STRUCTURAL, FUNCTIONAL AND EVOLUTIONARY ANALYSIS OF CHROMATE ION TRANSPORTER PROTEINS USING COMPUTATIONAL TOOLS AND TECHNIQUES

A

Thesis submitted

To

Orissa University of Agriculture and Technology, Bhubaneswar

In partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE IN ZOOLOGY

By

Sayeda Tahera Requab

Admission No-08ZOL/12

DEPARTMENT OF ZOOLOGY

COLLEGE OF BASIC SCIENCE AND HUMANITIES

ORISSA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

BHUBANESWAR-751003

2014

Advisor: Dr. Netaji Upadhyaya

Department of Zoology

College of Basic Science and Humanities

Orissa University of Agriculture and Technology, Bhubaneswar-3

CERTIFICATE – I

This is to certify that the thesis entitled, **"Structural, functional and evolutionary analysis of chromate ion transporter proteins using computational tools and techniques",** submitted in partial fulfillment of the requirements for the award of the degree of **Master of Science in Zoology** of the **Orissa University of Agriculture and Technology, Bhubaneswar**, is a faithful bonafide research work carried out by **Miss Sayeda Tahera Requab (Adm. No 08ZOL/12)** , under my guidance and supervision and that no part of this thesis has been submitted for any other degree or diploma or published in any form.

It is further certified that the help and sources of information availed of during the course of study have been duly acknowledged.

Place: Bhubaneswar

Date:

Dr. Netaji Upadhyaya

Chairman

Advisory committee

CERTIFICATE-II

This is to certify that the thesis entitled, **"Structural, functional and evolutionary analysis of chromate ion transporter proteins using computational tools and techniques",** submitted by **Miss SayedaTahera Requab (Adm. No 08ZOL/12)** to the **Orissa University of Agriculture and Technology, Bhubaneswar**, in partial fulfillment of the requirements for the degree of **Master of Science in Zoology,** has been approved by the Students' Advisory Committee after an oral examination on the same in collaboration with an External Examiner.

Advisory Committee

1. **Dr. Netaji Upadhyaya** — **Chairman**
   Professor, Department of Zoology
   College of Basic Science and Humanities

2. **Mr. Sukanta Kumar Pradhan** — **Member**
   Assistant Professor and Head

   Department of Bioinformatics

   Centre of Post Graduate Studies

3. **Dr. Choudhury Suryakanta Mishra** — **Member**
   Professor and Head, Department of Zoology

External Examiner

Name:

Designation:

Address:

Signature:

ACKNOWLEDGEMENT

Though this dissertation is an individual work, it could not have been complete with the support and blessing of many important people.

I would first like to thank my sincere reverence, obligation and deepest sense of gratitude to my guide Dr Netaji Upadhyaya, Professor, Department of Zoology, College of Basic Science and Humanities, Orissa University of Agriculture and Technology, Bhubaneswar for his keen supervision, help in every respect and encouragement throughout the period of this manuscript.

I express my sincere gratitude and obligation to Mr Sukanta kumar Pradhan, H.O.D, Department of Bioinformatics, OUAT, Bhubaneswar for his constant supervision, scholarly guidance, timely advice and help throughout the period of my work for preparing this thesis. He had allowed me to use his laboratory and software for various analyses related to my study.

My sincere obligation is due to Dr C.S.K Mishra, H.O.D, Department of Zoology,College of Basic Science and Humanities, O.U.A.T, Bhubaneswar for his co-opration, constant inspiration, guidance and encouragement during the period of work. His deep trust in me was the source of energy to the complete education.

I am grateful to Dr Bikash Chandra Panda, Director, College of Basic Science and Humanities, O.U.A.T, Bhubaneswar for his encouragement and co-operation.

I acknowledge, with highest sense of respect and regard to my parents especially my father-my first teacher, mentor and guide forever. His constant motivation has always pushed me through difficult situations and made my goal reachable. I cannot thank him enough.

It would not be an exaggeration to be thankful the vital support from my family and friends as phenomenal. I needed them the most when my other work had to be delegated in favour of the research I had to carry out. I am grateful to them all.

The thanks giving is incomplete without remembering help provided by the staff of the Department of Zoology, College of Basic Science and Humanities, O.U.A.T. I express my obligation to Dr Ashis K. Mohanty, Dr Amita Gupta, Professor and Mr

SriangaTilak Pattnaik, Asst. Professorfor their guidance and co-operation. I am thankful to Mr Rudhish Acharya, Demonstrator, and Mr Dushashan and Mr Pabitra for their help and co-operation. My thanks are due to all those persons of O.U.A.T who directly or indirectly contributed to this work.

A hearty gratitude-filled wish to all of them.

Date: **Sayeda Tahera Requab**

Name of the student:	**Miss Sayeda Tahera Requab**
Admission No:	**08ZOL/12**
Title of Thesis:	**Structural, functional and evolutionary analysis of chromate ion transporter proteins using computational tools and techniques**
Degree for which thesis is submitted:	**Master of Science in Zoology**
Name of the Department:	**Department of Zoology**
College & University:	**College of Basic Science and Humanities, O.U.A.T, Bhubaneswar-3**
Year of Submission:	**2014**
Name of the Advisor:	**Dr Netaji Upadhyaya, Professor, Department of Zoology**

ABSTRACT

A transporter protein is a protein that serves the function of moving other materials within an organism. Transporter proteins are vital to the growth and life of all living things. Transporter proteins in biological membranes may be divided into channels and carriers. Channels function as selective pores that open in response to a chemical or electrophysiological stimulus, allowing movement of a solute down an electrochemical gradient.Active carrier proteins use an energy producing process to translocate a substrate against a concentration gradient. Secondary active transporters use the movement of a solute down a concentration gradient to drive the translocation of another substrate across a membrane. ATP-binding cassette (ABC) transporters couple hydrolysis of adenosine triphosphate (ATP) to the translocation of various substrates across cell membranes. **ChrA is a membrane protein that confers resistance to the toxic ion chromate through the energy-dependent chromate efflux from the cytoplasm. In the protein databases, ChrA is a member of the chromate ion transporter (CHR) superfamily, composed of at least several dozens of members, distributed in the three domains of life. The aim of this work was to perform a phylogenetic analysis of the CHR superfamily. Chromate reduction is carried out by chromate reductases from diverse bacterial species generating Cr (III) that may be detoxified by other mechanisms. 358 numbers of chrA proteins were retrieved from NCBI protein database comprising different organisms (bacteria and fungus) out of which 237 are long chain chromate ion transporter (LCHR) (236-bacterial, 1-fungal) and rest 121 are short chain chromate ion transporter (SCHR) (all bacterial). The divergence studies among LCHR proteins and SCHR proteins were carried out. It was observed that in LCHR 237 proteins were grouped into 4 subgroups and in case of SCHR 121 proteins were grouped in to 6 subgroups.** There were 25 conserved sites and 638 variable sites in LCHR proteins . There were no conserved sites but the variable sites were 293 in SCHR proteins. Secondary structure (Alpha helix, extended strand, random coil) were predicted of the both LCHR and SCHR proteins in %. The % of alpha helix and random coil of chromate ion transporter proteins were more than the extended strand in both LCHR & SCHR proteins. Functional domains were predicted of all LCHR proteins. LCHR proteins contained chromate transporter domains (IPR014047 & IPR003370) and other domain like IPR006187. Most of the LCHR proteins contained both chromate transporter domains IPR014047 and

IPR003370. One LCHR protein of fungi that was *Coprinopsis cinerea* having amino acid length 546 and another LCHR protein of one organism that was *Kyrpidia tusciae* having amino acid length 252 contained only one chromate transporter domain IPR003370. But one LCHR protein of one organism that was *Arthrobacter sp.* having amino acid length 277 contained one chromate transporter domain IPR003370 and another claudin domain IPR006187. Functional domains were predicted of all SCHR proteins. SCHR proteins contained chromate transporter domain (IPR003370) and some other domains like IPR011006, IPR001789, IPR011991, IPR016032, IPR000792 and IPR024414. Most of the SCHR proteins contained chromate transporter domain IPR003370. One SCHR protein of one organism that was *Roseburia hominis* having amino acid length 129 contained uncharacterized domain IPR024414. But one SCHR protein of one organism that was *Corynebacterium diphtheriae* having amino acid length 199 contained five types of other domains IPR011006, IPR001789, IPR011991, IPR016032 and IPR000792. **The 3D structure of the protein chrA of** *Burkholderia pseudomallei (*AHK64337) was predicted using modeler by taking the suitable template of 2C5QA (*Saccaromyces cerevisiae*). The structural study implies that protein having more random coils, 6 alpha helixes, 6 beta sheets. The present study will help to understand the mechanism of chromate resistance of different micro-organisms.

Key Words: LCHR, SCHR, **phylogenetic analysis, chromate reductases, chromate ion transporter.**

Dr Netaji Upadhyaya — **ADVISOR**

Sayeda Tahera Requab — **AUTHOR**

CONTENTS

CHAPTER	PARTICULARS	PAGE NO
1	INTRODUCTION	1-6
2	REVIEW OF LITERATURE	7-16
3	MATERIALS & METHODS	17-33
4	RESULT & DISCUSSION	34-99
5	SUMMARY	100-101
	REFERENCE	

LIST OF FIGURES

FIGURE NO	PARTICULARS	PAGE
1.	Home page of NCBI	18
2.	Home page of MEGA	19
3.	Home page of GOR IV	20
4.	Home page of InterProScan 4	21
5.	Home page of BLASTP	24
6.	Home page for predicting 3D structure	25
7.	Home page of externally by SAVES SERVER	28
8.	Sequences of LCHR protein	35
9.	Alignment by muscle of LCHR protein	35
10.	Alignment view of LCHR protein	36
11.	Conserved site of LCHR protein	36
12.	Variable site of LCHR protein	37
13.	Phylogenetic analysis of LCHR protein	38
14.	Sequences of SCHR protein	39
15.	Alignment view of SCHR protein	39
16.	Conserved sites of SCHR protein	40
17.	Variable sites of SCHR protein	40
18.	Phylogenetic analysis of SCHR protein	41
19.	Comparison of the templates	95
20.	Alignment of target protein with the template	96
21.	3D structure of the target protein (AHK64337)	97
22.	Ramchandran Plot of the target protein	98
23.	Graphical representation of the model by verify 3D	98
24.	The Ramchandran Plot in .jpg format	99

LIST OF TABLES

TABLE NO	PARTICULARS	PAGE
1.	Prediction of secondary structure of LCHR protein	42-48
2.	Prediction of secondary structure of SCHR protein	48-51
3.	Functional analysis of domain of LCHR protein	55-76
4.	Functional analysis of domain of SCHR protein	85-94
5.	Homologos protein of (*Burkholderia pseudomallei*) target protein	95

LIST OF ABBREVIATIONS

NCBI	National Centre For Biotechnology Information.
CHR	Chromate Ion Transporter.
LCHR	Long Chain Sequence.
SCHR	Short Chain Sequence.
FASTA	Fast Alignment.
MEGA	Molecular Evolutionary Genetics Analysis.
GOR IV	Garnier-Osguthorpe-Robson IV.
BLAST	Basic Local Alignment Search Tool.
NumPy	Numerical Python

CHAPTER-I

1. INTRODUCTION

1. WHAT IS CHROMIUM?

Chromium is an element which has the symbol Cr and atomic number 24. It is the first element in group 6. It is a steely-gray, lustrous, hard and brittle metal which has a high melting point. It is found in about 0.0122% of the Earth's crust and it is also a very useful industrial metal.

Chromium is an essential mineral consumed through the diet. It is found in trace amounts in plant products, especially grain. Chromium exists in many natural foods including brewer's yeast, meats, potato, cheese, fresh fruits and vegetables.

It is called an "essential trace element" because very small amounts of chromium are necessary for human health. It maintains proper carbohydrate and lipid metabolism. Chromic ions bind to an oligopeptide chromodulin as a response to insulin. Then this chromic ion saturated chromodulin binds to insulin stimulated insulin receptor. There it activates the tyrosinase activity to amplify the action of insulin in transporting sugars into the cells. [1]

It slows the loss of calcium, so it helps prevent bone loss.

Chromium is good in small quantity but larger quantity can be extremely harmful to human because chromium^{3+} is mutagenic and genotoxic. [2] Chromium is well known toxic metal for microorganisms and plants.

2. FORM OF CHROMIUM

Chromium occurs in the environment predominantly in two valence states; trivalent chromium Cr (III) and hexavalent chromium Cr (VI). Exposure may occur from natural or industrial source of chromium. Natural source includes rocks, animals, plants, soil and volcanic dust and gases. The most important industrial sources of chromium are ferrochrome production industries. Chromium is present in atmosphere as trivalent chromium in the form of small particles or aerosols.

Cr (III) occurs naturally and is an essential nutrient required for normal glucose, protein and fat metabolism. It is stable and sparingly soluble in water and is relatively innocuous. It is labile and kinetically very slow to react or form complexes.

Cr (VI) is the second most stable and readily soluble in water. It is highly toxic, carcinogenic, mutagenic and teratogenic due to its strong oxidizing nature.

3. CHROMIUM TOXICITY

With the development of chromium uses, chromium toxicity and adverse effects of chromium compounds in human health were being discovered. Chromium enters the air, water and soil mostly in the chromium (III) and chromium (VI) forms. In air, chromium compounds are present mostly as fine dust particles which eventually settle over land and water. Chromium can strongly attach to soil and only a small amount can dissolve in water and move deeper in the soil to underground water.

Between trivalent and hexavalent chromium, hexavalent chromium causes both acute and chronic toxicity; trivalent chromium may show some toxicity in very large quantities. The most important toxic effects, after contact, inhalation, or ingestion of hexavalent chromium compounds are dermatitis, allergic and eczematous skin reactions, skin and mucous membrane ulcerations, perforation of the nasal septum, allergic

asthmatic reactions, bronchial carcinomas, gastro-enteritis, hepato-cellular deficiency and renal deficiency. [3]

4. BIOREMEDIATION OF CHROMIUM TOXICITY

Chromium is an important toxic environmental pollutant. Chromium pollution results largely from industrial activities, but other natural and anthropogenic sources also contribute to the problem. Plants that are exposed to environmental contamination by chromium are affected in diverse ways, including a tendency to suffer metabolic stress. The stress imposed by chromium exposure also extends to oxidative metabolic stress in plants that leads to the generation of active toxic oxygen free radicals. Such active free radicals degrade essential biomolecules and distort plant biological membranes. We describe sources of environmental chromium contamination, and provide information about the toxic impact of chromium on plant growth and metabolism. We address different phytoremediation processes that are being studied for use worldwide, in contaminated regions, to address and mitigate chromium pollution.

There has been a long history of attempts to successfully mitigate the toxic effects of chromium-contaminated soil on plants and other organisms. One common approach, the shifting of polluted soil to landfills, is expensive and imposes environmental risks and health hazards of its own. Therefore, alternative eco-friendly bioremediation approaches are much in demand for cleaning chromium-polluted areas. To achieve its cleaning effects, bioremediation utilizes living organisms (bacteria, algae, fungi, and plants) that are capable of absorbing and processing chromium residues in ways which amend or eliminate it. Phytoremediation (bioremediation with plants) techniques are increasingly being used to reduce heavy metal contamination and to minimize the hazards of heavy metal toxicity.

To achieve this, several processes, viz., rhizofiltration, phytoextraction, phyto-detoxification, phytostabilization, and phytovolatilization, have been developed and are showing utility in practice, or promise. Sources of new native hyper-accumulator plants for use at contaminated sites are needed and constitute a key goal of ongoing phytoremediation research programs. Such new plants are needed to enhance the attractiveness of phytoremediation as an effective, affordable, and eco-friendly technique to achieve successful clean-up of metal-contaminated sites worldwide. [4]

5. MECHANISM OF CHROMATE RESISTANCE

Chromium is a non-essential and well-known toxic metal for microorganisms and plants. The widespread industrial use of this heavy metal has caused it to be considered as a serious environmental pollutant. At the intracellular level, however, Cr (III) seems to be responsible for most toxic effects of chromium. Cr (VI) is usually present as the oxyanion chromate. Inhibition of sulfate membrane transport and oxidative damage to biomolecules are associated with the toxic effects of chromate in bacteria. Several bacterial mechanisms of resistance to chromate have been reported.

The best characterized mechanisms comprise efflux of chromate ions from the cell cytoplasm and reduction of Cr (VI) to Cr (III). Chromate efflux by the ChrA transporter has been established in *Pseudomonas aeruginosa* and *Cupriavidus metallidurans* (formerly *Alcaligenes eutrophus*) and consists of an energy-dependent process driven by the membrane potential. The CHR protein family, which includes putative ChrA orthologs, currently contains about 135 sequences from all three domains of life.

Chromate reduction is carried out by chromate reductases from diverse bacterial species generating Cr (III) that may be detoxified by other mechanisms. Most

characterized enzymes belong to the widespread NAD (P) H-dependent flavoprotein family of reductases. Several examples of bacterial systems protecting from the oxidative stress caused by chromate have been described. Other mechanisms of bacterial resistance to chromate involve the expression of components of the machinery for repair of DNA damage, and systems related to the homeostasis of iron and sulfur. [5]

6. CHROMATE ION TRANSPORTER PROTEIN

A transporter protein is a protein that serves the function of moving other materials within an organism. Transporter proteins are vital to the growth and life of all living things. Transporter proteins in biological membranes may be divided into channels and carriers. Channels function as selective pores that open in response to a chemical or electrophysiological stimulus, allowing movement of a solute down an electrochemical gradient.

Carrier proteins are proteins involved in the movement of ions, small molecules, or macromolecules, such as another protein, across a biological membrane. Carrier proteins are integral membrane proteins; that is, they exist within and span the membrane across which they transport substances. Active carrier proteins use an energy producing process to translocate a substrate against a concentration gradient. Secondary active transporters use the movement of a solute down a concentration gradient to drive the translocation of another substrate across a membrane. ATP-binding cassette (ABC) transporters couple hydrolysis of adenosine triphosphate (ATP) to the translocation of various substrates across cell membranes.

ChrA is a membrane protein that confers resistance to the toxic ion chromate through the energy-dependent chromate efflux from the cytoplasm. In the protein databases, ChrA is a member of the chromate ion transporter (CHR)

superfamily, composed of at least several dozens of members, distributed in the three domains of life. The aim of this work was to perform a phylogenetic analysis of the CHR superfamily. [6]

OBJECTIVES:

- To collect proteins belonging to CHR superfamily involve in chromate resistance.
- To perform a phylogenetic analysis of the CHR superfamily.
- To predict secondary structure of chromate ion transporter proteins.
- To find out the functional domains of chromate ion transporter proteins.
- To predict 3D structure of chromate ion transporter of ***Burkholderia pseudomallei.***

CHAPTER-II

2. REVIEW OF LITERATURE

1. Toxic effects of chromium and its compounds

Chromium was discovered in 1797 by Vauquelin. Numerous industrial applications raised chromium to a very important economic element. At the same time, with the development of its uses, the adverse effects of chromium compounds in human health were being defined. Trivalent chromium is an essential trace element in humans and in animals. Chromium as pure metal has no adverse effect. Little toxic effect is attributed to trivalent chromium when present in very large quantities. Both acute and chronic toxicity of chromium are mainly caused by hexavalent compounds. The most important toxic effects, after contact, inhalation, or ingestion of hexavalent chromium compounds are the following: dermatitis, allergic and eczematous skin reactions, skin and mucous membrane ulcerations, perforation of the nasal septum, allergic asthmatic reactions, bronchial carcinomas, gastro-enteritis, hepatocellular deficiency, and renal oligo anuric deficiency. Prevention of occupational risks, biological monitoring of workers, and treatment of poisoning are also reported. [3]

2. Attenuation of chromium toxicity by bioremediation technology

Chromium is an important toxic environmental pollutant. Chromium pollution results largely from industrial activities, but other natural and anthropogenic sources also contribute to the problem. Plants that are exposed to environmental contamination by chromium are affected in diverse ways, including a tendency to suffer metabolic stress. The stress imposed by Cr exposure also extends to oxidative metabolic stress in plants that leads to the generation of active toxic oxygen free radicals. Such active free radicals degrade essential biomolecules and distort plant biological membranes. In this chapter,

we describe sources of environmental chromium contamination, and provide information about the toxic impact of chromium on plant growth and metabolism. In addition, we address different phytoremediation processes that are being studied for use worldwide, in contaminated regions, to address and mitigate Cr pollution. There has been a long history of attempts to successfully mitigate the toxic effects of chromium-contaminated soil on plants and other organisms. One common approach, the shifting of polluted soil to landfills, is expensive and imposes environmental risks and health hazards of its own. Therefore, alternative eco-friendly bioremediation approaches are much in demand for cleaning chromium-polluted areas. To achieve its cleaning effects, bioremediation utilizes living organisms (bacteria, algae, fungi, and plants) that are capable of absorbing and processing chromium residues in ways which amend or eliminate it. Phytoremediation (bioremediation with plants) techniques are increasingly being used to reduce heavy metal contamination and to minimize the hazards of heavy metal toxicity. To achieve this, several processes, viz., rhizofiltration, phytoextraction, phytodetoxification, phytostabilization, and phytovolatilization have been developed and are showing utility in practice, or promise. Sources of new native hyperaccumulator plants for use at contaminated sites are needed and constitute a key goal of ongoing phytoremediation research programs. Such new plants are needed to enhance the attractiveness of phytoremediation as an effective, affordable, and eco-friendly technique to achieve successful clean-up of metal-contaminated sites worldwide. [4]

3. Mechanisms of bacterial resistance to chromium compounds

Chromium is a non-essential and well-known toxic metal for microorganisms and plants. The widespread industrial use of this heavy metal has caused it to be considered as a serious environmental pollutant. Chromium exists in nature as two main species, the trivalent form, Cr (III), which is relatively innocuous, and the hexavalent form, Cr (VI),

considered a more toxic species. At the intracellular level, however, Cr (III) seems to be responsible for most toxic effects of chromium. Cr (VI) is usually present as the oxyanion chromate. Inhibition of sulfate membrane transport and oxidative damage to biomolecules are associated with the toxic effects of chromate in bacteria. Several bacterial mechanisms of resistance to chromate have been reported. The best characterized mechanisms comprise efflux of chromate ions from the cell cytoplasm and reduction of Cr (VI) to Cr (III). Chromate efflux by the ChrA transporter has been established in *Pseudomonas aeruginosa* and *Cupriavidus metallidurans* (formerly *Alcaligenes eutrophus*) and consists of an energy-dependent process driven by the membrane potential. The CHR protein family, which includes putative ChrA orthologs, currently contains about 135 sequences from all three domains of life. Chromate reduction is carried out by chromate reductases from diverse bacterial species generating Cr (III) that may be detoxified by other mechanisms. Most characterized enzymes belong to the widespread NAD (P) H-dependent flavoprotein family of reductases. Several examples of bacterial systems protecting from the oxidative stress caused by chromate have been described. Other mechanisms of bacterial resistance to chromate involve the expression of components of the machinery for repair of DNA damage and systems related to the homeostasis of iron and sulfur. [5]

4. Phylogenetic analysis of the chromate ion transporter (CHR) superfamily

ChrA is a membrane protein that confers resistance to the toxic ion chromate through the energy-dependent chromate efflux from the cytoplasm. In the protein databases, ChrA is a member of the chromate ion transporter (CHR) superfamily, composed of at least several dozens of members, distributed in the three domains of life. The aim of this work was to perform a phylogenetic analysis of the CHR superfamily. An exhaustive search for ChrA homologous proteins was carried out at the

National Center for Biotechnology Information database. One hundred and thirty-five sequences were identified as members of the CHR superfamily [77 long-chain sequences, or bidomains (LCHR), and 58 short-chain sequences, or monodomains (SCHR)], organized mainly as tandem pairs of genes whose resultant proteins probably possess oppositely oriented membrane topology. LCHR sequences were split into amino and carboxyl domains and the resultant domains were aligned with the SCHR proteins. A phylogenetic tree was reconstructed using four different methods, obtaining similar results. The domains were grouped into three clusters: the SCHR proteins cluster, the amino domain cluster of LCHR proteins and the carboxyl domain cluster of LCHR proteins. These results, as well as differences in the genomic context of CHR proteins, enabled the proteins to be sorted into two families (SCHR and LCHR), and 10 subfamilies. Evidence was found suggesting an ancient origin of LCHR proteins from the fusion of two SCHR protein-encoding genes; however, some secondary events of fusion and fission may have occurred later. The separate distribution of the LCHR and SCHR proteins, differences in the genomic context in both groups and the fact that chromate transport has been demonstrated only in LCHR proteins suggest that the CHR proteins comprise two or more paralogous groups in the CHR superfamily. [6]

5. Genes related to chromate resistance by *Pseudomonas aeruginosa* PAO1

Chromate-hypersensitive mutants of the *Pseudomonas aeruginosa* PAO1 strain were isolated using transposon-insertion mutagenesis. Comparison of the nucleotide sequences of the regions interrupted in the mutants with the PAO1 genome revealed that the genes affected in three mutant strains were oprE (ORF PA0291), rmlA (ORF PA5163), and ftsK (ORF PA2615), respectively. A relationship of these genes with chromate tolerance has not been previously reported. No other phenotypic changes were

observed in the oprE mutant but its resistance to chromate was not fully restored by expressing the ChrA protein, which extrudes chromate ions from the cytoplasm to the periplasmic space. These data suggest that OprE participates in the efflux of chromate from the periplasm to the outside. Increased susceptibility of the rmlA mutant to the metals cadmium and mercury and to the anion-superoxide generator paraquat suggests a protective role of LPS against chromate toxicity. A higher susceptibility of the ftsK mutant to compounds affecting DNA structure (ciprofloxacin, tellurite, mitomycin C) suggests a role of FtsK in the recombinational repair of DNA damage caused by chromate. In conclusion, the *Pseudomonas aeruginosa* genome contains diverse genes related to its intrinsic resistance to chromate. Systems pertaining to the outer membrane (OprE), the cell wall (LPS), and the cytoplasm (FtsK) were identified in this work as involved in chromate protection mechanisms. [7]

6. Essential residues in the chromate transporter ChrA of *Pseudomonas aeruginosa*

The chrA gene of *Pseudomonas aeruginosa* plasmid pUM505 encodes the hydrophobic protein ChrA, which confers resistance to chromate by the energy-dependent efflux of chromate ions. Chromate-sensitive mutants were isolated by in vivo random mutagenesis. Transport experiments with cell suspensions of selected mutants showed that 51CrO4 (2-) extrusion was drastically lowered as compared to suspensions of the strain with the wild-type plasmid, confirming that the mutations affected a chromate efflux system. DNA sequence analysis showed that most point mutations affected amino acids clustered in the N-terminal half of ChrA, altering either cytoplasmic regions or transmembrane segments, and replaced residues moderately to highly conserved in ChrA homologs. PhoA and LacZ translational fusions were used to confirm the membrane topology at the N-terminal half of the ChrA protein. [8]

7. Effect of chromates on ciliated cells of rat tracheal epithelium

Exposure to chromates is an occupational hazard that may result in disease of the respiratory tract. Specific subcellular effects of exposure to this chemical that might be the basis for observed pathologic findings were assessed in tracheal organ culture and excised whole tracheas. Chromate concentrations of 10 mg/ml or more were found to induce ciliostasis within 20 minutes and to kill cells within 24 hours. Levels between 10mug/ml and 100 mug/ml were not ciliostatic but resulted in cytotoxicity after a delay of several hours. Ultrastructural alterations associated with chromate-induced cytotoxicity are similar to the progressive degenerative changes that result from other chemical injuries. An inhibitory effect of chromate on protein synthesis, RNA synthesis, or succinic dehydrogenase activity could not be demonstrated. The data suggest that damage to the plasmalemma by chromates may be responsible for the observed change. [9]

8. Resistance to apoptosis, increased growth potential, and altered gene expression in cells that survived genotoxic hexavalent chromium [Cr(VI)] exposure

Certain hexavalent chromium [Cr (VI)] compounds are known genotoxic respiratory carcinogens, which induce apoptosis as a predominant mode of cell death. Selection of cells that are resistant to apoptosis may be a factor in tumour progression. We developed sub-populations of telomerase-transfected human fibroblasts (BJ-hTERT) that survived a 99% clonogenically lethal exposure to Cr (VI) (B-5Cr). B-5Cr cells were markedly resistant to apoptosis induced by several agents and exhibited increased clonogenic survival, especially at apoptogenic doses. B-5Cr cells did not exhibit altered cellular uptake of Cr(VI) and retained a normal p53 response to Cr(VI) exposure. We conducted large-scale gene expression analysis at different time-points after a secondary genotoxic Cr (VI) insult in B-5Cr and BJ-hTERT cells using Affymetrix Genechip

human genome arrays. Cr (VI) exposure led to differential regulation of many genes, which affect a diverse set of cellular activities such as transcription, signal transduction, stress response, cell adhesion, DNA repair, apoptosis and cell cycle modulation. We compared Cr (VI)-induced altered gene expression in the B-5Cr cells to that in the parental cells and identified 223, 147 and 204 genes with at least a two-fold difference in expression at 4, 8 and 18 h after exposure, respectively. Cluster analysis by gene function revealed altered expression of genes involved in apoptosis, cell cycle regulation and DNA repair. Our data suggest an alteration in gene expression that may favor cell survival and/or incomplete DNA repair after genotoxic exposure. Selection of cells with altered expression of these genes may constitute the early stages of tumour progression. [10]

9. Mismatch repair proteins are activators of toxic responses to chromium-DNA damage

Chromium (VI) is a toxic and carcinogenic metal that causes the formation of DNA phosphate-based adducts. Cr-DNA adducts are genotoxic in human cells, although they do not block replication in vitro. Here, we report that induction of cytotoxicity in Cr (VI)-treated human colon cells and mouse embryonic fibroblasts requires the presence of all major mismatch repair (MMR) proteins. Cr-DNA adducts lost their ability to block replication of Cr-modified plasmids in human colon cells lacking MLH1 protein. The presence of functional mismatch repair caused induction of p53-independent apoptosis associated with activation of caspases 2 and 7. Processing of Cr-DNA damage by mismatch repair resulted in the extensive formation of gamma-H2AX foci in G(2) phase, indicating generation of double-stranded breaks as secondary toxic lesions. Induction of gamma-H2AX foci was observed at 6 to 12 h postexposure, which was followed by activation of apoptosis in the absence of significant G(2) arrest. Our results demonstrate

that mismatch repair system triggers toxic responses to Cr-DNA backbone modifications through stress mechanisms that are significantly different from those for other forms of DNA damage. Selection for Cr (VI) resistant, MMR-deficient cells may explain the very high frequency of lung cancers with microsatellite instability among chromate workers. [11]

10. Transcriptional inhibition by carcinogenic chromate: relationship to DNA damage

Hexavalent chromium compounds are carcinogenic to humans, are potent inducers of tumors in experimental animals, and can neoplastically transform cells in culture. In this study, the effects of sodium chromate on the expression of the 78-kDa glucose-regulated protein (GRP78) gene and on general transcription were investigated with respect to the DNA damage induced by this agent. DNA single-strand breaks, DNA-proteincross-links, and chromium-DNA adducts were present in CHO cells immediately after 2 h of treatment with sodium chromate. Subsequently, these types of damage were repaired at different rates. Single-strand breaks were essentially repaired after 8 h. By 24 h post treatment, no cross-links remained in cells exposed to 30 or 150 microM chromate, although cells treated with the 300-microM concentration still contained cross-links at that time. DNA-chromium adducts remained unrepaired for at least 32 h. The moderate constitutive level of GRP78 mRNA was not affected by chromate. Chromate did, however, suppress induction of this gene by tunicamycin in a concentration-and time-dependent manner. Thirty micromolar sodium chromate (96% survival), which caused the least DNA damage, had no effect on GRP78 induction, general RNA synthesis, or mRNA synthesis. Induction of GRP78 was suppressed immediately and 12 h after treatment with 150 microM chromate (54% survival), although there was a partial recovery of induction at 24 h after treatment, which

correlated with the repair of DNA-protein cross-links. In contrast, both total cytoplasmic RNA and mRNA synthesis were suppressed by approximately 60-75% for at least 32 h by 150 microM chromate. At the 300-microM concentration (8% survival), where DNA-protein cross-links persisted beyond 24 h, GRP78 induction was totally suppressed for at least 24 h, while total RNA and mRNA synthesis were suppressed by 80-90% for at least 32 h. Overall, the effects of chromate on GRP78 induction correlated most closely with the presence of DNA-protein cross-links, but suppression of total RNA and mRNA synthesis correlated with the presence of DNA-chromium adducts. These results indicate that chromate exerts differential effects on the induction of the GRP78 gene and on general transcription. [12]

11. Membrane topology of the chromate transporter ChrA of *Pseudomonas aeruginosa*

The membrane topology of the plasmid-encoded *Pseudomonas aeruginosa* ChrA protein, which effluxes chromate ions, was determined by the analysis of translational fusions with reporter enzymes alkaline phosphatase and beta-galactosidase. A novel 13-TMS (transmembrane segments) topology, with the N-terminus located in the cytoplasm and the C-terminus in the periplasmic space was consistent with the enzyme activities determined in both *Escherichia coli*_and *Pseudomonas aeruginosa*. Alignment of the two halves of ChrA showed significant sequence homology, with TMS I, II, III, IV, V and VI displaying similarity to TMS VIII, IX, X, XI, XII and XIII, respectively, although with opposite membrane orientations. This suggests that ChrA arose from the duplication of a gene encoding a 6-TMS ancestral protein, followed by the insertion of extra TMS VII. These data also suggest that the two halves of ChrA may carry out distinct functions for the transport of chromate. [13]

12. Chromate toxicity and the role of sulfur

The molecular mode(s)-of-action of the toxic metal chromium has yet to be fully resolved. This Mini review focuses on interactions between chromate and sulfur in biological systems. Cr binds sulfur ligands, with cysteine and glutathione having the capacity to aggravate or ameliorate Cr toxicity. Competition between chromate and sulfate for uptake and in metabolism provokes sulfur starvation, which can be growth limiting. Recent data indicate that sulfur deficiency determines protein damage-related Cr toxicity, due to mRNA mistranslation caused by Cr-induced S limitation. Sulfur deprivation could contribute to additional aspects of Cr toxicity, including oxidative DNA damage and Cr related disease. [14]

13. Antiparallel membrane topology of paired short-chain chromate transport proteins in *Bacillus subtilis*

Short-chain monodomain family comprises pairs of membrane proteins of about 200 amino acid residues each that belong to the chromate ion transporter (CHR) superfamily. The short-chain CHR homologous pair Chr3N/Chr3C from *Bacillus subtilis* strain 168 confers chromate resistance only when both proteins are expressed. Membrane topology of the Chr3N and Chr3C proteins was determined in Escherichia coli by the analysis of translational fusions with reporter enzymes alkaline phosphatase and β-galactosidase. Each short-chain CHR protein was found to consist of five transmembrane segments with antiparallel orientation between them. The C terminus of Chr3N is located in the cytoplasm, whereas the C terminus of Chr3C is located in the periplasm. In silico analyses suggest that this antiparallel arrangement is shared by all protein members of the short-chain CHR3 subfamily and that the two Chr3N/Chr3C proteins might carry out distinct functions for the transport of chromate. [15]

CHAPTER-III

3. MATERIALS AND METHODS

A. MATERIALS

I. Collections of chromate ion transporter protein

NCBI (National Centre for Biotechnology Information)

NCBI is part of the United States National Library of Medicine (NLM), a branch of the National Institute of Health. It is located in Bethesda, Maryland. The NCBI houses a series of databases relevant to biotechnology and biomedicine. Major databases include GenBank for DNA sequences and Pubmed, a bibliographic database for the biomedical literature. Other databases include the NCBI Epigenomics database. GenBank coordinates with individual laboratories and other sequence databases such as those of the European Molecular Biology Laboratory (EMBL) and the DNA Data Bank of Japan (DDBJ).

NCBI has grown to provide other databases in addition to GenBank. NCBI provides Gene, Online Mendelian Inheritance in Man, the Molecular Modeling Database (3D protein structures), dbSNP (a database of single-nucleotide polymorphisms), the reference sequence collection, a map of the human genome, and a taxonomy browser, and coordinates with the National Cancer Institute to provide the Cancer Genome Anatomy Project. The NCBI assigns a unique identifier (taxonomy ID number) to each species of organism.

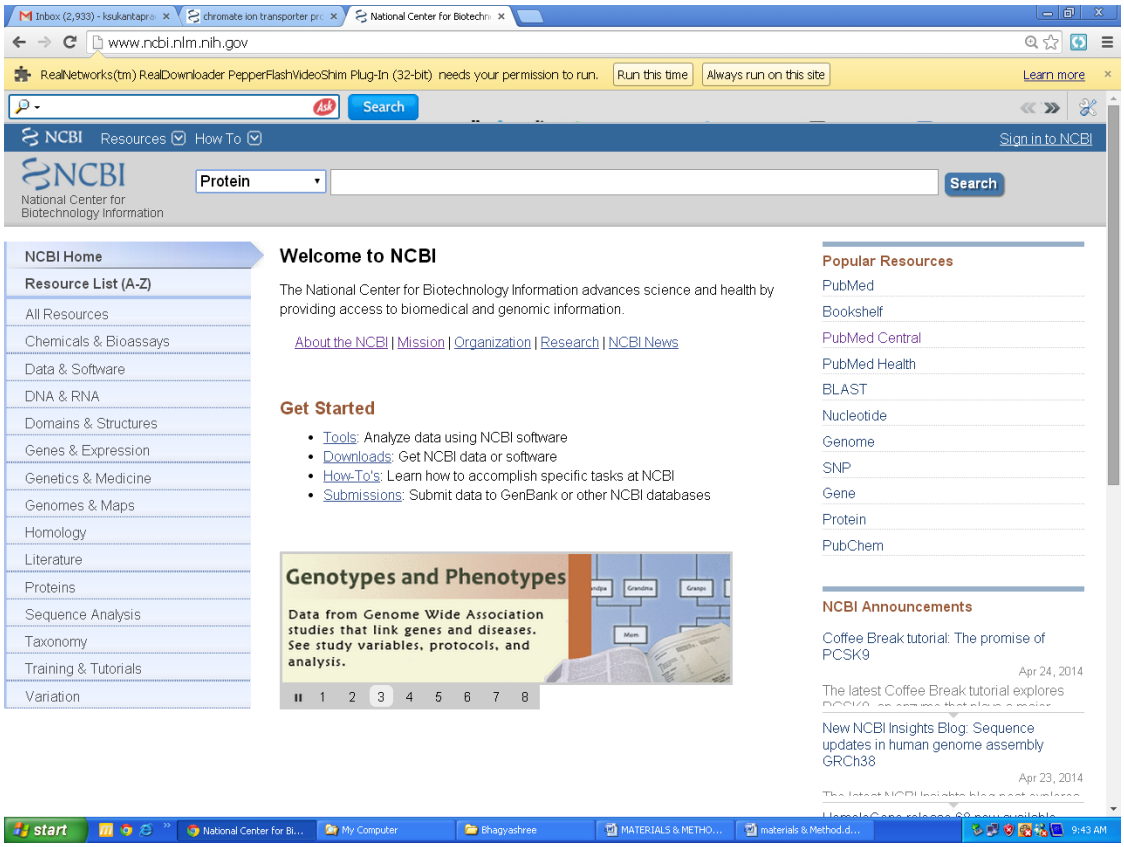


Fig. 1. Home page of NCBI

II. Protein divergence study

MEGA (Molecular Evolutionary Genetics Analysis)

MEGA is an integrated tool for conducting the sequence alignment, inferring phylogenetic trees, estimating divergence times, mining online databases, estimating rates of molecular evolution inferring ancestral sequences & testing evolutionary hypotheses. MEGA is used by biologist in a large number of laboratories for reconstructing for the evolutionary histories of species and inferring the extent and nature of the selective forces shaping the evolution of genes and species.

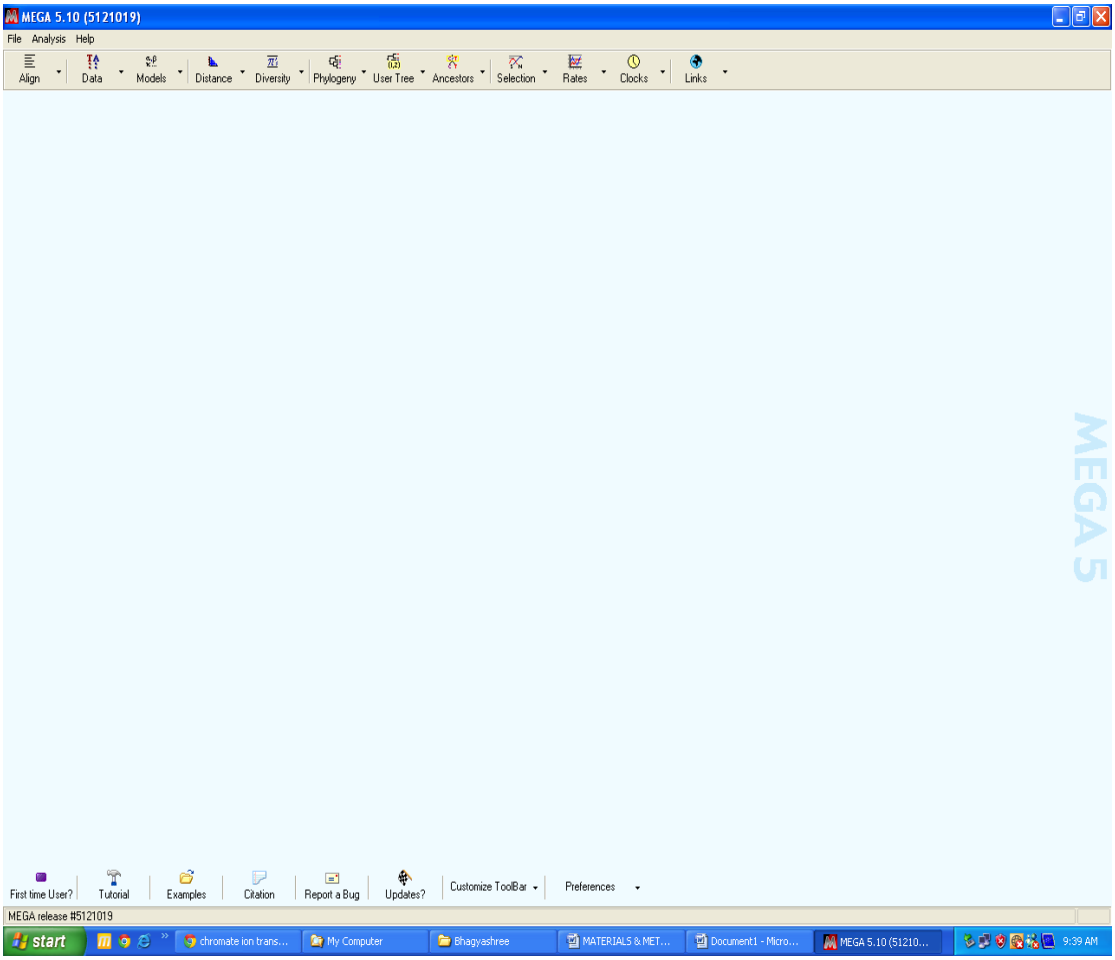


Fig. 2. Home page of MEGA

III . Secondary structure prediction

GORIV (Garnier-Osguthorpe-Robson IV)

GOR IV is the fourth version of GOR secondary structure prediction methods based on the information theory (Garnier et al., 1996). There is no defined decision constant. GOR IV uses all possible pair frequencies within the window of 17 amino acid residues. After cross validation on a data base of 267 proteins, the version IV of GOR has a mean accuracy 64.4% for a three state prediction.

Fig. 3. Home page of GOR IV

IV. Functional Analysis

InterProScan 4

InterProScan is a tool that combines different protein signature recognition methods from the InterPro consortium member databases into one resource. It is the tools for protein sequence classification and comparision. It is used to identify signatures from the InterPro member databases that is Pfam, PROSITE, PRINTS, ProDom, SMART, TIGRFAMs, PIR superfamily, SUPERFAMILY, Gene 3D and PANTHER. InterPro provides functional analysis of proteins by classifying them into families and predicting domains and important sites.

It is a bioinformatics tool that provides a one stop-shop for automated sequence analysis of both protein and nucleic acid. It offers the ability to identify both structural and functional regions of interest based upon methods and models that have been generated by a large number of member databases. These databases use a variety of

different bioinformatic techniques and algorithms, optimised for specific features types. It is able to offer the researcher the ability to quickly characterize a new or novel sequence with considerable confidence.

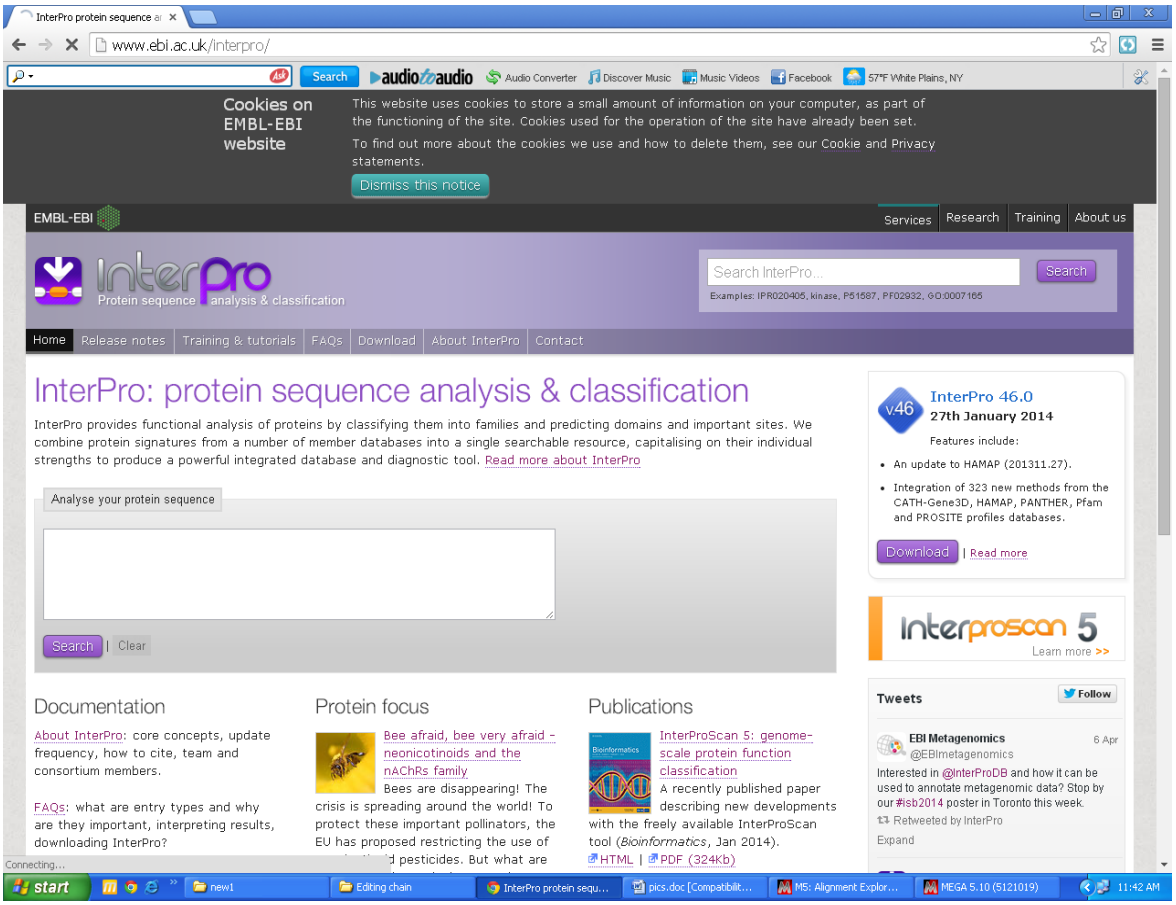


Fig. 4.Home page of InterProScan 4

V. 3D Structure prediction

(a) For finding out the homologs

BLASTP

In bioinformatics, BLAST for Basic Local Alignment Search Tool is algorithm for comparing primary biological sequence information, such as the amino acid sequences of different proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a

certain threshold. Different types of BLASTs are available according to the query sequences. For example, following the discovery of a previously unknown gene in the mouse, a scientist will typically perform a BLAST search of the human genome to see if humans carry a similar gene; BLAST will identify sequences in the human genome that resemble the mouse gene based on similarity of sequence. The BLAST program was designed by Stephen Altschul, Warren Gish, Webb Miller, Eugene Myers, and David J. Lipman at the NIH and was published in the Journal of Molecular Biology in 1990.

Input

Input sequences are in FASTA or Genbank format and weight matrix.

Output

BLAST output can be delivered in a variety of formats. These formats include HTML, plain text, and XML formatting. For NCBI's web-page, the default format for output is HTML. When performing a BLAST on NCBI, the results are given in a graphical format showing the hits found, a table showing sequence identifiers for the hits with scoring related data, as well as alignments for the sequence of interest and the hits received with corresponding BLAST scores for these. The easiest to read and most informative of these is probably the table.

If one is attempting to search for a proprietary sequence or simply one that is unavailable in databases available to the general public through sources such as NCBI, there is a BLAST program available for download to any computer, at no cost. This can be found at BLAST+ executables. There are also commercial programs available for purchase. Databases can be found from the NCBI site, as well as from Index of BLAST databases .

Process

Using a heuristic method, BLAST finds similar sequences, not by comparing either sequence in its entirety, but rather by locating short matches between the two sequences. This process of finding initial words is called seeding. It is after this first match that BLAST begins to make local alignments. While attempting to find similarity in sequences, sets of common letters, known as words, are very important. For example, suppose that the sequence contains the following stretch of letters, GLKFA. If a BLASTP was being conducted under default conditions, the word size would be 3 letters. In this case, using the given stretch of letters, the searched words would be GLK, LKF, KFA. The heuristic algorithm of BLAST locates all common three-letter words between the sequence of interest and the hit sequence, or sequences, from the database. These results will then be used to build an alignment. After making words for the sequence of interest, neighborhood words are also assembled. These words must satisfy a requirement of having a score of at least the threshold T, when compared by using a scoring matrix. One commonly used scoring matrix for BLASTP searches is BLOSUM62, although the optimal scoring matrix depends on sequence similarity. Once both words and neighborhood words are assembled and compiled, they are compared to the sequences in the database in order to find matches. The threshold score T determines whether or not a particular word will be included in the alignment. Once seeding has been conducted, the alignment, which is only 3 residues long, is extended in both directions by the algorithm used by BLAST. Each extension impacts the score of the alignment by either increasing or decreasing it. Should this score be higher than a pre-determined T, the alignment will be included in the results given by BLAST. However, should this score be lower than this pre-determined T, the alignment will cease to extend, preventing areas of poor alignment from being included in the

BLAST results. Note, that increasing the T score limits the amount of space available to search, decreasing the number of neighborhood words, while at the same time speeding up the process of BLAST.

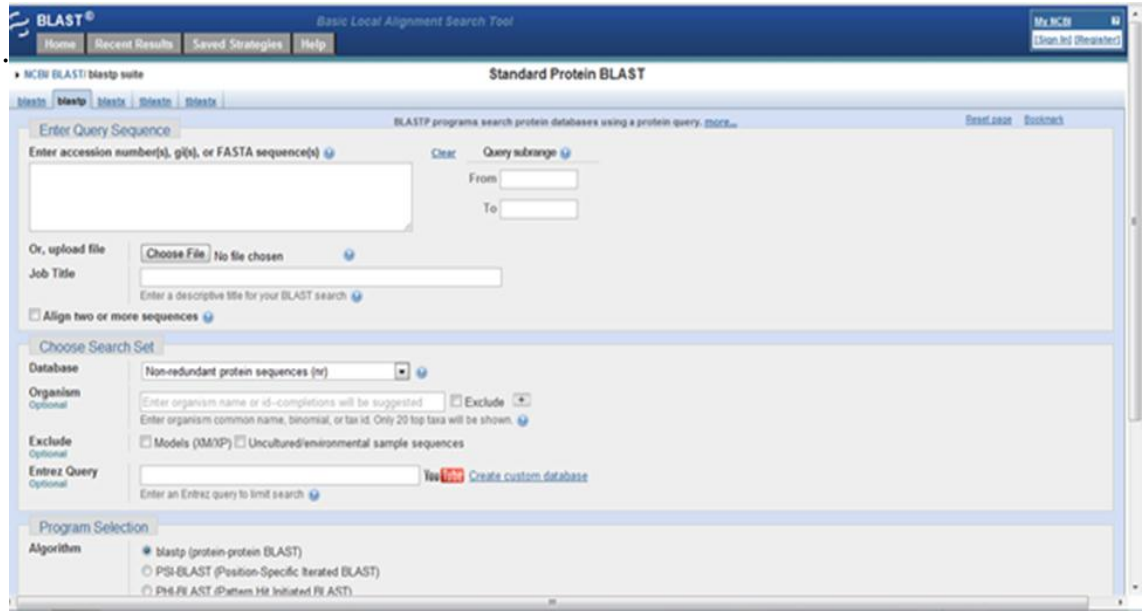


Fig. 5. Home page of BLASTP

(b) For predicting 3D Structure

Modeller

Modeller is a computer program that models three-dimensional structures of proteins and their assemblies by satisfaction of spatial restraints. Modeller is most frequently used for homology or comparative protein structure modeling. The user provides an alignment of a sequence to be modeled with known related structures and Modeller will automatically calculate a model with all non-hydrogen atoms.

More generally, the inputs to the program are restraints on the spatial structure of the amino acid sequence(s) and ligands to be modeled. The output is a 3D structure that satisfies these restraints as well as possible. Restraints can in principle be derived from a number of different sources. These include related protein structures (comparative modeling), NMR experiments (NMR refinement), rules of secondary structure packing

(combinatorial modeling), cross-linking experiments, fluorescence spectroscopy, image reconstruction in electron microscopy, site-directed mutagenesis, intuition, residue–residue and atom–atom potentials of mean force, etc.

The restraints can operate on distances, angles, dihedral angles, pairs of dihedral angles and some other spatial features defined by atoms or pseudo atoms. Presently, Modeller automatically derives the restraints only from the known related structures and their alignment with the target sequence.

A 3D model is obtained by optimization of a molecular probability density function (pdf). The molecular pdf for comparative modeling is optimized with the variable target function procedure in Cartesian space that employs methods of conjugate gradients and molecular dynamics with simulated annealing.

Modeller can also perform multiple comparisons of protein sequences and/or structures, clustering of proteins, and searching of sequence databases. The program is used with a scripting language and does not include any graphics. It is written in standard Fortran 90 and will run on Unix, Windows, or Mac computers.

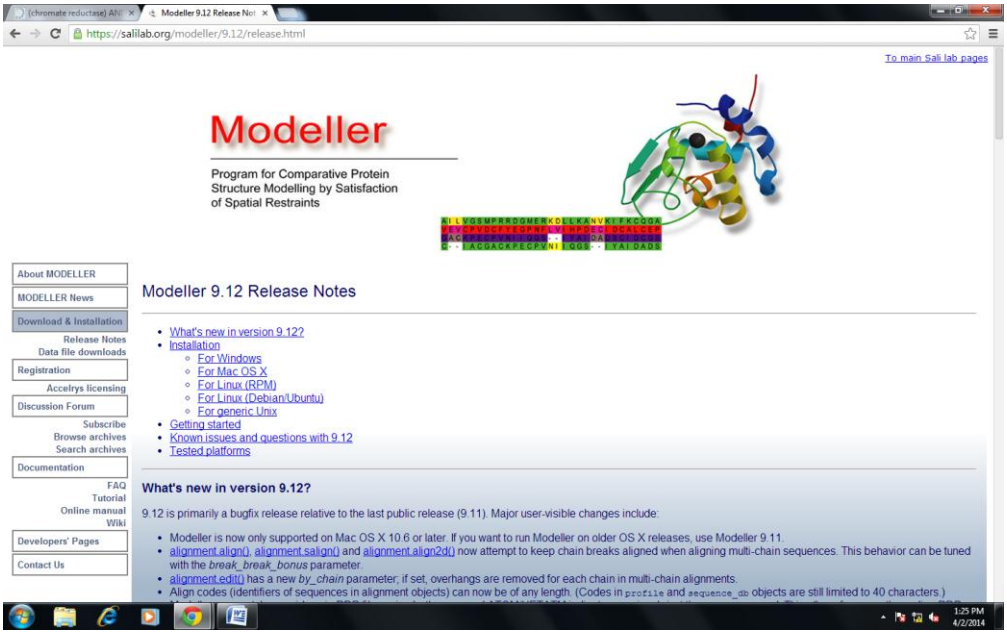


Fig. 6. Home page for predicting 3D structure

(c) For model evaluation

After a model is built, it is important to check it for possible errors. The quality of a model can be approximately predicted from the sequence similarity between the target and the template. Sequence identity above 30% is a relatively good predictor of the expected accuracy of a model. In Evaluation Modeller there are 2 types of evaluation can be carried out.

1. **Internally by Modeller**

- **Python2.6**

Python is a widely used general-purpose, high-level programming language. Its design philosophy emphasizes code readability, and its syntax allows programmers to express concepts in fewer lines of code than would be possible in languages such as C. The language provides constructs intended to enable clear programs on both a small and large scale. Python is often used as a scripting language, but is also used in a wide range of non-scripting contexts. Python interpreters are available for many operating systems.

Python is a multi-paradigm programming language: object-oriented programming and structured programming are fully supported, and there are a number of language features which support functional programming and aspect-oriented programming. It features a dynamic type system and automatic memory management.

- **NumPy-1.6.0b1-win32-superpack-python2.6.exe**

Numerical Python adds a fast and sophisticated array facility to the Python language. NumPy is the most recent and most actively supported package. NumPy is the fundamental package for scientific computing with Python. It contains among other things:

- a powerful N-dimensional array object.
- sophisticated (broadcasting) functions.
- tools for integrating C/C++ and Fortran code.
- useful linear algebra, Fourier transform, and random number capabilities.

NumPy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined. This allows NumPy to seamlessly and speedily integrate with a wide variety of databases.

- **matplotlib-1.0.1.win32-py2.6.exe**

matplotlib is a python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. matplotlib can be used in python scripts, the python and ipython shell, web application servers, and six graphical user interface toolkits.matplotlib tries to make easy things easy and hard things possible. matplotlib is a plotting library for the Python programming language and its NumPy numerical mathematics extension. It provides an object-oriented API for embedding plots into applications using general-purpose GUItoolkits like wxPython, Qt, or GTK+.The advantages of matplotlib are:

- Nice default plot styles: less code to polish the figure.
- Deep integration with Python.
- matlab style programming interface (this is an advantage for some, but a disadvantage for others).

2. Externally by SAVES SERVER

“Internal" evaluation of self-consistency checks whether or not a model satisfies the restraints used to calculate it. “External" evaluation relies on information that was not used in the calculation of the model.

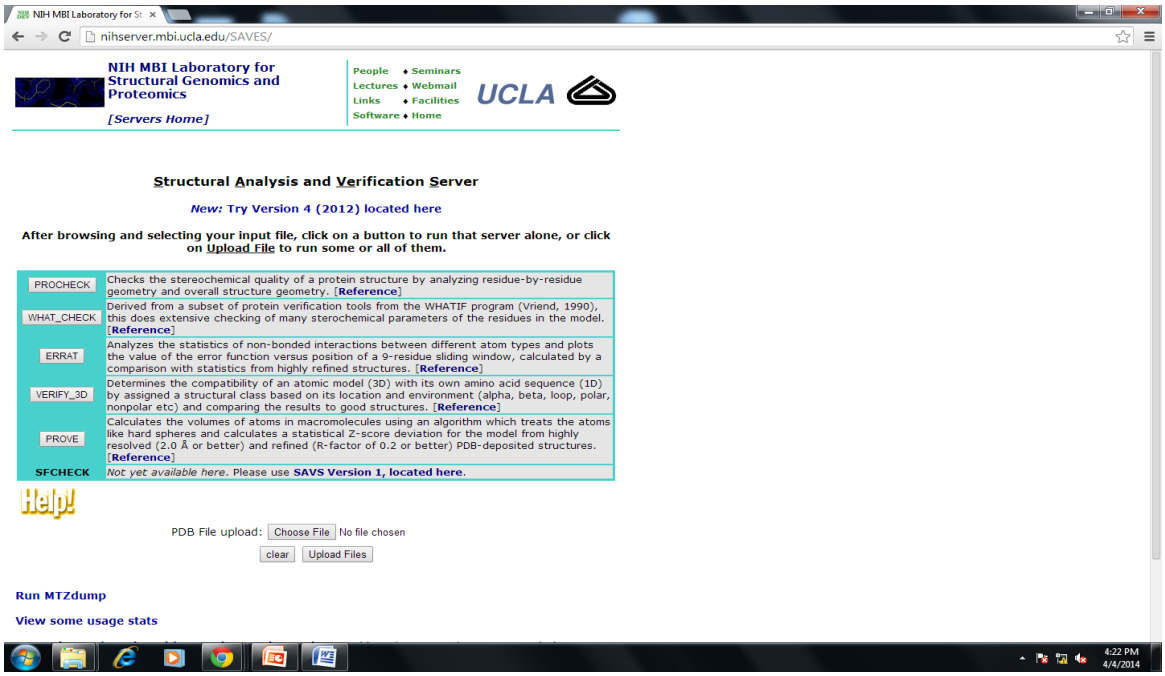


Fig. 7. Home page of externally by SAVES SERVER

B. METHODS

I. Collection of chromate ion transporter proteins from NCBI protein database

STEPS

- Opened the Home page of NCBI.
- Clicked on “All databases” and selected the “Protein” from the dropdown menu.
- Then wrote “Chr A” in the right search box and clicked on “Search”.
- There were about 9006 chromate ion transporter proteins of different organisms, but collected chromate ion transporter proteins of only individual organism.
- Double clicked on one organism from the top organisms.

- Then choosed single organism from a list of similar organism.
- Then clicked on the FASTA file of that organism and copied the protein sequences.
- Then saved this in the text file.
- Then these proteins are divided into two groups that is LCHR and SCHR based on amino acid length of the protein sequences (in bacterial SCHR-123-234, in bacterial LCHR-345-495 and in fungal LCHR-502-584 amino acids).
- In this way LCHR and SCHR protein sequences are separated into two new text files.

II. Phylogenetic divergence analysis among chromate ion transporter proteins using MEGA

STEPS

- The LCHR and SCHR protein sequences are edited.
- The protein LCHR and SCHR protein containing text files, we are converted into FASTA file.
- Then opened Home page of MEGA 5.1.
- Then opened file/session
    - Selected the long chain FASTA file.
    - Opened.
    - Then selected the align bottom.
- Edit select all.
- Alignment
    - Alignment by muscle.
    - Compute.

- Date export alignment
    - MEGA format.
    - Save with the same name and the file saved as .meg.
- Double click open the MEGA file
    - Select C for observing the conserved file.
    - Select V for observing the variable file.
- Phylogeny construct/text maximum likelihood tree method
    - Then select yes.
    - Compute.
- Then the tree is saved for observing divergence among the groups of proteins.
- The similar process is followed for constructing the phylogenetic tree for SCHR and LCHR protein.

III. For finding out secondary structure (Alpha helix, Extended strand and Random coil in %) of the proteins using GOR IV

STEPS

- Opened the Home page of GOR IV.
- Opened the text file of LCHR protein.
- Selected the protein sequences of one LCHR text file.
- Then copied the sequences.
- Then pasted them in the box of GOR IV.
- Then clicked the submit bottom.
- Then opened the result of the protein sequences.
- Then wrote the Alpha helix, extended strand and random coiling in %.

- Similar process is followed for finding out the result of secondary structure of all LCHR and SCHR protein sequences.

IV. For finding out the functional domain of the proteins using InterProScan 4

STEPS

- Opened the Home page of InterProScan 4.
- Double clicked on InterPro.
- Opened the text file of LCHR protein.
- Selected the protein sequences of one LCHR text file.
- Then copied the sequences.
- Then pasted them in the box of InterProScan 4.
- Then clicked the search bottom.
- Then wrote the functional domain.
- Similar process is followed for finding out functional domain of all LCHR and SCHR protein sequences.

V. Homology modeling for 3D structure prediction

(a) BLASTP

For finding out homologous

STEPS

- Opened the BLAST tool box.
- Then select the protein blast tool.
- Pasted the sequence in search box.
- Selected the protein data bank against the sequence.
- Then pressed the option BLAST.

(b) MODELLER

For prediction of 3D structure

STEPS

- Go to basic example folder which was presented within the MODELLER folder.
- Opened the basic example Ali file changed the protein name and sequences of previous protein and added the protein Id and sequences of protein which was taken for modeling.
- Opened the BLASTP web page. Sequences of the protein which was taken was pasted in the query box and pressed the submit button.
- The homologous protein of the query protein was shown in the result page. Then proteins were chosen. Then opened the pdb web page. The template proteins are entered in the query page of query box. Then downloaded the protein structure in the pdb text format and saved them in the basic example folder.
- Compare.py file was opened. The template proteins Id and chain are edited. Then comparre.py file was run in modeler software by gave the command of

  >mod9.12 compare.py. Then pressed the enter button.
- Log file was generated. By opening the file, template pdb file with their resolution and their resolution were shown in a dendrogram. The lower value resolution will be taken and crossed check there identity with the target. The appropriate template was selected.
- The target was aligned and with the selected template. Then edited where ever necessary. Then run the command >mod9.12 align2d.py.Two files were generated Ali file and pap file.

- Model-single.py file was opened and the required settings were changed. Then run the >mod9.12 model-single.py in modeler software. The log file of model-single was opened and choosed the pdb file of lower dope score.

(c) Externally by SAVE SERVER

Evaluating the predicted model

STEPS

- Save server web page was opened.
- Then choosed the selected pdb file and Procheck option was also choosed.
- By clicked on the Ramachandran plot option Ramachandran plot was shown.
- Verify 3D option was clicked and the result page was shown.

4. RESULT AND DISCUSSION

I. Chromate ion transporter protein

Three hundred and fifty-eight (358) non-redundant protein sequences were identified as members of the CHR family. Of these, three hundred and fifty-seven (357) proteins belong to bacteria and one (1) to fungi. Protein sequence members of the CHR family were not found in either animals or plants.

Two main sizes of CHR proteins were found bacterial monodomain proteins with a sequence length of 123-234 amino acids, called bacterial short-chain CHR (bacterial SCHR) and bacterial bidomain proteins with a sequence length of 345-495 amino acids, called bacterial long- chain CHR (bacterial LCHR).

Only bidomain LCHR protein sequences were identified in fungi (fungal LCHR).These showed longer length 502-584 amino acids because they possess an interdomain sequence longer than that of bacterial LCHR sequences.

In this way, we were identified two hundred and thirty-seven (237) LCHR proteins and one hundred and twenty-one (121) SCHR proteins. Of these 237 LCHR proteins, there were 236 LCHR protein belong to bacteria and 1 from fungi that was *Coprinopsis cinerea*having amino acid length 546. But one organism that was *Thiobacillus denitrificans* having amino acid length 507 that was found also bacteria.

II. Divergence study of chromate ion transporter protein

Long chain transporter (LCHR) protein

All the LCHR proteins of bacteria and fungi (237) were aligned by muscle. There were 25 conserved sites and 638 variable sites in LCHR proteins.

Phylogenetic divergence study was conducted of 237 LCHR proteins. It was observed that the total LCHR proteins diversified into four groups i.e group-1, group-2, group-3 and group-4.Group-1 (blue), group-2 (red), group-3 (green) and group-4 (pink) contained 123, 4, 32 and 78 chromate ion transporter proteins respectively. Most of the LCHR proteins contained in group-1 and less number of the LCHR proteins contained in group-2. One LCHR protein of fungi that was *Coprinopsis cinerea* having amino acid length 546 was found in group-3.

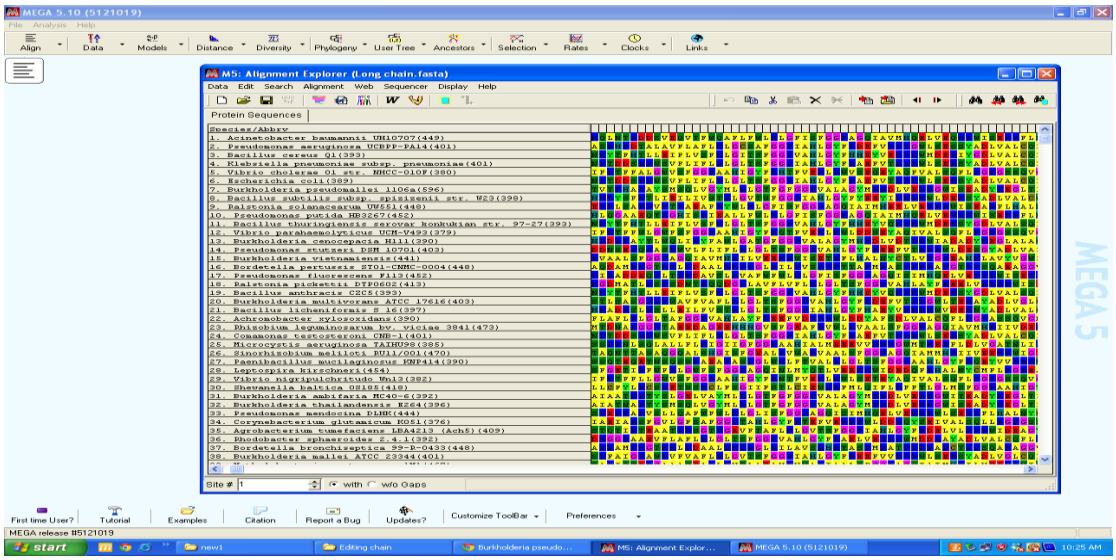


Fig. 8. Sequences of LCHR protein

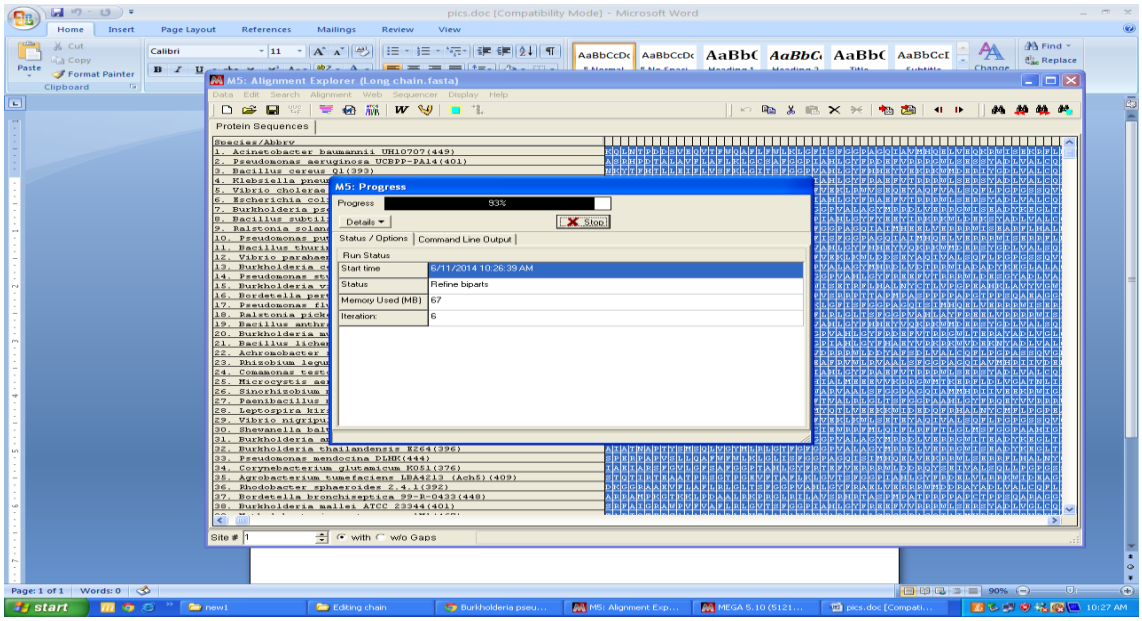


Fig. 9. Alignment by muscle of LCHR protein

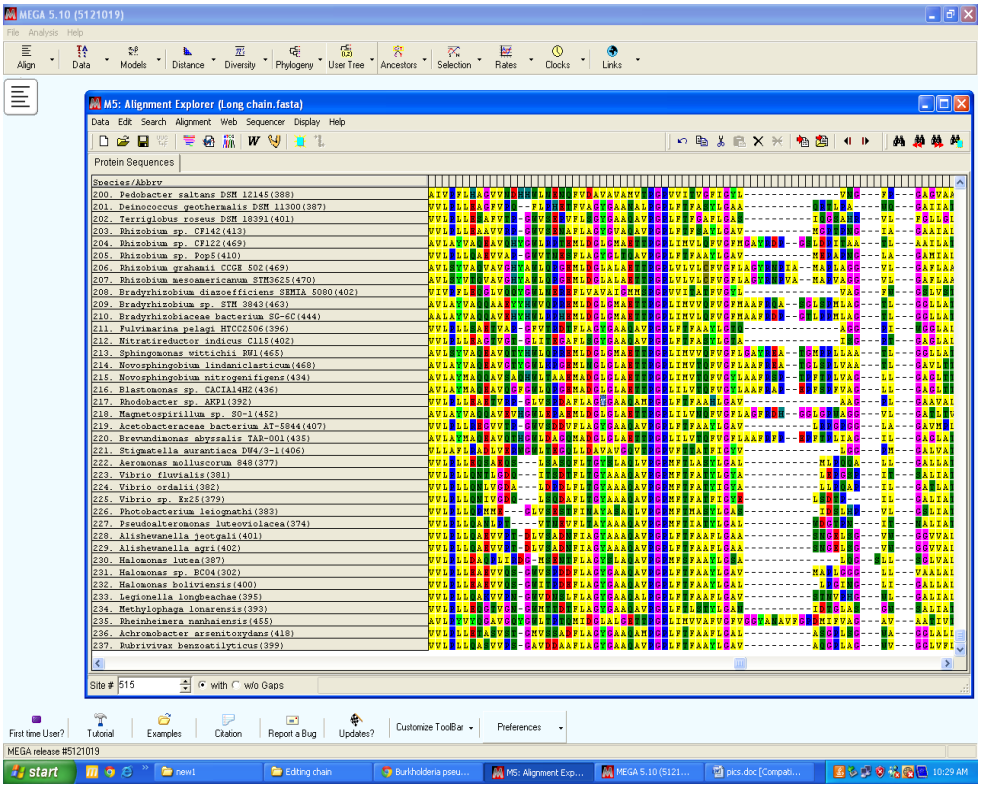


Fig. 10. Alignment view of LCHR protein

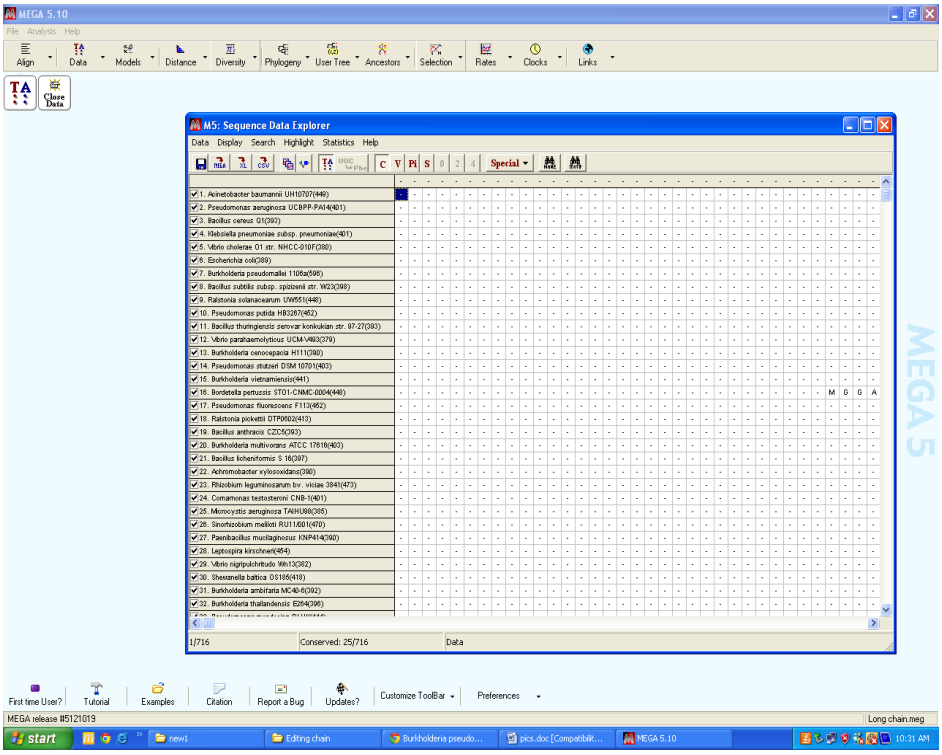


Fig. 11. Conserved site of LCHR protein

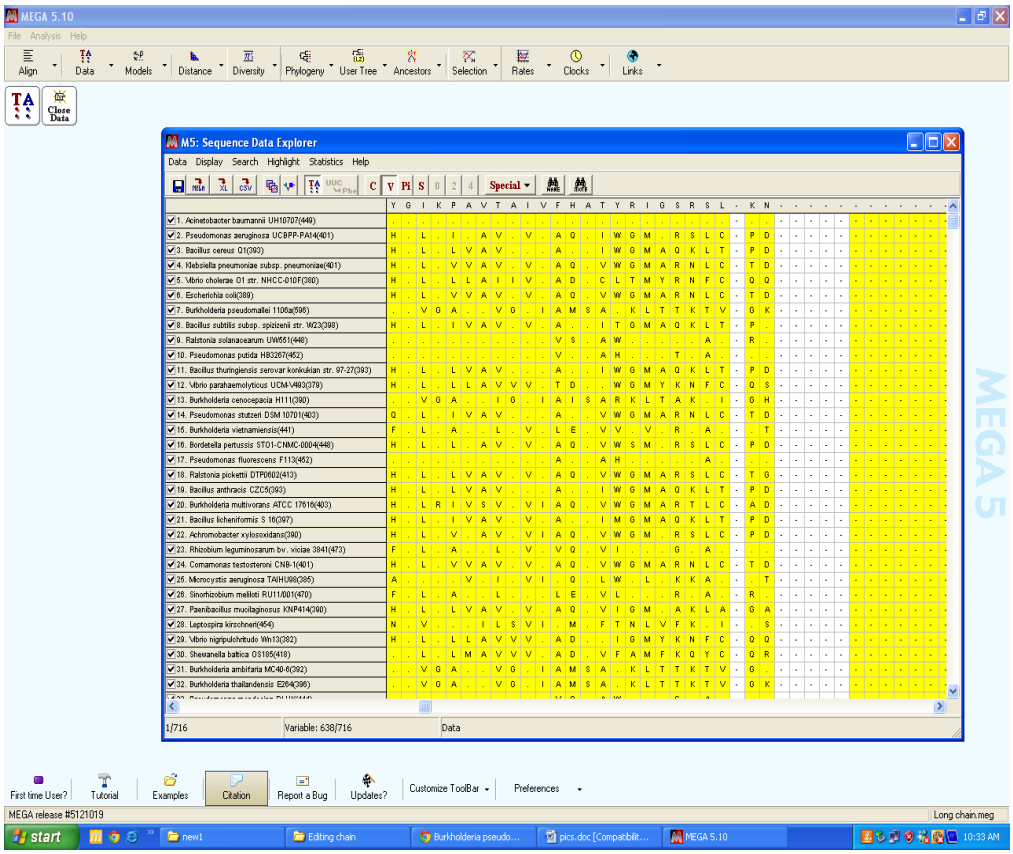


Fig. 12. Variable site of LCHR protein

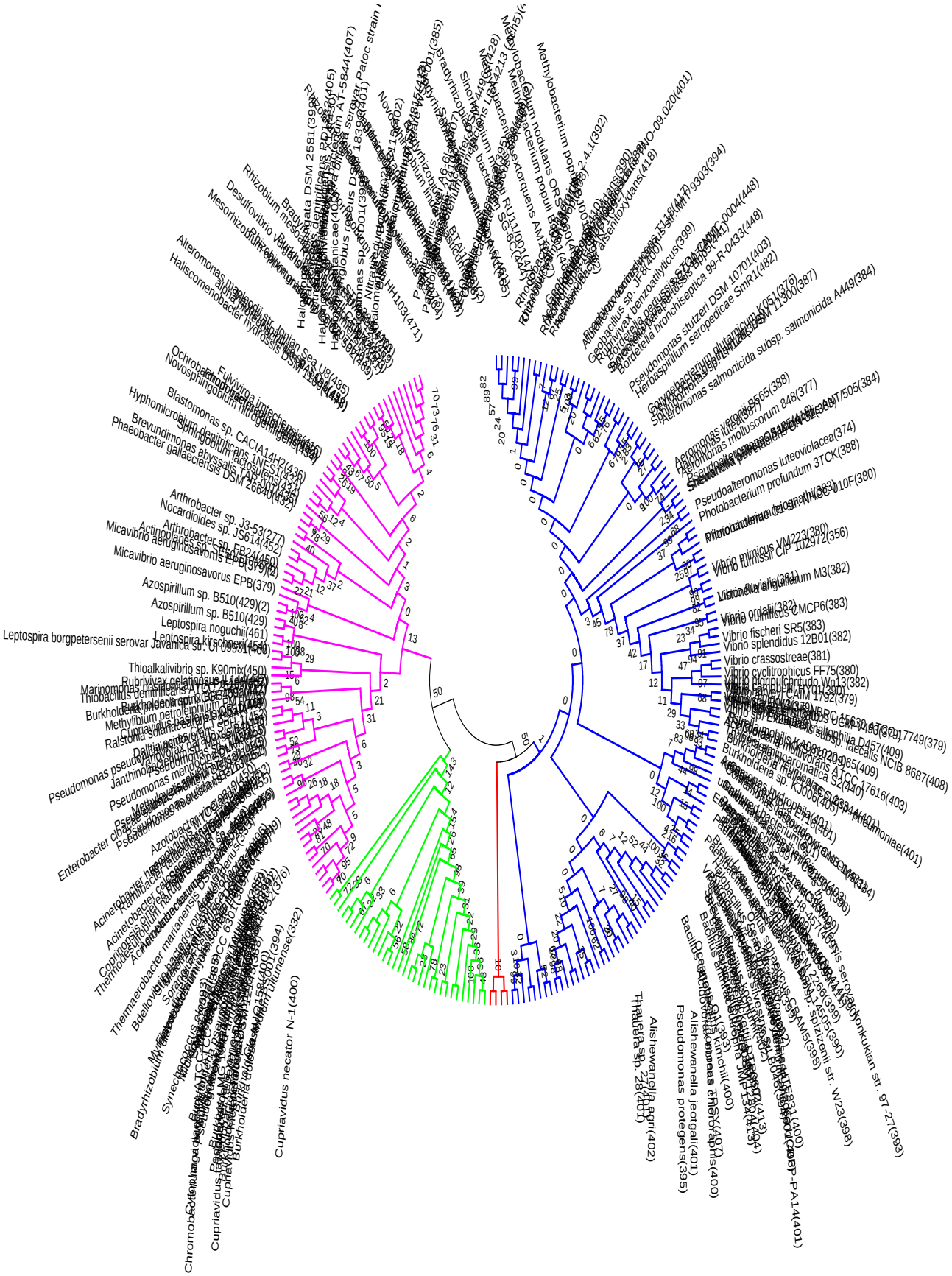


Fig. 13. Phylogenetic analysis of LCHR protein

Short chain transporter (SCHR) protein

All the SCHR proteins of bacteria (121) were aligned by muscle. There were no conserved sites but the variable sites were 293 in SCHR proteins.

Phylogenetic divergence was conducted of 121 SCHR proteins. It was observed that the total SCHR proteins diversified into six groups that were group-1, group-2, group-3, group-4, group-5 and group-6.Group-1 (red), group-2 (deep blue), group-3 (green) and group-4 (coffee), group-5 (light blue) and group-6 (pink) contained 53, 16, 45, 2, 1 and 4 chromate ion transporter proteins respectively. Most of the SCHR proteins contained in group-1 and less number of the SCHR proteins contained in group-5.

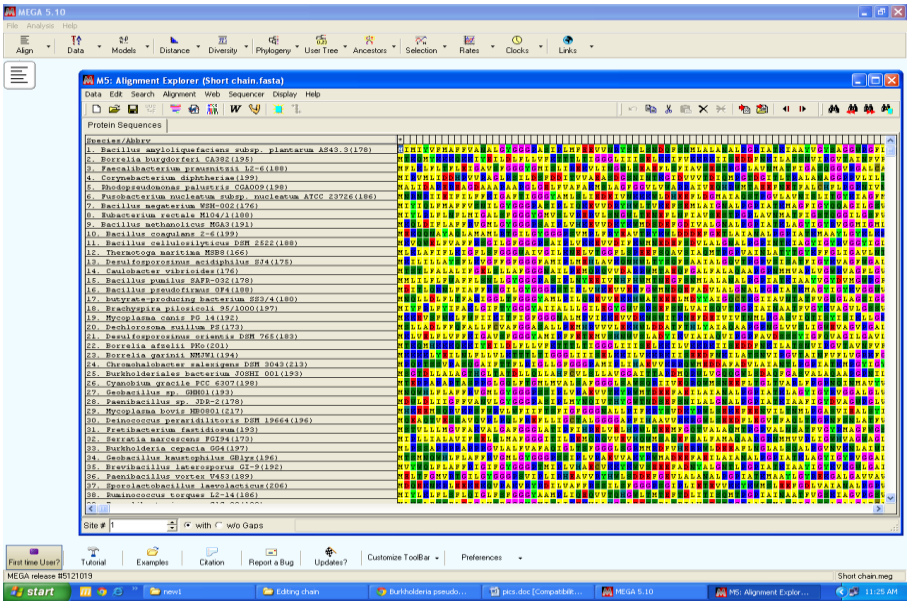


Fig. 14. Sequences of SCHR protein

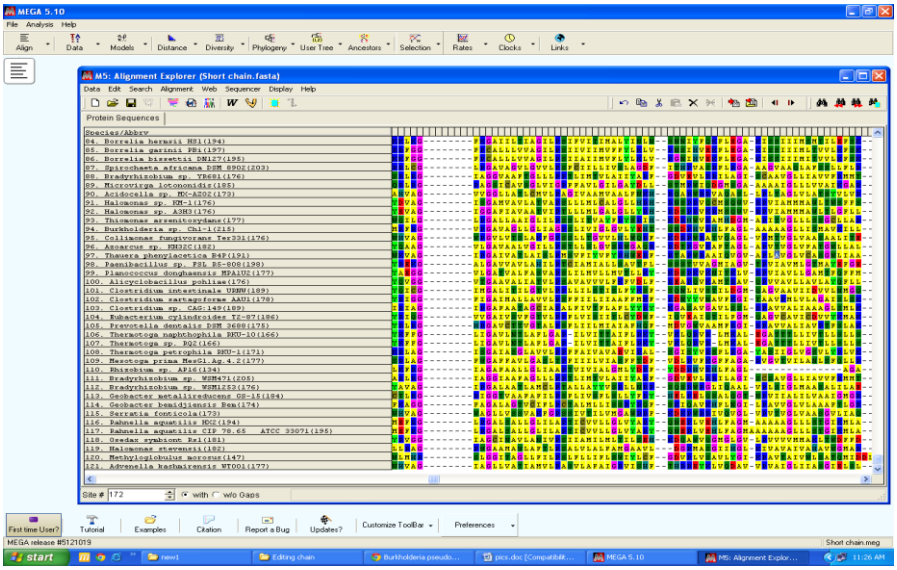


Fig. 15. Alignment view of SCHR protein

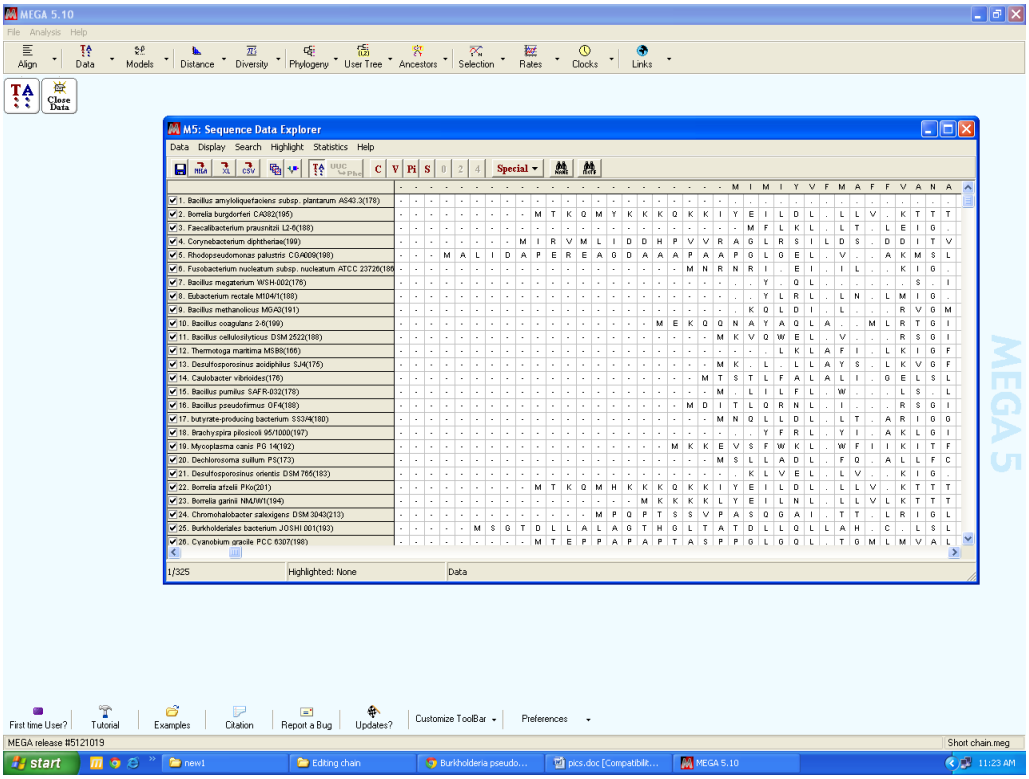


Fig. 16. Conserved sites of SCHR protein

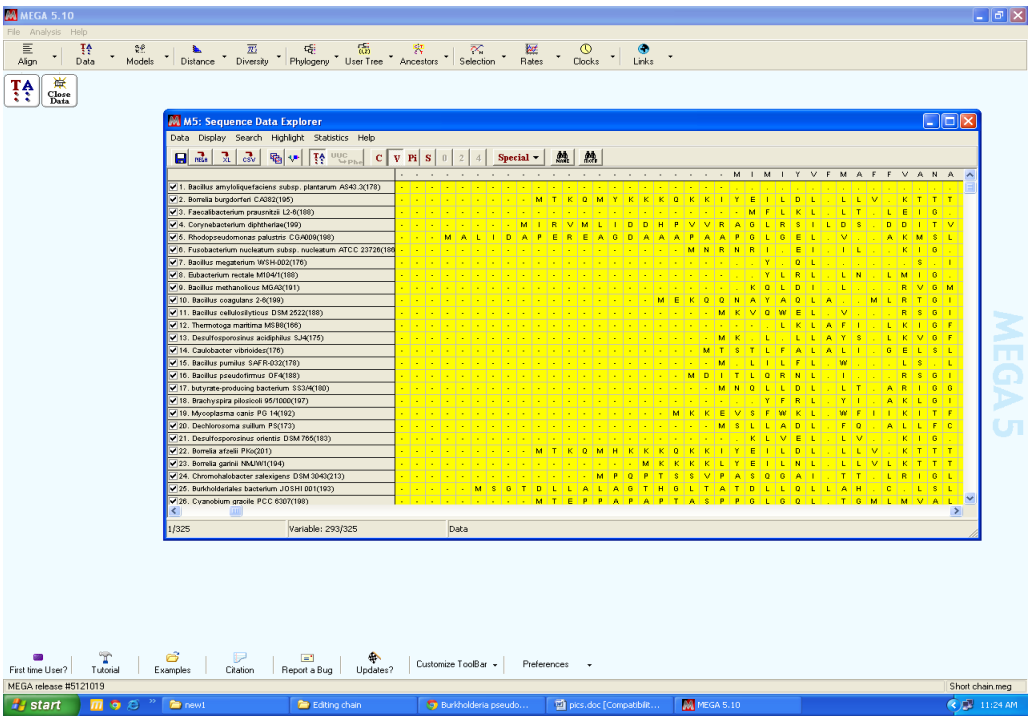


Fig. 17. Variable sites of SCHR protein

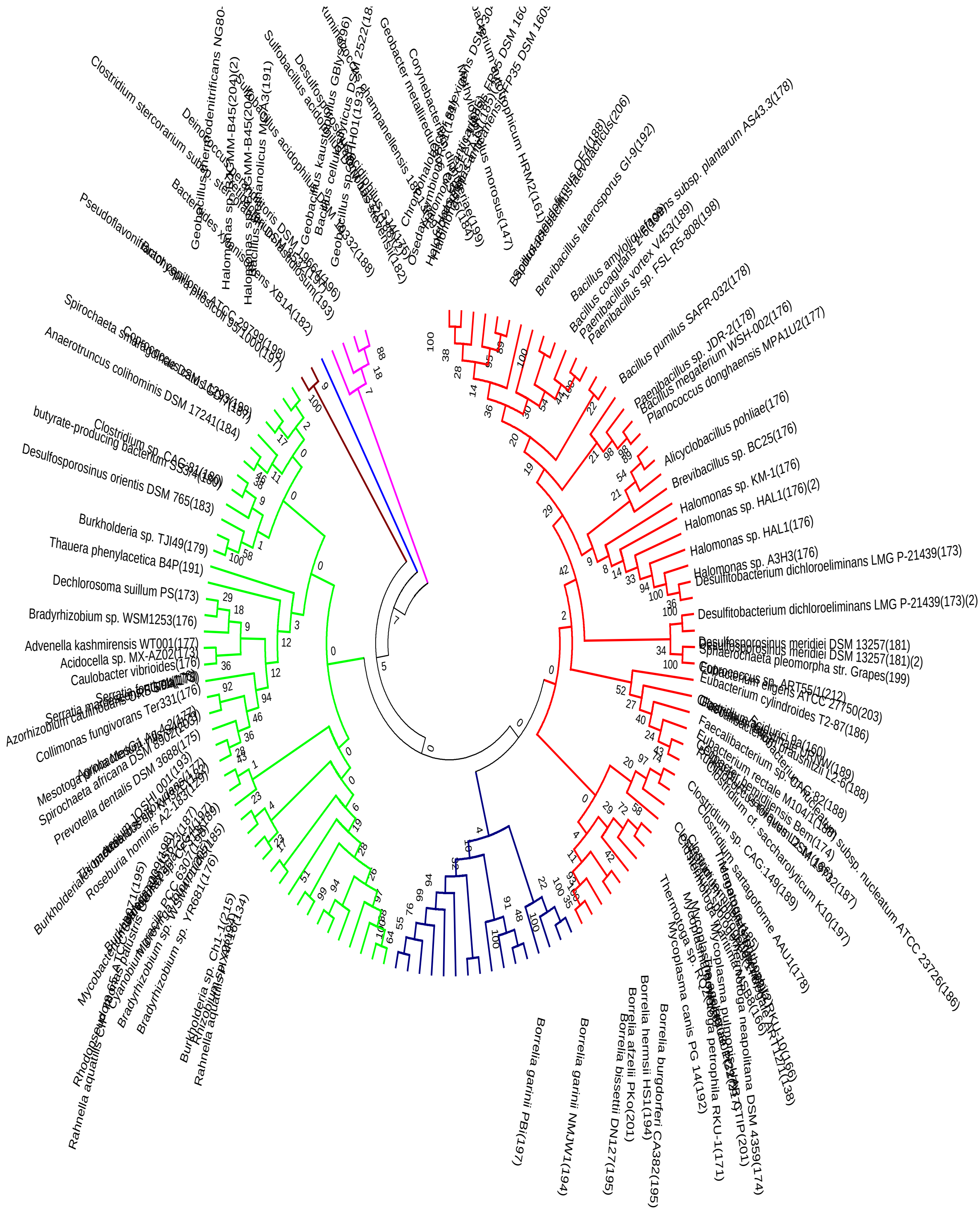


Fig. 18. Phylogenetic analysis of SCHR protein

III. Prediction of secondary structure of chromate ion transporter protein

Secondary structure (Alpha helix, extended strand and random coil) were predicted of the both LCHR and SCHR proteins in %. The % of alpha helix and random coil of chromate ion transporter proteins were more than the extended strand in both LCHR & SCHR proteins.

Long chain transporter (LCHR) protein

Table-1. Prediction of secondary structure of LCHR protein

SEQUENCE ID	ORGANISM NAME	LENGTH OF AMINO ACID	SECONDARY STRUTURE (Alpha helix, Extended strand, Random coil) (in %)
ETR55227.1	*Acinetobacter baumannii*	449	(42.09,15.81,42.09)
ABJ13559.1	*Pseudomonas aeruginosa*	401	(55.11,6.23,38.65)
ACM15429.1	*Bacillus cereus*	393	(36.39,17.81,45.80)
AHY00005.1	*Klebsiella pneumoniae*	401	(41.90,16.21,41.90)
EMQ66925.1	*Vibrio cholerae*	380	(46.32,16.32,37.37)
WP_012372815.1	*Escherichia coli*	389	(43.19,15.68,41.13)
ABN95479.1	*Burkholderia pseudomallei*	396	(45.71,16.16,38.13)
ADM37870.1	*Bacillus subtilis*	398	(46.48,14.82,38.69)
EAP72482.1	*Ralstonia solanacearum*	448	(44.20,18.30,37.50)
AGA73085.1	*Pseudomonas putida*	452	(43.14,15.71,41.15)
YP_039193.1	*Bacillus thuringiensis*	393	(36.90,16.54,46.56)
AHI98792.1	*Vibrio parahaemolyticus*	379	(33.77,16.09,50.13)
CDN64820.1	*Burkholderia cenocepacia*	390	(40.00,22.31,37.69)
AFN78404.1	*Pseudomonas stutzeri*	403	(50.87,7.44,41.69)
WP_021161752.1	*Burkholderia vietnamiensis*	441	(57.82,12.70,29.48)
ETI07845.1	*Bordetella pertussis*	448	(50.22,8.04,41.74)
AEV63938.1	*Pseudomonas fluorescens*	452	(45.58,14.60,39.82)
YP_008600505.1	*Ralstonia pickettii*	413	(52.54,9.69,37.77)
GAF00771.1	*Bacillus anthracis*	393	(36.90,16.79,46.31)
YP_001946933.1	*Burkholderia multivorans*	403	(54.09,10.92,34.99)
EWH21927.1	*Bacillus licheniformis*	397	(47.10,11.84,41.06)
WP_024068811.1	*Achromobacter xylosoxidans*	390	(50.51,11.03,38.46)
CAK08671.1	*Rhizobium leguminosarum*	473	(50.74,13.53,35.73)
ABM06203.1	*Comamonas testosteroni*	401	(41.90,16.21,41.90)
ELP55763.1	*Microcystis aeruginosa*	385	(47.53,12.99,39.48)
CDH84656.1	*Sinorhizobium meliloti*	470	(45.11,20.85,34.04)

AEI41168.1	*Paenibacillus mucilaginosus*	390	(48.72,8.72,42.56)
WP_004769390.1	*Leptospira kirschneri*	454	(35.24,24.01,40.75)
CCO52615.1	*Vibrio nigripulchritudo*	382	(45.81,13.61,40.58)
ABS09622.1	*Shewanella baltica*	418	(48.56,13.40,38.04)
ACB65671.1	*Burkholderia ambifaria*	392	(48.21,15.31,36.48)
ABC36222.1	*Burkholderia thailandensis*	396	(44.19,18.43,37.37)
EJO94496.1	*Pseudomonas mendocina*	444	(45.05,14.86,40.09)
YP_007561622.1	*Corynebacterium glutamicum*	376	(35.64,13.03,51.33)
AHK02891.1	*Agrobacterium tumefaciens*	409	(48.41,14.18,37.41)
YP_352338.1	*Rhodobacter sphaeroides*	392	(58.67,6.38,34.95)
KCV66631.1	*Bordetella bronchiseptica*	448	(51.56,7.59,40.85)
YP_103831.1	*Burkholderia mallei*	401	(46.88,17.71,35.41)
ACS39230.1	*Methylobacterium extorquens*	467	(42.83,16.70,40.47)
YP_006966045.1	*Aeromonas hydrophila*	401	(43.39,14.96,41.65)
AGU47511.1	*Variovorax paradoxus*	448	(40.62,19.64,39.73)
CDG75483.1	*Acinetobacter nosocomialis*	449	(35.63,20.94,43.43)
EEY78117.1	*Acinetobacter calcoaceticus*	448	(38.84,18.53,42.63)
EPR88996.1	*Acinetobacter haemolyticus*	450	(40.00,20.40,39.56)
AFJ86665.1	*Burkholderia sp.*	405	(59.01,8.64,32.35)
WP_017399575.1	*Acinetobacter pittii*	449	(42.32,16.70,40.98)
EAP94742.1	*Vibrio splendidus*	382	(37.17,19.90,42.93)
NP_694198.1	*Oceanobacillus iheyensis*	400	(45.75,14.50,39.75)
EPG55962.1	*Leptospira borgpetersenii*	488	(35.45,21.52,43.03)
WP_016560147.1	*Leptospira noguchii*	461	(24.73,29.72,45.55)
ACI91597.1	*Oligotropha carboxidovorans*	384	(48.18,11.98,39.84)
EEY44780.1	*Vibrio mimicus*	380	(45.26,16.58,38.16)
ADC70598.1	*Thioalkalivibrio sp.*	450	(31.11,26.89,42.00)
YP_161712.1	*Cupriavidus metallidurans*	374	(41.18,21.12,37.70)
ABR90147.1	*Janthinobacterium sp.*	477	(46.12,11.53,42.35)
AGK04154.1	*Meiothermus ruber*	349	(54.73,10.03,35.24)
ABC91487.1	*Rhizobium etli*	473	(48.41,14.16,37.42)
EXL01874.1	*Ochrobactrum anthropi*	408	(45.83,14.95,39.22)
YP_006184233.1	*Stenotrophomonas maltophilia*	409	(45.72,10.76,43.52)
NP_442845.1	*Synechocystis sp*	399	(36.59,21.30,42.11)
YP_002774542.1	*Brevibacillus brevis*	411	(43.31,15.57,41.12)
EPH11961.1	*Myroides odoratimimus*	379	(40.63,18.47,40.90)
YP_005191479.1	*Sinorhizobium fredii*	471	(44.80,16.35,38.85)

AEH84626.1	*Mesorhizobium opportunistum*	463	(49.24,17.06,33.69)
AFM60000.1	*Enterobacter cloacae*	455	(39.78,19.56,40.66)
AAO10171.1	*Vibrio vulnificus*	383	(45.95,15.40,38.64)
AEB48657.1	*Aeromonas veronii*	388	(48.45,13.66,37.89)
AGP93754.1	*Alteromonas macleodii*	485	(47.84,15.67,36.49)
EPY13218.1	*Paenibacillus alvei*	407	(43.49,19.66,36.86)
WP_005585600.1	*Clostridium ultunense*	332	(41.27,18.98,39.76)
EDL68584.1	*Vibrio campbellii*	390	(39.23,16.92,43.85)
YP_009133.1	*Desulfovibrio vulgaris*	445	(43.60,17.53,38.88)
AEI80435.1	*Cupriavidus necator*	400	(36.75,21.25,42.00)
WP_016704768.1	*Pseudomonas chlororaphis*	400	(48.50,11.25,40.25)
GAD72596.1	*Vibrio alginolyticus*	379	(38.52,15.30,46.17)
AFU46855.1	*Acidovorax sp.*	409	(45.72,10.76,43.52)
YP_004389086.1	*Alicycliphilus denitrificans*	406	(64.53,2.46,33.00)
EJO33504.1	*Achromobacter piechaudii*	388	(47.94,9.28,42.78)
AET87449.1	*uncultured bacterium*	401	(41.90,16.21,41.90)
ELK48494.1	*Halobacillus sp.*	399	(46.37,12.53,41.10)
ACL60922.1	*Methylobacterium nodulans*	469	(49.68,12.15,38.17)
AGK58336.1	*Hyphomicrobium denitrificans*	434	(45.16,17.28,37.56)
CDM42712.1	*Pseudomonas pseudoalcaligenes*	444	(49.55,12.16,38.29)
EHN70797.1	*Vibrio fischeri*	383	(37.34,18.02,44.65)
YP_003896680.1	*Halomonas elongata*	399	(44.36,13.28,42.36)
YP_001630202.1	*Bordetella petrii*	404	(53.71,5.94,40.35)
YP_002008176.1	*Cupriavidus taiwanensis*	400	(48.25,16.00,35.75)
YP_171835.1	*Synechococcus elongatus*	383	(47.52,12.01,40.47)
ABM77873.1	*Prochlorococcus marinus*	394	(54.57,11.93,33.50)
YP_006182093.1	*Halobacillus halophilus*	399	(41.60,14.04,44.36)
AGT31995.1	*Geobacillus sp.*	400	(58.25,9.00,32.75)
ACX66242.1	*Paenibacillus sp.*	409	(32.52,21.03,46.45)
ABL72111.1	*Paracoccus denitrificans*	405	(53.58,6.67,39.75)
AEX22965.1	*Vibrio sp.*	379	(43.80,13.98,42.22)
ABE37066.1	*Burkholderia xenovorans*	398	(40.45,20.60,38.94)
EON15238.1	*Pandoraea sp.*	410	(46.34,12.93,40.73)
ADJ62751.1	*Herbaspirillum seropedicae*	482	(50.83,8.71,40.46)
ENO82674.1	*Thauera sp.*	401	(53.87,8.98,37.16)
CCQ60486.1	*Crocosphaera watsonii*	389	(45.50,18.51,35.99)
YP_006463894.1	*Solibacillus silvestris*	394	(45.18,15.23,39.59)
YP_004437607.1	*Thermodesulfobium narugense*	368	(34.78,26.63,38.59)

EFH84111.1	*Ktedonobacter racemifer*	428	(46.26,11.68,42.06)
AHD09798.1	*Phaeobacter gallaeciensis*	432	(58.10,8.80,33.10)
ACO80512.1	*Azotobacter vinelandii*	453	(48.34,13.69,37.97)
ABM95481.1	*Methylibium petroleiphilum*	445	(36.18,23.37,40.45)
ABX35705.1	*Delftia acidovorans*	452	(43.58,15.93,40.49)
AEG94919.1	*Ramlibacter tataouinensis*	457	(47.70,12.25,40.04)
AAQ19822.1	*Alcaligenes faecalis*	408	(48.53,8.82,42.65)
EAY70264.1	*Burkholderia dolosa*	400	(38.25,24.00,37.75)
CAK28037.1	*Synechococcus sp*	401	(51.12,9.73,39.15)
YP_004101092.1	*Thermaerobacter marianensis*	450	(67.33,6.22,26.44)
AEV99229.1	*Niastella koreensis*	392	(38.78,22.19,39.03)
ABE62743.1	*Nitrobacter hamburgensis*	430	(46.74,9.53,43.72)
ABQ36693.1	*Bradyrhizobium sp*	463	(50.54,13.17,36.29)
AEQ52600.1	*Pelagibacterium halotolerans*	464	(44.18,20.69,35.13)
ACB82306.1	*Methylobacterium populi*	469	(40.51,18.98,40.51)
AGI71815.1	*Octadecabacter arcticus*	400	(40.00,12.50,47.50)
YP_003447441.1	*Azospirillum sp*	429	(44.06,16.32,39.63)
AGH97925.1	*Micavibrio aeruginosavorus*	379	(40.37,14.78,44.85)
AHC87436.1	*Pseudomonas monteilii*	453	(44.37,16.34,39.29)
WP_011061436.1	*Pseudomonas protegens*	395	(53.16,7.59,39.24)
CBL87892.1	*Pseudomonas sp.*	455	(38.68,24.84,36.48)
AFN31616.1	*Klebsiella oxytoca*	401	(41.90,16.21,41.90)
AGU57013.1	*Listonella anguillarum*	382	(42.15,16.23,41.62)
ERM60026.1	*Vibrio cyclitrophicus*	380	(35.00,19.74,45.26)
ABO91542.1	*Aeromonas salmonicida*	384	(54.95,10.94,34.11)
EGI73761.1	*Pseudoalteromonas haloplanktis*	384	(46.88,15.62,37.50)
YP_004481554.1	*Marinomonas posidonica*	452	(41.15,20.58,38.27)
ABQ36693.1	*Bradyrhizobium sp*	463	(50.54,13.17,36.29)
AEQ52600.1	*Pelagibacterium halotolerans*	464	(44.18,20.69,35.13)
ACB82306.1	*Methylobacterium populi*	469	(40.51,18.98,40.51)
AGI71815.1	*Octadecabacter arcticus*	400	(40.00,12.50,47.50)
YP_003447441.1	*Azospirillum sp*	429	(44.06,16.32,39.63)
AGH97925.1	*Micavibrio aeruginosavorus*	379	(40.37,14.78,44.85)
YP_001619890.1	*Sorangium cellulosum*	400	(52.75,6.50,40.75)
AFY01497.1	*Bdellovibrio bacteriovorus*	335	(55.52,11.34,33.13)
YP_005437080.1	*Rubrivivax gelatinosus*	447	(51.68,9.17,39.15)
YP_006897362.1	*Bordetella parapertussis*	441	(51.02,8.16,40.82)
AAZ64575.1	*Ralstonia eutropha*	413	(46.73,13.80,39.47)

EHP40784.1	*Cupriavidus basilensis*	464	(35.56,24.78,39.66)
ACC74843.1	*Burkholderia phymatum*	412	(50.24,11.65,38.11)
EIM94433.1	*Burkholderia terrae*	394	(42.39,17.77,39.85)
AFT85530.1	*Burkholderia phenoliruptrix*	442	(44.80,18.55,36.65)
YP_004227940.1	*Burkholderia sp*	442	(45.25,18.55,36.20)
ABE46762.1	*Polaromonas sp*	456	(43.42,16.45,40.13)
ACM33301.1	*Acidovorax ebreus*	407	(63.64,2.46,33.91)
ABD70536.1	*Albidiferax ferrireducens*	417	(40.77,8.39,50.84)
AAQ58596.1	*Chromobacterium violaceum*	393	(34.10,23.92,41.98)
ENO88273.1	*Thauera aminoaromatica*	440	(54.09,8.64,37.27)
ENO93942.1	*Thauera sp*	401	(51.37,8.98,39.65)
WP_020166052.1	*Methyloversatilis universalis*	445	(46.07,13.71,40.22)
WP_021478256.1	*Pseudogulbenkiania ferrooxidans*	386	(51.04,12.95,36.01)
YP_006963560.1	*Arthrobacter sp.*	277	(45.13,14.08,40.79)
ABL81046.1	*Nocardioides sp*	452	(43.14,15.04,41.81)
EWH33178.1	*Lysinibacillus sphaericus*	398	(44.22,19.60,36.18)
WP_007503781.1	*Caldalkalibacillus thermarum*	403	(51.36,13.90,34.74)
EQB34717.1	*Virgibacillus sp*	398	(45.98,12.56,41.46)
WP_009590732.1	*Paenibacillus sp.*	409	(34.96,19.80,45.23)
WP_007128576.1	*Paenibacillus lactis*	409	(46.21,15.65,38.14)
GAE04947.1	*Paenibacillus sp.*	407	(36.61,22.60,40.79)
WP_023559015.1	*Brevibacillus panacihumi*	402	(46.02,14.93,39.05)
EIM08287.1	*Planococcus antarcticus*	390	(34.36,21.28,44.36)
WP_008301397.1	*Bhargavaea cecembensis*	397	(37.28,16.62,46.10)
ADG06946.1	*Kyrpidia tusciae*	252	(35.71,20.63,43.65)
WP_009577949.1	*Fulvivirga imtechensis*	416	(35.82,22.36,41.83)
ACU05572.1	*Pedobacter heparinus*	392	(34.44,29.59,35.97)
YP_004448553.1	*Haliscomenobacter hydrossis*	447	(40.49,20.58,38.93)
XP_002911996.1	*Coprinopsis cinerea (Fungi)*	546	(32.42,22.53,45.05)
YP_005609514.1	*Bradyrhizobium japonicum*	461	(47.51,13.45,39.05)
ELT47282.1	*Ochrobactrum intermedium*	432	(55.09,11.34,33.56)
AAQ18188.1	*Rhodobacter capsulatus*	449	(62.58,8.02,29.40)
YP_005794667.1	*Ketogulonicigenium vulgare*	385	(53.25,10.39,36.36)
EJU14087.1	*Sphingomonas sp*	399	(56.39,8.02,35.59)
YP_003545570.1	*Sphingobium japonicum*	463	(49.24,15.33,35.42)
WP_021228313.1	*Sphingobium lactosutens*	434	(54.38,13.13,32.49)
AFK56631.1	*Tistrella mobilis*	409	(45.72,10.76,43.52)
EDP65385.1	*alpha proteobacterium*	448	(45.54,19.87,34.60)
EMR36372.1	*Vibrio harveyi*	379	(37.99,16.62,45.38)

WP_020334329.1	*Vibrio natriegens*	380	(41.58,15.53,42.89)
WP_017070361.1	*Vibrio crassostreae*	381	(37.27,18.37,44.36)
EEX38787.1	*Vibrio furnissii*	356	(57.02,7.30,35.67)
EAS42875.1	*Photobacterium profundum*	388	(41.24,13.92,44.85)
ABP76766.1	*Shewanella putrefaciens*	383	(48.83,15.67,35.51)
ABK50579.1	*Shewanella sp.*	455	(42.86,18.90,38.24)
EGP21262.1	*Halomonas sp.*	399	(47.12,12.53,40.35)
WP_009287634.1	*Halomonas titanicae*	400	(47.50,14.75,37.75)
AAZ98609.1	*Thiobacillus denitrificans*	507	(37.67,20.32,42.01)
AEV85163.1	*Actinoplanes sp.*	474	42.83,14.56,42.62)
ABK05889.1	*Arthrobacter sp.*	450	(42.89,20.00,37.11)
EID44387.1	*Geobacillus thermoglucosidans*	401	(46.38,15.96,37.66)
WP_017798519.1	*Oceanobacillus kimchii*	400	(48.00,14.00,38.00)
WP_018924065.1	*Salsuginibacillus kocurii*	402	(43.03,17.91,39.05)
ETT66405.1	*Paenibacillus sp.*	409	(35.45,19.56,44.99)
EMT53387.1	*Brevibacillus borstelensis*	405	(48.64,14.57,36.79)
WP_019122516.1	*Brevibacillus massiliensis*	401	(45.14,14.96,39.90)
ABZ92904.1	*Leptospira biflexa*	390	(38.97,18.46,42.56)
ABG59418.1	*Cytophaga hutchinsonii*	385	(27.53,30.65,41.82)
ABQ07421.1	*Flavobacterium johnsoniae*	376	(46.01,15.16,38.83)
EJG02233.1	*Flavobacterium sp.*	376	(37.23,17.55,45.21)
ADY52544.1	*Pedobacter saltans*	388	(30.15,29.64,40.21)
ABF44423.1	*Deinococcus geothermalis*	387	(48.58,10.85,40.57)
AFL88631.1	*Terriglobus roseus*	401	(42.39,11.72,45.89)
WP_007816752.1	*Rhizobium sp*	413	(57.63,6.78,35.59)
WP_007802676.1	*Rhizobium sp*	469	(46.91,16.63,36.46)
EJZ18077.1	*Rhizobium sp.*	410	(44.88,13.41,41.71)
EPE99336.1	*Rhizobium grahamii*	469	(51.39,15.78,32.84)
CCM74575.1	*Rhizobium mesoamericanum*	470	(51.91,12.77,35.32)
KAK31649.1	*Bradyrhizobium diazoefficiens*	402	(45.02,18.41,36.57)
WP_008974529.1	*Bradyrhizobium sp*	463	(49.46,14.47,36.07)
EGP06243.1	*Bradyrhizobiaceae bacterium*	444	(55.63,12.39,31.98)
EAU42843.1	*Fulvimarina pelagi*	396	(46.97,7.58,45.45)
EKF44211.1	*Nitratireductor indicus*	402	(43.53,11.94,44.53)
ABQ68897.1	*Sphingomonas wittichii*	465	(50.97,15.91,33.12)
WP_021236016.1	*Novosphingobium lindaniclasticum*	468	(51.92,10.68,37.39)
WP_008070462.1	*Novosphingobium nitrogenifigens*	434	(53.92,12.44,33.64)

WP_023838731.1	*Blastomonas sp.*	436	(51.38,12.39,36.24)
WP_009566963.1	*Rhodobacter sp.*	392	(7.65,5.61,36.73)
EME71470.1	*Magnetospirillum sp.*	452	(46.68,16.15,37.17)
WP_007437134.1	*Acetobacteraceae bacterium*	407	(48.40,7.13,44.47)
GAD60219.1	*Brevundimonas abyssalis*	435	(61.15,8.28,30.57)
ADO68171.1	*Stigmatella aurantiaca*	406	(47.29,14.29,38.42)
EOD53216.1	*Aeromonas molluscorum*	377	(63.40,7.96,28.65)
WP_020429815.1	*Vibrio fluvialis*	381	(49.34,11.29,39.37)
WP_017050438.1	*Vibrio ordalii*	382	(35.60,21.73,42.67)
ACY50816.1	*Vibrio sp.*	379	(41.42,12.66,45.91)
WP_023934304.1	*Photobacterium leiognathi*	383	(40.47,14.62,44.91)
WP_005491871.1	*Pseudoalteromonas luteoviolacea*	374	(35.29,20.86,43.85)
WP_008951656.1	*Alishewanella jeotgali*	401	(55.11,8.98,35.91)
WP_008983961.1	*Alishewanella agri*	402	(54.23,10.20,35.57)
WP_019020512.1	*Alishewanella agri*	387	(53.49,9.82,36.69)
EWG99922.1	*Halomonas sp*	302	(44.70,14.90,40.40)
WP_007113380.1	*Halomonas boliviensis*	400	(52.00,9.50,38.50)
WP_003636376.1	*Legionella longbeachae*	395	(39.49,14.18,46.33)
WP_009726276.1	*Methylophaga lonarensis*	393	(52.93,8.40,38.68)
WP_008221558.1	*Rheinheimera nanhaiensis*	455	(34.95,27.03,38.02)
WP_008163392.1	*Achromobacter arsenitoxydans*	418	(46.65,12.44,40.91)
WP_009857981.1	*Rubrivivax benzoatilyticus*	399	(51.38,10.28,38.35)

Short chain transporter (SCHR) protein

Table-2. Prediction of secondary structure of SCHR protein

SEQUENCE ID	**ORGANISM NAME**	**LENGTH OF AMINO ACID**	**SECONDARY STRUTURE (Alpha helix, Extended strand, Random coil) (in %)**
AFZ92424.1	*Bacillus amyloliquefaciens*	178	(51.69,8.43,39.89)
AGS66462.1	*Borrelia burgdorferi*	195	(48.72,17.44,33.85)
YP_007838646.1	*Faecalibacterium prausnitzii*	188	(37.77,18.09,44.15)
AAD45903.1	*Corynebacterium diphtheriae*	199	(54.27,11.06,34.67)
CAE28939.1	*Rhodopseudomonas palustris*	198	(57.58,8.59,33.84)
EFG95800.1	*Fusobacterium nucleatum*	186	(47.31,14.52,38.17)

AEN89059.1	*Bacillus megaterium*	176	(53.41,9.66,36.93)
YP_007840994.1	*Eubacterium rectale*	188	(37.23,24.47,38.30)
EIJ83658.1	*Bacillus methanolicus*	191	(40.31,18.32,41.36)
AEH52943.1	*Bacillus coagulans*	199	(48.74,12.56,38.69)
ADU29508.1	*Bacillus cellulosilyticus*	188	(29.26,25.00,45.74)
NP_229615.1	*Thermotoga maritima*	166	(40.36,22.29,37.35)
AFM41663.1	*Desulfosporosinus acidiphilus*	175	(48.00,20.00,32.00)
WP_004615133.1	*Caulobacter vibrioides*	176	(57.39,9.09,33.52)
ABV63928.1	*Bacillus pumilus*	178	(42.70,19.10,38.20)
ADC50166.1	*Bacillus pseudofirmus*	188	(35.11,26.06,38.83)
YP_007824800.1	*butyrate-producing bacterium*	180	(45.56,17.22,37.22)
ADK31722.1	*Brachyspira pilosicoli*	197	(34.52,22.34,43.15)
EIE39498.1	*Mycoplasma canis*	192	(18.23,34.38,47.40)
AEV25808.1	*Dechlorosoma suillum*	173	(50.29,14.45,35.26)
AET69926.1	*Desulfosporosinus orientis*	183	(36.07,28.69,34.97)
AEL69677.1	*Borrelia afzelii*	201	(53.23,14.93,31.84)
AFT83775.1	*Borrelia garinii*	194	(54.64,13.92,31.44)
ABE59094.1	*Chromohalobacter salexigens*	213	(43.19,18.78,38.03)
WP_009549950.1	*Burkholderiales bacterium*	193	(48.70,19.17,32.12)
AFY28759.1	*Cyanobium gracile*	198	(53.54,10.61,35.86)
AGE23918.1	*Geobacillus sp.*	193	(44.04,17.62,38.34)
ACT04510.1	*Paenibacillus sp.*	178	(30.34,26.97,42.70)
AFM52111.1	*Mycoplasma bovis*	217	(29.95,32.26,37.79)
AFZ69572.1	*Deinococcus peraridilitoris*	196	(45.92,5.10,48.98)
YP_007827638.1	*Fretibacterium fastidiosum*	193	(61.66,9.84,28.50)
AGB82453.1	*Serratia marcescens*	173	(57.80,14.45,27.75)
AFQ48990.1	*Burkholderia cepacia*	197	(40.61,23.86,35.53)
GAD12944.1	*Geobacillus kaustophilus*	196	(51.53,12.76,35.71)
CCF14649.1	*Brevibacillus laterosporus*	192	(36.46,26.04,37.50)
EFU42285.1	*Paenibacillus vortex*	189	(43.92,20.63,35.45)
WP_023509682.1	*Sporolactobacillus laevolacticus*	206	(36.41,22.33,41.26)
YP_007787904.1	*Ruminococcus torques*	186	(31.18,27.42,41.40)
WP_022256937.1	*Faecalibacterium sp.*	188	(48.40,13.30,38.30)
WP_004631164.1	*Clostridium termitidis*	195	(43.08,22.05,34.87)
AEV70027.1	*Clostridium clariflavum*	187	(24.60,28.34,47.06)
ACR73188.1	*Eubacterium eligens*	203	(41.87,24.63,33.50)
YP_001256815.1	*Mycoplasma agalactiae*	217	(24.88,36.87,38.25)
ADK80433.1	*Spirochaeta smaragdinae*	198	(33.84,26.26,39.90)

AEV29307.1	*Sphaerochaeta pleomorpha*	199	(41.71,25.63,32.66)
YP_007795149.1	*Bacteroides xylanisolvens*	182	(31.87,22.53,45.60)
YP_001525949.1	*Azorhizobium caulinodans*	176	(57.39,11.93,30.68)
ACM35260.1	*Agrobacterium vitis*	173	(26.59,27.17,46.24)
EGD06189.1	*Burkholderia sp.*	179	(51.96,15.08,32.96)
YP_007849682.1	*Clostridium cf.*	197	(33.50,21.83,44.67)
AFS77551.1	*Clostridium acidurici*	160	(46.25,18.12,35.62)
WP_021904986.1	*Clostridium sp*	180	(46.67,17.22,36.11)
ABX60169.1	*Halomonas anticariensis*	184	(23.91,41.30,34.78)
EHA15803.1	*Halomonas sp.*	176	(49.43,15.34,35.23)
EHK62369.1	*Halomonas sp.*	185	(34.59,30.81,34.59)
WP_021820586.1	*Halomonas sp.*	204	(43.63,20.10,36.27)
ACN14075.1	*Desulfobacterium autotrophicum*	161	(46.58,13.66,39.75)
ABX60169.1	*Halomonas anticariensis*	184	(23.91,41.30,34.78)
EHA15803.1	*Halomonas sp*	176	(49.43,15.34,35.23)
EHK62369.1	*Halomonas sp.*	185	(34.59,30.81,34.59)
WP_021820586.1	*Halomonas sp*	204	(43.63,20.10,36.27)
EJL79983.1	*Variovorax sp.*	189	(50.79,14.29,34.92)
AGB25619.1	*Mycobacterium smegmatis*	187	(41.71,13.90,44.39)
ABO68618.1	*Geobacillus thermodenitrificans*	193	(43.52,19.69,36.79)
EJL20165.1	*Brevibacillus sp*	176	(27.84,33.52,38.64)
YP_007770831.1	*Coprococcus catus*	187	(32.62,31.55,35.83)
YP_007795406.1	*Coprococcus sp.*	212	(37.26,23.58,39.15)
AEN97180.1	*Roseburia hominis*	129	(42.64,4.65,52.71)
AGC67747.1	*Clostridium stercorarium*	197	(12.18,36.04,51.78)
EDS13094.1	*Anaerotruncus colihominis*	184	(52.72,12.50,34.78)
KCZ68594.1	*Clostridium sporogenes*	178	(13.48,36.52,50.00)
AGA69406.1	*Desulfitobacterium dichloroeliminans*	173	(36.99,25.43,37.57)
AFQ42925.1	*Desulfosporosinus meridiei*	181	(35.91,18.78,45.30)
AEW06911.1	*Sulfobacillus acidophilus*	188	(31.91,26.60,41.49)
YP_007829706.1	*Ruminococcus champanellensis*	193	(48.19,22.80,29.02)
KCZ68594.1	*Clostridium sporogenes*	178	(13.48,36.52,50.00)
AGA69406.1	*Desulfitobacterium dichloroeliminans*	173	(36.99,25.43,37.57)
AFQ42925.1	*Desulfosporosinus meridiei*	181	(35.91,18.78,45.30)
AEW06911.1	*Sulfobacillus acidophilus*	188	(31.91,26.60,41.49)
EDN02096.1	*Pseudoflavonifractor capillosus*	198	(43.94,17.17,38.89)

YP_007834179.1	*Megamonas hypermegale*	138	(37.68,18.12,44.20)
NP_326559.1	*Mycoplasma pulmonis*	201	(46.77,11.44,41.79)
ACM22927.1	*Thermotoga neapolitana*	174	(36.21,18.97,44.83)
AAX16961.1	*Borrelia hermsii*	194	(48.97,15.46,35.57)
AAU07303.1	*Borrelia garinii*	197	(50.76,14.72,34.52)
AEL18626.1	*Borrelia bissettii*	195	(49.74,17.95,32.31)
AFG38607.1	*Spirochaeta africana*	203	(40.89,19.70,39.41)
WP_008134472.1	*Bradyrhizobium sp.*	176	(51.70,11.36,36.93)
WP_009491422.1	*Microvirga lotononidis*	185	(34.05,17.84,48.11)
EKN01360.1	*Acidocella sp.*	173	(52.60,12.72,34.68)
WP_010630089.1	*Halomonas sp.*	176	(40.34,22.73,36.93)
CDG52120.1	*Halomonas sp.*	176	(49.43,15.34,35.23)
YP_003624246.1	*Thiomonas arsenitoxydans*	177	(39.55,24.86,35.59)
EIF35615.1	*Burkholderia sp.*	215	(45.12,16.74,38.14)
AEK61791.1	*Collimonas fungivorans*	176	(42.05,17.61,40.34)
YP_007552597.1	*Azoarcus sp.*	182	(43.96,14.29,41.76)
ENO98478.1	*Thauera phenylacetica*	191	(40.31,21.99,37.70)
ETT35787.1	*Paenibacillus sp.*	198	(42.42,22.73,34.85)
EGA89899.1	*Planococcus donghaensis*	177	(38.42,14.69,46.89)
WP_018132767.1	*Alicyclobacillus pohliae*	176	(57.39,7.95,34.66)
ERK29329.1	*Clostridium intestinale*	189	(42.33,22.75,34.92)
EOR27970.1	*Clostridium sartagoforme*	178	(38.76,19.10,42.13)
CCY81219.1	*Clostridium sp*	189	(41.80,19.58,38.62)
YP_007797951.1	*[Eubacterium] cylindroides*	186	(39.78,24.19,36.02)
EGQ15609.1	*Prevotella dentalis*	175	(25.71,29.71,44.57)
ADA67097.1	*Thermotoga naphthophila*	166	(40.36,21.69,37.95)
ACB09353.1	*Thermotoga sp.*	166	(40.36,21.69,37.95)
ABQ47118.1	*Thermotoga petrophila*	171	(45.03,18.71,36.26)
YP_006347049.1	*Mesotoga prima*	177	(45.20,15.82,38.98)
WP_007691670.1	*Rhizobium sp.*	134	(47.01,8.96,44.03)
EHR05009.1	*Bradyrhizobium sp.*	205	(59.02,5.37,35.61)
EIG61708.1	*Bradyrhizobium sp.*	176	(48.30,17.05,34.66)
ABB32950.1	*Geobacter metallireducens*	184	(45.11,16.85,38.04)
ACH37741.1	*Geobacter bemidjiensis*	174	(56.90,10.92,32.18)
WP_021806463.1	*Serratia fonticola*	173	(45.66,24.86,29.48)
AFE58234.1	*Rahnella aquatilis*	194	(44.33,19.59,36.08)
AEX51981.1	*Rahnella aquatilis*	195	(49.74,17.44,32.82)
WP_019808580.1	*Osedax symbiont*	181	(30.94,27.07,41.99)
WP_016916520.1	*Halomonas stevensii*	182	(53.30,4.40,42.31)
WP_023495244.1	*Methyloglobulus morosus*	147	(35.37,18.37,46.26)
AFK63870.1	*Advenella kashmirensis*	177	(44.63,20.90,34.46)

IV. Functional analysis of chromate ion transporter protein

Long chain transporter (LCHR) protein

Functional domains were predicted of all LCHR proteins. LCHR proteins contained chromate transporter domains (IPR014047&IPR003370) and other domain like IPR006187. Most of the LCHR proteins contained both chromate transporter domains IPR014047 and IPR003370.One LCHR protein of fungi that was *Coprinopsis cinerea* having amino acid length 546 and another LCHR protein of one organism that was *Kyrpidia tusciae* having amino acid length 252 contained only one chromate transporter domain IPR003370. But one LCHR protein of one organism that was *Arthrobacter sp.* having amino acid length 277 contained one chromate transporter domain IPR003370 and another claudin domain IPR006187.

The function of all domains of LCHR proteins were given below:

- IPR014047 (Chromate transporter, long chain)
    - This entry represents the long chain chromate transporters [PMID: 2152903, PMID: 2180932,PMID: 17986256]. The protein reduces chromate accumulation and is essential for chromate resistance. They appear to have arisen from a gene fusion event of two short chain transporters [PMID: 19581367].

- IPR003370 (Chromate transporter)
    - This entry represents chromate transporters (CHR) [PMID: 2152903, PMID: 2180932]. These proteins reduce chromate accumulation and are essential for chromate resistance. They are composed of one or two copies of this region. The short-chain

CHR proteins form heterodimer transporters which efflux chromate ions from the cytoplasm, while the long chain CHR proteins appear to have arisen from a gene fusion event of two short chain transporters[PMID: 19581367].

- IPR006187(Claudin)

    - Zona occludens (ZO), or tight junctions (TJ) are specialised membrane domains found at the most apical region of polarised epithelial and endothelial cells. They create a primary barrier, preventing paracellular transport of solutes, and restricting the lateral diffusion of membrane lipids and proteins, thus maintaining cellular polarity [PMID: 9647647]. They also act as diffusion barriers within plasma membranes, creating and maintaining apical and basolateral membrane domains. Under freeze-fracture electron microscopy, TJs appear as a network of continuous anastomosing intramembranous strands. These strands consist mainly of claudins and occludin (IPR005417), which are transmembrane proteins that polymerise within plasma membranes to form fibrils [PMID: 11283726].

Recently, the molecular architecture of tight junctions has begun to be elucidated. One group of proteins thought to be major components of TJs is the claudin family [PMID: 10370242]. Immunofluorescence studies have shown that claudins are targeted to and incorporated into tight junctions [PMID: 9647647]. Furthermore, when claudins are introduced into cells that lack tight junctions,

networks of strands and grooves form at cell-cell contact sites that closely resemble native TJs [PMID: 9786950].

The claudin protein family is encoded by at least 17 human genes, with many homologues cloned from other species. Tissue distribution patterns for the claudin family members are distinct. Claudin-1 and -2, for example, are expressed at high levels in the liver and kidney, whereas claudin-3 mRNA is detected mainly in the lung and liver [PMID: 9647647, PMID: 9892664]. This suggests that multiple claudin family members may be involved in tight junction strand formation in a tissue-dependent manner.

Hydropathy analysis suggests that all claudins share a common transmembrane (TM) topology. Each family member is predicted to possess four TM domains with intracellular N and C termini. Although their C-terminal cytoplasmic domain sequences vary, most claudin family members share a common motif of -Y-V in this region. This has been postulated as a possible binding motif for PDZ domains of other tight junction-associated membrane proteins, such as ZO-1 (IPR005418).

Table-3. Functional analysis of domain of LCHR protein

SEQUENCE ID	ORGANISM NAME	LENGTH OF AMINO ACID	DOMAIN
ETR55227.1	*Acinetobacter baumannii*	449	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABJ13559.1	*Pseudomonas aeruginosa*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACM15429.1	*Bacillus cereus*	393	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AHY00005.1	*Klebsiella pneumoniae*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EMQ66925.1	*Vibrio cholerae*	380	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_012372815.1	*Escherichia coli*	389	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABN95479.1	*Burkholderia pseudomallei*	396	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ADM37870.1	*Bacillus subtilis*	398	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EAP72482.1	*Ralstonia solanacearum*	448	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

AGA73085.1	*Pseudomonas putida*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_039193.1	*Bacillus thuringiensis*	393	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AHI98792.1	*Vibrio parahaemolyticus*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CDN64820.1	*Burkholderia cenocepacia*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFN78404.1	*Pseudomonas stutzeri*	403	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_021161752.1	*Burkholderia vietnamiensis*	441	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ETI07845.1	*Bordetella pertussis*	448	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)
AEV63938.1	*Pseudomonas fluorescens*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_008600505.1	*Ralstonia pickettii*	413	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
GAF00771.1	*Bacillus anthracis*	393	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_001946933.1	*Burkholderia multivorans*	403	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

EWH21927.1	*Bacillus licheniformis*	397	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_024068811.1	*Achromobacter xylosoxidans*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CAK08671.1	*Rhizobium leguminosarum*	473	IPR014047(Chromate transporter, long chain IPR003370(Chromate transporter))
ABM06203.1	*Comamonas testosteroni*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ELP55763.1	*Microcystis aeruginosa*	385	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CDH84656.1	*Sinorhizobium meliloti*	470	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEI41168.1	*Paenibacillus mucilaginosus*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_004769390.1	*Leptospira kirschneri*	454	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CCO52615.1	*Vibrio nigripulchritudo*	382	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABS09622.1	*Shewanella baltica*	418	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACB65671.1	*Burkholderia ambifaria*	392	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

ABC36222.1	*Burkholderia thailandensis*	396	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EJO94496.1	*Pseudomonas mendocina*	444	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_007561622.1	*Corynebacterium glutamicum*	376	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AHK02891.1	*Agrobacterium tumefaciens*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_352338.1	*Rhodobacter sphaeroides*	392	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
KCV66631.1	*Bordetella bronchiseptica*	448	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_103831.1	*Burkholderia mallei*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACS39230.1	*Methylobacterium extorquens*	467	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)
YP_006966045.1	*Aeromonas hydrophila*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGU47511.1	*Variovorax paradoxus*	448	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CDG75483.1	*Acinetobacter nosocomialis*	449	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

EEY78117.1	*Acinetobacter calcoaceticus*	448	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EPR88996.1	*Acinetobacter haemolyticus*	450	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFJ86665.1	*Burkholderia sp.*	405	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_017399575.1	*Acinetobacter pittii*	449	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EAP94742.1	*Vibrio splendidus*	382	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
NP_694198.1	*Oceanobacillus iheyensis*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EPG55962.1	*Leptospira borgpetersenii*	488	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_016560147.1	*Leptospira noguchii*	461	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACI91597.1	*Oligotropha carboxidovorans*	384	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EEY44780.1	*Vibrio mimicus*	380	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ADC70598.1	*Thioalkalivibrio sp.*	450	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

YP_161712.1	*Cupriavidus metallidurans*	374	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABR90147.1	*Janthinobacterium sp.*	477	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGK04154.1	*Meiothermus ruber*	349	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABC91487.1	*Rhizobium etli*	473	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EXL01874.1	*Ochrobactrum anthropi*	408	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_006184233.1	*Stenotrophomonas maltophilia*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
NP_442845.1	*Synechocystis sp*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_002774542.1	*Brevibacillus brevis*	411	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EPH11961.1	*Myroides odoratimimus*	379	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)
YP_005191479.1	*Sinorhizobium fredii*	471	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEH84626.1	*Mesorhizobium opportunistum*	463	IPR014047(Chromate IPR003370(Chromate transporter)transporter, long chain)

AFM60000.1	*Enterobacter cloacae*	455	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AAO10171.1	*Vibrio vulnificus*	383	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEB48657.1	*Aeromonas veronii*	388	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGP93754.1	*Alteromonas macleodii*	485	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EPY13218.1	*Paenibacillus alvei*	407	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_005585600.1	*Clostridium ultunense*	332	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EDL68584.1	*Vibrio campbellii*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_009133.1	*Desulfovibrio vulgaris*	445	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEI80435.1	*Cupriavidus necator*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_016704768.1	*Pseudomonas chlororaphis*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
GAD72596.1	*Vibrio alginolyticus*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

AFU46855.1	*Acidovorax sp.*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004389086.1	*Alicycliphilus denitrificans*	406	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EJO33504.1	*Achromobacter piechaudii*	388	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AET87449.1	*uncultured bacterium*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ELK48494.1	*Halobacillus sp.*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACL60922.1	*Methylobacterium nodulans*	469	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGK58336.1	*Hyphomicrobium denitrificans*	434	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CDM42712.1	*Pseudomonas pseudoalcaligenes*	444	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EHN70797.1	*Vibrio fischeri*	383	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_003896680.1	*Halomonas elongata*	399	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)
YP_001630202.1	*Bordetella petrii*	404	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

YP_002008176.1	*Cupriavidus taiwanensis*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_171835.1	*Synechococcus elongatus*	383	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABM77873.1	*Prochlorococcus marinus*	394	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_006182093.1	*Halobacillus halophilus*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGT31995.1	*Geobacillus sp.*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACX66242.1	*Paenibacillus sp.*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABL72111.1	*Paracoccus denitrificans*	405	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEX22965.1	*Vibrio sp.*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABE37066.1	*Burkholderia xenovorans*	398	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EON15238.1	*Pandoraea sp.*	410	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ADJ62751.1	*Herbaspirillum seropedicae*	482	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

ENO82674.1	*Thauera sp.*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CCQ60486.1	*Crocosphaera watsonii*	389	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_006463894.1	*Solibacillus silvestris*	394	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004437607.1	*Thermodesulfobium narugense*	368	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EFH84111.1	*Ktedonobacter racemifer*	428	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AHD09798.1	*Phaeobacter gallaeciensis*	432	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACO80512.1	*Azotobacter vinelandii*	453	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABM95481.1	*Methylibium petroleiphilum*	445	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABX35705.1	*Delftia acidovorans*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEG94919.1	*Ramlibacter tataouinensis*	457	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AAQ19822.1	*Alcaligenes faecalis*	408	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)

EAY70264.1	*Burkholderia dolosa*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CAK28037.1	*Synechococcus sp*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004101092.1	*Thermaerobacter marianensis*	450	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEV99229.1	*Niastella koreensis*	392	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABE62743.1	*Nitrobacter hamburgensis*	430	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABQ36693.1	*Bradyrhizobium sp*	463	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEQ52600.1	*Pelagibacterium halotolerans*	464	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACB82306.1	*Methylobacterium populi*	469	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGI71815.1	*Octadecabacter arcticus*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_003447441.1	*Azospirillum sp*	429	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGH97925.1	*Micavibrio aeruginosavorus*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

AHC87436.1	*Pseudomonas monteilii*	453	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_011061436.1	*Pseudomonas protegens*	395	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CBL87892.1	*Pseudomonas sp.*	455	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFN31616.1	*Klebsiella oxytoca*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGU57013.1	*Listonella anguillarum*	382	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ERM60026.1	*Vibrio cyclitrophicus*	380	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABO91542.1	*Aeromonas salmonicida*	384	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EGI73761.1	*Pseudoalteromonas haloplanktis*	384	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004481554.1	*Marinomonas posidonica*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABQ36693.1	*Bradyrhizobium sp*	463	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEQ52600.1	*Pelagibacterium halotolerans*	464	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

ACB82306.1	*Methylobacterium populi*	469	IPR014047(Chromate transporter, long chain)IPR003370(Chromate transporter)
AGI71815.1	*Octadecabacter arcticus*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_003447441.1	*Azospirillum sp*	429	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AGH97925.1	*Micavibrio aeruginosavorus*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_001619890.1	*Sorangium cellulosum*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFY01497.1	*Bdellovibrio bacteriovorus*	335	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_005437080.1	*Rubrivivax gelatinosus*	447	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_006897362.1	*Bordetella parapertussis*	441	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AAZ64575.1	*Ralstonia eutropha*	413	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EHP40784.1	*Cupriavidus basilensis*	464	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACC74843.1	*Burkholderia phymatum*	412	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

EIM94433.1	*Burkholderia terrae*	394	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFT85530.1	*Burkholderia phenoliruptrix*	442	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004227940.1	*Burkholderia sp*	442	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABE46762.1	*Polaromonas sp*	456	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACM33301.1	*Acidovorax ebreus*	407	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABD70536.1	*Albidiferax ferrireducens*	417	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AAQ58596.1	*Chromobacterium violaceum*	393	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ENO88273.1	*Thauera aminoaromatica*	440	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ENO93942.1	*Thauera sp*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_020166052.1	*Methyloversatilis universalis*	445	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_021478256.1	*Pseudogulbenkiania ferrooxidans*	386	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

YP_006963560.1	*Arthrobacter sp.*	277	IPR003370(Chromate transporter) IPR006187(Claudin)
ABL81046.1	*Nocardioides sp*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EWH33178.1	*Lysinibacillus sphaericus*	398	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007503781.1	*Caldalkalibacillus thermarum*	403	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EQB34717.1	*Virgibacillus sp*	398	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_009590732.1	*Paenibacillus sp.*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007128576.1	*Paenibacillus lactis*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
GAE04947.1	*Paenibacillus sp.*	407	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_023559015.1	*Brevibacillus panacihumi*	402	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EIM08287.1	*Planococcus antarcticus*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008301397.1	*Bhargavaea cecembensis*	397	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

ADG06946.1	*Kyrpidia tusciae*	252	IPR003370(Chromate transporter)
WP_009577949.1	*Fulvivirga imtechensis*	416	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACU05572.1	*Pedobacter heparinus*	392	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_004448553.1	*Haliscomenobacter hydrossis*	447	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
XP_002911996.1	*Coprinopsis cinerea (Fungi)*	546	IPR003370(Chromate transporter)
YP_005609514.1	*Bradyrhizobium japonicum*	461	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ELT47282.1	*Ochrobactrum intermedium*	432	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AAQ18188.1	*Rhodobacter capsulatus*	449	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_005794667.1	*Ketogulonicigenium vulgare*	385	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EJU14087.1	*Sphingomonas sp*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
YP_003545570.1	*Sphingobium japonicum*	463	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_021228313.1	*Sphingobium lactosutens*	434	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

AFK56631.1	*Tistrella mobilis*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EDP65385.1	*alpha proteobacterium*	448	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EMR36372.1	*Vibrio harveyi*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_020334329.1	*Vibrio natriegens*	380	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_017070361.1	*Vibrio crassostreae*	381	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EEX38787.1	*Vibrio furnissii*	356	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EAS42875.1	*Photobacterium profundum*	388	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABP76766.1	*Shewanella putrefaciens*	383	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABK50579.1	*Shewanella sp.*	455	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EGP21262.1	*Halomonas sp.*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_009287634.1	*Halomonas titanicae*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

AAZ98609.1	*Thiobacillus denitrificans*	507	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AEV85163.1	*Actinoplanes sp.*	474	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABK05889.1	*Arthrobacter sp.*	450	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EID44387.1	*Geobacillus thermoglucosidans*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_017798519.1	*Oceanobacillus kimchii*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_018924065.1	*Salsuginibacillus kocurii*	402	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ETT66405.1	*Paenibacillus sp.*	409	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EMT53387.1	*Brevibacillus borstelensis*	405	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_019122516.1	*Brevibacillus massiliensis*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABZ92904.1	*Leptospira biflexa*	390	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABG59418.1	*Cytophaga hutchinsonii*	385	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

ABQ07421.1	*Flavobacterium johnsoniae*	376	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EJG02233.1	*Flavobacterium sp.*	376	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ADY52544.1	*Pedobacter saltans*	388	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABF44423.1	*Deinococcus geothermalis*	387	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
AFL88631.1	*Terriglobus roseus*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007816752.1	*Rhizobium sp*	413	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007802676.1	*Rhizobium sp*	469	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EJZ18077.1	*Rhizobium sp.*	410	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EPE99336.1	*Rhizobium grahamii*	469	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
CCM74575.1	*Rhizobium mesoamericanum*	470	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
KAK31649.1	*Bradyrhizobium diazoefficiens*	402	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

WP_008974529.1	*Bradyrhizobium sp*	463	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EGP06243.1	*Bradyrhizobiaceae bacterium*	444	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EAU42843.1	*Fulvimarina pelagi*	396	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EKF44211.1	*Nitratireductor indicus*	402	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ABQ68897.1	*Sphingomonas wittichii*	465	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_021236016.1	*Novosphingobium lindaniclasticum*	468	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008070462.1	*Novosphingobium nitrogenifigens*	434	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_023838731.1	*Blastomonas sp.*	436	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_009566963.1	*Rhodobacter sp.*	392	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EME71470.1	*Magnetospirillum sp.*	452	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007437134.1	*Acetobacteraceae bacterium*	407	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

GAD60219.1	*Brevundimonas abyssalis*	435	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ADO68171.1	*Stigmatella aurantiaca*	406	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
EOD53216.1	*Aeromonas molluscorum*	377	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_020429815.1	*Vibrio fluvialis*	381	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_017050438.1	*Vibrio ordalii*	382	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
ACY50816.1	*Vibrio sp.*	379	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_023934304.1	*Photobacterium leiognathi*	383	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_005491871.1	*Pseudoalteromonas luteoviolacea*	374	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008951656.1	*Alishewanella jeotgali*	401	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008983961.1	*Alishewanella agri*	402	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_019020512.1	*Alishewanella agri*	387	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

EWG99922.1	*Halomonas sp*	302	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_007113380.1	*Halomonas boliviensis*	400	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_003636376.1	*Legionella longbeachae*	395	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_009726276.1	*Methylophaga lonarensis*	393	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008221558.1	*Rheinheimera nanhaiensis*	455	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_008163392.1	*Achromobacter arsenitoxydans*	418	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)
WP_009857981.1	*Rubrivivax benzoatilyticus*	399	IPR014047(Chromate transporter, long chain) IPR003370(Chromate transporter)

Short chain transporter (SCHR) protein

Functional domains were predicted of all SCHR proteins. SCHR proteins contained chromate transporter domain (IPR003370) and some other domains likeIPR011006, IPR001789, IPR011991, IPR016032, IPR000792 and IPR024414. Most of the SCHR proteins contained chromate transporter domain IPR003370. One SCHR protein of one organism that was *Roseburia hominis* having amino acid length 129 contained uncharacterized domain IPR024414. But one SCHR protein of one organism that was *Corynebacterium diphtheriae* having amino acid length 199 contained five

types of other domains IPR011006, IPR001789, IPR011991, IPR016032 and IPR000792.

The function of all domains of SCHR proteins were given below:

- IPR003370 (Chromate transporter)
    - This entry represents chromate transporters (CHR) [PMID: 2152903, PMID: 2180932]. These proteins reduce chromate accumulation and are essential for chromate resistance. They are composed of one or two copies of this region. The short-chain CHR proteins form heterodimer transporters which efflux chromate ions from the cytoplasm, while the long chain CHR proteins appear to have arisen from a gene fusion event of two short chain transporters[PMID: 19581367].
- IPR011006 (CheY-like superfamily)
    - CheY is a member of the response regulator family in bacterial two-component signalling systems, where CheY receives the signal from the sensor partner, usually a histidine protein kinase. Signal transduction involves phosphotransfer, whereby the histidine kinase phosphorylates a conserved aspartate in the response regulator to activate responses to environmental signals [PMID: 14636076]. CheY is a single domain protein that folds into a compact globular unit with a flavodoxin-like fold consisting of three-layer alpha/beta/alpha sandwich with 21345 beta topology, where the phosphorylation region lies in a cavity.

Other members of the response regulator family contain a CheY-like receiver domain, which is often found N-terminal to a DNA-binding effector domain. Examples include NarL (nitrate/nitrite response regulator), NtrC (nitrogen regulatory protein C), Spo0A and Spo0F (sporulation response) from Bacillus, PhoA and PhoB cyclin-dependent kinases from Aspergillus, among others.

AmiR, the positive regulator of the amidase operon in Psuedomonas, is an unusual member of the bacterial response regulator family; AmiR is able to bind RNA and uses ligand-regulated activation rather than phopho-activation. It has a CheY-like fold at its N terminus, but contains two subdomains in a C-terminal extension, one forming a coiled-coil and the other a long alpha helix. As such AmiR may represent a new family of RNA-binding response regulators [PMID: 10508151].

CheY-like domains can be found in other protein families as well. Examples include the receiver domain of the ethylene receptor (ETR1) from Arabidopsis, which is involved in ethylene detection and signal transduction [PMID: 10647185]; the N-terminal wing' domain of ornithine decarboxylase from Lactobacilli, which catalyses the conversion of ornithine to putrescine at the beginning of the polyamine pathway [PMID: 10666573]. The N-terminal domain of the circadian clock protein, KaiA, from cyanobacteria, acts as a psuedo-receiver domain, but lacks the conserved aspartyl residue required for phosphotransfer in response regulators [PMID: 12438647].

- IPR001789 (Signal transduction response regulator, receiver domain)

Two-component signal transduction systems enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth

conditions [PMID: 16176121]. Some bacteria can contain up to as many as 200 two-component systems that need tight regulation to prevent unwanted cross-talk [PMID: 18076326]. These pathways have been adapted to response to a wide variety of stimuli, including nutrients, cellular redox state, changes in osmolarity, quorum signals, antibiotics, and more [PMID: 12372152]. Two-component systems are comprised of a sensor histidine kinase (HK) and its cognate response regulator (RR) [PMID: 10966457]. The HK catalyses its own auto-phosphorylation followed by the transfer of the phosphoryl group to the receiver domain on RR; phosphorylation of the RR usually activates an attached output domain, which can then effect changes in cellular physiology, often by regulating gene expression. Some HK are bifunctional, catalysing both the phosphorylation and dephosphorylation of their cognate RR. The input stimuli can regulate either the kinase or phosphatase activity of the bifunctional HK.

A variant of the two-component system is the phospho-relay system. Here a hybrid HK auto-phosphorylates and then transfers the phosphoryl group to an internal receiver domain, rather than to a separate RR protein. The phosphoryl group is then shuttled to histidine phosphotransferase (HPT) and subsequently to a terminal RR, which can evoke the desired response [PMID: 11934609,PMID: 11489844].

Bipartite response regulator proteins are involved in a two-component signal transduction system in bacteria, and certain eukaryotes like protozoa, that functions to detect and respond to environmental changes [PMID: 7699720]. These systems have been detected during host invasion, drug resistance, motility, phosphate uptake, osmoregulation, and nitrogen fixation, amongst others [PMID: 12015152]. The two-component system consists of a histidine protein kinase

environmental sensor that phosphorylates the receiver domain of a response regulator protein; phosphorylation induces a conformational change in the response regulator, which activates the effector domain, triggering the cellular response [PMID: 10966457]. The domains of the two-component proteins are highly modular, but the core structures and activities are maintained.

The response regulators act as phosphorylation-activated switches to affect a cellular response, usually by transcriptional regulation. Most of these proteins consist of two domains, an N-terminal response regulator receiver domain, and a variable C-terminal effector domain with DNA-binding activity. This entry represents the response regulator receiver domain, which belongs to the CheY family, and receives the signal from the sensor partner in the two-component system.

- IPR011991 (Winged helix-turn-helix DNA-binding domain)

Winged helix DNA-binding proteins share a related winged helix-turn-helix DNA-binding motif, where the "wings", or loops, are small beta-sheets. The winged helix motif consists of two wings (W1, W2), three alpha helices (H1, H2, H3) and three beta-sheets (S1, S2, S3) arranged in the order H1-S1-H2-H3-S2-W1-S3-W2 [PMID: 10679470]. The DNA-recognition helix makes sequence-specific DNA contacts with the major groove of DNA, while the wings make different DNA contacts, often with the minor groove or the backbone of DNA. Several winged-helix proteins display an exposed patch of hydrophobic residues thought to mediate protein-protein interactions.

Many different proteins with diverse biological functions contain a winged helix DNA-binding domain, including transcriptional repressors such as biotin repressor, LexA repressor and the arginine repressor [PMID: 1409631];

transcription factors such as the hepatocyte nuclear factor-3 proteins involved in cell differentiation, heat-shock transcription factor, and the general transcription factors TFIIE and TFIIF [PMID: 8248124, PMID: 11292844]; helicases such as RuvB that promotes branch migration at the Holliday junction, and CDC6 in the pre-replication complex [PMID: 12408833, PMID: 11030343]; endonucleases such as FokI and TnsA [PMID: 9214510]; histones; and Mu transposase, where the flexible wing of the enhancer-binding domain is essential for efficient transposition [PMID: 8577730].

- IPR016032 (Signal transduction response regulator, C-terminal effector)

  Two-component signal transduction systems enable bacteria to sense, respond, and adapt to a wide range of environments, stressors, and growth conditions [PMID: 16176121]. Some bacteria can contain up to as many as 200 two-component systems that need tight regulation to prevent unwanted cross-talk [PMID: 18076326]. These pathways have been adapted to response to a wide variety of stimuli, including nutrients, cellular redox state, changes in osmolarity, quorum signals, antibiotics, and more [PMID: 12372152]. Two-component systems are comprised of a sensor histidine kinase (HK) and its cognate response regulator (RR) [PMID: 10966457]. The HK catalyses its own auto-phosphorylation followed by the transfer of the phosphoryl group to the receiver domain on RR; phosphorylation of the RR usually activates an attached output domain, which can then effect changes in cellular physiology, often by regulating gene expression. Some HK are bifunctional, catalysing both the phosphorylation and dephosphorylation of their cognate RR. The input stimuli can regulate either the kinase or phosphatase activity of the bifunctional HK.

A variant of the two-component system is the phospho-relay system. Here a hybrid HK auto-phosphorylates and then transfers the phosphoryl group to an internal receiver domain, rather than to a separate RR protein. The phosphoryl group is then shuttled to histidine phosphotransferase (HPT) and subsequently to a terminal RR, which can evoke the desired response [PMID: 11934609,PMID: 11489844].

This entry represents a structural domain usually found at the C-terminal of bipartite response regulators. These proteins are known to bind to DNA and RNA polymerases, and their N-terminal receiver domain belongs to the CheY family. The C-terminal effector domain consists of a 3-helical bundle in an up-an-down arrangement with a right-handed twist. This domain occurs in:

- PhoB-like proteins, which includes PhoB [PMID: 12015152], OmpR [PMID: 8989318], and DrrB [PMID: 12837793]; these proteins contain a 4-stranded meander beta-sheet in the N-terminal extension.
- GerE-like proteins from the LuxR/UhpA family of proteins, which includes GerE [PMID: 11243786], TraR (quorum-sensing) [PMID: 12198141], NarL (nitrate/nitrite response regulator) [PMID: 9521685], and RcsB transcriptional regulator [PMID: 12740396]; these proteins contain an additional fourth helix in the C-terminal extension.
- Spo0A proteins [PMID: 11069648], which are elaborated with additional helices.

➢ IPR000792 (Transcription regulator LuxR, C-terminal)

This domain is a DNA-binding, helix-turn-helix (HTH) domain of about 65 amino acids, present in transcription regulators of the LuxR/FixJ family of response regulators. The domain is named after Vibrio fischeri luxR, a transcriptional activator for quorum-sensing control of luminescence. LuxR-type

HTH domain proteins occur in a variety of organisms. The DNA-binding HTH domain is usually located in the C-terminal region; the N-terminal region often containing an autoinducer-binding domain or a response regulatory domain. Most luxR-type regulators act as transcription activators, but some can be repressors or have a dual role for different sites. LuxR-type HTH regulators control a wide variety of activities in various biological processes.

The luxR-type, DNA-binding HTH domain forms a four-helical bundle structure. The HTH motif comprises the second and third helices, known as the scaffold and recognition helix, respectively. The HTH binds DNA in the major groove, where the N-terminal part of the recognition helix makes most of the DNA contacts. The fourth helix is involved in dimerisation of gerE and traR. Signalling events by one of the four activation mechanisms described below lead to multimerisation of the regulator. The regulators bind DNA as multimers [PMID: 11243786, PMID: 12740396, PMID: 12087407].

LuxR-type HTH proteins can be activated by one of four different mechanisms:

1) Regulators which belong to a two-component sensory transduction system where the protein is activated by its phosphorylation, generally on an aspartate residue, by a transmembrane kinase [PMID: 12352954, PMID: 12162958]. Some proteins that belong to this category are:

- Rhizobiaceae fixJ (global regulator inducing expression of nitrogen-fixation genes in microaerobiosis).
- *Escherichia coli* and *Salmonella typhimurium* uhpA (activates hexose phosphate transport gene uhpT).
- *E. coli* narL and narP (activate nitrate reductase operon).

- Enterobacteria rcsB (regulation of exopolysaccharide biosynthesis in enteric and plant pathogenesis).
- *Bordetella pertussis* bvgA (virulence factor).
- *Bacillus subtilis* coma (involved in expression of late-expressing competence genes).

2) Regulators which are activated, or in very rare cases repressed, when bound to N-acyl homoserine lactones, which are used as quorum sensing molecules in a variety of Gram-negative bacteria [PMID: 15255890]:

- *V. fischeri* luxR (activates bioluminescence operon).
- *Agrobacterium tumefaciens* traR (regulation of Ti plasmid transfer).
- *Erwinia carotovora* carR (control of carbapenem antibiotics biosynthesis).
- *E. carotovora* expR (virulence factor for soft rot disease; activates plant tissue macerating enzyme genes).
- *Pseudomonas aeruginosa* lasR (activates elastase gene lasB).
- *Erwinia chrysanthemi* echR and *Erwinia stewartii* esaR.
- *Pseudomonas chlororaphis* phzR (positive regulator of phenazine antibiotic production).
- *Pseudomonas aeruginosa* rhlR (activates rhlAB operon and lasB gene).

3) Autonomous effector domain regulators, without a regulatory domain, represented by gerE [PMID: 11243786]:

- *B. subtilis* gerE (transcription activator and repressor for the regulation of spore formation).

4) Multiple ligand-binding regulators, exemplified by malT [PMID: 11931562]:

- *E. coli* malT (activates maltose operon; MalT binds ATP and maltotriose).

➢ IPR024414 (Uncharacterised protein fam Prg I)

This family of bacterial proteins is functionally uncharacterised. Proteins in this family are typically between 116 and 146 amino acids in length. PrgI is found encoded on plasmids of *Enterococcus faecalis*, its function is not known.

Table-4. Functional analysis of domain of SCHR protein

SEQUENCE ID	ORGANISM NAME	LENGTH OF AMINO ACID	DOMAIN
AFZ92424.1	*Bacillus amyloliquefaciens*	178	IPR003370 (Chromate transporter)
AGS66462.1	*Borrelia burgdorferi*	195	IPR003370 (Chromate transporter)
YP_007838646.1	*Faecalibacterium prausnitzii*	188	IPR003370 (Chromate transporter)
AAD45903.1	*Corynebacterium diphtheriae*	199	IPR011006 IPR001789 IPR011991 IPR016032 IPR000792
CAE28939.1	*Rhodopseudomonas palustris*	198	IPR003370 (Chromate transporter)

EFG95800.1	*Fusobacterium nucleatum*	186	IPR003370 (Chromate transporter)
AEN89059.1	*Bacillus megaterium*	176	IPR003370 (Chromate transporter)
YP_007840994.1	*Eubacterium rectale*	188	IPR003370 (Chromate transporter)
EIJ83658.1	*Bacillus methanolicus*	191	IPR003370 (Chromate transporter)
AEH52943.1	*Bacillus coagulans*	199	IPR003370 (Chromate transporter)
ADU29508.1	*Bacillus cellulosilyticus*	188	IPR003370 (Chromate transporter)
NP_229615.1	*Thermotoga maritima*	166	IPR003370 (Chromate transporter)
AFM41663.1	*Desulfosporosinus acidiphilus*	175	IPR003370 (Chromate transporter)
WP_004615133.1	*Caulobacter vibrioides*	176	IPR003370 (Chromate transporter)
ABV63928.1	*Bacillus pumilus*	178	IPR003370 (Chromate transporter)
ADC50166.1	*Bacillus pseudofirmus*	188	IPR003370 (Chromate transporter)
YP_007824800.1	*butyrate-producing bacterium*	180	IPR003370 (Chromate transporter)
ADK31722.1	*Brachyspira pilosicoli*	197	IPR003370 (Chromate transporter)
EIE39498.1	*Mycoplasma canis*	192	IPR003370 (Chromate transporter)

AEV25808.1	*Dechlorosoma suillum*	173	IPR003370 (Chromate transporter)
AET69926.1	*Desulfosporosinus orientis*	183	IPR003370 (Chromate transporter)
AEL69677.1	*Borrelia afzelii*	201	IPR003370 (Chromate transporter)
AFT83775.1	*Borrelia garinii*	194	IPR003370 (Chromate transporter)
ABE59094.1	*Chromohalobacter salexigens*	213	IPR003370 (Chromate transporter)
WP_009549950.1	*Burkholderiales bacterium*	193	IPR003370 (Chromate transporter)
AFY28759.1	*Cyanobium gracile*	198	IPR003370 (Chromate transporter)
AGE23918.1	*Geobacillus sp.*	193	IPR003370 (Chromate transporter)
ACT04510.1	*Paenibacillus sp.*	178	IPR003370 (Chromate transporter)
AFM52111.1	*Mycoplasma bovis*	217	IPR003370 (Chromate transporter)
AFZ69572.1	*Deinococcus peraridilitoris*	196	IPR003370 (Chromate transporter)
YP_007827638.1	*Fretibacterium fastidiosum*	193	IPR003370 (Chromate transporter)
AGB82453.1	*Serratia marcescens*	173	IPR003370 (Chromate transporter)
AFQ48990.1	*Burkholderia cepacia*	197	IPR003370 (Chromate transporter)

GAD12944.1	*Geobacillus kaustophilus*	196	IPR003370 (Chromate transporter)
CCF14649.1	*Brevibacillus laterosporus*	192	IPR003370 (Chromate transporter)
EFU42285.1	*Paenibacillus vortex*	189	IPR003370 (Chromate transporter)
WP_023509682.1	*Sporolactobacillus laevolacticus*	206	IPR003370 (Chromate transporter)
YP_007787904.1	*Ruminococcus torques*	186	IPR003370 (Chromate transporter)
WP_022256937.1	*Faecalibacterium sp.*	188	IPR003370 (Chromate transporter)
WP_004631164.1	*Clostridium termitidis*	195	IPR003370 (Chromate transporter)
AEV70027.1	*Clostridium clariflavum*	187	IPR003370 (Chromate transporter)
ACR73188.1	*Eubacterium eligens*	203	IPR003370 (Chromate transporter)
YP_001256815.1	*Mycoplasma agalactiae*	217	IPR003370 (Chromate transporter)
ADK80433.1	*Spirochaeta smaragdinae*	198	IPR003370 (Chromate transporter)
AEV29307.1	*Sphaerochaeta pleomorpha*	199	IPR003370 (Chromate transporter)
YP_007795149.1	*Bacteroides xylanisolvens*	182	IPR003370 (Chromate transporter)
YP_001525949.1	*Azorhizobium caulinodans*	176	IPR003370 (Chromate transporter)

ACM35260.1	*Agrobacterium vitis*	173	IPR003370 (Chromate transporter)
EGD06189.1	*Burkholderia sp.*	179	IPR003370 (Chromate transporter)
YP_007849682.1	*Clostridium cf.*	197	IPR003370 (Chromate transporter)
AFS77551.1	*Clostridium acidurici*	160	IPR003370 (Chromate transporter)
WP_021904986.1	*Clostridium sp*	180	IPR003370 (Chromate transporter)
ABX60169.1	*Halomonas anticariensis*	184	IPR003370 (Chromate transporter)
EHA15803.1	*Halomonas sp.*	176	IPR003370 (Chromate transporter)
EHK62369.1	*Halomonas sp.*	185	IPR003370 (Chromate transporter)
WP_021820586.1	*Halomonas sp.*	204	IPR003370 (Chromate transporter)
ACN14075.1	*Desulfobacterium autotrophicum*	161	IPR003370 (Chromate transporter)
ABX60169.1	*Halomonas anticariensis*	184	IPR003370 (Chromate transporter)
EHA15803.1	*Halomonas sp*	176	IPR003370 (Chromate transporter)
EHK62369.1	*Halomonas sp.*	185	IPR003370 (Chromate transporter)
WP_021820586.1	*Halomonas sp*	204	IPR003370 (Chromate transporter)

EJL79983.1	*Variovorax sp.*	189	IPR003370 (Chromate transporter)
AGB25619.1	*Mycobacterium smegmatis*	187	IPR003370 (Chromate transporter)
ABO68618.1	*Geobacillus thermodenitrificans*	193	IPR003370 (Chromate transporter)
EJL20165.1	*Brevibacillus sp*	176	IPR003370 (Chromate transporter)
YP_007770831.1	*Coprococcus catus*	187	IPR003370 (Chromate transporter)
YP_007795406.1	*Coprococcus sp.*	212	IPR003370 (Chromate transporter)
AEN97180.1	*Roseburia hominis*	129	IPR024414 (Uncharacterised protein family Prg I)
AGC67747.1	*Clostridium stercorarium*	197	IPR003370 (Chromate transporter)
EDS13094.1	*Anaerotruncus colihominis*	184	IPR003370 (Chromate transporter)
KCZ68594.1	*Clostridium sporogenes*	178	IPR003370 (Chromate transporter)
AGA69406.1	*Desulfitobacterium dichloroeliminans*	173	IPR003370 (Chromate transporter)
AFQ42925.1	*Desulfosporosinus meridiei*	181	IPR003370 (Chromate transporter)
AEW06911.1	*Sulfobacillus acidophilus*	188	IPR003370 (Chromate transporter)
YP_007829706.1	*Ruminococcus champanellensis*	193	IPR003370 (Chromate transporter)

KCZ68594.1	*Clostridium sporogenes*	178	IPR003370 (Chromate transporter)
AGA69406.1	*Desulfitobacterium dichloroeliminans*	173	IPR003370 (Chromate transporter)
AFQ42925.1	*Desulfosporosinus meridiei*	181	IPR003370 (Chromate transporter)
AEW06911.1	*Sulfobacillus acidophilus*	188	IPR003370 (Chromate transporter)
EDN02096.1	*Pseudoflavonifractor capillosus*	198	IPR003370 (Chromate transporter)
YP_007834179.1	*Megamonas hypermegale*	138	IPR003370 (Chromate transporter)
NP_326559.1	*Mycoplasma pulmonis*	201	IPR003370 (Chromate transporter)
ACM22927.1	*Thermotoga neapolitana*	174	IPR003370 (Chromate transporter)
AAX16961.1	*Borrelia hermsii*	194	IPR003370 (Chromate transporter)
AAU07303.1	*Borrelia garinii*	197	IPR003370 (Chromate transporter)
AEL18626.1	*Borrelia bissettii*	195	IPR003370 (Chromate transporter)
AFG38607.1	*Spirochaeta africana*	203	IPR003370 (Chromate transporter)
WP_008134472.1	*Bradyrhizobium sp.*	176	IPR003370 (Chromate transporter)
WP_009491422.1	*Microvirga lotononidis*	185	IPR003370 (Chromate transporter)

EKN01360.1	*Acidocella sp.*	173	IPR003370 (Chromate transporter)
WP_010630089.1	*Halomonas sp.*	176	IPR003370 (Chromate transporter)
CDG52120.1	*Halomonas sp.*	176	IPR003370 (Chromate transporter)
YP_003624246.1	*Thiomonas arsenitoxydans*	177	IPR003370 (Chromate transporter)
EIF35615.1	*Burkholderia sp.*	215	IPR003370 (Chromate transporter)
AEK61791.1	*Collimonas fungivorans*	176	IPR003370 (Chromate transporter)
YP_007552597.1	*Azoarcus sp.*	182	IPR003370 (Chromate transporter)
ENO98478.1	*Thauera phenylacetica*	191	IPR003370 (Chromate transporter)
ETT35787.1	*Paenibacillus sp.*	198	IPR003370 (Chromate transporter)
EGA89899.1	*Planococcus donghaensis*	177	IPR003370 (Chromate transporter)
WP_018132767.1	*Alicyclobacillus pohliae*	176	IPR003370 (Chromate transporter)
ERK29329.1	*Clostridium intestinale*	189	IPR003370 (Chromate transporter)
EOR27970.1	*Clostridium sartagoforme*	178	IPR003370 (Chromate transporter)
CCY81219.1	*Clostridium sp*	189	IPR003370 (Chromate transporter)

YP_007797951.1	*[Eubacterium] cylindroides*	186	IPR003370 (Chromate transporter)
EGQ15609.1	*Prevotella dentalis*	175	IPR003370 (Chromate transporter)
ADA67097.1	*Thermotoga naphthophila*	166	IPR003370 (Chromate transporter)
ACB09353.1	*Thermotoga sp.*	166	IPR003370 (Chromate transporter)
ABQ47118.1	*Thermotoga petrophila*	171	IPR003370 (Chromate transporter)
YP_006347049.1	*Mesotoga prima*	177	IPR003370 (Chromate transporter)
WP_007691670.1	*Rhizobium sp.*	134	IPR003370 (Chromate transporter)
EHR05009.1	*Bradyrhizobium sp.*	205	IPR003370 (Chromate transporter)
EIG61708.1	*Bradyrhizobium sp.*	176	IPR003370 (Chromate transporter)
ABB32950.1	*Geobacter metallireducens*	184	IPR003370 (Chromate transporter)
ACH37741.1	*Geobacter bemidjiensis*	174	IPR003370 (Chromate transporter)
WP_021806463.1	*Serratia fonticola*	173	IPR003370 (Chromate transporter)
AFE58234.1	*Rahnella aquatilis*	194	IPR003370 (Chromate transporter)
AEX51981.1	*Rahnella aquatilis*	195	IPR003370 (Chromate transporter)

WP_019808580.1	*Osedax symbiont*	181	IPR003370 (Chromate transporter)
WP_016916520.1	*Halomonas stevensii*	182	IPR003370 (Chromate transporter)
WP_023495244.1	*Methyloglobulus morosus*	147	IPR003370 (Chromate transporter)
AFK63870.1	*Advenella kashmirensis*	177	IPR003370 (Chromate transporter)

V. 3D structure prediction of chromate ion transporter protein

Prediction of three dimensional structure of chromate ion transporter of *Burkholderia pseudomallei*

Preparation of ALI file of target protein (AHK64337)

```
>P1;AHK64337
sequence:AHK64337:::::::0.00: 0.00
MTIASVEAACCGERESLWALFKTVTGVSAVSWGGLAMMAQLERHYVEHERRIDPLSFADLVALAW
MMPGP
VGCNVAVQVGHALRGRAGAWIAGVASVLPFSAAMTVFAIFYQTPLVRSLASPVLLHHFAMVLAAL
IGLTW
FRQVRALVHAPLERVIAALATALLALAHNPAAFVAILAAAFAVGWLASGRKQGEALRLALPAREW
RLLAS
LALLIALFALPLPNEYESSLLWPRLAGAGMTLFGGGFSALPVLKSLFVTRSTGITEQDFMLAFTL
SPVSP
GPLLNVVPFLGYLEDGWRGALLSTVALFVPSGCLVIFARRHVERLKRHPRFASGMRVLRAATTAF
LAIAA
VRLVAKTPAEPMYWATGVIAWLCLARFKVPVYALYGAVAAACGGWLILAAHG*
```

The target protein *Burkholderia pseudomallei* and the sequences are shown in Ali file.

Find out the potential template

```
     Weighted pair-group average clustering based on a distance
matrix:



.---------------------------------------------------------- 2C5QA
@1.7    98.0000
        |
        |                                                        .-
-- 2JITA @3.1     7.0000
        |                                                        |
        |                                                     .----
-- 4LRMA @3.5    12.0000
        |                                                    |
.-------------------------------------------------------------
4G5JA @2.8

      +----+----+----+----+----+----+----+----+----+----+----+---
-+
   101.6400   85.2600   68.8800   52.5000   36.1200   19.7400
3.3600
         93.4500   77.0700   60.6900   44.3100   27.9300
11.5500


Total CPU time [seconds]                                    :
2.29
```

Fig. 19. Comparison of the templates

In this process we compared the templates of the target protein which are homologos with the target protein.

Table-5. Homologos protein of (*Burkholderia pseudomallei*) target protein

Protein name	**Protein ID**	**Protein length**	**Organism name**	**PDB ID**	**Organism name**	**Protein length**
Chromate transporter family protein	AHK64337	402	*Burkholderia pseudomallei*	2C5Q	*Saccaromyces cerevisiae*	240

By which we got resolution value of each of the templates from which we had choosen the template which have lowest resolution value and it should be considered as the best template to predict the 3D structure. The lowest resolution value template was 2C5Q Of *Saccaromyces cerevisiae.*

Target template alignment

```
>P1;2C5QA
structureX:2C5Q.pdb:   2 :A:+232 :A:MOL_ID  1; MOLECULE  RRAA-
LIKE PROTEIN YER010C; CHAIN  A, B, C, D, E, F; SYNONYM
HYPOTHETICAL 25.6 KDA PROTEIN IN NTF2-SRP1  INTERGENIC REGION;
ENGINEERED  YES:MOL_ID  1; ORGANISM_SCIENTIFIC  SACCHAROMYCES
CEREVISIAE; ORGANISM_COMMON  BAKER'S YEAST; ORGANISM_TAXID  4932;
EXPRESSION_SYSTEM  ESCHERICHIA COLI; EXPRESSION_SYSTEM_TAXID
562; EXPRESSION_SYSTEM_STRAIN  GOLD(DE3);
EXPRESSION_SYSTEM_VECTOR  PET9: 1.70: 0.16
SDLQKLQRFSTCDISDGLLNVYNIPT-------GG----------Y-----------
FPNLTAIS---P-PQ--N
--------------SSIVGTAYTVLFAPI--------DDP--R----PAV---------NYID--
---SV-----
PPNSILVLALEPHLQSQFH-P--FIKITQAMY--GGLMSTRAQ--------------YLKS--
NGTVVFGRIRDV
DEHRTLNHPVFA------YGVG-SCAP--KA--VVKAVG-TN------VQLKILTSDGVTQTIP-
---------G
DYIAGDN----NGIVRI------------PVQETDIS--KLVTYIEKSIEVDLLVSEDIK-----
-NGI------
-----------PAKQAQNDRRSVLKKYI*

>P1;AHK64337
sequence:AHK64337:     : :     : ::: 0.00: 0.00
MTIASVE-
AACCGERESLWALFKTVTGVSAVSWGGLAMMAQLERHYVEHERRIDPLSFADLVALAWMMPGPVG
CN
VAVQVGHALRGRAGAWIAGVASVLPFSAAMTVFAIFYQTPLVRSLASPVLLHHFAMVLAALIGLT
WFRQVRALVH
APLERVIAALATALLALAHNPAAFVAILAAAFAVGWLASGRKQGEALRLALPAREWRLLASLALL
IALFALPLPN
EYESSLLWPRLAGAGMTLFGGGFSALPVLKSLFVTRSTGITEQDFMLAFTLSPVSPGPLLNVVPF
LGYLEDGWRG
ALLSTVALFVPSGCLVIFARRHVERLKRHPRFASGMRVLRAATTAFLAIAAVRLVAKTPAEPMYW
ATGVIAWLCL
ARFKVPVYALYGAVAAACGGWLILAAHG*
```

Fig. 20. Alignment of target protein with the template

We got the alignment of target protein with the template (AHK64337-2C5QA) as an Ali file which helps us to do the model a single protein (target protein).

Model prediction

```
>> Summary of successfully produced models:
Filename                          molpdf     DOPE score    GA341
score
----------------------------------------------------------------
-----
AHK64337.B99990001.pdb         4068.10059   -37561.02734
0.22594
AHK64337.B99990002.pdb         3665.01172   -38285.99609
0.03954
AHK64337.B99990003.pdb         4061.22534   -39527.17188
0.03564
AHK64337.B99990004.pdb         3411.40186   -37979.36328
0.06694
AHK64337.B99990005.pdb         4276.06006   -37889.73047
0.08933

Total CPU time [seconds]                                    :
307.73
```

3D structure of the target protein (AHK64337)

It successfully produced five structures with the DOPE score, GA341 score, molpdf. The model which having the least DOPE score was choosen the AHK64337 .B99990003.pdb .

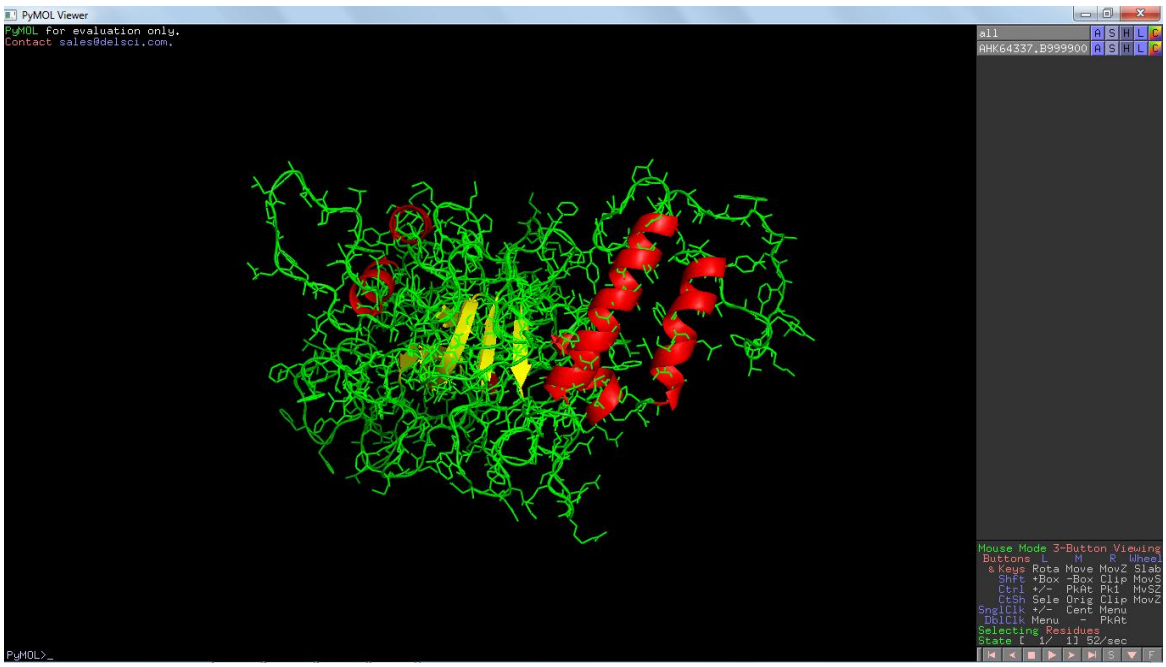
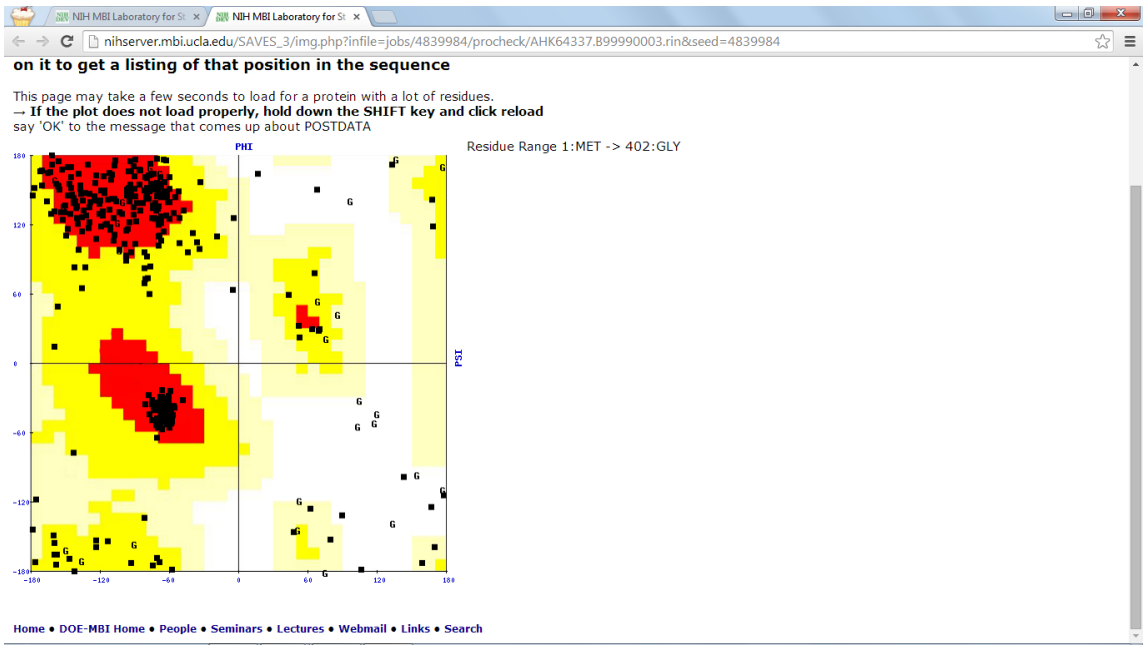


Fig. 21. 3D structure of target protein AHK64337 (*Burkholderia pseudomallei*)

This structures having six alpha helixes, six beta sheets and more random coils.

Model evaluation



I

Fig. 22. Ramchandran Plot of the target protein (*Burkholderia pseudomallei*)

The above figure showing Ramchandran Plot of the target protein (*Burkholderia pseudomallei*). In this it shows the total graphical representation of our target protein which containing the data of phi and psi angle also.

Fig. 23. Graphical representation of the model by verify 3D

It is the 3D plot of the target protein (*Burkholderia pseudomallei)* done in save server by verify 3D.

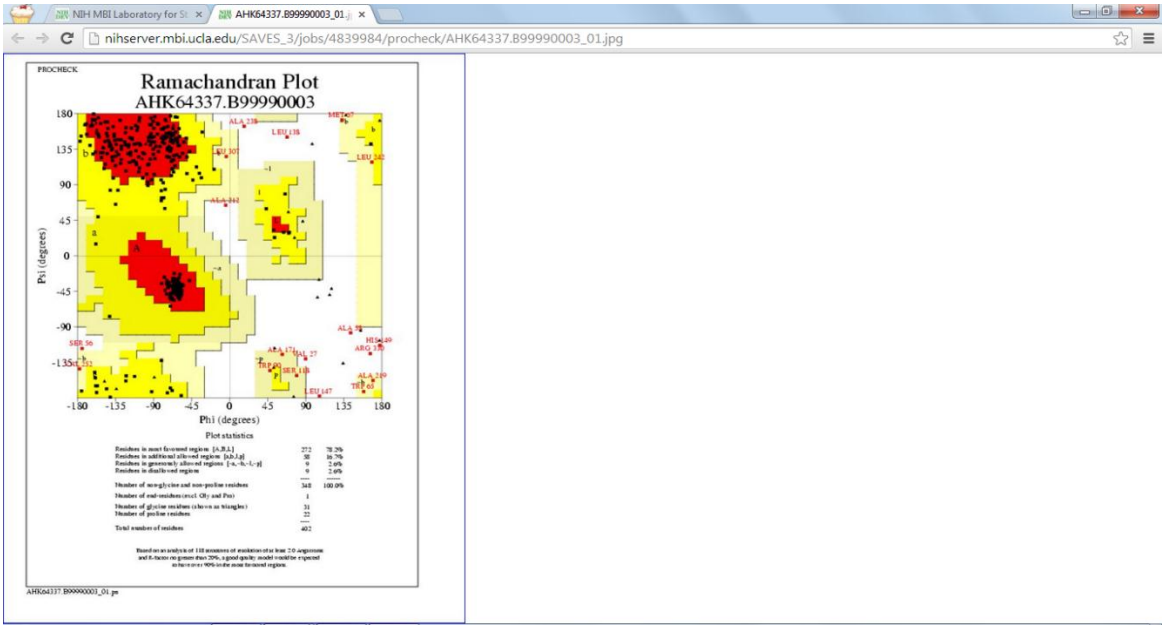


Fig. 24. The Ramchandran Plot in .jpg format

The Ramchandran Plot in .jpgformatdone by procheck in save srever. It was observed that 78% of the residues were in most favorable, in additional allowed region is 16.7%, residues and generally allowed region is 2.6%, and residues in disallowed region is 2.6% of total number of residues are 402.

5. SUMMARY

ChrA is a membrane protein that confers resistance to the toxic ion chromate through the energy-dependent chromate efflux from the cytoplasm. In the protein databases, ChrA is a member of the chromate ion transporter (CHR) superfamily, composed of at least several dozens of members, distributed in the three domains of life.

358 chrA proteins were retrieved from NCBI protein database comprising different organisms(bacteria and fungus) out of which 237 are long chain chromate ion transporter(LCHR)(236-bacterial, 1-fungal) and rest 121 are short chain chromate ion transporter (SCHR) (all bacterial).

The divergence studies among LCHR proteins and SCHR proteins were carried out. It was observed that in LCHR 237proteins were grouped into 4 subgroups and in case of SCHR 121 proteins were grouped in to 6 subgroups. There were 25 conserved sites and 638 variable sites in LCHR proteins .There were no conserved sites but the variable sites were 293 in SCHR proteins.

Secondary structure (Alpha helix, extended strand, random coil) were predicted of the both LCHR and SCHR proteins in %. The % of alpha helix and random coil of chromate ion transporter proteins were more than the extended strand in both LCHR & SCHR proteins. Functional domains were predicted of all LCHR proteins.

LCHR proteins contained chromate transporter domains (IPR014047 &IPR003370) and other domain like IPR006187. Most of the LCHR proteins contained both chromate transporter domains IPR014047 and IPR003370. One LCHR protein of fungi that was *Coprinopsis cinerea* having amino acid length 546 and another LCHR

protein of one organism that was *Kyrpidia tusciae* having amino acid length 252 contained only one chromate transporter domain IPR003370. But one LCHR protein of one organism that was *Arthrobacter sp.* having amino acid length 277 contained one chromate transporter domain IPR003370 and another claudin domain IPR006187. Functional domains were predicted of all SCHR proteins. SCHR proteins contained chromate transporter domain (IPR003370) and some other domains like IPR011006, IPR001789, IPR011991, IPR016032, IPR000792 and IPR024414. Most of the SCHR proteins contained chromate transporter domain IPR003370. One SCHR protein of one organism that was *Roseburia hominis* having amino acid length 129 contained uncharacterized domain IPR024414. But one SCHR protein of one organism that was *Corynebacterium diphtheriae* having amino acid length 199 contained five types of other domains IPR011006, IPR001789, IPR011991, IPR016032 and IPR000792.

The 3D structure of the protein chrA of *Burkholderia pseudomallei (*AHK64337) was predicted using modeler by taking the suitable template of 2C5QA (*Saccaromyces cerevisiae*). The structural study implies that protein having more random coils,6 alpha helixes,6 beta sheets. The present study will help to understand the mechanism of chromate resistance of different micro-organisms.

REFERENCES

1. Vincent JB. 2000. The biochemistry of chromium. *J Nutr.* **130**; 715-718
2. Diane M. Stearns. 2000. Is chromium a trace essential metal? *Biofactors.* **11**; 149-162
3. Francois Baruthio. 1992. Toxic effects of chromium and its compounds. *Biological Trace Element Research.* **32**; 145-153
4. Mohanty M & Patra HK. 2011. Attenuation of chromium toxicity by bioremediation technology. *Rev Environ Contam Toxicol*. **210**; 1-34
5. Ramírez Díaz MI, Díaz Pérez C, Vargas E, Riveros Rosas H, Campos García J & Cervantes C. 2008. Mechanisms of bacterial resistance to chromium compounds. *Biometals.* **21**; 321-332
6. Díaz Pérez C, Cervantes C, Campos García J, Julián Sánchez A & Riveros Rosas H. 2007. Phylogenetic analysis of the chromate ion transporter (CHR) superfamily. *FEBS J* **274**; 6215-6227
7. Rivera SL, Vargas E, Ramírez Díaz MI, Campos García J & Cervantes C. 2008. Genes related to chromate resistance by *Pseudomonas aeruginosa* PAO1. *Antonie Van Leeuwenhoek.* **94**; 299-305
8. Aguilera S, Aguilar ME, Chávez MP, López Meza JE, Pedraza Reyes M, Campos García J & Cervantes C. 2004. Essential residues in the chromate transporter ChrA of *Pseudomonas aeruginosa*. *FEMS Microbiol Lett.* **232**; 107-112
9. Mass MJ & Lane BP. 1976. Effect of chromates on ciliated cells of rat tracheal epithelium. *PubMArch Environ Health.* **31**; 96-100
10. Pritchard DE, Ceryak S, Ramsey KE, O Brien TJ, Ha L, Fornsaglio JL, Stephan DA & Patierno SR. 2005. Resistance to apoptosis, increased growth potential, and altered gene expression in cells that survived genotoxic hexavalent chromium [Cr(VI)] exposure. *Mol Cell Biochem.* **279**; 169-181

11. Peterson Roth E, Reynolds M, Quievryn G & Zhitkovich A. 2005. Mismatch repair proteins are activators of toxic responses to chromium-DNA damage. *Mol Cell Biol.* **25**; 3596-3607
12. Manning FC, Xu J & Patierno SR. 1992. Transcriptional inhibition by carcinogenic chromate relationship to DNA damage. *Mol Carcinog.* **6**; 270-279
13. Jiménez Mejía R, Campos García J & Cervantes C. 2006. Membrane topology of the chromate transporter ChrA of *Pseudomonas aeruginosa*. *FEMS Microbiol Lett.* **262**; 178-184
14. Holland SL & Avery SV. 2011. Chromate toxicity and the role of sulfur. *Metallomics.* **3**; 1119-1123
15. Martínez Valencia R, Reyes Cortés G, Ramírez Díaz MI, Riveros Rosas H & Cervantes C. 2012. Antiparallel membrane topology of paired short-chain chromate transport proteinsin Bacillus subtilis. *FEMS Microbiol Lett.* **336**; 113-121