

**Systematic approach to protein crystallization: emphasis on Vaccinia virus complement control protein (VCP).**

**Felix Adusei-Danso**



**Supervisors:**

**Prof. Girish Kotwal  
Prof. Trevor Sewell  
Dr. Mohamed Sayed**

## ABSTRACT

**Felix Adusei-Danso**

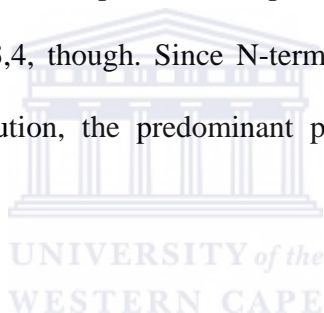
MSc minithesis, Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape.

This work examines the systematic approach to protein crystallization, exploring some of the techniques that have been developed to enhance the success rate of crystallization. The work was centered on two proteins; namely Vaccinia virus complement control protein (VCP) and glutamate dehydrogenase (GDH) from *Bacteriodes fragilis*. The crystal structures of the full length native VCP and VCP bound to heparin had already been determined. However, the structure of rVCP 2,3,4; a mutant VCP in which the first consensus repeat is deleted is not known. In the same way, the structure of GDH from *Bacteriodes fragilis* is not known; even though structures of other GDHs from different organisms have been determined.

Using the *Pichia pastoris* yeast expression system, rVCP 2,3,4 was expressed, concentrated and purified by heparin column affinity chromatography. rVCP 2,3,4 was eluted at 400mM NaCl, dialysed against 10mM Tris buffer and superconcentrated to a final concentration of 6.1mg/ml. Using Hampton screen 1 and 2, initial crystallization trials were setup on the protein via the hanging drop vapour diffusion method. Promising drops were optimised using pH gradient and concentration gradient. Only one of the

optimised drops led to the production of a single, pyramidal-shaped protein crystal. The protein crystallized in the F23 spacegroup and diffracted to a nominal resolution of 2.5Å.

The data was processed by DENZO and molecular replacement was used to phase the data but a correct solution could not be found. One of two reasons may account for why a correct solution could not be found. Firstly, rVCP 2,3,4 may fold in an entirely different shape from any of the VCP fragments used as probes. The second possibility might be that the protein in the crystal was not rVCP 2,3,4. Post crystallization SDS-PAGE gel conducted on the superconcentrated protein solution revealed that the solution used to set up the trials was not very pure. The predominant protein on the gel was in the right position expected of rVCP 2,3,4, though. Since N-terminal sequence analysis was not conducted on the protein solution, the predominant protein could not be positively identified as rVCP 2,3,4.



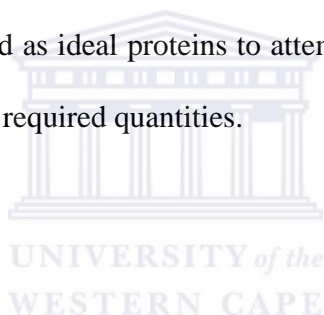
To further test if the protein in the crystal was rVCP 2,3,4, solvent content analysis was conducted on the crystal. Results from the solvent content analysis indicated that rVCP 2,3,4 could fit into the unit cell dimension of the crystal, making rVCP 2,3,4 a possible protein in the crystal.

Purified GDH in ammonium sulphate was processed, dialysed and concentrated to 14mg/ml. Initial crystallization trials were set up using Hampton screens 1 and 2 by hanging drop method. Showers of crystals obtained from the initial screens were optimised by the method of oils, *in situ* dilution of reservoir solution and pH gradient.

Single, large crystals were obtained from two drops, which diffracted to 6.5Å and 8.5Å respectively. For some unknown reasons, the data could not be indexed. Therefore, no further structural determination could be carried out on the data. It was suggested that the large size of a molecule of GDH (288 kDa) and the low-resolution data may play a role in the failure to index the data.

The crystal was taken to the synchrotron for data collection. However, data collected was not any better than that collected from the in-house X-ray diffractometer.

Vaccinia virus complement control protein (VCP) and glutamate dehydrogenase (GDH) from *Bacteriodes fragilis* served as ideal proteins to attempt crystallization as they could be produced and purified in the required quantities.



## **Dedication**

This work is dedicated to the fondest memory of my beloved father, George Kwaku Adusei, who saw the beginning of the compilation of this project but never lived to see the end of it.

(Requisat In Pacem)



## Declaration

I declare that “**Systematic approach to protein crystallization: emphasis on Vaccinia virus complement control protein**” is my own work and that it has not been submitted for any degree or examination in any other university. I also declare that the sources I have used or quoted have been indicated or acknowledge by complete reference.

**Felix Adusei-Danso**

August 15, 2006

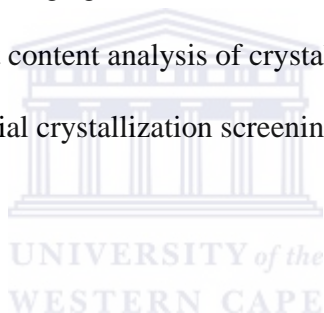
(Solemnity of the assumption of the Blessed Virgin)



Signed.....

## List of tables

- Table 3.1:** Results of initial crystallization trials of rVCP 2,3,4.
- Table 3.2:** A table showing the precipitants (and the concentration ranges) that gave observable results of the initial crystallization screening of rVCP 2,3,4.
- Table 3.3:** Summary of data collection statistics of rVCP 2,3,4 data.
- Table 3.4:** Results of molecular replacement of rVCP 2,3,3 using different molecular replacement methods.
- Table 3.5:** Results of molecular replacement using MOLREP involving fixing of modules and searching against other modules.
- Table 3.6:** Results of solvent content analysis of crystal under investigation.
- Table 3.7:** Results of the initial crystallization screening of GDH.



## List of figures

- Fig. 1.1:** Graph of structural biology success rate.
- Fig. 1.2:** Phase diagram of protein crystallization.
- Fig. 1.3:** Phase diagram of protein crystallization, comparing vapour diffusion and microbatch methods.
- Fig. 1.4:** Simplified diagram of the processes of the complement pathway.
- Fig. 1.5:** Crystal structures of VCP.
- Fig. 1.6:** NMR structures of VCP fragments.
- Fig. 1.7:** Simplified metabolic pathways of methanol.
- Fig. 1.8:** General expression vector of the *Pichia pastoris* expression system.
- Fig. 1.9:** The structure of NAD dependent GDH, showing the hexamer.
- Fig. 3.1:** A graph of predicted crystallization success rate (frequency) versus the likely (pI-pH) ranges of rVCP 2,3,4 crystallization.
- Fig. 3.2:** A graph of predicted crystallization success rate (frequency) versus the likely (pI-pH) ranges of GDH crystallization.
- Fig. 3.3:** SDS-PAGE gel of the expression of rVCP 2,3,4.
- Fig. 3.4:** SDS-PAGE gel for the purification of rVCP 2,3,4.
- Fig. 3.5:** SDS-PAGE gel of dialysis of rVCP 2,3,4.
- Fig. 3.6:** Phase separation of screen 1, 39, after 3 weeks.
- Fig. 3.7:** Crystalline precipitate of screen 1, 24 after 3 weeks.
- Fig. 3.8:** Amorphous precipitate of screen 1, 19 after 3 weeks.
- Fig. 3.9:** Amorphous precipitate of screen 1, 19 after 3 weeks.



- Fig. 3.10:** Salt crystals of screen 2, 25 after 3 weeks.
- Fig. 3.11:** Micro crystals from screen 1, 27 after 3 weeks.
- Fig. 3.12:** Salt crystal obtained from 25% Isopropanol, 0.1M Tris pH 8.5, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ .
- Fig. 3.13:** Salt crystal obtained from 28% PEG 400, 0.1M Na HEPES pH 7.5, 0.1M  $\text{CaCl}_2$  with 700 $\mu\text{l}$  of 2:1 (paraffin: silicone) oil.
- Fig. 3.14:** Salt crystals obtained from 35% PEG 8000, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ .
- Fig. 3.15:** True protein crystal obtained from 35% Isopropanol, 0.1M Na-HEPES pH 7.5, 0.3M  $\text{MgCl}_2$ .
- Fig. 3.16:** Diffraction pattern of a salt crystal.
- Fig. 3.17:** Diffraction pattern of protein crystal.
- Fig. 3.18:** Post crystallization SDS-PAGE gel of rVCP 2,3,4 solution.
- Fig. 3.19:** Showers of GDH crystals produced from 8% PEG 8000, 0.1M Tris pH 8.5, observed after 4 days.
- Fig. 3.20:** Showers of GDH micro-crystals produced from 1M  $\text{Li}_2\text{SO}_4$ , 0.1M Na Citrate pH 5.6, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ , observed after 8 days.
- Fig. 3.21:** Showers of GDH crystals produced from 30% PEG 8000, 0.1M Na Cacodylate pH 6.5, 0.2M  $(\text{NH}_4)_2\text{SO}_4$ , observed after 12 hours.
- Fig. 3.22:** Clear solution of GDH produced after 3hrs drop intervention time.
- Fig. 3.23:** GDH Crystals produced after 6hrs drop intervention time.
- Fig. 3.24:** GDH Crystals produced after 9hrs drop intervention time.
- Fig. 3.25:** Crystals of GDH grown from 1M  $\text{Li}_2\text{SO}_4$ , 0.1M Na Citrate pH 5.6, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ , with 800 $\mu\text{L}$  of 3:1 (Paraffin: Silicon) oil.

## ABBREVIATIONS

AOX	Alcohol oxidase
BMGY	Buffered Minimal Glycerol
BMMY	Buffered Minimal Methanol
C1q	Complement 1 q
C4	Complement 4
C3b	Complement 3b
CR1	Complement receptor type 1
C4BP	Complement 4b
CCP	Complement control protein
CCP4	Collaborative Computational Project Number 4
DAF	Decay accelerating factor
DHAS	Dihydroxyacetone synthase
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
fH	Factor H
GDH	Glutamate dehydrogenase
MAC	Membrane attack complex
MCP	Membrane cofactor protein
MMH	Minimal methanol histidine
NaCl	Sodium chloride
NADH	Nicotinamide adenine dinucleotide (reduced form)

NADPH	Nicotinamide adenine dinucleotide phosphate
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Ammonium sulphate
NMR	Magnetic nuclear resonance
PEG	Polyethylene glycol
RCA	Regulators of complement activation
RNA	Ribonucleic acid
SCR	Short consensus repeats
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TCA	Tricarboxylic acid
VCP	Vaccinia virus complement control protein
rVCP	Recombinant Vaccinia virus complement control protein
YEP	Yeast Extract Peptone
YNB	Yeast extract nitrogen base

## **Acknowledgements**

I wish to acknowledge and express my sincere gratitude to my supervisors; Prof. Girish Kotwal, Prof. Trevor Sewell and Dr. Mohamed Sayed for the opportunity they gave me to work with them. I appreciate the various ways by which they had contributed to the success of this project. I also want to thank Prof. David McIntosh through whose friendship I came to know about the Structural Biology program in South Africa.

I am also grateful to all students in the Structural Biology program, particularly my classmates (Jason van Rooyen, Jean Watermeyer, James Onyemata, Tim Frouws and Ndoriah Thuku) and Samuel Kojo Kwofie for all the encouragement, assistance and advice they gave me during the period of this project. I wish to acknowledge the kindness, friendliness and assistance of the students in the Medical Biotechnology Laboratory of UCT where I did a substantial amount of this project.

I also want to thank members of Council 40 of the Noble Order of Knights of Da Gama (Wynberg); particularly Brother Keith Arrow and his family for the diverse ways they helped made my stay in Cape Town an enjoyable one. I am also grateful to the Carnegie Cooperation (New York, USA) and the Polio Research Foundation (PRF) in South Africa for the financial assistance they provided me during the period of this project.

Finally, I want to say a big thank you to my wife, Freda Adusei-Danso, who stood by me day and night, encouraged and gave me the moral support I needed as I awaited desperately to see a crystal in my crystallization drops. She never gave up the fight!

## Table of contents

Abstract.....	i
Dedication.....	iv
Declaration.....	v
List of tables.....	vi
List of figures.....	vii
Abbreviations.....	ix
Acknowledgements.....	xi

### Chapter 1: Literature review

1.1 Poxviruses.....	1
1.2.1 The complement system.....	2
1.2.2 The classical pathway.....	2
1.2.3 The alternate pathway.....	5
1.3 Control of complement pathway.....	5
1.4.1 Vaccinia virus complement control protein (VCP).....	5
1.4.2 VCP inhibits the complement pathway.....	7
1.4.3 VCP has heparin activity.....	7
1.4.4 Therapeutic exploitation of the complement system.....	8
1.4.5 Crystal structure of VCP.....	8
1.4.6 Putative complement binding sites of VCP.....	11
1.4.7 NMR structures of VCP.....	11
1.4.8 rVCP 2,3,4.....	13
1.5.1 Expression of heterologous proteins using <i>Pichia pastoris</i> .....	14
1.5.2 Expression vector of <i>Pichia pastoris</i> .....	15
1.6.1 Glutamate dehydrogenase.....	17
1.6.3 Three dimensional structures of GDH.....	18
1.7 Protein crystallization.....	20
1.7.1 Crystallization is the rate limiting step in structural biology projects.....	20
1.7.2 Inhibition of crystallization by evolutionary negative design.....	22
1.7.3 Evidence of the negative evolutionary design of protein crystallization.....	22
1.8.1 The crystallization process.....	23
1.8.2 The nucleation state.....	25
1.8.3 Thermodynamics of nucleation of proteins.....	26
1.8.4 Formation of precipitates.....	29
1.8.5 Growth of crystals.....	29
1.9 Effect of pH and salt on protein crystallization.....	29
1.10.1 Methods of crystallization.....	30
1.10.2 Comparative advantages of microbatch and vapour diffusion methods.....	35
1.10.3 Problems with microbatch crystallization methods.....	35
1.10.4 Advantages of microbatch method.....	36
1.10.5 Adaptation of vapour diffusion to microbatch crystallization and vice versa....	37
1.11.1 Crystal optimization.....	39
1.11.2 Separation of nucleation from growth phase.....	39

1.11.3 Drop dilution as a means of control nucleation.....	40
1.11.4 The use of oils .....	41

## Chapter 2: Methodology

2.1 Prediction of likely pH ranges for crystallization of rVCP 2,3,4 and GDH.....	42
2.2 Expression of rVCP 2,3,4.....	42
2.2.1 Preparation of expression media.....	42
2.2.2 Growth of <i>Pichia pastoris</i> .....	44
2.2.3 Induction and harvesting.....	44
2.2.4 SDS-PAGE analysis.....	45
2.3 Concentration and purification of rVCP 2,3,4.....	45
2.4 Dialysis and supersaturation of rVCP 2,3,4.....	46
2.5 Preparation of GDH for crystallization.....	46
2.6 Crystallization of rVCP 2,3,4 and GDH.....	47
2.7 Optimisation of rVCP 2,3,4 and GHD trials.....	47
2.7.1 Introduction.....	47
2.7.2 The use of oils.....	47
2.7.3 Change in concentration of precipitating solution.....	48
2.8 Data collection and processing.....	48
2.9 Molecular replacement and refinement.....	50
2.10 Post crystallization SDS-PAGE gel to determine presence of impurities.....	51

## Chapter 3: Results and discussion

3.1 Prediction of likely crystallization pH.....	53
3.2 Expression of rVCP 2,3,4 in <i>Pichia pastoris</i> .....	56
3.3 Purification of rVCP 2,3,4.....	58
3.4 Dialysis of rVCP 2,3,4.....	61
3.5 Results and observed trend from the initial crystallization screening of rVCP 2,3,4.....	62
3.6 Optimization of the initial screening that produced precipitants.....	66
3.7 Optimisation of the initial screening that produced phase separation.....	67
3.8 Results from the optimisation of rVCP 2,3,4.....	69
3.9 Processing and phase determination of rVCP 2,3,4 data.....	76
3.9.1 Processing of rVCP 2,3,4 data.....	76
3.9.2 Phase determination using molecular replacement.....	77
3.9.3 Refinement of rVCP 2,3,4 data.....	80
3.9.4 Implication of the failure of molecular replacement of find a correct solution...81	
3.9.5 Post crystallization SDS-PAGE gel confirms presence of impurities.....	83
3.9.6 Solvent content analysis of the ‘disputed crystal’.....	85
3.10 Crystallization results of GDH.....	88
3.10.1 Results from initial screening.....	88
3.10.2 Optimisation of initial results of GDH trials.....	90
3.10.3 Optimisation of GDH crystal, using pH gradient.....	93
3.10.4 The use of oils to optimise GDH crystals.....	95

3.11 Data collection and processing of GDH.....96

Chapter 4 (conclusion and recommendations for further studies).....98

References.....101



## CHAPTER 1: LITERATURE REVIEW

### 1.1 Poxviruses

Poxviruses are large, cytoplasmic and complex DNA viruses; considered as one of the most evolutionary successful pathogens [Buller & Palumbo, 1991]. Unlike other DNA viruses like herpesviruses, poxviruses do not undergo latency. Instead, they express a variety of immunomodulatory proteins that assist them to evade the immune response of the host. They include vaccinia virus, cowpox virus, monkeypox virus, molluscum contagiosum and smallpox virus, the most deadly of them all.

Poxviruses are linear double stranded with a single stranded hairpin loop at the telomers. Close to the telomers are found the inverted terminal repeats, within which are found the open reading frames, which encode a variety of immunomodulatory proteins responsible for the pathogenesis of the virus. This portion of the poxviral genome had been termed 'non-essential' region because it can be dispensed without adversely affecting the replication of the viral genome [Kotwal & Moss, 1988(a)]. On the other hand, the central portion of the viral genome is called the essential region because this region contains the genes essential for replication and cannot be dispensed [Kotwal & Moss, 1988(a)].

Poxviruses encode a variety of immunomodulatory proteins, which act in one of several ways to influence their habitat and ensure the continuous and successful stay of the virus in the host [Kotwal, 2000]. These include those that block and inhibit the complement system, those that block cytokine biosynthesis or blind cytokine receptors, and those that



inhibit interferon signaling pathways, those that inhibit serine protease granzyme B and those that mimic the action of natural killer cell receptors [Kotwal, 2000].

### **1.2.1 The complement system**

The complement activation is the first line of defense of the body against microbial invasion. It consists of more than 30 plasma and membrane bound proteins that work in a sequential manner in order to bring about the elimination of invading pathogens [Kotwal, 1996; Reid, 1995]. There are three pathways by which complement activation can occur. These are the classical, alternate and lectin pathways (fig. 1.1). The two main pathways are the classical and alternate pathways.

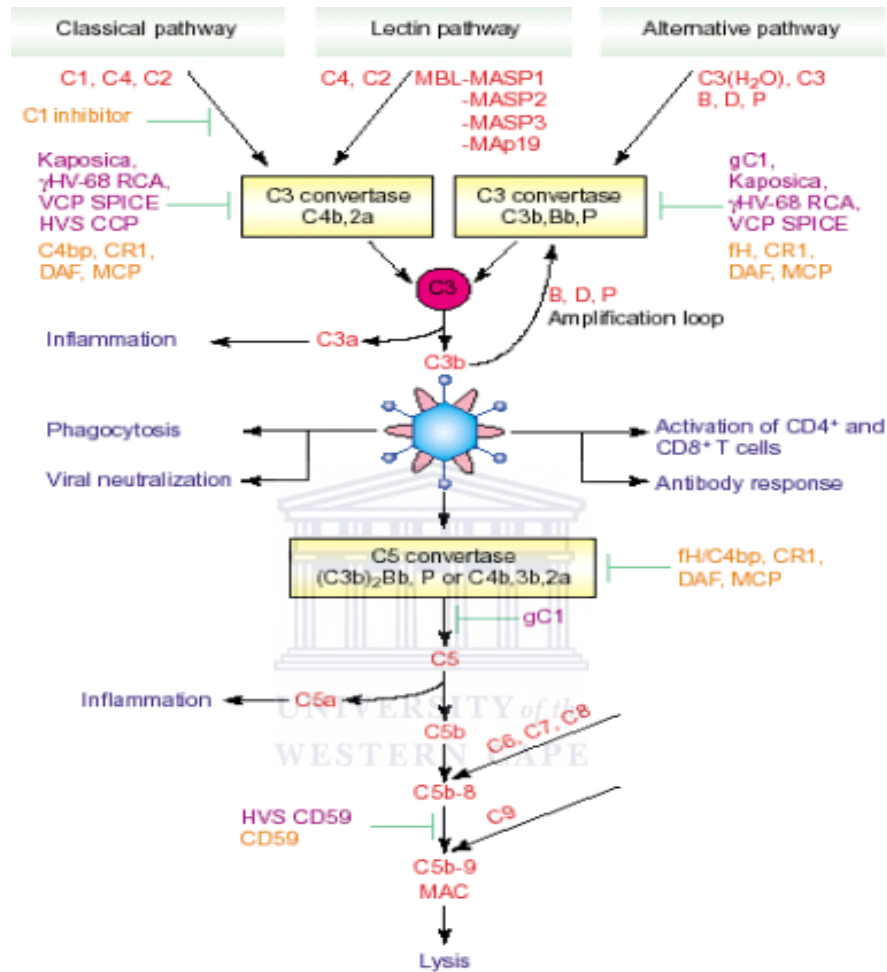
Irrespective of which pathway it takes, complement activation results in the elimination of the invading pathogens via a number of ways. These include neutralization and opsonization, lysis of infected cells and influx of chemotactic attractants (fig. 1.1). The ultimate result of all these effects is that they cause an increase in inflammatory immune response. Both the classical and alternative pathway results in the formation of the membrane attack complex (MAC) [Muller, 1988; Frank & Fries, 1989; Lambris *et al*, 1998; Welsh *et al*, 1975)].

### **1.2.2 The classical pathway**

The classical pathway of complement activation is also called the antibody dependent, because it is triggered by the specific antigen-antibody binding to complement proteins. The first step of the classical activation pathway involves binding of complement protein

C1q either to antibody–antigen complexes or, occasionally, directly to the surface of certain pathogens [Cooper *et al*, 1974; Ebenbichler *et al*, 1991; Ikeda *et al*, 1998; Spiller & Morgan, 1998]. C1q binding causes an activation of C1r, which then cleaves C1s to an active serine protease form; which in turn cleaves C4, thus causing a cascade of complement events (fig. 1.1).





**Fig. 1.1:** simplified diagram of the processes of the complement pathway; showing the three pathways. VCP inhibits both classical and alternate pathways by binding to C3 and C4; acting as cofactor for factor I mediated cleavage of C3 convertase and accelerating the decay of C3 convertase. [Mullck *et al*, 2003].

### **1.2.3 The alternate pathway**

The alternate pathway is a default process, and its activation is triggered by spontaneous and indiscriminate deposition of C3b on surfaces of host cells or foreign particles. C3b is produced at a significant rate in the plasma by spontaneous hydrolysis of C3 present in the plasma. After deposition of C3b, it causes a cascade of complement activation events unless down regulated by specific mechanisms. Irrespective of the pathway complement activation is triggered; all the pathways meet at the C3 level (fig. 1.1).

### **1.3 Control of complement pathway**

Because the complement system is such an important and potentially damaging system, it must be tightly controlled. Mammalian cells are protected from uncontrolled complement activation by a number of complement control proteins. These include complement receptor type 1 (CR1), decay accelerating factor (DAF), complement-4- binding protein (C4BP) and membrane cofactor protein (MCP) [Hoolers et al. 1985]. Complement control proteins had long been considered as strictly species dependent. Several studies [Van den Berg & Morgan, 1994; Rushmere *et al*, 1997; Perez *et al*, 2000] have however shown that some of the complement control proteins such as DAF tolerate a high degree of heterologous species.

#### **1.4.1 Vaccinia virus complement control protein (VCP)**

Vaccinia virus complement control protein (VCP) was the first soluble microbial immunomodulatory protein to be discovered [Kotwal & Moss, 1988b]. It is a 26kDa virokine and the major secretory protein of vaccinia virus infected cells [Kotwal & Moss,

1988(b)]. Structurally, it belongs to the group of regulators of complement activation (RCA) proteins. It is 263 amino acid in length before the cleavage of the signal sequence, structurally related to C4BP, and functionally related to CR1 [Kotwal 1996; Kotwal 1994]. It shares up to 38% sequence identity to the first four consensus repeats of C4BP [Kotwal & Moss, 1988(b); Kotwal & Moss, 1989]. VCP is a secreted protein encoded by an open reading frame (C3L) in the HindIII fragment of the vaccinia virus genome [Earl & Moss, 1989].

Regulators of complement activation proteins are composed of repeats of similar domains called short consensus repeats (SCR); otherwise called Sushi domains [Bork *et al*, 1996] or complement control protein (CCP) modules [Reid *et al.*, 1986]. Each repeat is made up of about 60 to 70 amino acids in length characterized by a motif with four invariant disulphide-bonded cysteines [Pangburn, 1986; Reid & Day, 1989]. Neighbouring modules are connected to each other by a linker sequence of varying lengths; usually four residues long; extending from the fourth cysteine in the preceding module to the first cysteine of the neighbouring module. Members in this family include complement-4-binding protein (C4BP), factor H (fH), complement receptor type 1 (CR1), decay accelerating factor (DAF), membrane cofactor protein (MCP), vaccinia complement control protein (VCP) and its orthopoxviral homologues [Kotwal & Moss, 1988(b)].

#### **1.4.2 VCP inhibits the complement system**

Unlike some other complement control proteins, VCP inhibits both the classical and alternate pathways at multiple sites [Kotwal *et al*, 1990]. VCP binds C3 and C4 and thus inhibits both classical and alternate pathways by blocking the formation of C3 and C5 convertases [Kotwal *et al*, 1990; Rossengard *et al*, 1999; Smith *et al*, 2003]. Even when C3 convertase is formed, VCP can also accelerate its (C3 convertase) decay back into its components [Kotwal *et al*, 1990; McKenzie *et al*, 1992]. Once the surfaces of cells are coated with C3b, they are destined for destruction. To prevent this from happening, VCP can also serve as a cofactor for factor I cleavage of C3b and C4b into their inactive form [Sahu *et al*, 1998]. VCP is a more powerful complement inhibitor than human C4BP [Kotwal, 1994].

#### **1.4.3 VCP has heparin binding activity**

Not many complement control proteins have heparin binding activity. With the exception of VCP, C4BP and fH, [Pangburn *et al*, 1991; Smith *et al*, 2000], no other complement binding proteins have been shown to bind to heparin. The ability of VCP to bind to heparin [Kotwal *et al*, 1998] contributes to the successful evasion of the vaccinia virus to the host immune system as well as blocking of xenoantibodies to endothelial cells [Al-Mohanna *et al*, 2000]. This is because by binding to heparin, VCP deprives the attachment of chemoattractant to heparin on the endothelial cells. Binding of heparin to chemoattractants is necessary for the localization and migration of the chemoattractants to the muscle, resulting in inflammation and pathogen elimination [Lalani & McFadden,

1997]. Heparin binding has been shown to occur at the positively charged patched C-terminal end of the molecule in SCR4 [Ganesh *et al*, 2004].

#### **1.4.4 Therapeutic exploitation of the complement system**

The development of therapeutic agents aimed at subverting and inhibition of the complement system is being vigorously pursued in conditions where uncontrolled complement activation is undesirable. These conditions include hyperacute rejection of xenogeneic transplants, Alzheimer's, biomaterial incompatibility injury, ischemia-reperfusion injury, restenosis and systematic lupus erythematoses [Makrides, 1998].

VCP has many properties that make it as an ideal candidate for its therapeutic development against a multiple of complement mediated diseases. These properties include its intrinsic solubility, its very high thermal stability, its smaller size, its ability to inhibit both the classical and alternate pathways and its ability to bind to heparin [Smith *et al*, 2002; Reynolds *et al*, 2000; Rosengard *et al*, 1999].

#### **1.4.5 Crystal structure of VCP**

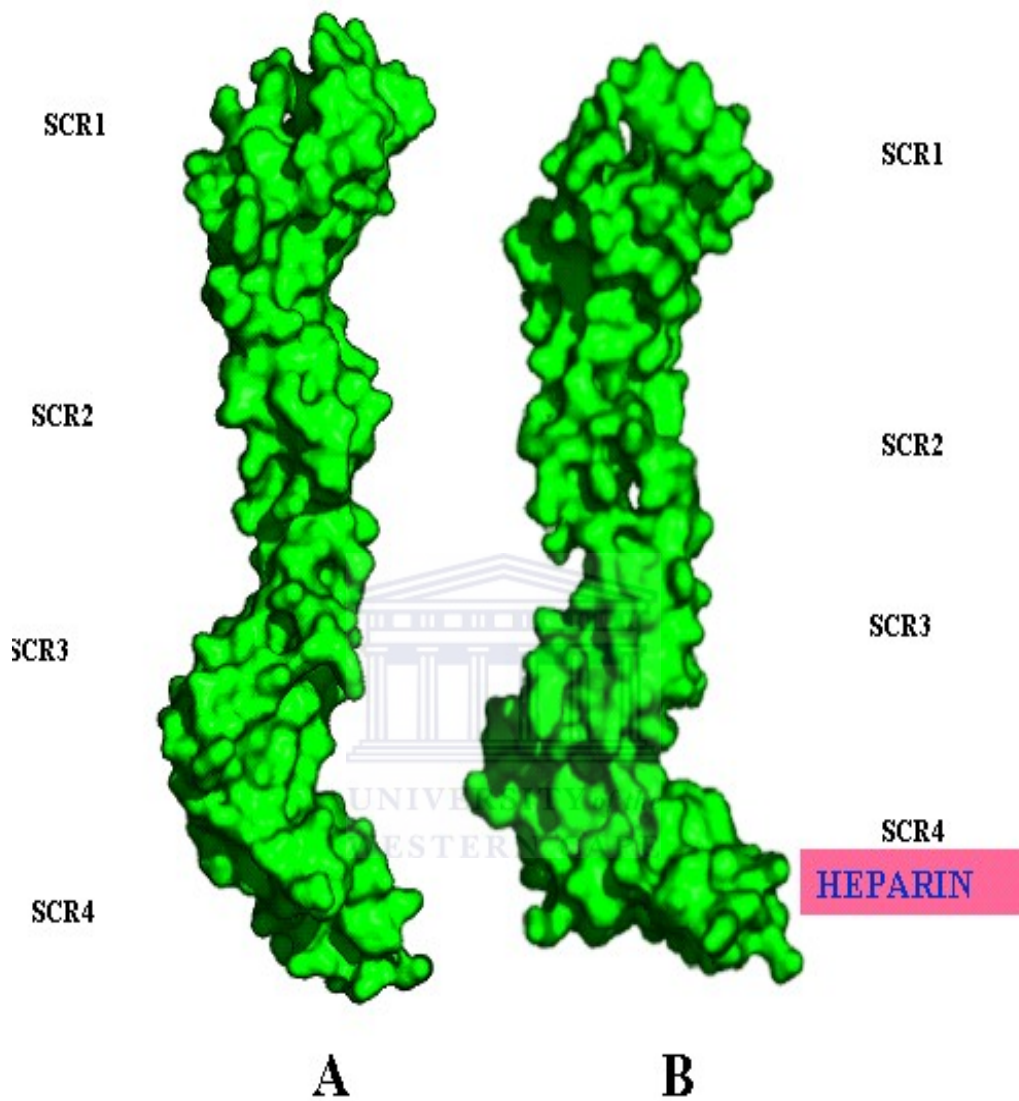
Crystal structures of both native [Murthy *et al.*, 2001] and VCP bound to heparin [Ganesh *et al*, 2004] have been solved (fig. 1.2). The relative orientations of the first 3 consensus modules (SCR 1-3) are very similar in both structures. There is, however, a significant difference in the orientation of the SCR 3-4. The tilt, twist and skew angles in the native VCP were 99, 3 and -37 degrees respectively [Murthy *et al*, 2001]. In the heparin bound complex, these values were 126, 31 and -51 respectively [Ganesh *et al*, 2004]. These

differences cause a motion of the SCR4 (in the heparin bound structure) towards the direction of the bound heparin, causing SCR 3 and 4 to come in close proximity to each other [Ganesh *et al*, 2004], with the SCR 3-4 linker serving as a hinge for the 29.5 degrees motion (fig 1.2 (b)).

The binding of heparin to VCP causes 22% increase in the surface area solvent accessibility. [Ganesh *et al*, 2004]. This substantial increase in the solvent accessibility suggests that heparin binding could cause an increase in the complement binding activities of VCP. The complex structure shows that heparin-binding site is located at the c-terminal end of the molecule (fig 1.2(b)), where there is a patch of positively charged lysine and arginine residues [Ganesh *et al*, 2004].







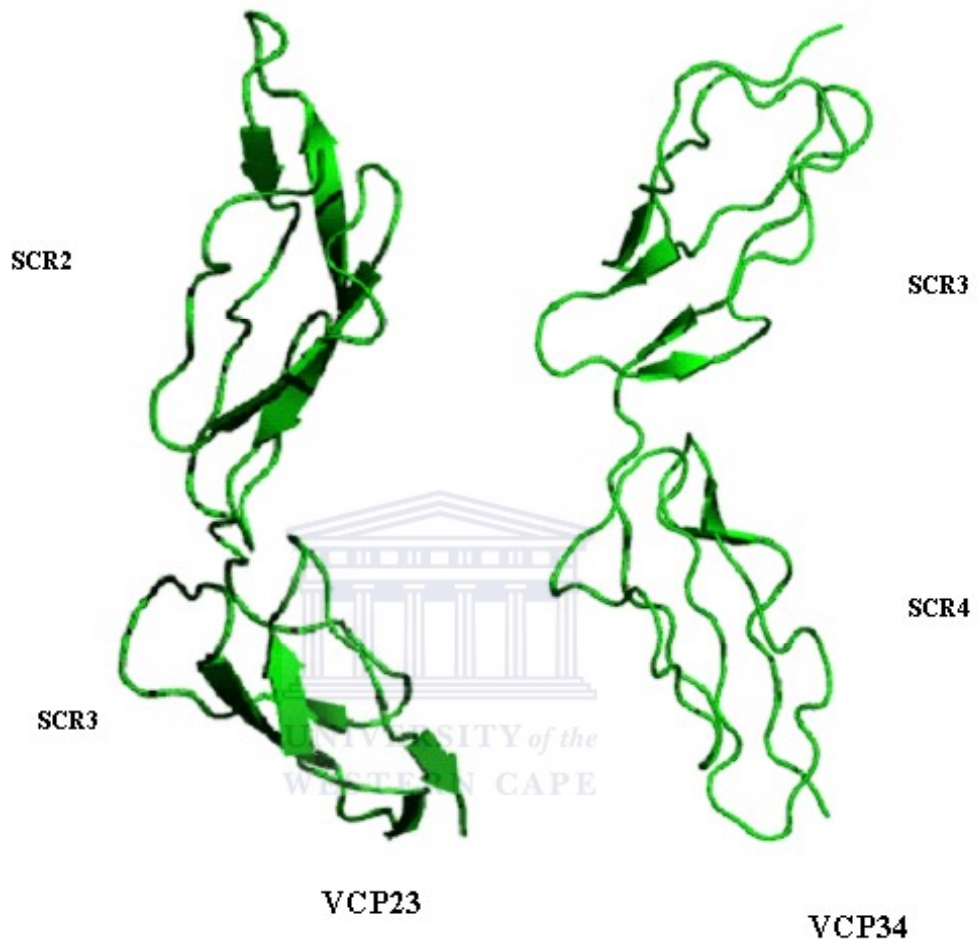
**Fig. 1.2:** Crystal structures of VCP, showing the spatial orientation of the modules. In the native structure (**A**), the modules are connected to neighbouring modules end to end with limited intermodular interactions. In the heparin bound complex VCP, (**B**), there is a greater interaction at the interface of SCR3-4. [Murthy *et al*, 2001; Ganesh *et al*, 2004]. The figure was generated by PYMOL molecular graphics system (2003) Delano Scientific, San Carlos, CA, USA.

#### **1.4.6 Putative complement binding sites of VCP**

Although the structure of VCP bound to the complement components like C3b has not been solved, the putative complement binding sites have been mapped [Murthy *et al.*, 2001]. This was made possible by the use of the information obtained from mutational studies of membrane cofactor protein (MCP) [Liszewski *et al.*, 2000], together with the high sequence conservation between MCP and VCP [Murthy *et al.*, 2001]. This region coincides with the highly concentrated positively charged region in the intermodular junction of SCR 1-2 as well as the lysine and histidine residues near the C-terminal of the SCR4 [Murthy *et al.*, 2001]. It thus appears that binding of VCP to heparin and complement proteins occur at the same region. All four CCP modules of VCP are required for complement binding [Rosengard *et al.*, 1999].

#### **1.4.7 NMR structures of VCP**

In addition to the known crystal structures of VCP, solution structures of fractions of the molecules have added to our understanding of the structural insights of VCP. The structures of two fragments of VCP have been solved. These are rVCP 3,4 [Wiles *et al.*, 1997] and rVCP2,3 [Henderson *et al.*, 2001].



**Fig. 1.3:** NMR structures of VCP fragments, showing the different intermodular interface extensions in rVCP 2,3 and rVCP 3,4. [Wiles *et al*, 1997; Henderson *et al*, 2001]. The figure was generated by PYMOL molecular graphics system (2003) Delano Scientific, San Carlos, CA, USA

Although the two solution structures show the same gross structural features of CCP modules, careful observation shows that there are some differences in spatial orientation of module 3 in the two structures [Henderson *et al.*, 2001]. There is a greater intermodular interaction between module 3 and 4 in rVCP 3,4 than it is between modules 3 and 2 in rVCP 2,3 [Henderson *et al.*, 2001].

The greater intermodular flexibility between modules 3 and 4 in the NMR structure (rVCP 3,4) looks very similar to that of the crystal structure of VCP bound to heparin. It appears, therefore, that the spatial orientation of a module of VCP (and CCP modules in general) depends, in part, on the presence or absence of neighboring modules, as has been suggested by several studies [Kirkitaдзе *et al.*, 1999(a, b, c & d); Henderson *et al.*, 2001].

#### **1.4.8 rVCP 2,3,4**

rVCP 2,3,4 is a truncated VCP in which the first consensus repeat is deleted. Lack of SCR1 makes rVCP 2,3,4 a 19.5kDa protein of 198 amino acid length. Residues 1-198 in rVCP 2,3,4 correspond to residues 65-263 in the full length VCP.

Lacking SCR1 makes rVCP 2,3,4 a non-complement binding protein since all 4 repeats are required for complement binding [[Rosengard *et al.*, 1999]. However, recent studies (Kotwal, unpublished) had shown that at high concentration of complement components such as C3b, rVCP 2,3,4 could interact and inhibit complement activation, probably by means of steric hindrance. Some fragments of VCP, including rVCP 1,2; rVCP 3,4 and rVCP 2,3,4 had been shown to have heparin binding activity [Smith *et al.*, 2000].

The structure of rVCP 2,3,4 would corroborate (or dispute) the notion that the overall structure of complement control proteins depends, in part, on its microenvironment [Kirkitadze *et al*, 1999(a, b, c & d); Henderson *et al*, 2001]. It would be interesting to see whether the absence of module 1 in rVCP 2,3,4 would tremendously affect its structure as compared to the full length VCP. It would also be interesting to see which region in the rVCP 2,3,4 does the interaction of the molecule with complement components occur.

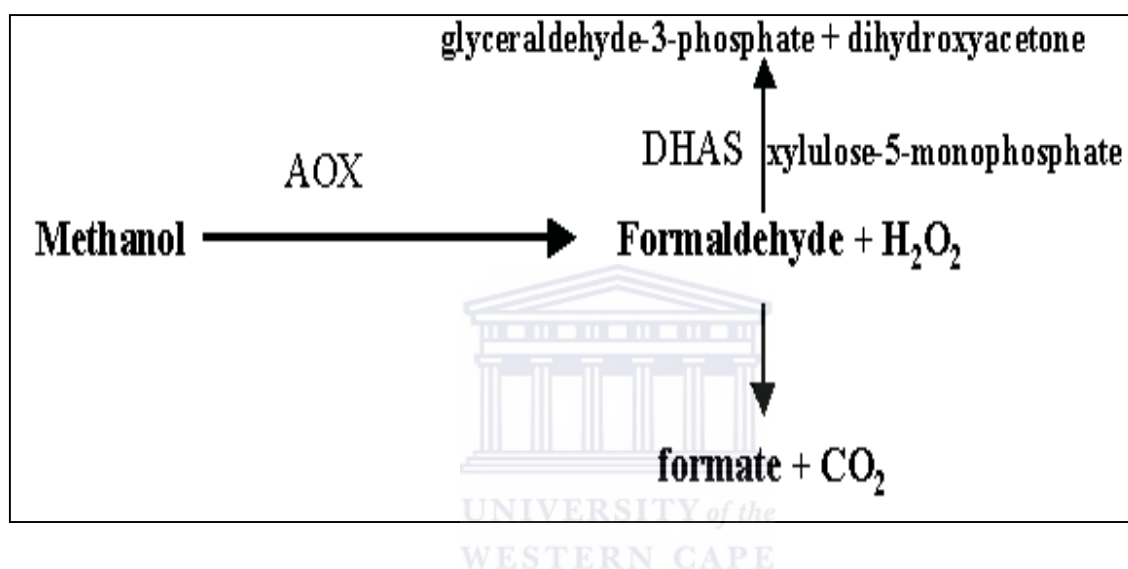
### **1.5.1 Expression of heterologous proteins using the *Pichia pastoris* system**

The *Pichia pastoris* expression system had become the choice for the expression of a number of proteins, including VCP. The increasing popularity of the use of *Pichia pastoris* system is due to its advantages over other well-characterized expression systems like the *Saccharomyces cerevisiae*. The advantages of the *Pichia* system include its possibility of obtaining high expression levels of both intracellular and extracellular proteins and its ability to perform posttranslational modifications.

*Pichia pastoris* expression system exploits the fact that some enzymes needed in methanol metabolism are produced in good quantities when the cells are grown on methanol as the only source of carbon [Veenhuis *et al*, 1993; Egli *et al*, 1980]. Two such enzymes are alcohol oxidase (AOX) and dihydroxyacetone synthase (DHAS).

The steps they catalyze are illustrated in fig. 1.4. AOX catalyses the first step of methanol utilization (namely its oxidation into formaldehyde and hydrogen peroxide). Part of the formaldehyde is further oxidized into formate and carbon dioxide (which serves as a source of

carbon for the energy requirement of the cell). The remaining formaldehyde enters another pathway in which DHAS catalyses the condensation of formaldehyde and xylulose-5-monophosphate eventually leading to the formation of glyceraldehyde-3-phosphate and dihydroxyacetone (fig.1.4)

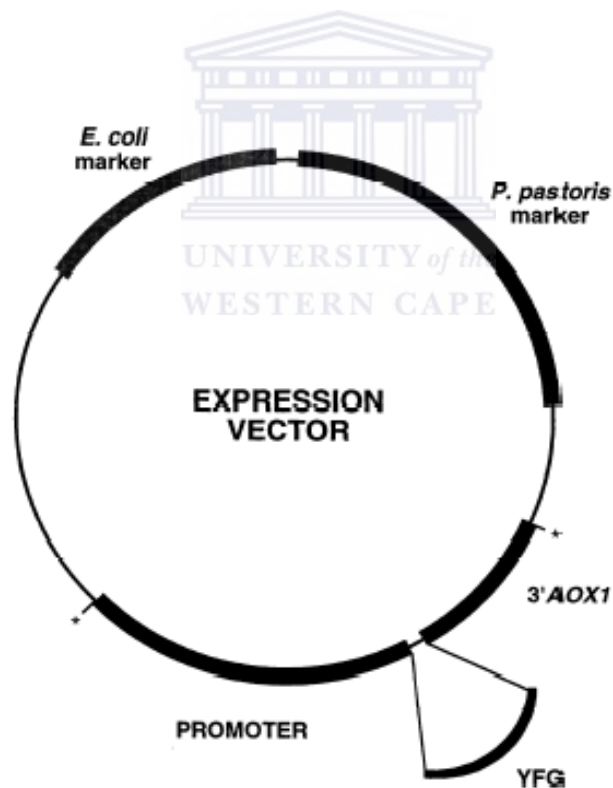


**Fig. 1.4:** Simplified metabolic pathways of methanol, showing the steps catalyzed by AOX and DHAS

### 1.5.2 Expression vector of *Pichia pastoris* system

In *Pichia pastoris*, two promoter genes (AOX1 and AOX2) encode alcohol oxidase but most of the alcohol oxidase activity is achieved by AOX1 [Tschopp *et al*, 1987; Ellis *et al*, 1985, Cregg *et al*, 1989]. In the expression vector of the *Pichia pastoris* system, the multiple cloning site for the insertion of the foreign gene is located between the AOX1 promoter and the terminator sequence (fig. 1.5).

Depending on whether the AOX genes are mutated or left intact, *Pichia pastoris* strains can further be classified into 3 types. The mut<sup>+</sup> (methanol utilization plus phenotype) is the wildtype. The KM71 strain has some mutations on the AOX1 [Cregg & Madden, 1987] and grows slowly on methanol and is therefore called mut<sup>s</sup> (methanol utilization slow phenotype) since it relies only on the weaker AOX2. Finally the mut<sup>-</sup> (methanol utilization minus phenotype) has mutations on both AOX1 and AOX2 genes and does not grow on methanol [Cregg *et al*, 1989].



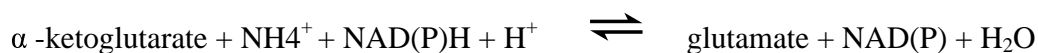
**Fig 1.5:** General expression vector of the *Pichia pastoris* expression system, showing the AOX1 promoter. The gene whose expression is to be induced is inserted at the multiple cloning site, indicated on the diagram as YFG (your favorite gene). [Cereghino & Cregg, 2000]

In spite of these mutations, the capability to induce expression at the AOX1 is retained in all three strains [Chiruvolu *et al*, 1997]. The mutated strains produce heterologous proteins better than the mut+ strains [Tschopp *et al*, 1987; Cregg *et al*, 1987; Chiruvolu *et al*, 1997] and they do not require large amount of methanol to induce expression.

Expression of foreign proteins using the *Pichia* system occurs in two major stages. In the first stage, the cells are grown on media containing glycerol as the only source of carbon. Protein expression is repressed at this stage but biomass accumulates. In the second phase, methanol is used to induce the expression of the protein.

### 1.6.1 Glutamate dehydrogenase

Glutamate dehydrogenase (GDH) is a mitochondrial enzyme and is one of the most important enzymes involved in the tricarboxylic acid (TCA) cycle in the metabolic pathway. It catalyses the reversible conversion of  $\alpha$ -ketoglutarate into glutamate by incorporating nitrogen into the  $\alpha$ -ketoglutarate in the reaction represented below:



[Garett & Grisham, 1995].

Glutamate dehydrogenase uses NADH or NADPH as a coenzyme. The reverse reaction furnishes carbohydrate for cells. The direction of the reaction depends on the concentration of the available

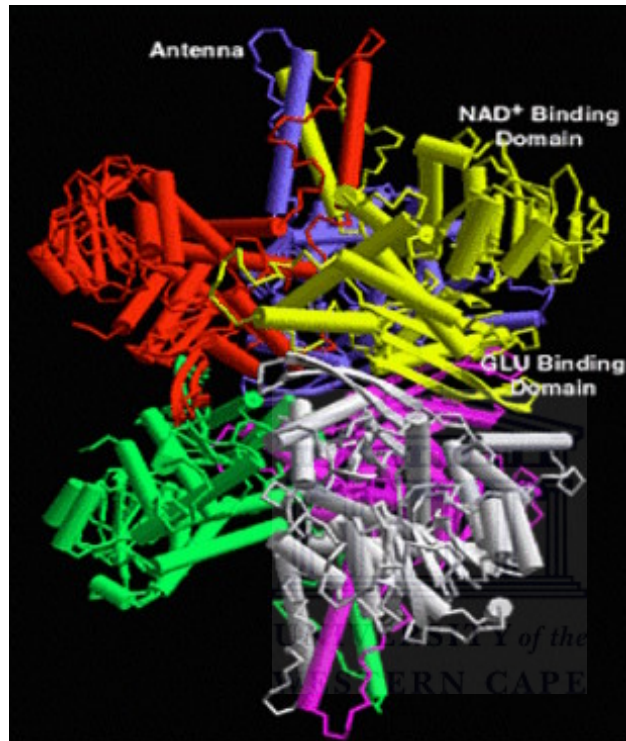


substrates,  $\alpha$ -ketoglutarate and glutamate [Joe *et al.*, 1994]. Both NADH and NADPH-dependent GDH have been found in *Bacteriodes fragilis* [Yammamoto *et al.*, 1987]; an obligate anaerobe found primarily in the abscesses of abdominal infections. The family of glutamate dehydrogenase enzymes shares similar structural and functional properties. Structurally, members in the family can be subdivided into the hexamers and tetramers, depending on whether the functionally active enzyme is made of six or four subunits of the enzyme, with the average molecular weight of a subunit being 48 kDa [Britton *et al.*, 1992]. The active site of the enzyme is usually located in a cleft created by the two monomer units involved in glutamate and NAD(P) binding.

### **1.6.2 3-D structures of glutamate dehydrogenase**

Several crystal structures of GDH from different sources (with or without their coenzymes) have been solved. Among them are those from *Clostridium symbiosum* (PDB ID: 1AUP) and bovine GDH (PDB ID: 1HWZ), two of the most studied GDH's. Sequence studies coupled with electron density maps of solved structures indicate that the site of substrate and coenzyme binding is similar in GDH.

An initial electron microscopy study on of *Bacteriodes fragilis* NADH-dependent GDH has indicated that it is a hexamer with a 3<sub>2</sub>-symmetry (unpublished results). The hexamer is arranged in two stacked trimers. These initial structural insights have neither been confirmed nor disproved, as there is no atomic resolution structure of the *Bacteriodes fragilis* NADH-dependent GDH yet.



**Fig. 1.6:** The structure of NAD dependent GDH, showing all the entire hexamer. The identical monomers are coloured differently. The active site is located in the cleft located between the glutamate binding domain and NAD binding domain. The figure was obtained from <http://www.danforthcenter.org/smith/gdh.htm>.

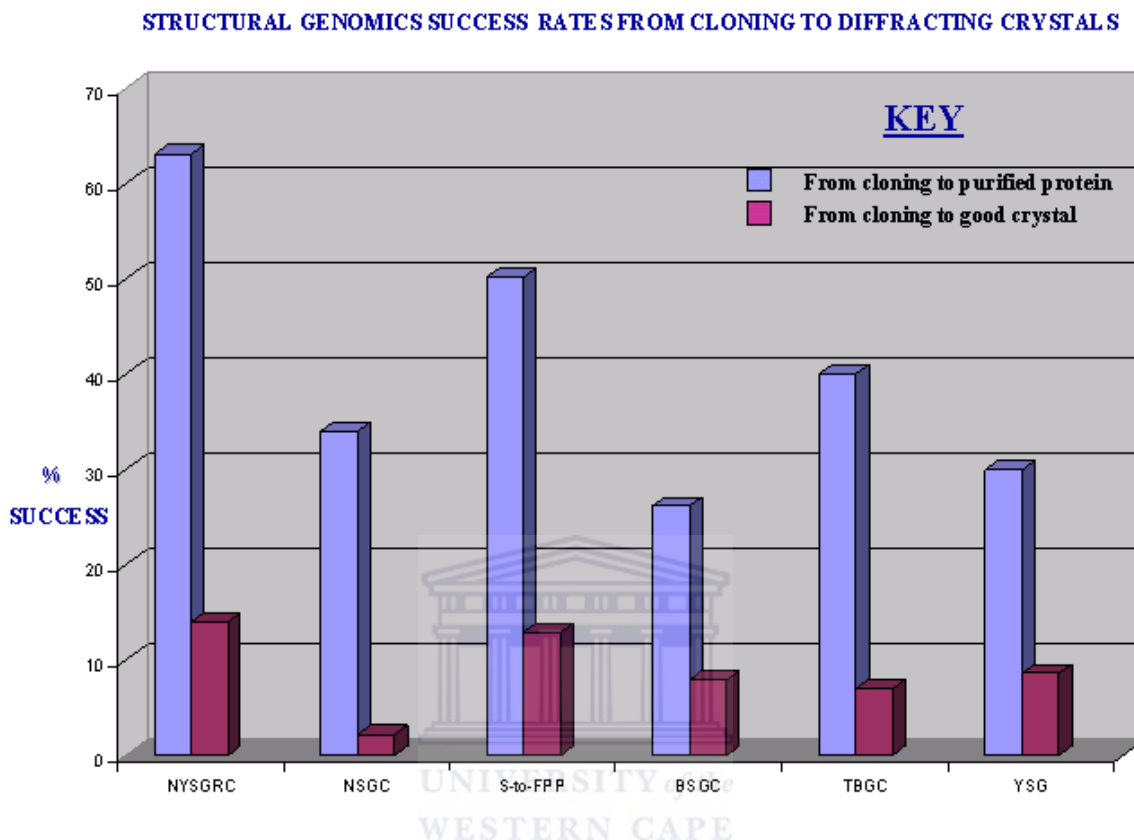
## **1.7. Protein crystallization**

### **1.7.1 Crystallization is the rate-limiting step in structural biology projects**

Structural biology projects involve several steps in order to get from gene to structure. The major steps involved in any structural biology projects are as follows: gene cloning, protein expression, protein purification, crystallization and structure determination. Structural biology projects worldwide have made great inroads in the steps prior to the crystallization process. There is tremendous success rate in the cloning, expression and purification steps. This however, had not been translated into comparable success rate in crystal formation. Crystallization is arguably considered the most difficult and rate determining step in structural biology [Vekilov & Chernov, 2002], although there could be some problems with protein expression and purification.

Figure 1.7 below shows the success rates of structural biology projects from gene cloning to crystal formation, taken from different projects worldwide. As seen from the fig. 1.7, on the average, more than 40% of all cloned target proteins are expressed and purified. On the other hand, the success rate from cloning to the formation of good, diffracting crystal is barely 7% (on the average). Some structural genomics projects (fig. 1.7) show as low as 3% success rate of producing diffracting crystals. In most crystallization trials, even if crystals are produced, their quality is not good enough to diffract to any appreciable resolution for structure determination [Chayen, 2002]. Several optimization procedures have to be tried in order to produce crystals that diffract well.

**Fig. 1.7:** Structural biology success rate from cloning to getting diffracting crystals



**ABBREVIATIONS & SOURCE:**

NYSGRG: New York Structural Genomics Research Consortium: [www.nysgrc.org](http://www.nysgrc.org)

NSGC: Northeast Structural Genomics Consortium: <http://www.nesg.org/>

S-to-FPP: Structure to Function Pilot project. <http://s2f.carb.nist.gov/>

BSGC: Berkeley Structural Genomics Centre: [http://www.strgen.org/status/progress\\_totals.html](http://www.strgen.org/status/progress_totals.html)

TBGC: TB Structural Genomics Consortium: <http://www.doe-mbi.ucla.edu/TB/currently.php>

YSG: Yeast Structural genomics. <http://genomics.eu.org/HAL/hal-public/targets.html>

### **1.7.2 Inhibition of crystallization by evolutionary negative design**

Recent technical innovations and methods have helped a great deal in saving time and labour and protein material; thus improving the success rate of crystallization [Bergfors, 1999]. These advances, notwithstanding, have not fully explained the fundamental problem of protein crystallization, the fact that most proteins have the natural tendency not to crystallize and have to be carefully induced to do so [Jonathan *et al*, 2004]. A current hypothesis suggests that most proteins have evolved to avoid aggregation (including crystallization) because aggregation is detrimental to the viability of most cells [Jonathan *et al*, 2004].

### **1.7.3 Evidence to support the negative evolutionary design of protein crystallization**

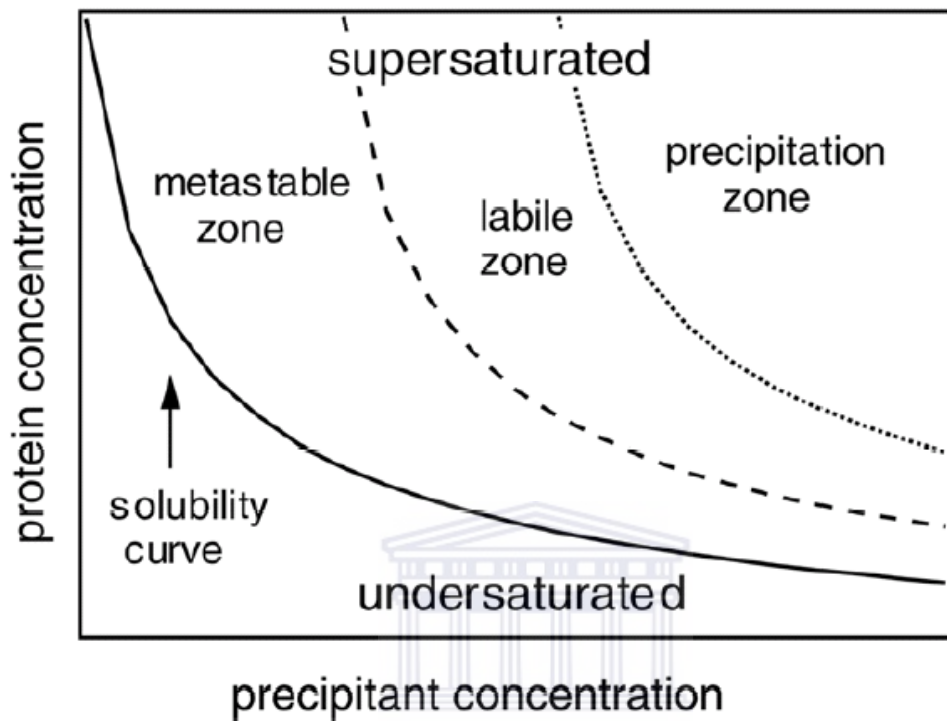
Protein aggregation in many cells presents a potential threat to the viability of the cell. For examples Alzheimer's diseases, systemic amyloidosis, Parkinson's disease, Huntington's disease and Creutzfeldt-Jakob disease are all caused by the aggregation of non-native protein structures [Kopito, 2000]. Not only misfolded proteins but also aggregation of proteins in their native states is also detrimental to the viability of some cells. For example, sickle cell anaemia is caused by the coalescing of a mutant form of haemoglobin into a well-ordered fibrillar aggregate in the red blood cells. In a few instances, certain diseases are caused by the production of protein crystals in cells. For examples, certain forms of cataract are caused by the crystallization of a mutant form of gamma crystallin [Pande *et al*, 2001] and certain forms of anaemia are caused by production of mutant haemoglobin [Vekilov *et al*, 2002].

Diseases associated with crystallization of proteins are not as common as those associated with non-native protein aggregates. This does not, however, mean that crystallization of proteins in cells do not pose a potential threat to the viability of most cells. One reason that may account for the differences in frequency is that the ordered crystals are more amenable to evolutionary control than the less ordered non-native protein aggregates [Jonathan *et al*, 2004].

Further evidence to support the negative evolutionary design of protein to avoid aggregation could be seen from a class of proteins with large  $\beta$  secondary content [Jonathan *et al*, 2004]. The edges of  $\beta$  sheets provide the natural sites for their association with other neighbouring beta sheets. This in turn can lead to the formation of extended structures found in amyloid plaque, for example [Jonathan *et al*, 2004]. Since proteins are evolved to avoid any form of aggregation, a number of negative design strategies have evolved to protect the edges of the beta sheets from forming extended structures [Richardson *et al*, 2002]. This includes the formation of beta barrels; which ensures a continuous beta sheet without any edges.

### **1.8.1 The crystallization process**

Crystallization is a phase phenomenon. The crystallization phase diagram is a map that represents the state transition as a function of variables that affect crystallization; such as temperature, pH, ionic strength, precipitant concentration, protein concentration etc. However, to simplify this multi-dimensional diagram, a 2-dimensional diagram representing the phase transition as a function of the concentrations of protein and precipitants are illustrated in fig. 1.8 below.



UNIVERSITY of the  
WESTERN CAPE

**Fig. 1.8:** Phase diagram showing the various phase changes during the process of crystallization. Aggregation begins to occur in the supersaturation region. Nucleation occurs at higher supersaturation state while growth occurs in lower metastable state. At very high supersaturation state, precipitate formation is favoured.

A phase transition diagram can broadly be divided into two main regions, namely the undersaturated and saturated region (fig 1.8). Protein molecules will continue to remain in solution in the undersaturation state. Hence, crystal formation does not occur within this region.

### 1.8.2 Nucleation state

When the solubility of a protein solution is brought above its solubility limit, it becomes supersaturated and conducive to aggregation of any form (crystals or precipitates) to occur. Within the saturated region, there are nucleation (or labile), metastable and precipitation zones (fig 1.8).

The nucleation zone (or labile zone) is located above the metastable zone. In theory, crystals are expected to grow when a solution reaches its supersaturation point. However, this rarely occurs in practice. Crystals are formed only when the protein concentration is at least three fold the amount needed to cross the threshold solubility [Chernov, 1997]. It is only at this high supersaturation state (nucleation zone) would nucleation begins to occur. In order to come out of solution and form a crystal, the protein molecules have to overcome a high activation energy barrier. The high supersaturation enables this high activation barrier to be overcome [Kashchiev, 2000].

Nucleation is a very slow process and can take a lot of time because of this high-energy barrier to cross. If the supersaturation is too low (in the metastable state), it would correspond to a very slow attainment of nucleation and crystals might not be formed within a reasonable length of time. Higher supersaturation can be induced using a variety of precipitants. The most successful precipitants used in protein crystallization include ammonium sulfate, polyethylene glycols, methylpentanediol; and chlorides of sodium, magnesium and calcium. If supersaturation is just right (in the nucleation zone), spontaneous nucleation will occur and crystals would be formed.



On the other hand, if supersaturation were too high above the rate required for nucleation, disordered aggregates and precipitates would be formed [Ivana, 2002].

During nucleation, the protein molecules interact and associate until a certain critical size aggregate is reached. The aggregation process is not fully understood but is known that for aggregates to lead to crystals, it must not occur before saturation [Mikol *et al.*, 1990; Wilson, 1990]. Therefore, presence of aggregates in protein solution before the crystallization process is highly detrimental and reduces the chances of crystallization.

### **1.8.3 Thermodynamics of nucleation of proteins:**

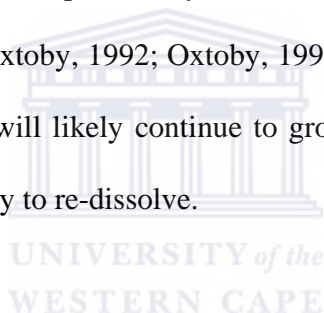
There is an energetic barrier that must be surmounted in order to induce the formation of stable nucleation. Nucleation of proteins occurs at high supersaturation levels because of this energy barrier. [Drenth & Haas, 1998]. Molecules of protein in solution (growth units) move freely in the solution and collide inelastically with each other. As collision occurs, some of the colliding molecules converge and forms a cluster, whose lifespan depends on the relative strength of attracting forces that keep the cluster together and repulsive forces that pull them apart.

Proteins are very heterogeneous macromolecules which exhibit definite hydrophobic and hydrophilic functional groups called patches, which are used in macrobond formation. However, not all the patches take part in the macrobond formation. Proteins have short ranged attractive forces. Therefore, macrobond formation occurs only between patches that are complementary [Chernov, 2003]. This could explain why nucleation constant rates of proteins are very high compared to small molecules; even though the energy barrier of crystal nucleation is of the same

order of magnitude for both protein and small molecule [Nanev, 2007]. The incidence of protein nucleation, therefore, depends in part on the number of patches with proper orientation.

If the attractive forces are stronger than the repulsive forces, the cluster stays together and continues to grow to form stable nuclei by the addition of single growth units [Feher & Kam, 1985]. On the other hand, if the repulsive forces are stronger than the attractive forces, the cluster re-dissolves.

At a certain cluster size, called the critical size, there is a balance between the attractive and repulsive forces. At the critical size, the probability of the cluster growing out of solution and re-dissolving into solution is equal [Oxtoby, 1992; Oxtoby, 1998]. Therefore, any cluster of growth units bigger than the critical size will likely continue to grow spontaneously while any cluster smaller than the critical size is likely to re-dissolve.



As can be seen from the phase diagram (figure 1.8), the metastable zone is bounded by two curves. The lower curve is called solubility curve and the upper curve is called the supersolubility curve. Along the solubility curve, the solution is in equilibrium thermodynamically. The probability of nucleation occurring along this curve is zero. In other words, induction time is infinite. Along the supersolubility curve (upper limit of the metastable zone), the probability of nucleation event occurring is one (thus induction time is 0). Within the metastable zone, nucleation would occur given enough time. Induction time for the metastable zone is therefore between zero and infinite. The length of the induction time depends on the rate

at which supersaturation is achieved. The faster the supersaturation, the shorter the induction time and vice versa.

The width of the metastable zone is determined by the position of the supersolubility curve and not by the solubility curve. In any given system, the solubility curve is fixed while the position of the supersolubility curve depends on the rate at which supersaturation occurs. If the system traverses the metastable zone too quickly, higher supersaturation levels will be reached before nucleation occurs, leading to the formation of amorphous precipitates.

The discussion thus far had assumed that nucleation is homogeneous where the probability of a given fluctuation is identical in the entire volume of the system. Theoretically, this occurs in highly pure crystallization solutions with no trace of foreign material. Under normal laboratory practices, this is often not achievable (and sometimes not desirable). Presence of foreign materials such as impurities, dust particles and container surface could cause local fluctuations within the system, resulting in heterogeneous nucleation. Heterogeneous nucleation results in the creation of a new phase which the available phase; thereby reducing the amount of work that needs to be done in order to create the critical cluster size. Consequently there is an increase in the probability of nucleation occurring within the local area compared to other regions in the bulk. This partly explains the use of nucleants to enhance nucleation of proteins. The magnitude of energy reduction due to presence of foreign surfaces is proportional to the wet angle [[Chernov, 1984].

#### **1.8.4 Formation of precipitates**

Since the rate of crystal growth is much slower than that of amorphous precipitate [Ivana, 2002], the protein should be brought to supersaturation very slowly in order to prevent the formation of amorphous precipitate. If supersaturation were attained too quickly, amorphous precipitates would be formed. Amorphous precipitation is predominantly favoured when the protein concentration is well over its saturation [Ivana, 2002]. Precipitate formation could be reduced by reducing either the concentration of precipitants or protein (or both) in order to bring about attainment of supersaturation more slowly.

#### **1.8.5 Growth of crystals**

Following nucleation, more protein molecules associate with the already formed nucleated sites, resulting in a decrease in concentration of the protein in solution. As the concentration of the protein decreases, the system is brought to a less supersaturated state, the metastable state (fig. 1.8). In the metastable stage, crystal growth occurs without further nucleation. Although the metastable state is a supersaturated state, the concentration of protein is not high enough for spontaneous nucleation to occur [Chernov, 1997]. The size of the crystals would depend, among other things, the number of nucleation sites upon which growth can occur, the concentration of the protein or presence of materials that can inhibit further growth.

#### **1.9 Effect of pH and salt on crystallization of protein**

Since saturation is a necessary requirement for protein crystallization, any factor that affects the solubility of protein can also affect crystallization. Protein solubility is highly influenced by the pH of the solution. Proteins have minimal solubility at the isoelectric point (pI); the pH at which all the positive species in a protein becomes equal to the negative species. Most proteins have

been crystallized around a pH of 6.5 even though the correlation between pI and the pH at which protein crystallization occurs is generally weak [Rupp, 2001]. Because of the strong dependence of solubility on pH, one of the strategies in (protein) optimization is to change the pH of the buffer.

Salts are responsible for ionic strength of a solution. In small concentrations, salts tend to increase the solubility of most proteins (salting in). However, in high salt concentrations, the solubility of protein decreases (salting out). In the presence of high salt concentration, the solubility of proteins is not dependent on pH as the salt screens the electrostatic interactions between proteins [Curtis et al, 1998].

### **1.10.1 Methods of crystallization**

Several crystallization methods can be employed to crystallize proteins. These include vapour diffusion method, batch method, free interface diffusion method and dialysis method [Ducruix & Giege, 1992]. The most frequently used methods are vapour diffusion and microbatch (adaptation of the batch method).

In vapour diffusion, a small droplet (usually 1-10 $\mu$ l) of protein sample is mixed with similar amount of crystallizing solution (containing the precipitant, buffer, salt or additive, if applicable). This is placed on a siliconised cover slip, inverted and sealed over 1ml of the precipitant solution. The difference in concentration between the reservoir and

the drop drives the system to equilibration via the vapour phase. This process concentrates the protein in the drop until saturation is reached, when conditions may become right for protein to come out of solution. Vapour diffusion can be achieved by hanging drop (fig 1.9), sitting drop (fig. 1.11) or the sandwich techniques (fig 1.10); depending on whether the drop is hanging from the cover slip, sitting on a microbridge [Harlos, 1992] or sandwiched in-between two cover slips [Fox & Karplus, 1993].



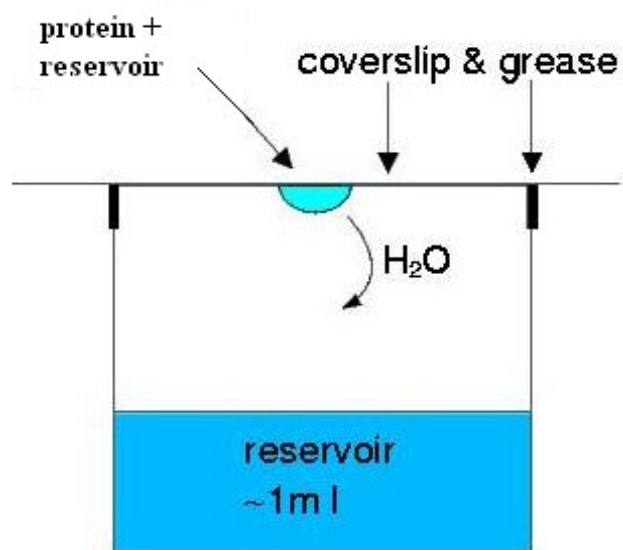


Fig. 1.9: Hanging drop vapour diffusion method of crystallization

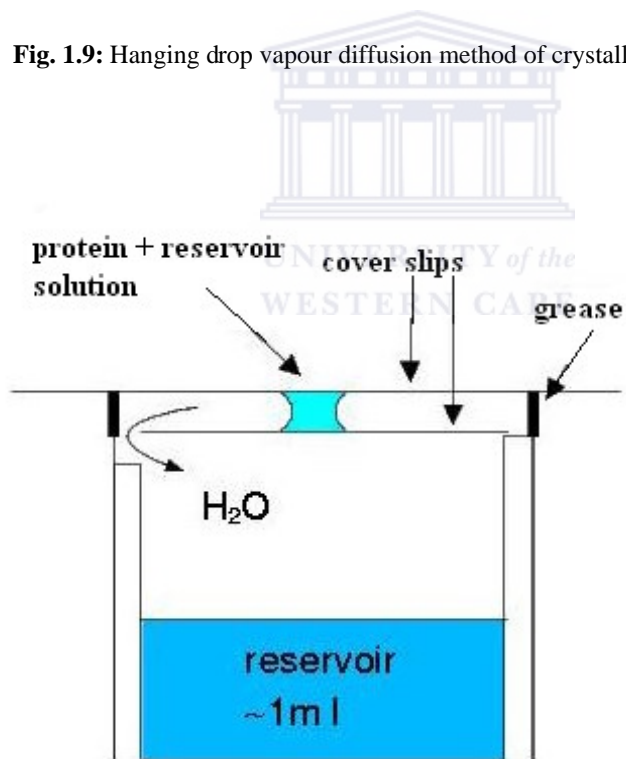
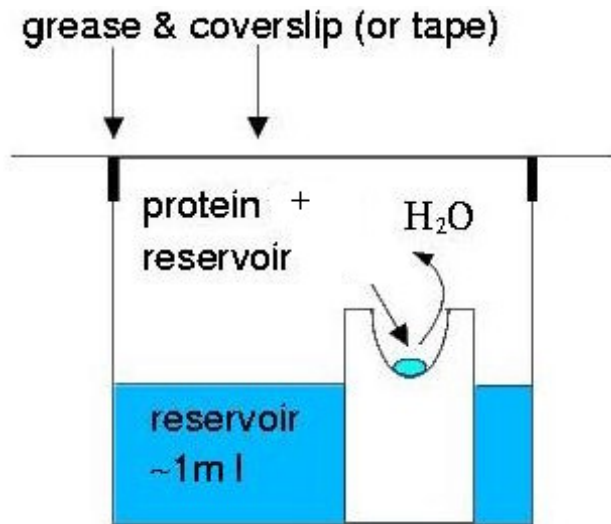
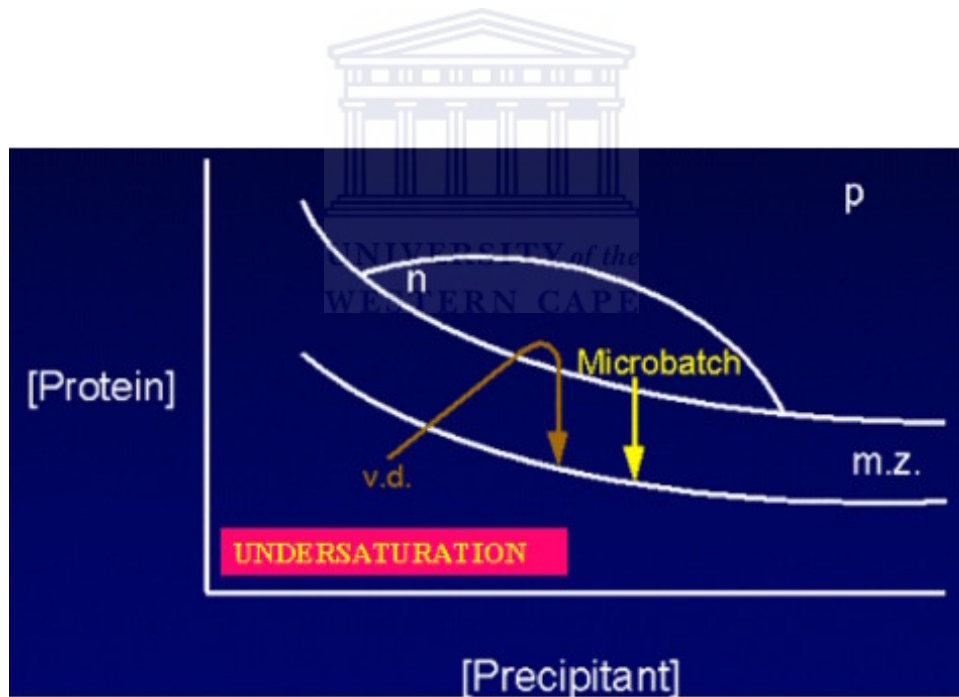


Fig. 1.10: Sandwich drop vapour diffusion method of crystallization



**Fig. 1.11:** Sitting drop vapour diffusion method of crystallization

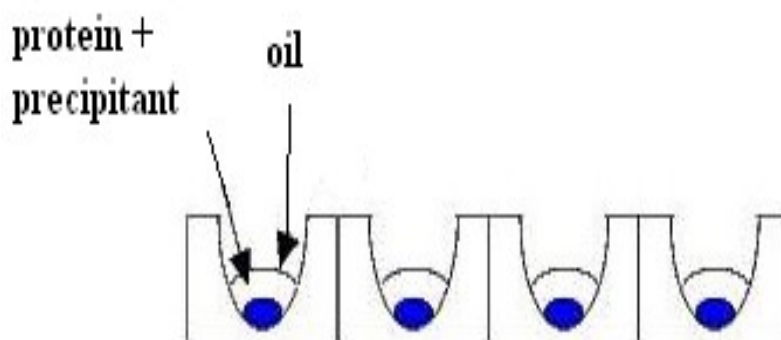


**Fig 1.12:** Phase diagram comparing protein concentration gradient of vapour diffusion (V.D) and microbatch processes. In V.D, supersaturation is achieved during the crystallization progress whiles in microbatch supersaturation is achieved at the beginning of the experiment. P = precipitation zone. M.Z = metastable zone. N = nucleation stage.



In batch crystallization, the protein is mixed with the precipitating agent at their final concentration so that supersaturation is achieved at the beginning of the experiment (fig.1.12). These conditions would not change until the protein begins to come out of solution as crystals or as precipitates. Thus, the system does not search through different conditions, unlike the vapour diffusion method. This is advantageous because it enables the experimenter to know precisely what conditions produced a crystal (if crystals are produced). One disadvantage of the batch crystallization is that it requires large quantities of materials.

The requirement of large quantities of materials used in the batch crystallization has been overcome by the modification employed in the microbatch method. Microbatch aims at reducing the consumption of sample by making use of small volumes of materials. As little as 0.5 $\mu$ l of material can be dispensed under oil in microbatch crystallization (fig. 1.13) [Chayen *et al*, 1990; Chayen *et al*, 1992; Chayen *et al*, 1994].



**Fig. 1.13:** Microbatch crystallization method

### **1.10.2 Comparative advantages and disadvantages of microbatch & vapour diffusion methods.**

There are no marked differences in the crystallization success rate between vapour diffusion and microbatch methods [Chayen *et al*, 1990]. Notwithstanding the fact that each method has its own merits and demerits, most of the proteins that have been crystallized by the vapour diffusion method had also been successfully crystallized by the microbatch method, sometimes with only a slight modification [Chayen, 1998]. Exploring the full coverage of both methods, therefore, seem the best way to succeed in obtaining crystals [Baldock *et al*, 1996].

### **1.10.3 Problems with microbatch crystallization method**

Microbatch method does not usually work in cases where volatile organic precipitant are used. Largely for this reason, microbatch methods had traditionally been excluded from the crystallization of membrane proteins because of the potential of detergent loss. The exclusion of membrane protein crystallization by the microbatch method is gradually fading away. A membrane protein has been crystallized by the microbatch method [Hankamer *et al*, 1992], but only as a last resort when all other methods had failed to yield crystals. Information gathered from that trails showed that detergent in membrane protein does not usually lost unless there is vigorous mixing between the oil and detergent. Furthermore, the presence of oil was thought to act as a stimulant, causing the slow absorption of detergents. This would enhance the reaching of supersaturation [Hankamer *et al*, 1992]. These findings suggest that microbatch method can be useful for membrane protein crystallization.

Microbatch method also has a problem of shock nucleation [Saridakis *et al*, 1994]. Since the protein is in contact with precipitating agent at their final high concentration, it can cause shock

nucleation, leading to the formation of showers of crystals, instead of a few large ones [Saridakis *et al*, 1994].

Harvesting crystals under oil in microbatch crystallization is more difficult than on the cover slide from vapour diffusion methods. This is because crystals under oil usually stick to the walls of the supporting vessel. However, advances made in comprehensive harvesting protocol [Shaw-Stewart & Conti, 1995] had marginally reduced this problem.

#### **1.10.4 Advantages of the microbatch method**

Microbatch method, microbatch methods provide a means to control mechanical shock and prevent the formation of heterogeneous crystals [Chayen *et al*, 1993; Blow *et al*, 1994; Chayen, 1996] since the drop is dispensed under oil. This also reduces the surface area of the drop available for contact with the walls of the containing vessel [Yonath *et al*, 1982].

Another area where microbatch has an advantage over vapour diffusion method is that it allows the use of precipitating agents such as polyethylene glycol and volatile solvents without any problem offered by the vapour diffusion method. In vapour diffusion, volatile precipitating agents may be absorbed [Yonath *et al*, 1982]. This may cause the drop volume to enlarge, causing dilution of protein concentration in the drop. This can cause already formed crystals to dissolve [Conti *et al*, 1996].

Setting up microbatch experiment is quicker, simpler and requires small quantities of materials compared to vapour diffusion methods. In addition, the crystallization drop is protected by the oil, resulting in the formation of stable crystals, which can be transported over long distance

without being damaged. The oil also protects the drop from any contamination like dust that may form on the surface of the drop. Stable crystals can be produced by the microbatch method since the contact between the crystal and oil makes it difficult for the crystal to be influenced by slight temperature changes [Chayen, 1999].

On one hand, the fact that vapour diffusion is a dynamic, self-searching process makes it advantageous over microbatch since equilibration and supersaturation are reached quite slowly [Luft *et al*, 1994; Luft *et al*, 1996]. On the other hand, the self-searching property of vapour diffusion makes the crystallization process control difficult to control once the experiment had started. The dynamic nature of vapour diffusion makes difficult the modeling of the crystallization process [Ataka & Tanaka, 1986; Ataka 1993; Saridakis *et al*, 1994]. Another advantage of vapour diffusion over microbatch method is that the concentration of the reservoir components can be altered and the process intervened by the experimenter in the course of the experiment, with very little or no disturbance to the drop [Yonath *et al*, 1982; Pryzbylska, 1989]. This can conveniently be done by transferring the drop to another reservoir of different concentration, [Chayen *et al*, 1989].

#### **1.10.5 Adaptation of vapour diffusion to microbatch crystallization and vice versa**

In the microbatch method, the crystallization components are mixed to their final concentration at the beginning of the experiment (fig. 1.12). This implies that there is negligible or no concentration changes once the drop is sealed under oil. Therefore, if a crystal is formed under such circumstances, the protein concentration can be assumed to be half the concentration of the original protein solution (since equal volumes of protein and precipitants solution were used in

the drop). Therefore, in adapting a microbatch method to vapour diffusion method, the concentration of protein solution to be used should be about half the concentration that was used to produce a crystal in the microbatch method. For example, if a protein concentration of 10mg/ml was used to produce a crystal in a microbatch method, this value must be reduced to about 5mg/ml to simulate similar crystallization conditions when adapted to a vapour diffusion method.

Conversely, when a protein and precipitants are mixed and sealed over a well in vapour diffusion method, the drop concentrates; making the protein concentration reach a final concentration approximately equal to its original concentration before mixing with precipitants. Therefore when a crystal is formed in vapour diffusion, the concentration in the drop can be fairly assumed to be equal to its original concentration. Hence, in adapting a vapour diffusion method to microbatch method, the protein concentration to be used should be twice the concentration that was used to produce a crystal in the vapour diffusion. However, it had been observed that in certain cases, where crystallization proceeds very rapidly, crystal formation occurs before equilibration is reached [Mikol *et al*, 1990]. In such cases, adaptation to the microbatch method requires significantly lower concentration of protein and precipitant [Chayen, 1998]. Concentrations of additives and buffers do not change significantly [Chayen, 1998] and can be used at the same concentration when the system is adapted from batch to vapour diffusion or vice versa.

### **1.11.1 Crystal optimization**

More often than not, initial crystallization trials yield too many, small sized crystals with low diffracting quality. Using these initial crystallization conditions, various crystallization methods and conditions can be explored to optimize the crystals, in an attempt to produce larger, diffracting crystals.

Larger crystals are better than smaller ones for structural determination for a number of reasons. A larger crystal scatters far more x-rays (relative to their background) than smaller crystal. Larger crystals, therefore, tend to have higher signal to noise ratios. At the refinement stage of data processing, larger crystals refine more easily than smaller crystals do, even when there is little difference in the R-factors between the two data sets [Rayment, 2002].

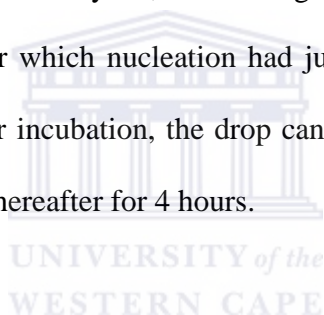
### **1.11.2 Separation of nucleation from growth phase**

One way to control the crystallization process, and therefore produce few, large crystals that diffract well, is to attempt to separate the nucleation phase from the growth phase. Several methods are available for achieving this goal. These include changing the crystallization temperature [Rosenberger *et al*, 1993; Haire, 1996], seeding and diluting the drop after nucleation had just taken place etc [Saridakis *et al*, 1994; Saridakis & Chayen, 2000]. These interventions are quite effective in improving crystal quality but they are very time consuming and based on trial and error (especially temperature changes [Saridakis *et al*, 2002]).

### 1.11.3 Drop dilution as a means to control nucleation

Over the past years, the determination of an ideal time for diluting a drop had been based on the physical appearance of the first crystal in the drop [Saridakis *et al*, 1994]. Unfortunately, the critical size of a crystal to be seen under the laboratory microscope is in the order of 5 microns. As such, by the time of the first visible crystal, the nucleation process might be too far advanced, resulting in the formation of showers of crystals and thus making it quite late to intervene [Saridakis *et al*, 1994].

In spite of this timing inaccuracy, if the crystallization process is intervened at different time intervals within the time of first visible crystal, there is a greater possibility of finding the right time of intervention; the time after which nucleation had just taken place. For example, if the first crystal is seen in 4 hours after incubation, the drop can be diluted after the first 30 min of incubation and then every 30 min thereafter for 4 hours.



The use of dynamic light scattering (DLS) these days had made the determination of the ideal time of dilution more accurate. DLS can resolve particles of size three orders of magnitude below an optical microscope [Saridakis *et al*, 2002]. Due to its sensibility to particle size variation and interactions with protein particles in solution [Schmitz, 1990] and its non-invasiveness, DLS had become an important tool in following the events of crystallization [Saridakis *et al*, 2002] and proved to be effective to predict accurately the right time of the beginning of nucleation [Malkin & McPherson, 1993].

#### **1.11.4. The use of oils**

The fundamental cause of excessive nucleation is that supersaturation is approached too quickly. One way to control the rate of nucleation is the use of oils. The use of oils has the effect of reducing the rate of evaporation of water, thus enabling supersaturation to be reached more slowly.

A combination of silicone and paraffin oil to achieve a balance between rapid evaporation and no evaporation at all had been proved very useful [Chayen, 1997]. Water evaporation through paraffin is negligible and this makes paraffin act as a sealant. On the other hand, silicone oil allows free diffusion of water. Therefore, by mixing paraffin and silicone, partial evaporation is achieved, the rate of which depends on the proportion of each oil used [D'Arcy *et al*, 1996].





## CHAPTER 2: METHODOLOGY

### 2.1 Prediction of likely pH ranges for crystallization of rVCP 2,3,4 & GDH

To improve the efficiency of screening and optimization, a program called CrysPred [Kantardjieff & Rupp, 2004; Kantardjieff *et al*, 2004] was used to predict the pH ranges most likely favourable for the crystallization of rVCP 2,3,4 and GDH based on its isoelectric point. This would help reduce materials and random experimentation during optimization of crystallization trials.

The program takes as input the amino acid sequence of a protein or its isoelectric point (pI) and outputs a graph of predicted success frequency of crystallization versus (pI-pH) values of crystallization success. This prediction is based on the correlation between crystallized proteins in the PDB and the pH units away from the pI at which they were crystallized [Kantardjieff & Rupp, 2004; Kantardjieff *et al*, 2004].

### 2.2 Expression of rVCP 2,3,4

#### 2.2.1 Preparation of expression media

rVCP 2,3,4 was expressed using the *Pichia pastoris* yeast expression system (*Pichia pastoris* catalogue 2000-2002). The following reagents were prepared as follows:

- 1M-phosphate buffer: 1M of  $\text{KH}_2\text{PO}_4$  (made up of 68g in 500ml) and 1M of  $\text{K}_2\text{HPO}_4$  (made up of 87.1g in 500ml) was mixed together, pH to 6.0 and autoclaved.

- 10X YNB (yeast extract nitrogen base): 67g of YNB was dissolved in 500ml distilled water, filter sterilized and stored at 4 °C in dark.
- 10X Glycerol: 50ml of glycerol was added to sufficient distilled water to a final volume of 500ml and autoclaved.
- 10X Methanol: 5ml of methanol was added to sufficient distilled water to make a final volume of 100ml, filter sterilized and stored at 4 °C.
- 500X Biotin: 0.02g of biotin was dissolved in 100ml of distilled water, filter sterilized and stored at 4 °C.
- Yeast Extract Peptone (YEP): 10g of yeast extract and 20g of peptone were dissolved in 700ml of distilled water and autoclaved.
- Buffered Minimal Glycerol/Methanol Complex (BMGY/BMMY): One litre of BMGY was formulated as follows: 700ml of YEP, 100ml of phosphate buffer, 100ml of 10X glycerol, 100ml of YNB and 2ml of 500X biotin. In BMMY, glycerol is replaced with methanol at the same concentration, all other ingredients remained same.

### **2.2.2 Growth of *Pichia pastoris* yeast cells**

Yeast cells were grown in buffered glycerol complex medium (BMGY). Briefly, single colonies of yeast cells previously grown on a minimal methanol histidine (MMH) plate were picked. The cells were then inoculated in 15ml BMGY in 250ml flask and incubated at 30°C in a shaker at 200 rpm. After 48 h, this starter culture was transferred and pre-induced into 100ml BMGY in a 2-litre flask and incubated further as before for another 48hrs. The cells were then harvested by centrifugation at 4000rpm at 4°C for 10 min. The cell pellets were collected and washed with distilled water before induction.

### **2.2.3 Induction and harvesting**

Induction was done in buffered minimal methanol media (BMMY). Washed cell pellets were suspended in 250ml BMMY and incubated at 30°C in a shaker incubator at 200 rpm. The culture was induced every 24 h for 96 h by adding 10X methanol to a final concentration of 1%. After 96 h of induction, the culture was centrifuged at 4000 rpm for 15 min at 4°C. The supernatant containing the protein was collected and filtered by passing it through a 0.22-micron filter. It was then analysed for rVCP 2,3,4 by SDS-PAGE gel and stored at 4°C.

### **2.2.4 SDS-PAGE analysis**

The presence of rVCP 2,3,4 was confirmed by 10% SDS-PAGE gel analysis. 15µl of the filtered supernatant was added to 5µl loading buffer and incubated for 10 min at 70°C. The sample was analyzed in 10% polyacrylamide gel electrophorised at 120V for about 1 h (until the moving front was at the base of the gel). The gel was silver stained and visualized.

### **2.3 Concentration and Purification of rVCP 2,3,4**

Supernatants from positive colonies were pooled together and concentrated before being purified. Concentration was achieved by Millipore Centricon centrifugal filters with 5 kDa cutoff; spinning at 2500 rpm. The concentrated supernatant was then purified using 1ml HiTrap heparin column (Pharmacia).

Briefly, the column was washed with 15ml of distilled water. The concentrated rVCP 2,3,4 supernatant was manually injected down the column at a rate of 1ml per minute. The column was then washed with 10ml of distilled water followed by 5 ml of 10mM Tris, pH 7.4. The column was then eluted at different NaCl concentrations ranging from 0.3M to 1.0M in order to establish the protocol for rVCP 2,3,4 elution in NaCl. Each NaCl portion contained 10mM Tris, pH 7.4. The column was then washed with distilled water and stored in 20% ethanol at 4°C. The different eluents were analyzed for the presence of rVCP 2,3,4 by 10% SDS-PAGE gel as described above. Fractions containing highly purified VCP were pooled together, dialyzed and concentrated.

### **2.4 Dialysis & superconcentration of rVCP 2,3,4**

To remove as much salt as possible from the purified rVCP 2,3,4, the sample was dialyzed against 10mM Tris at pH 7.4. Dialysis was done by a 4 kDa cutoff dialysis tube for 4 h at 4°C. After dialysis 10% SDS-PAGE gel was run on the sample to ascertain whether the protein was still in solution after dialysis.

Dialyzed rVCP 2,3,4 samples were then concentrated, using Millipore Centricon centrifugal filters with 5kDa cutoff as described before. The concentrated rVCP 2,3,4 was filtered to remove any precipitated solids prior to crystallization. The concentration of the dialysed, filtered, concentrated and purified rVCP 2,3,4 was estimated by its spectrophotometry absorbance at 280 nm, using NanoDrop ND-1000 V3.1.0 Spectrophotometer (Nanodrop Technologies Inc.) prior to setting up crystallization trials. The concentration was estimated to be 6.1 mg/ml. This solution was used to set up all crystallization trials.

## **2.5 Preparation of GDH for crystallization**

Recombinant GDH saturated in 70% ammonium sulfate was obtained from the laboratory of Professor Trevor Sewell of the electron microscopy unit of the University of Cape Town. Reconstitution of the rGDH was done as follows:

1 ml aliquot of rGDH was spun at 5000 x g, for 30 min at 4° C. The pellets were redissolved in 5ml of 10 mM Tris, pH 8.0 and dialysed against 1.0M Tris pH 8.0 by spin column desalting using a 20 ml Millipore Centricon centrifugal filters with 20 kDa cutoff. The desalted rGDH was concentrated using a 2 ml Millipore Centricon centrifugal filter. The concentration of protein was estimated by ND-1000 V3.1.0 Spectrophotometer (Nanodrop Technologies).

## **2.6 Crystallization of rVCP 2,3,4 and rGDH**

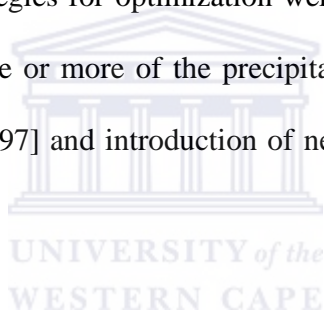
Initial crystallization trials both rVCP 2,3,4 and rGDH were set up with Hampton Research Crystal Screen 1 and 2 sparse matrix screens, using Linbro plates (Hampton). Hanging drops diffusion vapour method [Webber, 1991] was used for all the initial trials. The reservoir volume contained 1ml of the precipitating agents. The drop size was 2µl (1µl of protein sample and 1µl

of reservoir solution) for all drops. The plates were incubated at 18 °C on shelves free from vibration and other physical disturbances. The drops were observed immediately after set up and monitored daily for one week and then once every week thereafter using a LEICA light microscope. Photographs were taken and recorded by LEICA DFC320 camera.

## **2.7 Optimisation of rVCP 2,3,4 and GDH trials**

### **2.7.1 Introduction**

Following initial observations, trends were observed and promising drops were optimized. Depending on what class of precipitation was observed (i.e., precipitates, phase separation, micro- crystals etc), different strategies for optimization were used. These included an increase or decrease in concentration of one or more of the precipitation agents, variation of pH of the buffer, the use of oils [Chayen, 1997] and introduction of new precipitants or a combination of any of the above.



### **2.7.2 The use of oil**

Both paraffin (Hampton Research HR3-411) and silicon (Hampton Research HR3-415) oils were used, either alone or in several different combinations of 1:1, 1:2, 1:3, 1:4 or 2:3 (silicon oil: paraffin oil). Several different volumes of oils were also explored, ranging from 1ml to 200  $\mu$ l. The oils were used in order to bring about a gradual attainment of supersaturation.

Briefly, 1ml of the precipitating solution was discharged into the crystallization well. 1  $\mu$ l of this precipitating solution was taken and mixed with 1  $\mu$ l of the protein solution on a cover slide. As

quickly as practicable, the required volume of oil was dispensed onto the surface of the reservoir. The cover slide containing the crystallization drop was then quickly inverted over the plate and sealed.

### **2.7.3 Change in concentration of precipitating solution**

Precipitating solutions were either increased or decreased depending on what the initial observations were. Two methods were used to decrease the concentration of the precipitating solutions; either preparing a new reservoir solution of lower concentration or 'in situ' dilution of the existing reservoir solution. The latter procedure was used to ascertain the time nucleation is most likely to have occurred. It was used for situations in which uncontrolled nucleation had resulted in the formation of showers of crystals within a relatively short time.

In the '*in situ*' dilution method, the time of appearance of the first crystals was noted. More drops were set up under the same conditions as those that gave the showers of crystals. Within specific time intervals (depending on the time of first appearance of crystals), the cover slide of a drop was removed and specific amount of distilled water was added and mixed with the reservoir solution and the drop quickly re-sealed. The drop was then incubated under the same previous conditions and observed later.

## **2.8 Data Collection and processing**

X-ray diffraction data was collected from one rVCP 2,3,4 crystal and 2 rGDH crystals grown from different crystallization conditions. Data was collected at 100K from single crystals; using an in house X-ray diffractometer housed in the Biotechnology Department of the University of

the Western Cape. The X-ray source was Rigaku RUH3R copper rotating anode, producing X-rays at 22 mA and 40 kV. The images were recorded onto Rigaku R-axis IV plate camera. Cryostream was achieved by an X-stream 2000 cryo system. The crystal-detector distance was 100 mm for all the data collected.

rVCP 2,3,4 crystal was grown from 30% isopropanol, 0.1M Na HEPES pH 7.5 and 0.2M MgCl<sub>2</sub>. VCP crystal was briefly soaked in a solution containing the same reservoir conditions, plus 20% glycerol; acting as a cryoprotectant. The crystal was then flash-frozen in a cryostream prior to data collection. rVCP 2,3,4 data was collected to 2.5Å. A total of 720 images were collected at 0.5 degree oscillation and with 15 min exposure to cover an entire 360 degrees rotation. Attempts were made to reproduce this crystal, using rVCP 2,3,4 produced and purified from another batch of expression.

GDH data was collected on 2 crystals grown from different crystallization conditions. The first crystal was grown in 30% PEG 8000, 0.1M Tris pH 8.7, 0.2M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The drop was diluted by 50% after 6 hours of set up. The second crystal was grown from 1M Li<sub>2</sub>SO<sub>4</sub>, 0.1M Na Citrate pH 5.6, 0.5M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, with 800µl of 3:1 (Paraffin: Silicon) oil.

The first GDH crystal diffracted to a nominal resolution of 6.5Å while the second crystal diffracted to 8.5Å. The first crystal did not require any cryoprotectant solution but the second crystal was first briefly soaked in a cryoprotectant solution containing 25% glycerol. Both crystals were flash-frozen prior to data collection. In both GDH crystals, data were collected with 1° angle of oscillation, starting from 1°; with 20min exposure. 120 and 40 images were collected from the first and second crystals respectively.



Both the rVCP 2,3,4 and rGDHB data were processed using Denzo/Scalepack software [Otwinowski & Minor, 1997; Otwinowski, 1993]. The diffraction intensities were integrated and a program called TRUNCATE in the CCP4 suite [CCP4, 1994] was used to reduce the intensities into structure factor amplitudes.

## 2.9 Molecular replacement and refinement

Molecular replacement was done on the rVCP 2,3,4 data by four different methods in order to have a basis for comparison of results. The programs used were MOLREP within the CCP4 [Vagin & Teplyakov, 1997; CCP4, 1994], PHASER [Storoni *et al*, 2004; Read, 2001], EPMR2.5 [Kissinger *et al*, 1999] and CASPR, an online web server software [Jean-Baptiste *et al*, 2004]. Since the inherent intermodular flexibility and absence of one or more modules can have an effect on the overall structure of CCP modules [Kirkitaadze *et al*, 1999(a, b, c & d); Henderson *et al*, 2001], different individual VCP modules (and combination of modules) of VCP were used as rigid body probes. This included VCP 2, VCP 3, VCP 4, VCP 2,3; VCP 3,4; VCP 2,3,4 and the full length VCP (VCP 1,2,3,4). To validate the results of the molecular replacement process, a random protein (with similar size to one module of VCP) was selected from the PDB and used as a probe. The statistics obtained by this random probe would be compared to that obtained by the VCP probes in order to find out if the solutions obtained by the VCP probes are true or random solutions.

Next, some modules were fixed and searched against other modules; acting as probes as follows:

- (a) VCP 2 fixed against VCP 3
- (b) VCP 2 fixed against VCP 4

- (c) VCP 3 fixed against VCP 4
- (d) VCP 2 fixed against VCP 3 and then results fixed against VCP 4
- (e) VCP 2 fixed against VCP 4 and then results fixed against VCP 3
- (f) VCP 3 fixed against VCP 4 and then results fixed against VCP 2
- (g) VCP3 fixed against VCP 2 and then results fixed against Random probe

### **2.10 Post crystallization SDS-PAGE gel to determine the presence of impurities**

Following the failure of molecular replacement to produce a correct solution, it was assumed that there might be some impurity protein(s) in the rVCP 2,3,4 superconcentrated solution. 10% SDS-PAGE gel was run to investigate this aspect. Since none of the concentrated rVCP 2,3,4 solution was left unused at this stage, samples were collected from as many as 50 clear crystallization drops. This contained equal (or nearly so) amounts of protein solution and precipitating solution but the presence of the precipitating solution is not expected to affect the visualization of the protein on the gel. 40µl of this solution was loaded onto a 10% SDS-PAGE gel. The gel was stained and visualized by Commassie blue.



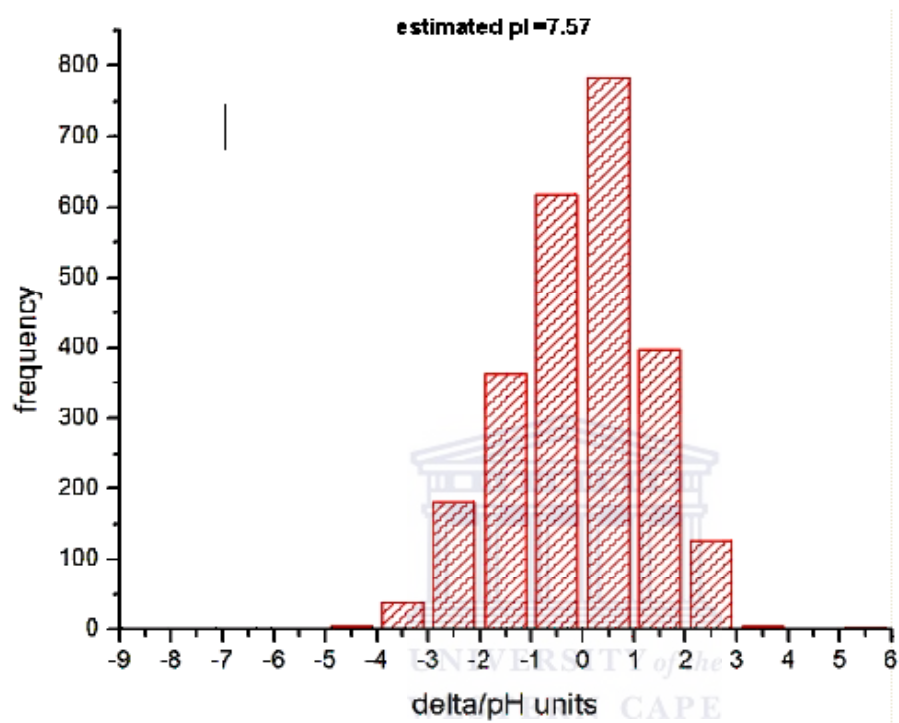
UNIVERSITY *of the*  
WESTERN CAPE

## Chapter 3: Results and discussion

### 3.1 Prediction of likely crystallization pH

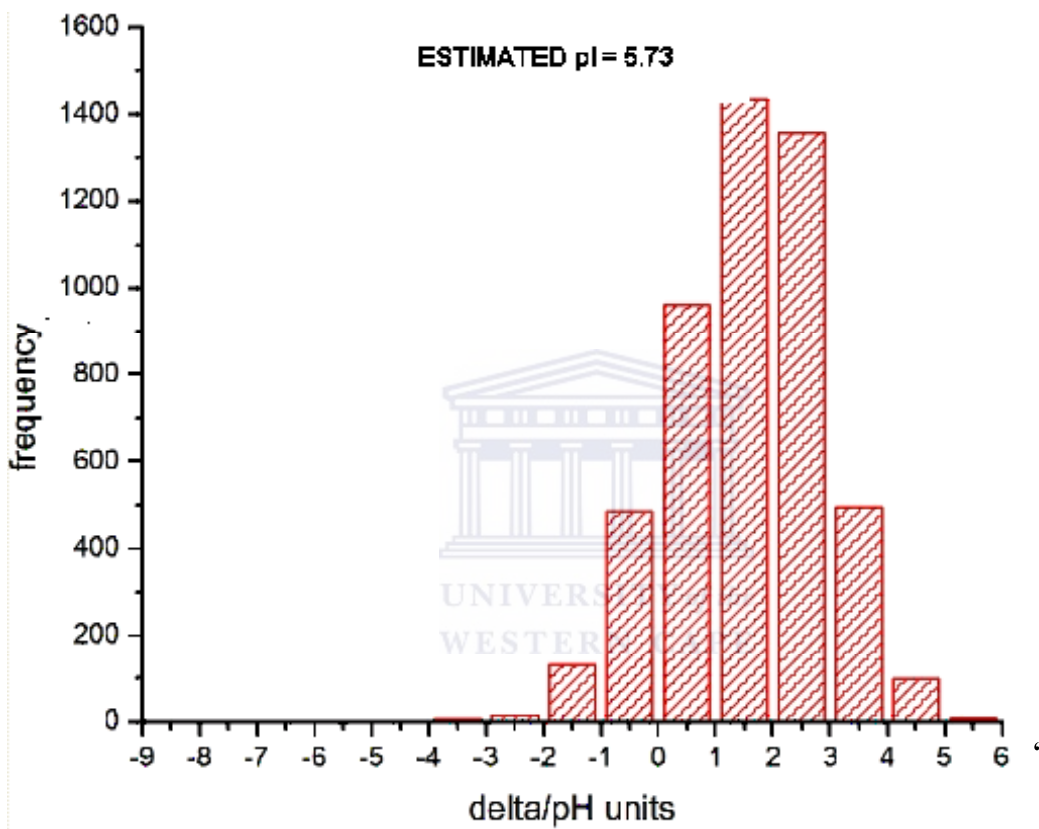
The graphs below (fig. 3.1 & fig. 3.2) represent the correlation between the (pI-pH) and the predicted success rate (frequency) of rVCP 2,3,4 and GDH crystallization as predicted by ‘CrysPred’ [Kantardjieff & Rupp, 2004; Kantardjieff *et al*, 2004]. CrysPred is a program for predicting the potential pH at which crystallization is most likely to occur, given the isoelectric point (pI) of the protein.





14

**Fig. 3.1:** A graph of predicted crystallization success rate (frequency) versus the likely (pI-pH) range of rVCP 2,3,4 crystallization.



**Fig. 3.2:** A graph of predicted crystallization success rate (frequency) versus the likely (pIpH) range of GDH crystallization.

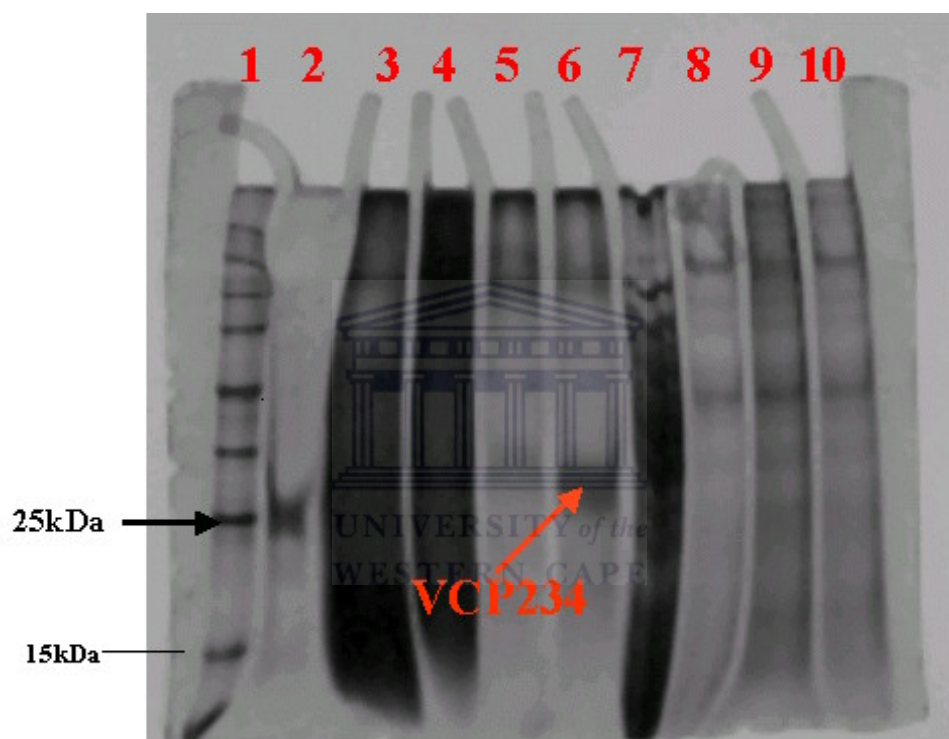
It is clear from fig. 3.1 that there is a greater likelihood of crystallization of rVCP 2,3,4 to occur within  $\pm 1$  pH units from its estimated pI. This is indicated by the two highest peaks, corresponding to the -1 (frequency of 650) and +1 (frequency of 750) values on the delta (pI-pH) axis. Given the calculated pI of rVCP 2,3,4 as 7.57, the predicted pH at which crystallization is most likely to occur is between pH 6.6-8.6. Outside these pH ranges, the predicted success rate for rVCP 2,3,4 crystallization is minimal.

Using the same analysis, it is clear from fig. 3.2 that GDH is most likely to crystallize within +2 to +3 pH units away from its pI. Given the estimated pI of GDH as 5.7, the most likely crystallization pH of GDH is predicted to be pH 6.5-8.5. Within these pH ranges, there is a very high-predicted success frequency of between 1400 and 1500, as seen from fig. 3.2.

These predictions would be of great help in designing optimisation of crystallization trials; particularly optimisation involving pH screening. Instead of screening random pH ranges, resulting in large usage of protein and crystallization materials, a narrow range of pH based on the CrysPred prediction would be screened first.

### **3.2 Expression of rVCP 2,3,4 in *Pichia pastoris***

The results of the rVCP 2,3,4 expression in *Pichia pastoris* are shown in the SDS-PAGE gel below.



**Fig. 3.3:** SDS-PAGE gel of the expression of rVCP 2,3,4. Lane 1 = molecular weight marker. Lane 2 = full length VCP. Lane 3-10 = supernatant from 8 single colonies used in expression

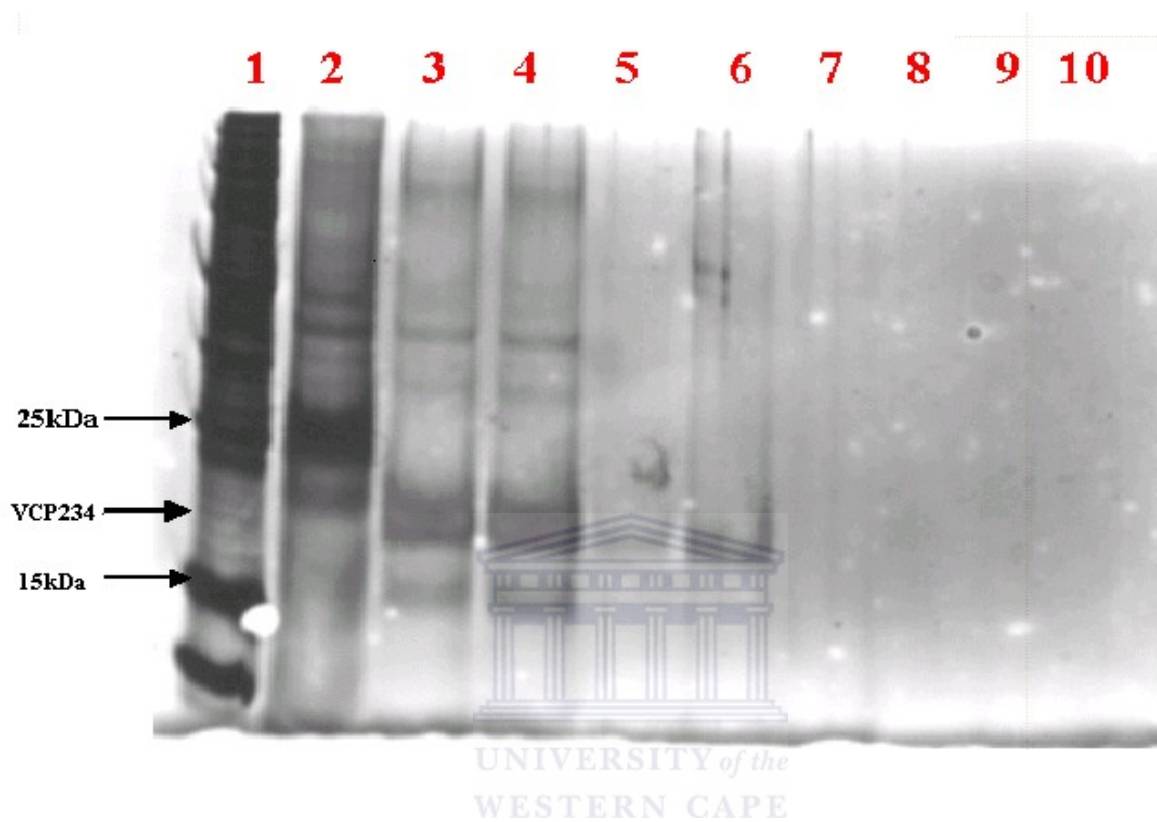


From fig. 3.3 above, it can be shown that there was some positive expression in some of the yeast single colonies. Colonies in lanes 5, 6 and 9 produced detectable amounts of rVCP 2,3,4; with the highest expression occurring in a colony in lane 6. However, the band of rVCP 2,3,4 on the gel was not at the expected position. With its molecular weight of about 19.5kDa, rVCP 2,3,4 was expected to be located below the full length VCP (26 kDa) seen in lane 2. However, it was located around the 30 kDa mark on the gel. In spite of this, its identity as rVCP 2,3,4 was not doubted. This is because experience gathered from the laboratory where the research was carried out had shown that VCP band on SDS-PAGE gel could sometimes migrate to a position other than its expected molecular weight. Based on this assumption, supernatant from colonies in lanes 5 and 6 were pooled together and used for further analysis.

There was no detectable expression in the colonies in lanes 3, 4, 8 and 7. Lane 10 had faint expression of rVCP 2,3,4 but also has many other prominent bands. The use of full length VCP in the SDS-PAGE gels of the rVCP 2,3,4 expression, purification and dialysis was to serve as an additional marker. While the full length VCP was pure (fig. 3.4, lane 2) the fractions used in the purification (fig. 3.4, lane 2) and dialysis (fig. 3.5, lane 2) were not as pure.

### **3.3 Purification of rVCP 2,3,4**

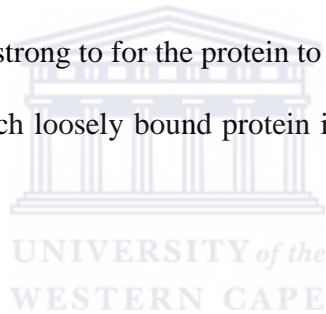
From fig. 3.4, it can clearly be seen that the pooled supernatant from fig. 3.3 indeed contained rVCP 2,3,4 (as seen in lane 3). Its position on the gel (between 25 kDa and 15 kDa) is a clear indication that the protein could be rVCP 2,3,4. The pooled rVCP 2,3,4 did not contain many impurity bands, as shown in fig. 3.4, lane 3. Although there were 3 visible bands on the gel, rVCP 2,3,4 band was far most intense than the 'contaminants' bands.



**Fig. 3.4:** SDS-PAGE gel for the purification of rVCP 2,3,4; using heparin column affinity. Lane 1= Molecular weight marker. Lane 2= Full length VCP. Lane 3 =Original (supernatant pooled from positive colonies). Lane 4 = Flow through. Lane 5 = First wash. Lane 6 = Elution with 400mM NaCl. Lane 7= Elution with 500mM. Lane 8 = Elution with 600mM NaCl. Lane 9 =Elution with 700mM NaCl. Lane 10 = Elution with 800mM NaCl.

It is also clear from fig. 3.4 that rVCP 2,3,4 elutes at 400 mM NaCl (lane 6). Beyond 400 mM NaCl elution, there was no presence of rVCP 2,3,4 at all. The flow through (fig. 3.4, lane 4) contained almost the same concentration of rVCP 2,3,4 as the original (lane 3). This was because the column was oversaturated since the original protein was highly concentrated. Given the fact that the 1ml Heparin column can bind a maximum of 1mg of rVCP 2,3,4, substantial amount of rVCP 2,3,4 was expected to be present on the flow through unbound.

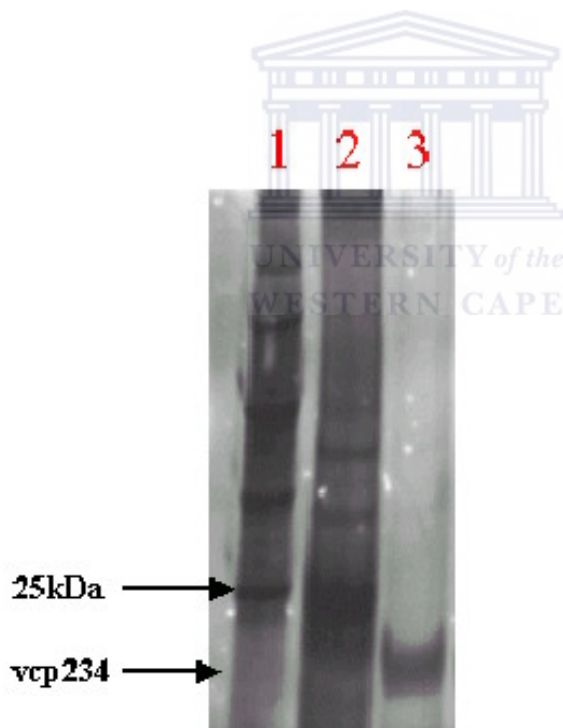
The first wash (lane 5) also contained purified rVCP 2,3,4. Presence of purified rVCP 2,3,4 in the wash could be due to the oversaturation of the column. With the heparin column being oversaturated with rVCP 2,3,4, it was not surprising that some rVCP 2,3,4 was loosely bound to the column. Such binding was too strong to for the protein to be in the flow through but too weak to be eluted in the salt elution. Such loosely bound protein is most likely be eluted in the wash fraction.



It can be concluded from fig. 3.4 that the purification process was successful; judging from the purified protein obtained in lane 5 and 6. The two most prominent ‘impurity’ bands in the original supernatant (lane 3 in fig. 3.4) (one above and one below the rVCP 2,3,4 band) are all removed in the purified portion. As a heparin binding protein [Smith *et al*, 2000], rVCP 2,3,4 binds to heparin, allowing all other proteins in the supernatant to pass through the heparin column unbound.

### 3.4 Dialysis of rVCP 2,3,4

The purified rVCP 2,3,4 was dialyzed against 10 mM Tris. The essence of dialysis was to remove salt from the protein solution. Although there was no means to estimate the final salt concentration in the protein solution after dialysis, it can be seen from fig. 3.5 that the protein was not lost during dialysis. Four hours of dialysis was expected to remove as much salt from the protein solution as possible. The aim of the dialysis step, though, was not to remove all traces of salt from the protein solution. In fact, presence of some amount of salt might as well enhance the crystal formation of rVCP 2,3,4, judging from the fact that the full length VCP was crystallized in the presence of 100mM NaCl [Murthy *et al*, 2001, Ganesh *et al*, 2004].



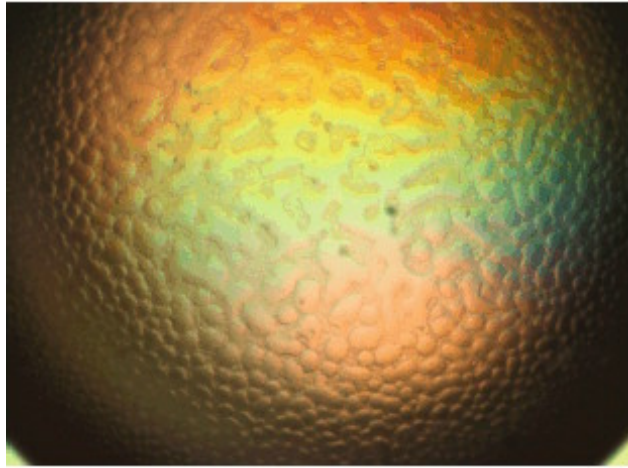
**Fig. 3.5:** SDS-PAGE gel of dialysis of VCP234. Lane 1= Molecular weight marker. Lane 2 = full length VCP. Lane 3 = Purified VCP234 after dialysis.

### 3.5 Results and observed trend from the initial crystallization screening of rVCP 2,3,4

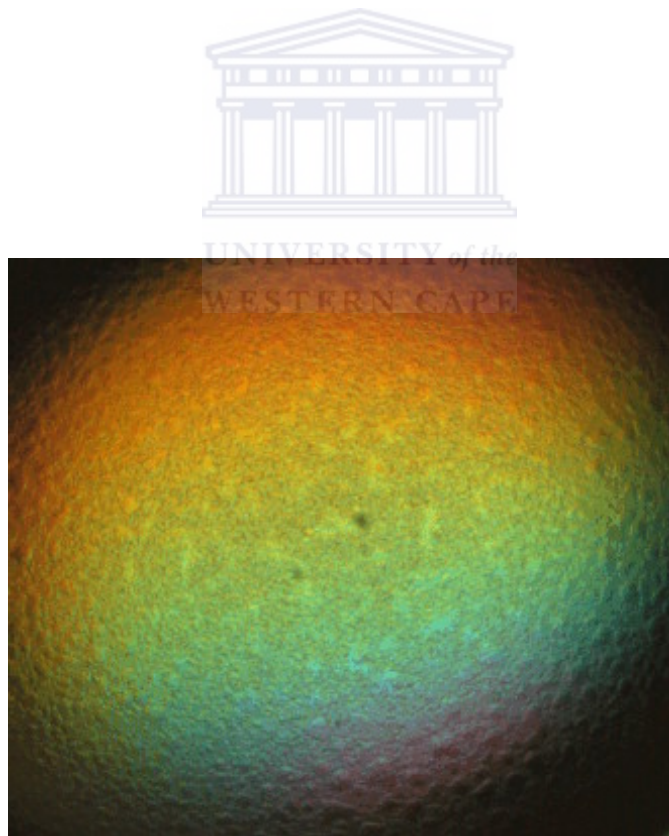
Initial screening was done using Hampton screen 1 and 2 hanging drop vapour diffusion method [McPherson, 1982]. After three weeks of setting up the trials, observations were tabulated as shown in table 3.1 and pictures of some of the drops shown in fig. 3.6 to fig. 3. 10.

DROP IDENTITY	CONDITIONS	OBSERVATION
Screen 1, 4	2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.1M Tris pH 8.5, 0.2M Na citrate	Crystalline precipitate
Screen 1, 14	28% PEG 400, 0.1M Na Hepes pH 7.5, 0.2M CaCl <sub>2</sub>	precipitate
Screen 1, 24	20% Isopropanol, 0.1M Na acetate pH 4.6, 0.2M CaCl <sub>2</sub>	Crystalline precipitate
Screen 1, 39	20% PEG 400, 0.1M Na Hepes pH 7.5, 2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Phase separation
Screen 1, 45	18% PEG 8000, 0.1M Na Cacodylate pH 6.5, 0.2M Zn acetate	Precipitate
Screen 1, 19	30% Isopropanol, 0.1M Tris pH 8.5, 0.2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Precipitate
Screen 1, 12	30% Isopropanol, 0.1M Na Hepes pH 7.5, 0.2M MgCl <sub>2</sub>	Micro crystals
Screen 1, 23	30% PEG 400, 0.1M Na Hepes pH 7.5, 0.2M MgCl <sub>2</sub>	Micro crystals
Screen 1, 22	30% PEG 4000, 0.1M Tris pH 8.5, 0.2M Na acetate	Precipitate
Screen 1, 6	30% PEG 4000, 0.1M Tris pH 8.5, 0.2M MgCl <sub>2</sub>	Precipitate
Screen 1, 30	30% PEG 8000, 0.2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Phase separation
Screen 1, 31	30% PEG 4000, 0.2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Phase separation
Screen 1, 8	30% Isopropanol, 0.1M Na Cacodylate pH 6.5, 0.2M Na citrate	Micro crystals
Screen 1, 27	20% Isopropanol, 0.1M Na Hepes pH 7.5, 0.2M Na citrate	Micro crystals
Screen 2, 25	1.8M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 0.1M MES pH 6.5, 0.01M CoCl <sub>2</sub>	Crystals
Screen 2, 30	30% Mp, 0.1M Na Hepes, pH 7.5, 0.5M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Phase separation
Screen 2, 28	20% PEG 10000, 0.1M Na Hepes, pH 7.5,	Precipitate
Screen 2, 39	3.4M 1,6 Hexandiol, 0.1M Tris pH 8.5, 0.2M MgCl <sub>2</sub>	Crystalline precipitate
Screen 2, 47	2M MgCl <sub>2</sub> , 0.1M Bicine pH 9.0	Precipitate

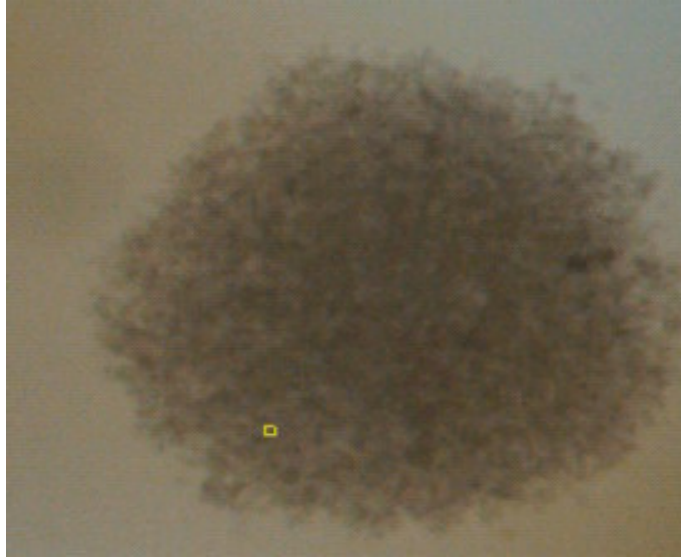
**Table 3.1:** Results of initial rVCP 2,3,4 crystallization trials from Hampton screen 1 and 2. Screen 1 is made up of 50 different formulated conditions (1-50) while screen 2 is made up of 48 conditions (1-48).



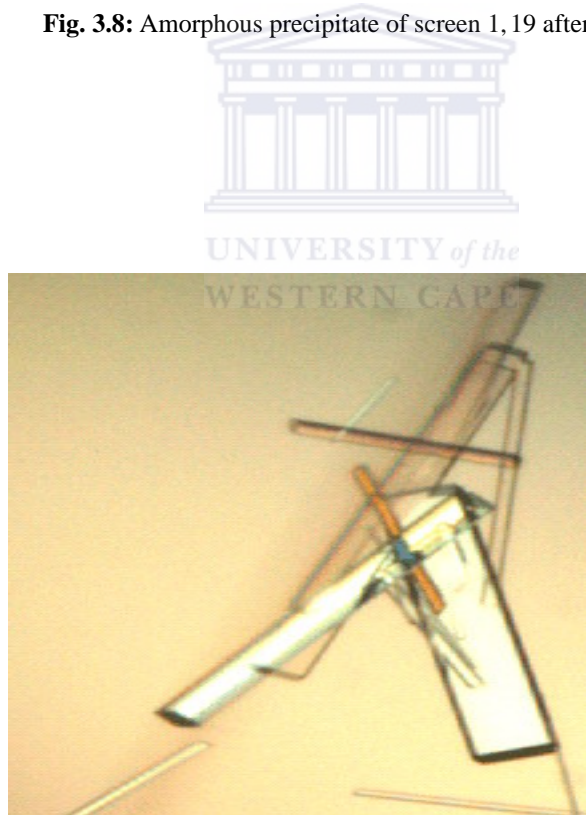
**Fig. 3.6:** Phase separation of screen 1, 39, after 3 weeks.



**Fig. 3.7:** Crystalline precipitate of screen 1, 24 after 3 weeks

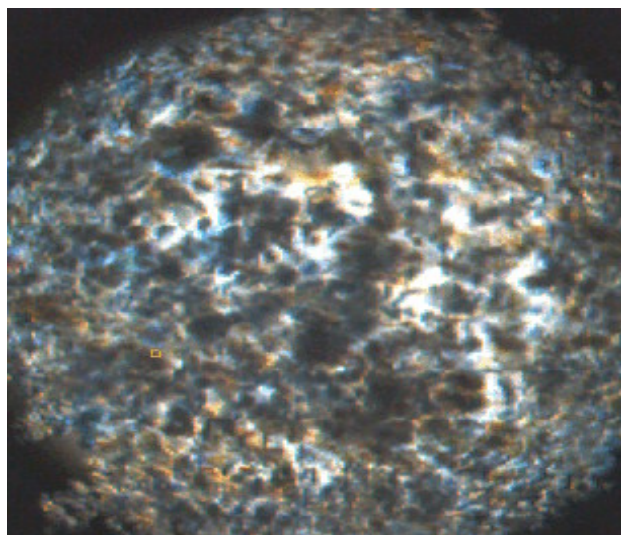


**Fig. 3.8:** Amorphous precipitate of screen 1, 19 after 3 weeks



**Fig.3.9:** Salt crystal of screen 2, 25 after 3 weeks.





**Fig. 3.10:** Micro crystals from screen 1, 27 after 3 weeks.

Although several different conditions of the initial screening resulted in some kind of precipitates, phase separation, micro crystals or salt crystals, a general trend was observed. Certain precipitants (within certain concentration ranges) were observed to have high frequency of producing some sort of precipitates, phase separation or micro-crystals. Optimisation of the initial trials took advantage of this trend.

<b>Potential precipitants</b>	<b>Concentration range</b>
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	0.2M-2.0M
PEG 400	20% -30%
Isopropanol	20% -30%
PEG8000	18% -20%
MgCl <sub>2</sub>	0.2M-2.0M
PEG 4000	30%
CaCl <sub>2</sub>	0.2M
Na citrate	0.2M

**Table 3.2:** A table showing the precipitants (and the concentration ranges) that produced a high frequency of observable results of the initial crystallization screening of rVCP 2,3,4.



### **3.6 Optimization of the initial screening that produced precipitates**

Drops that produced amorphous or crystalline precipitates were optimized by decreasing the concentration of one or more of the precipitants, varying pH or the use of oils; as described before. Combinations of the above techniques were also employed.

Precipitates are formed in the supersaturation region of the phase diagram (fig. 1.8). Precipitates are formed when saturation is approached too rapidly. This occurs when the attraction between the protein molecules is so strong that there is not sufficient time for the protein molecules to align or orientate themselves into well, ordered crystals. Precipitate formation is normally favoured when the precipitants or protein concentration is too high. Therefore, reducing the concentration of the precipitants and/or protein would provide a means for attainment of supersaturation more slowly; possibly leading to well ordered aggregation of protein molecules and resulting in crystal formation. Since the environment of a protein also influence its solubility, varying pH may also lead to a slower attainment of supersaturation, leading to crystal formation.

Two types of precipitates can be formed; amorphous or crystalline precipitates. Under the microscope, crystalline precipitates flicker (fig. 3.7) if the incident angle is slightly changed [Sica, 1996], while amorphous precipitates (fig. 3.8) do not. Observation of crystalline precipitates in a crystallization drop could be a step closer to the formation of crystal, although individual crystals are not seen in crystalline precipitates.

### 3.7 Optimization of the initial screening that produced phase separation

Drops producing phase separation were optimized by varying the pH of the buffer or by varying the concentration of one or more of the precipitating agents. Phase separation usually appears as numerous small, often clear droplets (fig. 3.6). Like precipitates, phase separation occurs at the supersaturation region of the phase diagram.

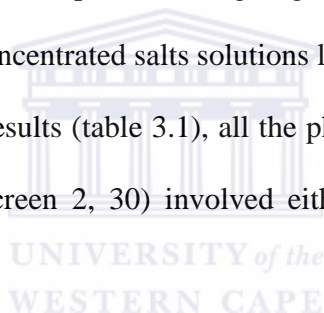
Phase separation is caused by the separation of two or more of the precipitating agents present in the mother liquor as a result of immiscibility with each other [Sica, 1996; McPherson, 1999]. As partition occurs between two components of the crystallization solution, small droplets are formed. Depending on how soluble the protein is in each of the two solvent components, the protein may sometimes become more concentrated in one phase than the other, resulting in very high supersaturation state in that component [Sica 1996, McPherson 1999; Ray & Bracker, 1996]. Since supersaturation is a prerequisite for the nucleation of crystals, crystals can sometime form in the protein-rich phase [Kuznestov *et al*, 2001] in phase separation drops. The exact mechanism by which phase separation liquid drops enhance protein crystallization is not fully understood [Asherie, 2004]. But it is generally believed that it is due to the higher supersaturation state reached by protein in the protein-rich phase [Asherie, 2004], or either due the fact that the surface of the crystal is wet by one of the liquid phase [ten Wolde & Frenkel, 1997].

Although phase separation can sometimes result in crystal production and therefore could be a good sign, crystals formed in phase separation drops can be very problematic during mounting. This is because the concentration in the protein-rich phase is usually unknown and very different

from its concentration in the mother liquor. The concentration of the crystallization components surrounding the crystal is usually higher than the concentration in the starting mother liquor. This presents a problem during the preparation of cryoprotectant solution.

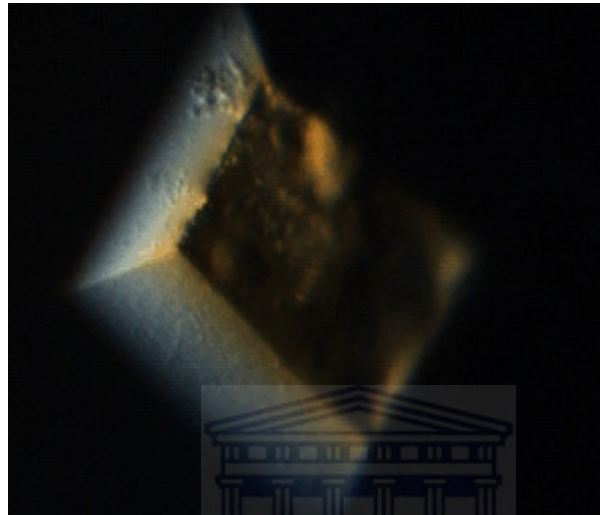
Certain steps, however, can be taken to correct the mounting problem of phase separation crystal. These include freezing the crystal directly in the drop [Xtal protocols: <http://www.xtal-protocols.de/drop/score5.html>]. Another means to go over the mounting problem is to add some protein solution in the mounting solution (cryoprotectant solution).

Phase separation usually occurs in drops containing organic solvents (such as PEG, MPD, ethanol and dioxane) and highly concentrated salts solutions like  $(\text{NH}_4)_2\text{SO}_4$  [Ray & Bracker, 1996]. From the above results (table 3.1), all the phase separation drops (screen 1, 39; screen 1, 30; screen 1, 31 and screen 2, 30) involved either PEG or MPD with  $(\text{NH}_4)_2\text{SO}_4$  consistent with expectation.



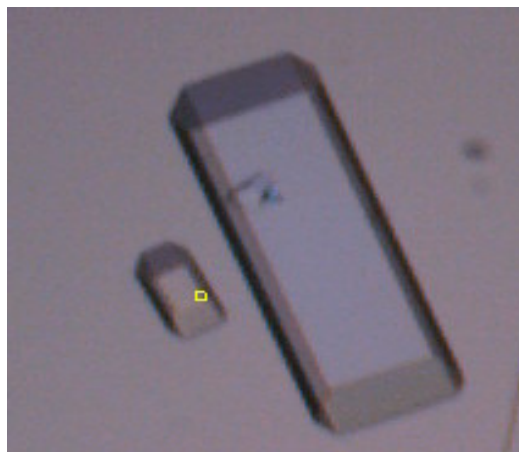
### 3.8 Results from the optimisation process of rVCP 2,3,4

The following are some of the results of the optimisation processes:

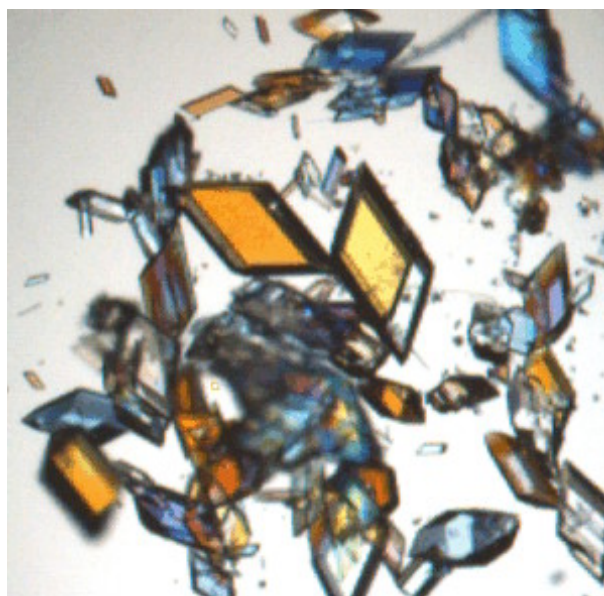


**Fig 3.11:** Salt crystal obtained from 25% Isopropanol, 0.1M Tris pH 8.5, 0.5M  $(\text{NH}_4)_2\text{SO}_4$

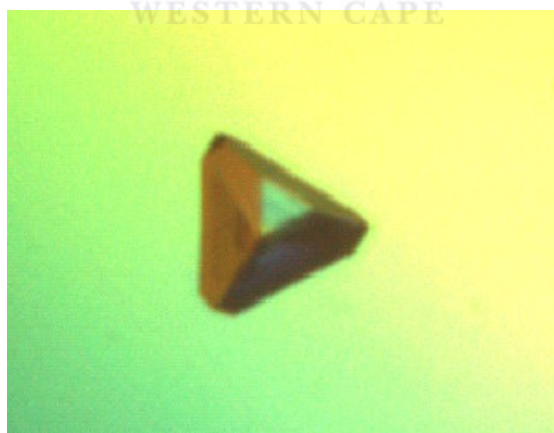
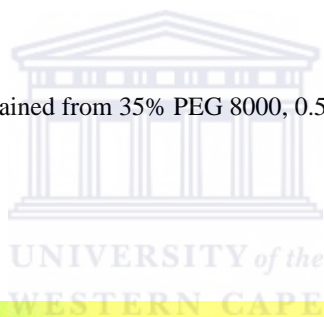
UNIVERSITY of the  
WESTERN CAPE



**Fig 3.12:** Salt crystal obtained from 28% PEGS 400, 0.1M Na Hepes pH 7.5, 0.1M  $\text{CaCl}_2$  with 700 $\mu\text{l}$  of 2:1 (paraffin: silicone) oil



**Fig. 3.13:** Salt crystals obtained from 35% PEG 8000, 0.5M(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>.



**Fig 3.14:** True protein crystal obtained from 35% Isopropanol, 0.1M NaHepes pH 7.5, 0.3M MgCl<sub>2</sub> at protein concentration of 6.1mg/ml. The crystal was produced after 27 days of set up and diffracted to 2.5Å.

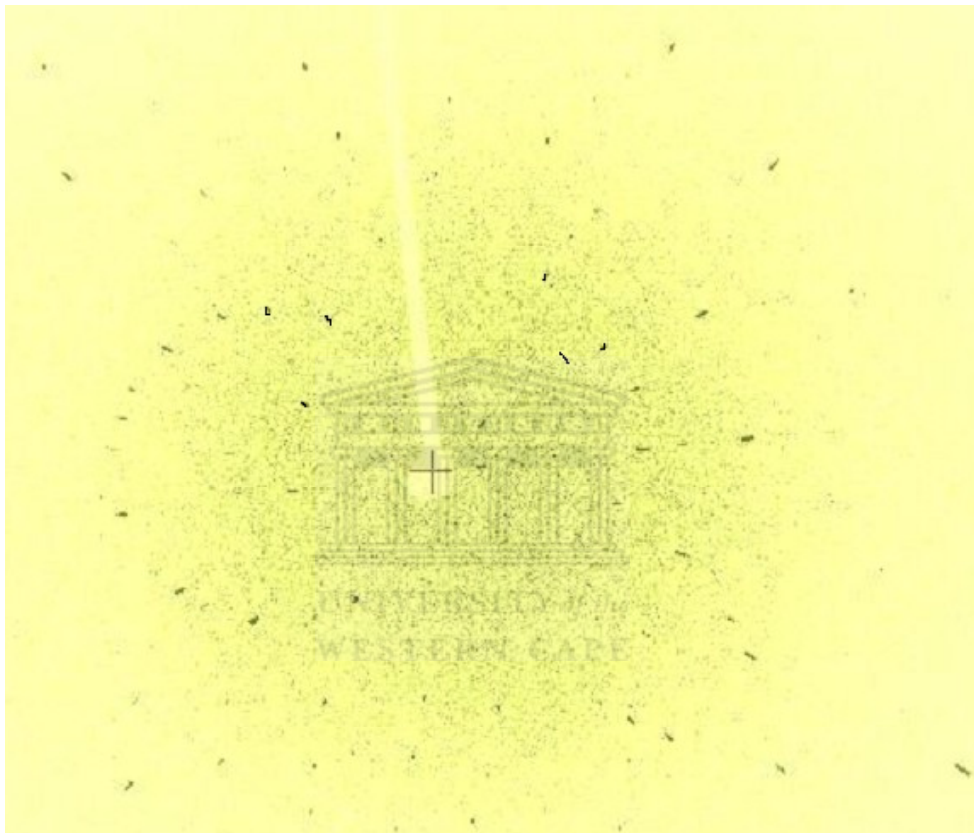
Most of the drops in the optimisation processes either resulted in salt crystals (figs. 3.11, 3.12 and 3.13) or no crystal at all. Formation of salt crystals is a common phenomenon in protein crystallization. Salts form crystals more easily than proteins do, and like protein crystals, salt crystals come in different forms and shape. The formation of salt crystals can arise from a number of factors.

Firstly, there is the possibility of a high NaCl concentration present in the protein solution after dialysis. If this were so, it would mean that dialysis failed to remove substantial amount of salt from the protein solution. Another possible source of salt could be from the precipitating (crystallization) solution itself. Salts like  $(\text{NH}_4)_2\text{SO}_4$ ,  $\text{CaCl}_2$  and  $\text{CoCl}_2$  could get crystallized even at reasonably low concentrations. As can be seen from figs 3.11, 3.12 and 3.13 and table 3.1, all the drops which resulted in salt crystals contained one of  $\text{CoCl}_2$ ,  $\text{CaCl}_2$  or  $(\text{NH}_4)_2\text{SO}_4$ .

Usually, it is pretty difficult to distinguish between the salt and protein crystals under the microscope. Salt crystals were positively confirmed by their diffraction pattern or by poking them with a fine needle. Protein crystals are more fragile than salt crystals and they tend to break apart on the slightest application of pressure whiles salt crystal need a relatively high pressure application to break it apart when poked.

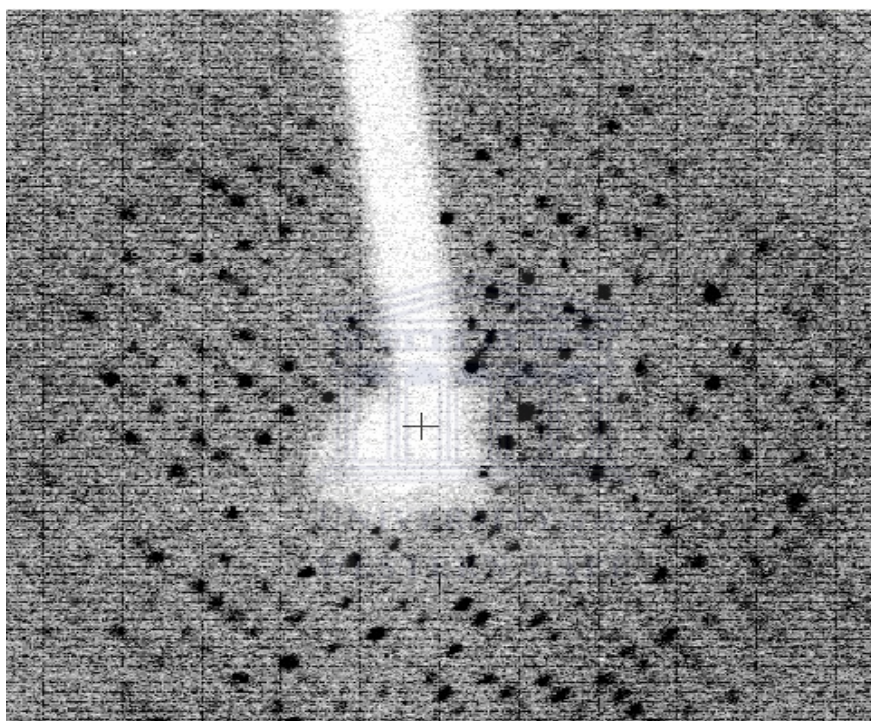
The diffraction pattern of salt crystals is very different from that of a protein crystal. Small molecules (like salts) have smaller Bragg distances compared to protein molecules. Therefore, salt crystals tend to give fewer reflections with widely separated spots ((fig. 3.15)). With large

Bragg distances, protein crystals produce more reflections (at the same crystal to detector distance) and the spots are closer to each other (fig. 3.16) than that of salt crystals.



**Fig 3.15:** Diffraction pattern of a salt crystal. The crystals was produced in 25% Isopropanol, 0.1M Tris pH 8.5, 0.5M  $(\text{NH}_4)_2\text{SO}_4$  in the rVCP 2,3,4 solution



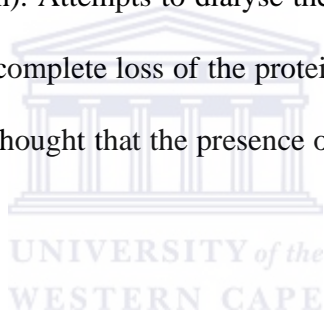


**Fig. 3.16:** Diffraction pattern of protein crystal (GDH). Crystal grown from 30% PEG 8000, 0.1M Tris pH 8.7, 0.2M  $(\text{NH}_4)_2\text{SO}_4$ .



Out of more than 750 drops in total, only one drop produced a protein crystal (fig. 3.14). Only one small but well formed crystal was produced in the drop. With only one crystal available, extreme caution was exercised not to lose it during the mounting process onto the X-ray beams.

Attempts made to reproduce the crystal were unsuccessful. Since the rVCP 2,3,4 solution from which the crystal was produced was entirely used up by the time the crystal was produced, a fresh batch of rVCP 2,3,4 solution was used in an attempt to reproduce the crystal. The two batches differed slightly in salt conditions. While the first batch was dialysed against 10mM Tris, the second batch was left in 250mM NaCl (the same salt concentration at which the protein was eluted from the purification column). Attempts to dialyse the second (new) batch of rVCP 2,3,4 against 10mM Tris resulted in the complete loss of the protein (the protein probably precipitated out of solution). It was, therefore, thought that the presence of the salt might be essential to keep the protein in solution.



It is not very clear why the first batch of rVCP 2,3,4 did well after dialysis with 10mM Tris while the second batch could stay in solution only in the presence of 250mM salt. One possible explanation could be that not all salt was dialysed out in the first batch. Possibly, the salt concentration was reduced from 400mM (in the elution) to about 250mM after dialysis; meaning that the solution that produced the crystal might after all contain some reasonable amount of salt (probably in the range of 250mM).

The second batch of rVCP 2,3,4 was eluted in 250mM salt, instead of 400mM in batch 1, for some unknown reasons. Attempt to dialyse the salt below the 250mM resulted in the loss of the

protein; probably by means of precipitation. The above observation could be explained on the assumption that rVCP 2,3,4 is stable only in the presence of some amount of salt (in the range of 250mM).

Crystallization trials, were therefore, set up on the new batch rVCP 2,3,4 solution containing 250mM NaCl. Several drops were set up around the conditions that originally produced a crystal in the first batch rVCP 2,3,4 solution. In spite of this, no crystal was formed in the new batch rVCP 2,3,4 after several weeks.

Failure to reproduce the crystal may be attributed to one or combination of several factors. Firstly, the presence of 250mM NaCl in second batch rVCP 2,3,4 might induce unfavourable conditions for crystallization to occur. This explanation works on the assumption that there was no salt in the first batch rVCP 2,3,4. Presence of salt in the second batch rVCP 2,3,4 solution might alter the ionic environment of the protein, making it very difficult to crystallize under similar conditions which originally produced a crystal.

Another explanation could be that the protein which was crystallized in the first batch may be absent in the solution of the second batch. If this were so, it might indicate that the crystallized protein was not a rVCP 2,3,4 protein but rather an 'impurity' protein. This 'impurity' protein might be absent in the second batch rVCP 2,3,4 solution; thus resulting in the irreproducibility of the crystal.

### 3.9 Processing and phase determination of rVCP 2,3,4 data

#### 3.9.1 Processing of rVCP 2,3,4 data

The statistics obtained from the processing of the rVCP 2,3,4 data are presented in table 3.3 below. The protein crystallized in the space group F23. This was different from the previously solved crystal structure of the native full length VCP or the VCP complex with heparin [Murthy *et al*, 2001, Ganesh *et al*, 2004]. In the full length VCP, two crystal forms were obtained; which crystallized in the space group P212121 and C222 [Murthy *et al*, 2001]. In the heparin bound complex, the VCP crystallized in the space group P2 [Ganesh *et al*, 2004]. The mosaicity was 0.38 and 93.9% completeness and 16.07 average redundancy (table 3.3).

---

Space group	F23		
Unit cell dimensions	123.99	123.99	123.99
	90.00	90.00	90.00
Resolution range	71.58	- 1.69	(1.75 - 1.69)
Total number of reflections	270193		
Number of unique reflections	16809		
Average redundancy	16.07		(2.25)
% Completeness	93.9		(44.7)
Mosaicity	0.38		
Rmerge	0.418		(0.929)
Reduced ChiSquared	0.05		(0.05)
Output <I/sigI>	1.2		(0.0)

---

**Table 3.3:** Summary of data collection statistics of rVCP 2,3,4 data. Note: Values in () are for the last resolution shell.

### 3.9.2 Phase determination using molecular replacement

Since crystal structures of VCP exist in the protein data bank, the obvious choice of phasing technique was molecular replacement. Results obtained in the molecular replacement of rVCP 2,3,4 are presented in tables 3.4 and 3.5 below.

PROBE	MOLREP		EPMR v 2.5		CASPR		PHASER
	Corr Coefficient / Rfactor		Corr Coefficient / Rfactor		Corr Coefficient / Rfactor		Z-score
VCP1234	0.09	/ 0.63	0.12	/ 0.53	0.13	/ 0.69	4.7
VCP234	0.11	/ 0.62	0.14	/ 0.60	0.11	/ 0.62	5.0
VCP23	0.12	/ 0.62	0.12	/ 0.61	0.13	/ 0.63	4.9
VCP34	0.09	/ 0.61	0.13	/ 0.56	0.12	/ 0.62	5.2
VCP2	0.13	/ 0.64	0.12	/ 0.59	0.11	/ 0.65	4.8
VCP3	0.12	/ 0.60	0.13	/ 0.64	0.09	/ 0.68	4.7
VCP4	0.11	/ 0.59	0.11	/ 0.67	0.13	/ 0.63	5.3
RANDOM	0.11	/ 0.63	0.13	/ 0.62	0.12	/ 0.64	4.9

**Table 3.4:** Results of molecular replacement of rVCP 2,3,4; using different molecular replacement methods and software.

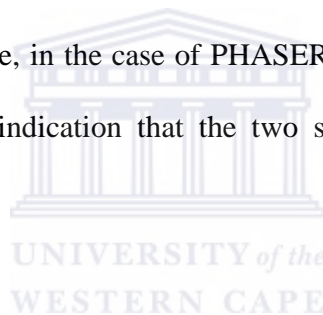
SEARCH PROBE	Correlation coefficient	Rfactor
Search with VCP 2, fix it and search against VCP 3	0.12	0.62
Search with VCP 2, fix it and search against VCP 4	0.11	0.61
Search with VCP 3, fix it and search against VCP 4	0.13	0.58
Fix VCP 2 and search against VCP 3; fix the resultant molecule and search against VCP 4	0.14	0.64
Fix VCP 2 and search against VCP 4; fix the resultant molecule and search against VCP 3	0.13	0.59
Fix VCP 3 and search against VCP 4; fix the resultant molecule and search against VCP 2	0.13	0.60
Fix VCP 3 and search against VCP 2; fix the resultant molecule and search against random molecule	0.12	0.59

**Table 3.5:** Results of molecular replacement using MOLREP involving fixing of modules and searching against other modules.

From tables 3.4 and 3.5 above, it can be seen that similar correlation coefficient and Rfactor values were obtained in both datasets. This means that fixing one module and searching against another module did not improve correlation coefficient and Rfactor values compared to using only one module. Generally, very high Rfactor values (in the range of 0.58-0.69) and very low correlation coefficient values (in the range of 0.09- 0.14) were obtained across board. These values are way out of accepted range of values for a correct solution. It can also be inferred from table 3.4 that none of the four different methods of molecular replacement used preferentially gave better correlation coefficient / z-score or Rfactor values. Finally, results from the search with a random probe did not give different correlation coefficient / z-score and Rfactor values compared with those obtained from VCP probes. For example, correlation coefficient and Rfactor values of

0.12 and 0.59 respectively obtained in the table 3.4 by the random probe are within the same range of values obtained from the VCP probes. Similarly, Z-score value of 4.9 obtained by the random search model was in the same order of values obtained by the VCP search models (table 3.4).

Molecular replacement exploits the fact that proteins with similar sequence identities (usually greater than 20%) fold in a similar manner. Therefore, positioning the probe (the protein with a known three-dimensional structure and with a reasonably high sequence identity) within the unit cell of the target crystal is expected to produce a model that matches very closely with the experimental, target model. The extent of ‘closeness’ is usually measured by the correlation coefficient and Rfactor (or Z- score, in the case of PHASER). A higher correlation coefficient / Z-score and low Rfactor are an indication that the two structures converge and hence fold similarly.



Cell content analysis [Matthew, 1968] conducted on the crystal predicted the presence of one molecule of the protein in the asymmetric unit with Matthew coefficient of 2.1 and solvent content of 39%. Molecular replacement was, therefore, expected to produce a correlation coefficient of 0.5 or higher (or Z-score of 8 and higher) for a correct solution [Kissinger *et al*, 1999; Storoni *et al*, 2004] with good VCP search probes. From tables 3.4 and 3.5, it is seen that the highest correlation factor achieved was 0.14; which is less than a third of what was expected of a correct solution. Similarly, the highest Z-score value achieved was 5.3, which falls into the category of ‘unlikely solution’ in PHASER [Storoni *et al*, 2004].

From these arguments, it could be concluded that molecular replacement did not find a correct solution.

### 3.9.3 Refinement

Though high correlation coefficient / Z-score and low Rfactor values are indicators of a successful finding of a solution in the molecular replacement, sometimes a lower correlation coefficient / Z-score and higher Rfactor values do not necessarily mean that a correct solution had not been found. The ultimate determinant of the achievement of a correct solution is refinement. A correct solution would successfully refine with a substantial decrease in Rfactor and Rfree.

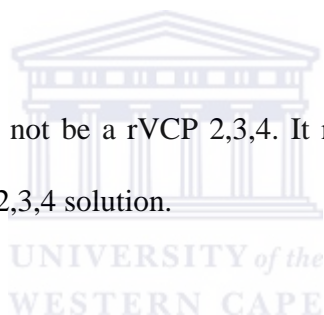
Attempts were made to refine the model generated by MOLREP, using a program called REFMAC5 [Murshodov *et al*, 1997]. Despite several attempts, refinement could not successfully complete. Consistent failure of the refinement step is an indication that the model generated by molecular replacement may be a random solution. To further confirm this, a look at the log file generated by refinement (even though the program could not successfully run to completion) showed that while Rfactor systematically decreased from 0.64 to 0.51 in 8 cycles of refinement, Rfree systematically increased from 0.49 to 0.67 (data not shown). This divergence (decrease in Rfactor and increase in Rfree) is a clear indication that the model generated by molecular replacement was a random solution. A correct solution was expected to cause a decrease in both Rfree and Rfactor.

Based on this observation and results obtained in the molecular replacement (tables 3.4 and 3.5), it was concluded that molecular replacement failed to find the correct solution.

#### **3.9.4 Implications of the failure to find a correct solution by molecular replacement**

Failure to find a correct solution by molecular replacement could be explained by one of the following two reasons:

- (a) rVCP 2,3,4 may fold very differently from any of the known VCP modules or combination of modules used for the molecular replacement. In other words, the structure of rVCP 2,3,4 may be entirely different from the known structures of VCP.
- (b) The protein in the crystal may not be a rVCP 2,3,4. It might have been produced from an 'impurity protein' in the rVCP 2,3,4 solution.



Various research [Kirkitaдзе *et al*, 1999(a, b, c & d); Henderson *et al*, 2001] had indicated that the environment of the neighbouring modules (presence or absence of neighbouring modules) might have an influence on the overall structure of VCP. Therefore, it will not be too surprising that the absence of module 1 in rVCP 2,3,4 would make the conformation of rVCP2,3,4 different from any of the VCP models used in the molecular replacement; resulting in very low correlation coefficients / Z-score.

This possibility, though not completely ruled out, does not sound very convincing. The absence of module 1 in rVCP2,3,4 is not expected to modify the overall structure to such an extent that it would fold entirely different from the known structure of the full length VCP or the structure of



any VCP modules combination. This is more so considering the fact that VCP modules used in the molecular replacement had very high sequence identity (more than 95% in some of them) to rVCP 2,3,4. The absence of module 1 in rVCP 2,3,4 may alter the overall structure of the molecule but the sushi / CCP domains making up the molecule would be expected to maintain their characteristics architectural identity. Therefore, molecular replacement was expected to find a correct solution for at least one of the VCP fragment if the protein in the crystal was a indeed a VCP.

The second possibility appears to be the reason for the failure of molecular replacement. If the supposedly rVCP 2,3,4 crystal was indeed a different protein, then ‘non-convergence’ in molecular replacement was expected; unless the sequence identity of the protein in the crystal, by any chance, happened to be close to that of VCP.

The possibility of the crystal coming from a different protein is supported by the presence of impurities seen in the SDS-PAGE gel of the post-crystallization gel (fig. 3.17). These impurities were not detected in the SDS-PAGE gel of the desalted solution prior to superconcentration; probably because their concentration was too low to be detected prior to superconcentration.

Perhaps the strongest argument to support the assumption that the crystal was from an ‘impurity’ protein is the correlation coefficient values of the random probe protein used in the molecular replacement (table 3.4 and 3.5). Molecular replacement with a random protein probe was carried out to find out how the correlation coefficient of the VCP probes compares with a random protein probe. Tables 3.4 and 3.5 show clearly that similar low correlation coefficient values

were obtained for all the VCP probes and the random probe. This is an indication that; like the random protein probe, the VCP probes do not have sequence identity to the protein in the crystal. This may suggest that the crystal might be a non-VCP 234 crystal.

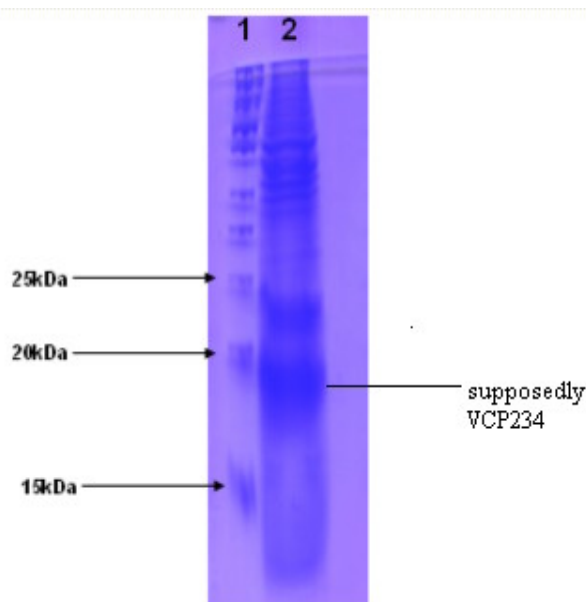
### **3.9.5 Post crystallization SDS PAGE gel confirms presence of impurity**

The post crystallization SDS PAGE gel of the VCP234 solution is represented in fig 3.17. The gel shows the presence of two dominant bands; one of which is just below the 20kDa mark (supposedly rVCP 2,3,4 with a molecular weight of about 19.5kDa) and one just below the 25kDa mark. Apart from these two bands, there are also a number of other faint bands of higher molecular weights. This confirms that the rVCP 2,3,4 solution used for crystallization was indeed impure; contrary to the observations on the silver stained gel of the purification and desalting steps (fig. 3.4 and 3.5 respectively). These impurities could not be seen prior to concentration of the protein solution. The supposedly rVCP 2,3,4 band is the most predominant (fig 3.17). Even though this band is located in the correct position expected for rVCP 2,3,4, we cannot say without doubt that the protein was indeed rVCP 2,3,4 since the protein was not sequenced prior to crystallization trials.

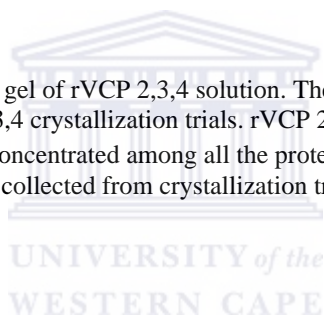
It should be noted that the gel was overloaded with 50 $\mu$ l solution; indicating that rVCP 2,3,4 solution present was about 25 $\mu$ l (assuming that equal amounts of rVCP 2,3,4 solution and precipitant were present in the loaded solution). Finally, it should also be noted that highly concentrated rVCP 2,3,4 (6.1mg/ml) was used in the gel. With such high concentration and volume, it is not surprising that all faint bands originally not visible in the silver stained gel prior to superconcentration (fig. 3.4 and 3.5) were now visible on the SDS-PAGE gel.

It is very likely that the crystallized protein is the one with the most dominant band (the supposedly VCP234; just below the 20kDa mark). This makes rVCP 2,3,4 a possible protein in the crystal. However, since sequence analysis was not conducted on the rVCP 2,3,4 solution to confirm its identity beyond all reasonable doubt, we cannot conclude that the crystal was rVCP 2,3,4. The protein in the other dominant band (just below 25kDa mark) could also be a possible protein that crystallized. It is most unlikely that the crystal was formed from any of the high molecular weight proteins with minor bands seen on the gel (fig 3.17) since they are present in minute quantities, compared to the other two bands.

Since an SDS-PAGE gel was not conducted on the concentrated rVCP 2,3,4 solution prior to setting up of crystallization trials, these impurities were not discovered prior to crystallization. The gel was run after molecular replacement failed to produce a correct solution. The essence of running this gel was to find out whether the final rVCP 2,3,4 solution used to set up crystallization trials had some impurities in them.



**Fig. 3.17:** Post crystallization SDS-PAGE gel of rVCP 2,3,4 solution. The gel was overloaded with 50 $\mu$ l of solution collected from clear drops of the rVCP 2,3,4 crystallization trials. rVCP 2,3,4 is expected to be the band just below the 20kDa mark and it is by far the most concentrated among all the proteins present in the solution. Lane 1= Molecular weight marker. Lane 2 = drops collected from crystallization trials.



### 3.9.6 Solvent content analysis of the 'disputed crystal'

The purpose of the solvent content analysis (table 3.6) was to test whether the cell dimensions of the crystals under investigation could possibly fit into the calculated unit cell of a rVCP 2,3,4 crystal. The analysis on the test crystal was done together with the crystals obtained from the solved structures of native VCP and VCP bound to heparin to provide a basis of comparison. Native VCP has been crystallized in 2 forms (form I and II in space group P212121 and C222 respectively [Murthy *et al*, 2001]. The heparin bound VCP was crystallized in space group P2 [Ganesh *et al*, 2004]. The density of protein was taken as 0.73 Dalton per cubic angstrom in all cases.

The volume occupied by mass was calculated as follows:

Molecular weight of protein \* # of asymmetric units in unit cell \* # of molecules per asymmetric unit

Density of protein

	Native VCP (form I crystal) Spacegroup P212121	Native VCP (form II crystal) Spacegroup C222	Heparin bound VCP Spacegroup P2	Crystal under investigation Spacegroup F23
Calculated solvent content (%)	72.2	73.9	76.6	39.1
Volume of unit cell (cubic Å)	918592	2940050	545459	1906162
Matthew coefficient	4.4	4.7	4	2.1
# of asymmetric units in a unit cell	4	8	2	48
# of molecules per asymmetric unit	2	3	2	1
Molecular weight (Dalton)	26000	26000	26000	19500
Volume of cell occupied by mass (cubic Å <sup>3</sup> )	284931	854794	142465	1249315
% of volume occupied by mass	31	29	26	65
Expected solvent content (%)	69	71	74	35
Margin of error	3.2	2.9	2.6	4.1

**Table 3.6:** Solvent content analysis of rVCP 2,3,4 crystal under investigation compared to authentic VCP crystals. Two native VCP crystal forms (I and II) as well as the heparin bound VCP were compared to the test crystal. [Murthy *et al*, 2001, Ganesh *et al*, 2004]. Solvent content and number of asymmetric units in unit cell were calculated by an online program:

[http://adelie.biochem.queensu.ca/~rlc/pfd/links/calcs/vm\\_calc.shtml](http://adelie.biochem.queensu.ca/~rlc/pfd/links/calcs/vm_calc.shtml)

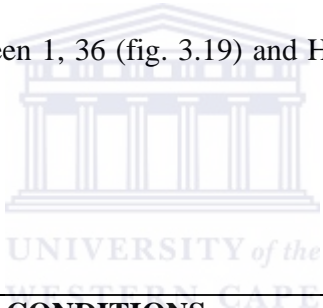
From table 3.6 above, it is clear that the crystal under investigation has a lower solvent content compared to the already solved VCP crystals. Solvent content in the test crystal was 39.1% while it was 72.2%, 73.9% and 76.6% in the form I, form II and heparin bound VCP crystals respectively. The large difference in solvent content could point to the assumption that the crystal under investigation was not a VCP crystal. On the other hand, the absence of module 1 in rVCP 2,3,4 could make the rest of the protein fold in such a way that it could exclude a lot of solvent, resulting in the low solvent content of 39.1% (table 3.6). Therefore, the solvent content alone is not sufficient to rule out the identity of the test crystal as a rVCP 2,3,4 crystal.

The possibility of a crystal fitting into the unit cell is estimated by the difference between the expected solvent content and calculated solvent content, (or the margin of error). The lower the margin of error, the higher the confidence that the crystal could fit into the unit cell. From the table 3.6, the error of margin for all the authentic VCP crystals was between 2.6% and 3.2%. The error of margin of the test crystal was 4.1%. In other words, if the crystal under investigation were indeed a rVCP 2,3,4, it was expected to have a solvent content of 35%, instead of 39% . Although the margin of error in the test crystal is higher than that of any of the authentic VCP crystal, the difference is not high enough to exclude the possibility that the test crystal could be rVCP 2,3,4 crystal. Therefore, based on the margin of error alone, it can be said that the crystal under investigation has right cell dimension that can approximately fit into a unit cell of rVCP 2,3,4. This makes rVCP 2,3,4 a possible protein in the crystal, even though the margin of error is slightly different from other authentic VCP crystals already solved.

### 3.10 Crystallization results of GDH

#### 3.10.1 Results from the initial screening.

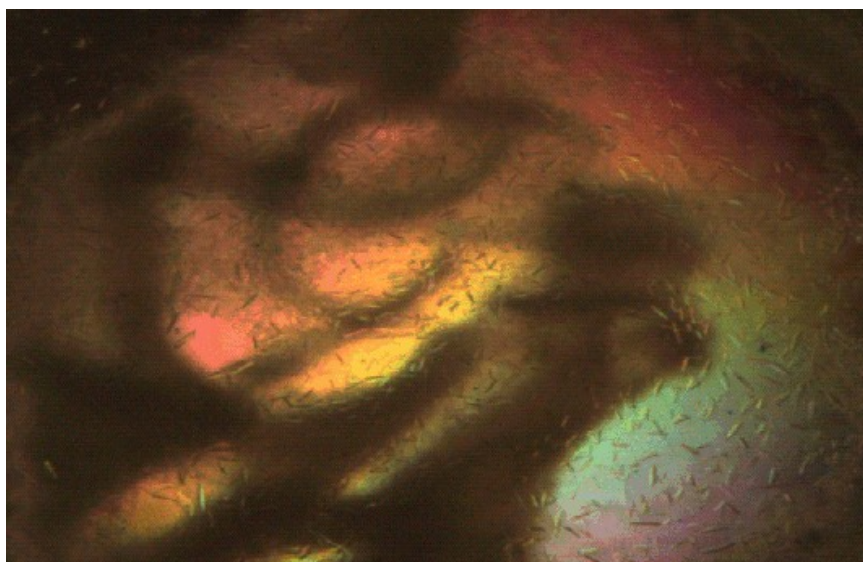
Like VCP 234, initial crystallization screening of GDH was carried out using Hampton screen 1 and 2. Results from the initial screening are presented in the table 3.7 and figures figs. 3.19-3.21 below. As can be seen from the figs. 3.19, 3.20 and 3.21, showers of protein crystals were produced in 3 of the drops within one week of set up. However, the quality of the crystals was not good enough to be used and needed further optimization. Getting crystals within such a short time in 3 drops from the initial screening was very encouraging since it defines a point from which optimization can be carried out. Upon careful observation, it could be seen that the crystals produced by Hampton screen 1, 36 (fig. 3.19) and Hampton screen 1, 15 (fig. 3.21) are plate shaped.



<b>DROP IDENTIFY</b>	<b>CONDITIONS</b>	<b>OBSERVATION</b>
SCREEN 1, 36	8% PEG 8000, 0.1M Tris pH 8.5	Showers of crystals
SCREEN 1, 38	1M Li <sub>2</sub> SO <sub>4</sub> , 0.1M Na acetate pH 5.6, 0.5M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Showers of crystals
SCREEN 1, 15	30% PEG 8000, 0.1M Na Cacodylate pH 6.5, 0.2M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Showers of crystals

**Table 3.7:** Results of the initial crystallization screening of GDH



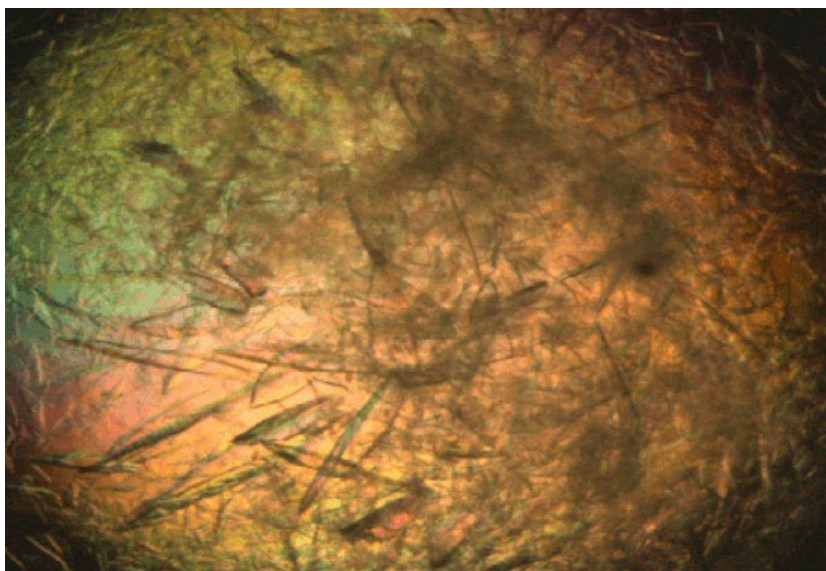


**Fig. 3.18:** Showers of crystals produced from 8% PEG 8000, 0.1M Tris pH 8.5, observed after 4 days.



**Fig. 3.19:** Showers of micro-crystals produced from 1M  $\text{Li}_2\text{SO}_4$ , 0.1M Na Citrate pH 5.6, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ , observed after 8 days.



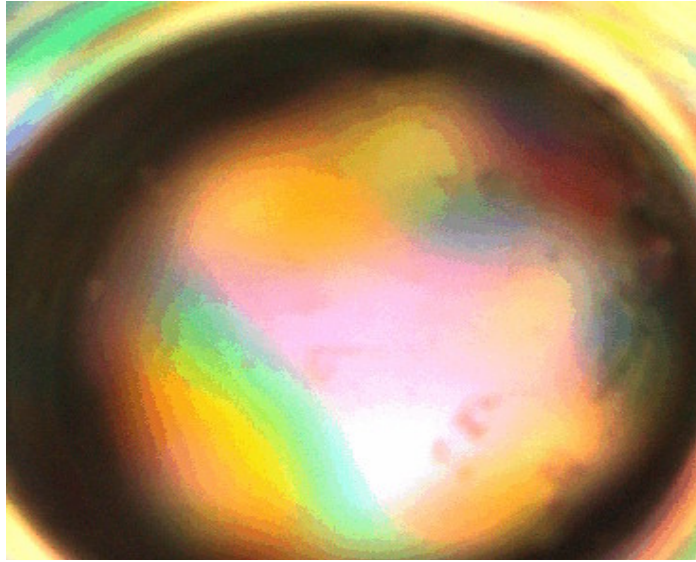


**Fig. 3.20:** Showers of crystals produced from 30% PEG 8000, 0.1M Na Cacodylate pH 6.5, 0.2M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, observed after 12 hours.

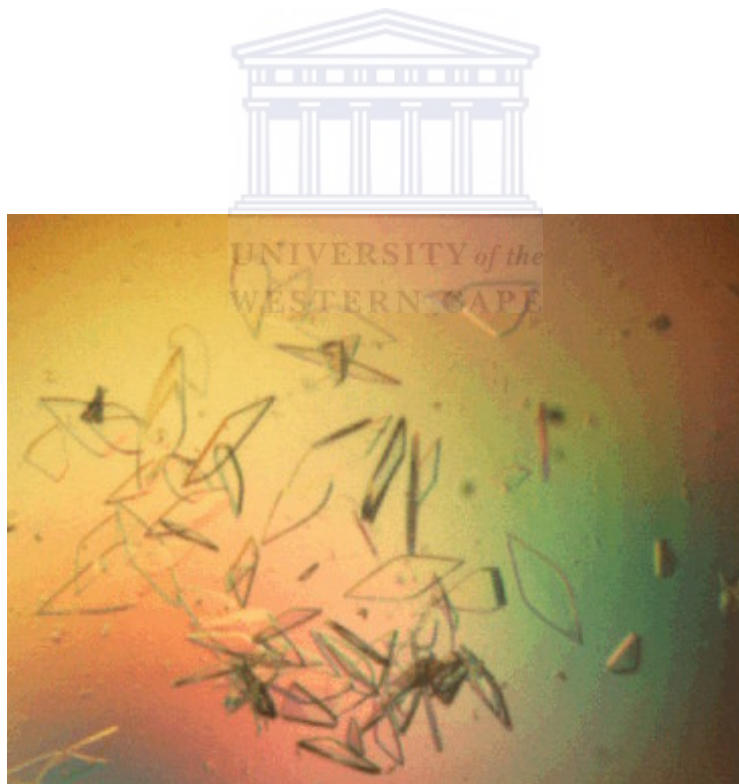
### 3.10.2 Optimization of initial results of GDH trials

Showers of crystals produced from Hampton screen 1, 15 (30% PEG 8000, 0.1M Na Cacodylate pH 6.5, 0.2M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, fig. 3.20 above) were first observed 12 hours after setting up the trials. Based on this duration of first appearance of crystals, the method of decoupling nucleation from growth using '*in situ*' time intervention was used to optimize the crystals as described previously.

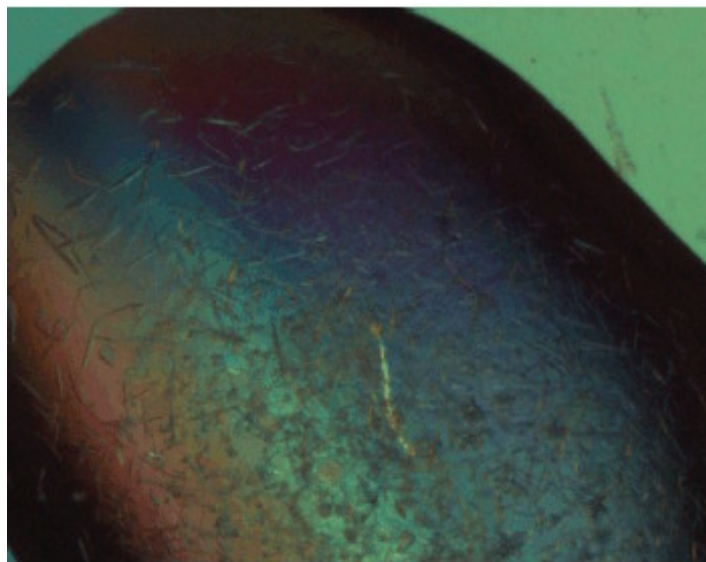
In brief, 3 drops under the same conditions were set up. The cover slide was removed and 500µl of distilled water was added to the reservoir after 3 hours and the drop re-sealed. The same procedure was carried out on the other two drops at 6 and 9 hours respectively. In this way, the concentrations of all precipitating solutions were decreased by half. The following are the results observed one week after the intervention.



**Fig. 3.21:** Drop intervention time = 3 hours after of set up.



**Fig. 3.22:** Drop intervention time = 6 hours after of set up.



**Fig. 3.23:** Drop intervention time = 9 hours after of set up.

Crystals were observed in fig. 3.22 and fig. 3.23 while the drop in fig. 3.21 remained clear. The quality of crystals produced after 6 hours intervention (fig. 3.22) was better than crystals produced after 9 hours intervention (fig 3.23). Bigger and fewer crystals were observed after 6 hours intervention as compared to 9 hours intervention.

The effect of decreasing the concentration of precipitants *in situ* is to bring the crystallization system from nucleation zone to metastable state; thus preventing further nucleation while encouraging growth of already nucleated sites (fig. 1.8). The results show that nucleation of the GDH crystals possibly occurred between 3 and 6 hours after setting up the drop. Prior to 3 hours of set up, nucleation had not occurred, thus resulting in a clear drop seen in fig. 3.21. Decreasing the precipitant concentration prior to nucleation would in no doubt result in a clear drop.

Between 6 and 9 h of set up, excessive nucleation had already occurred, making it impossible to produce fewer and larger crystals after decreasing the concentration. Decreasing the concentration of precipitant at the time excessive nucleation had already occurred is not expected to improve crystal quality. Decreasing the precipitant concentration *in situ* would enhance crystal quality only when it is done at the time nucleation had just started. The improved crystal quality seen in fig. 3.22 is a clear indication that nucleation occurred between 3 and 6 hours after the drop was set up. The exact time of nucleation, however, cannot be determined from the above results.

Although improved, the sizes of the optimized crystals (fig. 3.22) were too small to be shot onto the X-ray beam and needed further optimization. Attempts were made to further optimize the crystals by the same *in situ* decrease in concentration of precipitants. Based on the result obtained in fig. 3.22 (on assumption that nucleation occurred between 3 and 6 hours), the drop was intervened by the *in situ* decrease in concentration at 3, 4, 5 and 6 hours after set up as described above. However, results obtained (not shown) were all similar to that observed for 6 hours intervention (fig 3.22).

### **3.10.3 Optimization of GDH crystal, using pH gradient**

Following the failure of the *in situ* decrease in concentration method to improve the crystals obtained in fig. 3.22 above, another optimization method was employed; namely pH gradient. From the predicted crystallization pH of GDH by CrysPred (fig. 3.2), higher pH ranges of 6.9 to 9.0 were explored; using Tris or Na HEPES. CrysPred had predicted that GDH has a highest chance of crystallizing within pH ranges of 6.5-8.5 (fig. 3.2).

The same *in situ* time intervention procedure and crystallization conditions that produced the crystals in fig. 3.22 were followed. The only difference was that instead of pH 6.5, the pH of the buffer was screened from 6.9-9.0, in 0.3 units stepwise. After 6 h of set up, the reservoir concentration of each set up was decreased by half as described above.

After 8 days, single rectangular shaped (plate-like) crystals of appreciable sizes were obtained from the drop with Tris pH 8.7; which diffracted to about  $6.5\text{\AA}$  (fig. 3.24).

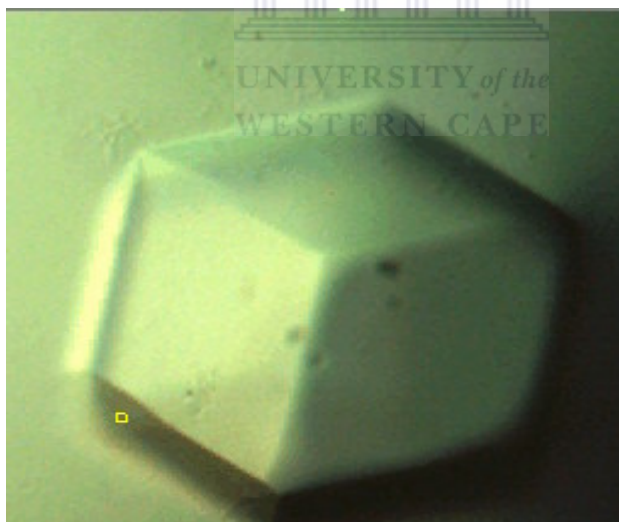


**3.24:** Crystals of GDH produced from 30% PEG 8000, 0.1M Tris pH 8.7 & 0.2M  $(\text{NH}_4)_2\text{SO}_4$ . The drop was diluted by 50% after 6 hours of set up. Crystals produced after 8 days and diffracted to  $6.5\text{\AA}$ .

### 3.10.4 The use of oils to optimize GDH crystals

Showers of micro-crystals produced from 1M  $\text{Li}_2\text{SO}_4$ , 0.1M Na Citrate pH 5.6, 0.5M  $(\text{NH}_4)_2\text{SO}_4$  (fig. 3.19) were optimized using the methods of oils. In brief, different volumes of different proportions of paraffin and silicon oils were poured on the surface of the reservoir, as described previously.

After 21 days of set up, single large crystals were observed in the drop containing 800 $\mu\text{l}$  of 3:1 (Paraffin: Silicon) oils. The crystal diffracted to about 8.5 $\text{\AA}$ . The crystals were very big compared to the rVCP 2,3,4 crystal (fig. 3.25). The dimensions of the crystals were, however, not measured, as the microscope was not equipped with a measuring device.



**Fig. 3.25:** Crystal of GDH grown from 1M  $\text{Li}_2\text{SO}_4$ , 0.1M Na Citrate pH 5.6, 0.5M  $(\text{NH}_4)_2\text{SO}_4$ , with 800 $\mu\text{L}$  of 3:1 (Paraffin: Silicon) oil. The crystal was produced 21 days after set up of the drop and diffracted to 8.5 $\text{\AA}$ .



### 3.11 Data collection and processing of GDH data

X-ray data were collected on two different GDH crystals (fig. 3.24 and 3.25). However, neither data set could be indexed. Therefore, processing could not go beyond image collection. The exact reason why the diffraction could not be indexed in both cases is not very clear but it appears that the low-resolution data (6.5Å and 8.5Å respectively) could account for this.

Producing GDH crystals was fairly easy. Three drops in the initial Hampton Screen 1 and 2 produced crystals of GDH at a protein concentration of 15mg/ml (table 3.7). However, producing high quality, well-diffracted crystals appears to be a difficult task. The inherent difficulty in the production of high-resolution GDH crystal may be due to the large size of the protein. Initial electron microscopy studies (unpublished) had suggested that there is a hexamer of protein in an asymmetric unit of GDH of *Bacteriodes fragilis*, with each monomer being about 48kDa in size [Britton *et al.*, 1992]. This comes to about 288kDa, the mass of protein in an asymmetric unit of the unit cell of the protein. With such a large size, crystal-packing interaction may be weakened, preventing the formation of well-ordered packing of the protein molecules. Ultimately, poorly ordered crystals would be formed, resulting in low-resolution diffraction.

Attempt was made to collect the data of the GDH crystal using a synchrotron at the European Synchrotron Radiation Facility (ESRF) in France. However, the data collected on the synchrotron was no better than the one collected on the in-house x-ray machine. Therefore, the data could not be further processed.



UNIVERSITY *of the*  
WESTERN CAPE



## Chapter 4

### Summary and suggestions for further work

In summary, a systematic approach was followed in an attempt to crystallize and solve the structures of rVCP 2,3,4 and rGDH. Although the final results were inconclusive, systematic methods were used in each of the 3 major steps; namely expression, crystallization and structural determination.

Results of rVCP 2,3,4 data were inconclusive. The major problem that made the results of rVCP 2,3,4 crystallization inconclusive hinges on the identity of the crystal produced in the trials. Results and analysis indicate that the crystal produced may or may not be that of rVCP 2,3,4. Since the crystal's identity is questionable and inconclusive, a wide range of suggestions for further work would be made to take into account of the two possibilities.

Firstly, the identity of the expression solution used for crystallization trials should be confirmed prior to crystallization trials. It is therefore suggested that any further work would attempt to conduct an N-terminal sequence analysis on the protein solution after expression and compare it to the sequence of rVCP 2,3,4. The sequence analysis should be repeated after the purification step; prior to crystallization trials even if VCP 2,3,4 is already confirmed after expression. This would remove any doubt about the success of the purification step.

Secondly, careful consideration should be given to producing very pure rVCP 2,3,4 solution devoid of any impurities. It is suggested that new, unused heparin columns should be used to bring about a very high degree of purity. The final purified solution should be visualized on the gel prior to setting up of the crystallization trials.

Despite several experiments, only one of them produced a protein crystal. Any further work would have to attempt to reproduce conditions both in terms of increasing the success rate and ease of crystal production. To this end, it is suggested that other crystallization screens (other than Hampton screen 1 and 2) should be explored. Other crystallization methods, such as high pressure crystallization, should also be explored.

If indeed the crystal produced was a rVCP<sub>2,3,4</sub>; failure of molecular replacement to find the correct solution might be an indication that rVCP<sub>2,3,4</sub> folds in a manner that is very different to that of any of the known VCP structures or fragments of VCP. Therefore, it is suggested that another phasing method (aside of molecular replacement) should be explored in any further work. In particular, heavy atom derivative methods (isomorphous replacement) should be explored.

Attempts should also be made to re-process the rVCP<sub>2,3,4</sub> data. The possibility that data was processed in the wrong spacegroup should not be ignored as this could certainly affect the success of finding a solution by Molecular Replacement. The use of POINTLESS, a new program within the CCP4 suite, could be very useful in determining

whether or not the crystal actually crystallized in F23 spacegroup. Processing the data in a spacegroup confirmed by POINTLESS and obtaining better statistics may lead to success of molecular replacement and also confirm that the crystal was indeed rVCP2,3,4.

Crystal production of rGDH was fairly easy and fast. Using Hampton screens 1 and 2, three different conditions produced authentic crystals. These initial crystals were successfully optimised systematically to produce bigger single crystals. The problem with the rGDH dataset is that the crystals were of low quality and diffracted to very low resolution. Whether this was the reason why the data could not be indexed or not is not clear.

It is, therefore, recommended that an attempt should be made to visit the rGDH data and re-process it. Apart from this, means should be devised to produce better quality crystal. To this end, it is suggested that other crystallization methods should be tried to crystallize it. Finally, other crystallization screens could also be explored.

## REFERENCES

**Al-Mohanna**, Parhar, R. & Kotwal, G. J. (2000). Vaccinia virus complement control protein is capable of protecting xenoendothelial cells from antibody binding and killing by human complement and cytotoxic cells. *Transplantation* **71**, 796–801.

**Ataka** M & Tanaka S (1986). The growth of large single crystals of lysozyme. *Biopolymers* **25**, 337–350.

**Ataka** M (1993). Protein crystal growth: an approach based on phase diagram determination. *Phase Transitions*. **45**, 205–219.

**Baker** PJ, Waugh ML, Wang XG, Stillman TJ, Turnbull AP, Engel PC & Rice D.W (1997). Determinants of substrate specificity in the superfamily of amino acid dehydrogenases. *Biochemistry* **36**, 16109–16115.

**Baldock** P, Mills V, & Shaw S. PD. (1996). A comparison of microbatch and vapor diffusion for initial screening of crystallization conditions. *J. Crystal Growth* **168**, 170–174.

**Bergfors** TM (1999). Protein crystallization: Techniques, Strategies & Tips (Lab Manual). International Uni. Line. 27–27.

**Blow DM, Chayen NE, Lloyd LF & Saridakis E (1994).** Control of nucleation of protein crystals. *Protein Sci.* **3**, 1638-1643.

**Bork P, Downing AK, Kieffer B & Campbell ID (1996).** Structure and distribution of modules in extracellular proteins. *Quarterly Reviews of Biophys.* **29**, 119–167.

**Britton, KL, Baker PJ, Rice DW & Stillman TJ (1992).** Structural relationship between the hexameric and tetrameric family of glutamate dehydrogenase. *Eur. J. Biochem.* **209**, 851-859.

**Buller RML & Palumbo GJ (1991)** Poxvirus pathogenesis. *Microbiol. Rev.* **55**, 80-122.

**CCP4:** Collaborative Computational Project, Number 4. 1994. "The CCP4 Suite: Programs for Protein Crystallography". *Acta Cryst.* **D50**, 760-763.

**Cereghino JL & Cregg JM (2000).** Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology reviews* **24**, 45-66.

**Chayen NE (1996).** A novel technique for container less protein crystallization. *Protein Engineering* **9**, 927-929.

**Chayen NE, Lloyd LF, Collyer CA & Blow DM (1989).** Trigonal crystals of glucose isomerase require thymol for their growth and stability. *J. Cryst. Growth.* **97**, 367-374.

**Chayen** NE, Radcliffe JW, Blow DM (1993). Control of nucleation in the crystallization of lysozyme. *Protein Sci* **2**, 113-118

**Chayen** N E, Shaw SPD & Baldock P (1994). New developments of the IMPAX small-volume automated crystallization system. *Acta Cryst.* **D50**, 456-458.

**Chayen** NE, Shaw-Steward PD & Blow DM (1992). Microbatch crystallization under oil-a new technique allowing many small-volume crystallization trials. *J. Crystal Growth* **122**, 176-180.

**Chayen** NE (1998). Comparative Studies of Protein Crystallization by Vapour-Diffusion and Microbatch Techniques *Acta Cryst.* **D54**, 8-15.

**Chayen** N E, Shaw SPD, Maeder DL & Blow DM (1990). An automated system for micro-batch protein crystallization and screening. *J. Appl. Crystallogr.* **23**, 297-302.

**Chayen** NE (1997). The role of oil in macromolecular crystallization. *Structure* **5**, 1269–1274.

**Chayen** NE (2002). *Trends Biotechnol.* **20(98)**. Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760-763.

**Chernov AA** (1997). Crystals built of biological macromolecules. Phys. Rep. **288**, 61–75.

**Chernov AA** (2003). Protein crystals and their growth. J. Struct. Biol. **142**, 3-21.

**Chernov A A** (1984). Growth of Crystals. Modern Crystallography III: Springer: Berlin.

**Chiruvolu V, Cregg JM & Meagher MM** (1997). **Recombinant protein production in an alcohol oxidase-defective strain of Pichia pastoris in fed-batch fermentations.** Enzyme Microb. Technol., **21**, 277-283.

**Conti E, Lloyd LF, Akins J, Franks NP, Brick P** (1996). Crystallization and preliminary diffraction studies of firefly luciferase from Photinus pyralis Acta Cryst. **D52**, 876-878

**Cooper NR, Jensen FC, Welsh RM, Jr & Oldstone MB** (1974). Lysis of RNA tumor viruses by human serum: direct antibody-independent triggering of the classical complement pathway. Journal of Experimental Medicine **144**, 970–984.

**Cregg JM, Madden KR, Barringer KJ, Thill GP & Stillman CA** (1989). Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris. Mol. Cell. Biol. **9**, 1316-1323.

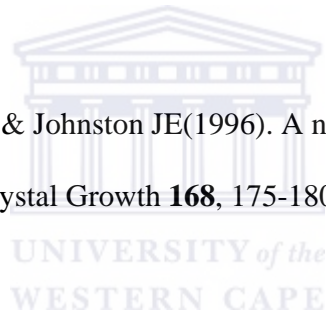
**Cregg JM & Madden KR** (1987). Pichia pastoris by gene Development of yeast transformation systems and construction of methanol-utilization-defective mutants of

disruption. In: *Biological Research on Industrial Yeasts* **2**, 1-18. (Stewart GG, Russell I, Klein RD & Hiebsch RR. Eds.) CRC Press, Boca Raton, FL

**Cregg JM**, Davis GR, Smiley BL, Cruze J, Torregrossa R, Velicelebi G, Thill GP & Kellaris PA (1987). High level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technology* **5**, 479-485.

**Curtis RA**, Prausnitz JM & Blanch HW (1998) Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnology & bioengineering* **57**, 11-21.

**D'Arcy A**, Elmore C, Stihle M & Johnston JE(1996). A novel approach to crystallising proteins under oil *Journal of Crystal Growth* **168**, 175-180.



**Drenth J**, Haas C (1998). Nucleation in Protein Crystallization. *Acta Cryst.* **D 54**, 867-872.

**Ducruix A** & Giege R (1992). *Crystallization of Nucleic acids and proteins. A practical approach.* Oxford: IRL Press Oxford University Press. 82-90.

**Earl PL** & Moss B (1989). Vaccinia virus; in *Genetics maps* (ed.) S J O'Brian (New York: Cold Spring Harbor Lab), **5<sup>th</sup>** edition, pp 1138–1148.



**Ebenbichler** CF, Thielens NM, Vornhagen R, Marschang P, Arlaud GJ & Dierich MP (1991). Human immunodeficiency virus type 1 activates the classical pathway of complement by direct C1 binding through specific sites in the transmembrane glycoprotein gp41. *J Exp Med* **174**, 1417–1424.

**Egli** T, Van Dijken JP, Veenhuis M, Harder W & Fiecher A (1980). Methanol metabolism in yeasts: Regulation of the synthesis of catabolic enzymes. *Arch. Microbiol.* **124**, 115-121.

**Ellis** SB, Brust PF, Koutz PJ, Waters AF, Harpold MM & Gingeras TR (1985). Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. *Mol. Cell. Biol.* **5**, 1111-1121.

**Fehér** G, Kam Z (1985). Nucleation and growth of protein crystals: General principles and assays. *Method Enzymol.* **114**, 77–112.

**Fox** K & Karplus A (1993). Crystallization of old yellow enzyme illustrates an effective strategy for increasing protein crystal size. *J Mol. Biol.* **234**, 502-507.

**Frank** MM & Fries LF (1989). Complement; in *Fundamental immunology* (ed.) W E Paul (New York: Raven), 679–701.

**Ganesh** VK, Smith SA, Kotwal GJ & Murthy KH (2004). Structure of vaccinia complement protein in complex with heparin and potential implications for complement regulation. PNAS **101**(24), 8924-8929.

**Gerrett** RH & Grisham CM (1995). BIOCHEMISTRY, Saunders College Publishing, NY, 598-603, 832-834.

**Geux** N & Peitsch MC (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modelling. Electrophoresis **18**, 2714-2732.

**Haire** LL F (1996). Strategies for protein crystal growth— screening and optimisation. PhD Thesis. University of London, page 144.

**Harlos** K (1992). Micro-bridge for sitting-drop crystallization. J Appl Cryst. **25**, 536-538.

**Hankamer** B, Chayen NE, De Las Rivas J & Barber J (1992). Robert Hill Symposium on Photosynthesis, 11-12.

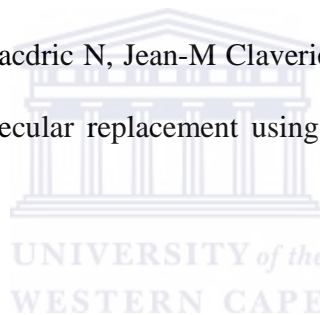
**Henderson** CE, Bromek K, Mullin NP, Smith BO, Uhrhin D & Barlow PN (2001). Solution Structure and Dynamics of the Central CCP Module Pair of a Poxvirus Complement Control Protein. Journal of Mol. Biol. **307**, 323-339.

**Holers VM, Cole JL, Lubin DM, Seya T & Atkinson J P** (1985). Human C3b- and C4b-regulatory proteins: a new multigene family; *Immunol. Today* **6**, 188–192.

**Ikeda F, Haraguchi Y, Jinno A, Iino Y, Morishita Y, Shiraki H & Hoshino H** (1998). Human complement component C1q inhibits the infectivity of cell-free HTLV-I. *J Immunol* **161**, 5712–5719.

**Ivana Kutá Smatanová** (2002). Crystallization of biological macromolecules. *Materials Structure*, **9**, 14-16

**Jean-Baptiste C, Karsten S, Cacdric N, Jean-M Claverie & Chantal A** (2004). CaspR: a web-server for automated molecular replacement using homology modelling. *Nucleic Acids Research*, **32**, 606-609.



**Jonathan PKD, Ard AL & Michele V** (2004). Inhibition of protein crystallization by evolutionary negative design. *Phys. Biol.* **1**, 9–13.

**Kantardjieff KA & Rupp B** (2004). Protein Isoelectric Point as a Predictor for Increased Crystallization Screening Efficiency. *Bioinformatics* **20(14)**, 2162-2168.

**Kantardjieff KA, Jamshidian M & Rupp B** (2004). Distributions of pI vrs pH provide strong prior information for the design of crystallization screening experiments. *Bioinformatics* **20(14)**, 2171-2174.

**Kashchiev D** (2000). Nucleation: Basic Theory with Applications, Butterworth-Heinemann, Oxford.

**Kirkitadze MD, Krych M, Uhrin D, Dryden D, Cooper A, Wang X, Hauhart R, Atkinson JP & Barlow PN (1999a)**. Independent melting modules and highly structured intermodular junctions within complement receptor type 1. *Biochemistry* **38**, 7019-7031.

**Kirkitadze MD, Handerson C, Price NC, Kelly SM, Mullin NP, Parkinson J, Dryden DTF & Barlow PN (1999b)**. Central modules of the Vaccinia virus complement control protein are not in an extensive contact. *Biochem. J.* **343**, 167-175.

**Kirkitadze MD, Krych M, Dryden DTF, Wang X, Atkinson JP, Kelly SM, Price NC & Barlow PN (1999c)**. Co-operativity between modules within C3b-binding site of complement receptor type 1. *FEBS Letters* **459**, 133-138.

**Kirkitadze MD, Jumel K, Krych M, Dryden D, Atkinson JP, Hardings S & Barlow PN (1999d)**. Combining ultracentrifugation with fluorescence to follow the unfolding of modules 16-19 of complement receptor type 1. *Prog. Polymer Colloid Sci.* **113**, 164-167.

**Kissinger RC, Gehlhaar KD & Fogel BD (1999)**. Rapid automated molecular replacement by evolutionary search. *Acta Crystallographica* **D55**, 484-491.

**Kopito RR** (2000). Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10**, 524-530.

**Kotwal GJ & Moss B** (1988a). Analysis of a large cluster of nonessential genes deleted from a vaccinia virus terminal transposition mutant. *Virology* **167**, 524-537.

**Kotwal GJ & Moss B** (1988b) Vaccinia virus encodes a secretory polypeptide structurally related to complement control proteins. *Nature* **335**, 176-178.

**Kotwal GJ** (1994). Purification of virokines using ultrafiltration. *Am. Biotech. Lab.* **12**, 76-77.

**Kotwal G J & Moss B** (1989). Vaccinia virus encodes two proteins that are structurally related to members of the plasma serine protease inhibitor superfamily; *J. Virol.* **63**, 600–696.

**Kotwal GJ** (2000). Poxviral mimicry of complement and chemokine system components: what is the end game? *Immunology Today* **21(5)**, 242-248

**Kotwal GJ** (1996). The great escape: immune evasion by pathogens. *Immunologist* **4/5**, 157-164

**Kotwal** GJ, Issacs ST, McKenzie R, Frank MM & Moss B (1990). Inhibition of the complement cascade by the major secretory protein of vaccinia virus *Science* **250**, 827–830.

**Kotwal** GJ, Reynolds D, Keeling K, Howard J & Justus DE (1998). Vaccinia virus complement control protein is a virokinin with lysozyme-like heparin binding activity: possible implications in prolonged evasion of host immune response. In: Talwar GP Nath I (Eds.), 10<sup>th</sup> international congress of immunology, New Delhi, India, 315-320.

**Kuznetsov** YG; Malkin AJ & McPherson A (2001). The liquid protein phase in crystallization: a case study-intact immunoglobulins. *Journal of Crystal Growth* **232**, 30-39.

**Lalani** AS & McFadden G (1997). Secreted poxvirus chemokine binding proteins. *J. Leukoc Biol.* **62**, 570–576.

**Lambris** JD, Sahu A & Wetsel R (1998). Chemistry and biology of C3, C4 and C5. In *The human complement system in health and disease*. Volanakis JE & Frank M, ed Marcel Dekker, Inc., New York, 83-118.

**Liszewski** MK, Leung M, Cui W, Subramanian VB, Parkinson J, Barlow PN, Manchester M & Atkinson JP (2000). Dissecting sites important for complement

regulatory activity in membrane cofactor protein (MCP; CD46). *J. Biol. Chem.* **275**, 37692–37701.

**Luft JR**, Arakali SV, Kirisits MJ, Kalenik J, Wawrzak I, Cody V, Pangborn WA & DeTitta GT (1994). A Macromolecular Crystallization Procedure Employing Diffusion Cells of Varying Depths as Reservoirs to Tailor the Time Course of Equilibration in Hanging and Sitting Drop Vapor Diffusion and Microdialysis Experiments, *J. Appl. Cryst.* **27**, 443-452.

**Luft JR**, Albright DT, Baird JK & DeTitta GT (1996). The Rate of Water Equilibration in Vapor-Diffusion Crystallizations: Dependence on the Distance from the Droplet to the Reservoir, *Acta Cryst.* **D52**, 1098-1106.

**Makrides SC** (1998). Therapeutic inhibition of the complement system. *Pharmacol. Rev.* **50**, 59–87.

**Malkin AJ** & McPherson A (1993). Light scattering investigations of protein and virus crystal growth: ferritin, apoferritin and satellite tobacco mosaic virus. *J. Cryst. Growth* **128**, 1232-1235.

**Matthews BW** (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.

**McKenzie R**, Kotwal GJ, Moss B, Hammer CH & Frank MM (1992). Regulation of complement activity by vaccinia virus complement-control protein. *J. Infect. Dis.* **166**, 1245–1250.

**McPherson A** (1999). *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press, New York.

**McPherson A** (1982). *Preparation and analysis of protein crystals*. New York: John Wiley.

**Mikol V**, Hirsch E & Giege R (1990). Diagnostic of precipitant for the biomolecule crystallization by quasi-elastic light scattering. *J. Mol. Biol.* **213**, 187-195.

**Muller-Eberhard HJ** (1988). Molecular organization and function of the complement system; *Annu. Rev. Biochem.* **57**, 321–347.

**Mullick J**, Kadam A & Sahu A (2003). Herpes pox viral complement control proteins: ‘the mask of self’. *Trends in Immunology*, **24**, 500-506.

**Murshudov G** (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst.* **D53**, 240-255.



**Murthy K, Smith SA, Ganesh VK, Judge KW, Mullin N, Barlow PN, Ogata CM & Kotwal, GJ (2001)** Crystal structure of a complement control protein that regulates both pathways of complement activation and binds heparin sulfate proteoglycans. *Cell* **104**, 301–311.

**Nanev C (2007)**. Protein crystal nucleation: recent notion. *Cryst. Res. Technol.* **42**, 4-12.

**Otwinowski Z & Minor W (1997)**. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

**Otwinowski Z (1993)**. In *Data Collection and Processing*, eds. Sawyer, L., Isaacs, N. & Bailey, S. (SERC, Daresbury Laboratory, Warrington, U.K.), pp. 55-62.

**Oxtoby DW (1992)** Homogeneous nucleation: theory and experiment. *J. Phys.: Condens. Matter*, **4**, 7627-7650.

**Oxtoby DW (1998)**. Nucleation of first order phase transition. *Acc. Chem. Res.* **31**, 91-97.

**Palmer RA & Niwa H (2003)**. X-ray crystallographic studies of protein-ligand interactions. *Biochemical Society Transaction* **31**, 973–979.

**Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King J & Benedek G B (2001)**. Crystal cataracts: Human genetics cataracts caused by protein crystallization *Proc. Natl Acad. Sci.* **98**, 6116-6120.

**Pangburn** MK (1986). Immunobiology of the Complement System, pp. 45–62. Edited by G. D. Ross. New York: Academic Press.

**Pangburn** MK, Atkinson MAL & Meri S (1991). Localization of the heparin-binding site on complement factor H. *J. Biol. Chem.* **266**, 16847–16853.

**Peitsch** MC (1996). PROMOD and Swiss-Model: Internet based tools for automated comparative protein modeling. *Biochemical Society Transaction* **24**, 274-279.

**Perez** de la L, Harris JM, Hinchliffe CL, Holt SJ, Rushmere NK & Morgan BP (2000). Pigs express multiple forms of decay-accelerating factor (CD55), all of which contain only three short consensus repeats. *J Immunol* **165**, 2563–2573.

**Przybylska** M (1989). A double cell for controlling nucleation and growth of protein crystals. *J. Appl. Crystallogr.* **22**, 115–118.

**Ray** WJ & Bracker CE (1986). Polyethylene glycol: catalytic effect on the crystallization of phosphoglucomutase at high salt concentration. *J. Crystal Growth* **76**, 562-576.

**Rayment** I (2002). Small-Scale Batch Crystallization of Proteins Revisited: An Underutilized Way to Grow Large Protein Crystals, *Ways & Means. Structure* **10**, 147–151.

**Read** RJ (2001). Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Cryst.* **D57**, 1373-1382.

**Reid** KBM (1995). The complement system: A major effector mechanism in humoral immunity. *Immunologist* **3**, 206-211.

**Reid** KBM, Bentley DR, Campbell RD, Chung LP, Sim RB, Kristensen T & Tack BF (1986). Complement-system proteins which interact with C3b or C4b: a superfamily of structurally related proteins. *Immun. Today* **7**, 230-234.

**Reid** KBM & Day AJ (1989). Structure-function relationships of the complement components. *Immunol. Today* **10**, 177-180.

**Reynolds** DN, Keeling KL, Molestina R, Srisatjaluk R, Butterfield JH, Ehringer W, Justus DE & Kotwal G J (2000). Heparin binding activity of vaccinia virus complement control protein confers additional properties of uptake by mast cells and attachment to endothelial cells; in *Advances in animal virology* (ed.) S Jameel (Villarreal: Science Publishers), 337–342.

**Richardson** JS Richardson DC (2002). Natural beta-sheet proteins use negative design. *Proc Natl. Acad Sci.* **5**, 2754-2759

**Rosenberger** F, Howard SB, Sowers JW, Nyce TA (1993). Temperature dependence of protein solubility — determination and application to crystallization in X-ray capillaries. *J. Cryst. Growth.***129**, 1-12.

**Rosengard** AM, Alonso LC, Korb LC, Baldwin WM, Sanfilippo F, Turka LA & Ahearn JM (1999). Functional characterization of soluble and membrane-bound forms of vaccinia virus complement control protein (VCP). *Mol. Immunol.* **36**, 685–697.

**Rupp** A (2001). Fundamentals of crystallization. (Notes for CSUF workshop January 2001). [http://www.llnl.gov/CCW/Fundamentals\\_of\\_Crystallization.htm](http://www.llnl.gov/CCW/Fundamentals_of_Crystallization.htm)

**Rushmere** NK, Tomlinson S & Morgan BP (1997). Expression of rat CD59: functional analysis confirms lack of species selectivity and reveals that glycosylation is not required for function. *Immunology* **90**, 640–646.

**Sahu** A, Isaacs SN, Soulika AM & Lambris JD (1998). Interaction of vaccinia virus complement control protein with human complement proteins: factor I mediated degradation of C3b to iC3b1 inactivates the alternative complement pathway. *J. Immunol.* **160**, 5596–5604.

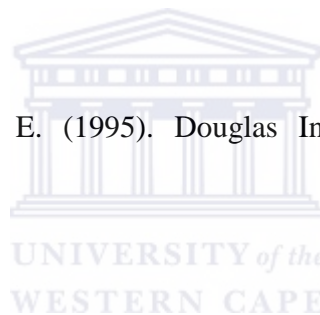
**Saridakis** E & Chayen NE (2000). Improving protein crystal quality by decoupling nucleation and growth in vapor diffusion. *Protein Science* **9**, 755-757.

**Saridakis** E, Dierks K, Moreno A, Dieckman MWM, Chayen NE (2002). Separating Nucleation and Growth in Protein Crystallization using Dynamic Light Scattering. *Acta Cryst.*, **D58**, 1597-1600.

**Saridakis** E, Shaw-Stewart PD, Lloyd LF & Blow DM (1994). Phase diagram and dilution experiments in the crystallization of carboxypeptidase G<sub>2</sub>. *Acta Crystallogr.* **D50**, 293–297.

**Schmitz** SK (1990). *An Introduction to Dynamic Light Scattering by Macromolecules*. New York Academic Press.

**Shaw-Stewart** PD & Conti E. (1995). Douglas Instruments Research Report 3, <http://douglas.co.uk/rep3.htm>



**Sica** F, Adinolfi S, Vitagliano L, Zagari A, Capasso S & Mazzarella L (1996). Cosolute effect on crystallization of two dinucleotide complexes of bovine seminal ribonuclease from concentrated salt solutions. *J. Crystal Growth* **168**,192-197.

**Smith** SA, Sreenivasan R, Krishnasamy G, Judge KW, Murthy KH, Arjunwadkar SJ, Pugh DR & Kotwal GJ (2003). Mapping of regions within the vaccinia virus complement control protein involved in dose-dependent binding to key complement components and heparin using surface plasmon resonance. *Biochim. Biophys. Acta.* **1650**, 30–39.

**Smith** SA, Mullin NP, Parkinson J, Shchelkunov SN, Totmentin AV, Loparev VN, Srisatjaluk R, Reynolds DN, Keeling KL & Justus DE (2000). Surface exposed conserved K/R-X-K/R sites and positive charge on poxviral complement control proteins contribute to heparin binding and to inhibition of molecular interactions with human endothelial cells: a novel mechanism for evasion of host defense. *J. Virol.* **74**, 5659–5666.

**Smith** SA, Krishnasamy G, Murthy KH, Cooper A, Bromek K, Barlow PN & Kotwal GJ (2002). Vaccinia virus complement control protein is monomeric, and retains structural and functional integrity after exposure to adverse conditions. *Biochim. Biophys. Acta* **1598**, 55–64.

**Spiller** OB & Morgan BP (1998). Antibody-independent activation of the classical complement pathway by cytomegalovirus-infected fibroblasts. *J Infect Dis* **178**, 1597–1603.

**Storoni** LC, McCoy AJ, Read RJ (2004). Likelihood-enhanced fast rotation functions. *Acta Cryst.* **D60**, 432-438.

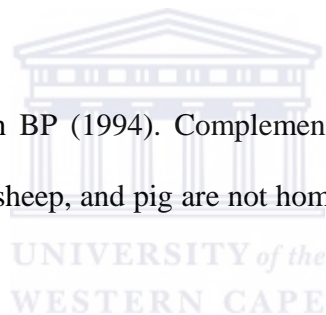
**ten Wolde** PR & Frenkel D (1997). Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations *Science* **277**, 1975-1978.

**Tschopp** JF, Brust PF, Cregg JM, Stillman CA & Gingeras TR (1987). Expression of the LacZ gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Res.* **15**, 3859-3876.

**Wilson** WW (1990). Monitoring crystallization experiments using the dynamic light scattering; assaying and monitoring protein crystallization in solution. *Methods of Comp Meth Enzymol.* **1**, 110-117.

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

**Van den Berg** CW & Morgan BP (1994). Complement-inhibiting activities of human CD59 and analogues from rat, sheep, and pig are not homologously restricted. *J Immunol* **152**, 4095-4101.



**Veenhuis** M, Van Dijken JP & Harder W (1983). The significance of peroxisomes in the metabolism of one-carbon compounds in yeast. *Adv. Microb. Physiol.* **24**, 1-82.

**Vekilov** PG, Feeling-Taylor AR, Petsev DN, Galkin O, Nagel R L & Hirsch RE (2002). Intermolecular Interactions, Nucleation, and Thermodynamics of Crystallization of Hemoglobin C. *Biophys. J.* **83**, 1147-1156.

**Vekilov** PG, Chernov AA (2002). In: Ehrenreich H, Spaepen F (Eds.), The Physics of Protein Crystallization, J. Solid State Physics, **57**, Academic Press, Amsterdam, 1-147.

**Welsh** RMJ, Cooper FCJ & Oldstone MB (1975). Human serum lyses RNA tumour viruses. Nature **275**, 612.

**Wiles** AP, Shaw G, Bright J, Perczel A, Campbell ID & Barlow PN. (1997). NMR studies of a viral protein that mimics the regulators of complement activation. J. Mol. Biol. **272**, 253-265.

**Yamamoto** I, Abe A. & Ishimoto M (1987). Properties of glutamate dehydrogenase purified from *Bacteroides fragilis*. J Biochem (Tokyo) **101**, 1391-1397.

**Yonath** A, Muessig J & Wittmann HG (1982). Parameters for crystal growth of ribosomal subunits, J Cell Biochem, **19**, 145-155.