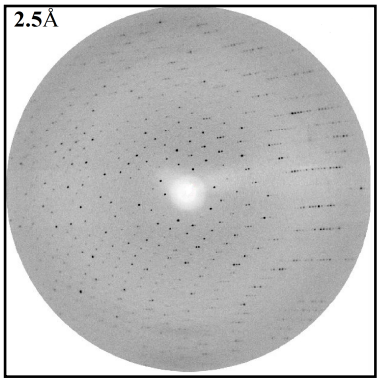

Figure 27: Diffraction image

Table 3: Statistics for the native crystal

Space group	$C222_1$
Unit-cell parameters (Å)	a = 88.0, b = 145.2, c = 218.9
Resolution range (Å)	50.0 – 2.46 (2.52 – 2.46)
Total Reflections	1311204 (3587)
Unique reflections	54602
Redundancy	7.8 (8.1)
Average I/σ (I)	38.6 (2.6)
R merge* (%)	8.4 (41.0)
Data Completeness (%)	98.8 (98.6)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

3.5.1 Sequence Homology

The already reported primary amino acid sequence [68] of the *Crocus vernus* showed a 33% sequence identity with *Scilla campanulata* agglutinin (SCA) [72].

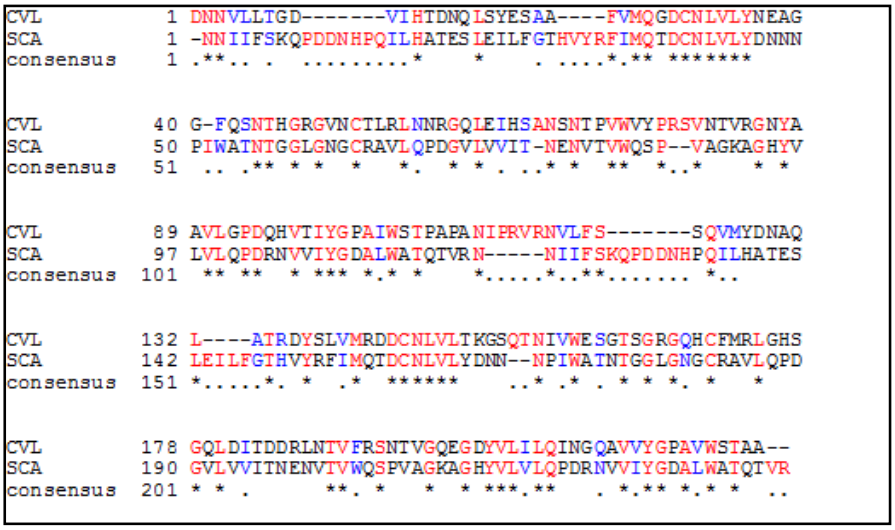

Figure 28: Sequence alignment between four subunits of *Crocus vernus* Lectin (CVL) and *Scilla campanulata* agglutinin (SCA). The figure was prepared using BOXSHADE server [79]

3.5.2 Structure Solution

Initially MR was applied to solve the phase problem using Molrep [73] from CCP4i suite [61]. A search model was constructed using the monomer of *Galanthus nivalis* (PDB entry, 1MSA) [74]. Different strategies were applied on this model e.g., using the original model as such, or the residual replacement with polyalanines or to use the most conserved part of the model after superposing different lectin models. Unfortunately statistical values were not satisfying and it was not possible to solve the phase problem by MR.

3.5.3 Heavy Metal Derivative Soaking

Lectin crystals were soaked with 1-5mM metal derivative solutions of Ethyl Mercuric Chloride, Mercuric Chloride, and Potassium tetrachloro palatinate II. The original drop size was 4µl and after adding the metal derivative solution, the final volume was 8µl. Soaking was performed overnight.

3.5.4 Metal Derivative Data Collection

Derivative diffraction data were collected on the tunable Consortium-Beam line X12, DESY Hamburg. The detector type used was a MarCCD 225 mm. Data collection was performed at 100 K and the ice production was prevented by soaking the crystal in

the same mother liquor with 20% Glycerol inside just before crystal mounting. A fluorescence scan (a Bruker/AXS (formerly Roentec) PX-XFlash energy-sensitive X-ray fluorescence detector) was used at the beam line to ensure the presence of the respective metal inside the crystals. Three data sets were collected at different wavelengths for the Mercuric Chloride soaked crystals at high energy remote, peak and low energy remote wavelengths. Native as well as derivative data were processed applying HKL 2000 using commands like ´´ANOMALOUS`` and ´´NO MERGE`` with modifier ´´original index``. The Multiple Isomorphous Replacement (MIR) strategy was applied using the SOLVE/RESOLVE crystallographic software.

Table 4: Statistics for the Mercury soaked crystal

Parameters	Low energy remote data
Space group	$C222_1$
Unit-cell parameters (Å)	a = 88.1, b = 144.2, c = 218.4
X-ray wavelength (Å)	1.01208
Resolution range (Å)	50.0 – 2.83 (2.91 – 2.83)
Energy (eV)	12250
Total Reflections	1189297 (2555)
Unique reflections	38550
Redundancy	6.9 (6.3)
Average I/σ (I)	12.8 (4.5)
R merge* (%)	11.5 (44.1)
Data Completeness (%)	99.3 (99.4)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 5: Statistics for the Mercury soaked crystal

Parameters	Peak data
Space group	$C222_1$
Unit-cell parameters (Å)	a = 88.4, b = 144.3, c = 219.3
X-ray wavelength (Å)	1.00581
Resolution range (Å)	50.0 – 2.70 (2.76 – 2.70)
Energy (eV)	12326
Total Reflections	1138291 (2573)
Unique reflections	38996
Redundancy	7.1 (6.0)
Average I/σ (I)	13.5 (2.5)
R merge* (%)	11.7 (41.8)
Data Completeness (%)	99.8 (100.0)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 6: Statistics for the Mercury soaked crystal

Parameters	High energy remote data
Space group	$C222_1$
Unit-cell parameters (Å)	a =89.8, b = 146.8, c = 222.6
X-ray wavelength (Å)	0.9918
Resolution range (Å)	50.0 – 2.80 (2.87 – 2.80)
Energy (eV)	12500
Total Reflections	1242789 (2394)
Unique reflections	36639
Redundancy	7.1 (5.7)
Average I/σ (I)	13.0 (3.9)
R merge* (%)	11.7 (47.9)
Data Completeness (%)	99.7 (99.8)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

3.5.5 Multiple Isomorphous Replacement (MIR)

The crystallographic software SOLVE/RESOLVE was applied for MIR but no sufficient phase information was generated.

3.5.6 Conclusions

In spite of the fact that the mercury sites have been detected by the software, still useful phase information was not generated because of the following possible reasons.

1. High number of molecules in the asymmetric unit
2. Heavy atoms did not bind at solid fixed positions

3.6 Search for New Lectin Crystals

Therefore, in search of new space group and cell dimensions, the monodispersive purified protein with optimized concentration was subjected to screening.

3.6.1 Robotic Screening

Robotic screening with the sitting drop method was performed using 480 different potential conditions in the NeXtal 96 well screening plates with 300 nanoliter spotted in each well. Five screening plates were placed at 20 ^{o}C and were analyzed after one week. A few potent conditions were identified after analyzing the plates with a visible microscope and a UV microscope to observe the fluorescence of protein crystals due to tryptophan presence in the primary sequence.

1) Condition: D12 4.0M Sodium Formate

2) Condition: D7 0.1M Hepes pH 7.5, 4.3M NaCl

3) Condition: B5 0.1M Sodium Cacodylate pH 6.5, 40%v/vMPD, 5%w/v PEG8000

4) Condition: D7 0.085M Mes pH 6.5, 1.7M Sodium Chloride, 15%Glycerol

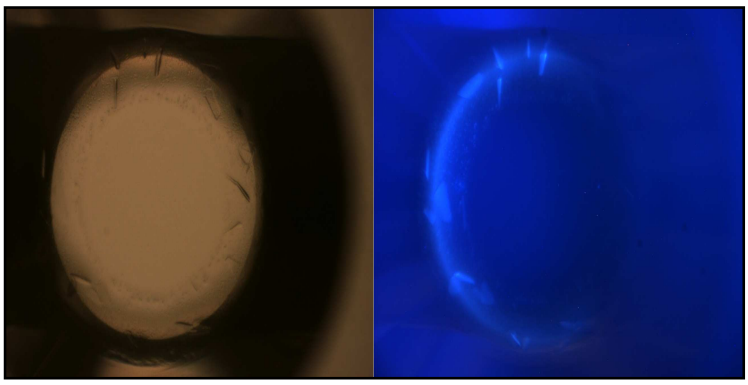

Figure 29: Visible and UV light microscopic photos showing crystals

3.6.2 Optimization of Lectin Crystallization

The lectin crystallization was optimized using hanging drop method with linbro plates and 2 + 2 µl drop size. The linbro plates like the screening plates were placed at 20 ^{o}C temperature. Good shaped crystals with suitable dimensions were grown using 4.0M Sodium Formate.

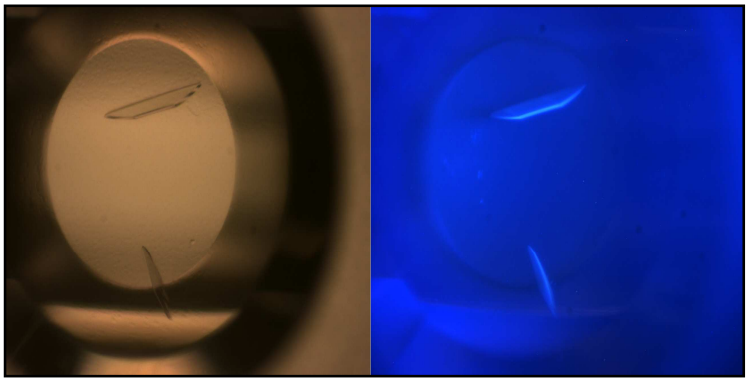

Figure 30: Visible and UV microscopic photos showing X-ray suitable crystals

3.7 Data Collection

Native diffraction data were collected at the Consortium-Beam line X13, DESY Hamburg. The detector type used was a Mar CCD 165 mm and the wavelength of the X-ray used was 0.8123 Å. Data collection was performed at 100 K and the ice production was prevented by soaking the crystal in the same mother liquor with 20% Glycerol inside just before crystal mounting. Sufficient data with a maximum resolution of 1.94 Å and a Phi of 1.0^{o} per image were collected. Again the data were processed using the HKL-2000

program package [60]. A new space group with smaller cell dimensions was found. The unit cell was primitive orthorhombic and the space group found is $P2_12_12$ with Cell dimensions a = 48.2 b = 98.0 c = 105.9 Å. The Mathews Coefficient V_M calculation [62] indicated the presence of one molecule in the asymmetric unit.

Table 7: Statistics of a native crystal

Space group	$P2_12_12$
Unit-cell parameters (Å)	a = 48.2 b = 98.0 c = 105.9
V_M ($Å^3$/ Da)	2.51
Solvent content (%)	51.06
Resolution range (Å)	30.0-1.94 (2.0 – 1.94)
Total Reflections	456372 (2695)
Unique reflections	41007
Redundancy	4.3 (4.3)
Average I/σ (I)	13.5 (3.84)
R merge* (%)	9.8 (43.2)
Data Completeness (%)	95.3 (95.6)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

3.7.1 Structure Solution

To solve the structure, phase information was tried to be obtained by Molecular Replacement (MR) using the programs AmoRe [75], Molrep [73] and Phaser [76] from CCP4i suite [61]. A search model was constructed using the monomer of *Galanthus nivalis* (PDB entry, 1MSA) [74]. Different strategies were applied on this model e.g., using the original model as such or the residual replacement with polyalanines or to use of the most conserved part of the model after superposing different Lectin models, but unfortunately statistical values were not satisfying and once again phase information could not be obtained.

3.7.2 Heavy Metal Derivative Soaking

A high number of Lectin crystals were soaked with 1-5mM metal derivative solutions as summarized in Table 1. The original drop size was 4µl and after adding the metal derivative solution, the final volume was 8µl. The soaking was performed overnight.

3.7.3 First Screening of Intact Unique Crystals

Since protein crystals because of their fragile nature are much prone to any typical destruction upon soaking with metal derivative solution, a preliminary screen was carried out on the in-house X-ray facility to check the quality of the soaked crystals. More than sixty Lectin crystals were exposed to X-rays. Few images were collected and processed for each crystal and around twenty crystals were found intact, which were immediately frozen in liquid nitrogen and kept in a Dewar.

3.7.4 Second Screening of X-ray Diffracting Crystals

The HKL Scenario 3 ´´Heavy Atom Search: Scale native plus a few frames of a potential derivative`` was used for a second screening in search of bound heavy atoms. All twenty frozen crystals were recovered through cryogenic techniques and were mounted on the Consortium-Beam line X13, DESY Hamburg. Three frames for each potential crystal with 1° Phi rotation were collected up to 3.5 Å resolutions. All data sets were processed through the HKL Package to produce the sca file and Mosflm [77] and Scala [78] to produce the electron density mtz files. The results of their statistical values are summarized in Table 8.

Table 8: Statistical values from screened heavy atom derivatives

No. of derivatives	Metal derivatives	Chemical formulae	Derivative data sca files production	Denzo Scenario 3: Chi square value $*\chi^2$
1	Mercuric chloride (1)	$HgCl_2$	Not possible	14.3
2	Mercuric chloride (2)	$HgCl_2$	yes	5.3
3	Mercuric chloride (3)	$HgCl_2$	yes	6.8
4	**Mercuric Acetate (1)**	**$(CH_3COO)_2Hg$**	**yes**	**9.4**
5	Mercuric acetate (2)	$(CH_3COO)_2Hg$	yes but poor data	3.2
6	Methyl mercuric chloride (1)	CH_3HgCl	Not possible	12.4
7	Methyl mercuric chloride (2)	CH_3HgCl	yes but poor data	3.9
8	Potassium tetrachloro platinate II (1)	K_2PtCl_4	yes	152.3
9	**Potassium tetrachloro platinate II (2)**	**K_2PtCl_4**	**yes**	**8.1**
10	**Potassium hexachloro platinate**	**K_2PtCl_6**	**yes**	**9.1**

* $\chi^2 \approx 10$ indicates a potential derivative

Successful candidates from the second screen were again preserved in liquid nitrogen applying cryogenic techniques.

3.7.5 Heavy Atom Derivative Data Collection

Derivative diffraction data were collected in approx. 36 hours after mounting the frozen screened crystals at the tunable Beam Line X12, DESY Hamburg. The detector type used was a MarCCD 225 mm. Data collection were performed at 100 K. The crystal was cryo protected before crystal mounting.

Table 9: Statistics of Mercuric acetate I soaked crystal

Parameters	High energy remote data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.4, b = 97.5, c = 105.0
X-ray wavelength (Å)	0.83550
Resolution range (Å)	30.0 – 2.05 (2.10 – 2.05)
Total Reflections	649835 (2266)
Unique reflections	34707
Redundancy	6.9 (6.8)
Average I/σ (I)	14.7 (5.0)
R merge* (%)	10.1 (42.7)
Data Completeness (%)	100.0 (100.0)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

A fluorescence scan (applying a Bruker/AXS (formerly Roentec) PX-XFlash energy-sensitive X-ray fluorescence detector) was performed at beamline to ensure the respective heavy atom inside the crystal. Two data sets were collected at different wavelengths for the Mercuric Chloride soaked crystals at high energy remote and peak wavelengths. Native as well as derivative data were processed applying HKL 2000 program using commands as ´´ANOMALOUS`` and ´´NO MERGE`` with modifier ´´original index``. Next Multiple Isomorphous Replacement (MIR) was applied using the SOLVE/RESOLVE crystallographic software.

Table 10: Statistics of Mercuric acetate I soaked crystal

Parameters	Peak data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.4, b = 97.5, c = 105.0
X-ray wavelength (Å)	1.00300
Resolution range (Å)	30.0 – 2.00 (2.05 – 2.00)
Total Reflections	672796 (2265)
Unique reflections	34749
Redundancy	7.3 (7.1)
Average I/σ (I)	24.9 (5.1)
R merge* (%)	8.8 (42.1)
Data Completeness (%)	99.6 (100.0)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i |Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 11: Statistics of the Potassium hexachloroplatinate soaked crystal

Parameters	High energy remote data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.2, b = 97.8, c = 106.6
X-ray wavelength (Å)	0.88550
Resolution range (Å)	30.0 – 2.59 (2.66 – 2.59)
Total Reflections	306241 (1367)
Unique reflections	20567
Redundancy	7.4 (7.2)
Average I/σ (I)	11.6 (4.85)
R merge* (%)	15.2 (44.3)
Data Completeness (%)	99.9 (100.0)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i |Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 12: Statistics of the Potassium hexachloroplatinate soaked crystal

Parameters	Peak data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.2, b = 97.8, c = 106.6
X-ray wavelength (Å)	1.06690
Resolution range (Å)	30.0 – 2.52 (2.59 – 2.52)
Total Reflections	300382 (1335)
Unique reflections	20564
Redundancy	7.0 (7.0)
Average I/σ (I)	13.0 (5.1)
R merge* (%)	13.7 (44.4)
Data Completeness (%)	99.9 (100.0)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i |Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 13: Statistics of the Potassium tetrachloroplatinate II soaked crystal

Parameters	High energy remote data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.3, b = 97.6, c = 105.3
X-ray wavelength (Å)	0.83550
Resolution range (Å)	30.0 – 2.40 (2.46 – 2.40)
Total Reflections	498310 (1299)
Unique reflections	20320
Redundancy	7.0 (6.6)
Average I/σ (I)	12.9 (3.95)
R merge* (%)	12.8 (49.4)
Data Completeness (%)	99.9 (99.1)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i |Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

Table 14: Statistics of the Potassium tetrachloroplatinate II soaked crystal

Parameters	Peak data
Space group	$P2_12_12$
Unit-cell parameters (Å)	a =48.3, b = 97.7, c = 105.3
X-ray wavelength (Å)	1.06512
Resolution range (Å)	30.0 – 2.52 (2.59 – 2.52)
Total Reflections	473355 (1315)
Unique reflections	20106
Redundancy	6.6 (5.7)
Average I/σ (I)	11.4 (3.27)
R merge* (%)	12.8 (46.8)
Data Completeness (%)	98.7 (98.2)

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_i | Ii\ \{hkl\} - \{I\ (hkl)\}\ | / \sum_{hkl}\sum_i Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

3.7.6 Phase Problem Solving Strategies

To combat the real bottleneck of the phase problem, following different approaches were applied using the above collected.

1. Multiple Isomorphous Replacement (MIR)
2. Single wavelength Anomalous Dispersion (SAD)
3. Two wavelength-Multi-wavelength Anomalous Dispersion (2W-MAD)
4. Single Isomorphous Replacement with Anomalous Signal (SIRAS)

SAD, 2W-MAD and SIRAS strategies were applied using the online web server of the program Auto Rickshaw provided by EMBL Hamburg, while MIR was applied through SOLVE/RESOLVE. The statistical values evaluating the success of the strategy are described in the Table 15.

Table 15: Summary of phasing statistics

Method Name	strategy	Statistical value	Refmac5	Comments
Multiple Isomorphous Replacement (MIR)	Solve/Resolve	Resolve. log FOM = 65 %	(R-factor/Rfree) 2.3 Å 0.427 / 0.478 FOM= 80 %	Very good
Single wavelength Anomalous Dispersion (SAD)	Auto rickshaw web server SHELXD	CC(All) 23.47 CC(weak) 12.90	(R-factor/Rfree) 2.3 Å 0.5769 / 0.5633 FOM = 18.2%	not good
Single Isomorphous Replacement with Anomalous Signal (SIRAS)	Auto rickshaw web server SHELXD	CC(All) 34.08 CC(weak) 24.51	(R-factor/Rfree) 2.3 Å 0.5360 / 0.5498 FOM = 27.6%	can improve
2 wavelength-Multi wavelength Anomalous Dispersion (2W-MAD)	Auto rickshaw web server SHELXD	CC(All) 27.94 CC(weak) 21.79	(R-factor/Rfree) 2.3 Å 0.5645 / 0.5780 FOM = 19.6%	not good

3.7.7 Multiple Isomorphous Replacement (MIR)

Phase information was produced successfully and a high quality electron density map was displayed using program Coot.

3.8 Model Building and Refinement

Substantial phase information was generated mainly via the platinum sites providing good statistical values. Model building and refinement was performed using the program Coot and Refmac5.

Table 16: Data collection and refinement statistics for the Lectin solution

Parameters	Lectin
Space group	$P2_12_12$
a (Å)	48.4
b (Å)	98.1
c (Å)	105.8
VM ($Å^3$/ Da)	2.5
Solvent content (%)	51
Completeness of data (%)	85.3 (85.6)
No. of total reflections	456372
No. of unique reflections	48106
Average I/sigma intensity	13.5 (3.84)
Resolution (Å)	30.0-1.94
Redundancy	4.3 (4.3)
*Rmerge (%)	9.8 (33.8)
No. of reflections used in refinement	1647
Rcrystal (%)	17.3 (20.3)
No. of reflections used in Rfree	100
Rfree (%)	22.5 (25.5)
Protein atoms	3368
Solvent atoms	394
Average B-factor ($Å^2$)	
Main-chain atoms	23.3
Side chain atoms	24.1
Solvent molecules	37.2
Other atoms ($HCOO^{-1}$)	43.1

Root mean square deviation	
Bonds (Å)	0.018
Bond angles (^{o})	1.845
Residues in regions of the Ramachandran plot (%)	
Most favored	82.1
Allowed	15.5
Disallowed	0.0
Generally allowed	1.9

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_{i}|Ii\ \{hkl\}-\{I\ (hkl)\}|/\sum_{hkl}\sum_{i}Ii\ (hkl)$, where $\{I\ (hkl)\}$ is the mean intensity of the observations $Ii\ (hkl)$ of reflection hkl.

3.8.1 Amino acid Sequence of *Crocus Vernus* Lectin (CVL)

The complete amino acid sequence deduced from the molecular structure of Lectin is given below in the form of single letter code.

Chain A

DNNVLLTGDV IHTDNQLSYE SAAFVMQGDC NLVLYNEAGG FQSNTHGRGV NCTLRLNNRG QLEIHSANSN TPVWVYPRSV NTVRGNYAAV LGPDQHVTIY GPAIWSTPAPA

Chain B

NIPRVRNVLF SSQVMYDNAQ LATRDYSLVM RDDCNLVLTK GSQTNIVWES GTSGRGQHCF MRLGHSGQLD ITDDRLNTVF RSNTVGQEGD YVLILQINGQ AVVYGPAVWSTAA

Chain C

DNNVLLTGDV IHTDNQLSYE SAAFVMQGDC NLVLYNEAGG FQSNTHGRGV NCTLRLNNRG QLEIHSANSN TPVWVYPRSV NTVRGNYAAV LGPDQHVTIY GPAIWSTPAA

Chain D

NIPRVRNVLF SSQVMYDNAQ LATRDYSLVM RDDCNLVLTK GSQTNIVWES GTSGRGQHCF MRLGHSGQLD ITDDRLNTVF RSNTVGQEGD YVLILQINGQ AVVYGPAVWS TA

3.9 Lectin-Mannose Complex

Data of a Lectin crystal soaked with 100mM Mannose were collected at the Consortium-Beamline X 13 and processed with the appropriate softwares to generate the map showing the electron density of Mannose sugar.

Table 17: Data collection and refinement statistics for the Lectin – Mannose complex

Parameters	Lectin/Mannose
Space group	$P2_12_12$
a (Å)	48.5
b (Å)	98.6
c (Å)	105.6
VM ($Å^3$/ Da)	2.4
Solvent content (%)	49
Completeness of data (%)	75.3 (76.7)
No. of total reflections	186208
No. of unique reflections	17436
Average I/sigma intensity	12.0 (3.67)
Resolution (Å)	30.0-2.55
Redundancy	4.6 (4.4)
Rmerge (%)	11.1 (45.3)
No. of reflections used in refinement	1768
Rcrystal (%)	18.5 (27.8)
No. of reflections used in Rfree	96
Rfree (%)	23.3 (33.4)
Protein atoms	3368
Solvent atoms	334
Average B-factor ($Å^2$)	
Main-chain atoms	25.58
Side chain atoms	26.97
Solvent molecules	40.29

Other atoms ($HCOO^{-1}$)	47.73
Root mean square deviation	
Bonds (Å)	0.024
Bond angles (^{o})	2.00
Residues in regions of the Ramachandran plot (%)	
Most favored	84.0
Allowed	14.3
Disallowed	0.0
Generally allowed	1.5

Values in parentheses are for the highest resolution shell.

*R merge $=\sum_{hkl}\sum_{i}|Ii\{hkl\}-\{I(hkl)\}|/\sum_{hkl}\sum_{i}Ii(hkl)$, where $\{I(hkl)\}$ is the mean intensity of the observations $Ii(hkl)$ of reflection hkl.

3.10 Description of the Molecule

Lectins are sugar specific proteins which can bind to various cell surface glycoproteins influencing different cellular functions like plant defense mechanisms [80]. Mannose-binding lectins are a well conserved super-family of lectins found in the monocotyledonous families. Normally these bulb lectins are very sugar specific and have conserved structural homologies. However still these vary in their biological roles e.g., Amaryllidaceae and Orchidaceae lectins are known to be potent inhibitors of HIV and other retroviruses [81] while Alliaceae lectins can bind to glycoproteins like invertases and allinases [82]. Hence the studies of 3-dimensional structures of lectins regarding their biological functional variations are of great importance. *Crocus vernus* Lectin (CVL) purified from *Crocus* bulbs is a monocot mannose-binding agglutinin and was characterized to be a tetramer comprising the four chains A, B, C and D with two dimers in the asymmetric unit. A tetramer is composed of two homo-dimers with chains A and B as one dimer and with chains C and D forming the other dimer. The two chains of a dimer share a primary amino acid sequence homology of around 38 %.

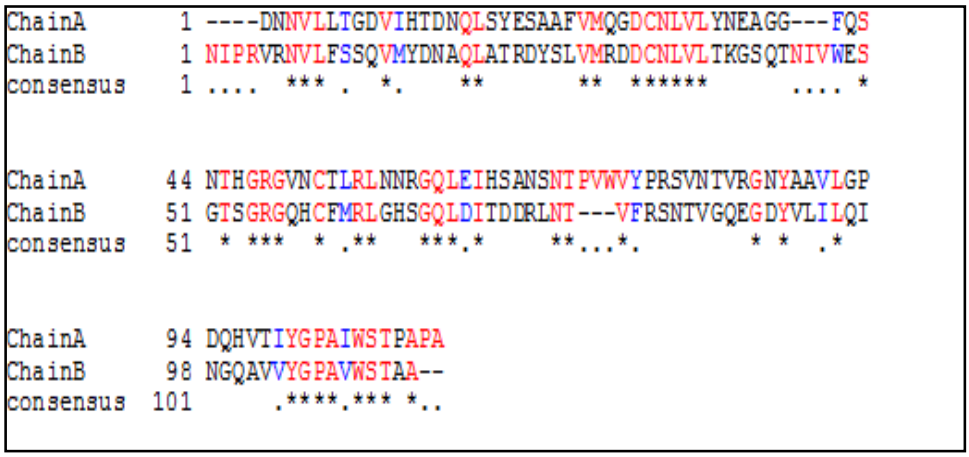

Figure 31: Sequence alignment between two chains of a CVL dimer

However the C-termini were flexible and weak or no electron density was observed. A total of 446 amino acids have been detected and sequenced with the exception of some surface residues from chains B and D, which are Ser42, Gln43 and Thr44 with no electron density. It should be made clear that the purified protein moved as two bands on a reducing SDS gel, clearly indicated two different molecular weight chains from the single dimer as already described above. However due to the flexibility of the C-termini almost an equal numbers of residues has been identified in the electron density. The gel pattern clearly showed that two chains of one dimer have different number of residues. The full length amino acid sequence for CVL has already been reported via cDNA clones [68] indicating a different number of chain residues in the dimers. Some of the primary sequence statistics has been summarized in the following table. The statistics has been prepared applying the ExPASy Proteomics Server [70].

Table 18: Statistical values of the primary amino acid sequence

Chain ID	Chain A	Chain B	Chain C	Chain D	Overall
Formula	$C_{525}H_{806}N_{154}O_{166}S_3$	$C_{543}H_{858}N_{160}O_{171}S_5$	$C_{525}H_{806}N_{154}O_{166}S_3$	$C_{543}H_{858}N_{160}O_{171}S_5$	$C_{2128}H_{3310}N_{626}O_{669}S_{16}$
No. of Amino acids	111	113	111	113	446
Molecular Weight (Da)	12027.2	12524.0	12027.2	12524.0	48880.4
Theoretical pI	5.66	6.06	5.66	6.06	5.81
Total No. of atoms	1654	1737	1654	1737	6749
No. of negative residues	8	10	8	10	36
No. of positive residues	5	9	5	9	28
No. of Cysteines	2	2	2	2	8

The final symmetry averaged MIR electron density map showed at some positions a different amino acid sequence than the published data for *Crocus vernus* agglutinin (CVA) [68], indicating an isoforms for this protein in the corms. The maximum sequence identity between CVL and CVA is approx. 90%. A similar homology relationship is reported for CVL and *Crocus sativus* agglutinin (CSA) [83]. Apart from the flexible C-

termini and some surface residues as indicated above, the rest of the sequence clearly showed deviations from the published data confirming the lectin to be a hetero-tetramer, as all the four chains are different from each other. Our data indicated different findings as isolectin CVL to be a tetramer with two homo-dimers e.g., in chains A & B, residues Arg59, Val80 are identical while reported sequence [68] is showing two different amino acids for both of these chains.

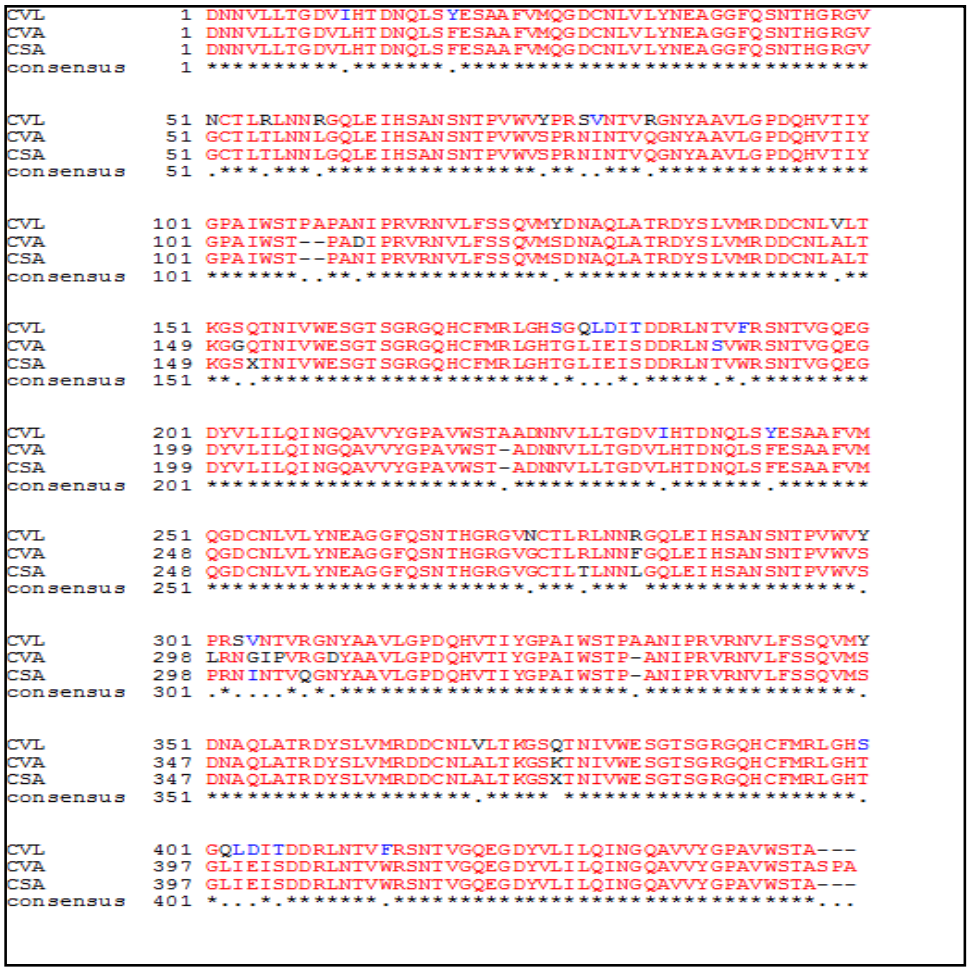

Figure 32: Sequence alignment between the four subunits of *Crocus vernus* Lectin (CVL) and *Crocus vernus* agglutinin (CVA) [68]. The flexible C-termini of CVA have been omitted.

Therefore it seems that CVL is comprised of two homo-dimers rather than a hetero-tetramer with four different chains. The deviations between the two sequences of *Crocus vernus* and also one from *Crocus sativus* are summarized in the Figure 32. These observations are explainable, if there are multiple lectins with little sequence differences.

As described before most of the deviations between the sequences are clearly identified in the electron density map.

A total of 339 bound water molecules, three glycerol molecules and eight formate ($HCOO^-$) molecules were identified in the refined 2 F_o-F_c electron density map.

The CVL tetramer in the asymmetric unit exhibits approximates 222 point symmetry (Figure 33).

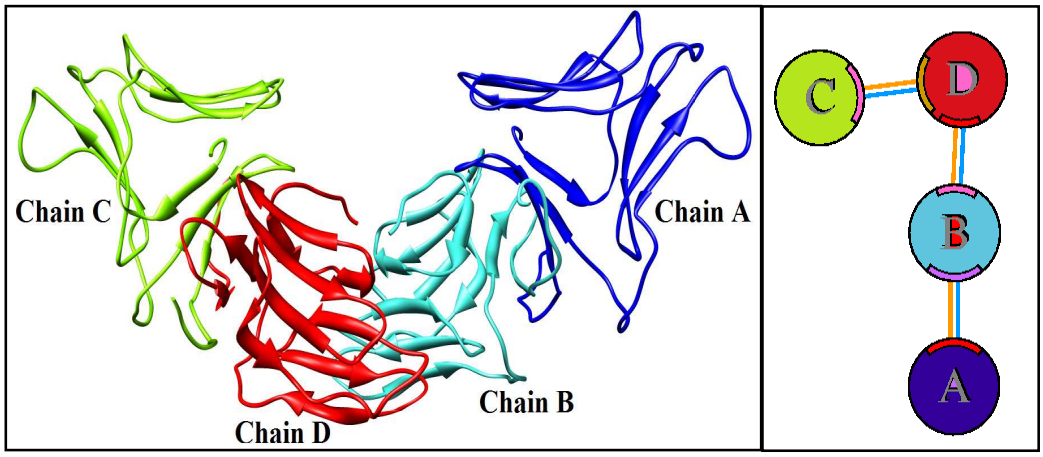

Figure 33: Tetrameric *Crocus vernus* Lectin (CVL) molecule. Prepared with Chimera [84]

The four chains are arranged in the forms of rings comprised of an octamer containing approximately a central 22 Å wide solvent channel (Figure 34). The outer surface of the tetramer and the hydrophilic nature of the channel is mainly build by Arg, Tyr and Pro residues. Both monomers of a dimer consist of three bundles of β-sheets which are oriented perpendicularly to the axis of the monomers. Each of these sheets forms one of the three domains. The three sequential sub domains are connected by loops to form a 12-stranded β-barrel in chains A & C. A12-stranded β-barrel is found in chains B and D. The three sub-domains are flat three-stranded, anti parallel β-sheets. Sub domain I contains both a N-terminal and C-terminal of the molecule and forms the interfacial region between subunits A and B (or C and D) (Figure 35). Its outer strand is donated from the dimer-related subunit as a result of C-terminal exchange. The proximity between the N-terminal end and this donated C-terminus strand from the dimer-related subunit leads one to speculate whether these monomers may have been covalently linked early during the evolution of this tetrameric molecule [74]. The centre of the β-barrel is filled with hydrophobic side chains (Figure 36). In chains A and C, they are Leu 5/Ile

11/Leu 17/Phe 24/Met 26/Leu 32/Leu 34/Phe 41/Leu 54/Leu 62/Ile 64/Trp 74/Try 76/Val 97, while in chains B and D they are Leu 9/Leu 28/Leu 36/Met 30/Trp 48/Met 61/Leu 63/Ile 71/Leu 93/Val 103.

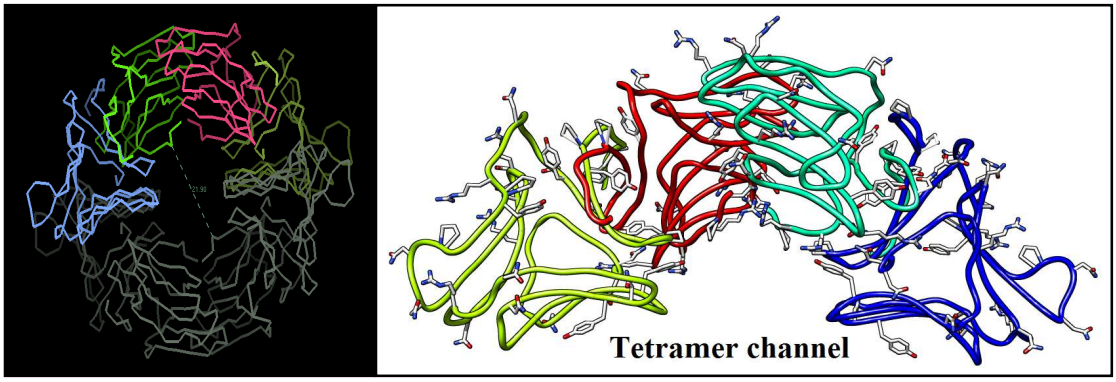

Figure 34: *Crocus vernus* Lectin (CVL) showing an octameric ring forming approx. 22 Å wide solvent channels

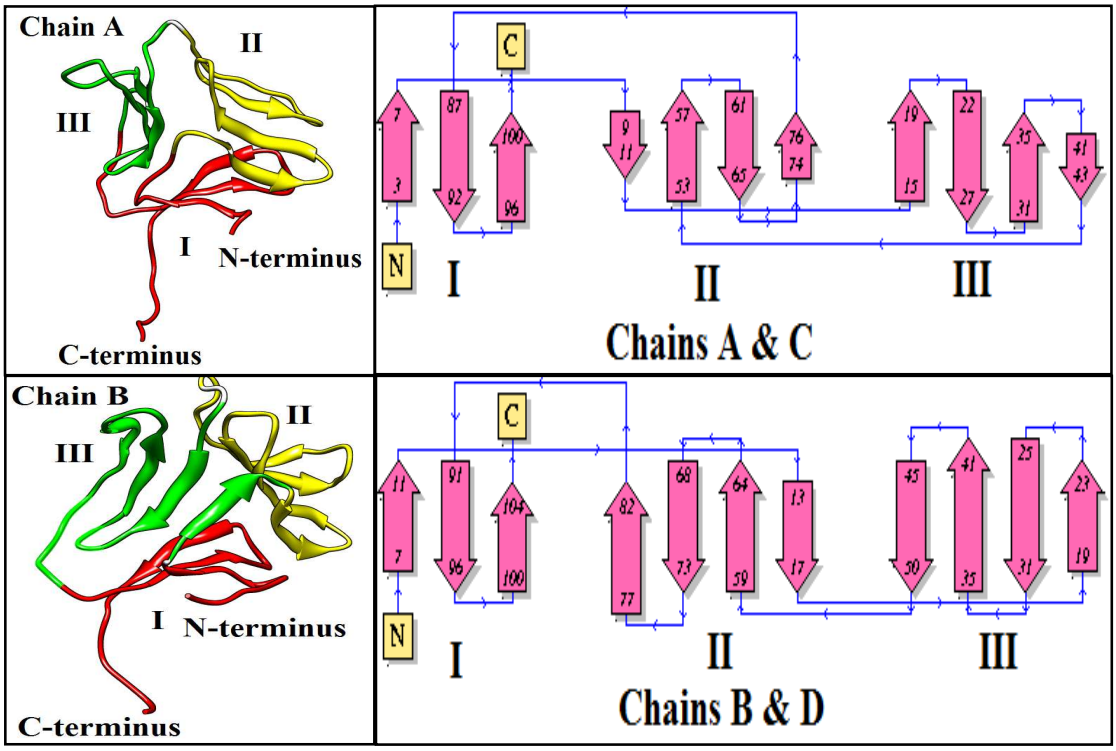

Figure 35: *Crocus vernus* Lectin (CVL) Chains A (C) & B (D) showing three sub -domains I, II & III.

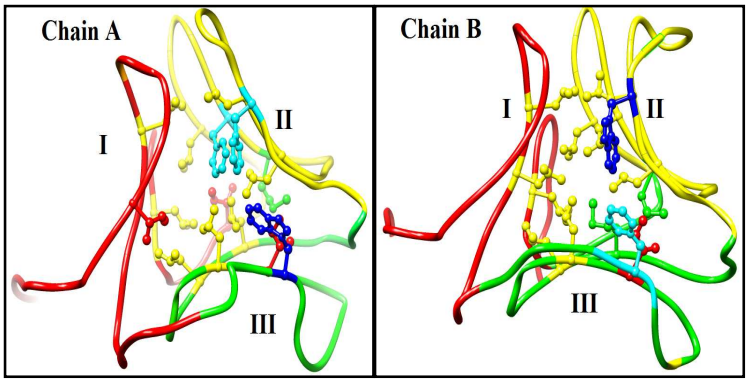

Figure 36: Chains A & B showing hydrophobic amino acids side chains filling the β-barrel channel

3.10.1 Disulfide Bonds

The protein is rich in Val, Asn, Gly and Leu and contains eight Cys residues in total, which form S-S disulfide bridge. There is one intra-chain disulfide bond in each chain providing stability to the chain. In chains A and C, the single intra chain disulphide bridge between Cys 30 and Cys 52 forms a link between sub domain II and III. The same is present for chain B and D between Cys34 and Cys 59 between sub domains II and III.

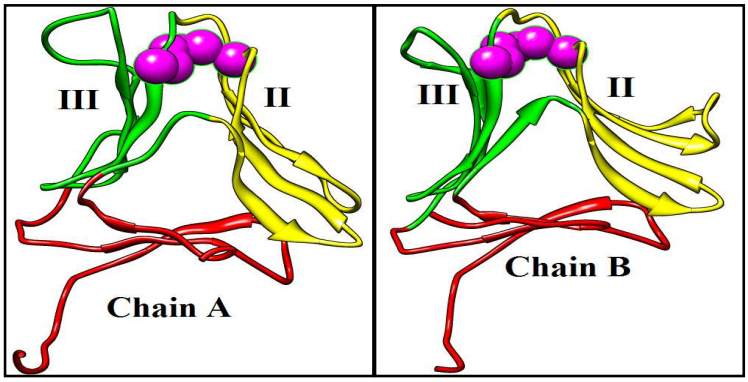

Figure 37 A & B: Disulfide bridges of chain A and B

The two chains of a dimer are stabilized through numerous hydrogen-bonds contacts and are strengthened additionally due to exchange of C-terminal strands between them in such a way that residues 96-112 of one chain cross over into other chain. A comparison of residues based on the interpretation of the *2Fo-Fc* electron density map revealed a loss of approx. 10-12 amino acids at the C-termini.

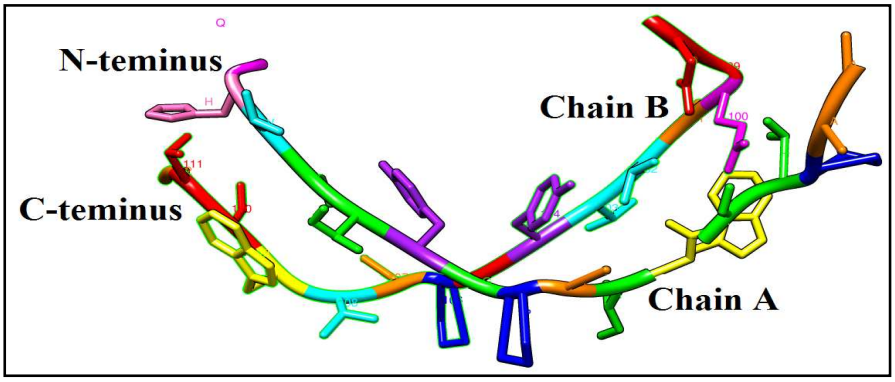

Figure 38: C-termini cross link of two chains of a dimer

3.10.2 Tertiary Structure

Although the primary sequence homology between CVL, SCA and *Galanthus nivalis* agglutinin GNA [74] is only 30% but still the structure homology is quite similar to each other for each of the CVL subunit. This indicates that this particular conserved structural fold is typical for all members of the bulb lectins super family [74]. The β-prism II fold with an approximate internal 3-fold axis is composed of three sub-domains with antiparallel β-sheets giving an appearance of a triangular prism. Since the strands are perpendicular to the 3-fold axis, the fold is different to the known fold of Jacalin specific lectins [85] where the strands are arranged parallel to the 3-fold axis [86]. These structural motifs have few common features but they possess neither much sequence homology nor any similarity in their functions [87].

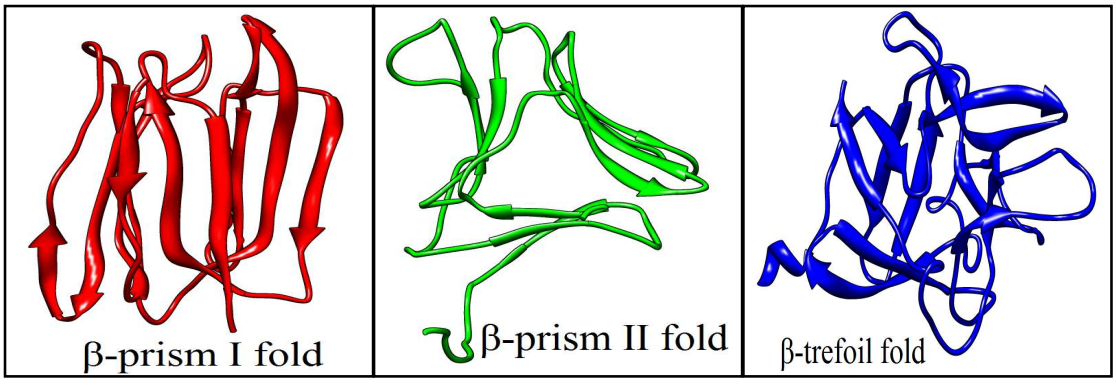

Figure 39: β-prism I fold: Jacalin monomer (PDB code 1JAC); β-prism II fold: CVL monomer (PDB code 3MEZ); β-trefoil: Amaranthin monomer (PDB code 1JLX)

3.10.3 Modes of Dimerization and Quaternary Structure

The four subunits of the CVL tetramer associated with 222 point symmetry showed numerous non-covalent interactions across the interfaces of the respective chains of the respective dimers. These are the hydrogen bonds between amino acid side-chains which actually stabilized the ultimate quaternary structure of the protein. An overall view of these interactions for the whole molecule can be seen in the figure below.

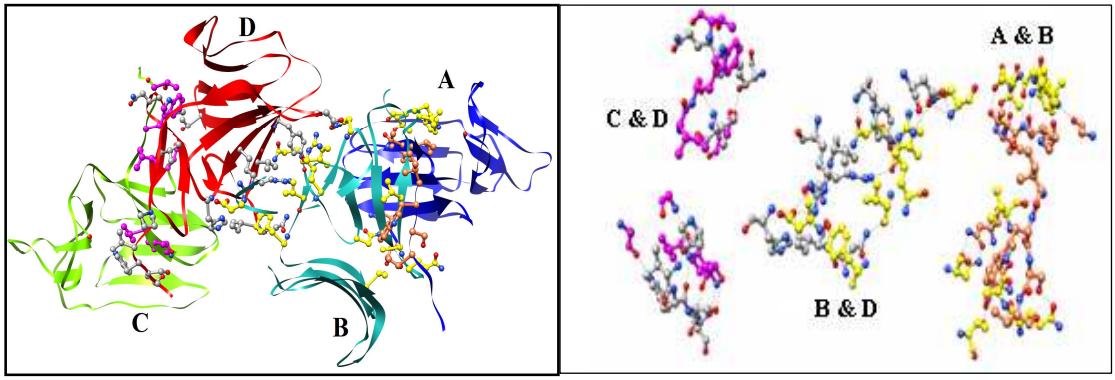

Figure 40: Tetramer showing three distinct patches of hydrogen bonds between the four chains

Hydrogen bonds are clustered in three distinct patches which mean they are mainly present on the interfaces of the C-termini of chains AB and CD forming their respective dimers and more importantly between the core of chains B and D resulting in the formation of a quaternary tetrameric structure. However it should also be made clear that to a lesser extent the N-termini have also been involved in the development of these hydrogen bonds. In addition, this typical pattern of interactions can easily explained on structural grounds the fact that why some members of this lectin family are dimers while others are still tetramers.

Since dimers AB and CD are identical, therefore we experienced identical number of the hydrogen bonds within these two dimers where the C-termini moved across each other giving rise to enormous number of these non-covalent interactions as depicted in the figure below.

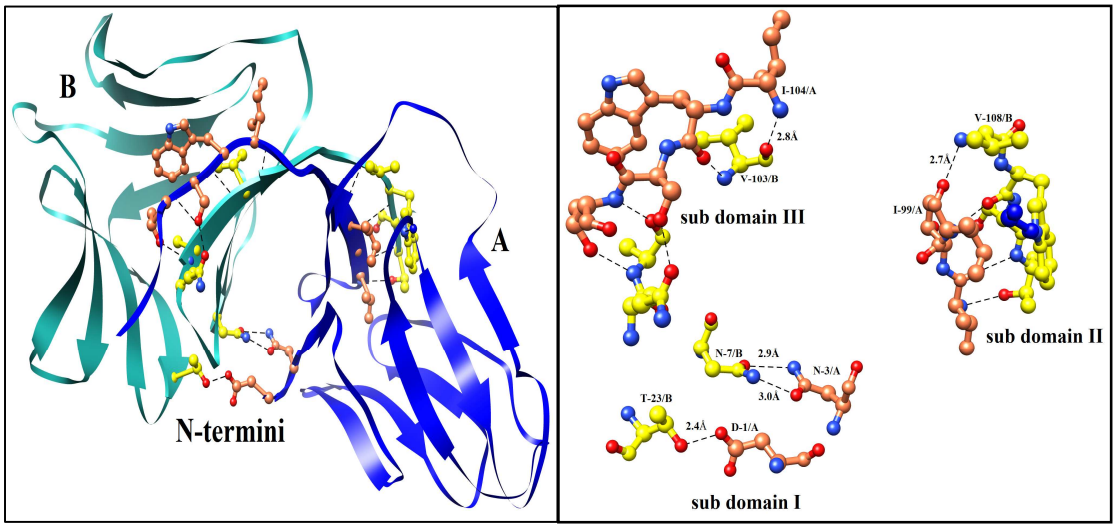

Figure 41: Figure showing the hydrogen bonded side chains and backbone amino acids of N & C-termini of the chains A & B (or C & D)

Chains B and D were identical chains and almost the same number and similar amino acid side chains from both sides developed the similar hydrogen bonded environment giving real contribution to the tetramer development.

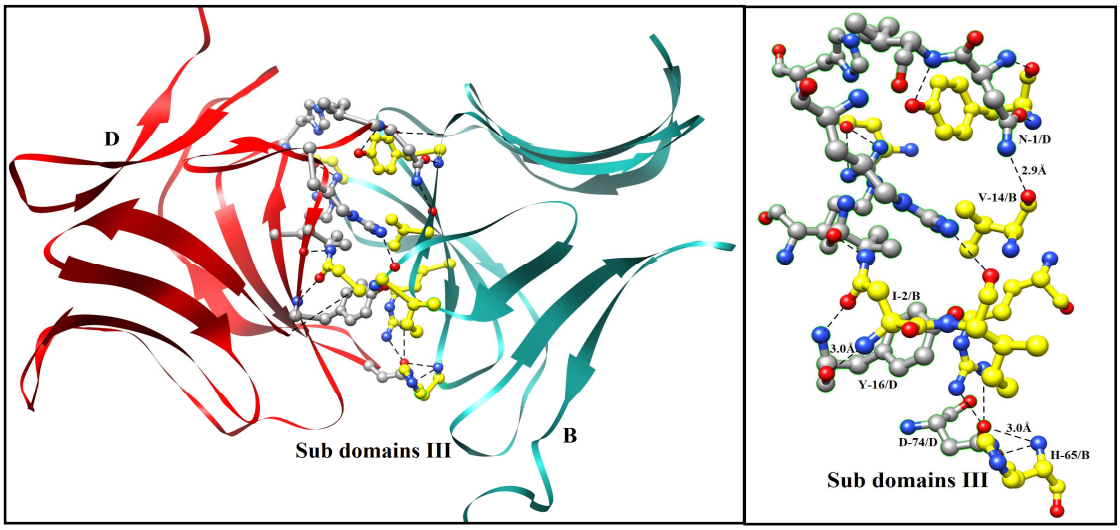

Figure 42: Figure showing the hydrogen bonded interactions between side chains and backbone amino acids of chains B & D

The total number of these hydrogen bonds between the amino acid side chains or backbone, contributing to the structure stabilization, is summarized in the following tables.

Table 19: Hydrogen bonds interactions across the interface of chains A & B (or C & D)

Nr.	Residue name	Atom name	Chain name		Residue name	Atom name	Chain name	Distance (Å)
1.	D-1	OD2	A	←→	T-23	OG1	B	2.7
2.	N-3	OD1	A	←→	N-7	ND2	B	2.9
3.	N-3	ND2	A	←→	N-7	OD1	B	3.0
4.	G-60	O	A	←→	W-109	NE1	B	2.9
5.	V-97	N	A	←→	T-111	OG1	B	3.0
6.	V-97	O	A	←→	T-111	N	B	2.9
7.	T-98	OG1	A	←→	S-110	OG	B	3.0
8.	T-98	OG1	A	←→	S-110	OG	B	3.0
9.	I-99	N	A	←→	W-109	O	B	2.7
10.	I-99	O	A	←→	V-108	N	B	3.0
11.	G-101	N	A	←→	P-106	O	B	2.8
12.	I-104	N	A	←→	V-103	O	B	2.8
13.	W-105	O	A	←→	V-103	N	B	2.9
14.	S-106	OG	A	←→	Q-100	OE1	B	3.0
15.	T-107	N	A	←→	A-101	O	B	2.9
16.	T-107	OG1	A	←→	A-101	N	B	2.8

In chains A and B (C & D), apart from the hydrogen bond interactions between the side chains of respective amino acids or between side chain and main backbone, few hydrogen bond interactions have been observed between the atoms of the main backbone chain. In particular interactions towards the C-termini are mostly of this type.

Table 20: Hydrogen bond interactions across the interface of chains B & D

Nr.	Residue name	Atom name	Chain name		Residue name	Atom name	Chain name	Distance **(Å)**
1.	N-1	N	B	←→	T-16	O	D	3.0
2.	N-1	OD1	B	←→	T-16	N	D	3.0
3.	N-1	ND2	B	←→	V-14	O	D	3.0
4.	I-2	N	B	←→	T-16	OH	D	3.0
5.	I-2	O	B	←→	R-4	NH1	D	2.8
6.	V-14	O	B	←→	N-1	ND2	D	2.9
7.	T-16	N	B	←→	N-1	OD1	D	2.9
8.	T-16	O	B	←→	N-1	N	D	3.0
9.	T-16	OH	B	←→	I-2	N	D	3.0
10.	R-62	NH1	B	←→	D-74	OD1	D	3.0
11.	R-62	NH2	B	←→	D-74	OD1	D	2.7
12.	H-65	N	B	←→	D-74	OD1	D	3.0
13.	D-74	OD1	B	←→	R-62	NH1	D	3.0
14.	D-74	OD1	B	←→	R-62	OH2	D	2.9
15.	D-74	OD1	B	←→	R-62	N	D	2.9

On the contrary, different type of hydrogen bonds has been found between chains B and D. There is only one hydrogen bonding between the atoms of the main backbone of the chains B and D e.g., Asparagine (N-1) of the chain B and Threonine (T-16) of chain D. The rest of the interactions are either found between the atoms of the side chains or between the atoms of side chain and the main backbone. The other difference from the inter-chain interactions of chains AB and CD, here no C-termini have been involved for the production of these interactions, but it is mainly from the inside of the primary sequence. Although it has already been discussed and clarified that the C-termini of all four chains are already been engaged in forming hydrogen bonds between the respective chains like chain A with B and chain C with D.

3.10.4 Structural Comparisons

It has already been shown that the primary sequence of CVL possess less than 40% sequence homology with the already reported Monocot mannose-binding lectins already deposited in the protein data bank. However, there is a relative structural similarity exist in this class of lectins even though the sequence homology is less.

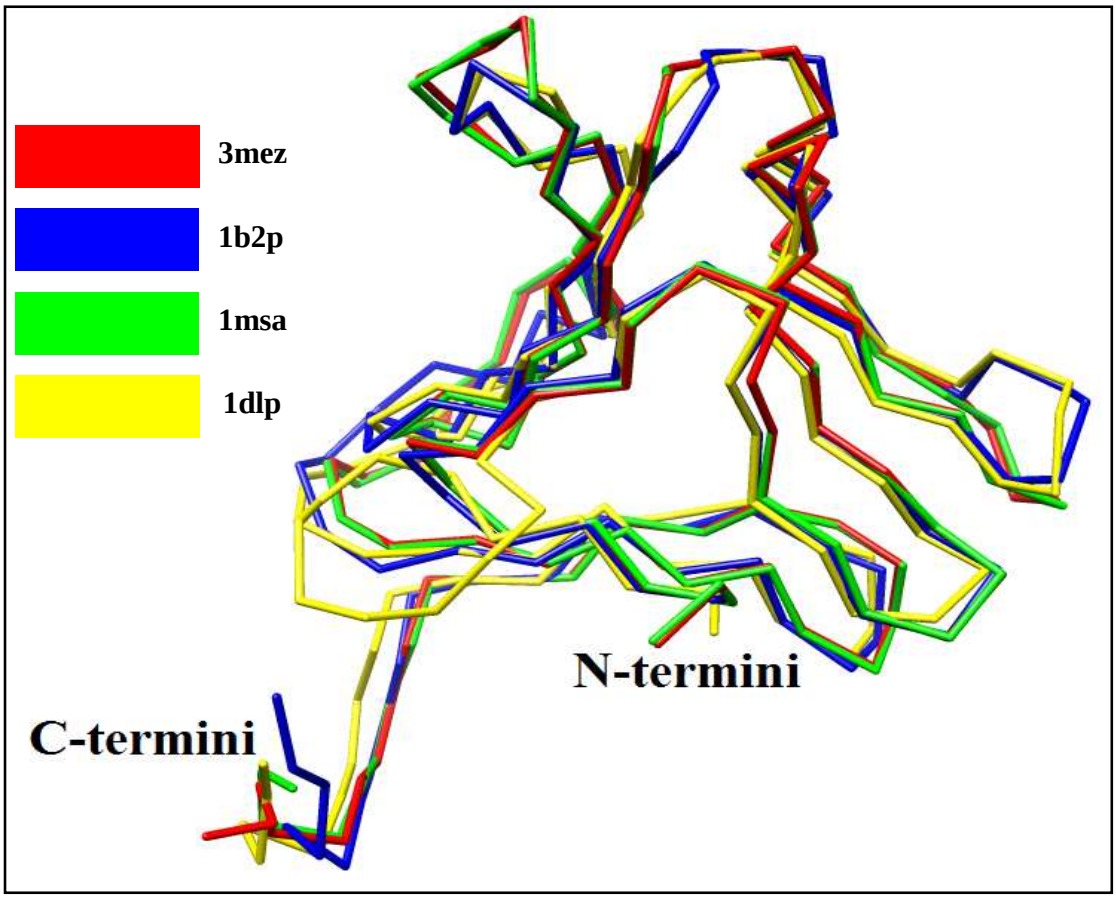

Figure 43: Superposed Monocot mannose-binding lectins showing structural similarity

In figure 43, four lectins from different sources have been superimposed to each other to show the high structural similarity between them. The Cα plot for structues has been shown and most notably is the 3mez which is the pdb code for CVL. The figure has been prepared through the online webserver ´´SuperPose`` [93] and Chimera. The superposition of the Cα atoms of 3mez with that of other members of the family gives a rmsd of 1.7, 0.8 and 1.7 Å for models1b2p, 1msa and 1dlp respectively.

The sequence homology between *Scilla campanulata* agglutinin (SCA) [72] and CVL is approx. 33%, but the structural similarity is considerably high. SCA is a homo-

dimer and both the chains are stabilized through the hydrogen bonds of the C-termini exchange with each other. This is typical to the lectins of this class and also as experienced in the case of CVL. SCA is considered to be a polypeptide of 244 amino acid residues and is folded into two distinct domains, 1 and 2, joined by a linker peptide which is difficult to confirm in the CVL structure, as the C-termini are not appearing and last 10-12 residues are missing.

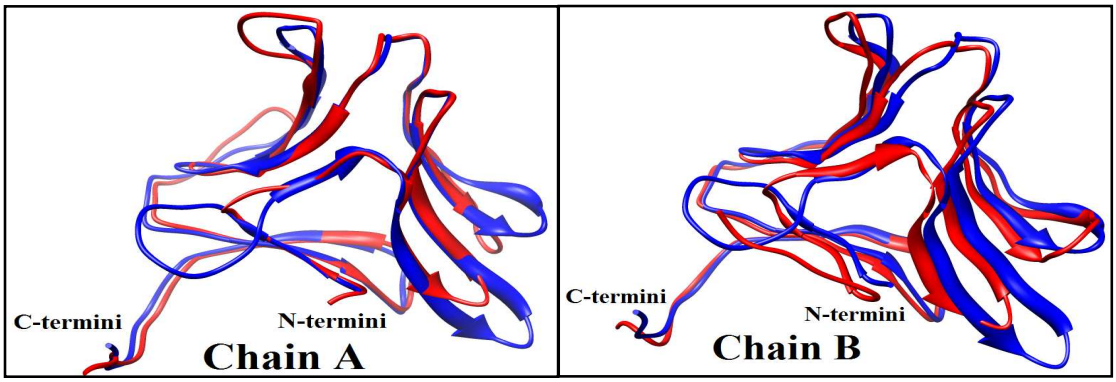

Figure 44: Structural superposition between CVL (3mez) and SCA (1b2p) proteins

4 SUMMARY AND OUTLOOK

Crocus vernus is specie of the *Crocus* genus in the Iridaceae family. As a result of evolutions, Plants have developed a defense mechanism of producing Pathogenesis-related proteins in their underground reproductive parts like seeds, bulbs, corms etc. These proteins are specifically induced during the pathological or related situations to combat the attacks of pathogenic organisms.

Chitinases and Lectins are two strong defense-related proteins of higher plants. Chitinases can digest glycoside bonds in chitin and as a result are of much importance against the fatal attacks of Fungi, insects etc. Similarly Lectins are sugar-binding proteins occurring in almost all types of life forms. Monocot mannose-binding lectins can bind to wide variety of micro-organisms and also found their importance in malignant cells therapy.

A family 18 chitinase from crocus bulbs has been successfully purified by a two-step chromatographic procedure involving a combination of ion-exchange and size exclusion chromatography. The enzyme activity was analyzed after purification by performing chitinase assay as described previously. The purity of the sample was confirmed by the presence of a single band on a 12 % SDS–PAGE stained with Coomassie Brilliant Blue G250 dye on which protein showed approximately 30000 Da molecular weight.

Blast analysis of this amino acid sequence indicated a 29 % identity with Narbonin, a seed protein from V*icia narbonensis* L. The purified protein was concentrated to about 16 mg ml^{-1} and immediately used for crystallization. Single crystals were obtained with 20 % (w/v) PEG 8000, 0.1M CHES at pH 9.0 by using the hanging drop method. Crystals diffracted synchrotron X-radiation to 2.0 Å resolutions. The crystal belongs to the space group C2. Matthews's coefficient calculations indicated two molecules per asymmetric unit, which corresponds to a packing parameter V_M of 2.62 $Å^3Da^{-1}$ and a solvent content of approx. 53 %.

The phase problem was solved by molecular replacement using the Narbonin structure (PDB code: 1nar) as search model. Although chitinase and Narbonin have

relatively low sequence identity (29%), both possess the classical TIM barrel fold that is conserved in all family 18 chitinases. The DIDYE motif of chitinase essential for its activity is modeled straightforwardly from the DIHYE motif of Narbonin in intial maps (not shown) and the molecular-replacement is now being utilized for further model building and refinement. To understand the biological function of this chitinase in detail, soaking and co-crystallization with potential inhibitors are planned as well.

The X-ray structure of a tetrameric lectin (CVL) from spring *Crocus* plant (*Crocus vernus*) was determined and refined to the final *R*-factor of 17.3% and R_{free} of 22.5%, respectively. The structure of CVL has been solved by Multiple Isomorphous Replacement (MIR), using Platinum and Mercury metal derivatives. CVL was characterized to be a tetramer comprising four chains A, B, C and D with two dimers in the asymmetric unit. The two chains of a dimer share approx. amino acid sequence homology of around 38 %. Although CBL showed a 33% primary amino acid sequence identity with *Scilla campanulata* agglutinin (SCA) structure, but much conserved structural homology has been observed. CVL primary amino acid sequence deduced from the crystallographic structure indicated few deviations from the reported cDNA based sequence which interestingly made CVL a homo-dimer contrary to the reported hetero-tetrameric nature of the protein. It could be possible that there are isoforms of this protein inside the corm. However, quite similar structure has been observed to that of the reported theoretical model based on the docking experiments. CVL is an exclusively β-sheeted structure with no alpha helices inside. The β-prism II fold with an approximate internal 3-fold axis is composed of three sub-domains with antiparallel β-sheets giving an appearance of a triangular prism. This particular conserved structural fold is typical for all members of the bulb lectins super family. The four chains are arranged in the forms of rings comprised of an octamer containing approx a central 22 Å wide solvent channel. Both monomers of a dimer consist of three bundles of β-sheets which are oriented perpendicularly to the axis of the monomers. There is one intra-chain disulfide bond in each chain providing stability to the chain. The two chains of a dimer are stabilized through numerous hydrogen-bonds contacts and are strengthened additionally due to exchange of C-terminal strands between them.

Co-crystallization experiment in the presence of large excess of D-Mannose sugar (100mM) was performed and data was collected. A single binding pocket has been observed (map not shown) and yet to be refined. More ligand binding experiments are under way to make the complex formation in the future.

5 REFERENCES

[1] H. Tillich, *Seedling morphology in Iridaceae: Indications for relationships within the family and to related families*, Flora, **198** (2003), 220–242.

[2] R. Dahlgren and K. Bremer, *Major clades of the angiosperms*, Cladistics, **1** (1985), 349-368.

[3] P. Goldblatt, *Phylogeny and classification of Iridaceae*, Annals of the Missouri Botanical Garden, **77 (**1990), no. 4, 121-124.

[4] P. Goldblatt, *Phylogeny and classification of the Iridaceae and the relationships of Iris.* Annali di Botanica, **1 (**2001), no. 2, 13–28.

[5] P. Goldblatt, *Family Iridaceae*, Flora of Tropical East Africa, **1** (1996), 223-226.

[6] P. Goldblatt and J. Manning, *The Iris Family: Natural History & Classification*. Portland: Timber Press, 2008.

[7] W. Hui, C. Yanmei, Z. Changqi, *Flavonoids of the Genus Iris (Iridaceae),* Mini Reviews in Medicinal Chemistry, **10** (2010), no. 7, 643-661.

[8] G. S. Nickerson, *Some Data on Plains and Great Basin Indian Uses of Certain Native Plants*, Tebiwa, **9** (1966), no. 1, 45–51.

[9] D. Bown, *Encyclopaedia of Herbs and their Uses*, Dorling Kindersley, London, 1995.

[10] A. Chevallier, *The Encyclopedia of Medicinal Plants*. Dorling Kindersley, London, 1996.

[11] S. Frello and J. S. Heslop-Harrison, *Chromosomal variation in Crocus vernus Hill (Iridaceae) investigated by in situ hybridization of rDNA and a tendemly repeated sequence,* Annals of Botany, **86** (2000), 317-322.

[12] D. J. Bowles, *Defense-related proteins in higher plants,* Annu. Rev. Biochem. **59** (1990), 873-907.

[13] B. Fritig, T. Heitz, and M. Legrand, *Antimicrobial proteins in induced plant defense,* Curr. Opin. Immunol, **10** (1998), 16-22.

[14] P. Jollès and R. A. A. Muzzarelli, *Chitin and Chitinases*, Birkhäuser, Basel, 1999.

[15] F. Brunner, A. Stintzi, B. Fritig and M. Legrand, *Substrate specificities of tobacco chitinases,* Plant J., **14** (1998), 225–234.

[16] I. A. Hoell, S. S. Klemsdal, G. Vaaje-Kolstad, S. J. Horn and V. G. H. Eijsink, Biochem. *Overexpression and characterization of a novel chitinase from Trichoderma atroviride strain P1*, Biophys. Acta, 1748 (2005), 180–190.

[17] D. B. Collinge, K. M. Kragh, J. D. Mikkelsen, K. K. Nielsen, U. Rasmussen and K. Vad, *Plant chitinases,* Plant J. **3** (1993), 31–40.

[18] J. Flach, P. E. Pilet and P. Jolles, *What's new in chitinase research?* Experientia, **48** (1992), 701–716.

[19] G. Davies and B. Henrissat, *Structures and mechanisms of glycosyl hydrolases,* Structure, **3** (1995), 853–859.

[20] P. J. Hart, H. D. Pfluger, A. F. Monzingo, T. Hollis and J. D. Robertus, *The refined crystal structure of an endochitinase from Hordeum vulgare L. seeds at 1.8 A resolution*, J. Mol. Biol. **248**, (1995), 402- 413.

[21] T. Hollis, A. F. Monzingo, K. Bortone, S. Ernst, R. Cox and J. D. Robertus, *The X-ray structure of a chitinase from the pathogenic fungus Coccidioides immitis,* Protein Sci. **9** (2000), 544–551.

[22] B. Synstad, S. Gaseidnes, D. M. van Aalten, G. Vriend, J. E. Nielsen and V. G. Eijsink, *Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase,* Eur. J. Biochem. **271**(2004), 253–262.

[23] A. F. Monzingo, E. M. Marcotte, P. J. Hart and J. D. Robertus, *Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core,* Nature Struct. Biol. **3** (1996), 133–140.

[24] E. Cabib, S. J. Silverman and J. A. Shaw, *Chitinase and chitin synthase 1: counterbalancing activities in cell separation of Saccharomyces cerevisiae,* J. Gen. Microbiol. **138** (1992), 97–102.

[25] H. Merzendorfer and L. Zimoch, *Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases*, J. Exp. Biol. **206** (2003), 4393–4412.

[26] J. A. Fuhrman and W. F. Piessens, *Chitin synthesis and sheath morphogenesis in Brugia malayi microfilariae,* Mol. Biochem. Parasitol. **17** (1985), 93–104.

[27] N. T. Keen and M. Yoshikawa, *β-1,3-Endoglucanase from Soybean Releases Elicitor-Active Carbohydrates from Fungus Cell Walls,* Plant Physiol. **11 (**1983), 460-465.

[28] J. G. Bishop, A. M. Dean and T. Mitchell-Olds, Expand+*Rapid evolution in plant chitinases: Molecular targets of selection in plant-pathogen coevolution,* Proc. Natl Acad. Sci. USA, **97** (2000), 5322–5327.

[29] S. Sakuda, A. Isogai, S. Matsumoto and A. Suzuki, *Search for microbial insect growth regulators II. Allosamidin, a novel insect chitinase inhibitor,* J. Antibiot. **40** (1987), 296–300.

[30] E. Sandor, T. Pusztahelyi, L. Karaffa, Z. Karanyi, I. Pocsi, S. Biro, and A. Szentirmai, *Allosamidin inhibits the fragmentation of Acremonium chrysogenum but does not influence the cephalosporin-C production of the fungus,* FEMS Microbiol. Lett. **164** (1998), 231–236.

[31] J. M. Vinetz, S. K. Dave, C. A. Specht, K. A. Brameld, B. Xu, R. Hayward and D. A. Fidock, *The chitinase PfCHT1 from the human malaria parasite Plasmodium falciparum lacks proenzyme and chitin-binding domains and displays unique substrate preferences,* Proc. Natl. Acad. Sci. USA, **96** (1999), 14061–14066.

[32] J. M. Vinetz, J. G. Valenzuela, C. A. Specht, L. Aravind, R. C. Langer, J. M. C. Ribeiro, and D. C. Kaslow**,** *Chitinases of the Avian Malaria Parasite Plasmodium gallinaceum, a Class of Enzymes Necessary for Parasite Invasion of the Mosquito Midgut,* J. Biol. Chem. **275** (2000), 10331–10341.

[33] Z. Zhu, T. Zheng, R. J. Homer, Y. K. Kim, N. Y. Chen, L. Cohn, Q. Hamid and J. A. Elias, *Acidic mammalian chitinase in asthmatic Th2 inflammation and IL-13 pathway activation,* Science, **304** (2004), 1678–1682.

[34] L. Barbieri, M. G. Battelli and F. Stirpe, *Ribosome inactivating proteins from plants,* Biochim. Biophys. Acta. **1154** (1993), 237–282.

[35] W. J. Peumans and E. J. M. Van Damme, *Lectins as plant defense proteins*. Plant Physiol. **109** (1995)**,** 347–352.

[36] A. D. Mehta and R. S. Boston, Ribosome-inactivating proteins, eds., *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants,* American Society of Plant Physiologists, Rockville, MD, (1998), 145–152.

[37] P. L. De Hoff, L. M. Brill and A. M. Hirsch, *Plant lectins: the ties that bind in root symbiosis and plant defense,* Mol. Genet. Genomics, **282** (2009),1–15.

[38] S. A. McMahon, J. L. Miller, J. A. Lawton, D. E. Kerkow, A. Hodes, M. A. M. Renom, S. Doulatov, E. Narayanan, A. Sali, J. F. Miller and P. Ghosh, *The C-type lectin fold as an evolutionary solution for massive sequence variation,* Nat. Struct. Mol. Biol. **12** (2005), 886–892.

[39] G. J. Arason, *Lectins as molecules in vertebrates and invertebrates*. Fish Shellfish Immunol. **6** (1996), 277–289.

[40] E. J. M. Van Damme, A. Barre, P. Rougé and W. J. Peumans, *Cytoplasmic/ nuclear plant lectins: a new story,* Trends Plant Sci. **9** (2004), 484–489.

[41] E. J. M. Van Damme, L. Nausicaa and W. J. Peumans, *Plant Lectins*. In: J. C. Kader, M. Delseny, (eds) Advances in botanical research, Elsevier Ltd, San Diego, **48** (2008)107–209.

[42] B. Xia, Z. S. Kawar, T. Ju, R. A. Alvarez, G. P. Sachdev and R. D. Cummings, *Versatile Xuorescent derivatization of glycans for glycomic analysis*. Nat. Methods, **2** (2005), 845–850.

[43] K. Michiels, E. J. M. Van Damme and G. Smagghe, *Plant-insect interactions: what can we learn from plant lectins?* Archives of Insect Biochemistry and Physiology,73 (2010), no. 4, 193–212.

[44] E. J. M. Van Damme, W. J. Peumans, A. Barre and P. Rougé, *Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological role,* Crit. Rev. Plant Sci. **17** (1998), 575-692.

[45] M. W. Turner, *The role of mannose-binding lectin in health and disease,* Molecular Immunology, **40** (2003), 423–429.

[46] R. Malhotra and R. B. Sim, *Collectins and viral infections*. Trends Microbiol. **3** (1995), 240–244.

[47] O. Neth, D. L. Jack, A.W. Dodds, H. Holzel, N. J. Klein and M.W. Turner, *Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition*, Infect. Immun. **68** (2000), 688–693.

[48] E. Gorelik, U. Galili and A. Raz, *On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis*, Cancer Metast. Rev. **20** (2001), 245–277.

[49] E. G. De Mejía and V. I. Prisecaru, *Lectins as bioactive plant proteins: a potential in cancer treatment*, Crit. Rev. Food Sci. Nutr. **45** (2005), 425–445.

[50] S. H. Choi, S. Y. Lyu and W. B. Park, *Mistletoe lectin induces apoptosis and telomerase inhibition in human A253 cancer cells through dephosphorylation of Akt*, Arch. Pharm. Res. **27** (2004), 68–76.

[51] J. C. Aub, B. H Sanford and L. H. Wang, *Reactions of normal and leukemic cell surfaces to a wheat germ agglutinin.* Proc. Natl. Acad. Sci. U.S.A. **54** (1965), 400–402.

[52] M. Burger and A. R. Goldberg, *Identification of a tumor-specific determinant of neoplastic cell surfaces*. Proc. Natl. Acad. Sci. U.S.A. **57** (1967), 359–366.

[53] M. Inbar and L. Sachs, *Interaction of the carbohydrate-binding protein concanavalin A with normal and transformed cells*, Proc. Natl. Acad. Sci. U.S.A, **63** (1969), 1418–1425.

[54] B. A. Sela, H. Lis, N. Sharon and L. Sachs, *Different locations of carbohydrate-containing sites at the surface membrane of normal and transformed mammalian cells*. J. Membr. Biol. **3** (1970), 267–279.

[55] Y. W. Lam and T. B. Ng, *A monomeric mannose-binding lectin from inner shoots of the edible* chive *(Allium tuberosum),* J. Protein Chem, **20** (2001), 361-366.

[56] E. Davidson, J. M. S. Forrest, I. M. Morrison and D. Stewart, *Mannose-specific plant lectins from plants as diagnostics, vaccines and tools for the elucidation of viral infection mechanisms in animals,* In Annual report for the Scottish crop research institute for 1999/2000, Edited by W.H.M. Smith and T.D. Heilbronn, SCRI, Dundee, Scotland, (2000), 125-128.

[57] U. K. Laemmli, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4,* Nature, **227** (1970), 680-685.

[58] R. M. Hewick, M. W. Hunkapiller, L. E. Hood and W. J. Dreyer, *A gas-liquid solid phase peptide and protein sequenator*, J. Biol. Chem. **256** (1981), 7990-7997.

[59] S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *"Basic local alignment search tool"*, J. Mol. Biol. **215** (1990), no. 2, 403–410.

[60] Z. Otwinowski and W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, Methods Enzymol, **276** (1997), 307–326.

[61] *Collaborative Computational Project, Number 4, "The CCP4 Suite: Programs for Protein Crystallography*", Acta Cryst. D, **50** (1994), 760-763.

[62] B. W. Matthews, *Solvent content of protein crystals*, J. Mol. Biol. **33** (1968), 491-497.

[63] T. C. Terwilliger and J. Berendzen. *"Automated MAD and MIR structure solution",* Acta Cryst. D, **55** (1999), 849-861.

[63] T. C. Terwilliger, "*Maximum likelihood density modification*", Acta Cryst. D, **56** (2000), 965-972.

[63] T. C. Terwilliger, *"Automated main-chain model building by template matching and iterative fragment extension*", Acta Cryst. D, **59** (2003), 38-44.

[64] S. Panjikar, V. Parthasarathy, V. S. Lamzin, M. S. Weiss, and P. A. Tucker, *Auto-Rickshaw- An automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment*, Acta Cryst. D, 61 (2005),449-457.

[65] S. Panjikar, V. Parthasarathy, V. Lamzin, M. S. Weiss and P. A. Tucker, *On the combination of molecular replacement and single anomalous diffraction phasing for automated structure determination*, Acta Cryst. D, **65** (2009), 1089-1097.

[66] P. Emsley and K. Cowtan, *"Coot: model-building tools for molecular graphics"* Acta Cryst. D, **60** (2004), 2126-2132.

[67] A. A. Vagin, R. S. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long and G. N. Murshudov, "*REFMAC5 dictionary: organisation of prior chemical knowledge and guidelines for its use*", Acta Cryst. D, **60** (2004), 2284-2295.

[68] E. J. M. Van Damme, C. H. Astoul, A. Barre, P. Rougé and W. J. Peumans, *Cloning and characterization of a monocot mannose-binding lectin from Crocus* vernus *(family Iridaceae)*, European Journal of Biochemistry**, 267** (2000), no. 16, 5067–5077.

[69] R. Castillo, L. Gomez-Gomez and J. A. Fernandez, *SafchiA is a new class of defence chitinase from Saffron (Crocus sativus L.),* Acta Horticulturae, **739** (2007), 195-202.

[70] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs,* Nucleic Acids Res. **25** (1997), 3389-3402.

[71] T. Yamagami, Y. Mine and M. Ishiguro, *Complete amino acid sequence of chitinase-a from bulbs of gladiolus (Gladiolus gandavensis),* Biosci. Biotechnol. Biochem. **62** (1998), 386-389.

[72] L. M. Wright, C. D. Reynolds, P. J. Rizkallah, A. K. Allen, E. J. Van Damme, M. J. Donovan and W. J. Peumans, *Structural characterization of the native fetuin-binding protein Scilla campanulata agglutinin: a novel two-domain lectin,* FEBS Lett. **468** (2000), no. 1, 19-22.

[73] A. Vagin, and A. Teplyakov, *MOLREP: an automated program for molecular replacement,* J. Appl. Cryst. **30** (1997), 1022-1025.

[74] G. Hester, H. Kaku, I. J. Goldstein and C. S. Wright, *Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family,* Nat. Struct. Biol. **2** (1995), no. 6, 472-479.

[75] J. Navaza, *AMoRe: an automated package for molecular replacement,* Acta Cryst. A, **50** (1994), 157-163.

[76] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L.C. Storoni and R. J. Read, *Phaser* crystallographic *software,* J. Appl. Cryst. **40** (2007), 658-674.

[77] A. G. W. Leslie, *Joint CCP4 + ESF-EAMCB,* Newsletter on Protein Crystallography, **26** (1992).

[78] P. R. Evans, "*Scaling and assessment of data quality*", Acta Cryst. D, **62** (2005), 72-82.

[79] K. Hofmann and M. D. Baron, *BOXSHADE: printouts from Aligned Protein or DNA Sequences, 2*001. ttp://bioweb.pasteur.fr/seqanal/interfaces/boxshade.html.

[80] W. I. Weis and K. Drickamer, *Structural basis of lectin-carbohydrate recognition,* Annu. Rev. Biochem. **65** (1996), 441-473.

[81] J. Balzarini, D. Schols, J. Neyts, E. V. Damme, W. Peumans and E. D. Clercq, *Alpha-(1-3)- and alpha-(1-6)-D-mannose-specific plant lectins are markedly inhibitory to human immunodeficiency virus and cytomegalovirus infections in vitro,* Antimicrob. Agents Chemother, **35** (1991), 410-416.

[82] A. Barre, E. J. M. V. Damme, W. J. Peumans and P. Rouge, *Structure-Function Relationship of Monocot Mannose-Binding Lectins,* Plant Physiology, **112** (1995), 1531-1540.

[83] J. Escribano, A. Rubio, M. Alvarez-Ortí, A. Molina and J. A. Fernández, *Purification and characterization of a mannan-binding lectin specifically expressed in corms of saffron plant (Crocus sativus L.)*, J. Agric. Food Chem., **48** (2000), 457-463.

[84] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *UCSF Chimera--a visualization system for exploratory research and analysis*, J. Comput. Chem., **25** (2004), 1605-1612.

[85] R. Sankaranarayanan, K. Sekar, R. Banerjee, V. Sharma, A. Surolia and M. Vijayan, *A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a beta-prism fold*, Nat Struct. Biol., **3** (1996), 596-603.

[86] G. Alexey, B. Murzina, A. M. Lesk and C. Chothia, *β-Trefoil fold: Patterns of structure and sequence in the Kunitz inhibitors interleukins-1β and 1α and fibroblast growth factors*, J. Mol. Biol., **223** (1992), 531-543.

[87] C. S. Wright, *New folds of plant lectins*, Curr. Opin. Struct. Biol., **7** (1997), 631-636.

[88] S. W. Provencher, *CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations*. Comput. Phys. Commun., **27** (1982), 213-229.

[89] A. T. Brunger, *Free R value: a novel statistical quantity for assessing the accuracy of crystal structures*. Nature, **355** (1992), 472–475.

[90] C. J. Huang, T. K. Wang, S. C. Chung and C. Y. Chen, *Identification of an antifungal chitinase from a potential biocontrol agent, Bacillus cereus.* J. Biochem. Mol. Biol., **38** (2005), 82-88.

[91] G. H. Renkema, R. G. Boot, A. O. Muiisers, W. E. D. Koopman and J. M. Aerts, *Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins*, J. Biol. Chem., **270** (1995), 2198-2202.

[92] R. Barone, F. D. Gregorio, M. A. Romeo, G. Schiliro and L. Pavone, *Plasma chitotriosidase activity in patients with β-thalassemia*, Blood Cells Mol. Dis., **25** (1998), 1-8.

[93] R. Maiti, G. H. V. Domselaar, H. Zhang and D. S. Wishart, *SuperPose: a simple server for sophisticated structural superposition*, Nucleic Acids Res., **32** (2004), 590-594.

List of Dangerous Materials

Compound	Symbols of hazardous materials	R-statements	S-statements
Acrylamide	T	45-46-20/21-25-36/38-43-48/23/24/25-62	53-45
Ampicillin	Xn	36/37/38-42/43	22-24-26-37
APS	O, Xn	8-22-36/37/38/-42/43	22-24-26-37
DMSO	Xi	36/38	26
DTT	Xn, Xi	22-36/38	36/37/39
Acetic acid (Essigsäure)	C	10-35	23.2-26-45
Ethanol	F	11	7-16
EtBr	T+	22-26-36/37/38-40	26-28.2-36/37-45
HCL	C	34-37	26-36/37/39-45
Methanol	T	61	26-36/37-39-45
NaOH	F, T	11-23/24/25-39/23/24/25	7-16-36/37-45
Ni-NTA Agarose	O, C	8-35	8-27-39-45
NaCl	F, Xi	11-36-67	7-16-24/25-26
2-Propanol	T	24/25-34	28.6-45
SDS	C	34-37	26-36/37/39-45
TEMED	Xn	22-36/38	22-24/25
Tris	F, C	11-20/21/22-35	3-16-26-29-36/37/39-45

Symbols of Dangerous Materials

E	Explosive (Explosionsgefährlich)
C	Caustic (ätzend)
F+	Extremely flammable (hochentzündlich)
Xi	Irritant (reizend)
O	Oxidizing (brandfördernd)
F	Highly flammable (leichtentzündlich)
T	Toxic (giftig)
T+	Very toxic (sehr giftig)
Xn	Harmful (gesundheitsschädlich)
N	Hazardous to the environment (umweltgefährlich)

ACKNOWLEDGEMENTS

‘‘O Almighty ALLAH (*Azza Wa Jalla*)’’ I acknowledge you because whenever I became desperate or confused, my only hope and motivation was to pray and you never leave me alone and guided me the true pathway. I would always love to say that ‘‘I Love You’’.

For every truth seeking person, the one and the only realistic model and torch of light guiding him towards the right direction is the greatest personality of world which is Holy Prophet MUHAMMAD (*Sallallaho Alaihe Wasallama*).

I express my deep sense of gratitude towards my supervisor, Professor Christian Betzel for his constant encouragement and technical advices during my whole research work. He provided me such a beautiful infrastructure in which I had executed all my research plans and experimental work smoothly and with all the facilities in hand. It was a wonderful experience to work under his supervision.

I also convey my thanks to Professor Wolfgang Voelter, Eberhard-Karls-Universität, Tübingen for his caring attitude during my PhD. Special thanks to Dr. Lars Redecke and Dr. Markus Perbandt for their technical guidance and suggestions during the protein purification phase and at time of diffraction data collection respectively. Dr. Friedrich Buck, Universitäts-Klinikum Hamburg-Eppendorf (UKE) helped me a lot for the N-terminal sequencing and Mass Spectrometry results which greatly facilitated my progress.

Many thanks for the Dr. Arne Meyer who guided me during the initial stages of my research work and helped me in planning my PhD projects which I ended successfully. Similarly, Chitinase project has been greatly facilitated and smoothened because of the enormous help of Miss Sadaf Iqbal. Bunches of thanks are forwarded to acknowledge her efforts.

I also appreciate my colleague Amr Negm in providing me the lab facilities for conducting sequencing experiments and for the pleasant company. Dominik Oberthuer looked deeply into my thesis and provided me valuable suggestions which resulted in much thesis improvement for which I am grateful to him.

Last but not least I would like to give heartiest thanks and deep sense of gratitude to the patience and constant encouragement of my family in Pakistan. My parents are always praying for my success and I cannot find any alternative for these constant wishes.

AHMED AKREM

Eidessstattliche Erklärung

Hiermit erkläre ich an Eides statt, diese Arbeit selbstständig und ohne fremde Hilfe verfasst, sowie keine anderen, als die von mir angegebenene Hilfsmittel verwendet zu haben.

Ferner versichere ich, dass ich noch keine Promotionsversuche an anderen Universitäten unternommen habe.

Hamburg, 25. November 2010

Ahmed Akrem

Lebenslauf

Name: Ahmed Akrem

Geburtstag: 30.01.1976

Geburtsort: Lahore, Pakistan

Ausbildung:

Degree	Year	Marks obtained	Grade/Div.	Subject	Board/University
Matric	1991	665/850	1st	Science Group	Lahore
F.Sc.	1994	672/1100	1st	Pre-Medical	Lahore
B.Sc.	1997	473/800	2nd	Botany, Chemistry, Zoology.	Punjab University
M.Sc.	1999	3.87/4.00	1st	Botany	University of Agriculture, Faisalabad, Pakistan.
M.Phil.	2001	4.2/5.00	1st	Biotechnology	Quaid-I-Azam University, Islamabad, Pakistan.

Immatrikulation Nr: 5888798

Publikation: Latif F, M. Asghar, R Saleem, **A Akrem** and R.L. Legge (2006). Purification and Characterization of a Xylanase Produced by *Chaetomium thermophile* NIBGE. *World J. Microbiol. Biotechnol.* **22:** 45-50.