PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
THE MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY DR. MOULAY TAHER. SAÏDA
FACULTY OF SCIENCES
DEPARTMENT: BIOLOGY

MEMORY

PRESENTED FOR OBTAINING THE MASTER'S DEGREE IN: BIOLOGY

SPECIALITY : Biochemistry and Cell Physiology

Presented by: Ould Ahmed Salem Abdellahi and **Khatir Mebarka**

Citric Acid Cycle Enzymes Research on the Function and Structure Relationship Bases

URL: http://bioinformaticstools.org/prjs/cacfms

Day: 03-07-2016

In front of the jury commission, composed by:

✓ Dr. Kefifa Abdelkrim	President
✓ Dr. Abdelkrim Barroukch	Examiner
✓ Dr. Abdelkrim Rachedi	Supervisor

2015/2016


بسم الله الرحمن الرحيم

Acknowledgment

We would like to thank everyone who has made his moral support to enrich our work.

We would like to express our special thanks of gratitude to our supervisor Dr. Abdelkrim Rachedi, who gave us the golden opportunity to do this wonderful project on the topic (Citric Acid Cycle Enzymes Research on the Function and Structure Relationship Bases), which also helped us in doing ~~a~~ lots of Research and we came to know about so many new things in the subject structural bioinformatics. We are really thankful to have acquired them.

Our thanks also go to the jury members for the great honor they gave us to examine this work.

We would also like to thank our parents and friends who helped us a lot in finalizing this project within the limited time frame.

Dedication

It's only with the help of God almighty who guided me the right way and my parents' blessings that I was able to complete this modest work which I would like to dedicate to:

- *My dear parents for their sacrifices and encouragement throughout the course of this study.*
- *My very dear brother: Dr. Ould Ahmed Salem Ahmedou.*
- *My very dear sisters: Fatima and Hind.*

And lastly to my partner in this work Khatir Mebarka and my dear colleagues: Bouha, Dah, Ely, Aicha, Mira and to my promotion.

Ould Ahmed Salem

Dedication

All praise to Allah, today we fold the days' tiredness and the errand summing up between the cover of this humble work.

To the utmost knowledge lighthouse, to our greatest and most honored prophet Mohamed - May peace and grace from Allah be upon him

To the spring that never stops giving, to my mother who weaves my happiness with strings from her merciful heart to my mother Zohra

To whom he strives to bless comfort and welfare and never stints what he owns to push me in the success way who taught me to promote life stairs wisely and patiently, to my dearest father.

To those who have demonstrated to me what is the most beautiful of my brothers and sisters: Mustapha; Dine; Hicham; Ahmed; Fatima; Houria; Bouchra(bibich); Mohamed; Zohra; Nafissa.

My dedication is also extended to the pure hearts and innocent souls in my life; Omar, Hadia, Taha, Said, Tidjani, Yassin, kawthar, Marwa, soundos, Rihab, Anfal, Maraim, Bouchra Allae.

My dedication also goes to the spirit that had inhabited my soul, my fiancé tayyib Abdou illah.

To the one who worked with me so hard in order to complete this work, to my friend and partner: Ould Ahmed Salem Abdellahi.

To friends: Zahra, Taymoucha, Halima, Mira, Aicha, Asma, Amina, Fouzia, Abdou, Houcine and to every student in the second year master of Biochemistry.

Mebarka

Summary

Chapter I: Literature Review

Summary

Summary

Summary

List of Figures

List of Figures

List of Figures

List of Figures

List of Tables

List of Tables

List of Abbreviations

AA: Amino acid.

AKG: 2-Oxoglutaric Acid.

ATH: 4-Hydroxy-Aconitate Ion.

BMRB: Biological Magnetic Resonance Bank.

CAC: Citric Acid Cycle metabolism pathway.

CacSFMs: Citric Acid Cycle Binding Structural and Functional Motifs.

CATH: Class Architecture Topology and Homology.

CIT: Citric Acid.

Cryo-EM: Cryo-Electron Microscopy.

FAD: Flavin Adenine Dinucleotide.

FLC: Citrate Anion.

FUM: Fumaric Acid.

GDP: Guanosine Diphosphate.

GTP: Guanosine Triphosphate.

H: Alpha-Helix.

ICT: Isocitric Acid.

L: Loop.

MLT: Malate Ion.

NAD: Nicotinamide Adenine Dinucleotide.

NMR: Nuclear Magnetic Resonance spectroscopy.

PDB: Protein Data Bank.

PDBe: Protein Data Bank Europe.

PDBJ: Protein Data Bank Japan.

RCSB: Research Collaboratory for Structural Bioinformatics.

S: Beta-Strand.

List of Abbreviations

SCA: Succinyl-Coenzyme A.

SCOP: Structural Classification of Proteins.

SDX: S-Citryldethia Coenzyme A.

SIN: Succinic Acid.

TEM: Transmission Electron Microscopy.

TRA: Aconitate Ion.

W: Water.

3D: Three-dimensional.

Abstract

Metabolism is important for cell survival which is undertaken by specific enzymes. Healthy cells usually have normal metabolism pathways while sick cells show abnormal pathways. Understanding the function of enzymes responsible for metabolism pathways is one way to better understand metabolic related illnesses.

As it has accumulated by a vast number of research studies in the field of structural biology and biology in general, that enzymes 3-dimensional (3D) structure dictates the function they do or, in other words, the function of enzymes necessitate a certain 3D structure where active site residues are exposed into favorable spatial environment and electro-chemical conditions that allow for the carrying out of the reactions needed in their biological function.

In this work, enzymes involved in the Citric Acid Cycle (CAC) metabolism pathway have been selected for undertaking a structural-bioinformatics study to try and discover some of the underlying bases behind the structure and function relationship necessary in their biological function.

It was necessary to use spatial structures of proteins from the international database known as the Protein Data Bank or PDB which stores and distributes structural data for proteins and enzymes in complex forms with substrates or ligands.

The study resulted in the **detection**, **identification** and **description** of a group of structural elements, termed in this study as Structural and Functional Motifs, which are deemed important in figuring out the bases behind the structure and function relationship in the CAC enzymes.

Keywords:

Proteins, Enzymes, Structure, Function, Krebs Cycle, Ligands, Structural and Functional Motifs, PDB, Databases, Structural Bioinformatics.

Abstract

الملخص

عملية الأيض مهمة لبقاء الخلية والتي يتدخل في إنجازها إنزيمات معينة متخصصة، حيث تسلك الخلايا السليمة عموما مسارات أيضية طبيعية في حين أن الخلايا المريضة قد تبدي مسارات غير طبيعية. إن السعي لفهم وظيفة الإنزيمات المسؤولة عن الأيض تمكن من التعامل الأفضل مع الأمراض المتعلقة بالإختلالات الأيضية.

كما أثبتت مجموعة من الدراسات والأبحاث في مجال البيولوجيات الهيكلية والبيولويحيا عموما أن التركيب ثلاثي الأبعاد للإنزيمات هو الهيكل الذي يتحكم في الوظيفة التي تقوم بها الإنزيمات والبروتينات على العموم أو بعبارة أخرى، فإن وظيفة الإنزيمات تتطلب بنية ثلاثية الأبعاد معينة حيث يتم استعراض الأحماض الأمينية الفعالة أو النشطة في مواقع مواتية تسمح بتحقيق التفاعلات اللازمة لوظائفها البيولوجية.

في هذا البحث تمت دراسة الإنزيمات المشاركة في حلقة كريبس أو دورة حمض ألستريك بهدف استكشاف الأسس الحيوية الكامنة وراء وظائفها البيولوجية وذلك باعتماد طرق البحث المعتمدة على المعلوماتية الحيوية الهيكلية.

كان من الضروري استخدام البنية الفراغية للبروتينات محملة من قاعدة البيانات الدولية المعروفة باسم بنك معلومات البروتين (PDB) الذي يخزن ويوزع المعلومات حول التراكيب الفراغية للبروتينات والإنزيمات في أشكال معقدة متفاعلة مع ركائز معينة أو بما يسمى بـالليجندات (Ligands).

وأسفرت الدراسة عن **الكشف وتحديد ووصف** مجموعة من العناصر الهيكلية (وتسمى هنا بالوحدات الهيكلية الوظيفية) التي تعتبر حجر الزاوية في الأسس وراء العلاقة بين التركيب الفراغي والوظيفة لدى الإنزيمات المسؤولة عن جملة التفاعلات المشكلة لحلقة كريبس.

الكلمات المفتاحية:

البروتين، التركيب الفراغي, الوظيفة، حلقة كريبس، الليجندات، الوحدات الهيكلية الوظيفية، قواعد البيانات، المعلوماتية الحيوية الهيكلية.

Abstract

Résumé

Le métabolisme est important pour la survie de la cellule et doit être réalisé par des enzymes spécifiques. Les cellules saines ont généralement des voies métaboliques normales tandis que les cellules malades montrent des voies anormales. La vision adopte par cette étude considère que pour comprendre les maladies liées au mauvais métabolisme, il faut d'abord une compréhension de la relation entre la fonction et la structure tridimensionnelle des enzymes responsables des voies du métabolisme.

Comme a été décrit par les études et les recherches dans le domaine de la biologie structurale et la biologie en général, que la structure tridimensionnelle des enzymes dictat leur fonction ou autrement dit, la fonction des enzymes nécessite une structure tridimensionnelle où les résidus du site actif sont exposés à des conditions favorables à la réalisation des réactions nécessaires pour leur fonction biologique.

Dans ce travail, les enzymes impliquées dans la voie du métabolisme du cycle de l'acide citrique (cycle de Krebs) ont été sélectionnées pour une étude utilisant les principes de la bioinformatique structurale pour essayer de découvrir les bases derrière la relation entre la structure et la fonction chez les enzymes et les protéines en général.

Il était nécessaire d'utiliser des structures spatiales des protéines à partir de la base de données internationale connue sous le nom « Protein Data Bank » ou PDB qui stocke et distribue la structure tridimensionnelle des protéines et des enzymes en formes de complexes avec des substrats ou ligands.

L'étude a abouti à la **détection, l'identification et la caractérisation** d'un groupe d'éléments structuraux (appelé ici comme les Motifs Structurales et Fonctionnelles) qui sont jugés dans ce travail comme importants vers la compréhension des bases derrière la relation « Structure et Fonction » des enzymes du cycle de Krebs.

Mots clés :

Protéines, enzymes, structure, fonction, cycle de Krebs, Ligands, Motifs Structurels et Fonctionnelles, PDB, Bases de Données, Bioinformatique Structurale.

General Introduction

General Introduction

Proteins are the most diverse biomolecules on Earth, performing many functions required for life. Protein functions are so diverse because of the many unique three-dimensional structures protein polymers form. Despite such variety, proteins also share several specific structural characteristics in their monomers.

The Ligands play an important role in several processes biological and make interactions with the proteins; this complex are stored in a base Global data (PDB) and regards the 3-dimensional structure of macromolecules to share an information on these structures with the scientific and international community.

The objective of this study is to analyze the binding site between the ligand and protein available in the database for identify, define and characterize the different structural units which can exist in the protein structures and the frequency of amino acids in each ligand studied.

This project is divided into three chapters:

- **The first chapter:** Literature Review

  It talks about the concepts of protein and their structural levels as well as about the Krebs cycle (Definition and Steps), PDB (Protein data Bank that stores an information on 3D structures of macromolecules and the end The other protein structures of banks (Scop) and (Cath).

- **The second chapter:** material and methods

  The tools and material used are presented.

- **In the third chapter:** presents result and discussion

  It talks about the results obtained in this study, their discussion and some important conclusions.

- And the end the work is completed by a **"General Conclusion"**.

Chapter I: Literature Review

I. Generality on proteins

A protein is a functional biological molecule that is made up of one or more polypeptides that are folded, coiled into a specific structure. Proteins are important macromolecules that serve as structural elements, transportation channels, signal receptors and transmitters, and enzymes. Proteins are linear polymers that are built up of the monomer units called amino acids. There are 20 different amino acids **(see Annex I)** and they are connected by a peptide bond between the carboxyl group and the amino group in a linear chain called a polypeptide **(Benjamin Cummings, 2008)**.

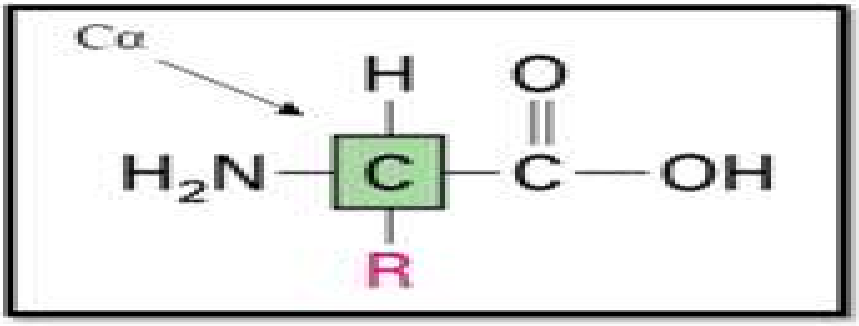


Figure 01: Configuration of Amino Acid.

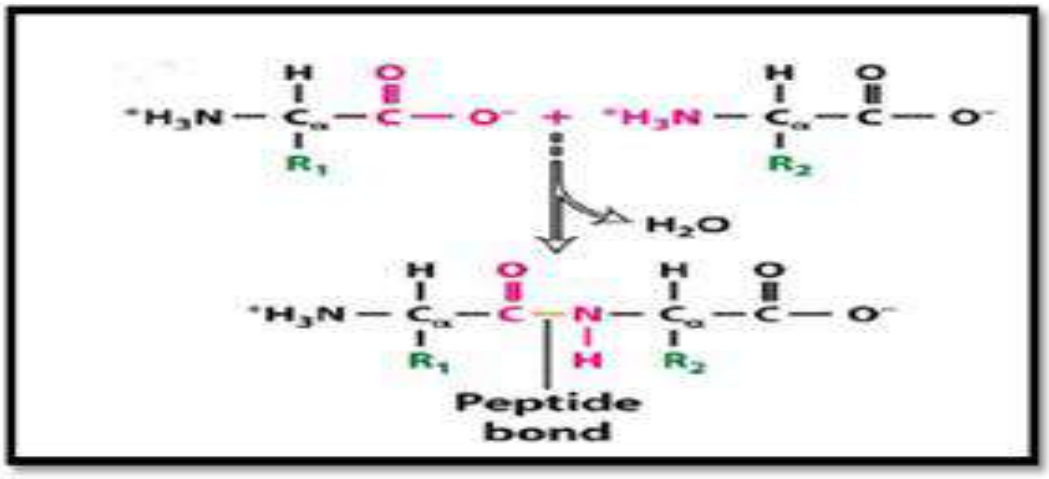


Figure 02: Peptide bonds connect amino acid into linear chains.

II. Levels of protein structure

There are four levels of protein structure:

II. 1. Primary Structure

The primary structure of a protein is the level of protein structure which refers to the specific sequence of amino acids. When two amino acids are in such a position that the carboxyl groups of each amino acid are adjacent to each other, they can be combined by undergoing a dehydration reaction which results in the formation of a peptide bond **(Benjamin Cummings, 2008)**.

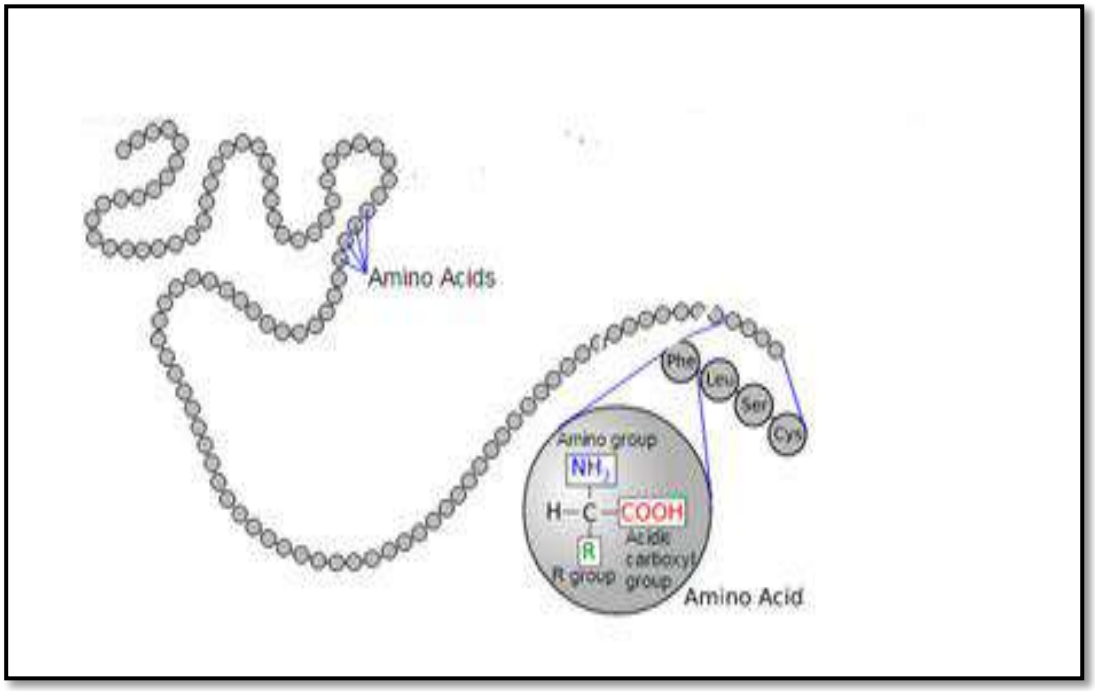


Figure 3: Primary structure of protein.

II. 2. Secondary structure

The three important secondary structures are α-helix, β-sheets, and β-turns. Also, the beta sheets can be parallel, antiparallel, beta sheets are more stable because the hydrogen bonds are at a nighty degree angles **(Perticaroli *S and al, 2013*)**.

II. 2. 1. Types of secondary structure (figure 2)

II. 2. 1. A. The alpha-Helix

Alpha (α) has a right-handed spiral conformation, in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues before it in the sequence.

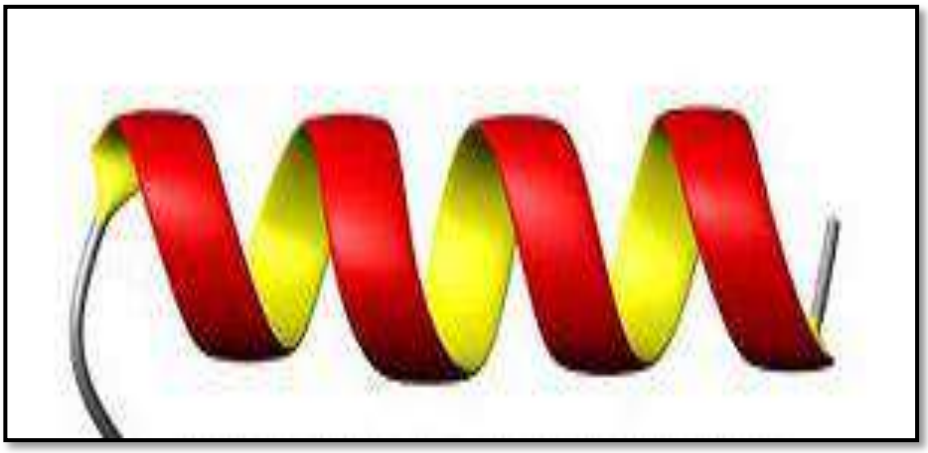


Figure 4: Alpha-Helix.

II. 2. 1. B. Beta-Sheet

Is a stretch of polypeptide chain, typically 3 to 10 amino acids long, with its backbone in an almost fully extended conformation, two or more parallel or antiparallel adjacent polypeptide chains of beta strand stabilized by hydrogen bonds form a beta-Sheet. For example, the proteins in silk have a beta-sheet structure.

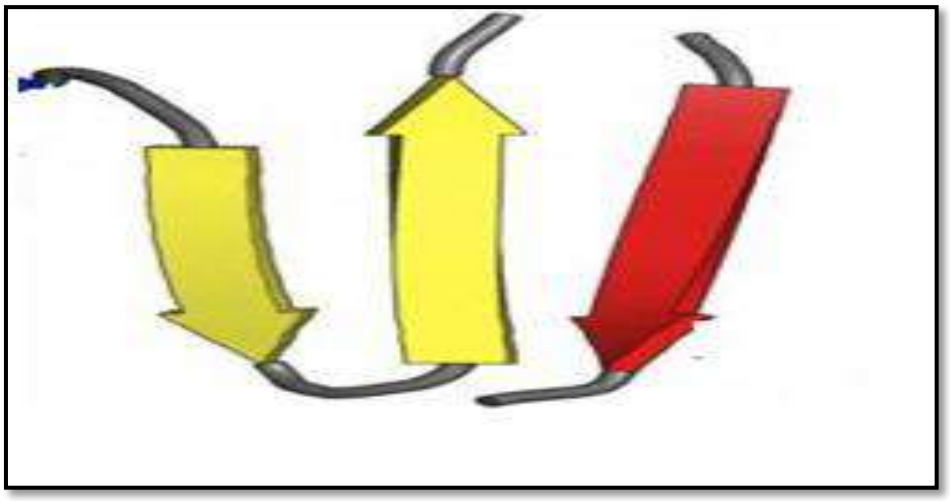


Figure 5: Beta-Sheet.

II. 2. 1. C. Turns and Loops

Can be defined by the close approach of two Cα atoms (<7A°) in a stretch of residues not folded into a common secondary structure. **Figure 3**: EGF (PDB-entry 1ivo) colored by element. The names of the different loops and side chains of Cys are shown.

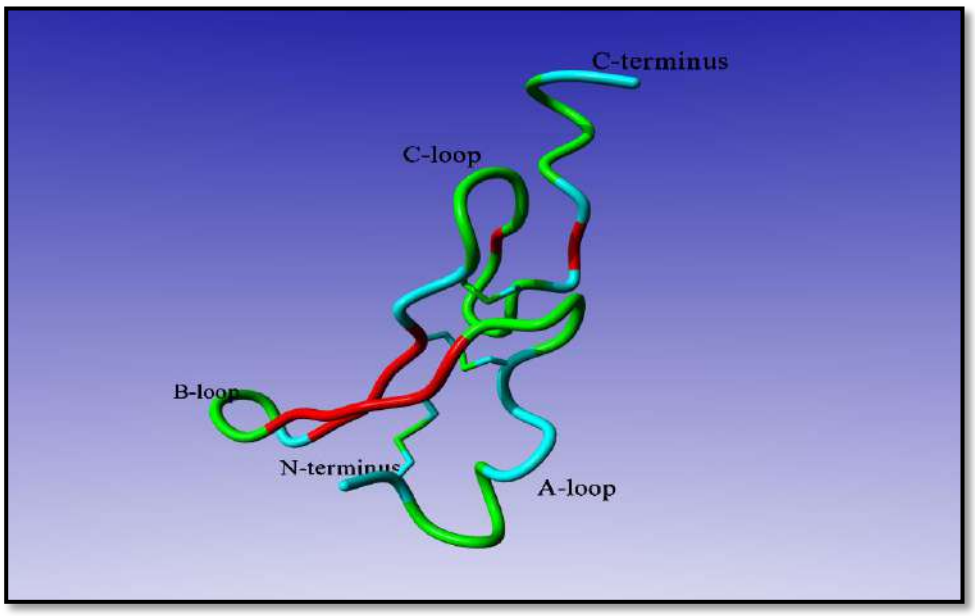


Figure 6: Turns and Loops.

II. 2. 2. Ramachandran Plot

Ramachandran plot a two-dimensional plot of the values of the backbone torsion angles phi and psi, with allowed regions indicated for conformations where there is no steric interference **(Deane C-M and al, 1999)**.

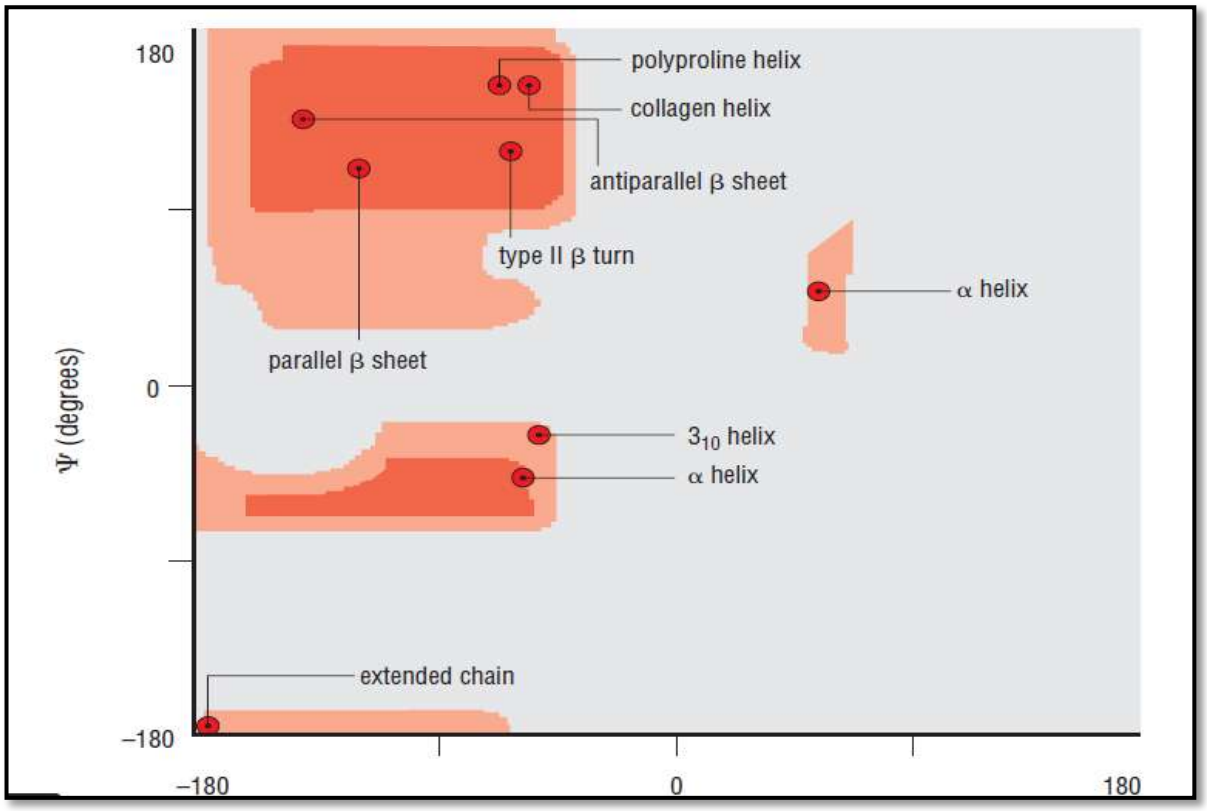


Figure 7: Importance and Determinants of Secondary Structure (Ramachandran plot).

II. 3.Tertiary structure

As the secondary structure becomes established due to the primary structure, a polypeptide folds and refolds upon itself to assume a complex three-dimensional shape called the protein tertiary structure. Tertiary structure is the overall shape of a polypeptide **(Benjamin Cummings, 2008)**.

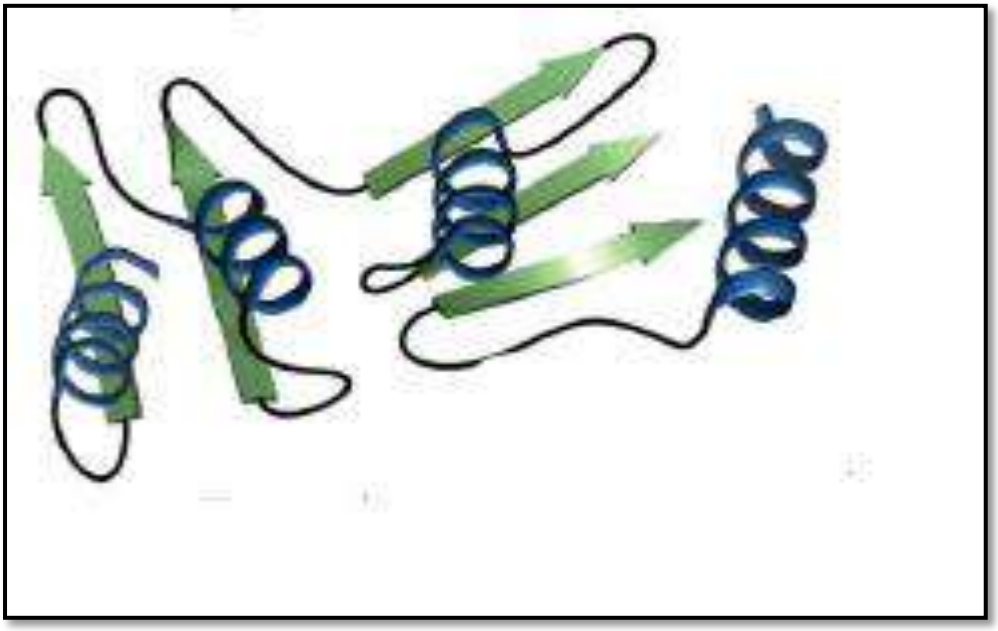


Figure 8: Tertiary structure.

II. 4. Quaternary Structure

While all proteins contain primary, secondary and tertiary structures, quaternary structures are reserved for proteins composed of two or more polypeptide chains.

Quaternary structures can also define as when more than one protein comes together to create either a dimer, trimer, tetramer, etc...**(Hector Viadiu, 2011)**.

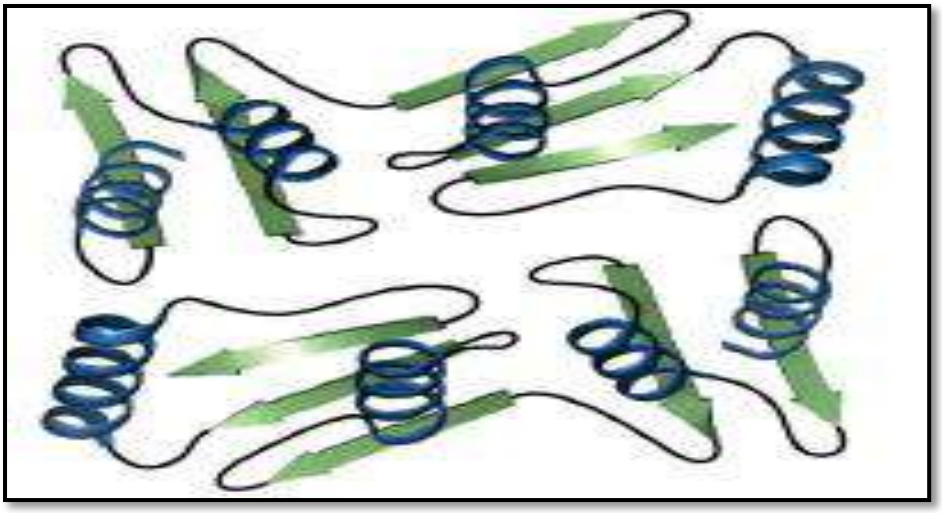


Figure 9: Quaternary Structure.

III. Structural Motifs in proteins

III. 1. Local Structural Motifs (Helices, Beta Strands & Sheets, Turns & Loops)

III. 1. A. alpha helices

Conformation	Phi	Psi	Omega
Alpha helix	-57	-47	180
3-10 helix	-49	-26	180
Pi-helix	57	-70	180

Table 1: Average conformational parameters of the most commonly found helical secondary structure elements. Types of alpha helices.

III. 1. B. Beta-Strands & Sheets

β-Sheets are structural configurations made of more than one β-stand held up by a number forces of which hydrogen bonds play major role. There are many types of β-Sheets:

- ✓ Parallel Beta-Sheet.
- ✓ Anti Parallel Beta-Sheet.
- ✓ Beta Barrel motif

- **Parallel beta-Sheet**: a beta sheet, formed from noncontiguous regions of the polypeptide chain, in which every strand runs in the same direction.

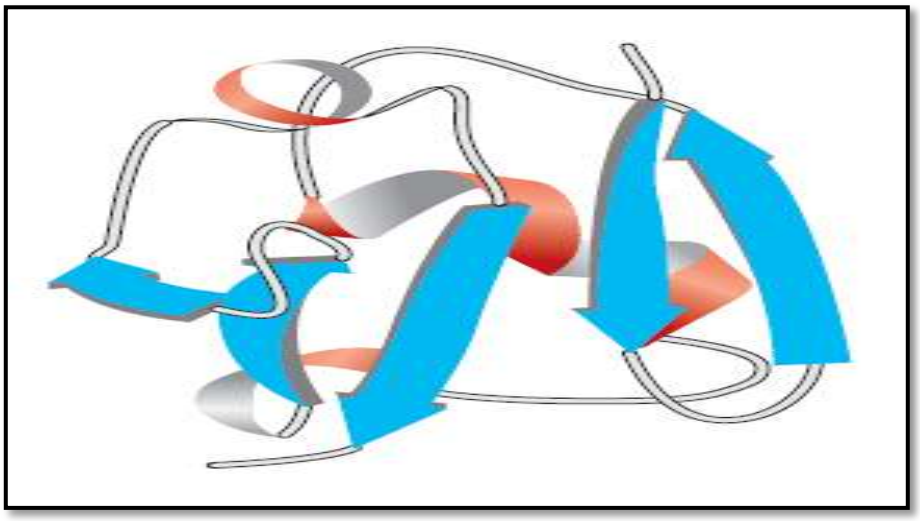


Figure 10: Parallel beta-Sheet.

- **Anti parallel beta-Sheet**: A beta sheet often formed from contiguous regions of the polypeptide chain, in which each strand runs in the opposite direction from its immediate neighbors **(Gellman S-H, 1998)**.

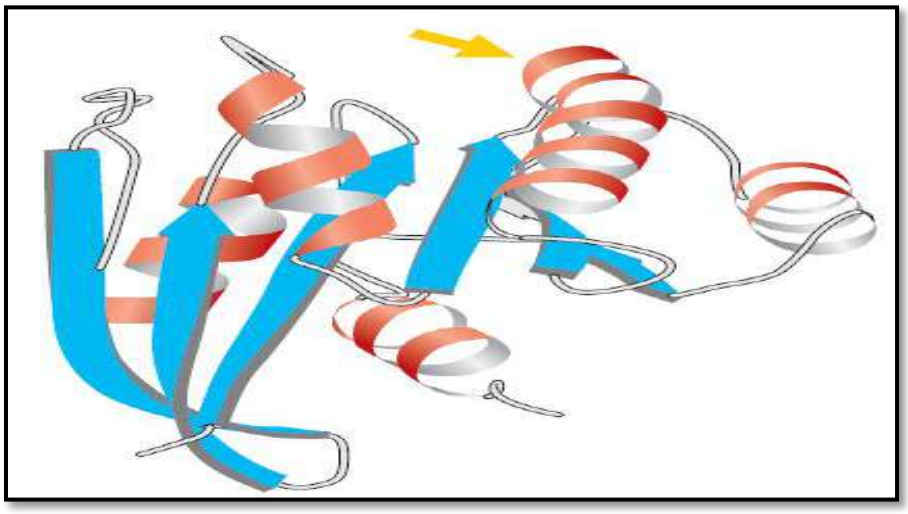


Figure 11: Anti parallel beta-Sheet in blue.

- **Beta Barrel motif:** A beta sheet in which the last strand is hydrogen bonded to the first strand, forming a closed cylinder **(Gellman S-H, 1998)**.

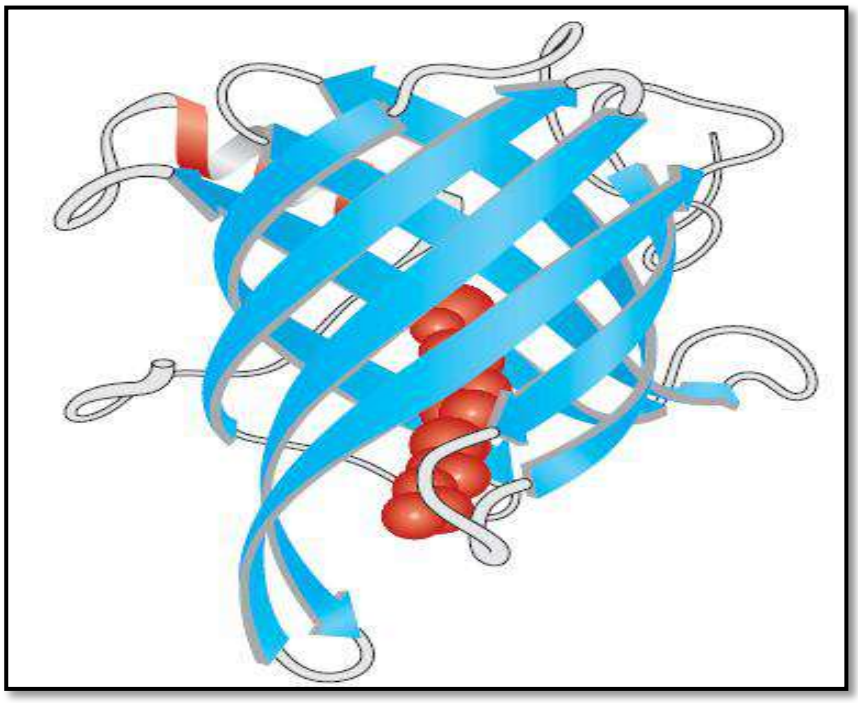


Figure 12: Beta Barrel motif.

III. 1. C. Turns

Turns are specific types of loops as they can be structurally described and classified. In general they can be distinct from loose loops in the cases when the closest Cα atoms are about 7A° apart **(Némethy, George, 1972)**.

The following summarize the existing types of turns:

- ✓ α – Turn: a hydrogen bond(s) formed between residues (amino acids) are separated by four residues (x → x+4).
- ✓ B – Turn: a hydrogen bond(s) between residues (amino acids) are separated by four residues (x → x+3).
- ✓ γ – Turn: a hydrogen bond(s) between residues (amino acids) are separated by four residues (x → x+2).
- ✓ π – Turn: a hydrogen bond(s) between residues (amino acids) are separated by four residues (x → x+5).
- ✓ Beta Bulges: Beta bulge loops are commonly occurring motifs in proteins and polypeptides consisting of five to six amino acids **(Milner-White, EJ, 1987)**.

III. 2. Functional Motifs

Functional motifs are structural motifs that are usually associated with specific and definite biological function **(Rachedi A, 2013, Tyson and al, 2010)**.

It is sequence or structural motif that is always associated with a particular biochemical function. An example is zinc finger motif **(Figure 13)** and the helix-turn-helix **(Figure 14) (Aitken A, 1999)**.

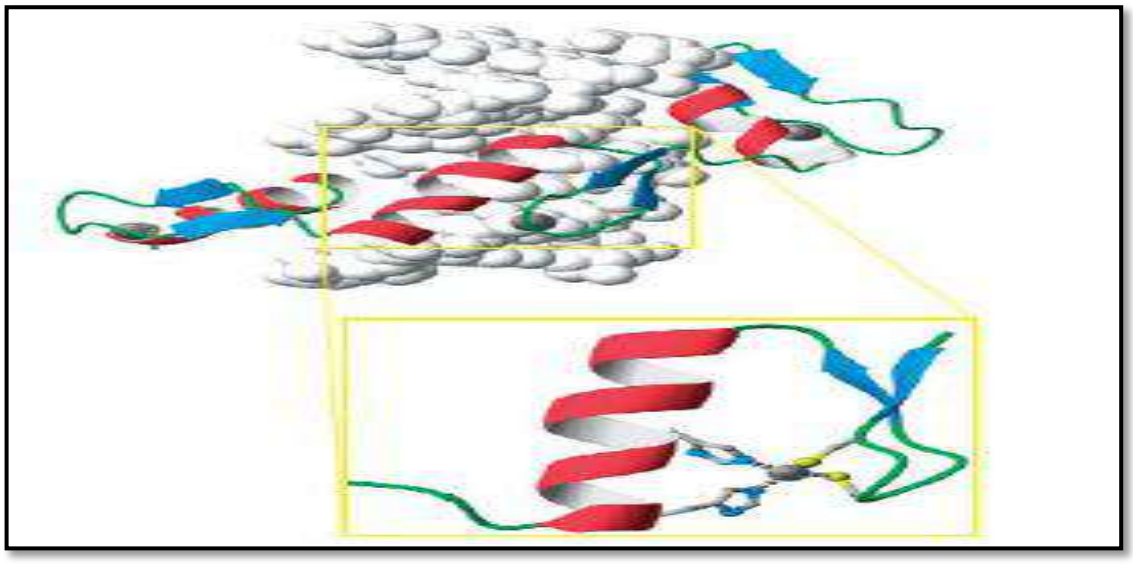


Figure 13: Zinc finger motif a fragment derive from a mouse gene regulatory protein is shown, with three zinc fingers bound spirally in the major groove of a DNA molecule. The inset shows the coordination of a zinc atom by characteristically spaced cysteine and histidine residues in a single zinc finger motif. The image is of Zif268. (PDB 1aay).

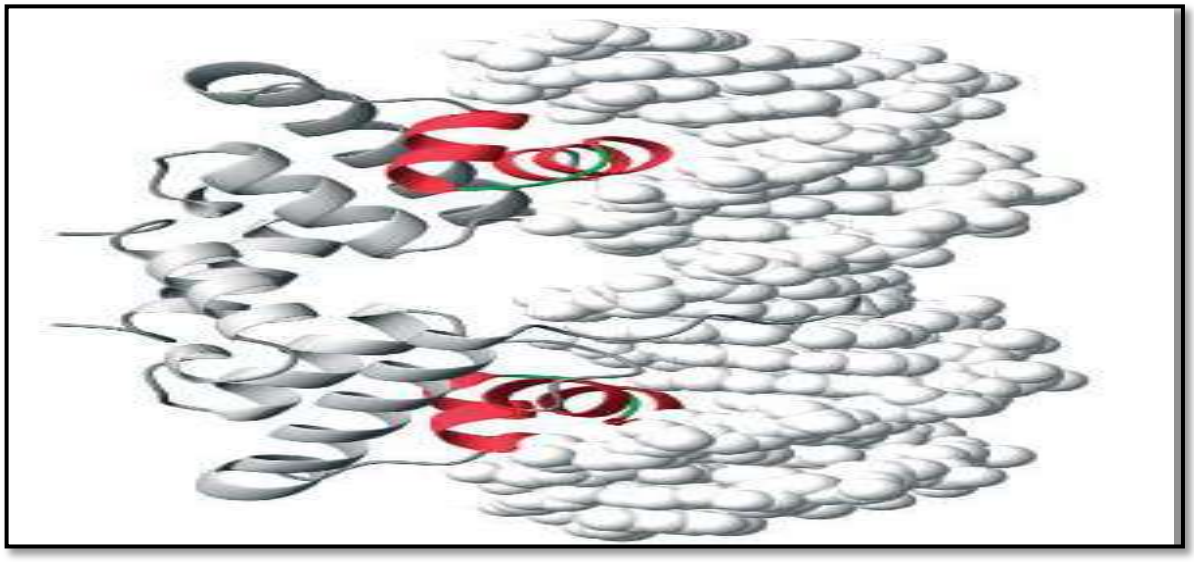


Figure 14: Helix-turn-helix The DNA-binding domain of the bacterial gene regulatory protein lambda repressor, with the two helix-turn-helix motifs shown in color. The helices closest to the DNA are the reading or recognition helices, which bind in the major groove and recognize specific gene regulatory sequences in the DNA. (PDB 1lmb)

IV. Experimental methods for determining protein structures

IV. 1. X-ray crystallography

IV. 1. 1. Definition

X-ray crystallography reveals the spatial structure of molecules by measuring how they scatter X-ray radiation when arranged in a crystal lattice. Two broad fields may be distinguished: small molecule crystallography deals with a small number of atom positions and typically well-ordered crystals, while macromolecular (usually protein) crystallography determines a much larger number of atomic positions, usually despite considerable crystalline disorder (**Drenth J, 1994)**.

Their common goal is to calculate the electron density distribution in the crystal from measured X-ray diffraction intensities. The electron density and diffraction intensities are represented by mathematical functions, which may be interconverted variously using the transform, convolution and complex product operations if the phases of the diffracted X-rays are known. Because the phases are not directly measurable, the dilemma known as the "phase problem" arises. Several methods exist that enable phase estimate (**Rhodes G, 2000)**.

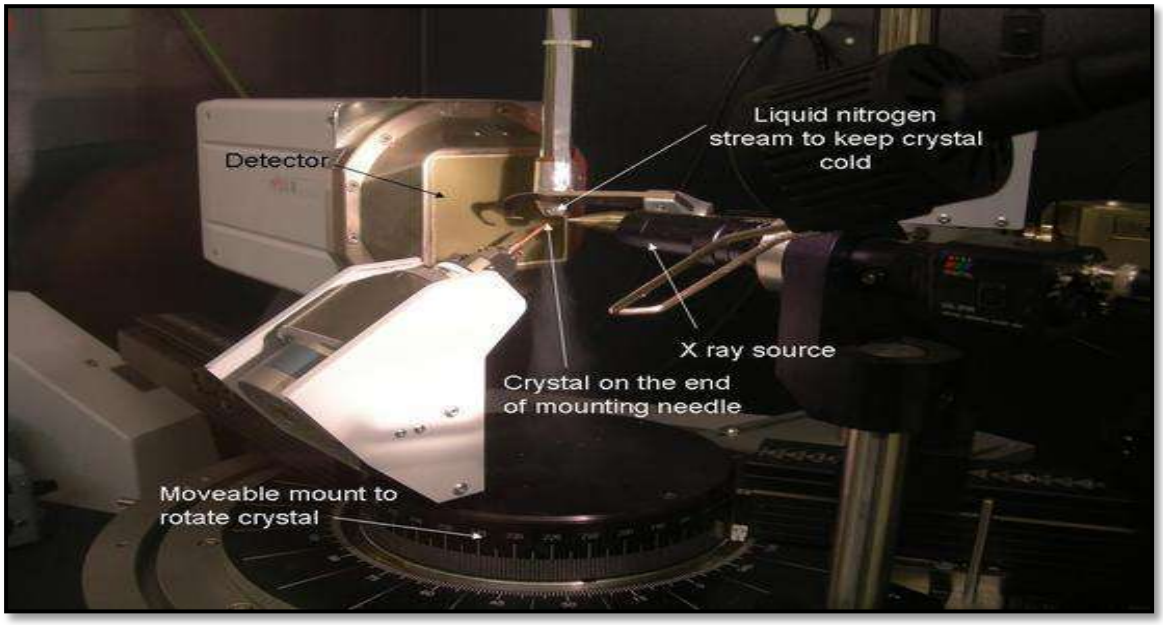


Figure 15: View of the entire machine and (right) a crystal mounted on a goniometer shown with the x-ray generator and detector **(Drenth J, 2007)**.

IV. 1. 2. Principle

X-ray crystallography has long been a vital method for studying the structure of proteins and other macromolecules. As the importance of proteins continues to grow, in fields from biochemistry and biophysics to pharmaceutical development and biotechnology, many researchers have found that knowledge of X-ray diffraction is an indispensable tool. In this new edition of his essential work, Dr. Jan Drenth, recognized internationally for his numerous contributions to crystallographic research, has provided an up-to-date and technically rigorous introduction to the subject.

Principles of Protein X-ray Crystallography provide the theoretical background necessary to understand how the structure of proteins is determined at atomic resolution. It is intended to serve as an introduction for graduate students, postdoctoral researchers, and established scientists who want to use protein crystallography in their own endeavors, or need to understand the subject in order to critically evaluate the literature. New additions to the book include a section on twinning, an additional chapter on crystal growth and a discussion of single-wavelength anomalous dispersion (SAD) **(Drenth J, 1994).**

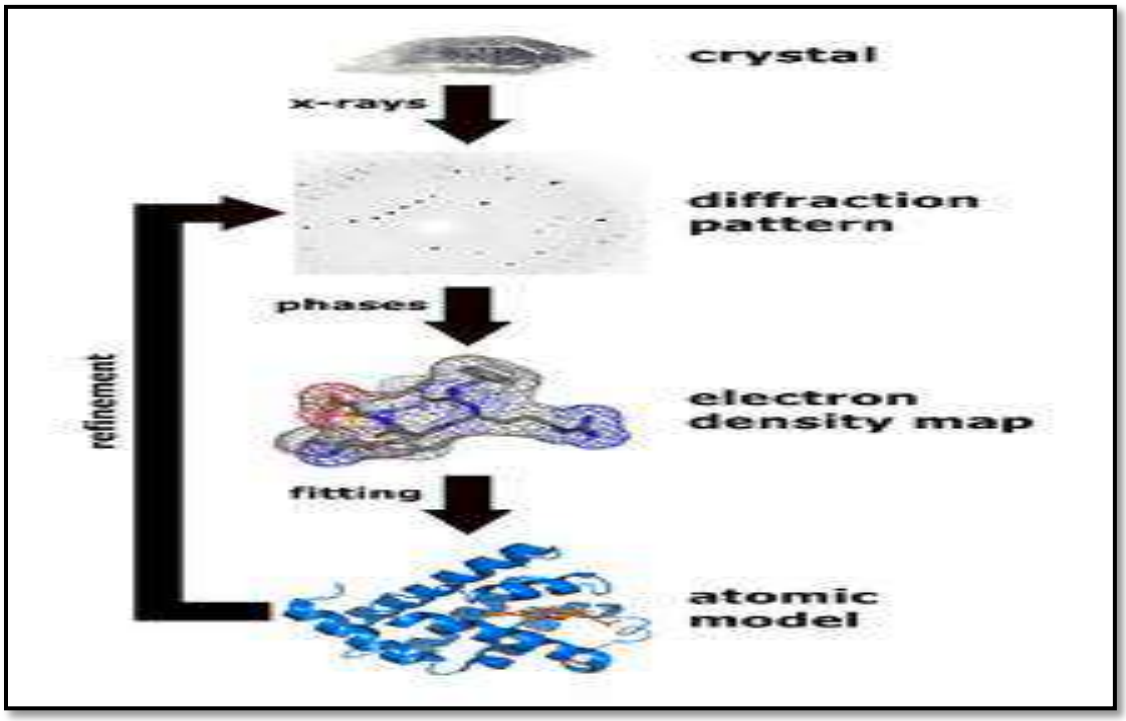


Figure 16: Workflow for solving the structure of a molecule by X-ray crystallography.

IV. 2. Nuclear magnetic resonance (NMR)

IV. 2. 1. Definition

Nuclear Magnetic Resonance (NMR) spectroscopy is an analytical chemistry technique used in quality control and research for determining content and purity of a sample as well as its molecular structure. For example, NMR can quantitatively analyze mixtures containing known compounds. For unknown com pounds, NMR can either be used to match against spectral libraries or to infer the basic structure directly. Once the basic structure is known, NMR can be used to determine molecular conformation in solution as well as studying physical properties at the molecular level such as conformation in solution as well as studying physical properties at the molecular level such as conformational exchange, phase changes, solubility, and diffusion. In order to achieve the desired results, a variety of NMR techniques are available. The basics of NMR are described here. **(Zeroka D, Hameka H-F, 1966)**.

Figure 17: High Field NMR.

IV. 2. 2. Principle

The principle of NMR imaging. The inductive sensing method of NMR signals applies to macroscopic samples. In homogeneous field intensity, all the nuclear spins located in the same environment resonate at the same Larmor frequency, regardless of their spatial location in the sample. To produce an image, that is to say locate the resonating nuclei in a sample, it is necessary to place the not homogeneous field in a sample, but in an inhomogeneous field, so that each ring may resonate at a frequency dependent on its position in the échantillon22. Practically, a field gradient is added to the main field in three orthogonal directions in order to locate spins in space. The spatial resolution is mainly determined by the intensity of the gradient, that is to say the amplitude of the magnetic field variation as a function of distance. Indeed, the variation in frequency between two points is directly proportional to this gradient: In its simplest version, just gradually change the frequency of detection to observe a particular region of space. However, this process is too long for practical applications, particularly in medicine. On modern imaging, the gradient fields are pulsed and processed by Fourier transform signals, in order to accelerate the time measurements **(Fernande D. Rochon Al, 2004)**.

IV. 3. Cryo-electron microscopy (Cryo-EM)

IV. 3. 1. Definition

Single-particle Cryo-electron microscopy (Cryo-EM), is an increasingly popular technique used by structural biologists to solve structures at atomic resolution. This technique complements x-ray crystallography because it reveals structural details without the need for a crystalline specimen. Through the examination of a frozen-hydrated specimen in vitreous (non-crystalline) ice, the specimen ultra structure, buffer and ligand distribution from its native state are maintained. Cryo-EM also complements structural studies using nuclear magnetic resonance (NMR) in that it enables the study of specimens larger than 150 kDa. Structural biologists frequently use Cryo-EM to study viruses, small organelles, and macromolecular biological complexes, as well as molecular interactions in supramolecular assemblies or machines **(Hoenger A, Boucher-Marquis C, 2011).**

Chapter I: Literature Review

During Cryo-EM, a transmission electron microscope (TEM) is used to record high resolution images of thousands to hundreds of thousands identical, but randomly oriented, particles (molecules) from each specimen. These images are then grouped, aligned, and averaged with image classification algorithms to distinguish between multiple orientations of the 3D molecule. With a well-behaved sample, Cryo-EM can solve molecular structures with a resolution approaching 2 Å; a resolution level that was inconceivable only a few years ago.

Figure 18: Machine of the Cryo- EM.

IV. 3. 2. Principle

Cryo-EM is based on transmission electron microscopy (TEM), including the basic steps of sample preparation, imaging, image processing, and structure analysis **(Figure19)**. In TEM, electrons generated by an electron gun are accelerated to very high energies in the high vacuum of the microscope column. Because high-speed electrons are deflected by magnetic fields, a series of electromagnetic lenses condense the electrons and then focus them on the specimen. The transmitted electrons are detected and form recorded images that are magnified by factors of 103/105. Computer programs then solve the detailed structure of the sample from the magnified images.

Biological samples present several technical challenges for TEM:

The high vacuum is not compatible with hydrated samples. - biological samples are mainly composed of light elements that are vulnerable to damage from high- energy electron- Light elements interact with electrons weakly, lowering image contrast **(Frank J, 2009)**.

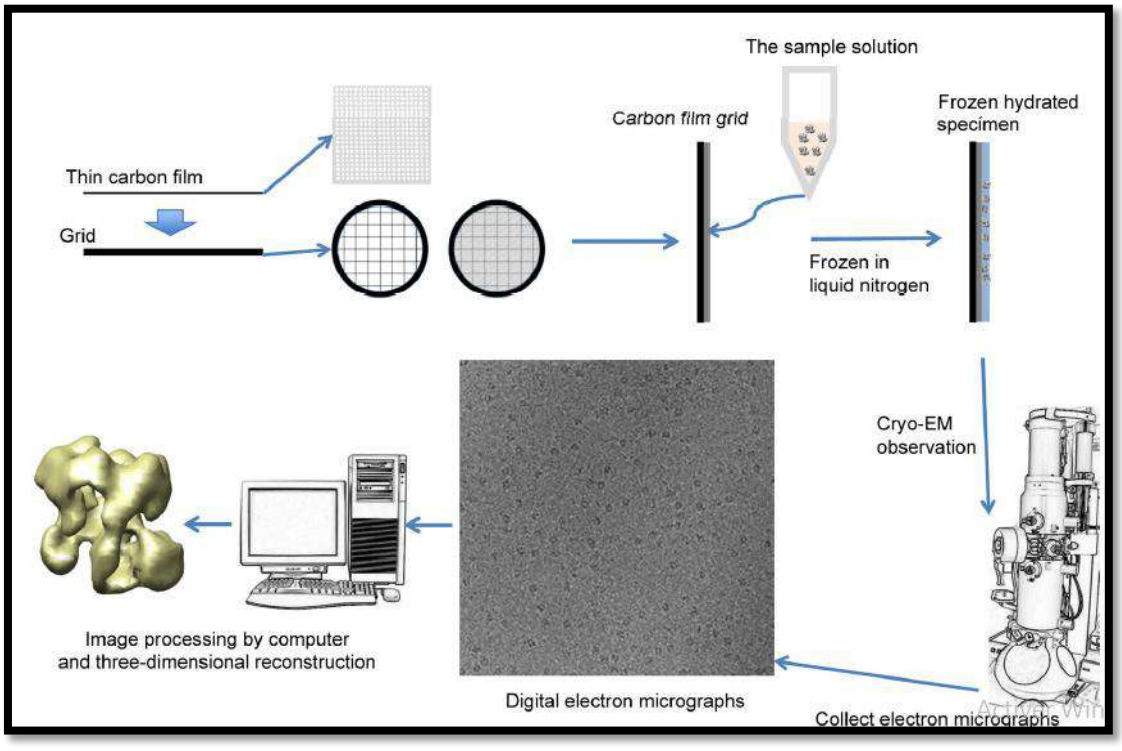


Figure 19: Workflow of Cryo-EM structure determination.

V. Protein data Bank (PDB)

V. 1. What's the PDB

Biologists and biochimistes use sequence databases, structure databases, literature databases, etc. The databases we will learn here is called the Protein databases (PDB). The PDB has all known 3D structures of proteins, DNAs and RNAs. To find the PDB on the web, type 'PDB' into Google, and go to the first link returned, which is:

http://www.rcsb.org/pdb/home/home.do

You need to download the Protein structures (i.e., the PDB files) that you are going to study, to your own computer. Each structure is in a PDB file with a name that does not carry much information. A PDB file is a simple text file with the(x; y; z) coordinates of all the atoms in the Protein **(Berman H-M and all, 2000).**

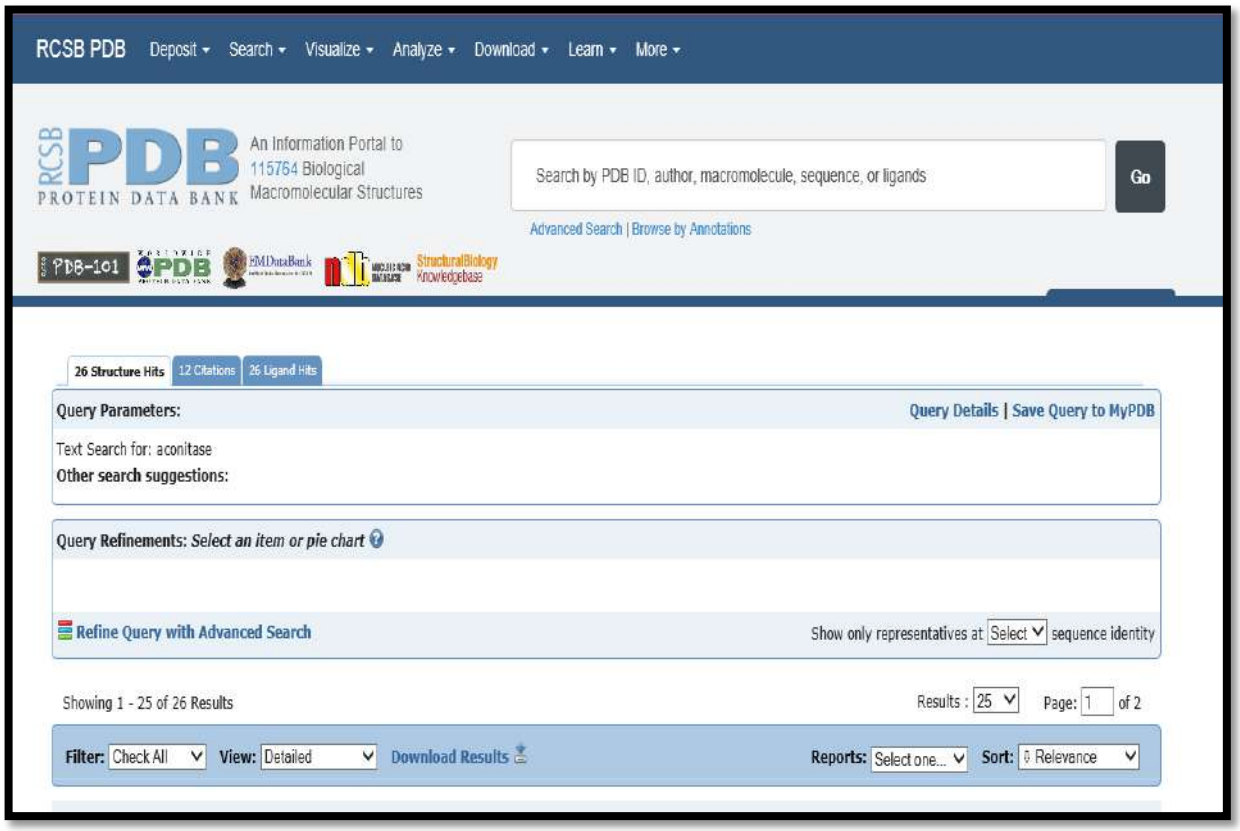


Figure 20: PDB – Protein Data Bank Page.

V. 2. Members PDB

The RCSB PDB is a member of the wwPDB, a collaborative effort with PDBe (UK), PDBj (Japan), and BMRB (USA) to ensure the PDB archive is global and uniform.

As the wwPDB archive keeper, the RCSB PDB updates the PDB archive atftp://ftp.wwpdb.org weakly. The structures included in each release are highlighted on the RCSB PDB home page and clearly defined on the FTP site. These sites are maintained 24 hours a day, seven days a week. A failover system automatically **(Berman H-M and al, 2003).**

Figure 21: Members PDB.

V. 2. 1. Protein Data Bank (RCSB)

Simple and advanced searching for macromolecules and Ligands, tabular reports, specialized visualization tools, sequence-structure comparisons, RCSB PDB Mobile, Molecule of the Month and other educational resources at PDB-101, and more.

V. 2. 2. Biological Magnetic Resonance Bank (BMRB)

Collects NMR data from any experiment and captures assigned chemical shifts, coupling constants, and peak lists for a variety of macromolecules; contains derived annotations such as hydrogen exchange rates, pKa values, and relaxation parameters.

V. 2. 3. Protein Data Bank Europe (PDBe)

Rich information about all PDB entries, multiple search and browse facilities, advanced services including PDBePISA, PDBe Fold and PDBe Motif, advanced visualisation and validation of NMR and EM structures, tools for bioinformaticians.

V. 2. 4. Protein Data Bank Japan (PDBj)

Supports browsing in multiple languages such as Japanese, Chinese, and Korean; identifies functionally or evolutionarily conserved motifs by locating and annotating sequence and structural similarities, tools for bioinformaticians, and more **(Berman H-M and al, 2003).**

VI. The other protein structures of banks

The classification and comparison of the more than 50'000 protein structures deposited in the PDB (Berman H-M and al) is an essential step to extract valuable knowledge from protein structure data **(Berman H-M and al, 2000)**.

Today, the two most prominent protein structure classification schemes are SCOP (**Andreeva A and dell, 2008)** and CATH (**Greene L-H and al, 2007**) both partition proteins into domains.

These domains are classified in a hierarchical manner:

VI. 1. SCOP (Structural Classification of Proteins)

The SCOP database is mainly based on expert knowledge and, on the first level of the hierarchy, defines four major classes namely all α, all β, α/β as well as α + β describing the content of secondary structure elements in the domain.

SCOP sorts protein domains into classes: Fold, super families, families.

VI. 2. CATH (Class Architecture Topology and Homology)

CATH starts at the class level defining three major classes of secondary structure content (all α, all β and α/β).

The four major levels of CATH are class: architecture, - topology, and homologous, -super family.

VI. Citric acid cycle (Krebs cycle)

VI. I. Definition

The citric acid cycle also known as the tricarboxylic acid cycle (TCA cycle), the Krebs cycle, or the Szent-Gyorgyi-Krebs cycle is a series of enzyme catalysed chemical reactions, which is of central importance in all living cells that use oxygen as part of cellular respiration. In eukaryotic cells, the citric acid cycle occurs in the matrix of the mitochondrion. The components and reactions of the citric acid cycle were established by seminal work from Albert Szent-Gyorgyi and Hans Krebs. In aerobic organisms, the maerianne is part of a metabolic pathway involved in the chemical conversion of carbohydrates, fats and proteins into carbon dioxide and water to generate a form of usable energy. Other relevant reactions in the pathway include those in glycolysis and pyruvate oxidation before the citric acid cycle and oxidative phosphorylation after it. In addition, it provides precursors for many compounds including some amino acids and is therefore functional even in cells performing fermentation. **(Raven, Peter, 2005)**.

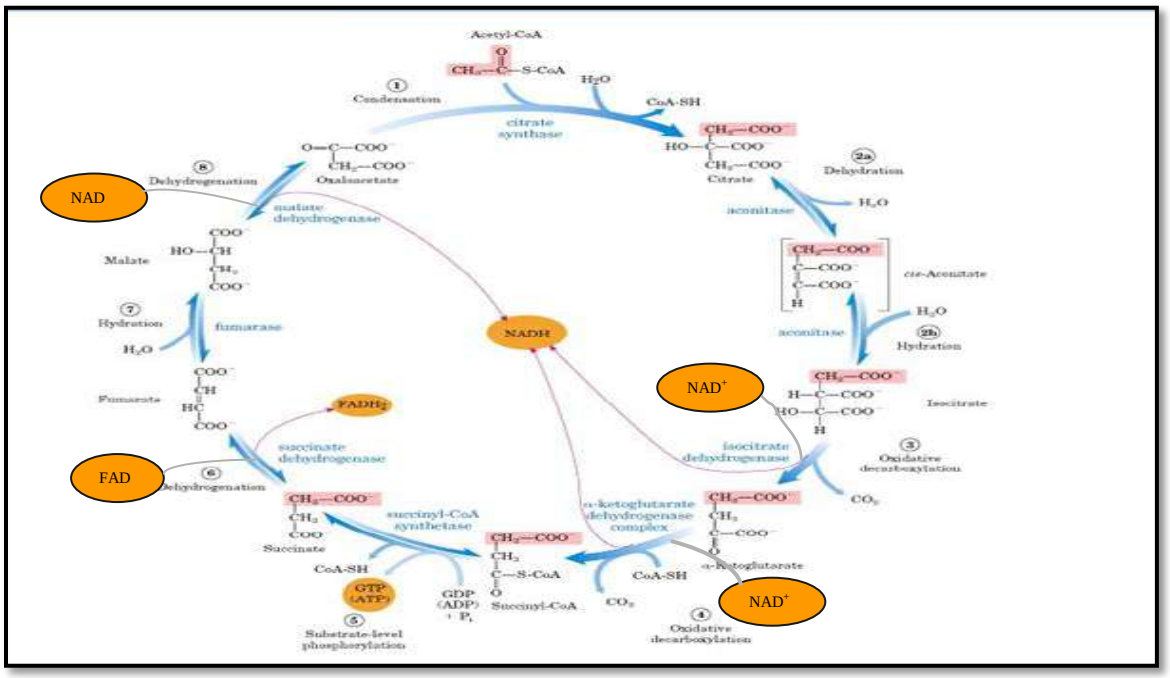


Figure 22: Citric acid cycle (Krebs cycle enzymes).

II. 2. Steps cycle enzymes

There are eight steps **(Raven, Peter, 2005)**

- ✓ **Step1:** The acetic acid subunit of acetyl CoA is combined with oxaloacetate to form a molecule of citrate through the enzyme citric synthase.

  The acetyl coenzyme A acts only as a transporter of acetic acid from one enzyme to another.

  After Step1, the coenzyme is released by hydrolysis. So that it may combine with another acetic acid molecule to begin the Krebs cycle again.

.

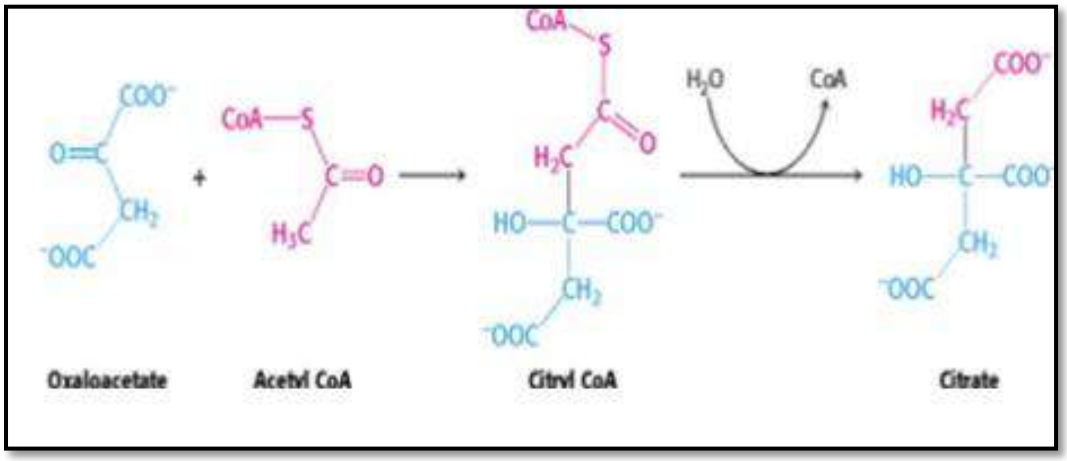


Figure 23: Formation of citrate.

- ✓ **Step2:** The citric acid molecule undergoes an isomerisation through the use of the enzyme Aconitase. A hydroxyl group and a hydrogen molecule are removed from the citrate structure in the form of water.

The two carbons form a double bond until the water molecule is added back.Only now, the hydroxyl group and hydrogen molecule are reversed with respect to the original structure of the citrate molecule. Thus, Isocitrate is formed.

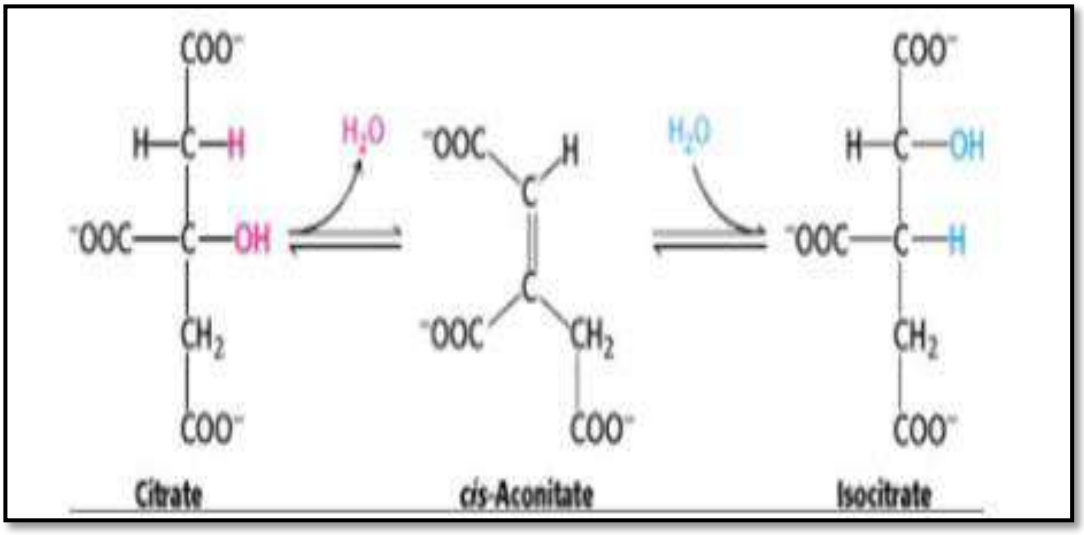


Figure 24: Formation of Isocitrate.

- ✓ **Step3:** In this step, the Isotrate molecule is oxidized by a NAD molecule through the use of the enzyme Isocitrate Synthase. The NAD molecule is reduced by the hydrogen atom and the hydroxyl group. The NAD binds with a hydrogen atom and carries off the other hydrogen atom leaving a carbonyl group. This structure is very unstable, so a molecule of CO2 is released creating alpha-Ketoglutarate

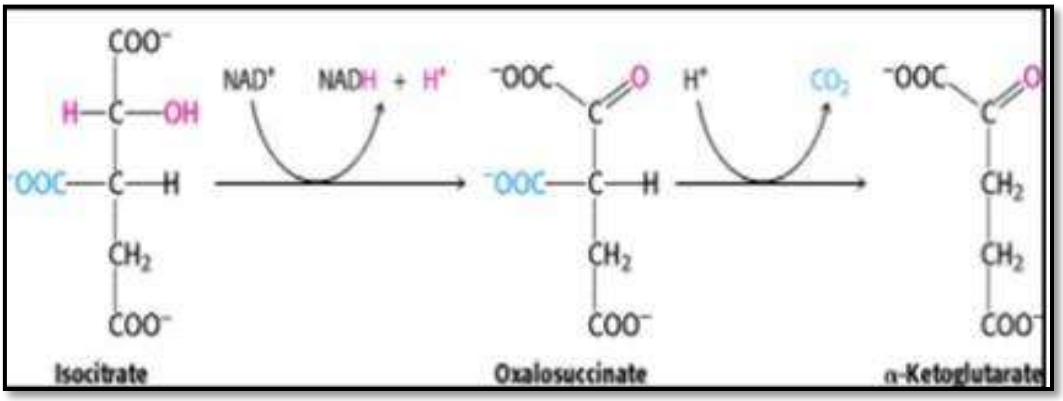


Figure 25: Formation of alpha Ketoglutarate.

- ✓ **Step 4 :** In this alpha-Ketoglutarate is oxidized through the use of the enzyme alpha-Ketoglutarate dehydrogenase.A molecule of NAD is reduced again to form NADH and leaves with another hydrogen.

  This instability causes a carbonyl group to be released as carbon dioxide and a thioester bond is formed in its place between the former alpha-Ketoglutarate and coenzyme A to create a molecule of Succinyl-coenzyme.

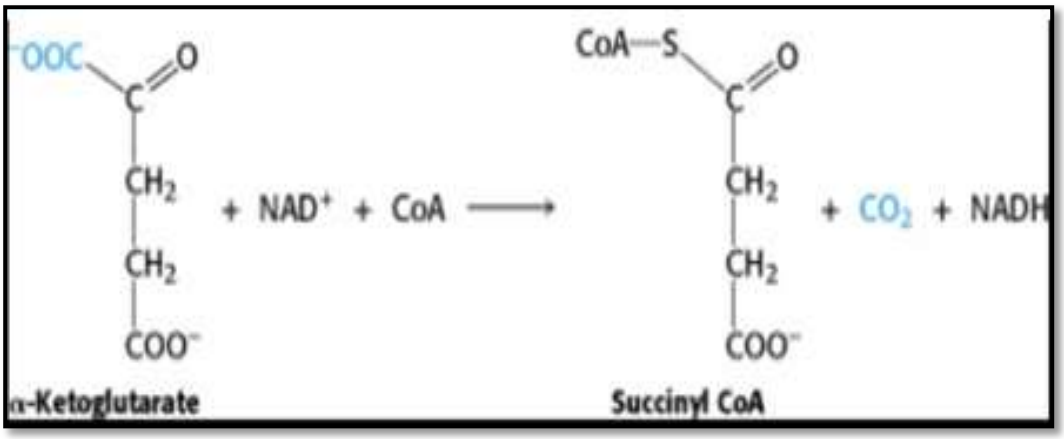


Figure 26: Formation of Succinyl-CoA.

- ✓ **Step 5:** A water molecule sheds is hydrogen atoms to coenzyme A. Then, a free-floating phosphate group displaces coenzyme A and forms a bond with the Succinyl complex by using the enzyme Succinyl-CoA synthase. The phosphate is then transferred to a molecule of GDP to produce an energy molecule of GTP. It leaves behind a molecule of Succinate.

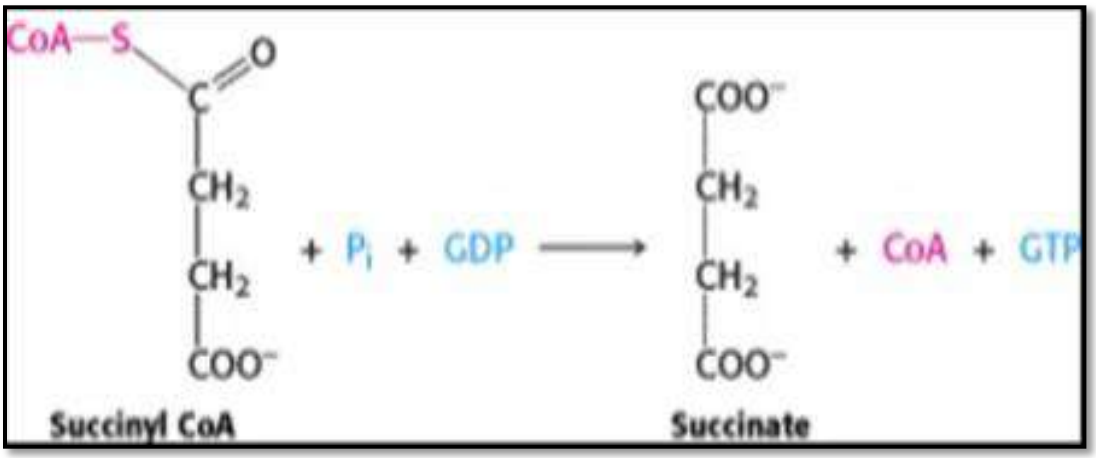


Figure 27: Formation of Succinate.

- ✓ **Step6:** In this step, Succinate is oxidized by a molecule of FAD (Flavin adenine dinucleotide) through the use of the enzyme Succinate Dehydrogenase. The FAD removes two hydrogen atoms from the Succinate and forces a double bond to form between the two carbon atoms, thus creating Fumarate.

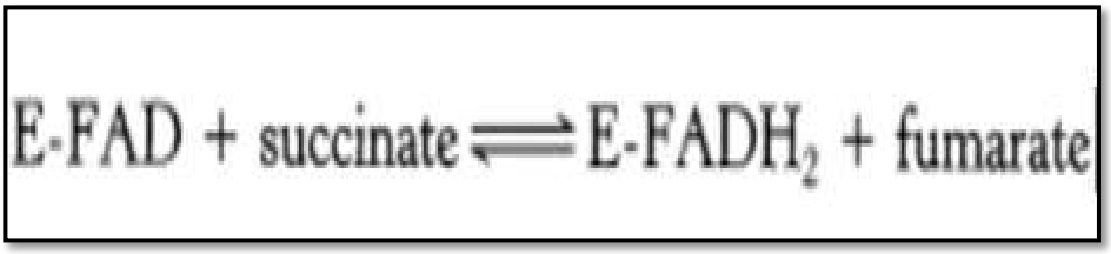


Figure 28: Formation of Fumarate.

- ✓ **Step7:** An enzyme adds water to the Fumarate molecule to form Malate through the use of the enzyme Fumarase.

  The Malate is created by adding one hydrogen atom to a carbon atom and then adding a hydroxyl group to a carbon next to a terminal.

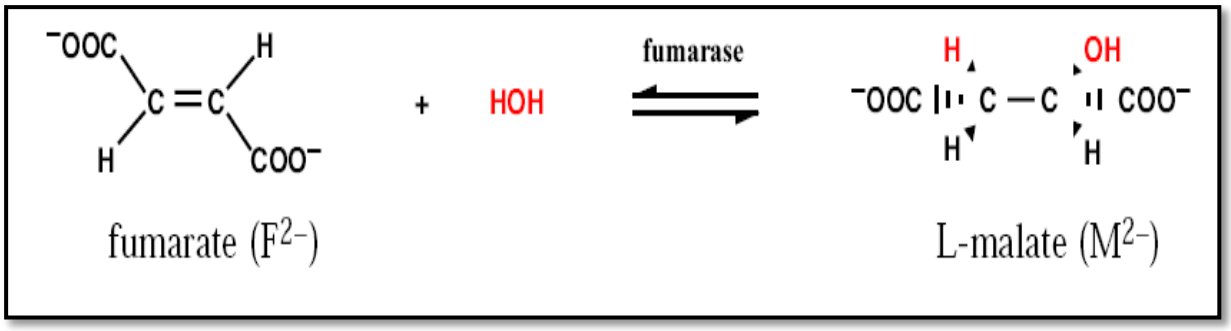


Figure 29: Formation of Malate.

✓ **Step8:** in this final step, the Malate molecule is oxidized by a NAD molecule through the use of the enzyme Malate Dehydrogenase. The carbon that carried the hydroxyl group is now converted into a carbonyl group.

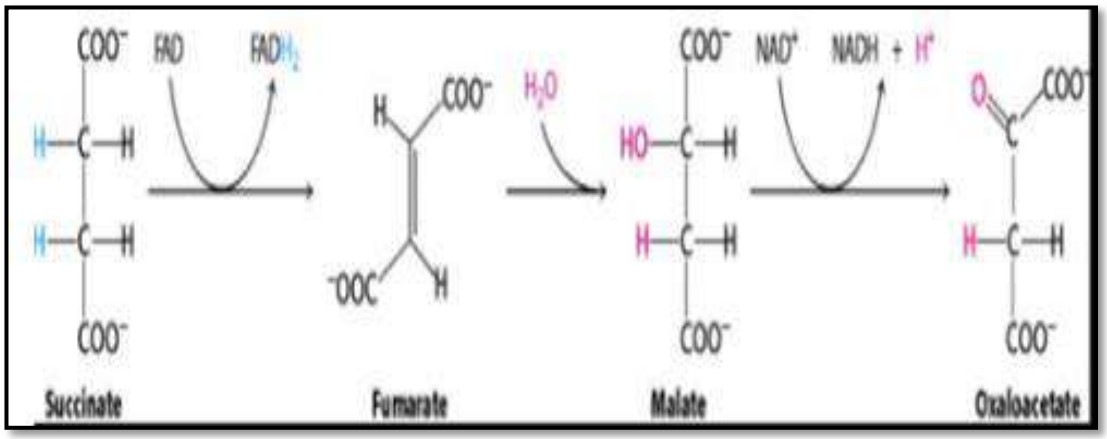


Figure 30: Rgeneration of Oxaloacetate.

❖ The end product is oxaloacetate which can then combine with acetyl-coenzyme A and begin the Krebs cycle all again.

Chapter II: Material and Methods.

Chapter II: Material and Methods

Objectives:

The objective of this study is to try and discover some of the basics of the relationship between the structure and function in the enzymes of the Krebs cycle or the Citric Acid Cycle (CAC).

Structural data related to the enzymes of the CAC cycle and their bound Ligands had to be located and prepared to properly carry out this structure-function relationship investigation.

The steps followed in achieving the goals of this study can be summarized in the following:

I. **Data preparation.**
II. **Data mining: Ligands Binding details calculation.**
III. **Flat-File Database: Binding details annotation and storing.**
IV. **Database Access.**

I. Data Preparation

As explained in the first chapter, macromolecular structure (DNA, RNA and Proteins) that are calculated at the different institutes around the world are send for annotation and storing to the PDB database. Generally the structure of proteins, which are the subject of interest in this project, are stored by the PDB in the form of entries that have codes of 4 digits **(see table below)**, these entry codes are also called identification codes or PDB IDs that are used to access the structural data provided in the relevant entries.

I. 1. The List of Protein Structures

All of the eight (8) enzymes involved in the CAC metabolic pathway, as explained in chapter I, have representative 3D-stuctures that are found in the PDB database. The structures are found in complex with their substrates or substrate analogues. For reasons to enrich the discovery that may exist more than one entry of each enzyme have been selected for the study whenever is possible and relevant.

Chapter II: Material and Methods

The PDB entries used this project amounted therefore to 32 ids. To facilitate the study, all of the PDB structures have been selected to be those determined by the X-ray crystallography method.

The table below represents the list of protein structures in complex with Ligands relevant to the CAC pathway. Resolution and R-factor which reflect the quality of the structures under study are also shown in the table.

Enzymes	Codes	Substrats	Méthode	Resolution	R-Value)
Citrate synthase	1AJ8 1IXE 2C6X 2P2W 2R9E 6CSC	CITRIC ACID CITRIC ACID CITRIC ACID CITRATE ANION S-CITRYLDETHIA COENZYME A CITRIC ACID	X-RAY DIFFRACTION	1.9A° 2.3A° 3.4A° 1.7A° 1.95A° 2.25A°	0.191 0.177 0.284 0.175 0.157 0.158
Aconitase	1ACO 1FGH 1L5J	ACONITATE ION 4-HYDROXY-ACONITATE ION ACONITATE ION	X-RAY DIFFRACTION	2.05A° 2.05A° 2.4A°	0.168 0.177 0.151
Isocitrate Dehydrogenase	1ITW 1XKD 2UXR 4AJ3 4AJB 4BNP	ISOCITRIC ACID	X-RAY DIFFRACTION	1.95A° 2.3A° 2.3A° 1.9A° 1.8A° 2.0A°	0.193 0.226 0.195 0.189 0.183 0.163
α- Ketoglutarate Dehydrogenase	3INM 4L06 4L04 4L03 4KZO 4AJR 4AJC 3o9Z 1CW4	2-OXOGLUTARIC ACID	X-RAY DIFFRACTION	2.1A° 2.28A° 2.87A° 2.1A° 2.2A° 2.69A° 2.3A° 1.45A° 2.1A°	0.220 0.224 0.200 0.189 0.178 0.160 0.175 0.194 0.184
SUCCINYL-CoA-Synthetase	2BWO 2BWN	SUCCINYL-COENZYME A SUCCINIC ACID	X-RAY DIFFRACTION	2.8 Å 2.1 Å	0.160 0.161
Succinate Dehydrogenase	2W8Q 4LH2 2DJL 3JU8	SUCCINIC ACID	X-RAY DIFFRACTION	2,4 1 ,67 1,38 1,82	0,224 0,164 0,165 0,158

Fumarase	4APB	FUMARIC ACID	X-RAY DIFFRACTION	1.94 Å	0.147
Malate Dehydrogenase	2dfd	MALATE ION	X-RAY DIFFRACTION	1.9 Å	0.194

Table 02: Protein/Enzyme structure (PDB codes) of the CAC pathway. Method of determination Resolution and R-factor.

I. 2. List of Ligands

The Table bellow represents all of the Ligands used in this project that make contacts with the proteins **(reported in Table 02).**

Ligand Name	Formula	Ligand ID	PDB ID
CITRIC ACID	C6 H8 O7	CIT	1AJ8, 1IXE, 2C6X, 6CSC
CITRATE ANION	C6 H5 O7 (3-)	FLC	2P2W
S-CITRYLDETHIA COENZYME A	C28 H46 N7 O22 P3	SDX	2R9E
ACONITATE ION	C6 H3 O6 (3-)	TRA	1ACO, 1L5J
4-HYDROXY-ACONITATE ION	C6 H3 O7 (3-)	ATH	1FGH
ISOCITRIC ACID	C6 H8 O7	ICT	1ITW, 1XKD, 2UXR, 4AJ3, 4AJB, 4BNP

SUCCINIC ACID	C4 H6 O4	SIN	2BWN, 2W8Q, 4LH2, 2DJL, 3JU8
FUMARIC ACID	C4 H4 O4	FUM	4APB
MALATE ION	C4 H5 O5 (1-)	MLT	2DFD
2-OXOGLUTARIC ACID	C5 H6 O5	AKG	3INM, 4L06, 4L04, 4L03, 4KZO, 4AJR, 4AJC, 3o9Z, 1CW4
SUCCINYL-COENZYME A	C25 H40 N7 O19 P3 S	SCA	2BWO

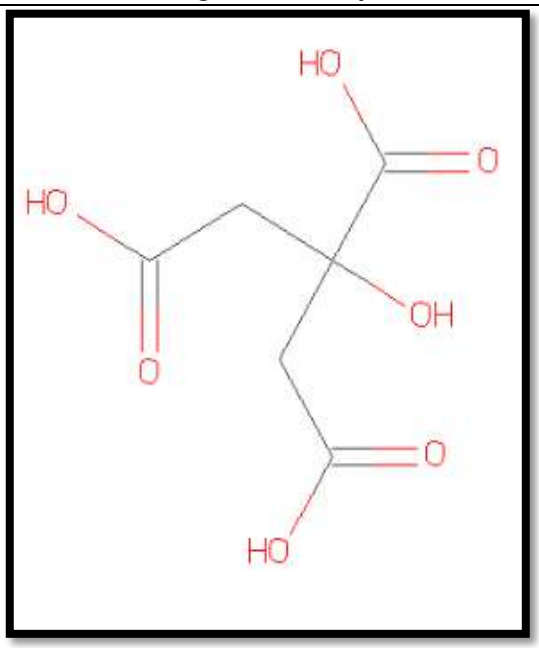
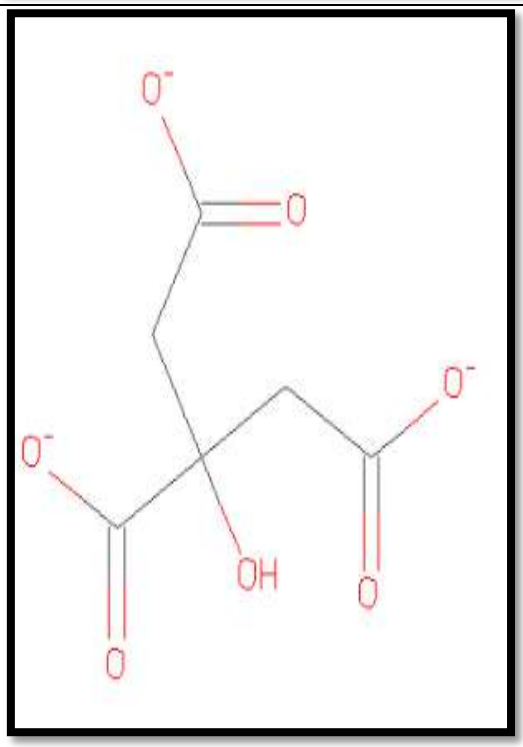
Table 03: List of Ligands, their PDB IDs, Name and general formulas.

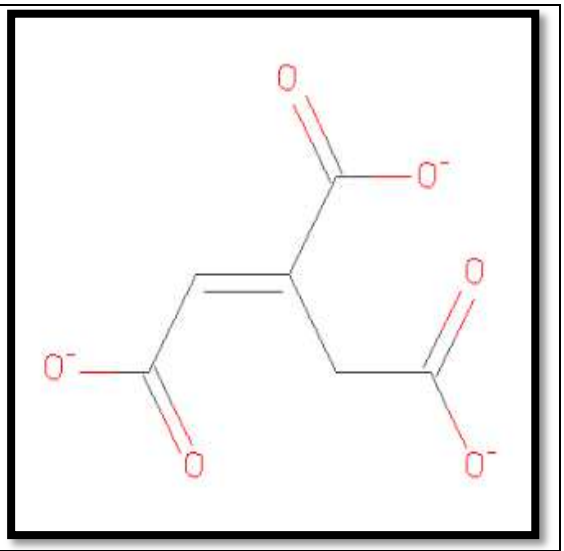
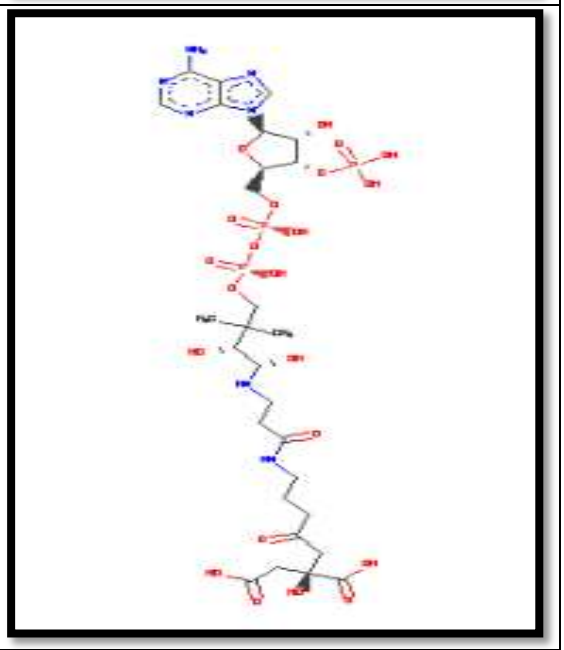
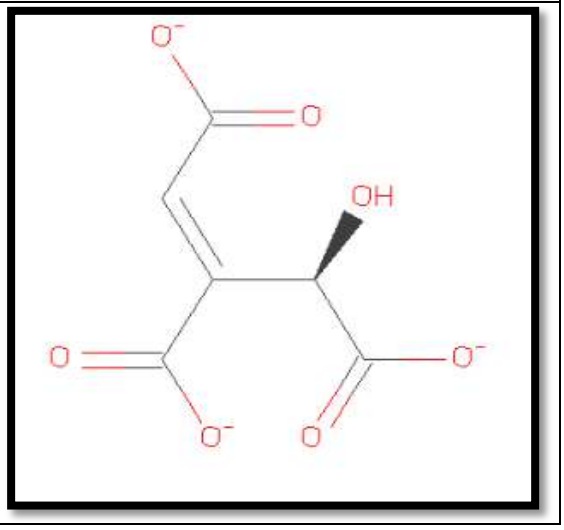
It should be noted that due to the difficulty of producing 3D-structures of enzymes bound to their natural substrates, analogue of these are used instead; hence some of the Ligands above are analogues to the natural substrates; for example the Ligands CIT, FLC and SDX are analogues to the natural substrate "Oxaloacetate". These analogues are, in addition, used by the structures' producers to study the various aspects of the enzymes binding sites and reaction dynamics.

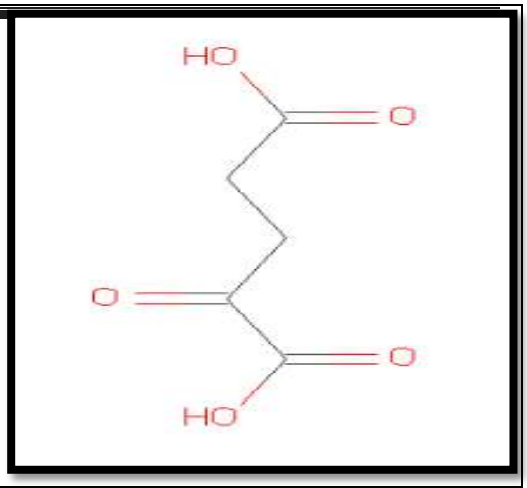
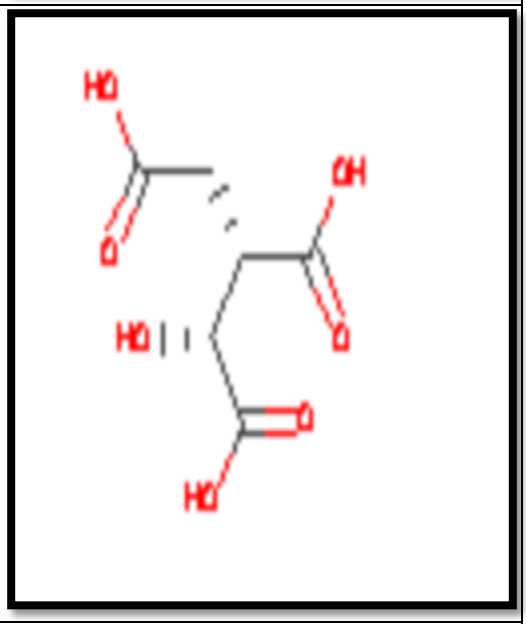
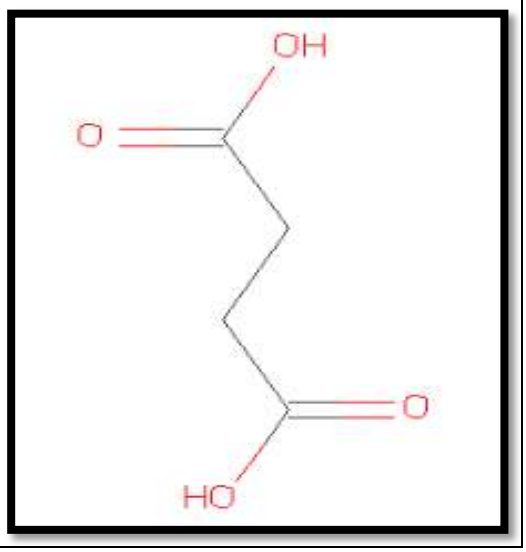
I. 3. Ligands Chemistry

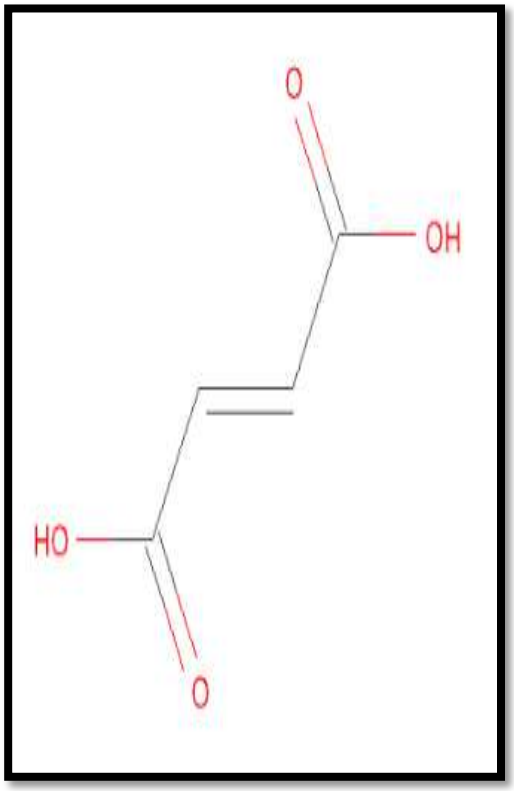
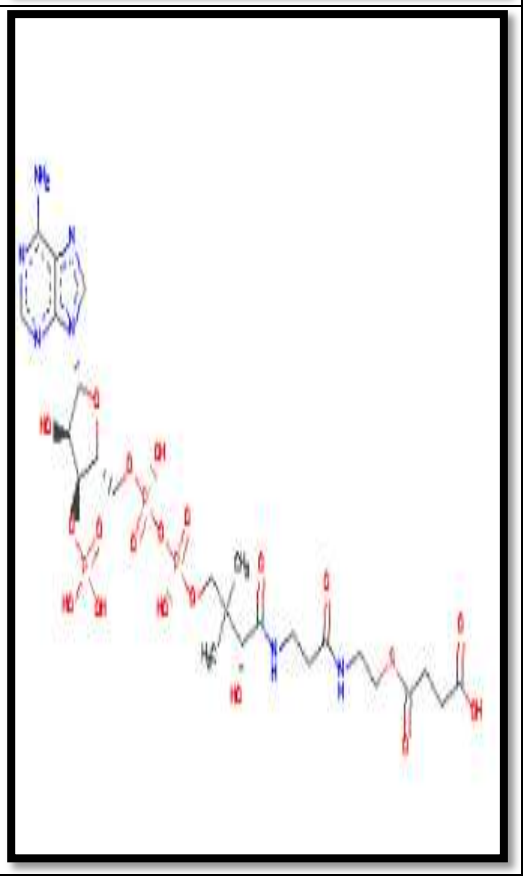
Here we choose the Ligands chemistry of Succinyl-coA synthase. .

For the others enzymes see:

Ligand Name	*Formula*	*Ligand ID*	*Ligand Chemistry*
CITRIC ACID	C6 H8 O7	CIT	
CITRATE ANION	C6 H5 O7 (3-)	FLC	

ACONITATE ION	C6 H3 O6 (3-)	TRA	
S-CITRYLDETHIA COENZYME A	C28 H46 N7 O22 P3	SDX	
4-HYDROXY-ACONITATE ION	C6 H3 O7 (3-)	ATH	

2-OXOGLUTARIC ACID	C5 H6 O5	AKG	
ISOCITRIC ACID	C6 H8 O7	ICT	
SUCCINYL-COENZYME A	C25 H40 N7 O19 P3 S	SCA	

FUMARIC ACID	C4 H4 O4	FUM	
SUCCINIC ACID	C4 H6 O4	SIN	

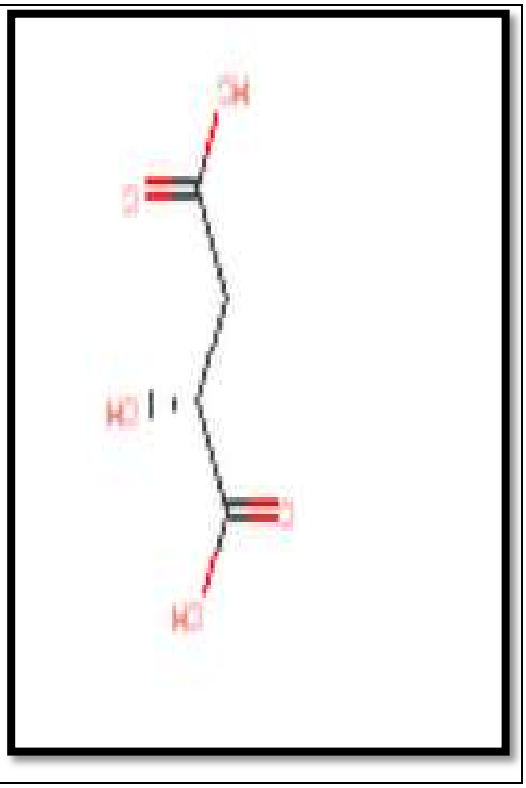
MALATE ION	C4 H5 O5 (1-)	MLT	

Table 04: Stereo chemistry of the Ligands under study.

II. Data mining: Ligands Binding details calculation

II. 1.Calculation of the Ligands Binding details:

The technique of Data Mining which is the systematic collection of data has been employed to collect the details of binding environment existing between the enzymes and their Ligands as found in the above selected PDB structures.

The Ligand Binding system (**Lgb**; http://bioinfrmaticstools.org/prjs/lgb) which is a lighter version from the Sequence Structure Function Server – SSFS **(A. Rachedi, 2011, Gloving A *et. al.*, 2005)** was used to calculate the binding environment details between the Ligands and the enzymes.

The table below, Table 05, is an example that represents the output of the **Lgb** for the enzyme Aconitase and its ligand the Aconitate Ion (ID: TRA) as found in the PDB entry 1aco. The **Lgb** is used as shown in **Figure 30**.

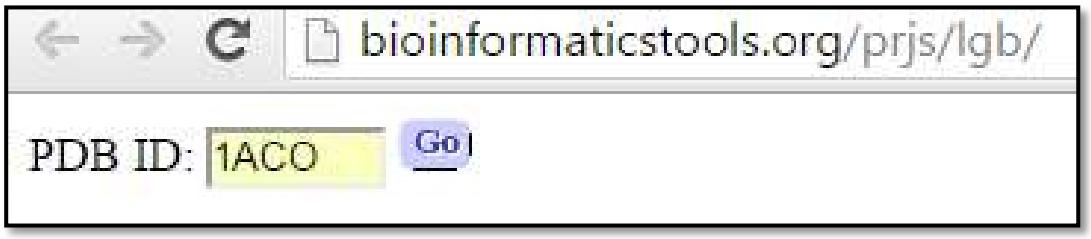


Figure 31: Capture the interface of the site **Lgb**.

Entry: 1aco	LYASE (CARBON-OXYGEN)

Protein-Ligand Environment

Protein Residues					Ligand				Bonds	
Chain	**SSelm**	**Name**	**Number**	**Atom**	**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	67-73 S: 1	GLN	72	CD	A	TRA	755	OA2	3.87	van der Waals
A	67-73 S: 1	GLN	72	OE1	A	TRA	755	OA2	3.93	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CA	3.7	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CB	3.96	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CAC	3.7	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	OA2	2.93	H.Bond
A	No SSE	ALA	74	CB	A	TRA	755	OA2	3.5	van der Waals
A	No SSE	THR	75	CG2	A	TRA	755	OA2	3.59	van der Waals
A	No SSE	HIS	101	CD2	A	TRA	755	OA2	3.94	van der Waals
A	No SSE	HIS	101	NE2	A	TRA	755	CG	3.56	H.Bond
A	No SSE	ASP	165	CA	A	TRA	755	OB2	3.33	van der Waals
A	No SSE	ASP	165	C	A	TRA	755	OB2	3.52	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	CBC	3.72	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	OB1	3.92	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	OB2	3.87	van der Waals
A	No SSE	ASP	165	OD1	A	TRA	755	CB	3.33	van der Waals
A	No SSE	ASP	165	OD1	A	TRA	755	CG	3.49	van der Waals
A	No SSE	ASP	165	OD1	A	TRA	755	CBC	3.32	van der Waals
A	No SSE	ASP	165	OD1	A	TRA	755	OB1	3.71	H.Bond
A	No SSE	ASP	165	OD1	A	TRA	755	OB2	3.69	H.Bond
A	No SSE	ASP	165	OD2	A	TRA	755	CBC	3.95	van der Waals
A	No SSE	ASP	165	OD2	A	TRA	755	OB1	3.75	H.Bond
A	166-174 H: 5	SER	166	N	A	TRA	755	CBC	3.47	H.Bond
A	166-174 H: 5	SER	166	N	A	TRA	755	OB1	3.45	H.Bond
A	166-174 H: 5	SER	166	N	A	TRA	755	OB2	2.82	H.Bond
A	166-174 H: 5	SER	166	CA	A	TRA	755	OB2	3.83	van der Waals
A	166-174 H: 5	SER	166	CB	A	TRA	755	OB1	3.52	van der Waals
A	166-174 H: 5	SER	166	CB	A	TRA	755	OB2	3.65	van der Waals

A	166-174 H: 5	SER	166	OG		A	TRA	755	OB1		2.7	H.Bond
A	166-174 H: 5	SER	166	OG		A	TRA	755	OB2		3.67	H.Bond
A	423-425 H: 5	ILE	425	CG2		A	TRA	755	CGC		3.91	van der Waals
A	423-425 H: 5	ILE	425	CG2		A	TRA	755	OG1		3.99	van der Waals
A	No SSE	ARG	447	NH1		A	TRA	755	CGC		3.68	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OB1		3.17	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OG1		3.38	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OG2		3.53	H.Bond
A	No SSE	ARG	452	CZ		A	TRA	755	OG1		3.75	van der Waals
A	No SSE	ARG	452	CZ		A	TRA	755	OG2		3.68	van der Waals
A	No SSE	ARG	452	NH1		A	TRA	755	CGC		3.7	H.Bond
A	No SSE	ARG	452	NH1		A	TRA	755	OG1		2.97	H.Bond
A	No SSE	ARG	452	NH1		A	TRA	755	OG2		3.67	H.Bond
A	No SSE	ARG	452	NH2		A	TRA	755	CGC		3.61	H.Bond
A	No SSE	ARG	452	NH2		A	TRA	755	OG1		3.68	H.Bond
A	No SSE	ARG	452	NH2		A	TRA	755	OG2		2.87	H.Bond
A	No SSE	ARG	580	CZ		A	TRA	755	CAC		3.99	van der Waals
A	No SSE	ARG	580	CZ		A	TRA	755	OA1		3.52	van der Waals
A	No SSE	ARG	580	CZ		A	TRA	755	OA2		3.81	van der Waals
A	No SSE	ARG	580	NH1		A	TRA	755	OA1		3.18	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	CAC		3.12	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	OA1		3.05	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	OA2		2.71	H.Bond
A	No SSE	SER	642	CA		A	TRA	755	CBC		3.86	van der Waals
A	No SSE	SER	642	CA		A	TRA	755	OB1		3.82	van der Waals
A	No SSE	SER	642	CA		A	TRA	755	OB2		3.83	van der Waals
A	No SSE	SER	642	C		A	TRA	755	OB2		3.92	van der Waals
A	No SSE	SER	642	CB		A	TRA	755	CBC		3.9	van der Waals
A	No SSE	SER	642	CB		A	TRA	755	OG1		3.44	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CA		2.9	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CB		3.17	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CAC		3.65	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CBC		3.29	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	OA1		3.32	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OB1		3.88	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OB2		3.43	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OG1		3.37	H.Bond
A	No SSE	SER	643	N		A	TRA	755	CBC		3.82	H.Bond

A	No SSE	SER	643	N		A	TRA	755	OB2		3.04	H.Bond
A	No SSE	SER	643	CA		A	TRA	755	OB2		3.86	van der Waals
A	No SSE	SER	643	CB		A	TRA	755	OB2		3.6	van der Waals
A	No SSE	SER	643	OG		A	TRA	755	CBC		3.97	van der Waals
A	No SSE	SER	643	OG		A	TRA	755	OB2		2.76	H.Bond
A	No SSE	ARG	644	CB		A	TRA	755	OA1		3.82	van der Waals
A	No SSE	ARG	644	CD		A	TRA	755	OA1		3.94	van der Waals
A	No SSE	ARG	644	NE		A	TRA	755	OA1		2.93	H.Bond
A	No SSE	ARG	644	CZ		A	TRA	755	OA1		3.49	van der Waals
A	No SSE	ARG	644	NH2		A	TRA	755	OA1		3.36	H.Bond
A	No SSE	ARG	644	NH2		A	TRA	755	OG1		3.1	H.Bond
A	Water	HOH	1000	O		A	TRA	755	CB		3.88	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CG		3.45	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CBC		3.39	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CGC		3.5	van der Waals
A	Water	HOH	1000	O		A	TRA	755	OB1		2.63	H.Bond
A	Water	HOH	1000	O		A	TRA	755	OG2		2.98	H.Bond
A	Water	HOH	1000	H1		A	TRA	755	CBC		3.66	
A	Water	HOH	1000	H1		A	TRA	755	OB1		2.97	
A	Water	HOH	1000	H1		A	TRA	755	OG2		3.93	
A	Water	HOH	1000	H2		A	TRA	755	CB		3.53	
A	Water	HOH	1000	H2		A	TRA	755	CG		3.34	
A	Water	HOH	1000	H2		A	TRA	755	CBC		2.82	
A	Water	HOH	1000	H2		A	TRA	755	CGC		3.17	
A	Water	HOH	1000	H2		A	TRA	755	OB1		1.83	, Covalent
A	Water	HOH	1000	H2		A	TRA	755	OB2		3.73	
A	Water	HOH	1000	H2		A	TRA	755	OG2		2.79	

Table 05: The binding environment details of the TRA bound the enzyme Aconitase (PDB ID: 1ACO) as calculated by the **Lgb** system.

The binding details shown in the table above and presented in the columns are explained in the following:

- **The set of columns "Protein residues":** These columns show the atoms of the enzyme residues (AA) that bind with the ligand. The residues are also annotated in terms of what secondary structure elements (helix, b-sheet or loop) they may belong to.
- **The set of columns "Ligand":** Atoms of the ligand that are interacted with the protein.
- **The column "Bonds":** Types of possible bonds and their lengths.

Using the **Lgb** system, the binding environment details of all the Ligands associated with the 32 PDB structures have been calculated, collected and stored into a system of organized files (see Flat-Files database section).

II. 2. The Ligand Binding details and Motifs construction:

As seen above in the binding details, residues in contact with the ligand may or may not belong to some secondary structure elements.

The secondary structure elements annotation, found in the table above, are used to create a pattern to describe the ligand binding sites in an abstract manner as follows:

- (67-73 S: 1): Represents the secondary structure beta-Sheet denoted as **S**.
- (No SSE 74): Represents the lack of secondary structure which means it's a loop region and is denoted as **L**.
- (166-174 H: 5): Represents the secondary structure alpha-Helix and is denoted as **H**.
- (Water 1000): Represents the cases when water molecules contribute in the ligand site binding. This is denoted as W put between brackets.

The pattern representing the binding site of the ligand TRA found in the table above is represented as follows: **SLLLHHLLL (W)**.

Since types of patterns seem to be associated with types of functions, refer **Chapter III**, these abstract patterns can be better denoted as Functional Motifs.

II. 3. Binding Motifs Graphical Representation:

Graphical representations of the motifs in the ligand binding sites have been generated by the RasMol molecular graphics program (Virsiuon2.6) **(Roger Sayle, 1997)** where the helices **(H)** are shown as Red ribbons, beta-Stands **(S)** as Yellow ribbons and Loop regions **(L)** as Light Grey strips.

For reasons of clarifying the ligand binding sites, three types of RasMol images were produced for each ligand binding case:

- Motif- only, see **Figure 31-b**.
- Motif + Ligand see **Figure 32-b**.
- Motif + Ligand + Binding Residues. See **Figure 33-b**.

In order to create the images for the graphical representation of the motifs in the binding sites, RasMol program uses a script language that tells it what to and how to represent the molecular data in the graphical mode as seen in the figures **Figure 31-a, Figure 32-a and Figure 33-a**.

✓ **Motif -only** in RasMol representation of the motif **SLLLHHLLL (W)** for the ligand TRA, PDB ID:1ACO, chain A **(from Table 5)**:

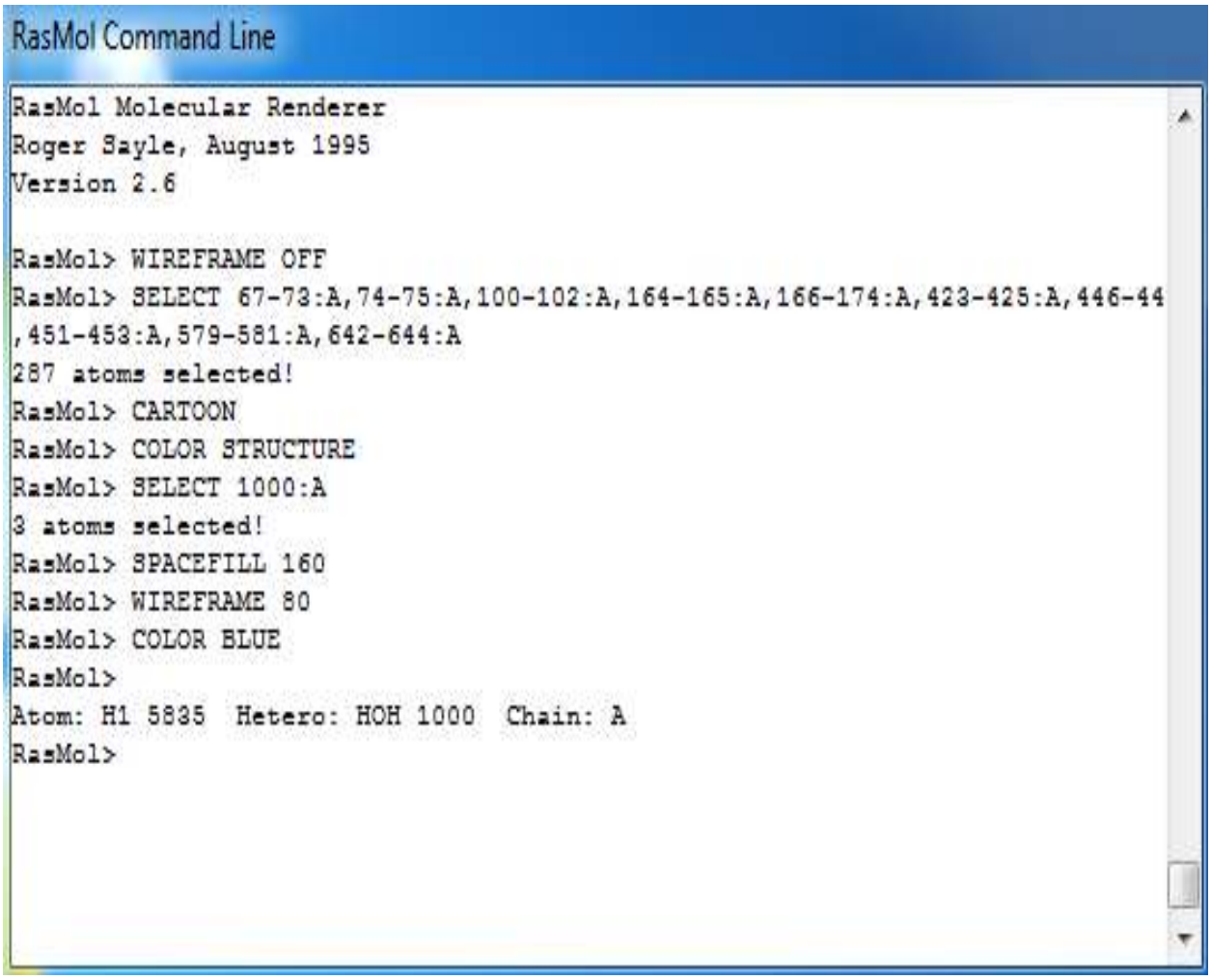


Figure 32-a: Capture of the RasMol script to create the motif representation shown in **Figure 32-b.**

Figure 32-B: Capture of RasMol representation of the binding motifs where the ligand TRA and binding residue are not shown in the case of Aconitase (PDB ID: 1ACO, chain A).

✓ **Motif + Ligand** in RasMol representation of the motif **SLLLHHLLL (W)** for the ligand TRA number 755, PDB ID:1ACO, Chain A **(from Table 5)**:

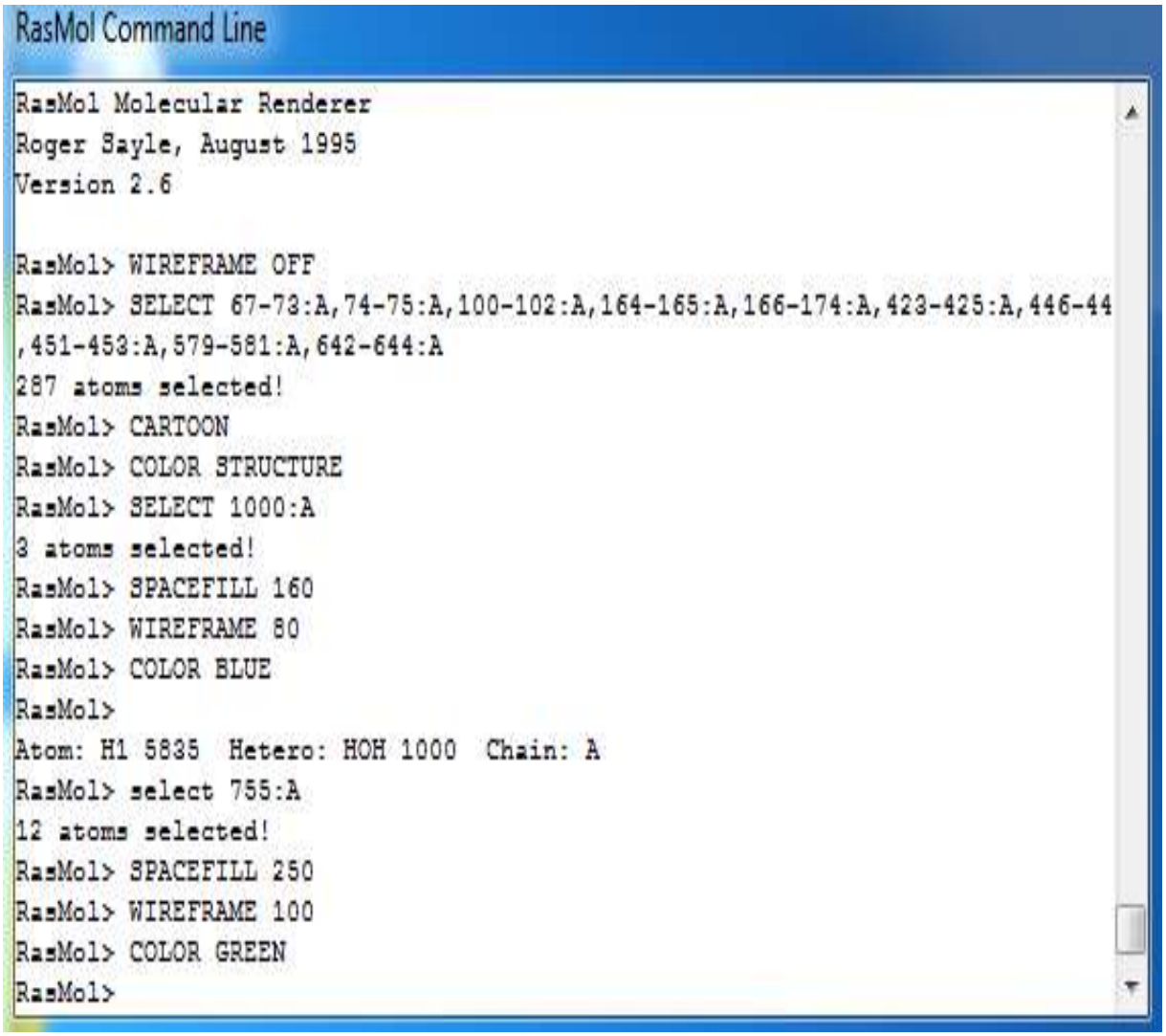


Figure 33-a: Capture of the RasMol script to create the motif representation shown in **Figure 33-b.**

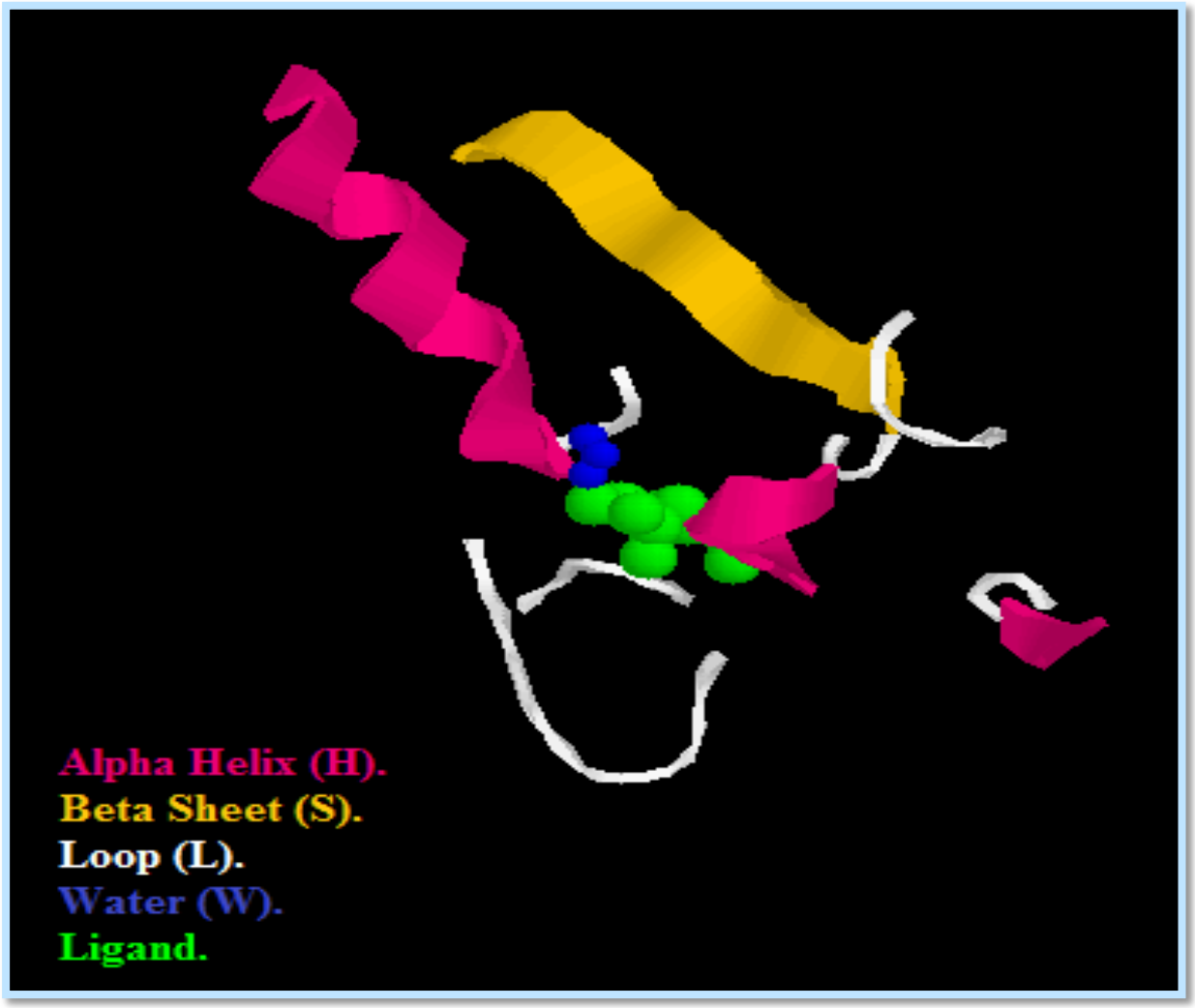


Figure 33-B: Capture of RasMol representation of the binding motifs where the binding residues are not shown but the ligand TRA is shown in the case of Aconitase (PDB ID: 1ACO, chain A).

✓ **Motif + Ligand + Binding Residues** in RasMol representation of the motif **SLLLHHLLL (W)** for the ligand TRA number 755, PDB ID:1ACO, Chain A **(from Table 5)**:

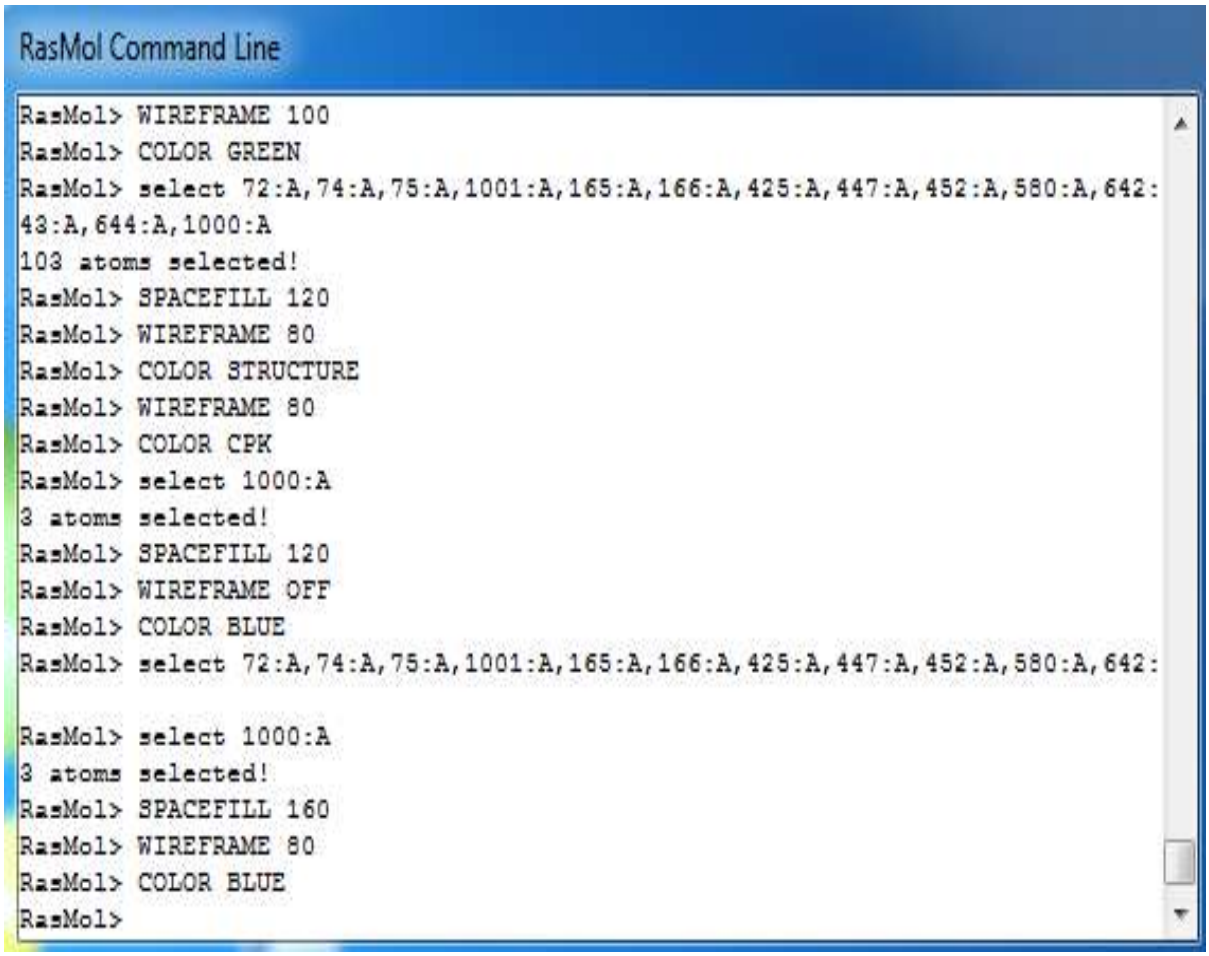


Figure 34-a: Capture of the RasMol script to create the motif representation shown in **Figure 34-b.**

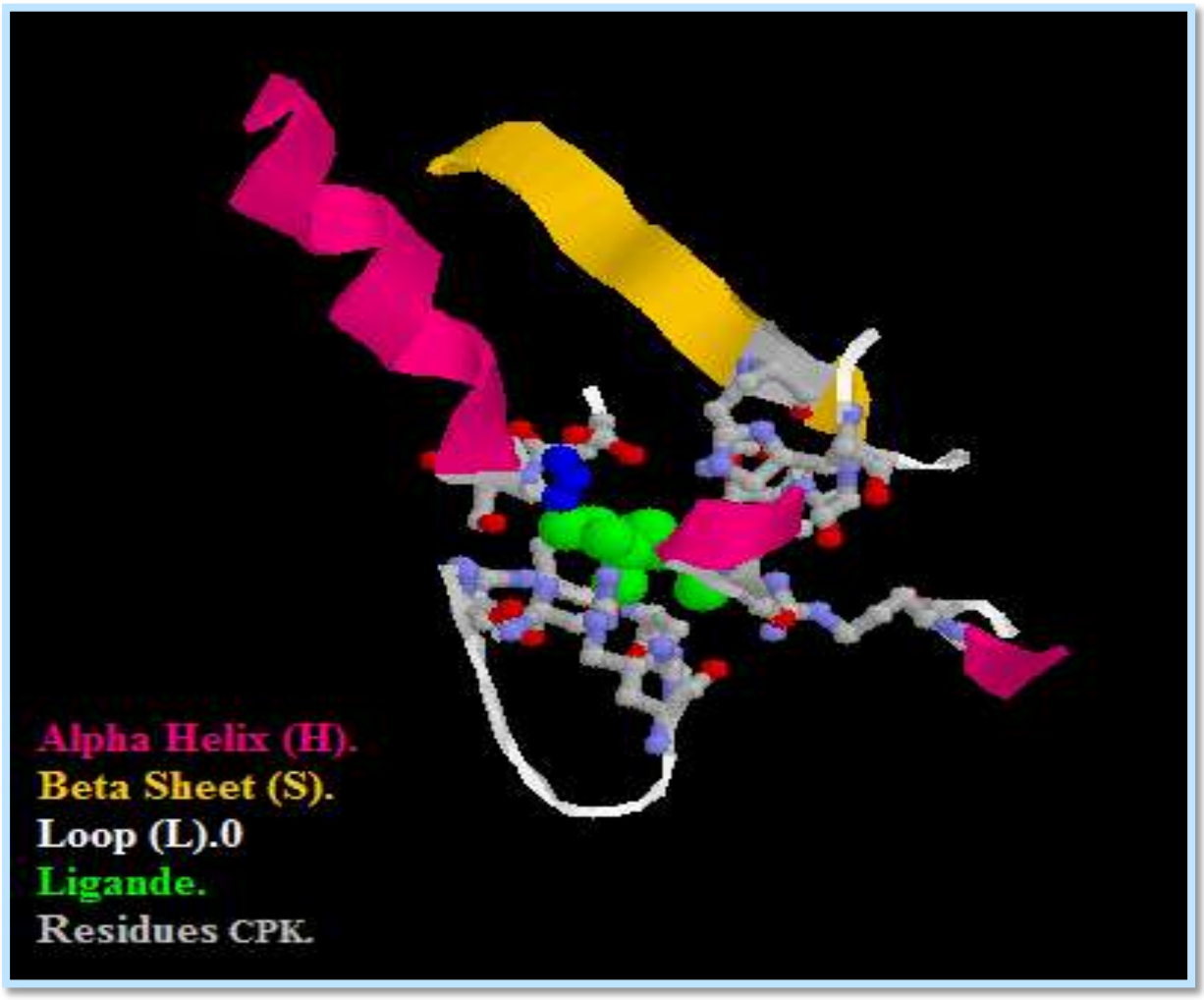


Figure 34-B: Capture of RasMol representation of the binding motifs where the binding residues and the ligand TRA are shown in the case of Aconitase (PDB ID: 1ACO, chain A).

It should be noted here that the three types of the graphical representations shown above for the case of the ligand TRA in complex with the enzyme Aconitase (PDB ID: 1aco, chain A) are done for all of the ligand binding instances in all of the enzyme complexes studied in this project, see **Index II**.

III. Flat-files database: Binding details annotation and storing

A Flat-Files database, which is a simple schema type of a database, has been created and to achieve this, the calculated binding details, dealt with above, have been stored into files named systematically so that the names reflect faithfully the origins of each set of calculated ligand binding details (see **Figure 34**). The files were then arranged based on the types of enzymes and PDB ids as shown in **Figure 35**. The same treatment has been applied when storing the RasMol graphical representations of the binding details as shown in **Figure 36**.

1wiret1ACO_A_TRA_A_755_ct.TXT | wiret1FGH_A_ATH_A_755_ct.TXT | wiret1L5J_A_TRA_A_866_ct.TXT | wiret1AJ8_A_CIT_A_1000_ct.TXT

	ID	FB	Chain	SSE	SSErg	Residues	Atom	Residues	Atom	Dist/Å	B.Type
2	02	1ACO	A	S: 1	67-73	72	CD	755	OA2	3.87	van der Waals
3			A	S: 1	67-73	72	OE1	755	OA2	3.93	H.Bond
4			A	S: 1	67-73	72	NE2	755	CA	3.7	H.Bond
5			A	S: 1	67-73	72	NE2	755	CB	3.96	H.Bond
6			A	S: 1	67-73	72	NE2	755	CAC	3.7	H.Bond
7			A	S: 1	67-73	72	NE2	755	OA2	2.93	H.Bond
8			A	L	73-76	74	CB	755	OA2	3.5	van der Waals
9			A	L	73-76	75	CG2	755	OA2	3.59	van der Waals
10			A	L	100-102	101	CD2	755	OA2	3.94	van der Waals
11			A	L	100-102	101	NE2	755	CG	3.56	H.Bond
12			A	L	164-166	165	CA	755	OB2	3.33	van der Waals
13			A	L	164-166	165	C	755	OB2	3.52	van der Waals
14			A	L	164-166	165	CG	755	CBC	3.72	van der Waals
15			A	L	164-166	165	CG	755	OB1	3.92	van der Waals
16			A	L	164-166	165	CG	755	OB2	3.87	van der Waals
17			A	L	164-166	165	OD1	755	CB	3.33	van der Waals
18			A	L	164-166	165	OD1	755	CG	3.49	van der Waals
19			A	L	164-166	165	OD1	755	CBC	3.32	van der Waals
20			A	L	164-166	165	OD1	755	OB1	3.71	H.Bond
21			A	L	164-166	165	OD1	755	OB2	3.69	H.Bond
22			A	L	164-166	165	OD2	755	CBC	3.95	van der Waals
23			A	L	164-166	165	OD2	755	OB1	3.75	H.Bond
24			A	H: 5	166-174	166	N	755	CBC	3.47	H.Bond
25			A	H: 5	166-174	166	N	755	OB1	3.45	H.Bond
26			A	H: 5	166-174	166	N	755	OB2	2.82	H.Bond
27			A	H: 5	166-174	166	CA	755	OB2	3.83	van der Waals
28			A	H: 5	166-174	166	CB	755	OB1	3.52	van der Waals
29			A	H: 5	166-174	166	CB	755	OB2	3.65	van der Waals
30			A	H: 5	166-174	166	OG	755	CBC	3.6	van der Waals
31			A	H: 5	166-174	166	OG	755	OB1	2.7	H.Bond
32			A	H: 5	166-174	166	OG	755	OB2	3.67	H.Bond
33			A	H: 5	423-425	425	CG2	755	CGC	3.91	van der Waals
34			A	H: 5	423-425	425	CG2	755	OG1	3.99	van der Waals
35			A	L	446-453	447	NH1	755	CGC	3.68	H.Bond
36			A	L	446-453	447	NH1	755	OB1	3.17	H.Bond
37			A	L	446-453	447	NH1	755	OG1	3.38	H.Bond
38			A	L	446-453	447	NH1	755	OG2	3.53	H.Bond

Figure 35: Notepad++ example of a ligand binding details file, named 1ACO_A_TRA_A_755_ct.TXT, which means that the calculations belong to the ligand named TRA, number 755 found in PDB ID: 1ACO protein chain A**(see Index III)**.

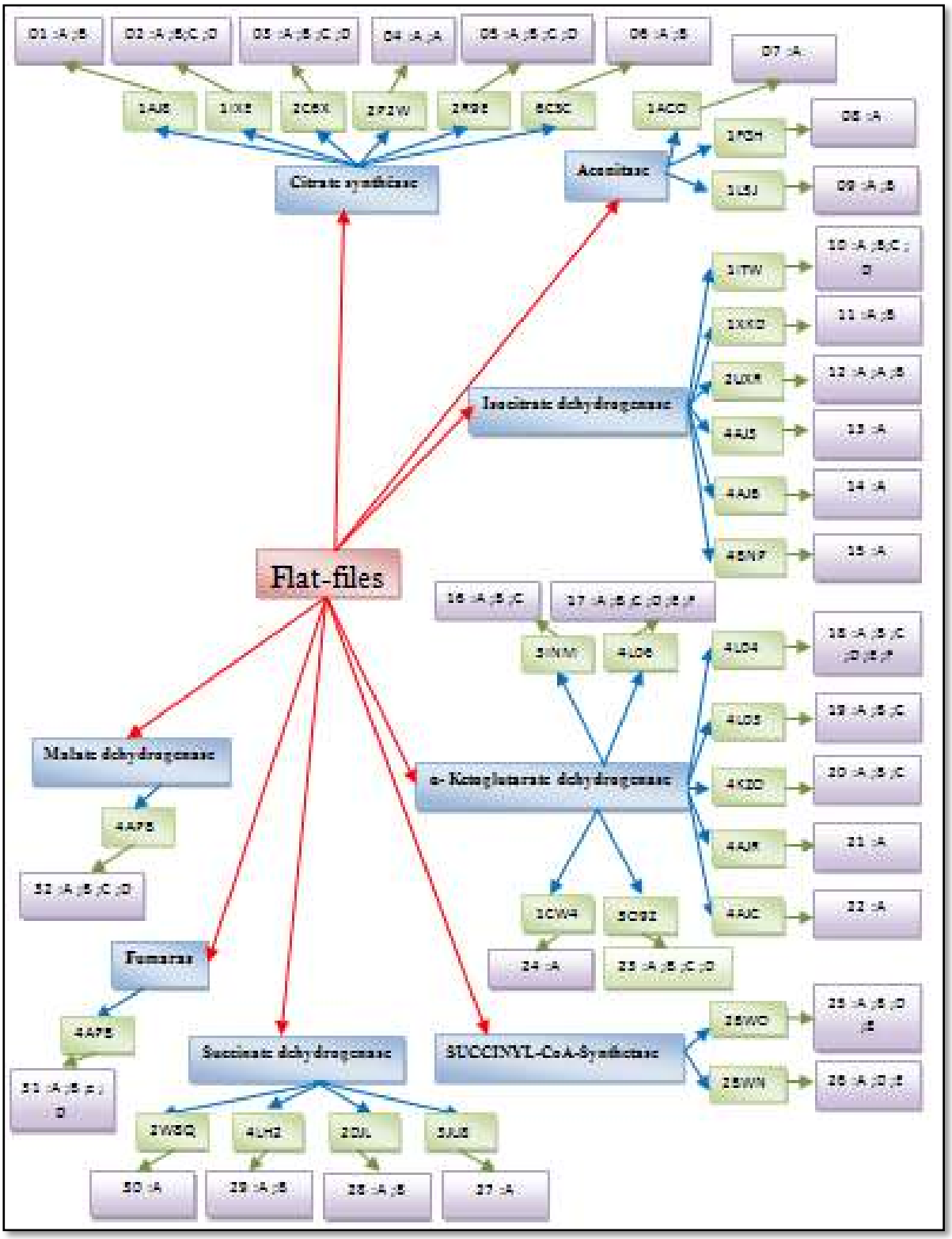


Figure 36: Flat-Files database schema shows the files of the ligand binding details are the arranged and stored.

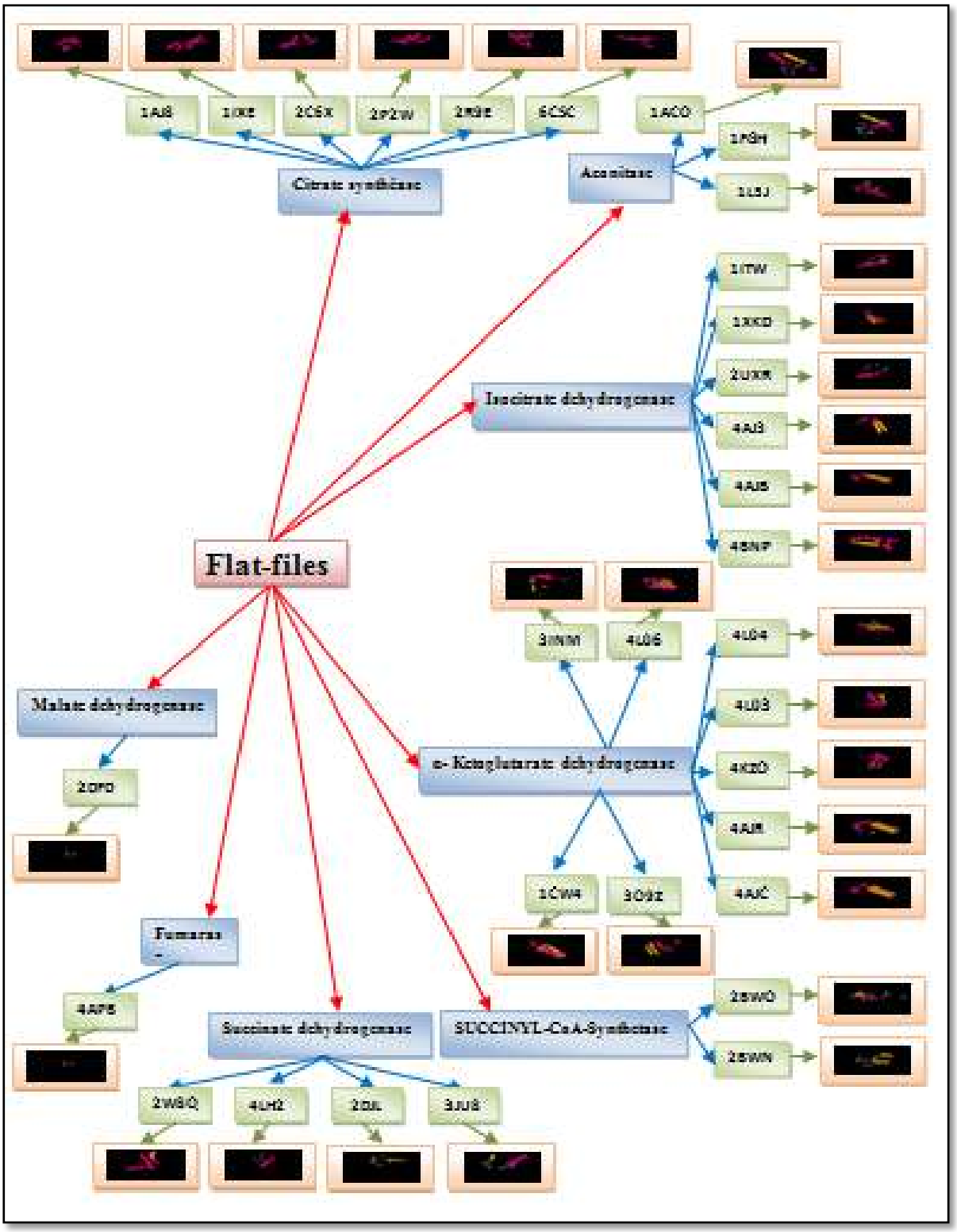


Figure 37: Flat-Files database schema shows the arrangement and storing of the ligand binding details files plus the image-files of the RasMol graphical representation.

IV. Online Database:

To share data with the local and international scientific communities, the Flat-Files database has been made available online by mounting it on the server BioinfornaticsTools **(Rachedi A., 2102)** by the supervisor of this project, who also wrote the necessary scripts to make it searchable.

This web version of the database has been named Citric Acid Cycle Binding Structural and Function motifs (CacSFMs).

CacSFMs database is accessible online via the following URL address:

http://bioinformaticstools.org/prjs/cacfms

Chapter III: Result and Discussion

I. Presentation of the results

I. 1. Data Base web access and Querying:

The online version of the database "CacSFMs" can be uploaded by invoking the URL address shown in previous chapter, section **VI. Online Database (page N° 76).** The figure below, no. 38, shows the interface of the "CacSFMs" which has been developed to allow the easy querying of data recorded in the database.

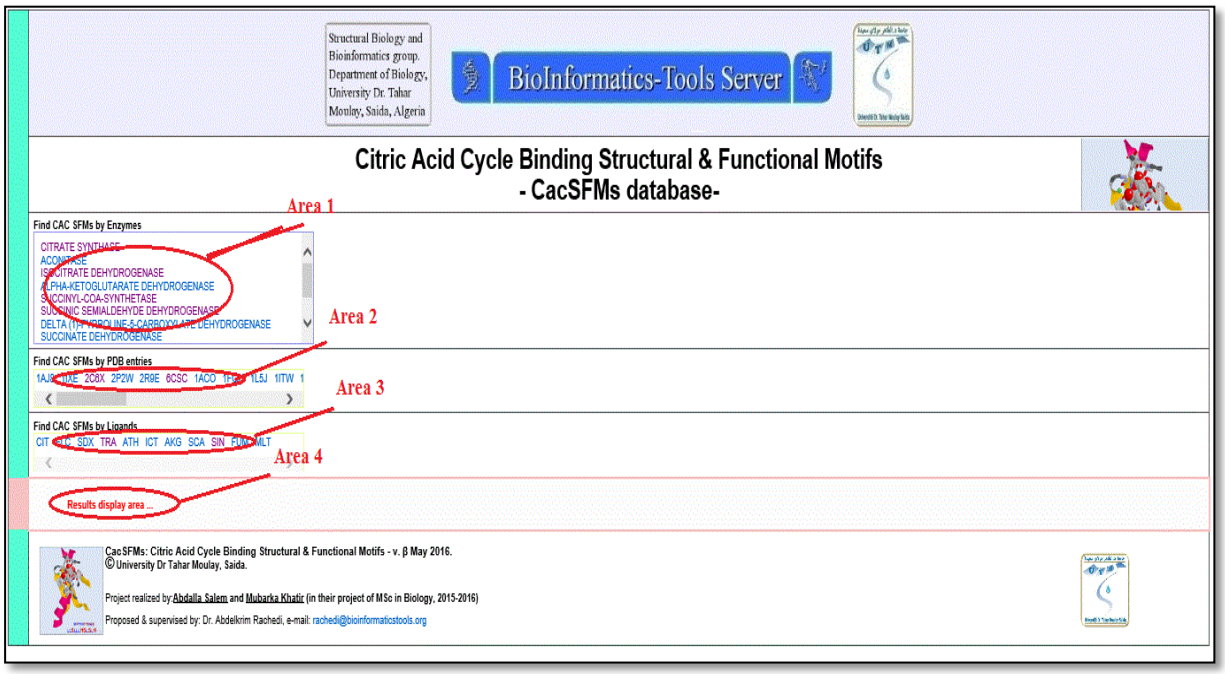


Figure 38: Capture of the interface page of the ”CacSFMs” database.

I. 2. Ways to search the Database and display Results:

As shown above in **Figure 38**, the CacSFMs interface allows 3 ways for search the database content. For clarity these ways of searching are highlighted and numbered in red color:

- **Area 1:** This list allows for querying by clicking on the CAC related enzymes.

- **Area 2:** This clickable list of PDB entries allows querying by PDB entry.
- **Area 3:** This clickable list of Ligands allows querying by ligand id.

Results of any search action taken by clicking on one of the areas above will be shown in the "**Results display area …**" (pink rectangle) as highlighted above in **Area 4**.

I. 3. Search by Enzymes or PDB entries:

As explained above, if the user clicks on the Enzymes or PDB entries lists, the search engine of the database will search accordingly and display the results as follows:

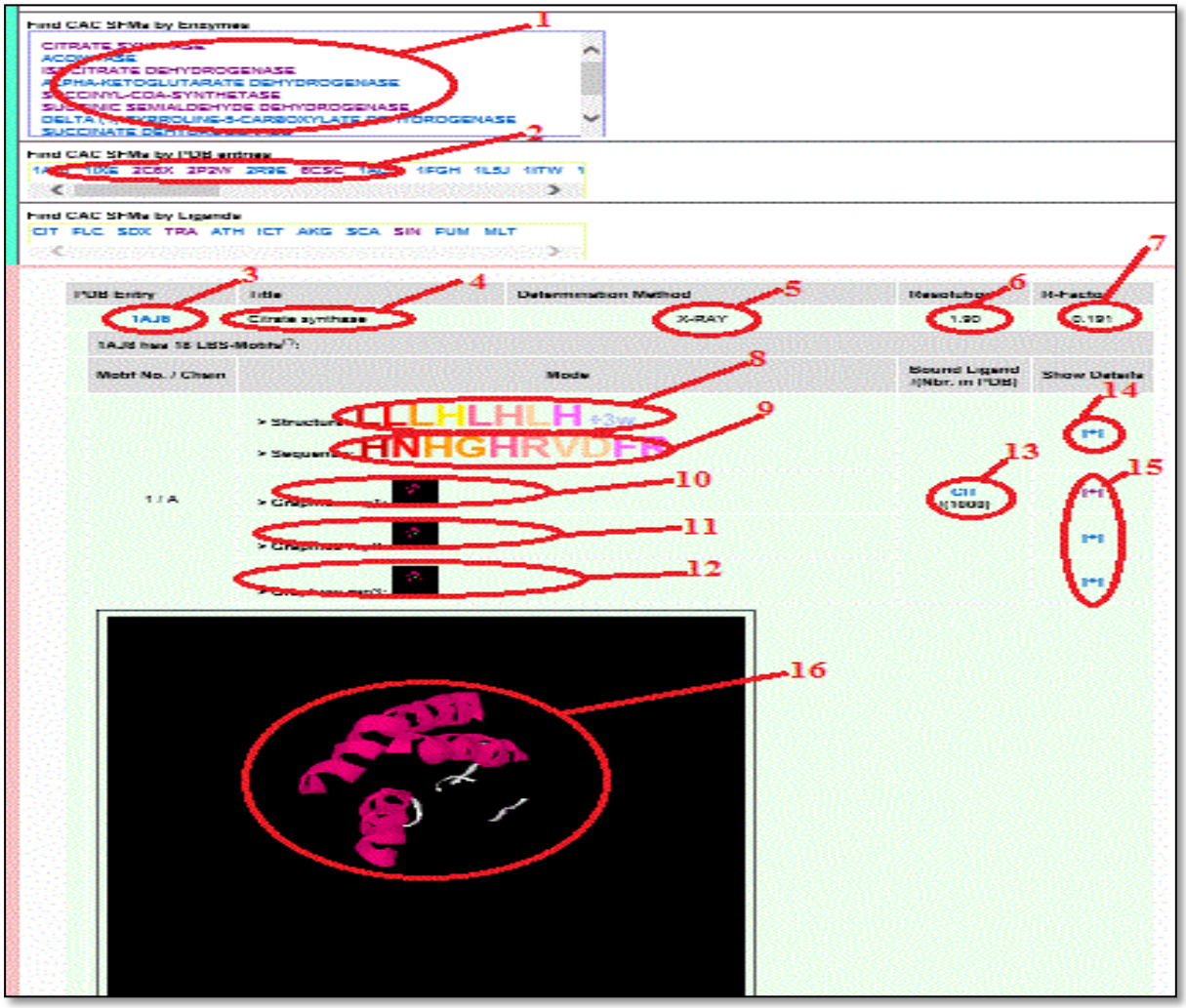


Figure 39: screenshot shows the result page after clicking on the Enzymes or PDB entries lists.

- **Area 1** represents the enzyme selected.

- **Area 2** represents the PDB entry selected (if applied).

Results are shown in the pink rectangle and explained in the following:

- **Area 3** shows the PDB entry.

- **Area 4** displays the enzyme name.

- **Areas 5, 6 and 7** display respectively the method of determination, resolution and R-Factor

- **Area 8** shows the binding motif and the number of water molecules involved.

- **Area 9** shows the Sequence of amino acids.

- **Area 10, 11 and 12** show respectively the images of Motifs-only, Motif+Ligand and -Motif+Ligand Binding residues.

- **Area 13** represents the ligand id and its number in the PDB.

- **Area 14** To be clicked to show the table for the Ligands binding details, see section "Ligand binding details Display".

- **Area 15** To be clicked to enlarge the motifs images (from RasMol).

- **Area 16** Represents enlarged image of motifs.

It's here noted that an important difference exists between searching by Enzymes and searching by PDB entry in that the first search could produce ligand binding results for more than one PDB. The search by PDB entry gives always results for only the selected PDB entry.

I. 4. Search by Ligand ID:

If the user clicks on the Ligands list, the search engine of the database will display the ligand binding results as follows:

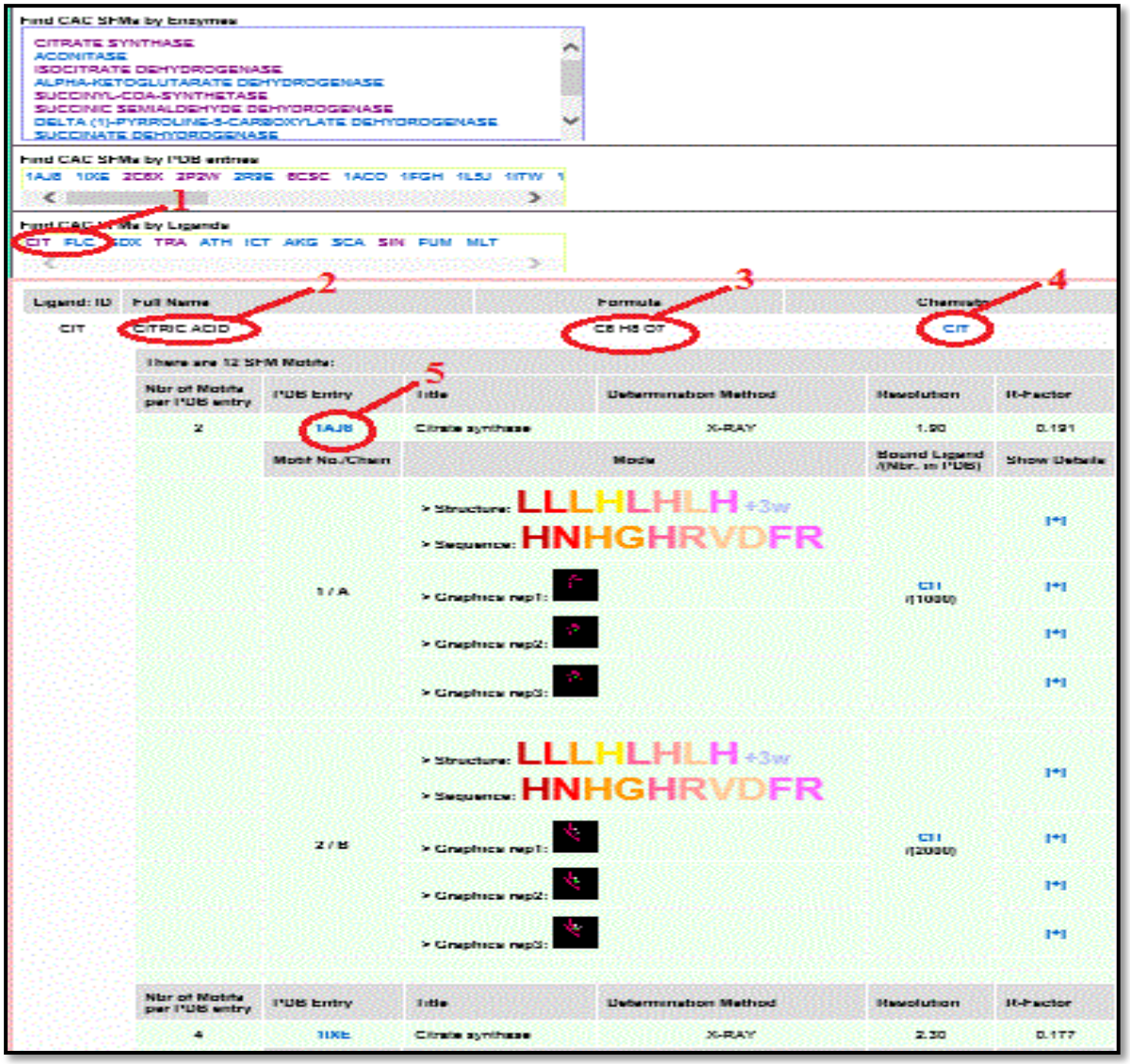


Figure 40: Screenshot shows the result page after the selection of a Ligand.

- **Area 1** represents the ligand selected.

- **Area 2** represents the scientific name of ligand.

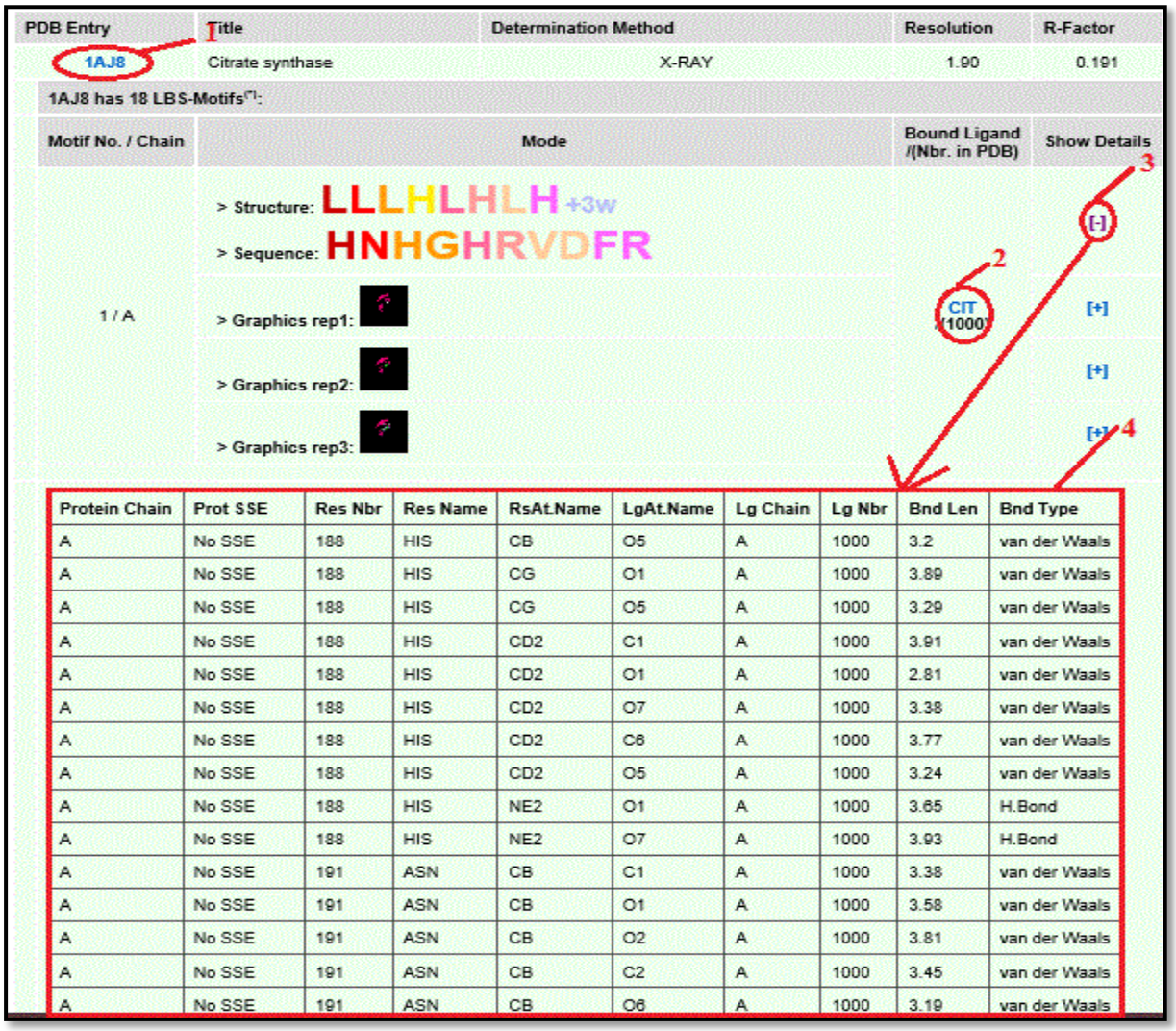
- **Area 3** represents the formula of ligand.

- **Area 4** represents the ligand chemistry.

- **Area 5** and above show the same data as explained in the last section **"Search by Enzymes or PDB entries" (page N° 79)**.

It's here noted that this method of searching by Ligands is more comprehensive compared to the other methods in that the ligand binding results are shown per PDB entries and per enzyme types.

I. 5. Ligand binding details display:



PDB Entry	Title	Determination Method	Resolution	R-Factor
1AJ8	Citrate synthase	X-RAY	1.90	0.191

1AJ8 has 18 LBS-Motifs[?]:

Motif No. / Chain	Mode	Bound Ligand /(Nbr. in PDB)	Show Details
1 / A	> Structure: LLLHLHLH +3w > Sequence: HNHGHRVDFR > Graphics rep1: > Graphics rep2: > Graphics rep3:	CIT (1000)	[-] [+] [+] [+]

Protein Chain	Prot SSE	Res Nbr	Res Name	RsAt.Name	LgAt.Name	Lg Chain	Lg Nbr	Bnd Len	Bnd Type
A	No SSE	188	HIS	CB	O5	A	1000	3.2	van der Waals
A	No SSE	188	HIS	CG	O1	A	1000	3.89	van der Waals
A	No SSE	188	HIS	CG	O5	A	1000	3.29	van der Waals
A	No SSE	188	HIS	CD2	C1	A	1000	3.91	van der Waals
A	No SSE	188	HIS	CD2	O1	A	1000	2.81	van der Waals
A	No SSE	188	HIS	CD2	O7	A	1000	3.38	van der Waals
A	No SSE	188	HIS	CD2	C6	A	1000	3.77	van der Waals
A	No SSE	188	HIS	CD2	O5	A	1000	3.24	van der Waals
A	No SSE	188	HIS	NE2	O1	A	1000	3.65	H.Bond
A	No SSE	188	HIS	NE2	O7	A	1000	3.93	H.Bond
A	No SSE	191	ASN	CB	C1	A	1000	3.38	van der Waals
A	No SSE	191	ASN	CB	O1	A	1000	3.58	van der Waals
A	No SSE	191	ASN	CB	O2	A	1000	3.81	van der Waals
A	No SSE	191	ASN	CB	C2	A	1000	3.45	van der Waals
A	No SSE	191	ASN	CB	O6	A	1000	3.19	van der Waals

Figure 41: screenshot shows the ligand binding details for CIT, number 1000, PDB ID 1AJ8.

- **Area 1** PDB entry

- **Area 2** The binding ligand CIT number 1000

- **Area 3** To be clicked to show the table for the Ligands binding details.

- **Area 4** Table display the binding details for the ligand CIT. For explanations of the binding details, see chapter II section "**Calculation of the Ligands Binding details**" **(page N° 61)**.

II. Binding Motifs Types and Properties

The total of 11 Ligands studied in this project bound with 83 protein chains found in the 32 PDB entries. This resulted in the total number of 83 motif instances which in turn boiled down to 34 unique binding motifs as shown below in **Table 06**:

Number	Enzymes	Ligands	PDB ID	Chain	Motifs	Unique Motifs
01	Citrate synthase	CIT	1AJ8	A	LLHLHLH	LLHLHLH LHLLHHH LHLLHLH LHLLHLHH LHLHHH LLLHLHH LL LHLHLHHLHH LHLHLHLHH LHLHLHHHLHH SLLLHHLLL(W) HLLHLLHL HLLHLHL SHLLHLLLL(3W) LHSLHHHL SSHH LHSSHHL LLHL LHSSHHSL LHSSHH HSSHH LHSHHL
				B		
			1IXE	A	LHLLHHH	
				B		
				C	LHLLHLH	
				D	LHLLHLHH	
			2C6X	A	LHLHHH	
				B		
				C		
				D		
			6CSC	A	LLLHLHH	
				B		
		FLC	2P2W	A	LHLLHHH	
				A	LL	
		SDX	2R9E	A	LHLHLHHLHH	
				B		
				C	LHLHLHLHH	
				D	LHLHLHHHLHH	
02	Aconitase	TRA	1ACO	A	SLLLHHLLL(W)	
			1L5J	A	HLLHLLHL	
				B	HLLHLHL	
		ATH	1FGH	A	SHLLHLLLL(3W)	
03	Isocitrate Dehydrogenase	ICT	1ITW	A	LHSLHHHL	
				B		

				C		LHSSHL LHSSHSL LHSSH LSLHHSL LSLHSLL SLSLLL HLHSHLL LHLL LSLSLL LHLLS L LHLHLHL
				D		
			1XKD	A	SSHH	
				B	LHSSHHL	
			2UXR	A	LLHL	
				A	LHSSHHL	
				B	LHSSHHSL	
			4AJ3	A	LHSSHH	
			4AJB	A	HSSHH	
			4BNP	A		
04	α- Ketoglutarate Dehydrogenase	AKG	3INM	A	LHSHHL	
				B		
				C		
			4L06	A	LHSSH	
				B		
				C		
				D		
				E		
				F		
			4L04	A	LHSSHSL	
				B	LHSSHL	
				C	LHSSH	
				D		
				E	LHSSHL	
				F		
			4L03	A	LHSSHL	
				B		
				C		
			4KZO	A	LHSHHL	
				B		
				C		
			4AJR	A	LHSSHHL	
			4AJC	A	LHSSHHL	
			3O9Z	A	LSLHHSL	
				B		
				C		
				D		
			1CW4	A	HSSHH	
05	Succinyl-CoA-Synthetase	SCA	2BWO	A	LSLHSLL	
				B		
				D		
				E		
		SIN	2BWN	A	SLSLLL	
				D		
				E		
06	Succinate Dehydrogenase	SIN	2W8Q	A	HLHSHLL	
			4LH2	A	LHLL	
				B		

06	Succinate Dehydrogenase	SIN	2DJL	A	LSLSLL	
				B		
			3JU8	A	LHLLS	
07	Fumarase	FUM	4APB	A	L	
				B		
				C		
				D		
08	Malate Dehydrogenase	MLT	2DFD	A	LHLHLHL	
				B		
				C		
				D		
Total	**08**	**11**	**32**	**83**	**83**	**34**

Table 06: The total and unique ligand binding motifs that are found associated with all of the 11 Ligands under study. The colored motifs reflect higher occurrence frequency **(see details below)**.

The following describes the Ligands binding motifs, their properties and 3D representations as per the enzymes and reaction involved in the CAC pathway:

II. 1. Citrate synthase

II. 1. A. Formation of Citrate:

Reaction:

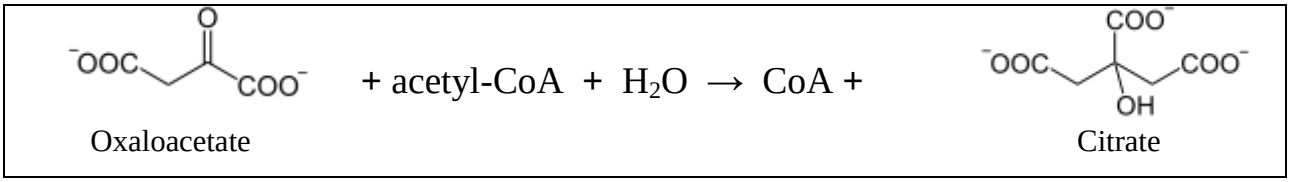


Figure 42: Formation of Citrate by the intervention of Citrate synthase enzyme.

II. 1. B. Citrate synthase Binding Motifs and their Properties:

The citrate synthase in all of the related PDB entries is represented by 18 protein chains where each is bound to an instant of the Oxaloacetate analogue; CIT, FLC and SDX thus the existence of 18 motifs, table 07:

- All motifs describe a mixture of **α-helices (H)** and **loop** regions **(L)** expect for the motifs denoted in the table as **LL**, in the case of the FLC ligand, instance number 402, in which case the motif is made of residues belonging to **loop region** only i.e. there is no secondary structure elements involved in the binding of the ligand in this case.
- The water molecules included in the motifs constructs are only those water molecules showing hydrogen bonding of less than 3Å. These waters are deemed biologically necessary in the reaction.

Citrate synthase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
CIT	1AJ8	A	LLHLHLH	2
		B		
	1IXE	A	LHLLHHH	2
		B		
		C	LHLLHLH	1
		D	LHLLHLHH	1
	2C6X	A	LHLHHH	4
		B		
		C		
		D		
	2CSC	A	LLLHLHH	2
		B		
FLC	2P2W	A	LHLLHHH(3W)	1
		A	LL(5W)	1
SDX	2R9E	A	LHLHLHHLHH	2
		B		
		C	LHLHLHLHH	1
		D	LHLHLHHHLHH	1

Table 07: Ligands bound to Citrate synthase, binding motifs and the motif's occurrence frequency.

III. 1. C. Motifs 3D-Graphical Representation:

- As seen above, in the case of the citrate synthase, there are 6 PDB entries: 1AJB, 1IXE, 2C6X, 6CSC, 2P2W, 2R9E binding a total of three ligand that are analogues to the natural substrate.
- The images below, table 08, represent the RasMol 3D-graphical representation of the binding motif associated with the PDN entry 2P2W.
- The 3D representation of all the other binding Motifs related to the other PDB entries (1AJB, 1IXE, 2C6X, 6CSC and 2R9E) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II.**

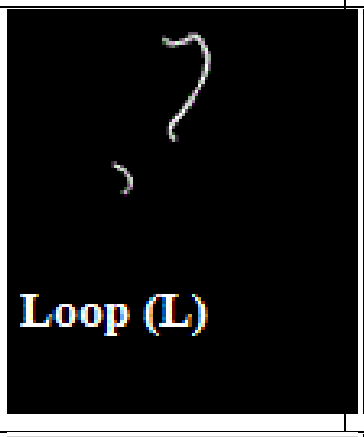
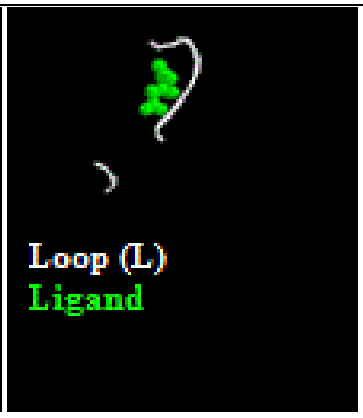
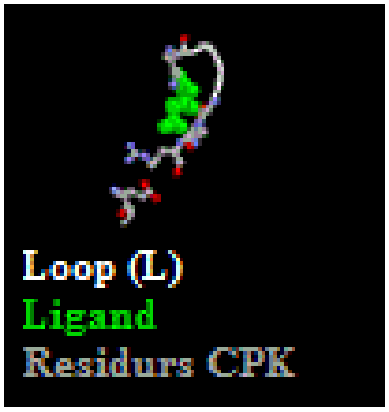
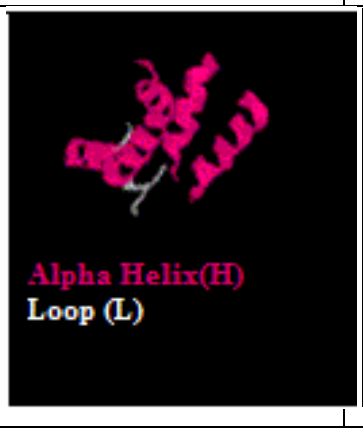
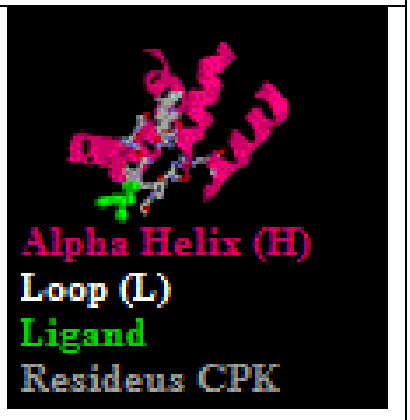
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
2P2W	A	FLC_{401}			
	A	FLC_{402}			

Table 08: 3D representation of the binding motifs associated with Citrate synthase for the case of PDB entry ID: 2P2W.

II. 2. Aconitase

II. 2. A. Formation of Aconitate:

Reaction:

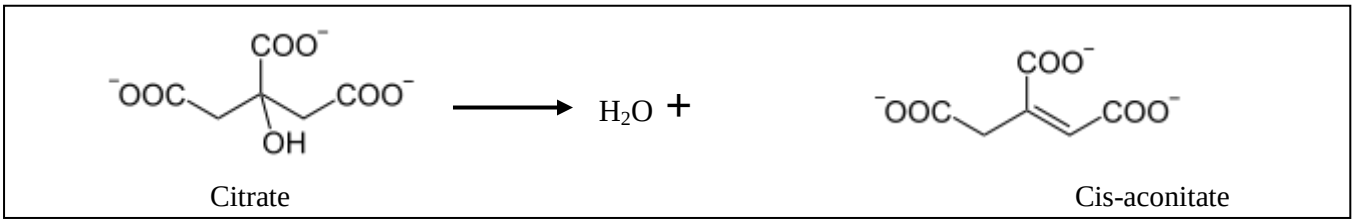


Figure 43: Formation of Aconitate by the intervention of Aconitase enzyme.

II. 2. B. Aconitase Binding Motifs and their Properties:

The Aconitase in all of the related PDB entries are represented by 04 protein chains where each is bound to an instant of the Citrate analogue; TRA and ATH thus the existence of 04 motifs, Table 09:

- Two motifs describe a mixture of **α-helices (H)**, **β-stands (S)** and **loop** regions **(L)** and the others Motifs are mixture of **α-helices (H)**, and **loop** regions **(L)** only.
- The water molecules included in the motifs constructs are only those water molecules showing hydrogen bonding of less than 3Å. These waters are deemed biologically necessary in the reaction.

Aconitase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
TRA	1ACO	A	SLLLHHLLL(W)	1
	1L5J	A	HLLHLLHL	1
		B	HLLHLHL	1
ATH	1FGH	A	SHLLHLLLL(W)	1

Table 09: Ligands bound to Aconitase, binding motifs and the motif's occurrence frequency.

II. 2. C. Motifs 3D-Graphical Representation:

- As seen above, in the case of the citrate synthase, there are **3** PDB entries: 1ACO, 1L5J, 1FGH binding a total of two ligand that are analogues to the natural substrate.
- The images below, table 10, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entries 1ACO (Chain A) and 1L5J (Chain A).
- The 3D representation of all the other binding Motifs related to the other chains of the PDB entry, 1L5J, and PDB entry (1FGH) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

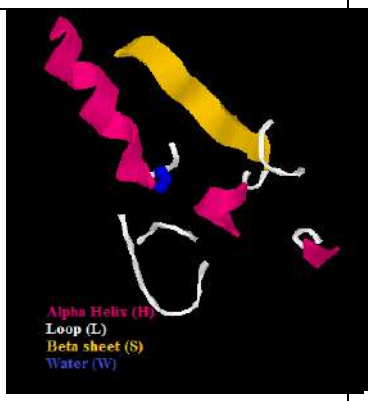
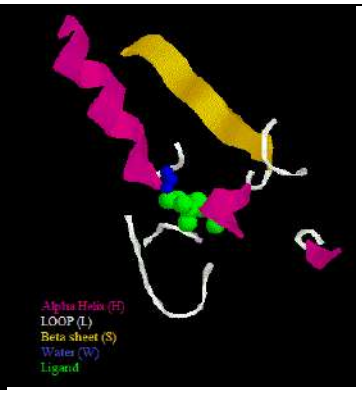
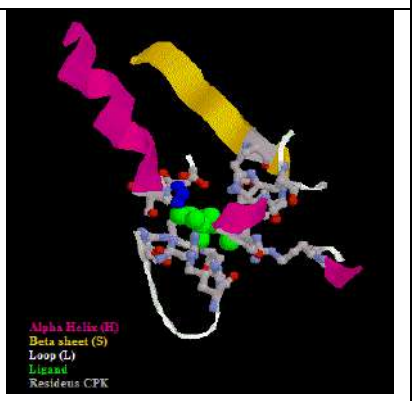
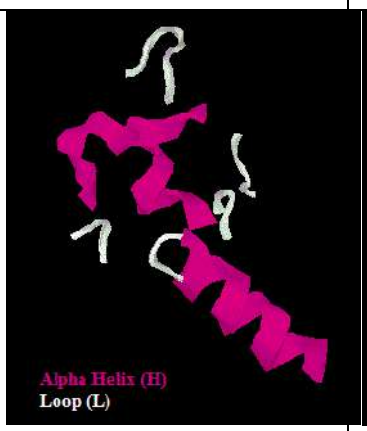
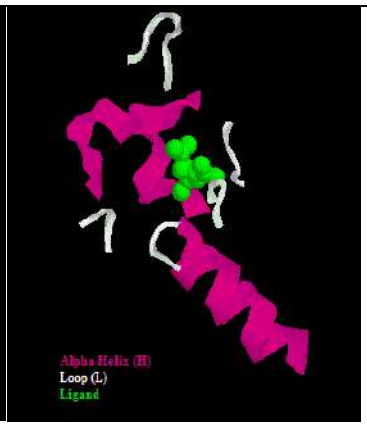
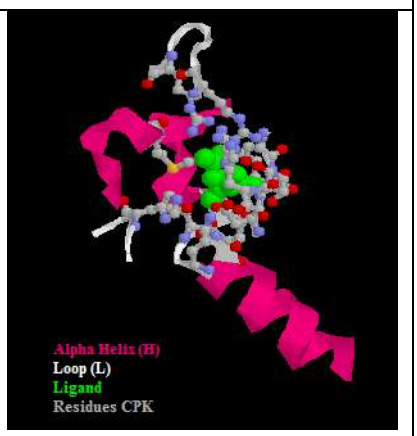
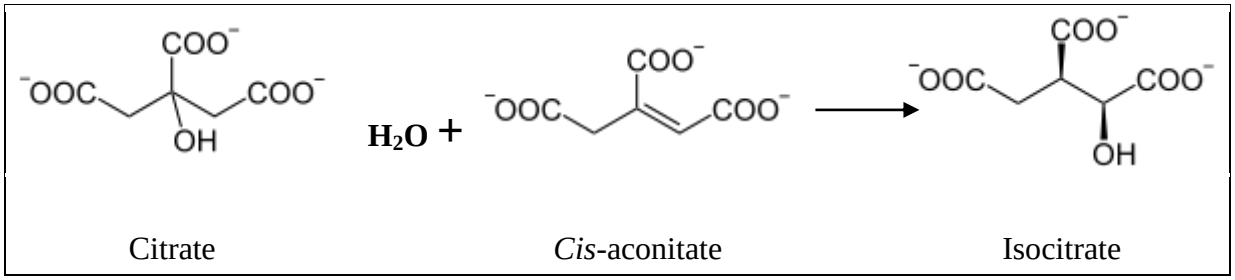
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
1ACO	A	TRA			
1L5J	A	TRA			

Table 10: 3D representation of the binding motifs associated with Aconitase for the cases of the PDB entries 1ACO (chain A) and 1L5J (chain A).

II. 3. Isocitrate dehydrogenase

II. 3. A. Formation of Isocitrat:

Reaction :



Citrate *Cis*-aconitate Isocitrate

Figure 44: Formation of Isocitrate by the intervention of Isocitrate Dehydrogenase enzyme.

II. 3. B. Isocitrate Dehydrogenase Binding Motifs and their Properties:

The Isocitrate Dehydrogenase in all of the related PDB entries is represented by 12 protein chains where each is bound to an instant of the Cis-aconitate analogue; ICT thus the existences of 12 motifs, table 11:

- Eight motifs describe a mixture of **α-helices** (**H**), **β-strands (S)** and **loop** regions (**L**). For the other cases, the motifs are divided as follows:
  - Three motifs are mixture of **α-helices** (**H**) and **β-strands** (**S**).
  - One motif is mixture of **α-helices** (**H**) and **loop** regions (**L**).

Isocitrate Dehydrogenase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
ICT	1ITW	A	LHSLHHHL	4
		B		
		C		
		D		
	1XKD	A	SSHH	1
		B	LHSSHHL	2
	2UXR	A	LHSSHHL	
		A	LLHL	1
		B	LHSSHHSL	1
	4AJ3	A	LHSSHH	1
	4AJB	A	HSSHH	2
	4BNP	A		

Table 11: Ligands bound to Isocitrate Dehydrogenase, binding motifs and the motif's occurrence frequency.

II. 3. C. Motifs 3D-Graphical Representation:

- As seen above, in the case of the Isocitrate Dehydrogenase, there are **6** PDB entries: 1ITW, 1XKD, 2UXR, 4AJ3, 4AJB, 4BNP all of which binds the ICT ligand which is an analogue of the natural substrate.
- The images below, table 12, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entries 1ITW (Chain A and Chain B).
- The 3D representation of all the other binding Motifs related to the other chains of PDB entry, 1ITW, and PDB entries (1XKD, 2UXR, 4AJ3, 4AJB, 4BNP) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

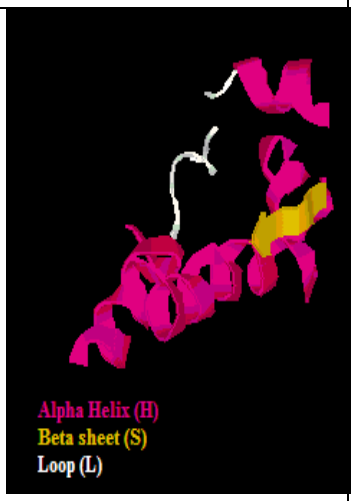
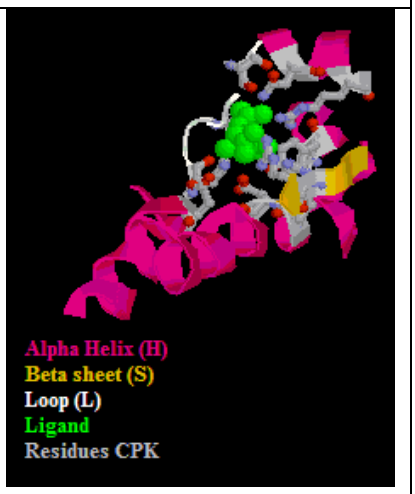
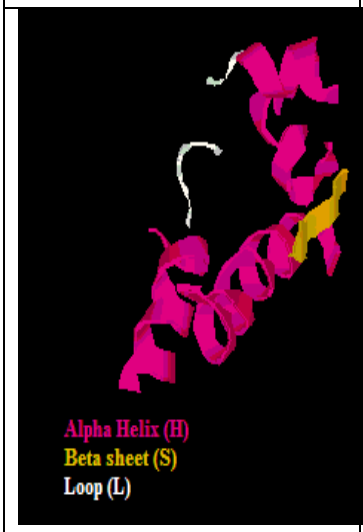
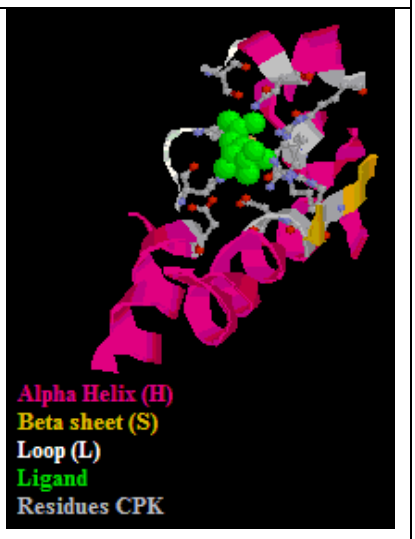
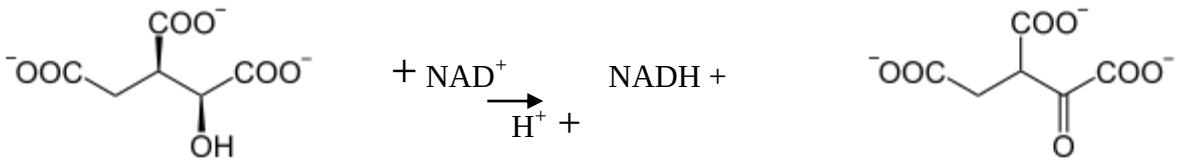
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
1ITW	A	ICT			
	B				

Table 12: 3D representation of the binding motifs associated with Isocitrate Dehydrogenase for the case of the PDB entry 1ITW (chain A & B).

II. 4. α-Ketoglutarate Dehydrogenase

II. 4. A. Formation of α-Ketoglutarate:

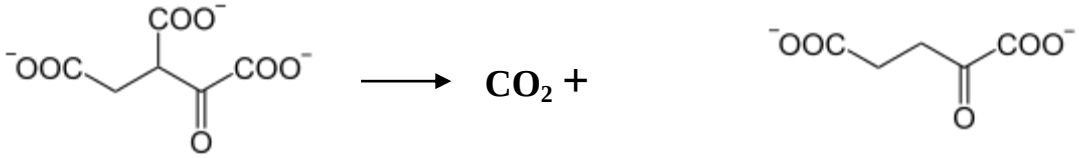
✓ **Step 1:**



Isocitrate — Oxalosuccinate

Figure 45-a: The Oxydation of Isocitrate into oxalosuccinate.

✓ **Step 2:**



Oxalosuccinate — α-Ketoglutarate

Figure 45-b: The Decarboxylation of Oxalosuccinate into the final product α-Ketoglutarate.

❖ **Overall reaction:**

Figure 46: Formation of alpha-Ketoglutarate by the intervention of the α-Ketoglutarate Dehydrogenase enzyme.

II. 4. B. α-Ketoglutarate Dehydrogenase Binding Motifs and their Properties:

The α-Ketoglutarate Dehydrogenase in all of the related PDB entries is represented by 28 protein chains where each is bound to an instant of the Isocitrate analogue; AKG thus the existences of 28 motifs, table 13:

- All motifs describe a mixture of **α-helices** (**H**), **β-stands (S)** and **loop** regions **(L)** expect one motif which is made of a mixture of **α-helices** (**H)** and **β-strands (S)** as shown with case of the entry 1CW4.

α-Ketoglutarate Dehydrogenase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
AKG	3INM	A	LHSHHL	3
		B		
		C		
	4L06	A	LHSSHL	6
		B		
		C		
		D		
		E		
		F		
	4L04	A	LHSSHSL	1
		B	LHSSHL	1
		C	LHSSH	2
		D		
		E	LHSSHL	5
		F		
	4L03	A	LHSSHL	
		B		
		C		
	4KZO	A	LHSHHL	3
		B		
		C		
	4AJR	A	LHSSHHL	2
	4AJC	A	LHSSHHL	

	3O9Z	A	LSLHHSL	4
		B		
		C		
		D		
	1CW4	A	HSSHH	1

Table 13: Ligands bound to **α**-Ketoglutarate Dehydrogenase, binding motifs and the motif's occurrence frequency.

II. 4. C. Motifs 3D-Graphical Representation:

- As seen above, in the case of the Alpha-Ketoglutarate Dehydrogenase, we have **09** PDB entries: 3INM, 4LO6, 4LO4, 4LO3, 4KZO, 4AJR, 4AJC, 3O9Z and binding at the same ligand AKG which is analogues to the natural substrate.
- The images below, table 14, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entries 3INM (Chain A) and 4LO6 (Chain A).
- The 3D representation of all the other binding Motifs related to the other chains of 3INM, 4LO6 and PDB entries (4LO4, 4LO3, 4KZO, 4AJR, 4AJC, 3O9Z, 1CW4) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

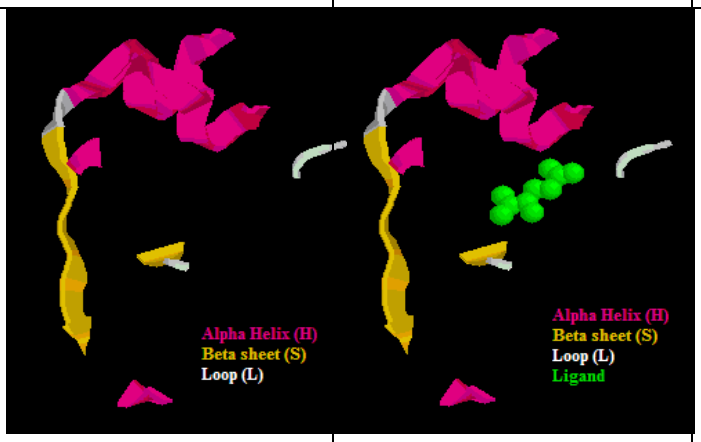
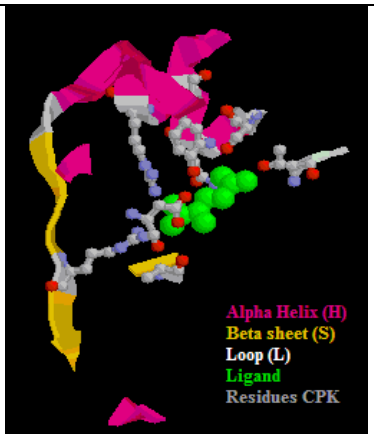
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
3INM	A	AKG			
4L06	A				

Table 14: 3D representation of the binding motifs associated with **α**-Ketoglutarate Dehydrogenase for the cases of the PDB entries 3INM and 4L06 (both chain A).

II. 5. Succinyl-CoA-Synthetase

II. 5. A. Formation of Succinyl-CoA:

Reaction:

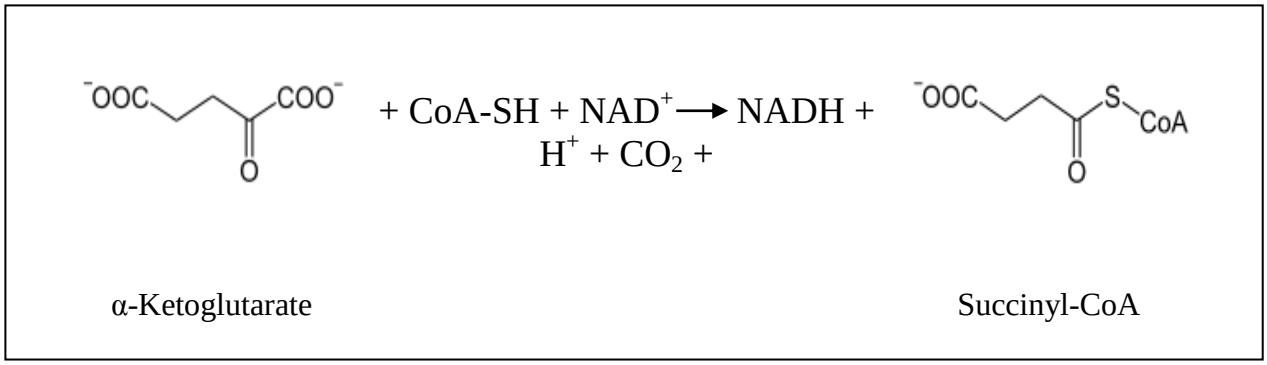


Figure 47: Formation of Succinyl-CoA by the intervention of Succinyl-CoA-Synthetase enzyme.

II. 5. B. Succinyl-CoA-Synthetase Binding Motifs and their Properties:

The Succinyl-CoA-Synthetase in all of the related PDB entries is represented by 07 protein chains where each is bound to an instant of the α-Ketoglutarate analogue; SCA and SDX thus the existence of 07 motifs, **table 15**:

- Four motifs describe a mixture of **α-helices (H), β-strands (S)** and **loop** regions **(L).** The rest of the motifs are a mixture of **β-strands (S)** and **loop** regions **(L)**.

Succinyl-CoA-Synthetase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
SCA	2BWO	A	LSLHSLL	4
		B		
		D		
		E		
SIN	2BWN	A	SLSLLL	3
		D		
		E		

Table 15: Ligands bound to Succinyl-CoA-Synthase, binding motifs and the motif's occurrence frequency.

II. 5. C. Motif 3D-Graphical representation:

- As seen above, in the case of the **Succinyl-CoA-Synthetase**, there are **02** PDB entries: 2BWO, 2BWN binding a total of two Ligands that are analogues to the natural substrate.
- The images below, table16, represent the RasMol 3D-graphical representation of the binding motifs associated with the PDB entries 2BWO (Chain A) and 2BWN (Chain A).
- The 3D representation of all the other binding Motifs related to the other chains of PDB entries (2BWO, 2BWN) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

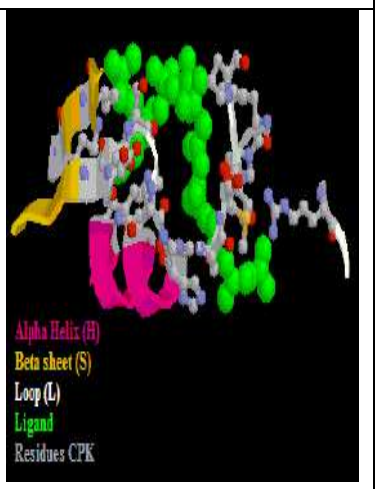
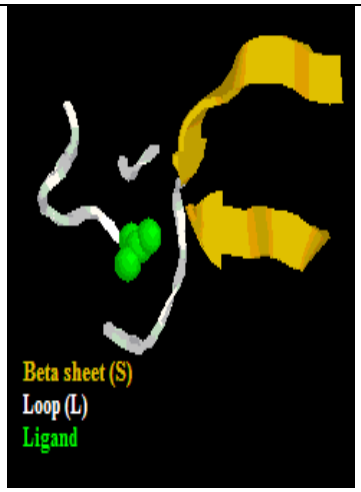
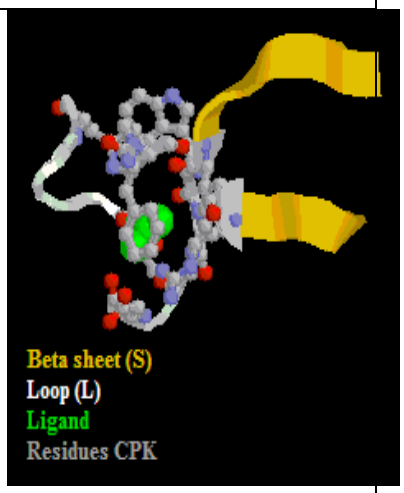
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
2BWO	A	SCA			
2BWN	A	SIN			

Table 16: 3D representation of the binding motifs associated with Isocitrate Dehydrogenase for the cases of the PDB entries 2BWO and 2BWN (both chain A).

II. 6. Succinate Dehydrogenase

II. 6. A. Formation of Succinate:

Reaction:

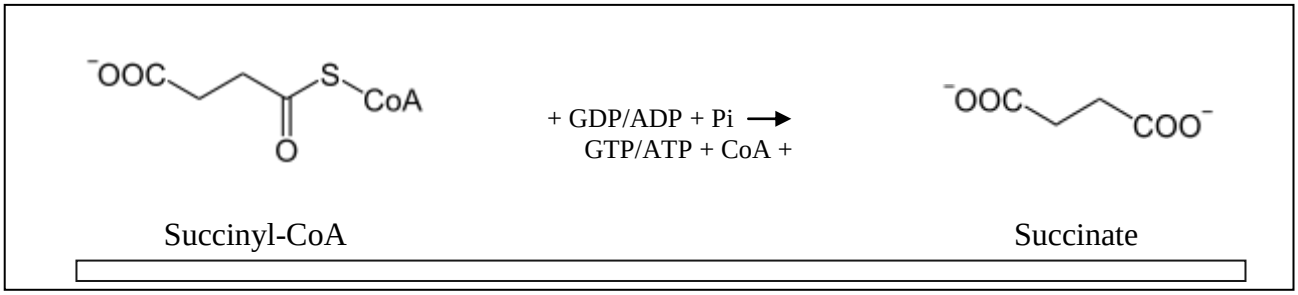


Figure 48: Formation of Succinate by the intervention of Succinate Dehydrogenase enzyme.

II. 6. B. Succinate Dehydrogenase Binding Motifs and their Properties:

The Succinate Dehydrogenase in all of the related PDB entries is represented by 06 protein chains where each is bound to an instant of the Citrate analogue; SIN thus the existence of 06 motifs, table 17:

- Two motifs describe a mixture of **α-helices (H), β-strands (S)** and **loop** regions **(L).** The other motifs are divides into the following:
  - Two motifs are mixture of **α-helices (H)**, and **loop** regions **(L)**.
  - Two motifs are mixture of **β-strands (H)** and **Loop** regions **(L)**.

Succinate Dehydrogenase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Frequency**
SIN	2W8Q	A	HLHSHLL	1
	4LH2	A	LHLL	2
		B		
	2DJL	A	LSLSLL	2
		B		
	3JU8	A	LHLLS	1

Table 17: Ligands bound to Succinate Dehydrogenase, binding motifs and the motif's occurrence frequency.

II. 6. C. Motif 3D-Graphical representation:

- As seen above, in the case of the Succinate Dehydrogenase, there are **04** PDB entries: 2W8Q, 4LH2, 2DJL, 3JU8 binding at the same ligand SIN which is analogues to the natural substrate.
- The images below, table 18, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entry 4LH2.
- The 3D representation of all the other binding Motifs related to the other PDB entries (2W8Q, 2DJL, and 3JU8) and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

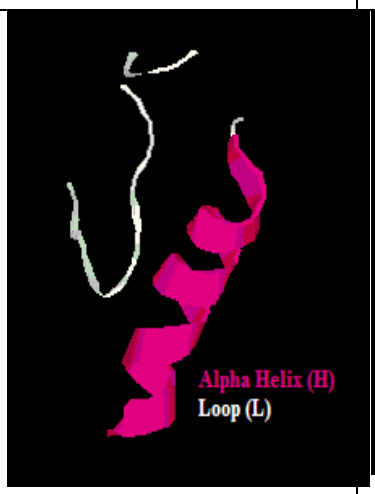
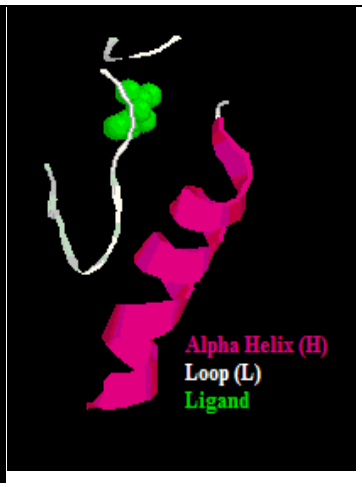
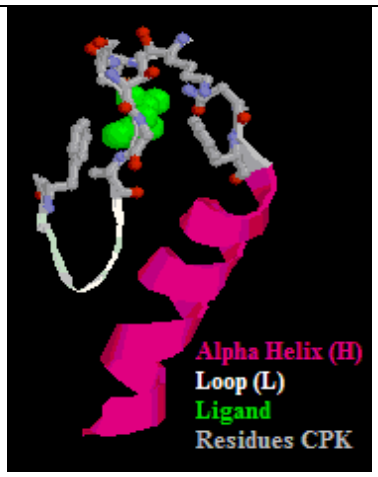
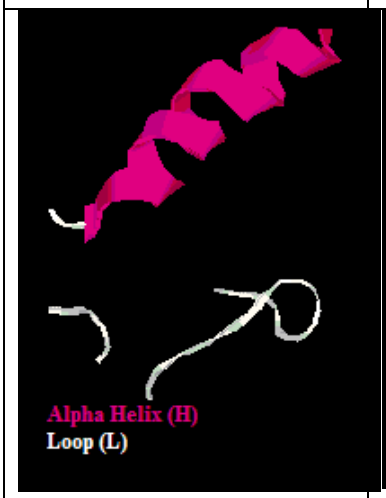
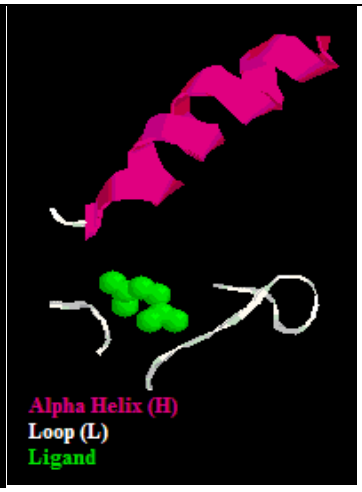
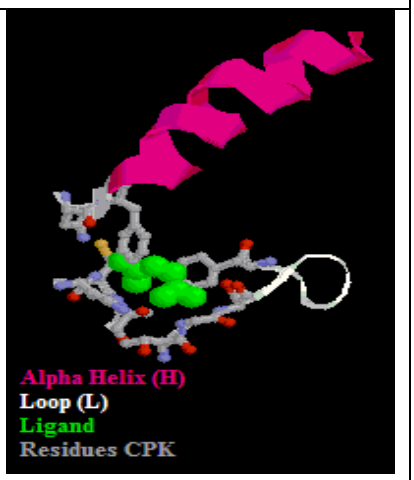
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
4LH2	A	SIN			
	B				

Table 18: 3D representation of the binding motifs associated with Succinate Dehydrogenase for the cases of the PDB entry 4LH2 (chain A & B).

II. 7. Fumarase

II. 7. A. Formation of Fumarate:

Reaction:

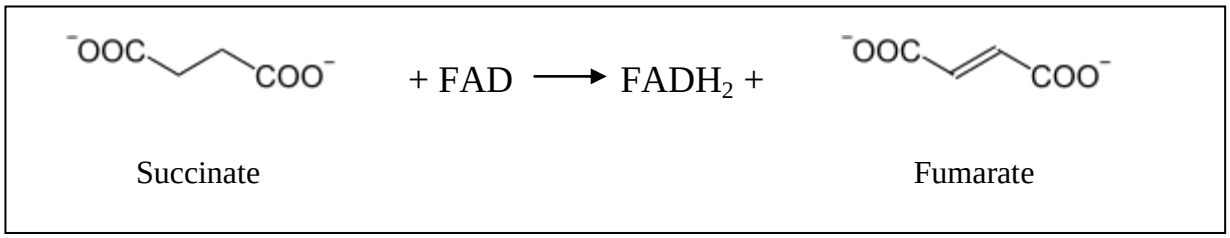


Figure 49: Formation of Fumarate by the intervention of Fumarase enzyme.

II. 7. B. Fumarase Binding Motifs and their Properties:

The Fumarase in the PDB entry 4APB is represented by 04 protein chains where each is bound to an instant of the Succinate analogue; FUM thus the existence of 04 motifs, table 19:

- All of the binding motifs in this case describe a motif made entirely of **loop** regions (**L**) i.e. there is no secondary structure elements involved in the binding of the ligand in this case.

Fumarase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrence Fréquence**
FUM	4APB	A	L	4
		B		
		C		
		D		

Table 19: Ligands bound to Fumarase, binding motifs and the motif's occurrence frequency.

II. 7. C. Motif 3D-Graphical representation :

- As seen above, in the case of the Fumarase, there is **01** PDB entry: 4APB binding a single ligand FUM which is an analogue to the natural substrate.
- The images below, table 20, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entry 4APB (Chain A).
- The 3D representation of all the other binding Motifs related to the other chains of PDB entry, 4APB, and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

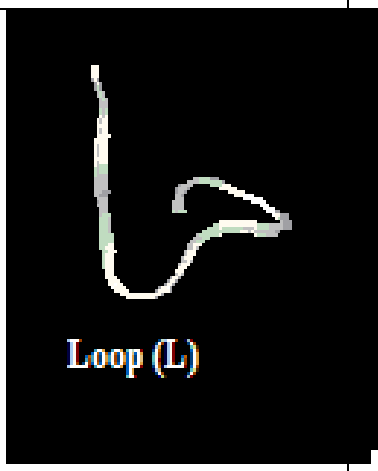
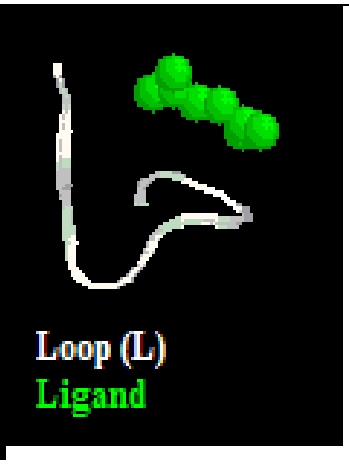
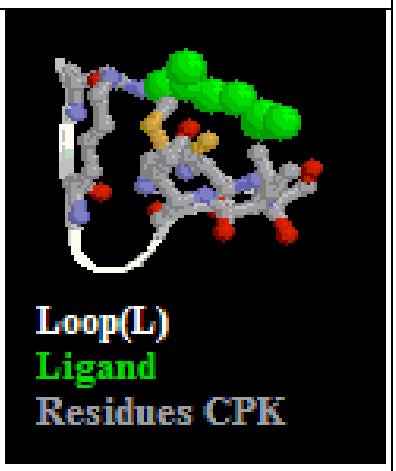
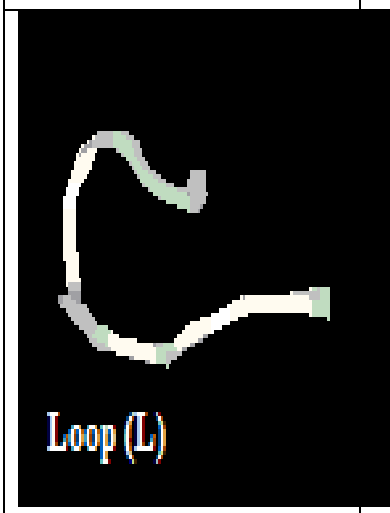
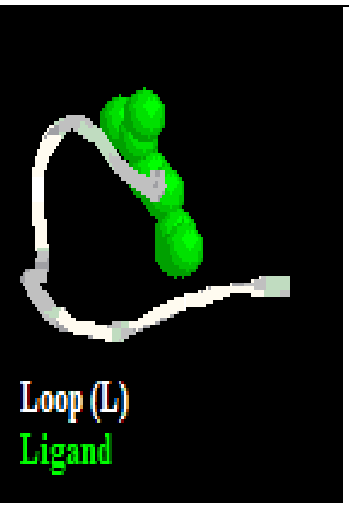
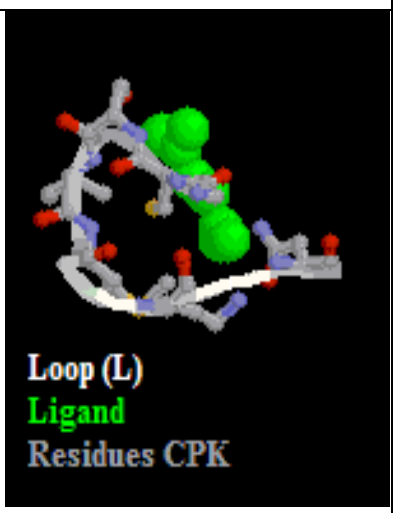
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
4APB	A	FUM			
	D				

Table 20: 3D representation of the binding motifs associated with Fumarase for the case of the PDB entry 4APB (chains A & D).

II. 8. Malate Dehydrogenase

II. 8. A. Formation of L-Malate:

Reaction:

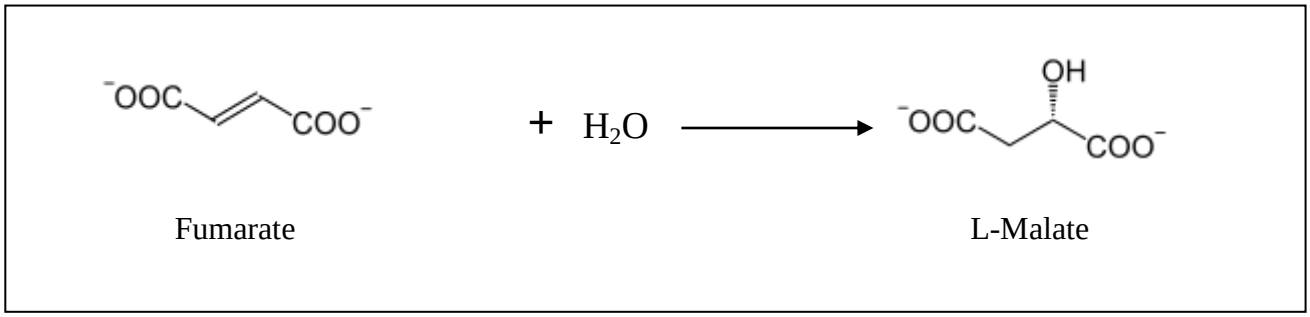


Figure 50: Formation of Malate by the intervention of Malate Dehydrogenase enzyme.

II. 8. B. Malate Dehydrogenase Binding Motifs and their Properties:

The Malate Dehydrogenase in the PDB entry 2DFD is represented by 04 protein chains where each is bound to an instant of the Malate analogue; MLT thus the existence of 04 motifs, table 21.

- All motifs describe a mixture of **alpha-helices** (**H**) and **loop** regions (**L**).

Malate Dehydrogenase				
Ligand	**PDB ID**	**Chain**	**Motif**	**Occurrences Frequency**
MLT	2DFD	A	LHLHLHL	4
		B		
		C		
		D		

Table 21: Ligands bound to Malate Dehydrogenase, binding motifs and the motif's occurrence frequency.

II. 8. C. Motif 3D-Graphical representation:

- As seen above, in the case of the Malate Dehydrogenase, there is **01** PDB entry: 2DFD binding a single ligand MLT which is an analogue of the natural substrate.
- The images below, table 22, represent the RasMol 3D-graphical representation of the binding motif associated with the PDB entry 2DFD (Chain A & Chain B).
- The 3D representation of all the other binding Motifs related to the other chains in the PDB entry, 2DFD, and their bound Ligands have been created and are stored in the Flat-Files database and in the online version CacSFMs, see also **Index II**.

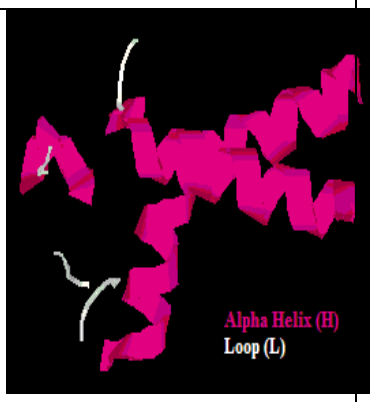
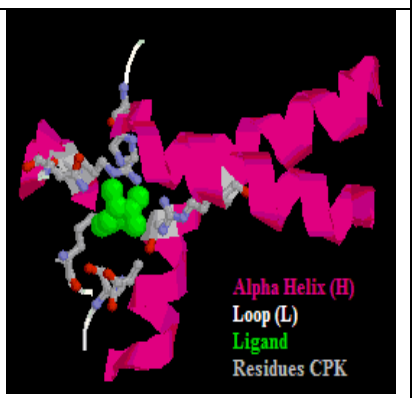
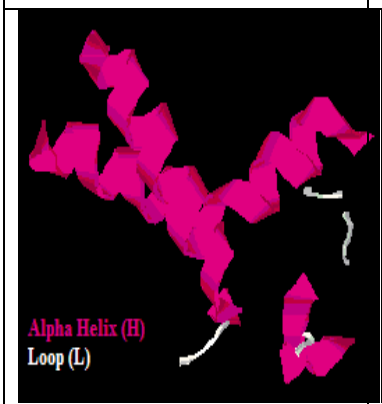
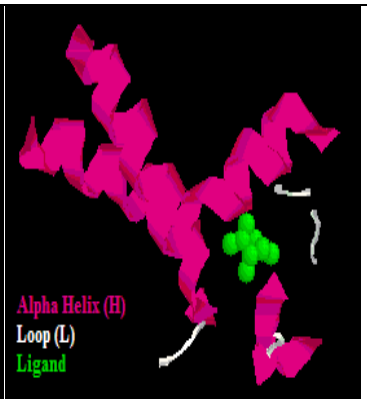
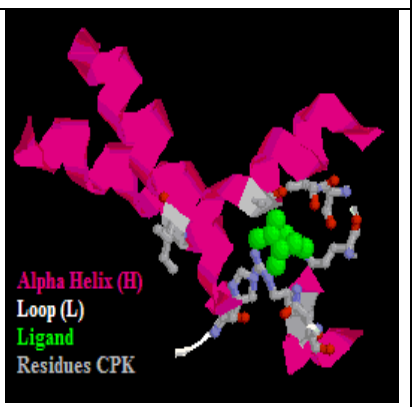
PDB ID	Chain	Ligand	Motifs	Motif+Ligand	Motif+Ligand+Residues
2DFD	A	MLT			
	B				

Table 22: 3D representation of the binding motifs associated with Malate Dehydrogenase for the case of the PDB entry 4APB (chains A & B).

III. Binding Residue types and Ligands:

The binding environment details calculated for the 11 Ligands, as reported in **Chapter II** and **Index III**, have been used to explore the types of the residues (amino acids while peptide bound each other in the protein chain) that **actually do the act of binding the Ligands** in the structural motifs described above.

The table below, no. 23, summarizes the types of residues per bound Ligands:

Ligand ID	Amino acides
CIT	HIS ; ASN ;GLY ;ARG ;VAL ;ASP ;PHE ;MET ;GLU
FLC	GLU ; ASN ; HIS ;GLY ;ARG ;ASP ;PHE ; LYS
SDX	HIS ;PRO ;LEU ;GLY ;ALA ;ARG ;MET ;PHE ;ILE ;ASN ;ASP ;LYS
TRA	GLN ; ALA ; THR ; HIS ;ASP ;SER ;ILE ;ARG ;MET
ATH	GLN ; ALA ; THR ; HIS ;ASP ;SER ;ILE ;ARG ;SER
ICT	SER ;ASN ;ARG ;LYS ;MET ;ASP ;TYR ;ALA ;THR ;VAL ;LEU ;HIS ;GLY ;GLU
AKG	THR ;SER ;ASN ;ARG ;TYR ;ASP ;ALA ;GLU ;LEU ;LYS ;ILE ;HIS ;PHE ; VAL
SCA	ARG ;SER ;ASP ;ASN ;HIS ;ALA ;ILE ;LYS ;MET ;PHE ;PRO ;THR ;LEU
SIN	ARG ;PHE ; PRO ;VAL ;TRP ;CYS ;GLY ;ASP ;THR ;HIS ;TYR ;GLU ;ALA ;SER ;ASN ;LYS ;MET ;LEU
FUM	GLY ; CYS ; SER ; ILE ; MET ; LYS ; ASN
MLT	ARG ; ASN ;HIS ;GLY ;SER ;ALA ;LEU

Table 23: Residues type shown arranged by their bound Ligands. The residues are colored per hydrophobicity, see **figure 51**.

Based on the information above it seems that the binding residues governing the chemical environment of the binding motifs are mostly hydrophilic with some low level of hydrophobicity. For implications of this properties see **General Conclusion**.

Water Affinity		PK		Name
Highly	1	3.1		Isoleucine
Hydrophobic	2	2.5		Phenylalanine
	3	2.3		Valine
	4	2.2		Leucine
	5	1.1		Methionine
	6	1.0		Tryptophan
	7	1.0		Alanine
	8	0.67		Glycine
	9	0.17		Cysteine
	10	0.08		Tyrosine
	11	-0.29		Proline
	12	-0.75		Threonine
	13	-1.1		Serine
	14	-1.7		Histidine
	15	-2.6		Glutamate
	16	-2.7		Asparagine
	17	-2.9		Glutamine
	18	-3.0		Aspartate
Highly	19	-4.6		Lysine
Hydrophilic	20	-7.5		Arginine

Figure 51: Amino acids hydrophicity scale and coloring **(Kaiser E, 1970).**

IV. Motifs classification and Ligands Binding Tendency:

The Analysis curried out above including the properties of the motifs suggest that the binding motifs associated with the enzymes involved in the CAC pathway can be classified into the following families:

IV. 1. α/Loop family:

This family includes all motifs that contain α-Helix (**H**) and Loop (**L**) elements, like the following:

LLHLHLH, LHLLHHH, LHLLHLH, LHLLHLHH, LHLHHH, LLLHLHH, LHLHLHLHH, LHLHLHHHLHH, HLLHLLHL, LHHLHLH, LLHL, LHLL, LHLHLHL

The Ligands that tend to bind this family of motifs are: CIT, FLC, SDX, TRA, ICT, SIN and MLT.

IV. 2. α/β family:

This family includes all motifs that contain α-Helix (**H**) and β-strand (**S**) elements, like the following:

SSHH, HSSHH

The Ligands that tend to bind this family of motifs are: ICT and AKG.

IV. 3. β/Loop family:

This family includes all motifs that contain β-strand (**S**) and Loop (**L**) elements, like the following:

SLSLLL, LSLSLL

The Ligands that tend to bind this family of motifs are: SIN.

IV. 4. α/β/Loop family:

This family includes all motifs that contain α-Helix (**H**), β-strand (**S**) and Loop (**L**) elements, like the following:

SLLLHHLLL, SHLLHLLLL, LHSLHHHL, LHSSHHL, LHSSHHSL, LHSSHH, LHSHHL, LHSSHL, LHSSHSL, LHSSH, LSLHHSL, LSLHSLL, HLHSHLL, LHLLS

The Ligands that tends to bind this family of motifs are: TRA, ATH, ICT, SCA and SIN.

IV. 5. Loop-only family:

This family includes all motifs that made of Loop (**L**) elements, like the following:

LL, L

The ligands that tend to bind this family of motifs are: FLC and FUM.

Note that the water (W) character has been removed from the motifs annotation, in above classification, for clarity and the fact that the water molecules are not part of protein structure.

The classification above show clearly that contrary to the common misunderstanding that loop regions serve only as connectors between secondary structure elements (usually endowed with biological function) in the overall three-dimensional structure of proteins. The **Loop-only family** of the Ligands binding motifs gives the indication that loop regions may play important roles in protein biological functions.

The information provided by the motifs classification, done above, versus the Ligands binding tendency and the properties of the residues responsible for the actual binding of the Ligands, outlined also above, can be instrumental in the process of designing new drugs depending on the type of binding motif targets and the binding tendency of the drugs, for example in cases where *de novo* drug design is needed for treating pathogenic metabolism situations.

General Conclusion

General Conclusion

This project has set out to try and contribute in understanding the basis of Structure-Function relationship in macromolecules; protein is the case of this study. This relationship seems to be **coded by the amino acids** that compose enzymes and proteins in general though the **protein folding** process.

As shown in the various analysis and deductions made in the **Results and Discussions (Chapter III)**, this project has identified, defined and characterized the protein structural elements dubbed here as the binding motifs together with the binding residues (amino acids) that are directly involved in the ligand binding process and hence the function of the enzymes associated with the citric acid cycle.

The structural elements (α-helices, β-strand and loops) in the define and characterized binding motifs are seen by this study as providing the structural support on which the functional elements, i.e. the residues, can reside.

The definition of the ligand binding sites in the form of structural motifs and storing it into a the database can be very useful in finding similar motifs in other protein which may be of the same and/or different function. This helps lots of studies for protein classifications, taxonomy, phylogenetics and homology molecular modeling.

This study has also deducted the important role of secondary structure and more importantly the crucial role of the non-secondary structure regions (loops) in the biological function of protein.

However, the conclusions related to the motifs classification and Ligands binding tendency and their potential usage are limited to the enzymes of the CAC pathway under study in this project. to find out if the conclusions can be generalized, there is the need to corroborate this kind of study by analyzing larger sets of enzymes associated with different types of metabolic pathways.

References

References

- **Aitken. A**. Protein consensus sequence motifs.(Mol. Biotechnol. **1999**, 12:241–253**).**
- **Andreeva A, Howorth D, Chandonia J, Brenner S, Hubbard T, Chothia C, Murzin A**: Data growth and its impact on the SCOP database: new developments. Nucleic Acids Res **2008**, ***36:*** D419–425. 10.1093/nar/gkm993.
- **Benjamin Cummings**, Biology, Eight Edition, Pearson, **2008**.
- **Berman H-M**, **K. Henrick**, **Nakamura H (2003)** Announcing the worldwide Protein Data Bank Nature Structural Biology 10 (12): 980.
- **Berman H-M**, **Westbrook J**, **Feng Z**, **Gilliland G**, **Bhat T**, **Weissig H**, **Shindyalov I**, **Bourne P**: The Protein Data Bank**.** Nucleic Acids Research **2000**, ***28:*** 235–242. 10.1093/nar/28.1.235.
- **Deane, C.M. et al**. Carbonyl-carbonyl interactions stabilize the partially allowed Ramachandran conformations of asparagines and aspartic acid. Protein Eng. **1999**, 12:1025–1028.
- **Drenth J**. Principles of Protein x-ray Crystallography, 3rd edition. **2007**, Springer Science + Business Media, LLC. pg. 14.
- **Drenth J (1994)** Principles of Protein X-ray Crystallography. Springer-Verlag, New York.
- **Drenth J (1994)** Principles of Protein X-ray Crystallography. Springer-Verlag, New York (pages: 332).
- **Fernande D. Rochon**, **Viorel Buculei**, **Inorganica Chimica** Acta 357 **(2004)** 2218–2230.
- **Frank J**. Single-particle reconstruction of biological macromolecules in electron microscopy--30 years. Q Rev Biophys. **2009**; 42:139–58.
- **Gellman**, **S.H**. Minimal model systems for beta sheet secondary structure in proteins. Curr.Opin.Chem.Biol. **1998**, 2:717–725.
- **Golovin A**, **Dimitropoulos D**, **Oldfield T**, **Rachedi A**, **Henrick K**., MSDsite: a database search and retrieval system for the analysis and viewing of bound Ligands and active sites. Proteins. **2005** Jan 1; 58(1):190-9.

References

- **Greene L-H**, **Lewis T-E**, **Addou S**, **Cuff A**, **Dallman T**, **Dibley M**, **Redfern O**, **Pearl F**, **Nambudiry R**, **Reid A**, **Sillitoe I**, **Yeats C**, **Thornton JM**, **Orengo CA**. The CATH domain structure database: new protocols and classification levels give a more comprehensive resource for exploring evolution. Nucleic acids research **2007**, (35 Database):D291–7. 10.1093/nar/gkl95.
- **Hector Viadiu**. Protein Composition and Structure, UCSD, **2011**.
- **Hoenger A**, **Bouchet-Marquis C**. Cellular tomography. Adv Protein Chem Struct Biol. **2011**; 82:67–90.
- **Kaiser E**, **Colescott R-L**, **Bossinger C-D**, P.I. *Analytical Biochemistry*, *Volume 34*, *Issue 2*, ***April 1970***, *Pages 595-598*
- **Milner-White**, **EJ (1987)**. "Beta bulges within loops as recurring features of protein structure". Biochimie Biopsy's Acta 911 (2): 261–265.doi:10.1016/0167-4838 (87)90017-3. PMID 3801498.
- **Némethy**, **George**; **Printz**, **Morton P (1972)**. "The-Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the β-Turn". Macromolecules 5 (6): 755. Doi: 10.1021/ma60030a017.
- **Perticaroli S**, **Nickels JD**, **Ehlers G**, **O'Neill H**, **Zhang Q**, **Sokolov AP (2013)**. "Secondary structure and rigidity in model proteins". Soft Matter **9** (40): 9548.
- **Rachedi A**, Masters Teaching Material at Saida University, **2013**, http://www.bioinformaticstools.org/masters/
- **Rachedi A**, Sequence, Structure and Function Server, **2011**, http://www.bioinformaticstools.org/ssfs/
- **Rachedi A**, Bioinformatics Tools Sever, **2012**, http://www.bioinformaticstools.org
- **Raven**, **Peter**. Biology. Boston: Twayne Publish ers, **2005**. http://www.incolor.com/mcanaday/Krebdiag.GIF (image).
- **Rhodes G (2000)** Crystallography made Crystal Clear, 2nd edn. Academic Press.
- **Roger Sayle** v2.6 features added May, **1997**: RasMol v2.6Quick Reference Card(c) Copyright1994. http://www.bio.cmu.edu/Courses/BiochemMols/RasFrames/TOC.HTM
- **Tyson and al**, Functional motifs in biochemical reaction networks. Annul Rev Phys Chem. **2010**; 61: 219–240.

References

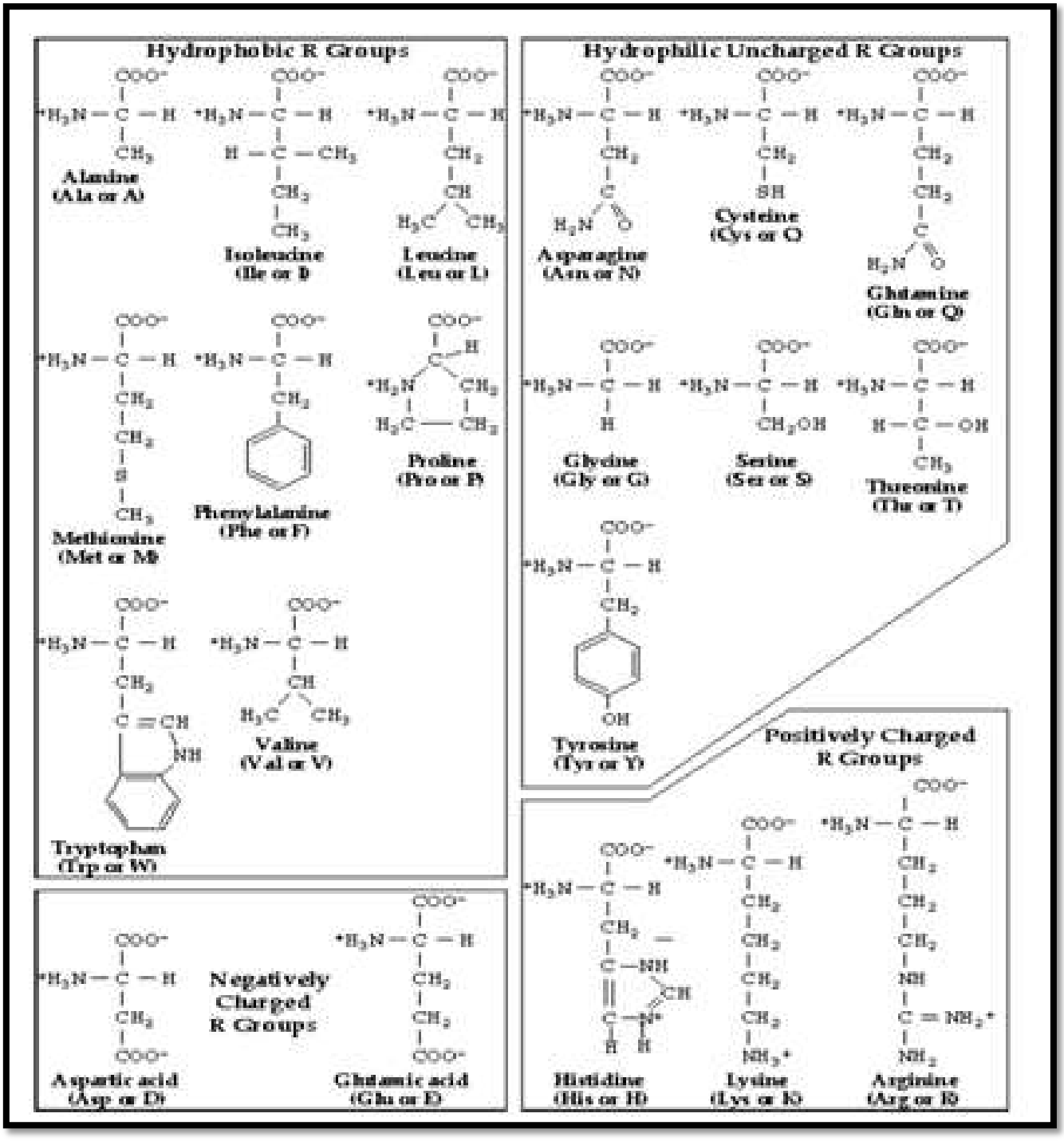
- **Zeroka D**, **Hameka F-H**, Calculation of magnetic shielding constants of diatomic Molecules. I. General theory, J. Chem. Phys. 1966, 45, 300–311

Index

Index I

Index I

Structure of the 20 amino acids:

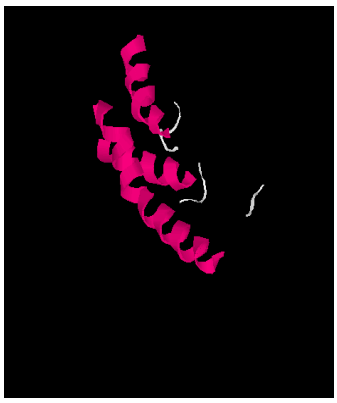


Structure of the 20 (alpha-amino acids used in synthe proteins sizing. Below each amino acid are its 3-and 1-letter abbreviation (derived from **Elseth** and **Baumgardner**, Principle of Modern Genetics **(1995)**.West Pub)

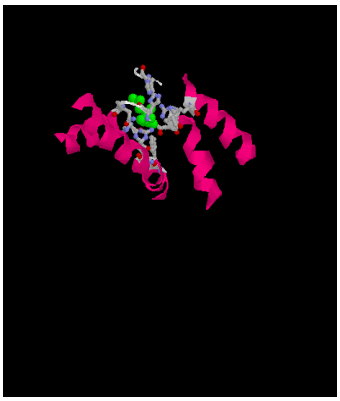
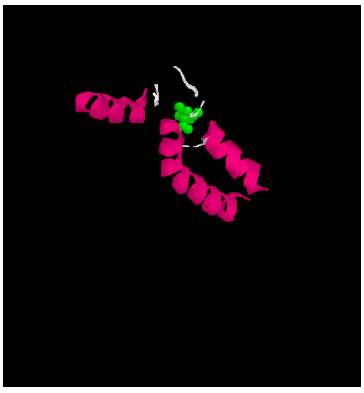
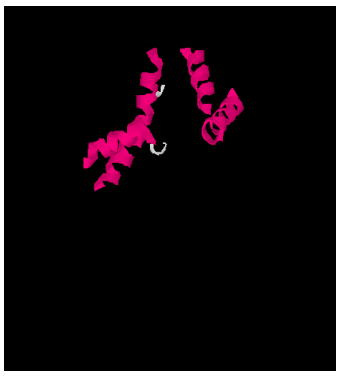
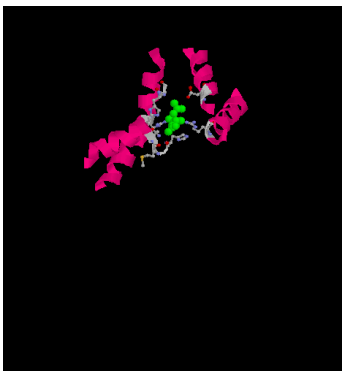
Index II

Index II

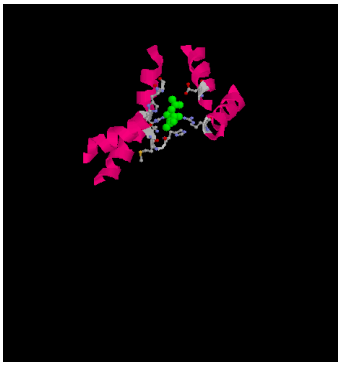
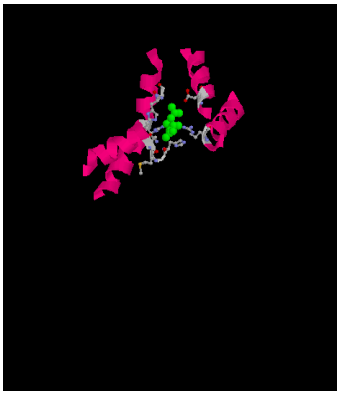
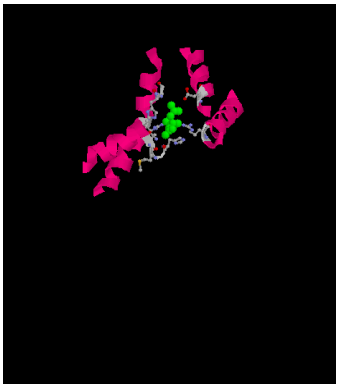
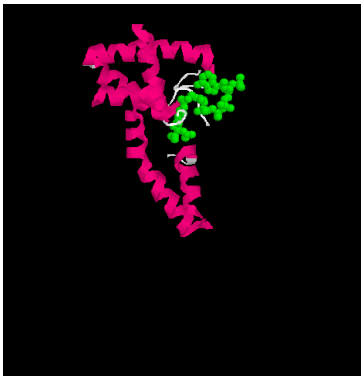
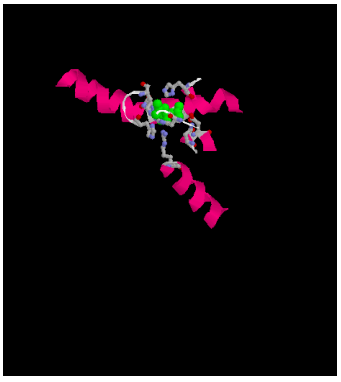
01-Citrate Synthetase:

Codes	Motifs	Motifs + Ligand	Motifs + Ligand + Residues
1AJ8-A-CIT-1000			
1AJ8-B-CIT-2000			
1IXE-A-CIT-405			

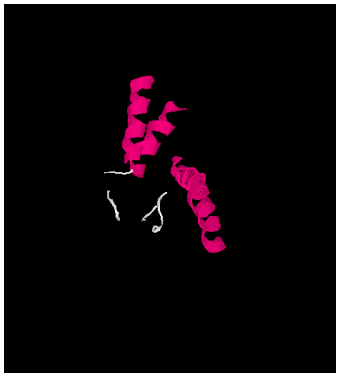
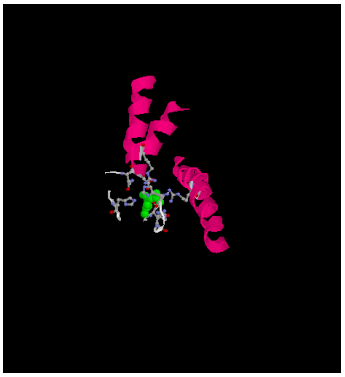
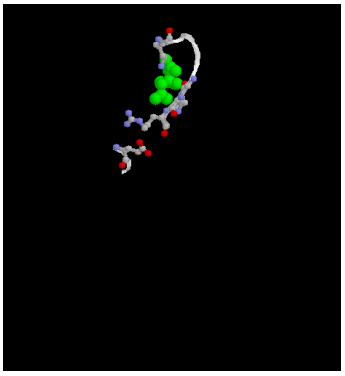
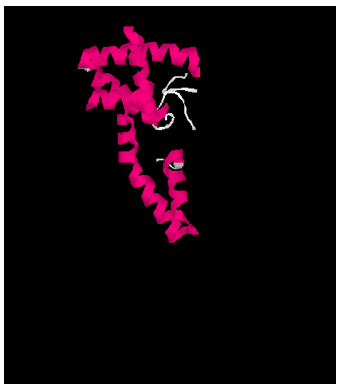
Index II

1IXE-B-CIT-406			
1IXE-C-CIT-407			
1IXE-D-CIT-408			
2C6X-A-CIT-1366			

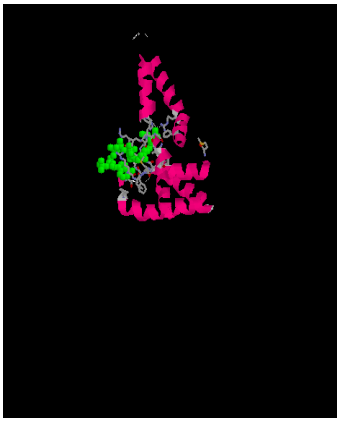
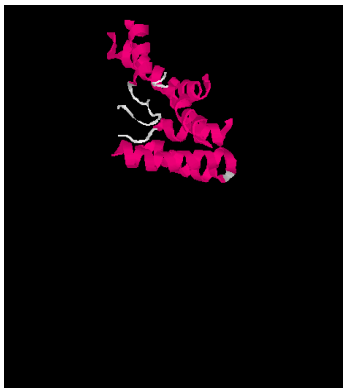
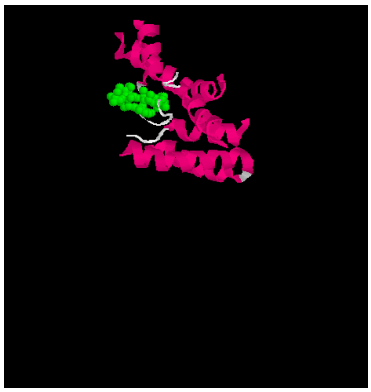
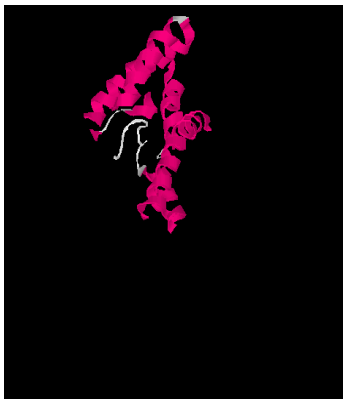
Index II

2C6X-B-CIT-1366			
2C6X-C-CIT-1366			
2C6X-D-CIT-1366-C			
6CSC-A-CIT-701			

Index II

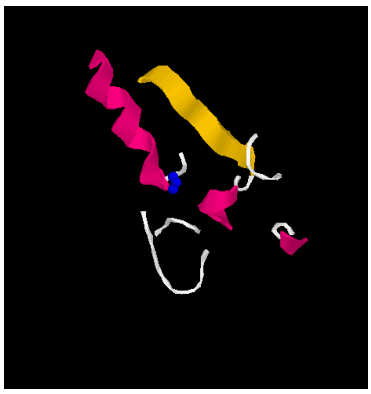
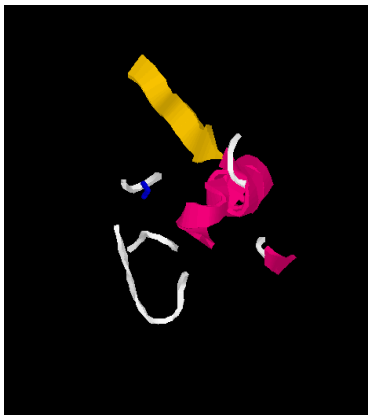
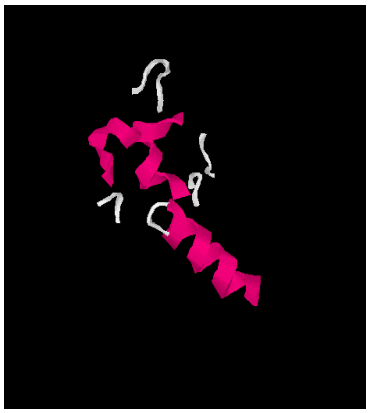
6CSC-B-CIT-701			
2P2W-A-FLC-402			
2P2W-A-FLC-401			
2R9E-A-SDX-700			

Index II

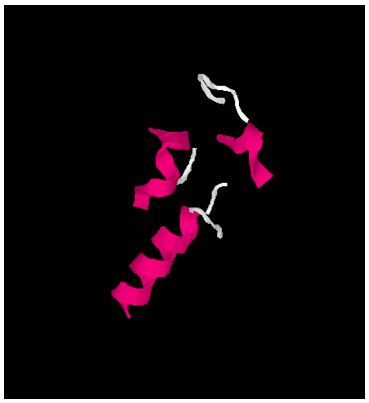
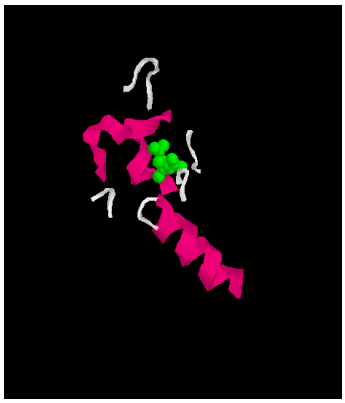
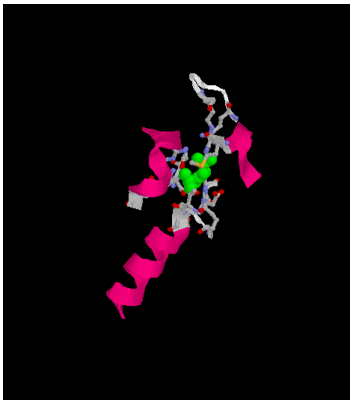
2R9E-B-SDX-701			
2R9E-C-SDX-702			
2R9E-D-SDX-703			

Index II

02-ACONITASE:

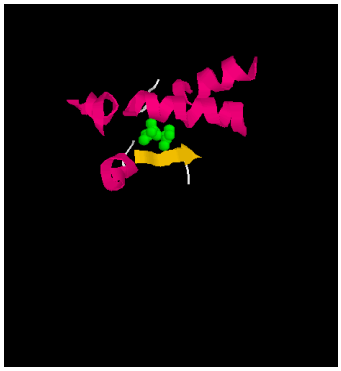
CODES	MOTIFS	MOTIFS + LIGAND	MOTIFS + LIGAND + RESIDUES
1ACO-A-TRA-755			
1FGH-A-ATH-755			
1L5J-A-TRA-866			

Index II

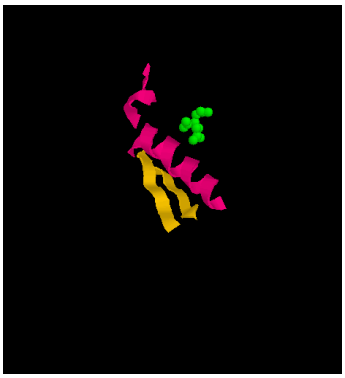
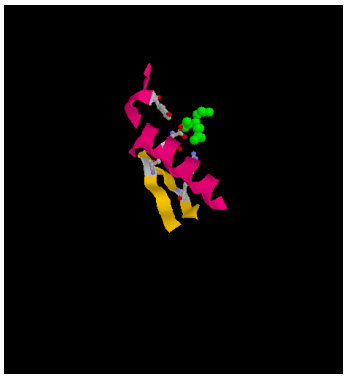
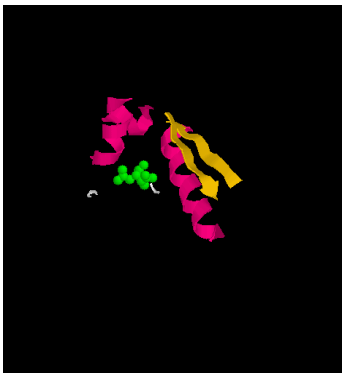
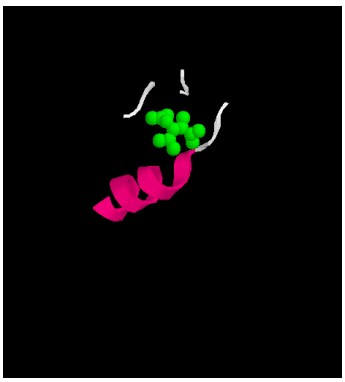
1L5J-B-TRA-867			

Index II

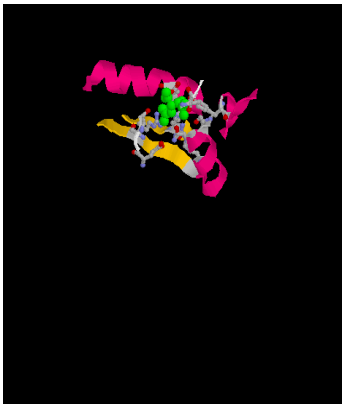
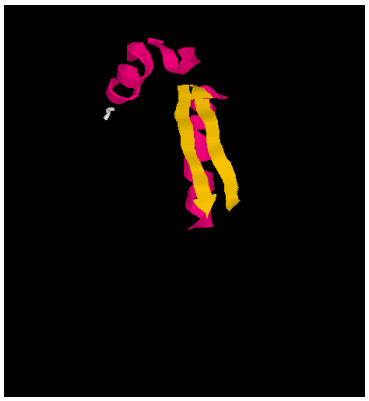
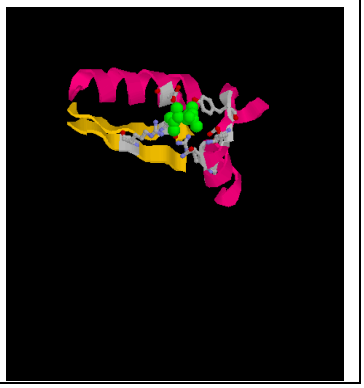
03-ISOCITRATE DEHYDROGENASE :

CODES	MOTIFS	MOTIFS + LIGAND	MOTIFS + LIGAND + RESIDUES
1ITW-A-ICT-743			
1ITW-B-ICT-744			
1ITW-C-ICT-743			

Index II

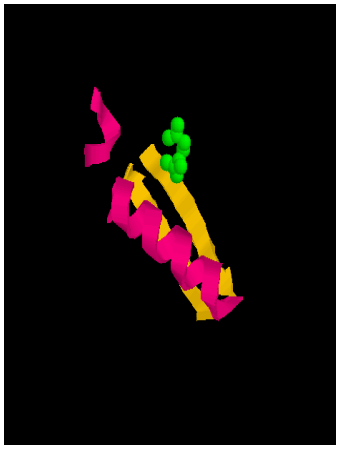
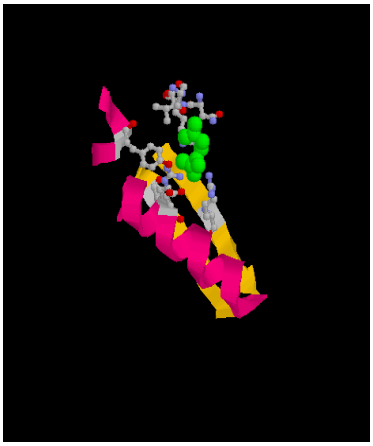
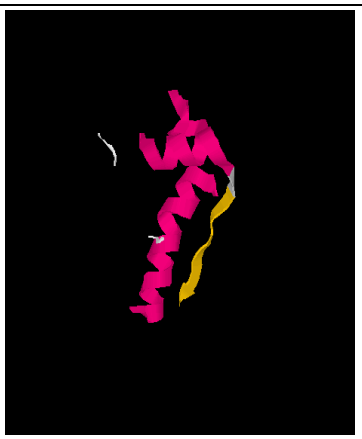
1ITW-D-ICT-743			
1XKD-A-ICT-1002			
1XKD-B-ICT-1004			
2UXR-A-ICT-1404			

Index II

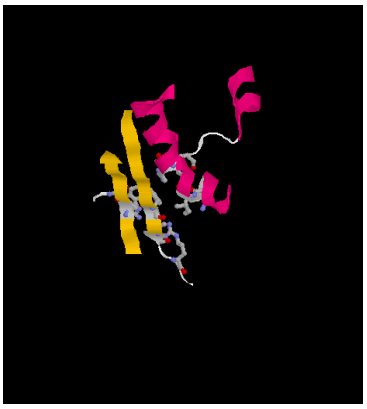
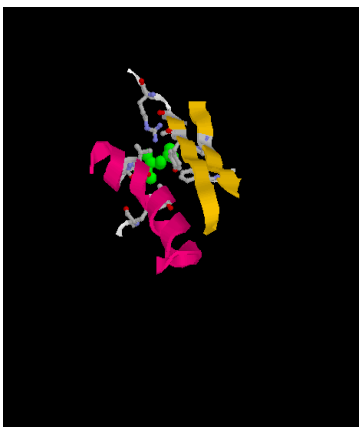
2UXR-B-ICT-1397			
4AJ3-A-ICT-1418			
4AJB-A-ICT-502			
4BNP-A-ICT-502			

Index II

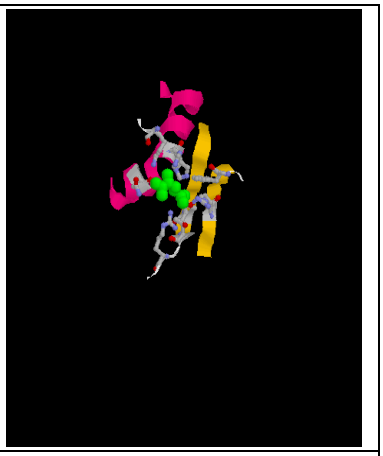
04- α- Ketoglutarate Dehydrogenase :

CODSE	*MOTIFS*	*MOTIFS + LIGAND*	*MOTIFS + LIGAND + RESIDUES*
1CW4-A-AKG-417			
3INM-A-AKG-511			
3INM-B-AKG-511			

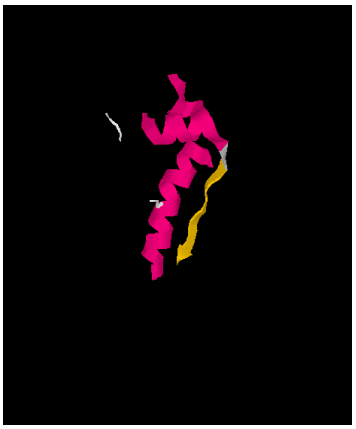
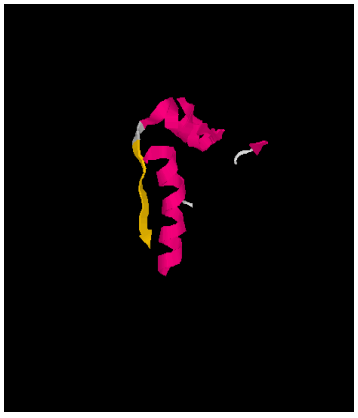
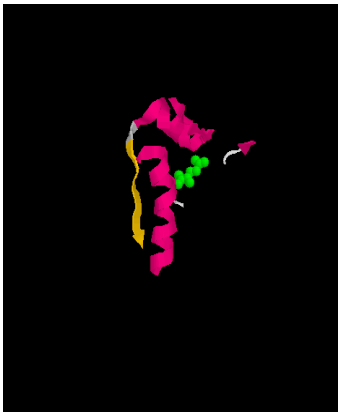
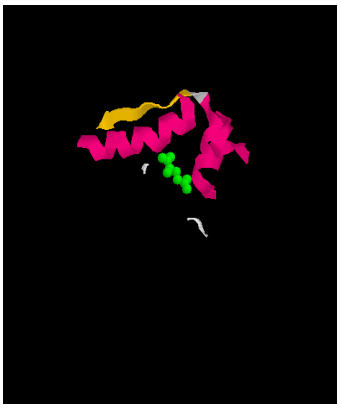
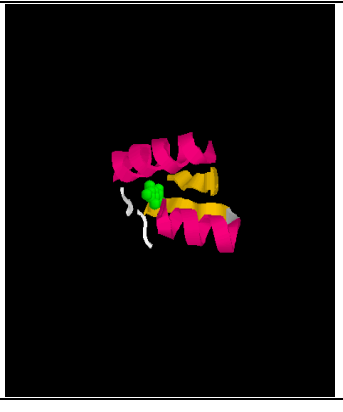
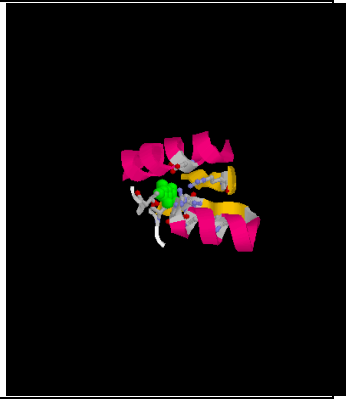
Index II

3INM C-AKG-511			
309Z-A-AKG-312			
309Z-B-AKG-312			
3O9Z-C-AKG-312			

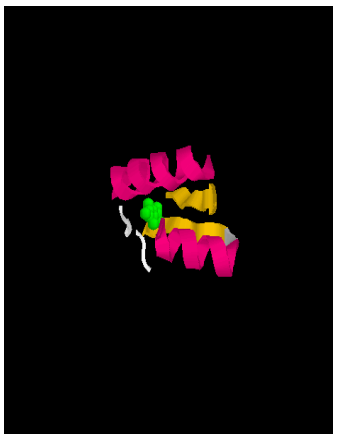
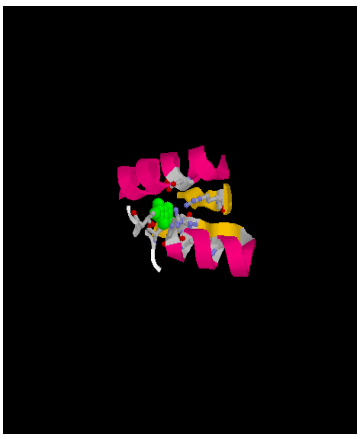
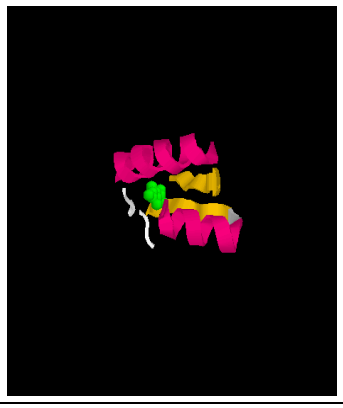
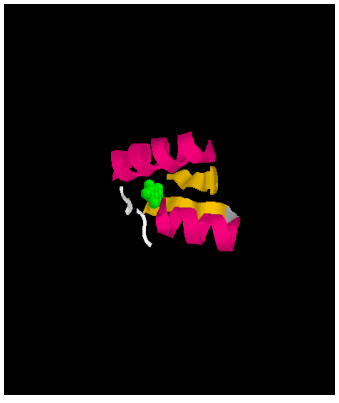
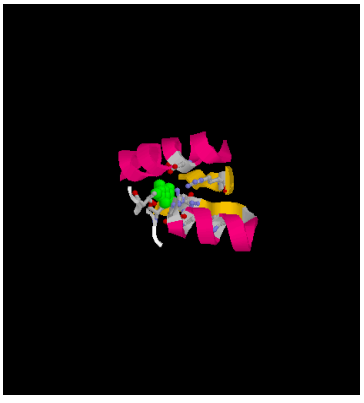
Index II

3O9Z-D-AKG-312			
4AJC-A-AKG-502			
4AJ4-A-AKG-1418			

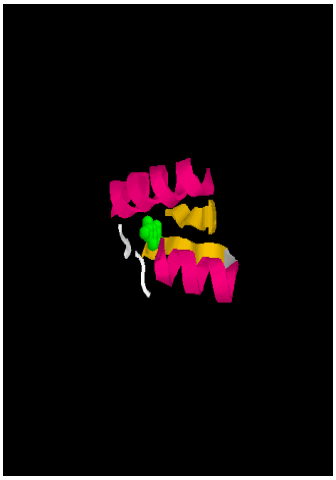
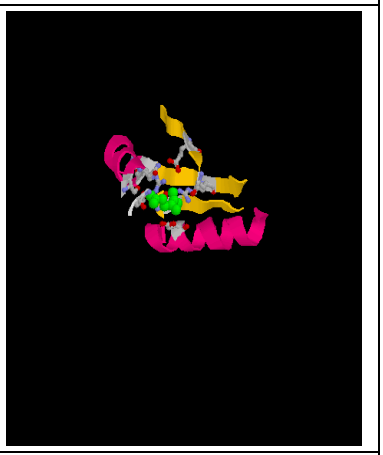
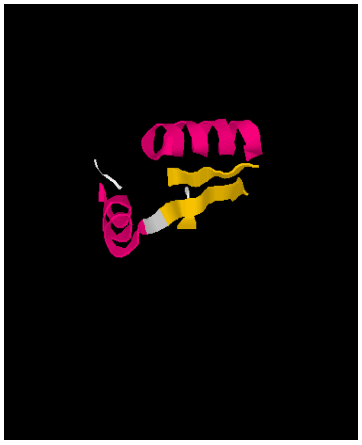
Index II

4KZO-A-AKG-503			
4KZO-B-AKG-503			
4KZO-C-AKG-503			
4L06-A-AKG-503-a			

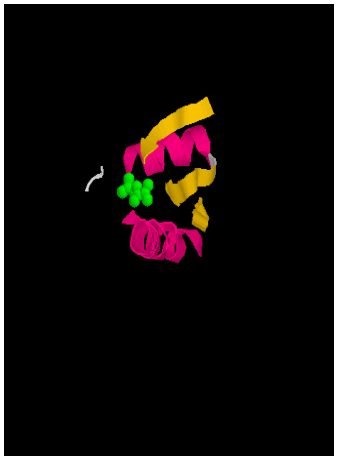
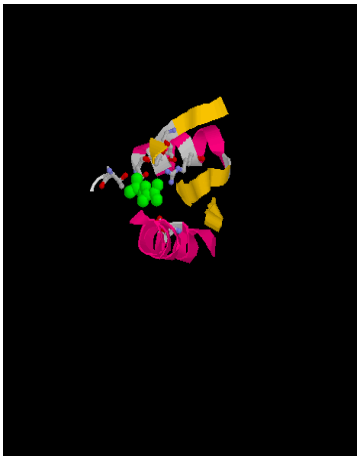
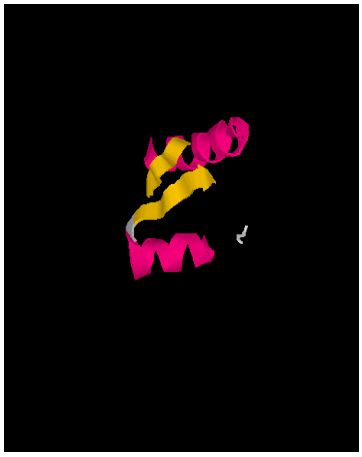
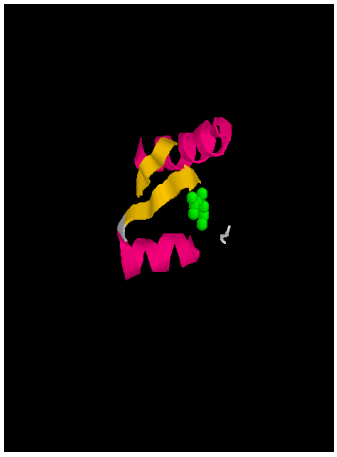
Index II

4L06-B-AKG-503			
4L06-C-AKG-503			
4L06-D-AKG-503			
4L06-E-AKG-503			

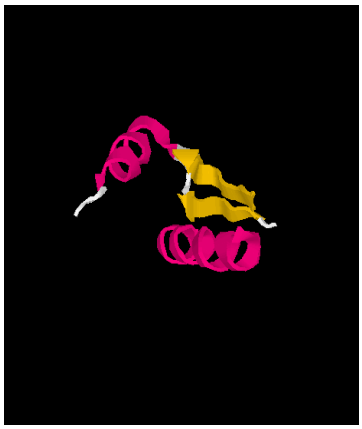
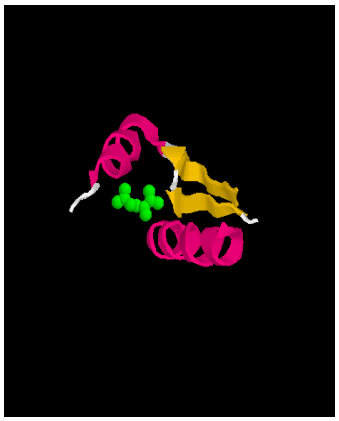
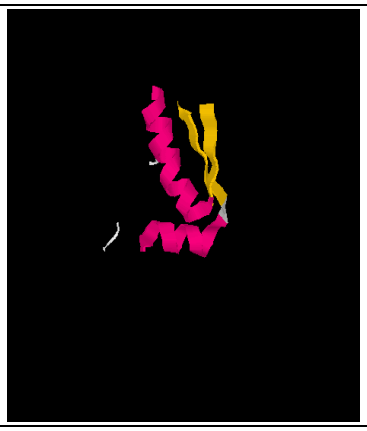
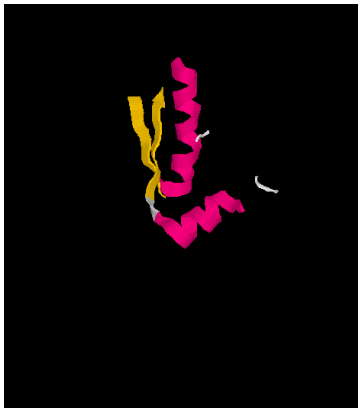
Index II

4L06-F-AKG-503			
4LO4-A-AKG-501			
4LO4-B-AKG-501			

Index II

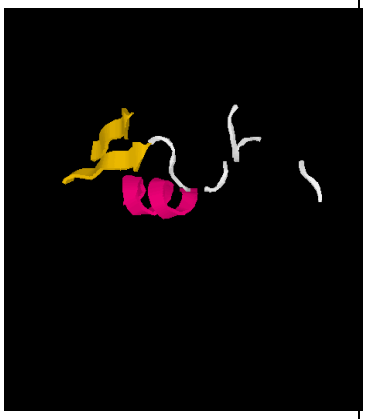
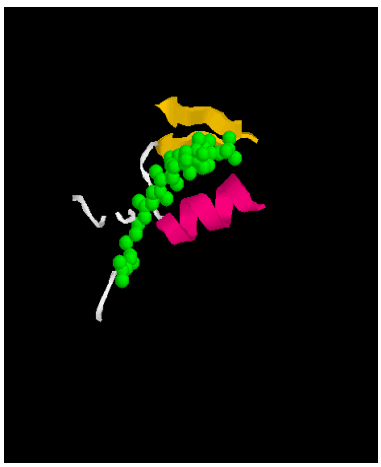
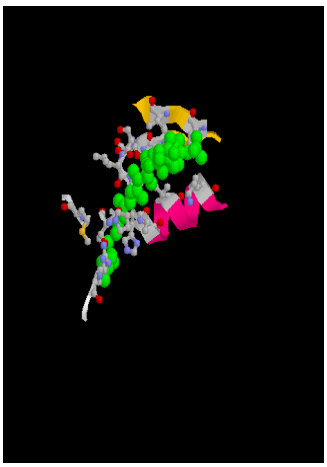
4LO4-C-AKG-501			
4LO4-D-AKG-501			
4LO4-E-AKG-501			

Index II

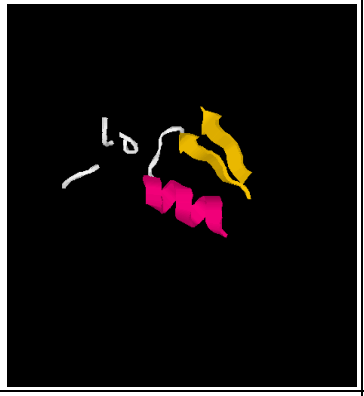
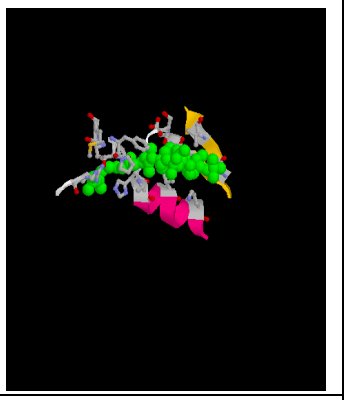
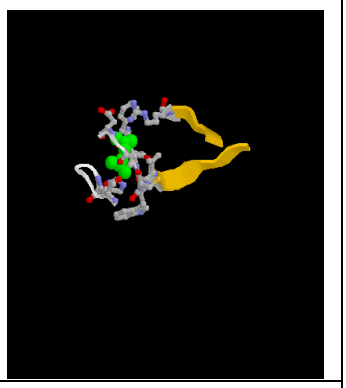
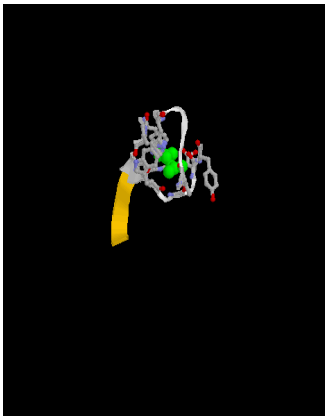
4L04-F-AKG-501			
4LO3-A-AKG-504			
4LO3-B-AKG-503			
4LO3-C-AKG-503			

Index II

05- SUCCINYL-COA-SYNTHETASE:

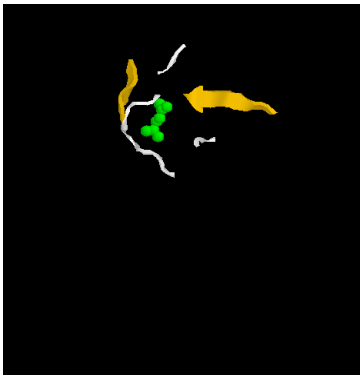
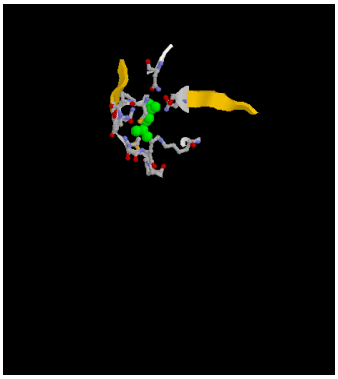
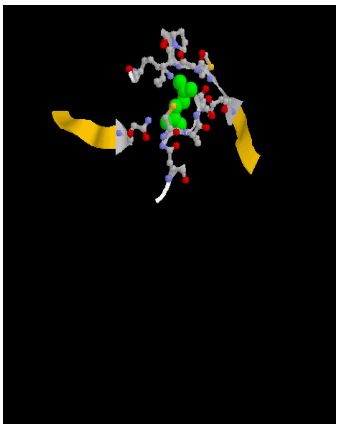
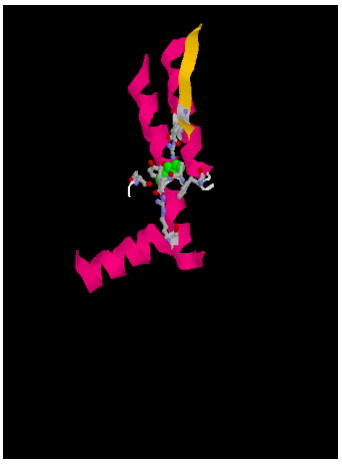
CODES	MOTIFS	MOTIFS + LIGAND	MOTIFS + LIGAND + RESIDUES
2BW0-A-SCA-500			
2BWO-B-SCA-500			
2BWO-D-SCA-500			

Index II

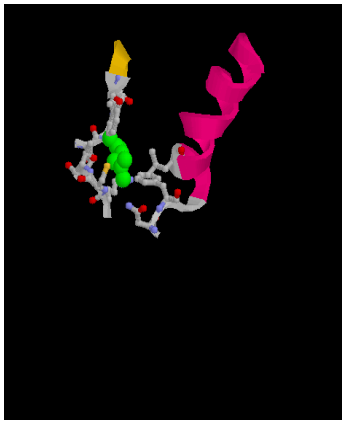
2BWO-E-SCA-500			
2BWN-A-SIN-1398			
2BWN-D-SIN-1399			
2BWN-E-SIN-1400			

Index II

06- Succinate Dehydrogenase:

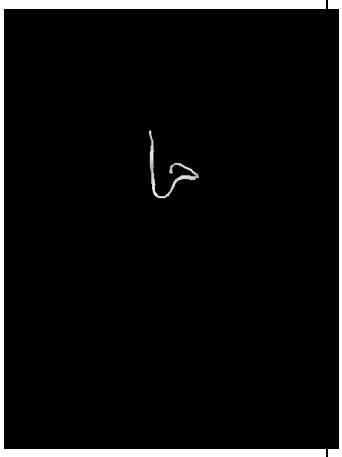
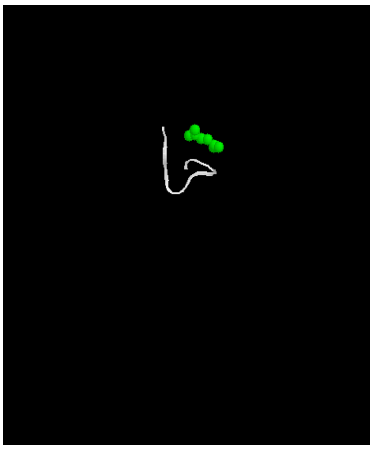
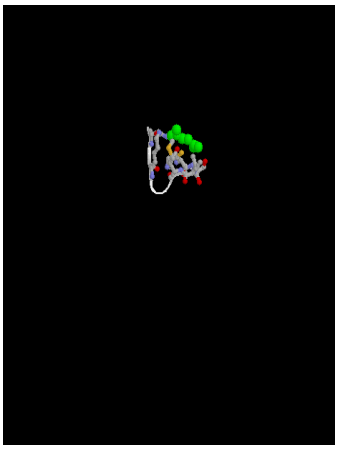
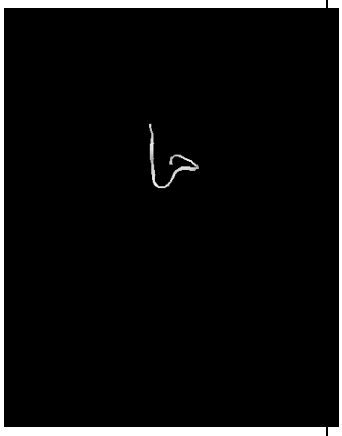
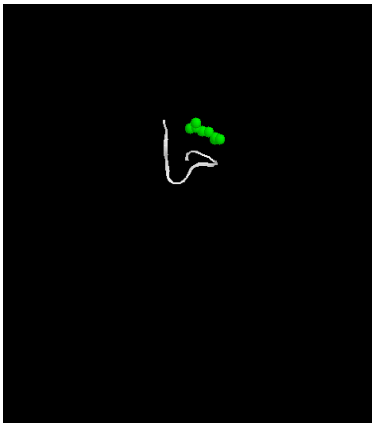
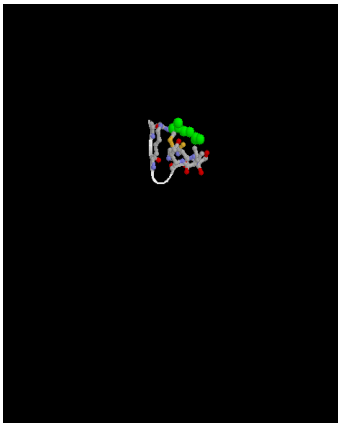
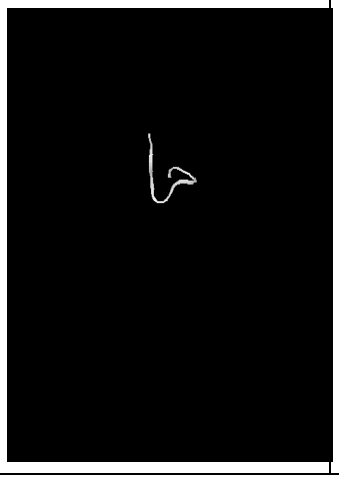
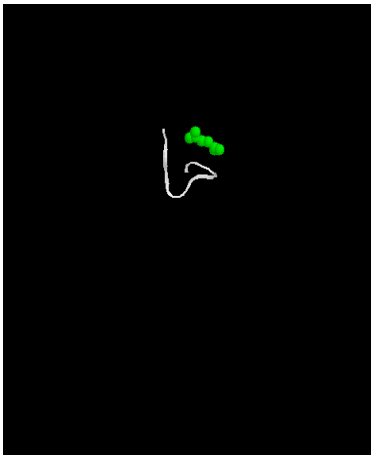
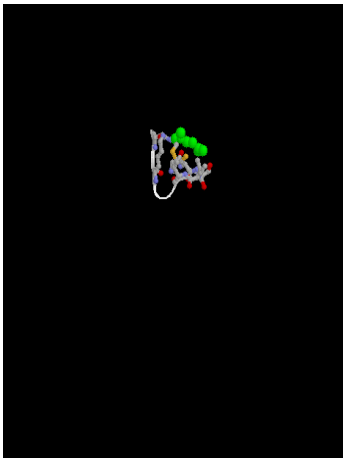
Codes	Motifs	Motifs + Ligand	Motifs + Ligand + Residues
2DJL-A-SIN-1370			
2DJL-B-SIN-2370			
2W8Q-A-SIN-3001			

Index II

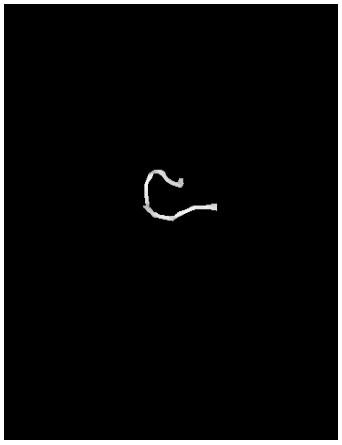
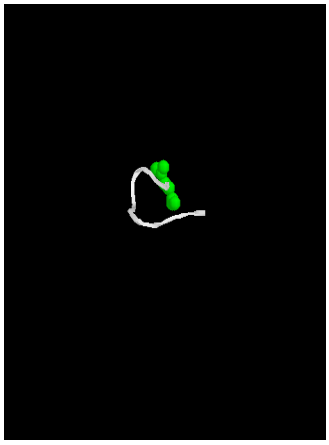
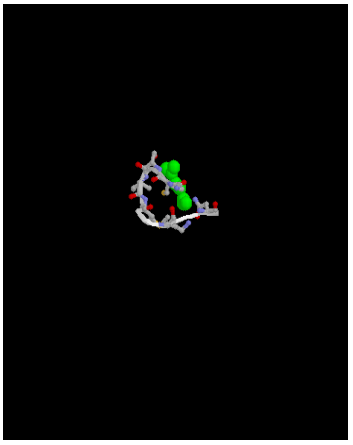
3JU8-A-SIN-504			
4LH2-A-SIN-602			
4LH2-B-SIN-602			

Index II

07- FUMARASE:

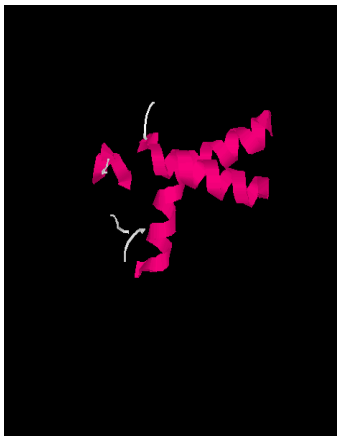
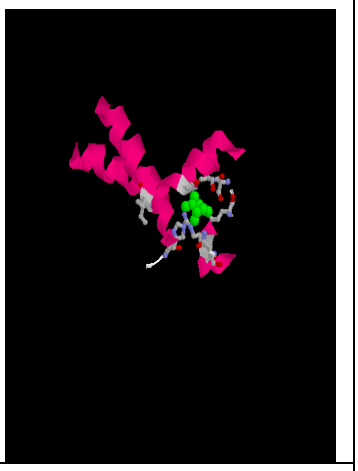
CODES	MOTIFS	MOTIFS + LIGAND	MOTIFS + LIGAND + RESIDUES
4APB-A-FUM-1469			
4APB-B-FUM-1469			
4APB-C-FUM-1469			

Index II

4APB-D-FUM-B-1471			

Index II

08- MALATE DEHYDROGENASE :

CODES	MOTIFS	MOTIFS + LIGAND	MOTIFS + LIGAND + RESIDUES
2DFD-A-MLT-3104			
2DFD-B-MLT-3102			
2DFD-C-MLT-3103			

Index II

2DFD-D-MLT-3101			

Index III

Index III

01-Citrate synthase :

Entry: 1aj8	LYASE									
Protein-Ligand Environment										
Protein Residues					**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**	**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	No SSE	HIS	188	CB	A	CIT	1000	O5	3.2	van der Waals
A	No SSE	HIS	188	CG	A	CIT	1000	O1	3.89	van der Waals
A	No SSE	HIS	188	CG	A	CIT	1000	O5	3.29	van der Waals
A	No SSE	HIS	188	CD2	A	CIT	1000	C1	3.91	van der Waals
A	No SSE	HIS	188	CD2	A	CIT	1000	O1	2.81	van der Waals
A	No SSE	HIS	188	CD2	A	CIT	1000	O7	3.38	van der Waals
A	No SSE	HIS	188	CD2	A	CIT	1000	C6	3.77	van der Waals
A	No SSE	HIS	188	CD2	A	CIT	1000	O5	3.24	van der Waals
A	No SSE	HIS	188	NE2	A	CIT	1000	O1	3.65	H.Bond
A	No SSE	HIS	188	NE2	A	CIT	1000	O7	3.93	H.Bond
A	No SSE	ASN	191	CB	A	CIT	1000	C1	3.38	van der Waals
A	No SSE	ASN	191	CB	A	CIT	1000	O1	3.58	van der Waals
A	No SSE	ASN	191	CB	A	CIT	1000	O2	3.81	van der Waals
A	No SSE	ASN	191	CB	A	CIT	1000	C2	3.45	van der Waals
A	No SSE	ASN	191	CB	A	CIT	1000	O6	3.19	van der Waals

Index III

A	No SSE	ASN	191	CB		A	CIT	1000	O6		3.19	van der Waals
A	No SSE	ASN	191	CG		A	CIT	1000	O6		3.14	van der Waals
A	No SSE	ASN	191	OD1		A	CIT	1000	O6		3.53	H.Bond
A	No SSE	ASN	191	ND2		A	CIT	1000	O6		3.49	H.Bond
A	No SSE	HIS	223	CA		A	CIT	1000	O3		3.3	van der Waals
A	No SSE	HIS	223	C		A	CIT	1000	O3		3.11	van der Waals
A	No SSE	HIS	223	O		A	CIT	1000	O3		3.28	H.Bond
A	No SSE	HIS	223	CG		A	CIT	1000	O3		3.89	van der Waals
A	No SSE	HIS	223	CD2		A	CIT	1000	C2		3.73	van der Waals
A	No SSE	HIS	223	CD2		A	CIT	1000	C4		3.61	van der Waals
A	No SSE	HIS	223	CD2		A	CIT	1000	C5		3.66	van der Waals
A	No SSE	HIS	223	CD2		A	CIT	1000	O3		2.98	van der Waals
A	No SSE	HIS	223	NE2		A	CIT	1000	C2		3.52	H.Bond
A	No SSE	HIS	223	NE2		A	CIT	1000	C4		3.81	H.Bond
A	No SSE	HIS	223	NE2		A	CIT	1000	O3		3.98	H.Bond
A	No SSE	HIS	223	NE2		A	CIT	1000	O6		3.29	H.Bond
A	224-237 H: 5	GLY	224	N		A	CIT	1000	O3		3.59	H.Bond
A	No SSE	HIS	262	CD2		A	CIT	1000	O7		3.58	van der Waals
A	No SSE	HIS	262	CE1		A	CIT	1000	C1		3.61	van der Waals
A	No SSE	HIS	262	CE1		A	CIT	1000	O1		3.22	van der Waals
A	No SSE	HIS	262	CE1		A	CIT	1000	O2		3.66	van der Waals
A	No SSE	HIS	262	NE2		A	CIT	1000	C1		3.19	H.Bond
A	No SSE	HIS	262	NE2		A	CIT	1000	O1		2.96	H.Bond

Index III

A	No SSE	HIS	262	NE2		A	CIT	1000	O2		3.54	H.Bond
A	No SSE	HIS	262	NE2		A	CIT	1000	C2		3.92	H.Bond
A	No SSE	HIS	262	NE2		A	CIT	1000	C3		3.72	H.Bond
A	No SSE	HIS	262	NE2		A	CIT	1000	O7		2.98	H.Bond
A	270-282 H: 1	ARG	271	CZ		A	CIT	1000	O7		3.65	van der Waals
A	270-282 H: 1	ARG	271	CZ		A	CIT	1000	O5		3.59	van der Waals
A	270-282 H: 1	ARG	271	NH1		A	CIT	1000	O7		3.62	H.Bond
A	270-282 H: 1	ARG	271	NH1		A	CIT	1000	C6		3.95	H.Bond
A	270-282 H: 1	ARG	271	NH1		A	CIT	1000	O5		2.75	H.Bond
A	270-282 H: 1	ARG	271	NH2		A	CIT	1000	O7		2.8	H.Bond
A	270-282 H: 1	ARG	271	NH2		A	CIT	1000	O5		3.57	H.Bond
A	No SSE	VAL	311	CG1		A	CIT	1000	O4		3.8	van der Waals
A	No SSE	ASP	312	OD2		A	CIT	1000	C5		3.46	van der Waals
A	No SSE	ASP	312	OD2		A	CIT	1000	O3		3.41	H.Bond
A	No SSE	ASP	312	OD2		A	CIT	1000	O4		3.38	H.Bond
A	326-351 H: 5	PHE	333	CE2		A	CIT	1000	C4		3.95	van der Waals
A	326-351 H: 5	PHE	333	CZ		A	CIT	1000	C4		3.79	van der Waals
A	326-351 H: 5	ARG	337	CZ		A	CIT	1000	O5		3.77	van der Waals
A	326-351 H: 5	ARG	337	CZ		A	CIT	1000	O6		3.62	van der Waals
A	326-351 H: 5	ARG	337	NH1		A	CIT	1000	C6		3.71	H.Bond
A	326-351 H: 5	ARG	337	NH1		A	CIT	1000	O5		3.72	H.Bond
A	326-351 H: 5	ARG	337	NH1		A	CIT	1000	O6		2.8	H.Bond

Index III

A	326-351 H: 5	ARG	337	NH2		A	CIT	1000	O5		2.93	H.Bond
A	326-351 H: 5	ARG	337	NH2		A	CIT	1000	O6		3.55	H.Bond
A	Water	HOH	3025	O		A	CIT	1000	C1		3.53	van der Waals
A	Water	HOH	3025	O		A	CIT	1000	O2		2.97	H.Bond
A	Water	HOH	3025	O		A	CIT	1000	C2		3.33	van der Waals
A	Water	HOH	3030	O		A	CIT	1000	C2		3.81	van der Waals
A	Water	HOH	3030	O		A	CIT	1000	C5		3.63	van der Waals
A	Water	HOH	3030	O		A	CIT	1000	O3		3.15	H.Bond
A	Water	HOH	3035	O		A	CIT	1000	C1		3.55	van der Waals
A	Water	HOH	3035	O		A	CIT	1000	O2		2.6	H.Bond
A	Water	HOH	3194	O		A	CIT	1000	C5		3.33	van der Waals
A	Water	HOH	3194	O		A	CIT	1000	O3		3.42	H.Bond
A	Water	HOH	3194	O		A	CIT	1000	O4		2.57	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

02-Aconitase :

Entry: 1aco	LYASE(CARBON-OXYGEN)									
Protein-Ligand Environment										
Protein Residues					**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**	**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	67-73 S: 1	GLN	72	CD	A	TRA	755	OA2	3.87	van der Waals
A	67-73 S: 1	GLN	72	OE1	A	TRA	755	OA2	3.93	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CA	3.7	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CB	3.96	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	CAC	3.7	H.Bond
A	67-73 S: 1	GLN	72	NE2	A	TRA	755	OA2	2.93	H.Bond
A	No SSE	ALA	74	CB	A	TRA	755	OA2	3.5	van der Waals
A	No SSE	THR	75	CG2	A	TRA	755	OA2	3.59	van der Waals
A	No SSE	HIS	101	CD2	A	TRA	755	OA2	3.94	van der Waals
A	No SSE	HIS	101	NE2	A	TRA	755	CG	3.56	H.Bond
A	No SSE	ASP	165	CA	A	TRA	755	OB2	3.33	van der Waals
A	No SSE	ASP	165	C	A	TRA	755	OB2	3.52	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	CBC	3.72	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	OB1	3.92	van der Waals
A	No SSE	ASP	165	CG	A	TRA	755	OB2	3.87	van der Waals

Index III

A	No SSE	ASP	165	OD1		A	TRA	755	CG		3.49	van der Waals
A	No SSE	ASP	165	OD1		A	TRA	755	CBC		3.32	van der Waals
A	No SSE	ASP	165	OD1		A	TRA	755	OB1		3.71	H.Bond
A	No SSE	ASP	165	OD1		A	TRA	755	OB2		3.69	H.Bond
A	No SSE	ASP	165	OD2		A	TRA	755	CBC		3.95	van der Waals
A	No SSE	ASP	165	OD2		A	TRA	755	OB1		3.75	H.Bond
A	166-174 H: 5	SER	166	N		A	TRA	755	CBC		3.47	H.Bond
A	166-174 H: 5	SER	166	N		A	TRA	755	OB1		3.45	H.Bond
A	166-174 H: 5	SER	166	N		A	TRA	755	OB2		2.82	H.Bond
A	166-174 H: 5	SER	166	CA		A	TRA	755	OB2		3.83	van der Waals
A	166-174 H: 5	SER	166	CB		A	TRA	755	OB1		3.52	van der Waals
A	166-174 H: 5	SER	166	CB		A	TRA	755	OB2		3.65	van der Waals
A	166-174 H: 5	SER	166	OG		A	TRA	755	CBC		3.6	van der Waals
A	166-174 H: 5	SER	166	OG		A	TRA	755	OB1		2.7	H.Bond
A	166-174 H: 5	SER	166	OG		A	TRA	755	OB2		3.67	H.Bond
A	423-425 H: 5	ILE	425	CG2		A	TRA	755	CGC		3.91	van der Waals
A	423-425 H: 5	ILE	425	CG2		A	TRA	755	OG1		3.99	van der Waals
A	No SSE	ARG	447	NH1		A	TRA	755	CGC		3.68	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OB1		3.17	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OG1		3.38	H.Bond
A	No SSE	ARG	447	NH1		A	TRA	755	OG2		3.53	H.Bond
A	No SSE	ARG	452	CZ		A	TRA	755	OG1		3.75	van der Waals

Index III

A	No SSE	ARG	452	NH2		A	TRA	755	OG1	3.68	H.Bond
A	No SSE	ARG	452	NH2		A	TRA	755	OG2	2.87	H.Bond
A	No SSE	ARG	580	CZ		A	TRA	755	CAC	3.99	van der Waals
A	No SSE	ARG	580	CZ		A	TRA	755	OA1	3.52	van der Waals
A	No SSE	ARG	580	CZ		A	TRA	755	OA2	3.81	van der Waals
A	No SSE	ARG	580	NH1		A	TRA	755	OA1	3.18	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	CAC	3.12	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	OA1	3.05	H.Bond
A	No SSE	ARG	580	NH2		A	TRA	755	OA2	2.71	H.Bond
A	No SSE	SER	642	CA		A	TRA	755	CBC	3.86	van der Waals
A	No SSE	SER	642	CA		A	TRA	755	OB1	3.82	van der Waals
A	No SSE	SER	642	CA		A	TRA	755	OB2	3.83	van der Waals
A	No SSE	SER	642	C		A	TRA	755	OB2	3.92	van der Waals
A	No SSE	SER	642	CB		A	TRA	755	CBC	3.9	van der Waals
A	No SSE	SER	642	CB		A	TRA	755	OG1	3.44	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CA	2.9	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CB	3.17	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CAC	3.65	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	CBC	3.29	van der Waals
A	No SSE	SER	642	OG		A	TRA	755	OA1	3.32	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OB1	3.88	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OB2	3.43	H.Bond
A	No SSE	SER	642	OG		A	TRA	755	OG1	3.37	H.Bond
A	No SSE	SER	643	N		A	TRA	755	CBC	3.82	H.Bond

Index III

A	No SSE	SER	643	N		A	TRA	755	OB2		3.04	H.Bond
A	No SSE	SER	643	CA		A	TRA	755	OB2		3.86	van der Waals
A	No SSE	SER	643	CB		A	TRA	755	OB2		3.6	van der Waals
A	No SSE	SER	643	OG		A	TRA	755	CBC		3.97	van der Waals
A	No SSE	SER	643	OG		A	TRA	755	OB2		2.76	H.Bond
A	No SSE	ARG	644	CB		A	TRA	755	OA1		3.82	van der Waals
A	No SSE	ARG	644	CD		A	TRA	755	OA1		3.94	van der Waals
A	No SSE	ARG	644	NE		A	TRA	755	OA1		2.93	H.Bond
A	No SSE	ARG	644	CZ		A	TRA	755	OA1		3.49	van der Waals
A	No SSE	ARG	644	NH2		A	TRA	755	OA1		3.36	H.Bond
A	No SSE	ARG	644	NH2		A	TRA	755	OG1		3.1	H.Bond
A	Water	HOH	1000	O		A	TRA	755	CB		3.88	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CG		3.45	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CBC		3.39	van der Waals
A	Water	HOH	1000	O		A	TRA	755	CGC		3.5	van der Waals
A	Water	HOH	1000	O		A	TRA	755	OB1		2.63	H.Bond
A	Water	HOH	1000	O		A	TRA	755	OG2		2.98	H.Bond
A	Water	HOH	1000	H1		A	TRA	755	CBC		3.66	
A	Water	HOH	1000	H1		A	TRA	755	OB1		2.97	
A	Water	HOH	1000	H1		A	TRA	755	OG2		3.93	
A	Water	HOH	1000	H2		A	TRA	755	CB		3.53	
A	Water	HOH	1000	H2		A	TRA	755	CG		3.34	
A	Water	HOH	1000	H2		A	TRA	755	CBC		2.82	
A	Water	HOH	1000	H2		A	TRA	755	CGC		3.17	

Index III

A	Water	HOH	1000	H2		A	TRA	755	OB1		1.83	, Covalent
A	Water	HOH	1000	H2		A	TRA	755	OB2		3.73	
A	Water	HOH	1000	H2		A	TRA	755	OG2		2.79	

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

03-ISOCITRATE DEHYDROGENASE :

Entry: 4aj3	OXIDOREDUCTASE									
Protein-Ligand Environment										
Protein Residues					**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**	**Chain**	**Name**	**Number**	**Atom**	**Distance /Å**	**Possible Bond Type**
A	No SSE	THR	105	CB	A	ICT	1418	O3	3.15	van der Waals
A	No SSE	THR	105	OG1	A	ICT	1418	C5	3.47	van der Waals
A	No SSE	THR	105	OG1	A	ICT	1418	O3	2.6	H.Bond
A	No SSE	THR	105	OG1	A	ICT	1418	O4	3.58	H.Bond
A	113-122 H: 1	SER	113	CB	A	ICT	1418	O4	3.38	van der Waals
A	113-122 H: 1	SER	113	OG	A	ICT	1418	C5	3.51	van der Waals
A	113-122 H: 1	SER	113	OG	A	ICT	1418	O3	3.82	H.Bond
A	113-122 H: 1	SER	113	OG	A	ICT	1418	O4	2.49	H.Bond
A	113-122 H: 1	ASN	115	CB	A	ICT	1418	O4	3.44	van der Waals
A	113-122 H: 1	ASN	115	CG	A	ICT	1418	O4	3.76	van der Waals
A	113-122 H: 1	ASN	115	ND2	A	ICT	1418	C5	3.42	H.Bond
A	113-122 H: 1	ASN	115	ND2	A	ICT	1418	O3	3.46	H.Bond
A	113-122 H: 1	ASN	115	ND2	A	ICT	1418	O4	3.1	H.Bond
A	113-122 H: 1	ARG	119	CZ	A	ICT	1418	O1	3.52	van der Waals
A	113-122 H: 1	ARG	119	CZ	A	ICT	1418	O6	3.66	van der Waals
A	113-122 H: 1	ARG	119	NH1	A	ICT	1418	O1	3.64	H.Bond

Index III

A	113-122 H: 1	ARG	119	NH1		A	ICT	1418	C6		3.65	H.Bond
A	113-122 H: 1	ARG	119	NH1		A	ICT	1418	O6		2.71	H.Bond
A	113-122 H: 1	ARG	119	NH2		A	ICT	1418	C1		3.67	H.Bond
A	113-122 H: 1	ARG	119	NH2		A	ICT	1418	O1		2.56	H.Bond
A	113-122 H: 1	ARG	119	NH2		A	ICT	1418	O6		3.76	H.Bond
A	126-132 S: -1	ARG	129	CZ		A	ICT	1418	O2		3.59	van der Waals
A	126-132 S: -1	ARG	129	NH1		A	ICT	1418	O2		3.26	H.Bond
A	126-132 S: -1	ARG	129	NH2		A	ICT	1418	C1		3.3	H.Bond
A	126-132 S: -1	ARG	129	NH2		A	ICT	1418	O1		2.86	H.Bond
A	126-132 S: -1	ARG	129	NH2		A	ICT	1418	O2		3.03	H.Bond
A	148-154 S: -1	ARG	153	CZ		A	ICT	1418	O2		3.96	van der Waals
A	148-154 S: -1	ARG	153	CZ		A	ICT	1418	O6		3.5	van der Waals
A	148-154 S: -1	ARG	153	NH1		A	ICT	1418	C1		3.79	H.Bond
A	148-154 S: -1	ARG	153	NH1		A	ICT	1418	O1		3.97	H.Bond
A	148-154 S: -1	ARG	153	NH1		A	ICT	1418	O2		3.04	H.Bond
A	148-154 S: -1	ARG	153	NH1		A	ICT	1418	O6		3.41	H.Bond
A	148-154 S: -1	ARG	153	NH2		A	ICT	1418	C6		3.67	H.Bond
A	148-154 S: -1	ARG	153	NH2		A	ICT	1418	O5		3.87	H.Bond
A	148-154 S: -1	ARG	153	NH2		A	ICT	1418	O6		2.74	H.Bond

Index III

A	157-162 H: 5	TYR	160	CE2		A	ICT	1418	C6		3.58	van der Waals
A	157-162 H: 5	TYR	160	CE2		A	ICT	1418	O5		3.04	van der Waals
A	157-162 H: 5	TYR	160	CE2		A	ICT	1418	O6		3.53	van der Waals
A	157-162 H: 5	TYR	160	CZ		A	ICT	1418	C6		3.82	van der Waals
A	157-162 H: 5	TYR	160	CZ		A	ICT	1418	O5		3.02	van der Waals
A	157-162 H: 5	TYR	160	CZ		A	ICT	1418	O6		3.84	van der Waals
A	157-162 H: 5	TYR	160	OH		A	ICT	1418	C6		3.13	van der Waals
A	157-162 H: 5	TYR	160	OH		A	ICT	1418	O5		2.2	H.Bond
A	157-162 H: 5	TYR	160	OH		A	ICT	1418	O6		3.37	H.Bond
A	302-317 H: 1	ASP	307	CG		A	ICT	1418	O2		3.96	van der Waals
A	302-317 H: 1	ASP	307	CG		A	ICT	1418	O7		3.93	van der Waals
A	302-317 H: 1	ASP	307	OD1		A	ICT	1418	C1		3.77	van der Waals
A	302-317 H: 1	ASP	307	OD1		A	ICT	1418	O2		2.83	H.Bond
A	302-317 H: 1	ASP	307	OD1		A	ICT	1418	C2		3.99	van der Waals
A	302-317 H: 1	ASP	307	OD1		A	ICT	1418	O7		3.07	H.Bond
A	302-317 H: 1	ASP	307	OD1		A	ICT	1418	O5		3.89	H.Bond
A	Water	HOH	2156	O		A	ICT	1418	O2		3.31	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

04-α- Ketoglutarate dehydrogenase :

Entry: 3inm	OXIDOREDUCTASE										
Protein-Ligand Environment											
Protein Residues						**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**		**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	No SSE	THR	77	CB		A	AKG	511	O3	3.35	van der Waals
A	No SSE	THR	77	OG1		A	AKG	511	C5	3.55	van der Waals
A	No SSE	THR	77	OG1		A	AKG	511	O3	2.78	H.Bond
A	No SSE	THR	77	OG1		A	AKG	511	O4	3.61	H.Bond
A	No SSE	THR	77	CG2		A	AKG	511	O3	3.94	van der Waals
A	94-104 H: 1	SER	94	CB		A	AKG	511	O4	3.51	van der Waals
A	94-104 H: 1	SER	94	OG		A	AKG	511	C5	3.48	van der Waals
A	94-104 H: 1	SER	94	OG		A	AKG	511	O3	3.77	H.Bond
A	94-104 H: 1	SER	94	OG		A	AKG	511	O4	2.53	H.Bond
A	94-104 H: 1	ASN	96	CB		A	AKG	511	O4	3.69	van der Waals
A	94-104 H: 1	ASN	96	CG		A	AKG	511	O4	3.76	van der Waals
A	94-104 H: 1	ASN	96	ND2		A	AKG	511	C3	3.93	H.Bond
A	94-104 H: 1	ASN	96	ND2		A	AKG	511	C5	3.84	H.Bond
A	94-104 H: 1	ASN	96	ND2		A	AKG	511	O4	2.89	H.Bond
A	94-104 H: 1	ARG	100	CZ		A	AKG	511	O1	3.3	van der Waals
A	94-104 H: 1	ARG	100	NH1		A	AKG	511	O1	3.16	H.Bond

Index III

A	94-104 H: 1	ARG	100	NH2		A	AKG	511	C1		3.87	H.Bond
A	94-104 H: 1	ARG	100	NH2		A	AKG	511	O1		2.65	H.Bond
A	106-111 S: -1	ARG	109	CZ		A	AKG	511	O2		3.52	van der Waals
A	106-111 S: -1	ARG	109	NH1		A	AKG	511	O1		3.84	H.Bond
A	106-111 S: -1	ARG	109	NH1		A	AKG	511	O2		3.66	H.Bond
A	106-111 S: -1	ARG	109	NH2		A	AKG	511	C1		3.57	H.Bond
A	106-111 S: -1	ARG	109	NH2		A	AKG	511	O1		3.87	H.Bond
A	106-111 S: -1	ARG	109	NH2		A	AKG	511	O2		2.55	H.Bond
A	136-141 H: 5	TYR	139	OH		A	AKG	511	C2		3.74	van der Waals
A	136-141 H: 5	TYR	139	OH		A	AKG	511	C3		3.17	van der Waals
A	136-141 H: 5	TYR	139	OH		A	AKG	511	C4		3.52	van der Waals
A	136-141 H: 5	TYR	139	OH		A	AKG	511	C5		3.9	van der Waals
A	136-141 H: 5	TYR	139	OH		A	AKG	511	O4		3.86	H.Bond
A	270-286 H: 1	ASP	275	CG		A	AKG	511	C1		3.67	van der Waals
A	270-286 H: 1	ASP	275	CG		A	AKG	511	O2		3.38	van der Waals
A	270-286 H: 1	ASP	275	CG		A	AKG	511	C2		3.76	van der Waals
A	270-286 H: 1	ASP	275	CG		A	AKG	511	O5		3.66	van der Waals
A	270-286 H: 1	ASP	275	OD1		A	AKG	511	C1		3.32	van der Waals
A	270-286 H: 1	ASP	275	OD1		A	AKG	511	O2		2.76	H.Bond
A	270-286 H: 1	ASP	275	OD1		A	AKG	511	C2		3.41	van der Waals

Index III

A	270-286 H: 1	ASP	275	OD 1		A	AKG	511	O5	3.04	H.Bond
A	270-286 H: 1	ASP	275	OD 2		A	AKG	511	C1	3.74	van der Waals
A	270-286 H: 1	ASP	275	OD 2		A	AKG	511	O2	3.86	H.Bond
A	270-286 H: 1	ASP	275	OD 2		A	AKG	511	C2	3.51	van der Waals
A	270-286 H: 1	ASP	275	OD 2		A	AKG	511	O5	3.54	H.Bond
A	No SSE	ALA	308	O		A	AKG	511	O2	3.9	H.Bond
A	Water	HOH	427	O		A	AKG	511	O2	3.69	H.Bond
A	Water	HOH	427	O		A	AKG	511	O5	3.36	H.Bond
A	Water	HOH	430	O		A	AKG	511	O5	3.39	H.Bond
A	Water	HOH	430	O		A	AKG	511	C4	3.34	van der Waals
A	Water	HOH	431	O		A	AKG	511	C4	3.4	van der Waals
A	Water	HOH	431	O		A	AKG	511	C5	3.29	van der Waals
A	Water	HOH	431	O		A	AKG	511	O3	2.43	H.Bond
A	Water	HOH	432	O		A	AKG	511	C5	3.7	van der Waals
A	Water	HOH	432	O		A	AKG	511	O3	3.73	H.Bond
A	Water	HOH	432	O		A	AKG	511	O4	3.3	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

05-SUCCINYL-COA-SYNTHETASE:

Entry: 2bwo	NH_1TRANSFERASE										
Protein-Ligand Environment											
Protein Residues						**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**		**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	No SSE	ARG	21	CZ		A	SCA	500	OS5	3.72	van der Waals
A	No SSE	ARG	21	NH1		A	SCA	500	CS4	3.62	H.Bond
A	No SSE	ARG	21	NH1		A	SCA	500	OS4	3.76	H.Bond
A	No SSE	ARG	21	NH1		A	SCA	500	OS5	2.79	H.Bond
A	No SSE	ARG	21	NH2		A	SCA	500	OS5	3.93	H.Bond
A	134-138 S: 1	SER	137	CB		A	SCA	500	C2	3.85	van der Waals
A	134-138 S: 1	SER	137	CB		A	SCA	500	N1	3.32	H.Bond
A	134-138 S: 1	SER	137	OG		A	SCA	500	C2	3.59	van der Waals
A	134-138 S: 1	SER	137	OG		A	SCA	500	N1	2.74	H.Bond
A	134-138 S: 1	SER	137	OG		A	SCA	500	C6	3.59	van der Waals
A	134-138 S: 1	SER	137	OG		A	SCA	500	N6	3.57	H.Bond
A	134-138 S: 1	ASP	138	C		A	SCA	500	N6	3.71	H.Bond
A	134-138 S: 1	ASP	138	O		A	SCA	500	C6	3.96	van der Waals
A	134-138 S: 1	ASP	138	O		A	SCA	500	N6	2.79	H.Bond
A	No SSE	SER	139	CA		A	SCA	500	N6	3.54	H.Bond
A	No SSE	SER	139	C		A	SCA	500	N6	3.5	H.Bond
A	No SSE	SER	139	O		A	SCA	500	N6	3.16	H.Bond
A	No SSE	LEU	140	O		A	SCA	500	OP1	3.99	H.Bond

Index III

A	No SSE	ASN	141	O		A	SCA	500	CP4		3.55	van der Waals
A	No SSE	ASN	141	O		A	SCA	500	CP3		3.99	van der Waals
A	No SSE	ASN	141	O		A	SCA	500	OP1		3.86	H.Bond
A	No SSE	HIS	142	C		A	SCA	500	CP4		3.78	van der Waals
A	No SSE	HIS	142	CD2		A	SCA	500	CP1		2.98	van der Waals
A	No SSE	HIS	142	CD2		A	SCA	500	OS1		3.9	van der Waals
A	No SSE	HIS	142	NE2		A	SCA	500	CP1		3.21	H.Bond
A	No SSE	HIS	142	NE2		A	SCA	500	CS1		3.91	H.Bond
A	No SSE	HIS	142	NE2		A	SCA	500	OS1		3.05	H.Bond
A	143-151 H: 1	ALA	143	N		A	SCA	500	CP4		3.2	H.Bond
A	143-151 H: 1	ALA	143	N		A	SCA	500	CP3		3.8	H.Bond
A	143-151 H: 1	ALA	143	N		A	SCA	500	NP1		3.9	H.Bond
A	143-151 H: 1	ALA	143	CA		A	SCA	500	CP4		3.79	van der Waals
A	143-151 H: 1	ILE	146	CG1		A	SCA	500	C5		3.69	van der Waals
A	143-151 H: 1	ILE	146	CG1		A	SCA	500	N1		3.71	H.Bond
A	143-151 H: 1	ILE	146	CG1		A	SCA	500	C6		3.48	van der Waals
A	143-151 H: 1	ILE	146	CG1		A	SCA	500	N6		3.85	H.Bond
A	143-151 H: 1	ILE	146	CG2		A	SCA	500	N9		3.86	H.Bond
A	143-151 H: 1	ILE	146	CG2		A	SCA	500	C4		3.65	van der Waals
A	143-151 H: 1	ILE	146	CG2		A	SCA	500	N3		3.76	H.Bond
A	143-151 H: 1	ILE	146	CD1		A	SCA	500	C5		3.57	van der Waals
A	143-151 H: 1	ILE	146	CD1		A	SCA	500	N7		3.66	H.Bond
A	143-151 H: 1	ILE	146	CD1		A	SCA	500	C6		3.67	van der Waals

Index III

A	143-151 H: 1	ILE	146	CD1		A	SCA	500	N6		3.76	H.Bond
A	143-151 H: 1	LYS	150	CD		A	SCA	500	O4'		3.87	van der Waals
A	143-151 H: 1	LYS	150	NZ		A	SCA	500	C4'		3.95	H.Bond
A	156-159 S: 1	LYS	156	CD		A	SCA	500	N3		3.96	H.Bond
A	156-159 S: 1	LYS	156	NZ		A	SCA	500	O2'		3.72	H.Bond
A	156-159 S: 1	LYS	156	NZ		A	SCA	500	O3'		3.67	H.Bond
A	156-159 S: 1	LYS	156	NZ		A	SCA	500	O33		3.97	H.Bond
A	156-159 S: 1	ILE	158	CD1		A	SCA	500	N9		3.9	H.Bond
A	156-159 S: 1	ILE	158	CD1		A	SCA	500	C4		3.6	van der Waals
A	156-159 S: 1	ILE	158	CD1		A	SCA	500	C5		3.79	van der Waals
A	156-159 S: 1	ILE	158	CD1		A	SCA	500	N3		3.88	H.Bond
A	156-159 S: 1	ILE	158	CD1		A	SCA	500	O2'		3.46	van der Waals
A	No SSE	SER	189	OG		A	SCA	500	CP1		3.86	van der Waals
A	No SSE	SER	189	OG		A	SCA	500	S		3.38	
A	No SSE	MET	190	CE		A	SCA	500	S		3.06	van der Waals
A	No SSE	MET	190	CE		A	SCA	500	CS1		3.98	van der Waals
A	No SSE	MET	190	CE		A	SCA	500	CS2		3.94	van der Waals
A	No SSE	PHE	363	CE1		A	SCA	500	CP7		3.96	van der Waals
A	No SSE	PHE	363	CE1		A	SCA	500	CP8		3.53	van der Waals
A	No SSE	PRO	364	C		A	SCA	500	NP2		3.92	H.Bond

Index III

A	No SSE	PRO	364	O		A	SCA	500	CP7		3.69	van der Waals
A	No SSE	PRO	364	O		A	SCA	500	OP3		3.94	H.Bond
A	No SSE	PRO	364	O		A	SCA	500	CP6		3.31	van der Waals
A	No SSE	PRO	364	O		A	SCA	500	NP2		2.71	H.Bond
A	No SSE	PRO	364	O		A	SCA	500	CP5		3.23	van der Waals
A	No SSE	THR	365	O		A	SCA	500	CP3		3.94	van der Waals
A	No SSE	THR	365	O		A	SCA	500	OP1		3.79	H.Bond
A	No SSE	THR	365	O		A	SCA	500	NP1		3.72	H.Bond
A	No SSE	THR	365	O		A	SCA	500	CP2		3.25	van der Waals
A	No SSE	THR	365	CB		A	SCA	500	OS5		3.42	van der Waals
A	No SSE	THR	365	OG1		A	SCA	500	CS3		3.73	van der Waals
A	No SSE	THR	365	OG1		A	SCA	500	CS4		3.37	van der Waals
A	No SSE	THR	365	OG1		A	SCA	500	OS5		2.44	H.Bond
A	No SSE	THR	365	CG2		A	SCA	500	CP2		3.6	van der Waals
A	No SSE	THR	365	CG2		A	SCA	500	S		3.58	van der Waals
A	No SSE	THR	365	CG2		A	SCA	500	CS2		3.66	van der Waals
A	No SSE	THR	365	CG2		A	SCA	500	CS3		3.75	van der Waals
A	No SSE	THR	365	CG2		A	SCA	500	OS5		3.55	van der Waals
A	No SSE	PRO	367	N		A	SCA	500	OP3		3.83	H.Bond
A	No SSE	PRO	367	CG		A	SCA	500	OP3		3.87	van der Waals
A	Water	HOH	2029	O		A	SCA	500	CPB		3.12	van der Waals

Index III

A	Water	HOH	2029	O		A	SCA	500	CPA		1.94	van der Waals, Covalent
A	Water	HOH	2029	O		A	SCA	500	CP9		2.73	van der Waals
A	Water	HOH	2029	O		A	SCA	500	CP8		2.6	van der Waals
A	Water	HOH	2029	O		A	SCA	500	CP6		2.02	van der Waals
A	Water	HOH	2029	O		A	SCA	500	OP2		3.04	H.Bond
A	Water	HOH	2029	O		A	SCA	500	NP2		2.59	H.Bond
A	Water	HOH	2030	O		A	SCA	500	CP7		3.65	van der Waals
A	Water	HOH	2030	O		A	SCA	500	OP3		3.95	H.Bond
A	Water	HOH	2030	O		A	SCA	500	CP6		2.28	van der Waals
A	Water	HOH	2030	O		A	SCA	500	OP2		2.64	H.Bond
A	Water	HOH	2030	O		A	SCA	500	NP2		1.34	H.Bond, Covalent
A	Water	HOH	2030	O		A	SCA	500	CP4		1.59	van der Waals, Covalent
A	Water	HOH	2030	O		A	SCA	500	CP3		2.61	van der Waals
A	Water	HOH	2030	O		A	SCA	500	OP1		3.47	H.Bond
A	Water	HOH	2030	O		A	SCA	500	NP1		3.17	H.Bond
A	Water	HOH	2031	O		A	SCA	500	CP3		3.08	van der Waals
A	Water	HOH	2031	O		A	SCA	500	OP1		3.42	H.Bond
A	Water	HOH	2031	O		A	SCA	500	NP1		2.03	H.Bond
A	Water	HOH	2031	O		A	SCA	500	S		2.13	
A	Water	HOH	2031	O		A	SCA	500	CS1		3.31	van der Waals
A	Water	HOH	2031	O		A	SCA	500	OS1		3.83	H.Bond
A	Water	HOH	2032	O		A	SCA	500	C5'		2.17	van der Waals
A	Water	HOH	2032	O		A	SCA	500	O5'		1.32	H.Bond, Covalent
A	Water	HOH	2032	O		A	SCA	500	O11		1.29	H.Bond, Covalent

Index III

A	Water	HOH	2032	O		A	SCA	500	O12		1.85	H.Bond, Covalent
A	Water	HOH	2032	O		A	SCA	500	O6		2.08	H.Bond
A	Water	HOH	2032	O		A	SCA	500	P2		3.39	
A	Water	HOH	2032	O		A	SCA	500	O21		3.77	H.Bond
A	Water	HOH	2032	O		A	SCA	500	O22		3.9	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

06-SUCCINATE DEHYDROGENASE:

Entry: 2w8q	OXIDOREDUCTASE										
Protein-Ligand Environment											
Protein Residues						**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**		**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	149-171 H: 1	TYR	159	CE2		A	SIN	3001	O1	3.51	van der Waals
A	No SSE	PHE	206	CE1		A	SIN	3001	C3	3.64	van der Waals
A	208-222 H: 1	ARG	213	NH1		A	SIN	3001	C2	3.81	H.Bond
A	208-222 H: 1	ARG	213	NH2		A	SIN	3001	C1	3.86	H.Bond
A	208-222 H: 1	ARG	213	NH2		A	SIN	3001	O1	3.15	H.Bond
A	208-222 H: 1	ARG	213	NH2		A	SIN	3001	C2	3.64	H.Bond
A	302-307 S: 1	GLU	306	OE1		A	SIN	3001	C4	3.53	van der Waals
A	302-307 S: 1	GLU	306	OE1		A	SIN	3001	O3	3.69	H.Bond
A	320-334 H: 1	ARG	334	CZ		A	SIN	3001	O2	3.74	van der Waals
A	320-334 H: 1	ARG	334	NH1		A	SIN	3001	O2	3.46	H.Bond
A	320-334 H: 1	ARG	334	NH2		A	SIN	3001	C1	3.9	H.Bond
A	320-334 H: 1	ARG	334	NH2		A	SIN	3001	O2	3.13	H.Bond
A	No SSE	ALA	340	N		A	SIN	3001	O3	3.29	H.Bond
A	No SSE	ALA	340	CA		A	SIN	3001	O3	3.82	van der Waals
A	No SSE	ALA	340	CB		A	SIN	3001	O3	3.35	van der Waals
A	No SSE	VAL	341	CG1		A	SIN	3001	C3	3.98	van der Waals
A	No SSE	VAL	341	CG1		A	SIN	3001	C4	3.75	van der Waals

Index III

A	No SSE	VAL	341	CG2		A	SIN	3001	O2		3.8	van der Waals
A	No SSE	VAL	341	CG2		A	SIN	3001	C3		3.91	van der Waals
A	No SSE	SER	498	OG		A	SIN	3001	C1		3.68	van der Waals
A	No SSE	SER	498	OG		A	SIN	3001	O1		3.74	H.Bond
A	No SSE	SER	498	OG		A	SIN	3001	O2		2.94	H.Bond
A	No SSE	PHE	504	CE2		A	SIN	3001	O2		3.99	van der Waals

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

07- FUMARASE:

Entry: 4apb	HYDROLASE										
Protein-Ligand Environment											
Protein Residues						**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**		**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
A	No SSE	GLY	317	O		A	FUM	1469	C4	3.72	van der Waals
A	No SSE	GLY	317	O		A	FUM	1469	O1	3.73	H.Bond
A	No SSE	CYS	318	CA		A	FUM	1469	C4	3.6	van der Waals
A	No SSE	CYS	318	CA		A	FUM	1469	C5	3.96	van der Waals
A	No SSE	CYS	318	CA		A	FUM	1469	C6	3.77	van der Waals
A	No SSE	CYS	318	CA		A	FUM	1469	O8	3.35	van der Waals
A	No SSE	CYS	318	C		A	FUM	1469	C6	3.97	van der Waals
A	No SSE	CYS	318	C		A	FUM	1469	O8	3.62	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	C2	3.64	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	C4	3.19	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	C5	3.66	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	C6	3.96	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	O1	3.59	van der Waals
A	No SSE	CYS	318	CB		A	FUM	1469	O8	3.99	van der Waals
A	No SSE	CYS	318	SG		A	FUM	1469	C2	3.21	van der Waals

Index III

A	No SSE	CYS	318	SG		A	FUM	1469	C4		2.63	van der Waals
A	No SSE	CYS	318	SG		A	FUM	1469	C5		2.54	van der Waals
A	No SSE	CYS	318	SG		A	FUM	1469	C6		3.03	van der Waals
A	No SSE	CYS	318	SG		A	FUM	1469	O1		3.75	
A	No SSE	CYS	318	SG		A	FUM	1469	O3		3.8	
A	No SSE	CYS	318	SG		A	FUM	1469	O7		3.55	
A	No SSE	CYS	318	SG		A	FUM	1469	O8		3.62	
A	No SSE	SER	319	N		A	FUM	1469	C6		3.26	H.Bond
A	No SSE	SER	319	N		A	FUM	1469	O7		3.49	H.Bond
A	No SSE	SER	319	N		A	FUM	1469	O8		2.91	H.Bond
A	No SSE	SER	319	CA		A	FUM	1469	O8		3.96	van der Waals
A	No SSE	SER	319	CB		A	FUM	1469	O7		3.68	van der Waals
A	No SSE	SER	319	CB		A	FUM	1469	O8		3.85	van der Waals
A	No SSE	SER	319	OG		A	FUM	1469	C6		3.45	van der Waals
A	No SSE	SER	319	OG		A	FUM	1469	O7		2.66	H.Bond
A	No SSE	SER	319	OG		A	FUM	1469	O8		3.58	H.Bond
A	No SSE	ILE	320	N		A	FUM	1469	O7		3.87	H.Bond
A	No SSE	ILE	320	CG2		A	FUM	1469	O7		3.69	van der Waals
A	No SSE	MET	321	CE		A	FUM	1469	C2		3.72	van der Waals
A	No SSE	MET	321	CE		A	FUM	1469	O3		3.2	van der Waals
A	No SSE	LYS	324	CD		A	FUM	1469	O1		3.74	van der Waals
A	No SSE	LYS	324	CE		A	FUM	1469	O1		3.66	van der Waals
A	No SSE	LYS	324	NZ		A	FUM	1469	C2		3.49	H.Bond
A	No SSE	LYS	324	NZ		A	FUM	1469	O1		2.65	H.Bond

Index III

A	No SSE	LYS	324	NZ		A	FUM	1469	O3		3.64	H.Bond
A	No SSE	ASN	326	CG		A	FUM	1469	O1		3.52	van der Waals
A	No SSE	ASN	326	OD1		A	FUM	1469	O1		3.58	H.Bond
A	No SSE	ASN	326	ND2		A	FUM	1469	C2		3.75	H.Bond
A	No SSE	ASN	326	ND2		A	FUM	1469	O1		2.67	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.

Index III

08-MALATE DEHYDROGENASE :

Entry: 2dfd	OXIDOREDUCTASE										
Protein-Ligand Environment											
Protein Residues						**Ligand**				**Bonds**	
Chain	**SSelm**	**Name**	**Number**	**Atom**		**Chain**	**Name**	**Number**	**Atom**	**Distance/Å**	**Possible Bond Type**
B	No SSE	ARG	86	CB		B	MLT	3102	O5	3.82	van der Waals
B	No SSE	ARG	86	CG		B	MLT	3102	O5	3.41	van der Waals
B	No SSE	ARG	86	CD		B	MLT	3102	O5	3.64	van der Waals
B	No SSE	ARG	86	NE		B	MLT	3102	C4	3.6	H.Bond
B	No SSE	ARG	86	NE		B	MLT	3102	O4	3.56	H.Bond
B	No SSE	ARG	86	NE		B	MLT	3102	O5	2.82	H.Bond
B	No SSE	ARG	86	CZ		B	MLT	3102	O4	3.64	van der Waals
B	No SSE	ARG	86	CZ		B	MLT	3102	O5	3.69	van der Waals
B	No SSE	ARG	86	NH2		B	MLT	3102	C4	3.67	H.Bond
B	No SSE	ARG	86	NH2		B	MLT	3102	O4	2.84	H.Bond
B	No SSE	ARG	86	NH2		B	MLT	3102	O5	3.68	H.Bond
B	91-94 H: 5	ARG	92	CD		B	MLT	3102	O3	3.76	van der Waals
B	91-94 H: 5	ARG	92	NE		B	MLT	3102	C2	3.92	H.Bond
B	91-94 H: 5	ARG	92	NE		B	MLT	3102	O3	2.72	H.Bond
B	91-94 H: 5	ARG	92	NE		B	MLT	3102	C4	3.85	H.Bond
B	91-94 H: 5	ARG	92	NE		B	MLT	3102	O4	3.76	H.Bond
B	91-94 H: 5	ARG	92	CZ		B	MLT	3102	O1	3.9	van der Waals

Index III

B	91-94 H: 5	ARG	92	CZ		B	MLT	3102	O3		3.46	van der Waals
B	91-94 H: 5	ARG	92	CZ		B	MLT	3102	O4		3.69	van der Waals
B	91-94 H: 5	ARG	92	NH2		B	MLT	3102	C1		3.58	H.Bond
B	91-94 H: 5	ARG	92	NH2		B	MLT	3102	O1		2.86	H.Bond
B	91-94 H: 5	ARG	92	NH2		B	MLT	3102	O3		3.34	H.Bond
B	91-94 H: 5	ARG	92	NH2		B	MLT	3102	O4		3.32	H.Bond
B	No SSE	ASN	124	CG		B	MLT	3102	O3		3.86	van der Waals
B	No SSE	ASN	124	OD1		B	MLT	3102	O3		3.94	H.Bond
B	No SSE	ASN	124	ND2		B	MLT	3102	C2		3.59	H.Bond
B	No SSE	ASN	124	ND2		B	MLT	3102	O3		2.99	H.Bond
B	No SSE	ASN	124	ND2		B	MLT	3102	C3		3.55	H.Bond
B	152-168 H: 1	ARG	158	CZ		B	MLT	3102	O1		3.63	van der Waals
B	152-168 H: 1	ARG	158	CZ		B	MLT	3102	O2		3.7	van der Waals
B	152-168 H: 1	ARG	158	NH1		B	MLT	3102	C1		3.48	H.Bond
B	152-168 H: 1	ARG	158	NH1		B	MLT	3102	O1		2.75	H.Bond
B	152-168 H: 1	ARG	158	NH1		B	MLT	3102	O2		3.59	H.Bond
B	152-168 H: 1	ARG	158	NH2		B	MLT	3102	C1		3.64	H.Bond
B	152-168 H: 1	ARG	158	NH2		B	MLT	3102	O1		3.64	H.Bond
B	152-168 H: 1	ARG	158	NH2		B	MLT	3102	O2		2.93	H.Bond
B	No SSE	HIS	182	CD2		B	MLT	3102	O3		3.47	van der Waals
B	No SSE	HIS	182	CE1		B	MLT	3102	O1		3.51	van der Waals
B	No SSE	HIS	182	CE1		B	MLT	3102	O3		3.69	van der Waals
B	No SSE	HIS	182	NE2		B	MLT	3102	C1		3.91	H.Bond
B	No SSE	HIS	182	NE2		B	MLT	3102	O1		3.44	H.Bond
B	No SSE	HIS	182	NE2		B	MLT	3102	C2		3.51	H.Bond
B	No SSE	HIS	182	NE2		B	MLT	3102	O3		2.69	H.Bond

Index III

B	201-224 H: 1	GLY	216	CA		B	MLT	3102	C1		3.84	van der Waals
B	201-224 H: 1	GLY	216	CA		B	MLT	3102	O1		3.67	van der Waals
B	201-224 H: 1	GLY	216	CA		B	MLT	3102	O2		3.42	van der Waals
B	201-224 H: 1	GLY	216	CA		B	MLT	3102	O4		3.41	van der Waals
B	201-224 H: 1	GLY	216	C		B	MLT	3102	O4		3.55	van der Waals
B	201-224 H: 1	GLY	216	O		B	MLT	3102	O4		3.63	H.Bond
B	No SSE	SER	228	CB		B	MLT	3102	O5		3.97	van der Waals
B	No SSE	ALA	229	CB		B	MLT	3102	O2		3.77	van der Waals
B	Water	HOH	3405	O		B	MLT	3102	C3		3.48	van der Waals
B	Water	HOH	3405	O		B	MLT	3102	C4		3.41	van der Waals
B	Water	HOH	3405	O		B	MLT	3102	O5		2.57	H.Bond
B	Water	HOH	3432	O		B	MLT	3102	O1		3.87	H.Bond
B	Water	HOH	3528	O		B	MLT	3102	C1		3.65	van der Waals
B	Water	HOH	3528	O		B	MLT	3102	O2		2.51	H.Bond
B	Water	HOH	3528	O		B	MLT	3102	C3		3.72	van der Waals
B	Water	HOH	3528	O		B	MLT	3102	C4		3.59	van der Waals
B	Water	HOH	3528	O		B	MLT	3102	O4		3.28	H.Bond

Note: The rest of binding details are stored in the Flat-Files database and in the online version of the database.