

Sl. No.	Practical title	Page no	Signature (Internal Supervisor)
1	1A) Chemical Data Retrieval using DrugBank. 1B) Chemical Data Retrieval using PubChem. 1C) Chemical Data Retrieval using ZINC	04	
2	2A) Element analysis in ChemSkech. 2B) Introduction to ChemSketch 2C) Marvin Sketch	76	
3	3A) Active site prediction using CASTp 3B) Active site prediction using PDBsum	101	
4	Molecular Docking Using Argus Lab 4A) Argus Docking 4B) GA Dock	118	
5	Comparative Molecular Docking Using Argus Lab	135	
6	Virtual Screening using iGemDock	143	
7	Protein Docking using HEX	145	
8	8A) Structural Based pharmacophores using PharmMapper . 8B) Ligand Based Pharmacophore using PharmaGist.	161	
9	Toxicity Analysis using SwissADME.	165	
10	10 A)Gene expression data analysis using GEO database and GEO2R tool. 10 B)ArrayExpress	183	
11	Molecular Dynamics Using HyperChem 11A) Study of small molecule molecular dynamics using HyperChem 11B) Study of Peptide wild and mutant	214	

Practical no: 1A)

Aim: Chemical Data Retrieval using DrugBank

Theory: DrugBank is a pharmaceutical knowledge base that is enabling major advances across the data-driven medicine industry. The knowledge base consists of proprietary authored content describing clinical level information about drugs such as side effects and drug interactions, as well as molecular level data such as chemical structures and what proteins a drug interacts with. DrugBank offers a suite of products powered by the DrugBank Platform and has customers located around the world crossing multiple industries including precision medicine, electronic health records, drug development and regulatory agencies. DrugBank also provides DrugBank Online as a free-to-access resource for academic research and is used by millions of pharmacists, pharmacologists, health professionals and pharmaceutical researchers every year.

Methodology:

My disease: Glioma

My drug:- Gliadel, Carmustine

Step 1: Go to the homepage of DrugBank and type the name in search bar:

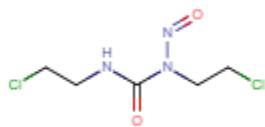
The screenshot shows the DrugBank Online website. At the top, there is a navigation bar with links for DRUGBANK Online, Browse, COVID-19, Search, Interaction Checker, Downloads, Solutions, and About. A banner at the top says "Search over 500,000 drugs & drug products on DrugBank Online". Below the banner is a search bar containing the text "gliadel". Underneath the search bar are four colored buttons: blue (Drugs), red (Targets), green (Pathways), and orange (indications). The main content area below the search bar has a heading "Drug knowledge for all your needs".

Step 2:

The screenshot shows the DrugBank product page for Carmustine. The URL in the address bar is https://go.drugbank.com/drugs/DB00262. The page title is "Carmustine". On the left, there is a sidebar with a navigation menu: Identification, Pharmacology, Interactions, Products, Categories, Chemical Identifiers, References, Clinical Trials, Pharmacoeconomics, Properties, and Spectra. The main content area displays the following details for Carmustine:
Summary: Carmustine is an alkylating agent used in the treatment of various malignancies, including brain tumours and multiple myeloma, among others.
Brand Names: Bioms, Gliadel
Generic Name: Carmustine
Drugbank Accession Number: DB00262
Background: A cell-cycle phase nonspecific alkylating antineoplastic agent. It is used in the treatment of brain tumors and various other malignant neoplasms. (From Martindale, The Extra Pharmacopoeia, 30th ed, p462) This substance may reasonably be anticipated to be a carcinogen according to the Fourth Annual Report on Carcinogens (NTP 85-002, 1985). (From Merck Index, 11th ed)
At the bottom of the page, there are links for "Watch" and "Star".

Title: Carmustine

a) Drug structure:-



- Chemical formula:- $C_5H_9Cl_2N_3O_2$
- InChI Key:- DLGOEMSEDOSKAD-UHFFFAOYSA-N
- Smile:-
ClCCNC(=O)N(CC(Cl)Cl)N=O

• Properties: (Experimental and calculated both):-

PROPERTY	VALUE	SOURCE
melting point (°C)	31 °C	PhysProp
water solubility	4000 mg/L (at 25 °C)	MERCK (1989)
logP	1.53	HANSCH,C ET AL. (1995)

PROPERTY	VALUE	SOURCE
Water Solubility	1.53 mg/mL	ALOGPS
logP	1.24	ALOGPS
logP	1.02	ChemAxon
logS	-2.2	ALOGPS
pKa (Strongest Acidic)	11.96	ChemAxon
pKa (Strongest Basic)	-5.3	ChemAxon
Physiological Charge	0	ChemAxon

b) Pharmacokinetic/ADME:-

The screenshot shows the DrugBank Online interface. The left sidebar has a dark theme with pink highlights for 'Properties' and 'Predicted Properties'. The main content area displays a table titled 'Predicted ADMET Features' with columns for PROPERTY, VALUE, and PROBABILITY. The table lists various pharmacokinetic and metabolic properties and their predicted values and probabilities.

PREDICTED ADMET FEATURES	PROPERTY	VALUE	PROBABILITY
	Human Intestinal Absorption	+	1.0
	Blood Brain Barrier	+	0.9533
	Caco-2 permeable	-	0.5621
	P-glycoprotein substrate	Non-substrate	0.7552
	P-glycoprotein inhibitor I	Non-inhibitor	0.797
	P-glycoprotein inhibitor II	Non-inhibitor	0.8778
	Renal organic cation transporter	Non-inhibitor	0.8177
	CYP450 2C9 substrate	Non-substrate	0.7656
	CYP450 2D6 substrate	Non-substrate	0.8491
	CYP450 3A4 substrate	Non-substrate	0.672
	CYP450 1A2 substrate	Non-inhibitor	0.9045
	CYP450 2C9 inhibitor	Non-inhibitor	0.907
	CYP450 2D6 inhibitor	Non-inhibitor	0.9231
	CYP450 2C19 inhibitor	Non-inhibitor	0.9025

c) Pharmacodynamic:-

Carmustine is one of the nitrosoureas indicated as palliative therapy as a single agent or in established combination therapy with other approved chemotherapeutic agents in treatment of brain tumors, multiple myeloma, Hodgkin's disease, and non-Hodgkin's lymphomas. Although it is generally agreed that carmustine alkylates DNA and RNA, it is not cross resistant with other alkylators. As with other nitrosoureas, it may also inhibit several key enzymatic processes by carbamoylation of amino acids in proteins.

d) Target information anyone:-

- Kind:- Nucleotide
- Organism:- Humans
- Pharmacological action:- Yes
- Actions:-

DNA is the molecule of heredity, as it is responsible for the genetic propagation of most inherited traits. It is a polynucleic acid that carries genetic information on cell growth, division, and function. DNA consists of two long strands of nucleotides twisted into a double helix and held together by hydrogen bonds. The sequence of nucleotides determines hereditary characteristics. Each strand serves as the template for subsequent DNA replication and as a template for mRNA production, leading to protein synthesis via ribosomes.

General Function:- Wnt-protein binding

The screenshot shows a web browser window with multiple tabs open at the top, including 'Dashboard', 'Course: Pharm...', 'Carmustine: U...', 'Research-Pres...', 'List of 4 Malign...', 'is drug gliadel...', 'Tum Hi...', and a blank tab. The main content area is the 'DRUGBANK Online' interface. The navigation bar includes 'DRUGBANK Online', 'Browse', 'COVID-19', 'Search', and 'About'. A search bar at the top right contains the text 'function'. Below the search bar, there's a dropdown menu set to 'Drugs' and a magnifying glass icon. A banner at the top says 'New Webinar + Q&A: Learn how DrugBank improves and accelerates early-stage drug discoveries' with a 'Register Now' button. On the left, a sidebar lists categories: Identification, Pharmacology, Interactions, Products, Categories, Chemical Identifiers, References, Clinical Trials, Pharmacoeconomics, Properties, Spectra, and Targets (3). The main panel displays the details for '2. Glutathione reductase, mitochondrial'. The table includes the following data:

Kind	Protein	General Function	Nadp binding
Organism	Humans	Specific Function	Maintains high levels of reduced glutathione in the cytosol.
Pharmacological action	Yes	Gene Name	GSR
Actions	Inhibitor	Uniprot ID	P00390
		Uniprot Name	Glutathione reductase, mitochondrial
		Molecular Weight	56256.565 Da

Below the table, under 'References', are two articles:

1. Akella SS, Harris C: Pyridine nucleotide flux and glutathione oxidation in the cultured rat conceptus. *Reprod Toxicol*. 1999 May-Jun;13(3):203-13. [Article]
2. Kirsch JD, Yi AK, Spitz DR, Krieg AM: Accumulation of glutathione disulfide mediates NF-kappaB activation during immune stimulation with CpG DNA. *Antisense Nucleic Acid Drug Dev*. 2002 Oct;12(5):327-40. [Article]

e) **Human Toxicity any two:-** The oral LD₅₀s in rat and mouse are 20 mg/kg and 45 mg/kg, respectively. Side effects include leukopenia, thrombocytopenia, nausea. Toxic effects include pulmonary fibrosis (20-0%) and bone marrow toxicity.

f) Patent information anyone:-

Not Available

g) similar structure :

Screenshot of the DRUGBANK Online search results page for a similar structure to DB00262.

The search results show two entries:

Score	Chemical Structure	Drug Name / CAS Number	Chemical Formula / Mono mass
1.0		Carmustine 154-93-8 approved investigational	C ₅ H ₉ Cl ₂ N ₃ O ₂ Mono mass: 213.007181961
0.724		2-chloroethyl-3-sarcosinamide-1-nitrosourea 81965-43-7 investigational	C ₆ H ₁₁ ClN ₄ O ₃ Mono mass: 222.0519679

Below the results, there is a search bar and system status information.

h) References any :-

FDA Approved Drug Products: BiCNU (carmustine) for intravenous injection [Link]

New Webinar + Q&A: Learn how DrugBank improves and accelerates early-stage drug discoveries [Register Now](#)

General References 1. FDA Approved Drug Products: BiCNU (carmustine) for intravenous injection [Link]

External Links Human Metabolome Database HMDB0014407
KEGG Drug D00254
KEGG Compound C06873
PubChem Compound 2578
PubChem Substance 46506980
ChemSpider 2480
BindingDB 50015950
RxNav 2105
ChEBI 3423
ChEMBL CHEMBL513
ZINC ZINC000003830387
Therapeutic Targets Database DAP000041

Step 1: Go to the homepage of DrugBank and type the name in search bar:

Announcing DrugBank Affiliate Partnerships for biotech and pharma consultants! [Learn More](#)

Drugs glioma

Small Molecule Drugs

Approved Nutraceutical Elicit

Investigational Withdrawn Experimental

Market Availability U.S. Canada E.U.

Displaying drugs 1 - 25 of 2705 in total

Step 2: For Glioma we got 24 hits:

Enjoying DrugBank? Help us test our new medication search! [Take me there](#)

Displaying all 24 drugs

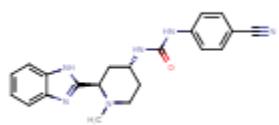
pembrolizumab

Anti-PD-1 antibody which is a fully humanized monoclonal antibody that blocks the PD-1 receptor. It is a bispecific antibody that displays immune activity, it inhibits cell proliferation, invasion and migration and induces apoptosis in cancer cells. It can inhibit protein phosphorylation, dephosphorylation, plasma membrane...
Detailed Description ... It inhibits cell proliferation, invasion and migration and induces apoptosis in glioma cells...

Gliangi

Gliangi, also known as MRGPRE, is a small molecule helping regulating blood vessels under the group of the...
Detailed Description ... It inhibits cell proliferation, invasion and migration and induces apoptosis in glioma cells...

i) Drug structure:-



Glasdegib

i) Chemical formula:- C₂₁H₂₂N₆O

- InChI Key:- SFNSLLSYNZWZQG-VQIMIIIECSA-N
- Smile:-
CN1CC[C@H](C[C@H]1C1=NC2=CC=CC=C2N1)NC(=O)NC1=CC=C(C=C1)C#N
- Physiochemical properties (Experimental and calculated both):-

Properties			
State			
Experimental Properties	PROPERTY	VALUE	SOURCE
	melting point (IC)	-214 °C	"M202"
	boiling point (IC)	Decomposes at 214 °C	"M202"
	water solubility	0.02 mg/mL (in the form of di-HCl monohydrate salt)	Munchhof M., et al. (2011) ACS Med Chem Lett.
	logP	2.28	Munchhof M., et al. (2011) ACS Med Chem Lett.
	Caco-2 permeability	0.0000596	Munchhof M., et al. (2011) ACS Med Chem Lett.
	pKa	6	Munchhof M., et al. (2011) ACS Med Chem Lett.

Predicted Properties	PROPERTY	VALUE	SOURCE
	Water Solubility	0.0469 mg/mL	ALOGPS
	logP	2.48	ALOGPS
	logP	2.28	ChemAxon
	logS	-3.9	ALOGPS
	pKa (Strongest Acidic)	11.39	ChemAxon
	pKa (Strongest Basic)	6.67	ChemAxon
	Physiological Charge	0	ChemAxon
	Hydrogen Acceptor Count	4	ChemAxon
	Hydrogen Donor Count	3	ChemAxon
	Polar Surface Area	96.84 Å ²	ChemAxon
	Rotatable Bond Count	3	ChemAxon
	Refractivity	108.26 m ³ mol ¹	ChemAxon
	Polarizability	40.73 Å ³	ChemAxon
	Number of Rings	4	ChemAxon
	Hydrogen Acceptor Count	4	ChemAxon
	Hydrogen Donor Count	3	ChemAxon
	Polar Surface Area	96.84 Å ²	ChemAxon
	Rotatable Bond Count	3	ChemAxon
	Refractivity	108.26 m ³ mol ¹	ChemAxon
	Polarizability	40.73 Å ³	ChemAxon
	Number of Rings	4	ChemAxon
	Bioavailability	1	ChemAxon
	Rule of Five	Yes	ChemAxon
	Ghose Filter	Yes	ChemAxon
	Veber's Rule	No	ChemAxon
	MDDR-like Rule	No	ChemAxon

Predicted ADMET Features Not Available

k) Pharmacokinetic/ADME:- Not Available

I) Pharmacodynamic:-

In preclinical studies, glasdegib achieved a significant reduction in leukemic stem cell burden in xenograft models and a reduction in cell population expressing leukemic stem cell markers.²

In clinical trials, glasdegib demonstrated a marked downregulation of more than 80% of the expression of glioma-associated transcriptional regulator GL11 in skin. In this same study 8% of the studied individuals with acute myeloid leukemia achieved morphological complete remission while 31% achieved stable disease state.²

The latest clinical trial proved glasdegib to generate an overall survival of 8.3 months which was almost double to what has been observed in patients under low-dose cytarabine treatment. As well, there have been reports of dose-dependent QTc prolongation in patients administered with glasdegib.

m) Target information anyone:-

- Kind:- Protein
- Organism:- Humans
- Pharmacological action:- Yes
- Actions:- Inhibitor
- General Function:- Wnt-protein binding
- Specific Function:- G protein-coupled receptor that probably associates with the patched protein (PTCH) to transduce the hedgehog's proteins signal.
- Gene Name:-SMO
- Uniprot ID :-
[Q99835](#)
- Uniprot Name :- Smoothened homolog
- Molecular Weight :- 86395.95 Da

g) **Human Toxicity any two:-** Glasdegib is not mutagenic in bacterial reverse mutation assays and is not clastogenic in *in vitro* chromosome aberration assays. In fertility studies, glasdegib has the potential to impair reproductive function in males due to the production of testicular changes such as hypospermatogenesis. [Label](#)

Overdose by glasdegib starts at 640 mg/day and shows to present nausea, vomiting, dehydration, fatigue, and dizziness. In case of overdose, symptomatic treatment and ECG monitoring are advised. [Label](#)

The reported oral LD50 in rat of gladegeb administered in triacetin is reported to be of 3000 mg/kg.⁹

n) Patent information anyone:-

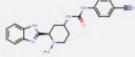
Prices	Not Available		
Patents	Show 10	entries	Search
PATENT NUMBER	PEDIATRIC EXTENSION	APPROVED	EXPIRES (ESTIMATED)
US8431597		2013-04-30	2028-06-29
US8148401		2012-04-03	2031-01-30
US9044748		2019-09-17	2036-04-13

Showing 1 to 3 of 3 entries

PROPERTIES

State	Solid
-------	-------

o) similar structure :

DB11978 Score: 1.0		Glasdegib 1095173-27-5 approved investigational	C ₂₁ H ₂₂ N ₆ O Mono mass: 374.185509352
-----------------------	---	---	--

p) References any :-

Synthesis Reference	Munchhof MJ, Li Q, Shavnya A, et al. Discovery of PF-04449913, a Potent and Orally Bioavailable Inhibitor of Smoothened. ACS Med Chem Lett. 2012;3(2):106-11.
General References	
1.	Munchhof MJ, Li Q, Shavnya A, Borzillo GV, Boyden TL, Jones CS, LaGreca SD, Martinez-Alsina L, Patel N, Pelletier K, Reiter LA, Robbins MD, Tkatchevic GT. Discovery of PF-04449913, a Potent and Orally Bioavailable Inhibitor of Smoothened. ACS Med Chem Lett. 2012;3(2):106-11.
2.	Misami Y, Minami H, Miyamoto T, Woshimoto G, Kobayashi Y, Moriwaki K, Oomori Y, Kobayashi M, Kudo M, Chan G, Weeraratna A, Ono C, Shau MN, Fujii Y, Zheng X, Nace T, Phate I. Phase I study of glasdegib (PF-04449913), an oral smoothened inhibitor, in Japanese patients with select hematologic malignancies. Cancer Sci. 2017 Aug;108(8):1629-1633. doi: 10.1111/cas.13285. Epub 2017 Jun 19. [PubMed]
3.	Irvine DA, Copland M. Targeting hedgehog in hematologic malignancy. Blood. 2012 Mar 8;119(10):2196-204. doi: 10.1182/blood-2011-10-38752. Epub 2012 Jan 5. [PubMed]
4.	Lam JL, Vaz A, Hee B, Liang Y, Yang X, Shan M. Metabolism, excretion and pharmacokinetics of [(14)C]glasdegib (PF-04449913) in healthy volunteers following oral administration. Xenobiotica. 2017 Dec;47(12):1064-1076. doi: 10.3109/00434429.2016.1261307. Epub 2017 Jan 1. [PubMed]
5.	Bethesda (2006). Drugs and Lactation Database. National Library of Medicine.
6.	FDA news. [Link]
7.	NH. [Link]
8.	FDA Approved Drug Products: Daurismo (glasdegib) oral tablets. [Link]
9.	Pfizer Glasdegib product information. [File]

Result: : Hence chemical data retrieval of our drug Carmustine was done using the online platform DrugBank which includes the drug's Pharmacology, Interactions, Clinical Trials, Pharmacoeconomics, physiochemical properties and targets.

Practical No: 1B

Aim : Chemical Data Retrieval using PubChem

Theory : PubChem is an open chemistry database at the National Institutes of Health (NIH). “Open” means that you can put your scientific data in PubChem and that others may use it. Since the launch in 2004, PubChem has become a key chemical information resource for scientists, students, and the general public. Each month our website and programmatic services provide data to several million users worldwide. PubChem mostly contains small molecules, but also larger molecules such as nucleotides, carbohydrates, lipids, peptides, and chemically-modified macromolecules. We collect information on chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, toxicity data, and many others. PubChem records are contributed by hundreds of data sources. Examples include: government agencies, chemical vendors, journal publishers, and more

Methodology:

First go to the home page of PubChem

My disease: Glioma

My drug:- Gliadel, Carmustine

The screenshot shows a Microsoft Edge browser window with the URL <https://pubchem.ncbi.nlm.nih.gov>. The page header includes links for COVID-19 Information, Public health information (CDC), Research information (NIH), SARS-CoV-2 data (NCBI), Prevention and treatment information (HHS), and Español. The NIH National Library of Medicine logo is visible. The main content area features a blue hexagonal background pattern. The title 'Explore Chemistry' is prominently displayed, followed by the subtext 'Quickly find chemical information from authoritative sources'. A search bar contains the query 'glioma'. Below the search bar is a table with three columns: 'Compound' (Glisema), 'Gene' (Glia maturation factor), and 'Taxonomy' (Lama glama). At the bottom of the screen, there is a taskbar with icons for File, Home, Stop, Refresh, Back, Forward, and Stop. A system tray shows the date and time as 27-10-2021 22:11, and a battery icon indicating 24% charge.

Step 2: We got 12 number of hits for glioma:

The screenshot shows the PubChem website interface. In the search bar at the top, the query 'glioma' is entered. Below the search bar, there are tabs for Substances (12), Genes (30), Proteins (21), Taxonomy (1), BioAssays (4,661), Literature (62,858), and Patents (1,263). The 'Substances' tab is selected. The main content area displays 12 results for 'glioma'. One result is highlighted: 'LGI1; Leucine-Rich, Glioma Inactivated 1; S17609; NM_005097'. It includes details such as Substance SID: 160721730, Data Source: Life Technologies, Applied Biosystems, Ambion, External ID: s17609, and Data Source Category: siRNA Reagent Vendors. The deposit date is 2013-02-07 and the last modified date is 2014-03-02. On the right side, there are options to 'Search in Entrez' and 'Push to Entrez'. The Windows taskbar at the bottom shows other open applications like WhatsApp, Mail, and a browser.

The screenshot shows a detailed view of the substance record for LGI1. The URL in the browser is https://pubchem.ncbi.nlm.nih.gov/substance/160721730. The page title is 'LGI1'. The left panel contains a table with the following information:

PubChem SID	160721730
Source	Life Technologies, Applied Biosystems, Ambion
External ID	s17609
Source Category	siRNA Reagent Vendors
Version	4
Status	Live
Dates	Modify: 2014-03-02 Available: 2013-11-24 Deposit: 2013-02-07

A note below the table states: 'Please note that the substance record is presented as provided to PubChem by the source (depositor).'. A link to 'PubChem' is also present. The right panel shows a 'CONTENTS' sidebar with sections: Title and Summary, Identity, Depositor Comments, Biological Test Results, Entrez Crosslinks, and Information Sources. The 'Title and Summary' section is currently active. The Windows taskbar at the bottom shows the same applications as the previous screenshot.

Gliadel

The screenshot shows the PubChem website interface. At the top, there is a navigation bar with links to Dashboard, Course: Pharmacoinformatics, YouTube channel, PubChem, and other sections. Below the navigation bar is the PubChem logo and a menu with About, Blog, Submit, and Contact options. The main header features the text "Explore Chemistry" and "Quickly find chemical information from authoritative sources". A search bar at the top contains the query "gliadel". Below the search bar, the results are displayed in two columns: "Compound" and "Gene". In the "Compound" column, results include Gliadel, Gliadel Wafer, Glimel, Priadel, GLIADIN, and Gliadin peptide CT-1. In the "Gene" column, results include Glial Lazarillo, glial cells missing, glial cells missing homolog 1, glial cells missing homolog 2, Glia maturation factor, and glia maturation factor beta. The background of the page features a blue hexagonal geometric pattern.

gliadel

Compound	Gene
Gliadel	Glial Lazarillo
Gliadel Wafer	glial cells missing
Glimel	glial cells missing homolog 1
Priadel	glial cells missing homolog 2
GLIADIN	Glia maturation factor
Gliadin peptide CT-1	glia maturation factor beta

Type here to search

SEARCH FOR

Gliadel

Treating this as a text search.

COMPOUND BEST MATCH

Carmustine; 154-93-8; 1,3-Bis(2-Chloroethyl)-1-Nitrosourea; BCNU; Carmustin; Gliadel; Nitruman; Carmubris; ...

Compound CID: 2578

MF: C5H9Cl2N2O2 MW: 214.05g/mol

IUPAC Name: 1,3-bis(2-chloroethyl)-1-nitrosourea

Isomeric SMILES: C(CCl)NC(=O)N(CCCl)N=N

InChIKey: DLGOEMSEDOSKAD-UHFFFAOYSA-N

InChI: InChI=1S/C5H9Cl2N3O2/c6-1-3-8-5(11)10(9-12)4-2-7/h1-4H2,(H,8,11)

Create Date: 2005-03-25

[Summary](#) [Similar Structures Search](#) [Related Records](#) [PubMed \(MeSH Keyword\)](#)

Compounds (1) **Substances** (16) **Literature** (144)

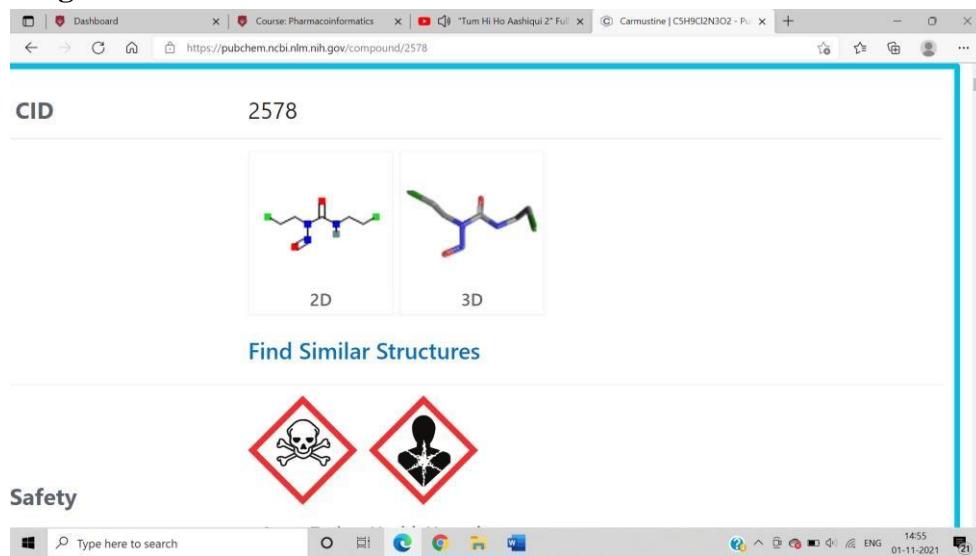
Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary pages is not searched. [Read More...](#)

a) Summary from [Pubchem](#):-

Carmustine (BCNU) is a parenterally administered alkylating agent used alone and in combination with other antineoplastic agents in the treatment of several forms of cancer including leukemias, lymphomas, and breast, testicular, ovarian, gastric and pancreatic cancer. Carmustine therapy is associated with minor transient serum enzyme elevations and has been linked to cases of acute liver injury including cholestatic hepatitis and acute veno-occlusive disease.

b) Title: Carmustine

c) Drug structure:-



- **Molecular Formula** C₅H₉Cl₂N₃O₂
- **Inchi key**:- DLGOEMSEDOSKAD-UHFFFAOYSA-N
- **Smile**:- C(CCl)NC(=O)N(CCCl)N=N

d) Properties:-

- Computed properties:-

The screenshot shows the PubChem website for Carmustine (Compound). The URL is <https://pubchem.ncbi.nlm.nih.gov/compound/2578#section=Chemical-and-Physical-Properties>. The page title is "PubChem Carmustine (Compound)". The main content area is titled "3 Chemical and Physical Properties" and "3.1 Computed Properties". A table lists various properties with their values and sources:

Property Name	Property Value	Reference
Molecular Weight	214.05	Computed by PubChem 2.1 (PubChem release 2021.05.07)
XLogP3	1.5	Computed by XLogP3 3.0 (PubChem release 2021.05.07)
Hydrogen Bond Donor Count	1	Computed by Cactvs 3.4.8.18 (PubChem release 2021.05.07)
Hydrogen Bond Acceptor Count	3	Computed by Cactvs 3.4.8.18 (PubChem release 2021.05.07)
Rotatable Bond Count	4	Computed by Cactvs 3.4.8.18 (PubChem release 2021.05.07)
Exact Mass	213.0071819	Computed by PubChem 2.1 (PubChem release 2021.05.07)
Monoisotopic Mass	213.0071819	Computed by PubChem 2.1 (PubChem release 2021.05.07)
Topological Polar Surface Area	61.8 Å ²	Computed by Cactvs 3.4.8.18 (PubChem release 2021.05.07)
Heavy Atom Count	12	Computed by PubChem
Formal Charge	0	Computed by PubChem
Complexity	156	Computed by Cactvs 3.4.8.18 (PubChem release 2021.05.07)
Isotope Atom Count	0	Computed by PubChem

The right sidebar shows a table of contents with "3 Chemical and Physical Properties" highlighted. The bottom status bar shows the date as 01-11-2021 and the time as 14:57.

- Experimental Properties

The screenshot shows the PubChem website for Carmustine (Compound). The URL is <https://pubchem.ncbi.nlm.nih.gov/compound/2578#section=Experimental-Properties>. The page title is "PubChem Carmustine (Compound)". The main content area is titled "3.2 Experimental Properties" and "3.2.1 Physical Description". It states that 1,3-bis(2-chloroethyl)-1-nitrosourea is an orange-yellow solid (NTP, 1992). It also mentions the National Toxicology Program, Institute of Environmental Health Sciences, National Institutes of Health (NTP), 1992. National Toxicology Program Chemical Repository Database, Research Triangle Park, North Carolina. The "CAMEO Chemicals" section indicates it is a solid. The "Human Metabolome Database (HMDB)" section is also mentioned. The right sidebar shows a table of contents with "3 Chemical and Physical Properties" highlighted. The bottom status bar shows the date as 01-11-2021 and the time as 14:57.

- Pharmacology :-

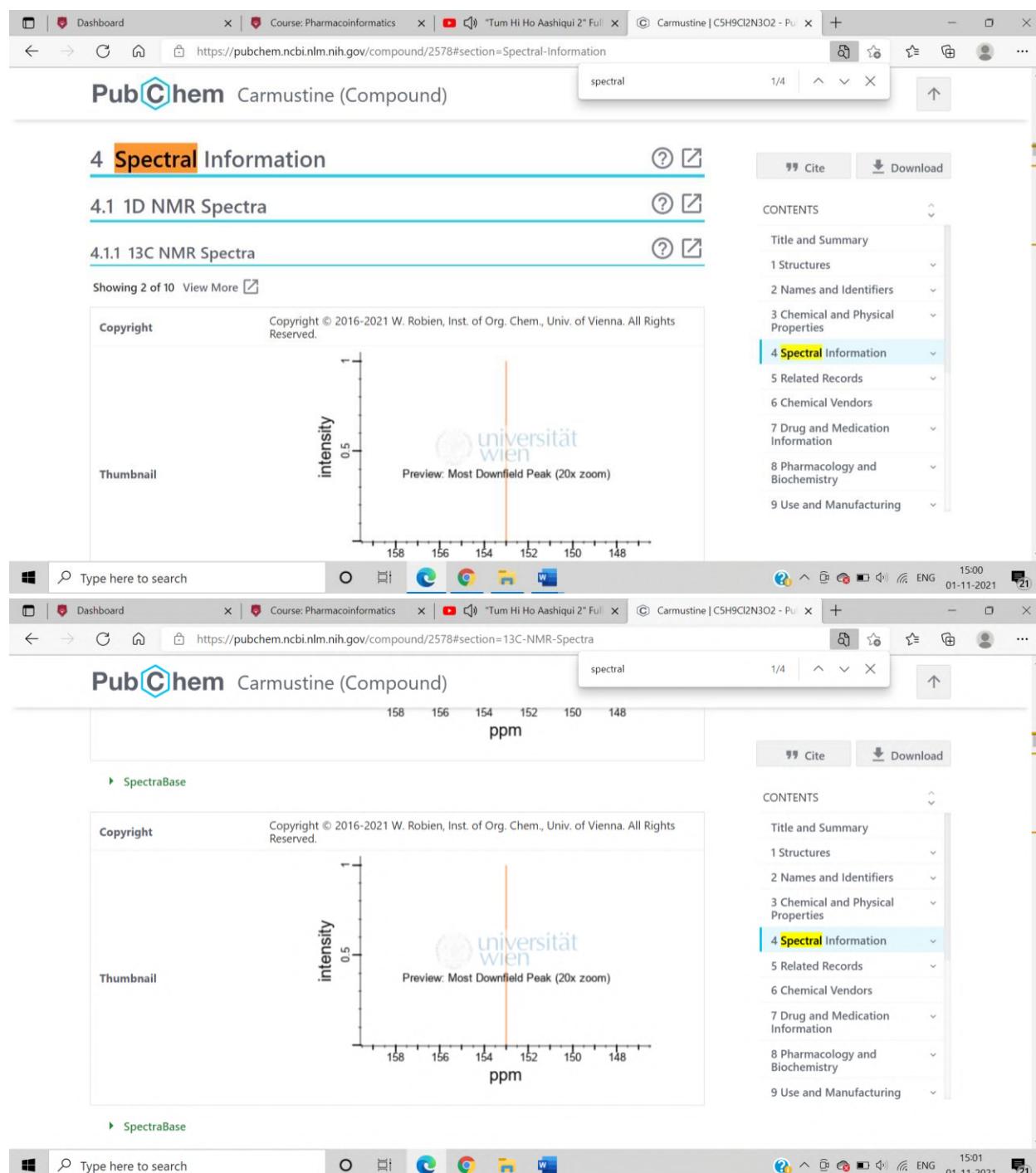
Carmustine is one of the nitrosoureas indicated as palliative therapy as a single agent or in established combination therapy with other approved chemotherapeutic agents in treatment of brain tumors, multiple myeloma, Hodgkin's disease, and non-Hodgkin's lymphomas. Although it is generally agreed that carmustine alkylates DNA and RNA, it is not cross

resistant with other alkylators. As with other nitrosoureas, it may also inhibit several key enzymatic processes by carbamoylation of amino acids in proteins.

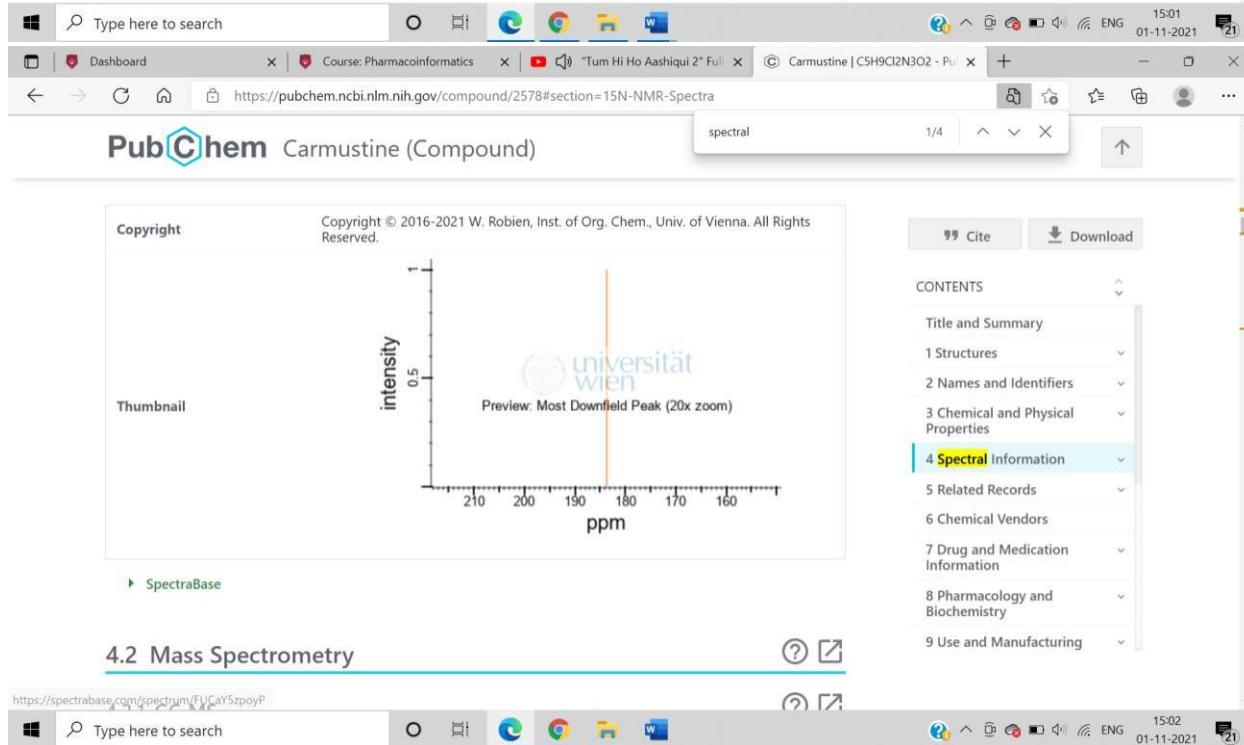
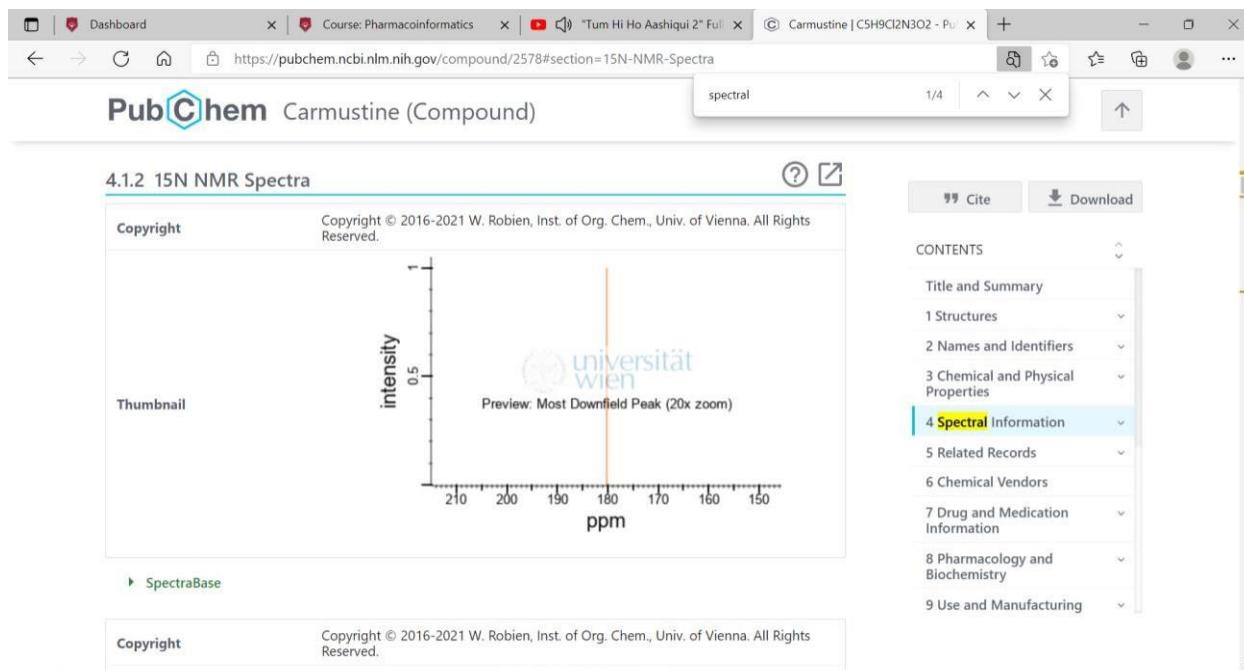
- **Target information anyone:-**
- **Action:-** Carmustine causes cross-links in DNA and RNA, leading to the inhibition of DNA synthesis, RNA production and RNA translation (protein synthesis). Carmustine also binds to and modifies (carbamoylates) [glutathione](#) reductase. This leads to cell death.

I) Spectral data any two:-

→ 1D NMR Spectra



1) 15N NMR Spectra



2)

Mass Spectrometry

The screenshot shows the PubChem compound page for Carmustine (Compound). The URL is <https://pubchem.ncbi.nlm.nih.gov/compound/2578#section=Mass-Spectrometry>. The page displays the following information:

PubChem Carmustine (Compound)

4.2 Mass Spectrometry

4.2.1 GC-MS

Showing 2 of 3 View More

NIST Number	248168
Library	Main library
Total Peaks	88
m/z Top Peak	108
m/z 2nd Highest	63
m/z 3rd Highest	106

Thumbnail Carmustine
El mass spectrum, top peaks displayed

abundance

15:02 01-11-2021

CONTENTS

- Title and Summary
- 1 Structures
- 2 Names and Identifiers
- 3 Chemical and Physical Properties
- 4 Spectral Information** (selected)
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing

PubChem Carmustine (Compound)

NIST Mass Spectrometry Data Center

NIST Number	131703
Library	Replicate library
Total Peaks	66
m/z Top Peak	63
m/z 2nd Highest	108
m/z 3rd Highest	42

Thumbnail Carmustine
El mass spectrum, top peaks displayed

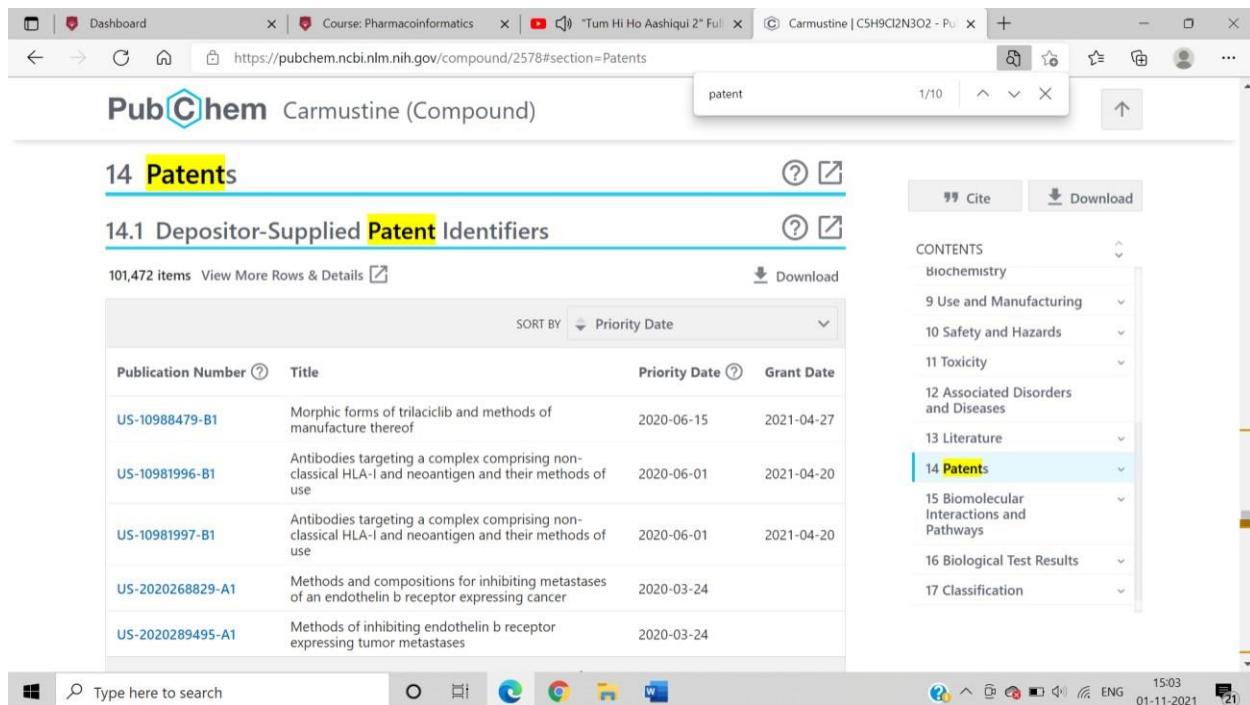
abundance

15:02 01-11-2021

CONTENTS

- Title and Summary
- 1 Structures
- 2 Names and Identifiers
- 3 Chemical and Physical Properties
- 4 Spectral Information** (selected)
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing

j) Patent information anyone:-
Depositor-Supplied Patent Identifiers



The screenshot shows the PubChem website for Carmustine (Compound). The search bar at the top contains the word "patent". Below the search bar, there are two tabs: "Cite" and "Download". A sidebar on the right lists various sections: CONTENTS, Biochemistry, 9 Use and Manufacturing, 10 Safety and Hazards, 11 Toxicity, 12 Associated Disorders and Diseases, 13 Literature, 14 Patents (which is selected and highlighted in blue), 15 Biomolecular Interactions and Pathways, 16 Biological Test Results, and 17 Classification.

14 Patents

14.1 Depositor-Supplied Patent Identifiers

101,472 items View More Rows & Details

Publication Number	Title	Priority Date	Grant Date
US-10988479-B1	Morphic forms of trilaciclib and methods of manufacture thereof	2020-06-15	2021-04-27
US-10981996-B1	Antibodies targeting a complex comprising non-classical HLA-I and neoantigen and their methods of use	2020-06-01	2021-04-20
US-10981997-B1	Antibodies targeting a complex comprising non-classical HLA-I and neoantigen and their methods of use	2020-06-01	2021-04-20
US-2020268829-A1	Methods and compositions for inhibiting metastases of an endothelin b receptor expressing cancer	2020-03-24	
US-2020289495-A1	Methods of inhibiting endothelin b receptor expressing tumor metastases	2020-03-24	

k) similar structure :-

SIMILAR 2/5

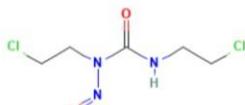
PubChem Carmustine (Compound)

1 Structures

1.1 2D Structure

Find Similar Structures Get Image Download

Chemical Structure Depiction



PubChem

CITE DOWNLOAD

CONTENTS

- Title and Summary
- 1 Structures
- 2 Names and Identifiers
- 3 Chemical and Physical Properties
- 4 Spectral Information
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing

15:04 01-11-2021

SIMILAR 2/5

PubChem Carmustine (Compound)

1.2 3D Conformer

Find Similar 3D Structures Get Image Download

Interactive Chemical Structure Model

- Ball and Stick
- Sticks
- Wire-Frame
- Space-Filling
- Show Hydrogens
- Animate



PubChem

2 Names and Identifiers

CITE DOWNLOAD

CONTENTS

- Title and Summary
- 1 Structures
- 2 Names and Identifiers
- 3 Chemical and Physical Properties
- 4 Spectral Information
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing

15:04 01-11-2021

L) References any 5 :-

Springer Nature References

PubChem Carmustine (Compound)

13.2 Springer Nature References

7,219 items View More Rows & Details

SORT BY Relevance

Thumbnail	Title	Publication Name	Publication Date	PMID
	Evaluation of toxicity of carmustine with or without bevacizumab in patients with recurrent or progressive high grade gliomas	Journal of Neuro-Oncology	2019	31432377
	Use of Gliadel (BCNU) Wafer in the Surgical Treatment of Malignant Glioma: A 10-Year Institutional Experience	Annals of Surgical Oncology	2008	18636295
	Inhibition of carboxyethylphosphoramido mustard formation from 4-hydroxycyclophosphamide by carmustine	AAPS PharmSci	1999	

CONTENTS

- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing
- 10 Safety and Hazards
- 11 Toxicity
- 12 Associated Disorders and Diseases
- 13 Literature**
- 14 Patents
- 15 Biomolecular Interactions and Pathways
- 16 Biological Test Results

Thieme References and Wiley References

PubChem Carmustine (Compound)

13.3 Thieme References

2 items View More Details

SORT BY Publication Date

Title	Publication Name	Publication Date	DOI
Chloro-, Bromo-, and Iodoalkanes (Vol. 35)	Science of Synthesis	2007	
Carmustine	Pharmaceutical Substances	2003	

Thieme Chemistry

13.4 Wiley References

1 item

SORT BY Publication Date

Title	Publication Name	Publication Date	DOI
Self-Assembled Gels for Biomedical Applications	Chem. Asian J.	2011	10.1002/asia.201000592

CONTENTS

- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing
- 10 Safety and Hazards
- 11 Toxicity
- 12 Associated Disorders and Diseases
- 13 Literature**
- 14 Patents
- 15 Biomolecular Interactions and Pathways
- 16 Biological Test Results

Conclusion:

Hence chemical data retrieval of our drug Carmustine was done using the online platform PubChem which includes the drug's Pharmacokinetics, pharmacodynamics, physiochemical properties, similar structures, toxicity, targets, spectral data etc

Practical No:1C

Aim: Chemical Data Retrieval using ZINC

Theory:

The ZINC Database contains commercially available compounds for structure based virtual screening. It currently has about 35 million compounds that can simply be purchased. It is provided in ready-to-dock, 3D formats with molecules represented in biologically relevant forms. It is available in subsets for general screening as well as target-, chemotype- and vendor-focused subsets. ZINC is free for everyone to use and download at the website zinc.docking.org. This database and service is provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California San Francisco. ZINC was originally designed for target based virtual screening (docking), and this remains its primary focus. However, ZINC is also useful for many other things, including: finding a compound to purchase, downloading a library in SMILES format for ligand based virtual screening, find compounds by similarity to a starting compound, find compound ANNOTATED for a particular target, find compounds PREDICTED for a particular target and much more. ZINC12:

Methodology:

My disease: Glioma

My drug:- Gliadel, Carmustine

Homepage of zinc database:-

The screenshot shows the ZINC15 homepage. At the top, there's a navigation bar with links for ZINC, Substances, Catalogs, Tranches, Biological, More, and About. Below the header, the main content area features a large "ZINC15" logo. To the right of the logo, a text box states: "ZINC is provided by the Irwin and Shoichet Laboratories in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). We thank NIGMS for financial support (GM71896)." Another text box below it provides citation information: "To cite ZINC, please reference: Sterling and Irwin, J. Chem. Inf. Model. 2015 <http://pubs.acs.org/doi/abs/10.1021/acs.jcim.5b00559>. You may also wish to cite our previous papers: Irwin, Sterling, Mysinger, Bolstad and Coleman, J. Chem. Inf. Model., 2012 DOI: 10.1021/ci3001277 or Irwin and Shoichet, J. Chem. Inf. Model. 2005;45(1):177-82 PDF, DOI."

Getting Started

- Getting Started
- What's New
- About ZINC 15 Resources
- Current Status / In Progress
- Why are ZINC results "estimates"?

Explore Resources

Chemistry

Ask Questions

You can use ZINC for **general** questions such as

- How many substances in current clinical trials have PAINS patterns? (150)
- How many natural products have names in ZINC and are not for sale? (9296) get them as SMILES, names and calculated logP
- How many endogenous human metabolites are there? (47319) and how many of these can I buy? (8271) How many are FDA approved drugs? (94)
- How many compounds known to aggregate are in current clinical

ZINC15 News

- 2018-02-14 - ZINC reaches 213,235,528 purchasable leadlike 3D!
- 2018-02-13 - ZINC reaches 736,001,654 purchasable molecules 2D!
- 2018-01-14 - Klara Anu is born! Welcome Klara Anu, sister to Lisa!
- 2018-01-01 - Chinzo Dandar joins our team. Welcome Chinzo! Follow us on twitter @chem4biology Known limitations What's new

At the bottom, there's a search bar with the placeholder "Type here to search" and a toolbar with various icons. The status bar at the bottom right shows "14:10 01-11-2021".

Go to search bar and search any data you want to retrieve:

The screenshot shows the ZINC substances search interface. At the top, there is a navigation bar with links for ZINC, Substances, Catalogs, Tranches, Biological, More, and About. Below the navigation bar is a search bar with the placeholder "Search for Substances" and a "Search" button. To the left of the search bar is a "Shopping List" button. Above the search bar are buttons for Help, Examples, Browse, Table, Subsets, and Shopping List. On the right side of the search bar is a "Search" button. Below the search bar are two sections: "Search Using One" and "Search Using Many". The "Search Using One" section contains a text input field for "ZINC ID, SMILES, SMARTS, or InChI" and a large grid of chemical structure fragments. The "Search Using Many" section contains a text input field for "One Identifier per Line" and an "OR Upload a File" section with a "Choose File" button. Below these sections are "Allow Lookups" and "Match Tolerance" settings. The "Allow Lookups" section includes checkboxes for ZINC ID, Structure, Names, Suppliers, and Analogs. The "Match Tolerance" section includes checkboxes for Retired IDs, Charge, Scaffold, Full Text, and Accept Multiple Results. At the bottom of the interface is a Windows taskbar with a search bar, pinned icons for WhatsApp, ZINC, and other applications, and system status indicators.

So we have found 1825113472 number of similar structures of Zinc:-

The screenshot shows the ZINC substance search results page. The search term "1825113472" is entered in the search bar at the top. Below the search bar is a toolbar with buttons for page navigation, download, filters, and lookup. The main area displays a grid of 12 chemical structures, each with its ZINC ID and name. The structures include ZINC7 (Acetamidoeugenol), ZINC10 (Acifran), ZINC11 (Aditoprim), ZINC12 (Adrafinil), ZINC14 (Allylprodine), ZINC15 (Alanine Mustard), ZINC16, ZINC17 (Albendazole Oxide), ZINC18 (Albutoin), ZINC21 (Alminoprofen), ZINC22 (Alprenolol), and ZINC23. Each structure is shown with its SMILES representation. A message "Waiting for stats.g.doubleclick.net..." is visible at the bottom of the grid. The interface is identical to the one shown in the previous screenshot, with a Windows taskbar at the bottom.

Search Gliadel:-

The screenshot shows the ZINC15 substances search interface. At the top, there is a navigation bar with links for WhatsApp, ZINC, Substances, Catalogs, Tranches, Biological, More, and About. Below the navigation bar, the main title is "Substances". A search bar contains the query "gliadel". Underneath the search bar, there are two sections: "Search Using One" and "Search Using Many". The "Search Using One" section includes a text input field for "ZINC ID, SMILES, SMARTS, or InChI" and a complex chemical structure editor with various buttons and a periodic table sidebar. The "Search Using Many" section includes a text input field for "ZINC ID, SMILES, SMARTS, InChI or Supplier Code" and an "OR Upload a File" section with a "Choose File" button. Below these sections are "Allow Lookups" and "Match Tolerance" checkboxes. The bottom part of the interface shows the search results for "gliadel", listing one result: "ZINC3830387 Bcnu". It displays the chemical structure of the compound, which is a substituted amine derivative. The results page includes standard navigation controls (prev/next, page number, download, filters) and a search bar at the bottom.

Screenshot of a web browser showing the ZINC database entry for substance ZINC3830387 (Bcnu). The URL is zinc.docking.org/substances/ZINC000003830387/. The page displays the following information:

- Substance Details:**
 - Added:** 2005-09-30
 - Availability:** In-Stock
 - Since:** 2015-08-07
 - Mwt:** 214.052
 - logP:** 1.157
 - Download:** Available
- Mol Formula:** C5H9Cl2N3O2
- Rings:** 0
- Heavy Atoms:** 12
- Hetero Atoms:** 7
- Fraction sp³:** 0.80
- Tranche:** BDGA

The chemical structure of Bcnu is shown as a 2D diagram with atoms colored by element (C=black, H=white, N=blue, O=red, Cl=green).

Below the main details, there is a section for "Available 3D Representations" which is currently "Waiting for referrer.disqus.com...".

At the bottom of the page, the Windows taskbar shows the search bar and system icons.

So here in Tranche you will get similar compounds:-

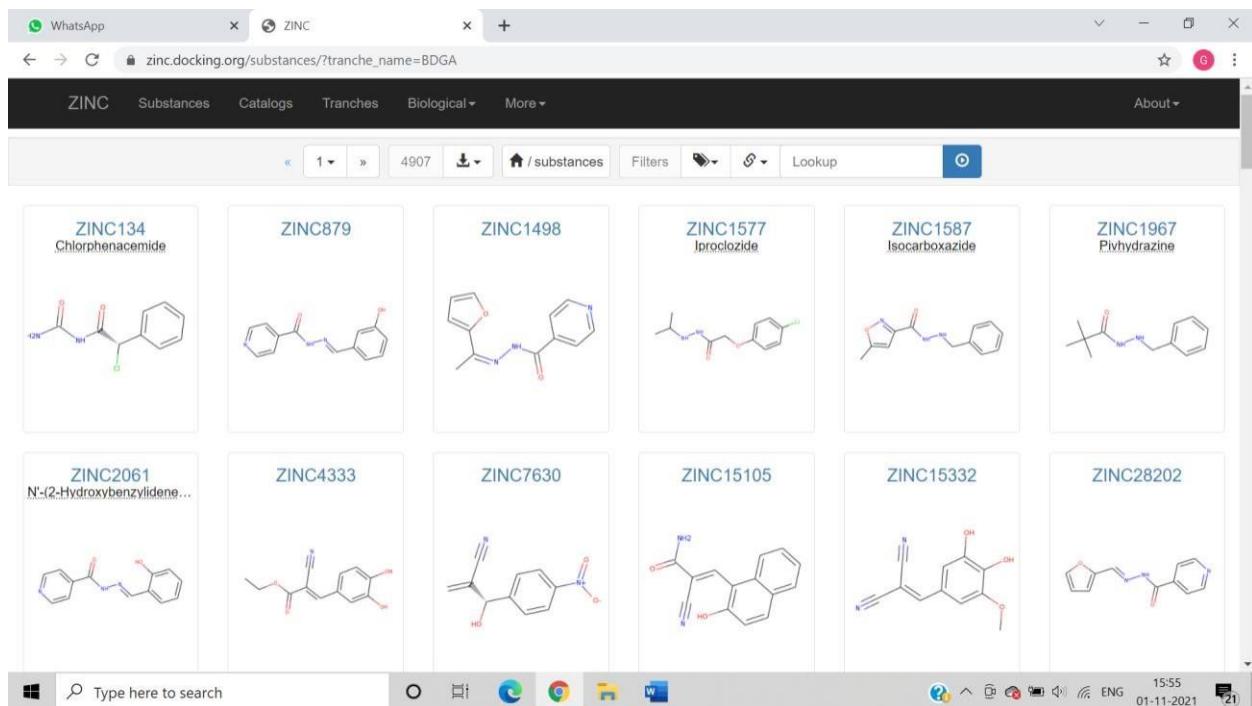
Screenshot of a web browser showing the ZINC database entry for substance ZINC3830387 (Bcnu). The URL is zinc.docking.org/substances/ZINC000003830387/. The page displays the following information:

- Substance Details:**
 - Added:** 2005-09-30
 - Availability:** In-Stock
 - Since:** 2015-08-07
 - Mwt:** 214.052
 - logP:** 1.157
 - Download:** Available
- Mol Formula:** C5H9Cl2N3O2
- Rings:** 0
- Heavy Atoms:** 12
- Hetero Atoms:** 7
- Fraction sp³:** 0.80
- Tranche:** BDGA

The chemical structure of Bcnu is shown as a 2D diagram with atoms colored by element (C=black, H=white, N=blue, O=red, Cl=green).

Below the main details, there is a section for "Available 3D Representations" which is currently "Waiting for referrer.disqus.com...".

At the bottom of the page, the Windows taskbar shows the search bar and system icons.



We have 52 vendors:-

The screenshot shows the vendor information for substance ZINC3830387 (Bcnr). It lists 52 vendors and 31 annotated catalogs. Below is a table of activities based on ChEMBL 20.

Gene Name	Class / Organism	pKi (L.E.)	Observations
GSR: Glutathione reductase, mitochondrial	enzyme / reductase (E)	5.09 (0.59)	1

SEA Predictions based on ChEMBL 20

There is currently no predicted activity for this compound. You may want to consider using [SEA](#) to predict it yourself.

The screenshot shows the ZINC15 web interface. At the top, there are tabs for WhatsApp, ZINC3830387 (Bcnr), and a search bar. Below the tabs, the URL is zinc.docking.org/substances/ZINC00003830387/. The main navigation menu includes ZINC, Substances, Catalogs, Tranches, Biological, More, ChEMBL 21, and CHEMBL513. A dropdown menu for BioSynth is open. The page title is "Activities based on ChEMBL 20". A table shows one observation: GSR: Glutathione reductase, mitochondrial enzyme / reductase (E) with pKi (L.E.) 5.09 (0.59). Below this, a section for "SEA Predictions based on ChEMBL 20" is shown, stating "There is currently no predicted activity for this compound. You may want to consider using [SEA](#) to predict it yourself." A "Run SEA" button is available. On the left, there's a "Interesting Analogs" section with four categories: Endogenous, Metabolites, Natural Products, and Aggregators, all showing "None Found". On the right, a "Framework of this compound" section shows the structure of ZINC3830387, which is a 1,2-diaminoimidazole derivative. The bottom of the screen shows a Windows taskbar with a search bar, icons for File Explorer, Google Chrome, and Microsoft Word, and system status indicators.

Scaffold of this Compound

Scaffold of this Compound

This compound has no scaffold by definition (no rings).

Substance Rings

This compound has no patterns of concern (or they have not yet been computed).

Smile notification:-

O=NN(CCC)C(=O)NCCCC

JSME Molecular Editor by Peter Ertl and Bruno Bienfait

One Identifier per Line

ZINC ID, SMILES, SMARTS, InChI or Supplier Code

OR Upload a File

Choose File No file chosen

Allow Lookups

ZINC ID Structure Names Suppliers Analogs Slow!

Match Tolerance

Retired IDs Charge Scaffold Full Text
 Accept Multiple Results

Subsets to Check

Nothing selected

Results

Output Format Summary Table Search Many

Than go to search with default:-

The screenshot shows the ZINC12 Substance Search interface. In the center, there is a JSME Molecular Editor window displaying the chemical structure of the target molecule: O=NN(CCCl)C(=O)NCCCl. To the left of the editor is a vertical element key for atoms: C, N, O, S, F, Cl, Br, I, Y, X. Below the editor are search dropdowns for "Search with" (set to "Default") and "Nothing selected". On the right side, there are sections for "Allow Lookups" (checkboxes for ZINC ID, Structure, Names, Suppliers, Analogs), "Match Tolerance" (checkboxes for Retired IDs, Charge, Scaffold, Full Text, Accept Multiple Results), and "Subsets to Check" (dropdown set to "Nothing selected"). At the bottom right, there is an "Output Format" dropdown set to "Summary Table" and a "Search Many" button.

Result:-

The screenshot shows the detailed information for substance ZINC3830387 (Bcnu). The title "ZINC3830387 (Bcnu)" is at the top. Below it is a table with columns: Added (2005-09-30), Availability (In-Stock), Since (2015-08-07), Mwt (214.052), logP (1.157), and Download. To the right of the table is a 2D chemical structure of the molecule. Below the table are sections for Mol Formula (C5H9Cl2N3O2), Rings (0), Heavy Atoms (12), Hetero Atoms (7), Fraction sp³ (0.80), and Tranche (BDGA). Further down are sections for SMILES (O=NN(CCCl)C(=O)NCCCl), InChI (InChI=1S/C5H9Cl2N3O2/c6-1-3-8-5(11)10(9-12)4-2-7/h1-4H2,(H,8,11)), and InChI Key (DLGOEMSEDOBKAD-UHFFFAOYSA-N). At the bottom, there is a section for Available 3D Representations with various properties like pH range, Net charge, H-bond donors, H-bond acceptors, tPSA, Rotatable bonds, Apolar desolvation, Polar desolvation, and Download.

ZINC Substances Catalogs Tranches Biological More About

Run search for more Run search for more Run search for more Run search for more

Drugs Find More **In Man** Find More **Bioactives** Find More **Purchasable** Find More, Unsorted **Nothing else with this framework**

ZINC3830387 bcnu Identity ZINC3830387 bcnu Identity ZINC3830387 bcnu Identity ZINC3830387 bcnu Identity

Scaffold of this Compound

This compound has no scaffold by definition (no rings)

Type here to search

Now go to search with 40% similarity:-

ZINC Substances Catalogs Tranches Biological More About

Choose File No file chosen

Allow Lookups

ZINC ID Structure Names Suppliers Analogs Slow!

Match Tolerance

Retired IDs Charge Scaffold Full Text
 Accept Multiple Results

Subsets to Check

Nothing selected

Results

Output Format Summary Table Search Many

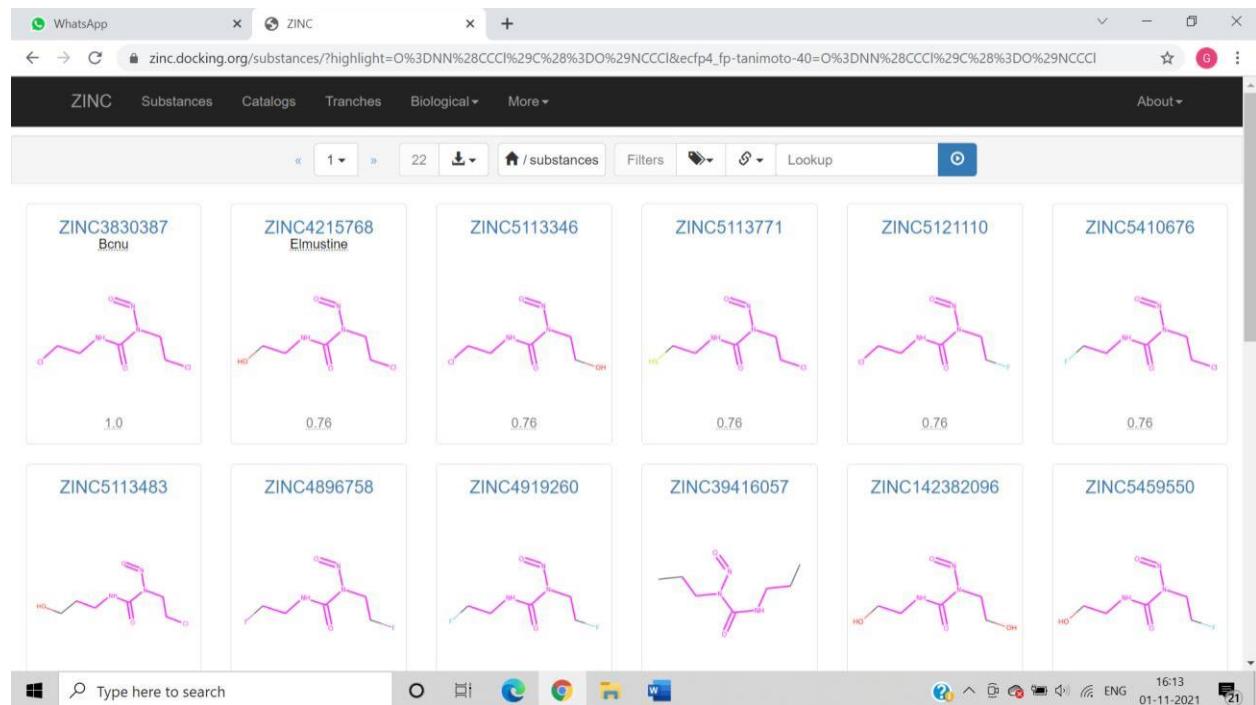
Search with Nothing selected

Default Similarity - 40 Similarity - 30 Substructure Smarts

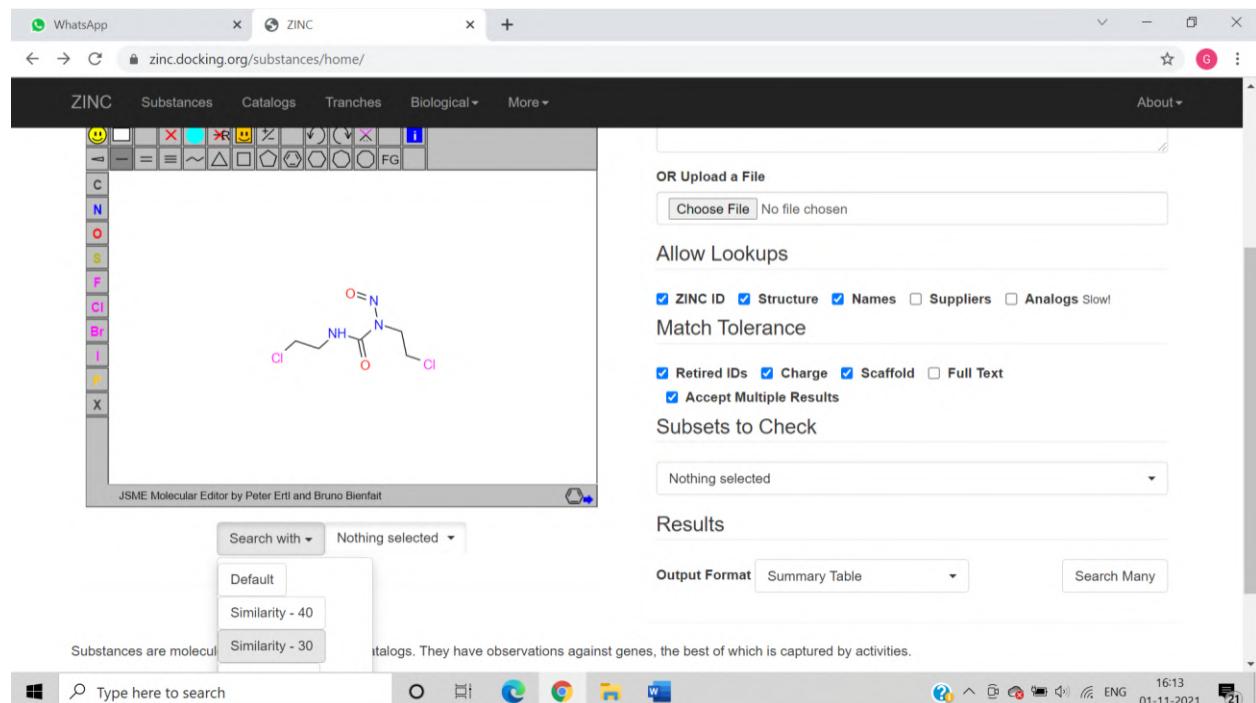
Substances are molecular catalogs. They have observations against genes, the best of which is captured by activities.

Type here to search

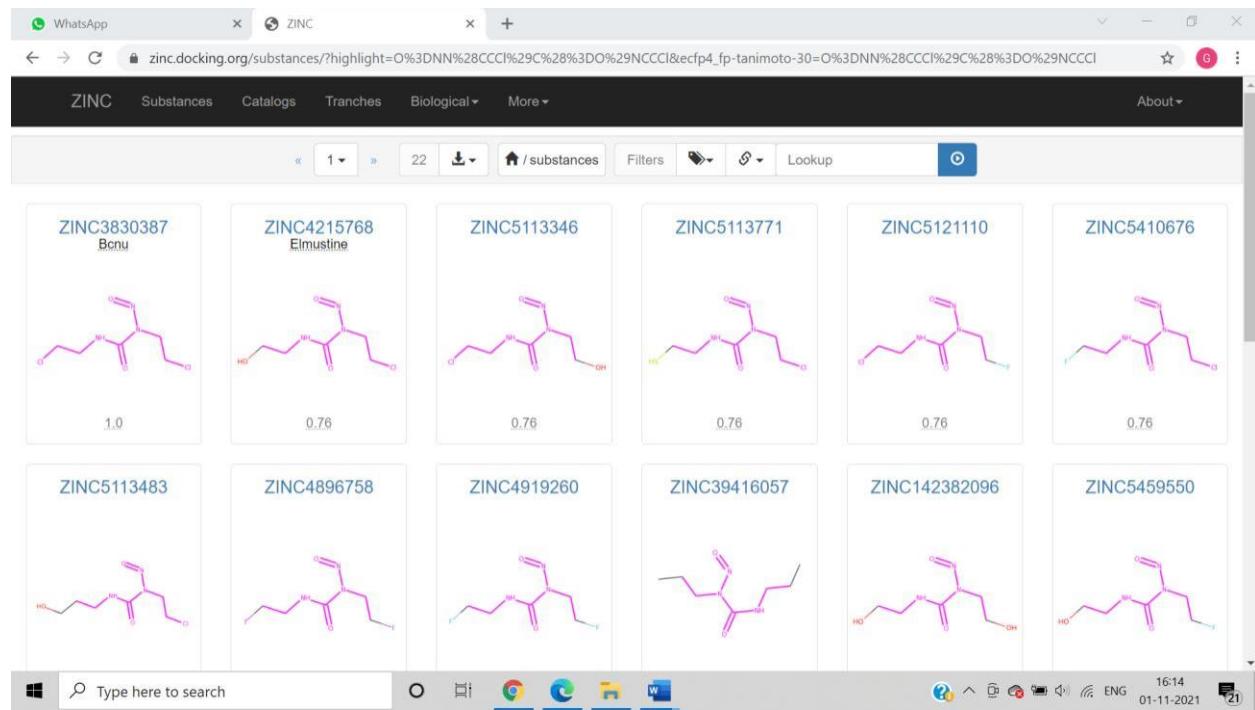
Result:-



Now search with 30% similarity:-



Result:-



Now go to nothing selected and select In trials and choose 40% similarity:-

All Subsets
Agent
Bb
Boutique
For Sale
In Stock
Not For Sale
Now
On Demand
Wait OK
Fda
In Cells
In Cells Only
In Man
In Man Only
In Trials
Retired IDs

Search with ▾ Compounds that have been investigated, including drugs ▾

OR Upload a File Choose File No file chosen

Allow Lookups

ZINC ID Structure Names Suppliers Analogs Slow!

Match Tolerance

Retired IDs Charge Scaffold Full Text
 Accept Multiple Results

Subsets to Check

Nothing selected

Results

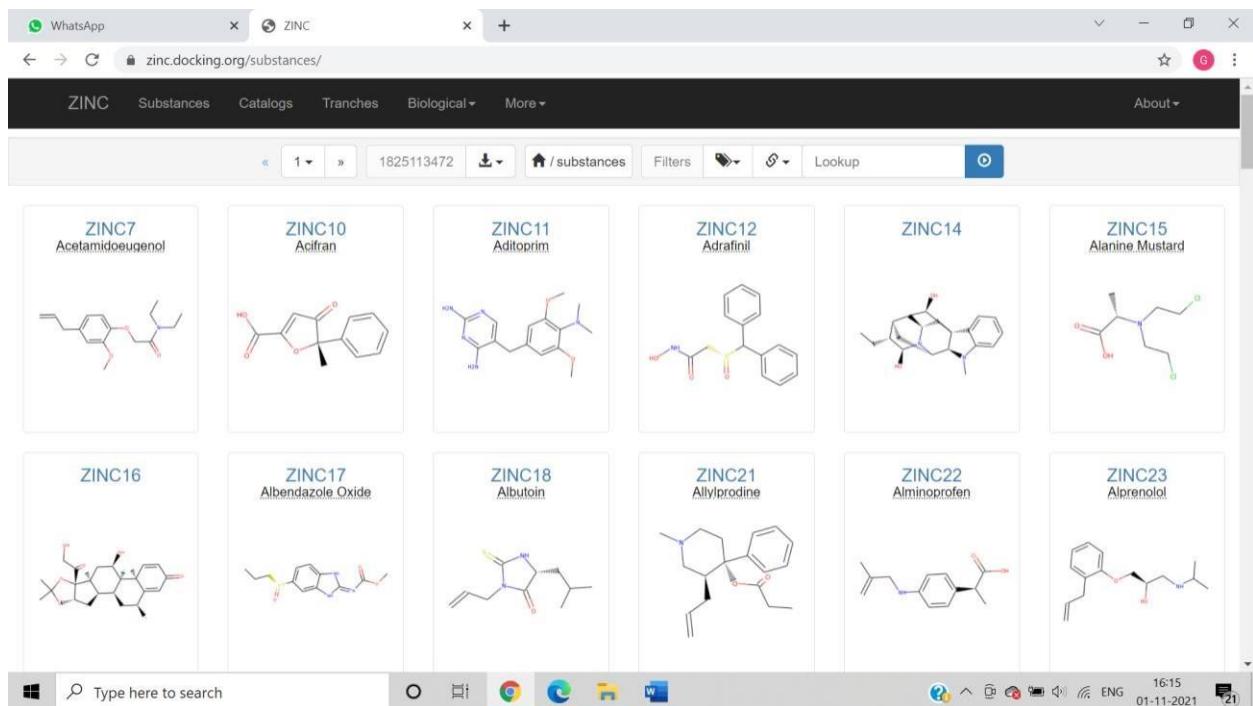
Output Format Summary Table Search Many

Substances are molecules that are loaded from catalogs. They have observations against genes, the best of which is captured by activities.

The screenshot shows the ZINC15 substances search interface. On the left, there is a JSME Molecular Editor window displaying a chemical structure of 2-(2-chloroethyl)-N,N-dimethylbenzimidazole-1,3(2H,3H)-dione. Below the editor are search filters: 'Search with' dropdown set to 'Compounds that have been investigated, including drugs', and a 'Subsets' dropdown menu with options: Default, Similarity - 40, Similarity - 30, Substructure, and Smarts. The main search area has a 'Match Tolerance' section with checkboxes for ZINC ID, Structure, Names, Suppliers, and Analogs (disabled). It also includes sections for Retired IDs, Charge, Scaffold, Full Text, and Accept Multiple Results. A 'Subsets to Check' dropdown is set to 'Nothing selected'. The results section is currently empty. At the bottom, there are links for Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896, Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs, and UC Regents.

Result:-

The screenshot shows the ZINC15 substances subsets search interface. It displays two identical search results panels. Each panel has a 'Get Total' button, a download icon, a URL link to '/substances / subsets / in-trials', a 'Filters' button, and a 'Lookup' button. The results section is currently empty. At the bottom, there are links for Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896, Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs, and UC Regents.



Go to catlogs:-

The screenshot shows a web browser window with the URL zinc.docking.org/catalogs/home/. The page title is "Catalogs". It features a search bar with "Search for Catalogs" and a "Search" button. Below the search bar, a message states: "Catalogs contain substances, which inherit properties of purchasability, biogenicity and bioactivity from their catalog membership." At the bottom of the page, there is footer text including links for Acknowledgements, Usage, Why are ZINC results "estimates?", Terms of use, Privacy policy, and Support. It also mentions NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs, and UC Regents. The page was originally generated at 2021-11-01 04:06:46.596495 in 0.02003s on zinc.docking.org using ZINC15.0 20210303.1.

Search gliadel components:-

The screenshot shows the ZINC Catalogs homepage. At the top, there are three tabs: WhatsApp, ZINC AbovChem, and ZINC. Below the tabs, a navigation bar includes links for ZINC, Substances, Catalogs, Tranches, Biological, More, and About. A search bar at the top right contains the query "gliadel". Below the search bar, a message states: "Catalogs contain substances, which inherit properties of purchasability, biogenicity and bioactivity from their catalog membership." A footer section includes links for Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs and UC Regents, and a note about the original generation date.

The screenshot shows the ZINC Catalogs search results page for the query "gliadel". The interface is identical to the homepage, with the search term "gliadel" entered in the search bar. The results table shows one entry with the number "1" in the first column and a "Get Total" button in the second column. The third column contains a link to "/catalogs". The fourth column has buttons for Filters and Sort. The fifth column displays the search term "gliadel". The footer information is identical to the homepage.

This screenshot shows the same ZINC Catalogs search results page for "gliadel", but the results table is displayed in a different format. It shows two rows of results, each with a "Get Total" button, a "/catalogs" link, and a "gliadel" search term. The rest of the interface and footer are identical to the previous screenshots.

Go to abo chem:-

The screenshot shows the ZINC platform interface with several tabs and sections:

- Top Navigation:** WhatsApp, zinc.docking.org/catalogs/, ZINC AbovChem, ZINC.
- Header:** ZINC, Substances, Catalogs, Tranches, Biological, More, About.
- Toolbar:** Back, Forward, Home, Catalogs, Filters, Lookup.
- Supplier Logos:** OOI CHEMICAL, 1PlusChem (1), A2B, SYNTHESIS^{A2Z}, SYNTHESIS^{A2Z}, aablocks.
- Supplier Details:** 001 Chemical, 1PlusChem LLC, A2B Chem Building Blocks, A2Z Synthesis, A2Z BB, AA Blocks.
- Supplier Logos:** AbovChem, Accela ChemBio Inc., AnalytiCon DISCOVERY, AnalytiCon DISCOVERY, AnalytiCon DISCOVERY, AnalytiCon DISCOVERY.
- Supplier Details:** AbovChem, Accela ChemBio Inc., AnalytiCon Discovery, AnalytiCon Discovery Natural Derivatives, AnalytiCon Discovery NP, AnalytiCon Discovery NP BB.
- Search Bar:** Type here to search.
- Bottom Navigation:** WhatsApp, zinc.docking.org/catalogs/abovchem/, ZINC AbovChem, ZINC AbovChem.
- Content Area:** AbovChem Catalog Properties, Last ZINC Import, Interesting Substance Subsets, Useful Links, Sole Source Substances.

Screenshot of a web browser showing a collection of ZINC substances. The page displays a grid of six rows and two columns of chemical structures. Each card contains the ZINC ID, substance name, and a 2D chemical structure.

ZINC ID	Substance Name	Chemical Structure
ZINC23	Alprenolol	
ZINC53	Aspirin	
ZINC61	Baclofen	
ZINC83		
ZINC96	Dexbrompheniramine	
ZINC99	Bucetin	
ZINC128	Carteolol	
ZINC133	Chlormezanone	
ZINC164	Climbazole	
ZINC189		
ZINC242	Doxylamine	
ZINC257	Esmolol	

Below the grid is a search bar and the Windows taskbar.

Screenshot of a web browser showing the details for ZINC61 (Baclofen). The page includes a chemical structure, various identifiers, and a table of properties.

ZINC61 (Baclofen)

In: anodyne bb fda for-sale in-stock

Google Wikipedia PubMed

Added	Availability	Since	Mwt	logP	Download
2005-09-27	In-Stock	2015-08-07	213.664	1.857	

Mol Formula	Rings	Heavy Atoms	Hetero Atoms	Fraction sp ³	Tranche
C10H12ClNO2	1	14	4	0.30	BDAA

SMILES: NC[C@H](CC(=O)O)c1ccc(Cl)cc1

InChI: InChI=1S/C10H12ClNO2/c11-9-3-1-7(2-4-9)8(6-12)5-10(13)14/h1-4,8H,5-6,12H2,(H,13,14)/t8-/m0/s1

InChI Key: KPYSYYIEGFHWV-QMMMGPOBSA-N

Available 3D Representations

N-Substituted
H-bond acceptor
H-bond donor
Rotatable bonds
Apolar
Polar

Waiting for accounts.google.com...

Below the 3D representation is a search bar and the Windows taskbar.

So we have 90 vendors:-

The screenshot shows the ZINC16 database interface. At the top, there are tabs for WhatsApp, ZINC AbovChem, and ZINC61 (Baclofen). Below the tabs, the main navigation bar includes links for ZINC, Substances, Catalogs, Tranches, Biological, More, and About.

Vendors (90 Total)

	pH range	Net charge	H-bond donors	H-bond acceptors	tPSA	Rotatable bonds	Apolar desolvation	Polar desolvation	Download
Reference		0	1	2	67	4	4.81	-48.5	Download

Annotated Catalogs (40 Total)

Illuminating the Druggable Genome Screening Library	1	Prestwick-2-A-6, Spectrum-10-B-5		
MicroSource Pharmakon	1	01500135		
MicroSource Spectrum	1	01500135		
MicroSource US Drugs	1	01500135		
MLSMR	1	855960, 858232		
Prestwick Chemical	1	Prestw-85		
SMDC Pharmakon	1	129991		
Ambinter	1	Amb10845652, Amb19930345, Amb4243162, Amb621641		
BindingDB.org	1	24182, 50032964		
ChEBI	1	CHEBI:2972		

At the bottom, the URL is https://zinc.docking.org/catalogs/pharmek/, the page title is BL 20, and the status bar shows 2 Total Observations 1641 01-11-2021 21.

One of them is:

Pharmeks

The screenshot shows the ZINC16 database interface for the Pharmeks catalog. At the top, there are tabs for WhatsApp, ZINC AbovChem, and ZINC Pharmeks. Below the tabs, the main navigation bar includes links for ZINC, Substances, Catalogs, Tranches, Biological, More, and About.

Pharmeks

In: for-sale in-stock now wait-ok

Contact Information

Phone	+7 (095) 702-9648
Fax	+7 (095) 702-9648
Website	www.pharmeks.com
Email	sales@pharmeks.com

Catalog Properties

Purchasability	In-Stock
Building Blocks	No
Activity Level	Unspecified
Biogenicity Level	Unspecified

Last ZINC Import

Version	2018-09-11
Last Loaded On	2018-09-11
Original Catalog Size	360084
Compounds Removed	264
Est. Sole Supplier	8332
Depleted Entries	Unknown (Browse)

Useful Links

- [Browse Substances](#)
- [Browse Catalog items](#)
- [Browse Protomers](#)
- [Genes \(With active compounds in this catalog\)](#)

Sole Source Substances

Interesting Substance Subsets

Browse All

Establishing secure connection...

Type here to search

1642 01-11-2021 21

Tranches:-

The screenshot shows a web-based application for managing chemical tranches. At the top, there's a navigation bar with links for WhatsApp, ZINC, Substances, Catalogs, Tranches, Biological, More, and About. Below the navigation is a toolbar with buttons for Rep., 2D, 3D, Read, Standard, Purch., Wait OK, pH, N/A, Charge, N/A, and various filters like Rep., 2D, 3D, Read, Standard, Purch., Wait OK, pH, N/A, Charge, N/A, and a download icon.

The main content area displays a table titled "Molecular Weight (up to, Daltons)". The table has columns for LogP (up to) values from -1 to 4.5, and rows for different molecular weights. The columns include: 200, 250, 300, 325, 350, 375, 400, 425, 450, 500, >500, and Totals, by LogP. The data shows the count of molecules for each LogP range and their total count across all LogP ranges.

Molecular Weight (up to, Daltons)												
	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	29,293	204,598	784,279	1,125,069	2,321,356	854,208	300,607	128,558	99,872	86,323	5,615	5,939,778
0	142,690	1,067,035	3,992,760	5,372,590	10,975,901	3,784,188	1,767,726	775,279	606,137	558,305	3,798	29,046,409
1	376,413	3,284,847	13,196,175	17,023,840	34,876,129	12,665,806	7,279,946	3,517,752	2,892,839	2,688,060	7,987	97,809,794
2	497,750	5,391,816	25,622,912	32,914,848	67,733,100	28,989,280	19,267,814	10,563,987	9,000,177	8,721,010	20,894	208,723,588
2.5	189,326	2,643,678	14,831,118	19,486,349	40,600,593	20,281,809	15,126,147	9,325,848	8,120,159	7,879,918	21,325	138,506,270
3	108,266	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,728	12,045,788	10,696,251	10,674,949	33,982	143,848,280
3.5	48,705	1,336,320	10,135,959	14,349,999	29,671,752	21,055,698	18,737,428	13,954,286	12,511,433	12,736,846	54,896	134,593,322
4	15,100	613,109	6,131,454	8,128,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529	13,378,049	82,058	100,238,996
4.5	1,993	170,043	2,873,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431	12,208,182	113,230	75,539,989

Go to In stock:-

This screenshot shows the same ZINC Tranches interface, but the table is titled "Minimum Purchasability". It includes an additional column for "In-Stock" status. The columns for LogP (up to) and Totals, by LogP remain the same as the previous table.

Minimum Purchasability												
	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	29,293	204,598	784,279	1,125,069								5,939,778
0	142,690	1,067,035	3,992,760	5,372,590								29,046,409
1	376,413	3,284,847	13,196,175	17,023,840								97,809,794
2	497,750	5,391,816	25,622,912	32,914,848	67,733,100	28,989,280	19,267,814	10,563,987	9,000,177	8,721,010	20,894	208,723,588
2.5	189,326	2,643,678	14,831,118	19,486,349	40,600,593	20,281,809	15,126,147	9,325,848	8,120,159	7,879,918	21,325	138,506,270
3	108,266	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,728	12,045,788	10,696,251	10,674,949	33,982	143,848,280
3.5	48,705	1,336,320	10,135,959	14,349,999	29,671,752	21,055,698	18,737,428	13,954,286	12,511,433	12,736,846	54,896	134,593,322
4	15,100	613,109	6,131,454	8,128,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529	13,378,049	82,058	100,238,996
4.5	1,993	170,043	2,873,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431	12,208,182	113,230	75,539,989

Result:

The screenshot shows a grid of 12 chemical structures from the ZINC12 substances subset. Each card includes the ZINC ID, substance name, and a 2D chemical structure. The substances include Caffeine (ZINC1084), Theobromine (ZINC2151), and several tranches such as ZINC8151, ZINC8153, ZINC8155, ZINC8657, ZINC10969, and ZINC18276.

ZINC ID	Substance Name	Chemical Structure
ZINC1084	Caffeine	<chem>CN1C=NC2=C1C(=O)C(=O)N(C2=O)C</chem>
ZINC1169		<chem>CC[C@H]1[C@@H](CO)[C@H](C1=O)N2C[C@H]1[C@H](O[C@H]1C[C@H](O)[C@H](O)[C@H]1O)C2=O</chem>
ZINC1447		<chem>CN1C=NC2=C1C(=O)C(=O)N(C2=O)C</chem>
ZINC2151	Theobromine	<chem>CN1C=NC2=C1C(=O)C(=O)N(C2=O)C</chem>
ZINC3976		<chem>C[C@H]1[C@H](O)[C@H](O)[C@H]1O</chem>
ZINC5168		<chem>CC1[C@H]2[C@H](O)[C@H]1O[C@H]2O</chem>
ZINC8151		<chem>CC[C@H]1[C@@H](O)[C@H](O)[C@H](O)[C@H]1O</chem>
ZINC8153		<chem>CC[C@H]1[C@@H](O)[C@H](O)[C@H](O)[C@H]1O</chem>
ZINC8155		<chem>CC[C@H]1[C@@H](O)[C@H](O)[C@H](O)[C@H]1O</chem>
ZINC8657		<chem>C1=CN=C2=C1C(=O)C(=O)N2C</chem>
ZINC10969		<chem>C1=CN=C2=C1C(=O)C(=O)N2C</chem>
ZINC18276		<chem>CC[C@H]1[C@H](O)[C@H](O)[C@H](O)[C@H]1O</chem>

To download:

The screenshot shows a modal window titled "Download Tranches" over the ZINC Tranches home page. The modal displays a table of selected substances and allows filtering by LogP and modification date. A "Download" button is visible at the bottom right.

1,721 (Non-Empty) Tranches Selected (968,412,837 Substances)									
	200	250							Totals, by LogP
-1	29,293	2,000							86,323
0	142,690	1,000							5,615
1	376,413	3,200							5,939,778
2	497,750	5,300							558,305
2.5	189,326	2,000							7,987
3	108,266	2,000							29,046,409
3.5	48,705	1,336,320	10,135,959	14,349,999	29,671,752	21,055,698	18,737,428	13,954,286	21,325
4	15,100	613,109	6,131,454	8,128,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529
4.5	1,993	170,043	2,873,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431
									12,208,182
									113,230
									75,539,989

To have the fragments compounds:-

Molecular Weight (up to, Daltons)												
	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	29,293	204,598	784,279	1,125,089	2,321,356	854,208	300,807	125,508	80,872	66,323	20,000	233,891
0	142,690	1,067,035	3,902,760	5,372,590	10,075,901	3,784,168	1,767,726	775,279	606,137	558,305	20,000	1,209,725
1	376,413	3,284,847	13,106,175	17,023,840	34,876,129	12,665,800	7,279,946	3,517,752	2,892,839	2,686,060	1,000	3,661,260
2	497,750	5,391,816	25,622,912	32,914,848	67,733,100	28,989,280	19,267,814	10,583,987	9,000,177	8,721,010	20,000	5,889,566
2.5	189,326	2,643,678	14,831,118	19,486,349	40,600,593	20,281,809	15,126,147	9,325,848	8,120,159	7,870,918	20,000	2,833,004
3	108,266	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,728	12,045,788	10,696,251	10,674,049	20,000	2,183,600
3.5	48,705	1,336,320	10,135,958	14,349,099	29,871,752	21,055,898	18,737,428	13,954,286	12,511,433	12,736,846	20,000	1,385,025
4	50,700	813,108	6,131,854	8,128,568	12,531,547	15,472,307	16,892,846	14,129,428	12,864,529	13,370,049	32,058	0
4.5	50,700	813,108	2,573,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431	12,208,182	111,250	0

To have Goldilocks:-

Molecular Weight (up to, Daltons)												
	200	250	300	325	350	375	400	425	450			Totals, by LogP
-1	29,293	204,598	784,279	1,125,089	2,321,356	854,208	300,807	125,508	80,872	66,323	20,000	233,891
0	142,690	1,067,035	3,902,760	5,372,590	10,075,901	3,784,168	1,767,726	775,279	606,137	558,305	20,000	1,209,725
1	376,413	3,284,847	13,106,175	17,023,840	34,876,129	12,665,800	7,279,946	3,517,752	2,892,839	2,686,060	1,000	3,661,260
2	497,750	5,391,816	25,622,912	32,914,848	67,733,100	28,989,280	19,267,814	10,583,987	9,000,177	8,721,010	20,000	5,889,566
2.5	189,326	2,643,678	14,831,118	19,486,349	40,600,593	20,281,809	15,126,147	9,325,848	8,120,159	7,870,918	20,000	2,833,004
3	108,266	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,728	12,045,788	10,696,251	10,674,049	20,000	2,183,600
3.5	48,705	1,336,320	10,135,958	14,349,099	29,871,752	21,055,898	18,737,428	13,954,286	12,511,433	12,736,846	20,000	1,385,025
4	50,700	813,108	6,131,854	8,128,568	12,531,547	15,472,307	16,892,846	14,129,428	12,864,529	13,370,049	32,058	0
4.5	50,700	813,108	2,573,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431	12,208,182	111,250	0

WhatsApp ZINC zinc.docking.org/tranches/home/#

ZINC Substances Catalogs Tranches Biological More About

Rep. 2D 3D React. Standard Purch. Wait OK pH N/A Charge N/A

Molecular Weight (up to, Daltons)

	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	86,258	204,500	784,270	3,125,080	2,321,156	858,200	3,000,007	1,205,530	90,872	88,323	0	0
0	142,800	1,807,035	3,902,760	5,372,590	10,075,301	3,784,188	1,767,726	775,276	808,137	558,108	0	0
1	370,443	3,264,847	13,106,175	17,023,040	34,076,129	12,065,800	1,279,940	3,517,752	2,892,020	2,605,060	0	0
2	497,720	5,391,810	25,622,912	32,914,848	67,733,100	20,960,280	19,267,814	10,563,387	9,000,177	8,721,010	0	126,270,860
2.5	100,320	2,843,878	14,831,118	19,486,349	40,600,593	20,281,800	15,126,147	9,325,840	8,120,159	7,870,918	0	74,918,060
3	100,320	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,726	12,045,788	10,696,251	10,674,049	0	68,372,125
3.5	40,700	1,336,320	10,135,950	14,349,990	29,671,752	21,055,696	16,737,426	13,954,286	12,511,433	12,730,846	0	0
4	100,320	613,100	6,131,254	8,126,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529	13,376,049	0	0
4.5	100,320	170,043	2,373,064	4,632,330	7,889,352	10,950,424	12,773,295	12,356,636	11,582,431	12,200,182	0	0

Biological process:

Organism:

WhatsApp ZINC zinc.docking.org/tranches/home/#

ZINC Substances Catalogs Tranches Biological More About

Rep. 2D 3D React. Standard Purch. Wait OK pH N/A Charge N/A

Entities

Organisms

Major Target Classes

Minor Target Classes

Molar Weight (up to, Daltons)

	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	29,210	204,500	321,156	858,200	3,000,007	1,205,530	90,872	88,323	0	0	0	0
0	142,800	1,807,035	10,075,301	3,784,188	1,767,726	775,276	808,137	558,108	0	0	0	0
1	370,443	3,264,847	12,065,800	1,279,940	3,517,752	2,892,020	2,605,060	0	0	0	0	0
2	497,720	5,391,810	25,622,912	32,914,848	67,733,100	20,960,280	19,267,814	10,563,387	9,000,177	8,721,010	0	126,270,860
2.5	100,320	2,843,878	14,831,118	19,486,349	40,600,593	20,281,800	15,126,147	9,325,840	8,120,159	7,870,918	0	74,918,060
3	100,320	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,726	12,045,788	10,696,251	10,674,049	0	68,372,125
3.5	40,700	1,336,320	10,135,950	14,349,990	29,671,752	21,055,696	16,737,426	13,954,286	12,511,433	12,730,846	0	0
4	100,320	613,100	6,131,254	8,126,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529	13,376,049	0	0
4.5	100,320	170,043	2,373,064	4,632,330	7,889,352	10,950,424	12,773,295	12,356,636	11,582,431	12,200,182	0	0

Homo Sapiens:-

The screenshot shows a web browser window with the URL zinc.docking.org/organisms/home/. The page title is "Organism class (organism)". A search bar contains the text "homo sapiens". Below the search bar, a message states: "Each gene is classified by the organism class to which it belongs." At the bottom, there is a footer with links to Acknowledgements, Usage, Why, Terms of use, Privacy policy, and Support.

The screenshot shows a web browser window with the URL zinc.docking.org/organisms/search/?q=homo+sapiens. The page title is "Organisms". A search bar contains the text "homo sapiens". Below the search bar, there is a table with columns: Code, Name, # Genes, # Orthologs, # Observations, # Substances, # Purchasable, and # Predictions. The table has one row for "homo sapiens". At the bottom, there is a footer with links to Acknowledgements, Usage, Why, Terms of use, Privacy policy, and Support.

The screenshot shows a web browser window with the URL <https://zinc.docking.org/organisms/>. The page title is "Organisms". A search bar contains the text "homo sapiens". Below the search bar, there is a table with columns: Code, Name, # Genes, # Orthologs, # Observations, # Substances, # Purchasable, and # Predictions. The table has one row for "homo sapiens". At the bottom, there is a footer with links to Acknowledgements, Usage, Why, Terms of use, Privacy policy, and Support.

ZINC Substances Catalogs Tranches Biological More About

Code	Name	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
A	Archaea	3	3	51	27	22	80153
B	Bacteria	369	525	10457	7625	1498	18561828
E	Eukaryotes	2864	4290	845949	391374	42305	178124739
U	Unknown	2	2	3	3	1	174
V	Viruses	73	112	19111	13445	1220	4289521

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 04:56:45.044258 in 0.05285s on zinc.docking.org using ZINC15.0.20210303.1

Go to Eukaryotes:-

ZINC Substances Catalogs Tranches Biological More About

Code	Name	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
A	Archaea	3	3	51	27	22	80153
B	Bacteria	369	525	10457	7625	1498	18561828
E	<u>Eukaryotes</u>	2864	4290	845949	391374	42305	178124739
U	Unknown	2	2	3	3	1	174
V	Viruses	73	112	19111	13445	1220	4289521

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 04:56:45.044258 in 0.05285s on zinc.docking.org using ZINC15.0.20210303.1

https://zinc.docking.org/organisms/eukaryotes/

Type here to search

17:26 01-11-2021

WhatsApp ZINC zinc.docking.org/organisms/eukaryotes/

ZINC Substances Catalogs Tranches Biological More About

Organism: Eukaryotes (E) - page under construction

- Number of genes: **2864**
- Number of observations: **845949**
- Number of orthologs: **4290**
- Number of predictions: **178124739**
- Number of purchasable: **42305**
- Fulltext: **E Eukaryotes**
- Public_identifier: **Eukaryotes**
- Token: **Eukaryotes**

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 04:57:37.542Z in 0.03954s on zinc.docking.org using ZINC15.0.26215003.1

17:26 01-11-2021

Go to eukaryotes , substances:

WhatsApp ZINC zinc.docking.org/organisms/eukaryotes/substances/

ZINC Substances Catalogs Tranches Biological More About

1 » Get Total Download / organisms / eukaryotes / substances Filters Lookup

ZINC10 Acifran	ZINC23 Alprenolol	ZINC53 Aspirin	ZINC61 Baclofen	ZINC71 Benzorzone	ZINC75 Benzoclidine
ZINC76 Benzoin	ZINC83	ZINC92 Brolamfetamine	ZINC96 Dexbrompheniramine	ZINC103 (s)-Bulbocapnine	ZINC128 Carteolol

17:27 01-11-2021

Go to viruses:

The screenshot shows a web browser window with the URL zinc.docking.org/organisms/home/. The page title is "Organism class (organism)". A search bar at the top right contains the word "viruses". Below the search bar, a message states: "Each gene is classified by the organism class to which it belongs." At the bottom of the page, there is a footer with links to Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs and UC Regents. The status bar at the bottom of the browser window shows the time as 17:27 and the date as 01-11-2021.

Go to genes in viruses:

The screenshot shows a web browser window with the URL zinc.docking.org/organisms/search/?q=viruses. The page title is "Organism class (organism)". A search bar at the top right contains the word "viruses". Below the search bar, a table displays the following data:

Code	Name	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
V	Viruses	73	112	19111	13445	1220	4289521

At the bottom of the page, there is a footer with links to Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs and UC Regents. The status bar at the bottom of the browser window shows the time as 17:28 and the date as 01-11-2021. The address bar at the very bottom of the browser window shows the URL https://zinc.docking.org/organisms/search/?sort=num_genes&q=viruses.

ZINC Substances Catalogs Tranches Biological More About

Name	Description	Organism	Sub Class (Major Class)	Orthologs	Observations	Substances	Purchasable	Predicted
NA	Neuraminidase	Viruses	hydrolase (enzyme)	14	977	480	20	23041
M	Matrix protein 2	Viruses	IC-other (ion channel)	1	32	32	12	1025
DPOL_HHV11	DNA polymerase catalytic subunit	Viruses	transferase (enzyme)	1	6	6	4	31646
TK	Thymidine kinase	Viruses	enzyme-other (enzyme)	5	302	147	35	20453
TAT_HV112	Protein Tat	Viruses	TF-other (Transcription factor)	1	58	49	3	3374
NS4A	Non-structural protein 4A	Viruses	protease (enzyme)	1	1	1	0	4879
UL80	Capsid scaffolding protein	Viruses	protease (enzyme)	1	208	182	4	176053

Now go to major drug classes:

ZINC Substances Catalogs Tranches Biological More About

Name	Description	Organism	Orthologs	Observations	Substances	Purchasable	Predicted
NA	Neuraminidase	Viruses	14	977	480	20	23041
M	Matrix protein 2	Viruses	1	32	32	12	1025
DPOL_HHV11	DNA polymerase catalytic subunit	Viruses	1	6	6	4	31646
TK	Thymidine kinase	Viruses	5	302	147	35	20453
TAT_HV112	Protein Tat	Viruses	1	58	49	3	3374
NS4A	Non-structural protein 4A	Viruses	1	1	1	0	4879
UL80	Capsid scaffolding protein	Viruses	1	208	182	4	176053

Protein Kinase:

So go to 13 enzymes and you will find kinase:

Name	# Sub Classes	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
adhesion	1	7	11	534	292	32	79415
auxiliary transport protein	3	8	14	643	458	182	497749
cytosolic other	1	39	58	5664	4162	461	2859674
enzyme	13	1942	2819	412921	205504	23889	107386614
epigenetic regulator	3	View target classes of the enzyme major class		103	6250	2856	510
ion channel	3	152	246	34563	22500	2638	34121578
membrane other	1	6	12	301	271	30	166343
membrane receptor	7	289	670	307365	143352	13445	79804362
Nuclear-other	1	6	8	1053	784	68	176523
Secreted	1	41	52	913	757	175	3733709
Structural	1	7	9	482	417	161	310275

Name	Description	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
hydrolase	enzyme.hydrolase	16	49	15769	7889	685	1514885
PDE	enzyme.phosphodiestererase	23	41	8588	5497	335	1681533
reductase	enzyme.reductase	45	107	30554	16689	2395	2644799
kinase	enzyme.kinase	409	524	107015	46542	3186	12211656
transferase	enzyme.transferase	19	38	3809	2594	250	589297
protease	enzyme.protease	246	373	103371	54213	2731	31253665
lyase	enzyme.lyase	13	23	20075	5086	1278	3177912
p450	enzyme.cytochrome p450	40	53	18009	10538	1953	2617047
isomerase	enzyme.isomerase	8	25	1953	1749	162	2684907
enzyme-other	enzyme.other	1068	1521	98383	59690	9171	56116883
ligase	enzyme.ligase	8	10	252	188	21	60735

So in kinase we have 409 genes:

Name	Description	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
hydrolase	enzyme.hydrolase	16	49	15769	7889	685	1514885
PDE	enzyme.phosphodiestererase	23	41	8588	5497	335	1681533
reductase	enzyme.reductase	45	107	30554	16689	2395	2644799
kinase	enzyme.kinase	409	524	107015	46542	3186	12211656
transferase	enzyme.transferase	19	38	3809	2594	250	589297
protease	enzyme.protease	246	373	103371	54213	2731	31253665
lyase	enzyme.lyase	13	23	20075	5086	1278	3177912
p450	enzyme.cytochrome p450	40	53	18009	10538	1953	2617047
isomerase	enzyme.isomerase	8	25	1953	1749	162	2684907
enzyme-other	enzyme.other	1068	1521	98383	59690	9171	56116883
ligase	enzyme.ligase	8	10	252	188	21	60735

Name	Description	Organism	Sub Class (Major Class)	Orthologs	Observations	Substances	Purchasable	Predicted
ABL1	Tyrosine-protein kinase ABL1	Eukaryotes	kinase (enzyme)	2	1886	994	186	193897
EGFR	Epidermal growth factor receptor	Eukaryotes	kinase (enzyme)	2	5947	3972	550	344091
INSR	Insulin receptor	Eukaryotes	kinase (enzyme)	3	537	486	61	39033
ERBB2	Receptor tyrosine-protein kinase erbB-2	Eukaryotes	kinase (enzyme)	2	1765	1447	121	281210
PDGFRB	Platelet-derived growth factor receptor beta	Eukaryotes	kinase (enzyme)	3	1409	1062	160	968919
IGF1R	Insulin-like growth factor 1 receptor	Eukaryotes	kinase (enzyme)	3	1624	1061	94	89272
KIT	Mast/stem cell growth factor	Eukaryotes	kinase (enzyme)	2	1444	774	192	317447

Minor classes:

Screenshot of the ZINC15 database interface showing minor classes for the ERBB2 gene.

The URL is zinc.docking.org/genes/ERBB2/

The "Biological" dropdown menu is open, showing the "Activity" option selected.

# Orthologs	# Observations	# Substances	# Purchasable	# Predicted
2	1765	121	281210	

Highest Affinity Substances

Four substances are listed with their pKi values and ZINC IDs:

- ZINC13682441: pKi (L_E) : 9.30(0.41)
- ZINC115937836: pKi (L_E) : 9.30(0.39)
- ZINC22935337: pKi (L_E) : 9.15(0.39)
- ZINC22451877: pKi (L_E) : 9.15(0.39)

Chemical structures for each substance are shown below their respective ZINC IDs.

Screenshot of the ZINC15 database interface showing subclasses for the kinase enzyme.

The URL is zinc.docking.org/subclasses/search/?q=kinase

The search term "kinase" is entered in the search bar.

Name	Description	# Genes	# Orthologs	# Observations	# Substances	# Purchasable	# Predictions
kinase	enzyme.kinase	409	524	107015	46542	3186	12211656

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 09:13:26 (434289) in 0.160246s on zinc.docking.org using ZINC15.0.20710003.1

Sp we have 409 genes:

Name	Description	Organism	Sub Class (Major Class)	Orthologs	Observations	Substances	Purchasable	Predicted
ABL1	Tyrosine-protein kinase ABL1	Eukaryotes	kinase (enzyme)	2	1886	994	186	193897
EGFR	Epidermal growth factor receptor	Eukaryotes	kinase (enzyme)	2	5947	3972	550	344091
INSR	Insulin receptor	Eukaryotes	kinase (enzyme)	3	537	486	61	39033
ERBB2	Receptor tyrosine-protein kinase erbB-2	Eukaryotes	kinase (enzyme)	2	1765	1447	121	281210
PDGFRB	Platelet-derived growth factor receptor beta	Eukaryotes	kinase (enzyme)	3	1409	1062	160	968919
IGF1R	Insulin-like growth factor 1 receptor	Eukaryotes	kinase (enzyme)	3	1624	1061	94	89272
KIT	Mast/stem cell growth factor	Eukaryotes	kinase (enzyme)	2	1444	774	192	317447

Activities:

Activities are the best observed affinity (observation) of each substance substance for each gene gene over all orthologs.

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 05:14:40.514096 in 0.03032s on zinc.docking.org using ZINC15 0.20/1000.1

Screenshot of a web browser showing the ZINC database interface at zinc.docking.org/activities/. The page displays a list of substances based on biological activity. The first substance listed is MGAM: Maltase-glucoamylase, intestinal, with a pKi (L.E.) of 8.64 (0.67) and an observation count of 1. Its structure is shown as ZINC170116778, featuring a complex organic molecule with a silicon atom and multiple hydroxyl groups. Below it is another entry for MGAM with a pKi (L.E.) of 8.40 (0.51) and an observation count of 1, labeled ZINC170116766.

Go and apply filter:

100 nm

Screenshot of the same web browser showing the application of a filter. A dropdown menu under the 'Filters' button is open, showing options for 'Affinity' (100Nm, 10Nm, 10Um, 1Nm, 1Um) and 'Availability' (For Sale, Not For Sale). The '100Nm' option is selected. The results are updated to show only substances active at 100 nM or better. The first entry remains the same (MGAM, pKi 8.64), but the second entry (MGAM, pKi 8.40) is no longer visible.

Observations:

Screenshot of the ZINC Observations homepage (zinc.docking.org/observations/home/):

The page title is "Observations". The top navigation bar includes links for WhatsApp, ZINC, Substances, Catalogs, Tranches, Biological, More, and About.

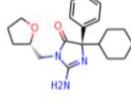
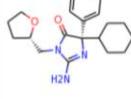
Below the navigation bar is a search bar with tabs for Help, Examples, Browse, Subsets, and a search input field with a "Search" button.

A message states: "Observations connect substances to genes (and vice versa) The best affinity observation for each gene - substance pair is captured in activity."

Acknowledgements, Usage, Why are ZINC results "estimates?", Terms of use, Privacy policy, Supported by NIGMS via GM71896, Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-01 05:17:20 984373 in 0.02963s on zinc.docking.org using ZINC15 0.20210303.1

Second screenshot shows the detailed observations table for BACE2_HUMAN:

Target	Description	Gene / Class / Organism	pKi (L.E.)	Substance	Structure	Citation
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948789		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948791		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948793		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010

Apply filter 10 μM

The screenshot shows a web browser window with the URL zinc.docking.org/observations/subsets/10uM/. The page displays a table of substances targeting BACE2_HUMAN. The columns include Target, Description, Gene / Class / Organism, pKi (L_uE_u), Substance, Structure, and Citation. Three rows are shown, each with the same substance information and a different chemical structure. The structures are complex organic molecules. The citation for all three entries is 49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010.

Target	Description	Gene / Class / Organism	pKi (L _u E _u)	Substance	Structure	Citation
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948789		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948791		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010
BACE2_HUMAN	Beta secretase 2	BACE2 / enzyme / protease (E)	6.77 (0.38)	ZINC000034948793		49463: PMID19959359 Bioorg. Med. Chem. Lett. 2010

Prediction:

The screenshot shows a web browser window with the URL zinc.docking.org/predictions/home/. The page title is "SEA Predictions". It features a navigation bar with links for Help, Examples, Browse, Table, Subsets, and a search bar with a "Search" button. Below the navigation bar, a message states: "SEA predictions connect substances to genes (and vice versa), using the Similarity Ensemble Approach." At the bottom of the page, there is a footer with links for Acknowledgements, Usage, Why are ZINC results "estimates"? Terms of use, Privacy policy, Supported by NIGMS via GM71896, Questions, Discussion, Bug reports, Feature requests, Irwin and Shoichet Labs, and UC Regents. The page also includes a timestamp: Originally generated at 2021-11-01 05:20:00, 149164 in 0.02967s on zinc.docking.org using ZINC15.0.20210303.1.

Result:

The screenshot shows a grid of 12 chemical structures, each with its target protein name and ZINC ID. The targets are: SSTR5 (ZINC1443503872), BIRC2 (ZINC1443504142), BIRC2 (ZINC1443504143), ROCK1 (ZINC1443504175), KCNQ2 (ZINC1443504355), KCNQ2 (ZINC1443504592), BIRC2 (ZINC1443504597), BIRC2 (ZINC1443504599), HSD11B1 (ZINC1443503849), HSD11B1 (ZINC1443503850), SSTR5 (ZINC1443504100), and SSTR5 (ZINC1443504101). Each structure is a complex organic molecule with various functional groups.

ZINC1443504599

ZINC1443504599

In: anodyne boutique for-sale

Google Wikipedia PubMed

Added	Availability	Since	Mwt	logP	Download
2018-08-07	For-Sale	2018-08-07	318.417	1.294	Download

Mol Formula	Rings	Heavy Atoms	Hetero Atoms	Fraction sp ³	Tranche
C18H26N2O3	3	23	5	0.61	DDAE

SMILES: CO[C@H](C(=O)NC1CN(Cc2ccc(C)c2)C1)C1CCCC1

InChI: InChI=1S/C18H26N2O3/c1-23-17(15-3-2-4-15)18(22)19-16-10-20(11-16)9-13-5-7-14(12-21)8-6-13/h5-8,15-17,21H,2-

InChI Key: OUYVMOZGQFCPKO-QGZVFWFLSA-N

Available 3D Representations

Waiting for zinc.docking.org...

Given P value:

The screenshot shows the ZINC1443504599 substance page. At the top, it says "There is no known activity for this compound." Below this, the "SEA Predictions based on ChEMBL 20" section displays the following table:

Gene	Description	Target Class	P-Value	Max Tc
SIGMAR1	Sigma non-opioid intracellular receptor 1	membrane receptor / Membrane-other	19	41
BIRC2	Baculoviral IAP repeat-containing protein 2	enzyme / enzyme-other	45	41

Below the table, there are sections for "Interesting Analogs" and "Framework of this compound". The framework section shows the chemical structure: CC1(C)C2=C1C(=O)N(C3CCCC3)C2Cc4ccccc4O.

Tool Compound:

The screenshot shows the ZINC1443504599 substance page with the "Biological" dropdown menu open. The menu includes options like Entities, Activity, and Tool Compounds. The "Activity" section is currently selected and shows a list of items: Activities, Observations, Predictions, and Tool Compounds.

Screenshot of a web browser showing the ZINC database search results for PDE5A. The search term 'PDE5A' is entered in the search bar.

Gene	Description	ZINC ID	Quality (Ordinal)	Image
PDE5A	cGMP-specific 3',5'-cyclic phosphodiesterase	ZINC00000274067 (In-Stock)	76 (#2)	
PDE5A	cGMP-specific 3',5'-cyclic phosphodiesterase	ZINC00003993855 (In-Stock)	324 (#2)	
PDE5A	cGMP-specific 3',5'-cyclic phosphodiesterase	ZINC000019796168 (In-Stock)	425 (#2)	

Below the table, the Windows taskbar shows the search bar, pinned icons for Google Chrome, File Explorer, and Microsoft Word, and the system tray indicating the date and time (01-11-2021, 17:53).

PDE5A (cGMP-specific 3',5'-cyclic phosphodiesterase)

Screenshot of the ZINC database page for PDE5A.

PDE5A (cGMP-specific 3',5'-cyclic phosphodiesterase)

In: eukaryotic liganded pde purchasable

3D models for docking

- DB2 format db2.gz
- mol2 format mol2.gz
- SDF format sdf.gz

Relations

# Orthologs	# Observations	# Substances	# Purchasable	# Predicted
4	2149	1773	159	286096

Highest Affinity Substances

Browse All (1773)

Establishing secure connection...

Windows taskbar at the bottom showing the search bar, pinned icons for Google Chrome, File Explorer, and Microsoft Word, and the system tray indicating the date and time (01-11-2021, 17:53).

Drugs Substances

Screenshot of the ZINC website showing Drugs Substances. The page displays four cards for Vardenafil, Sildenafil, Cialis, and Stendra.

ZINC18324776 Vardenafil
C[C@H]1[C@@H](C[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C)C(=O)N5[C@H]6[C@H]7[C@H]([C@H]6C(=O)N8CC[C@H]8C)C=C7C

ZINC19796168 Sildenafil
CC[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1C

ZINC3993855 Cialis
C[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1C

ZINC11677857 Stendra
CC[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1Cc5ccccc5

In-Stock Substances [Browse All \(108\)](#)

Screenshot of the ZINC website showing In-Stock Substances. The page displays four cards for Vardenafil, ZINC45336750, ZINC3963610, and ZINC589513.

ZINC18324776 Vardenafil
C[C@H]1[C@@H](C[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C)C(=O)N5[C@H]6[C@H]7[C@H]([C@H]6C(=O)N8CC[C@H]8C)C=C7C

ZINC45336750
CC[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1C

ZINC3963610
CC[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1C

ZINC589513
CC[C@H]1[C@H]2[C@H]3[C@H]([C@H]2C(=O)N4CC[C@H]4C)C=C3C[C@H]1Cc5ccccc5

Natural Products Substances [Browse All \(18\)](#)

Natural Products Substances

The screenshot shows a web browser window with the ZINC12 homepage. The URL in the address bar is zinc.docking.org/genes/PDE5A/. The page title is "Natural Products Substances". A navigation bar at the top includes links for "ZINC", "Substances", "Catalogs", "Tranches", "Biological", "More", and "About". Below the title, there is a button labeled "Browse All (18)". The main content area displays four chemical structures in cards:

- ZINC4199939: Sophoflavescenol. SMILES: O=C1C(O)=CC2=C(C=C1)C(O)=CC=C2C3=CC=CC=C3
- ZINC2107885: A complex polycyclic compound. SMILES: CCCC1=C2C(=O)N3C4=C(C=C3C=C4)C=C5C=C(C=C5)C=C2
- ZINC2107888: A complex polycyclic compound. SMILES: CCCC1=C2C(=O)N3C4=C(C=C3C=C4)C=C5C=C(C=C5)C=C2
- ZINC34199141: Icariside II. SMILES: CC=CC[C@H]1[C@@H](O)[C@H](O)[C@H](O)[C@H](O)[C@H]1Oc2cc3c(cc2O)C(=O)c4ccccc4N3

Below the cards, there is a section titled "Organism" with a search bar containing "Type here to search". The status bar at the bottom right shows "17:54 01-11-2021".

Go to zinc 12 homepage:

The screenshot shows the ZINC12 homepage. The URL in the address bar is zinc12.docking.org/. The page features a large "ZINC 12" logo. A message encourages users to switch to ZINC15. The main content area includes a "Molecule of the Minute" section for entry 95115733, which shows a complex polycyclic compound. A search bar at the bottom allows users to enter "ZINC ID, Drug Name, SMILES, Catalog, Vendor Code, Target" and includes links for "Structure/Draw", "Physical Properties", "Catalogs & Vendors", "ZINC IDs", "Targets", "Rings", and "Combination". The status bar at the bottom right shows "17:57 01-11-2021".

Go to search properties:

The screenshot shows the ZINC12 website with the URL <https://zinc12.docking.org/search/properties>. The main navigation bar includes links for About, Search, Subsets, Help, Social, and a dropdown menu for 'By: Text Structure Properties Catalogs'. A message box highlights the 'Properties' link. Below the navigation, there is a text area with a message about ZINC15 being superior to ZINC12. To the right, there is a 'Molecule of the Minute' section featuring a complex organic molecule with labels for Mo, Et, and H. At the bottom, there is a search bar and a 'Run Query' button.

This screenshot shows the same ZINC12 website as above, but with various filters applied in the search interface. The 'Properties' tab is selected in the top navigation. The search parameters include: 50 items per page, Format: Overview, Representations: Default, Purchasability: Purchasable, and a Predefined Subset set to 'Everything'. The search results are displayed as a series of sliders for different chemical properties: Molecular Weight [g/mol] (150.52 to 301.04), xlogP (-0.05 to 2.10), Net Charge (-1 to 1), Rotatable Bonds (22 to 29), Polar Surface Area [Å²] (0.00 to 200.00), Hydrogen Donors (0 to 10), Hydrogen Acceptors (0 to 20), Polar Desolvation [kcal/mol] (-400.00 to 1.00), and Apolar Desolvation [kcal/mol] (-100.00 to 40.00). A 'Run Query' button is visible at the bottom right of the search interface.

S WhatsApp X | ZINC X | Search Results | ZINC Is Not Com X + v - ⌂ X

← → 🔒 zinc12.docking.org/results/properties

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory docking.org Not Authenticated – sign in

Active cart: Temporary Cart (0 items)

ZINC 12

About Search Subsets Help Social Quick Search Bar... Go

Results Query Details

Back 1 Next Page Size: 50 Overview Default Purchasable Add All Refresh

1. 4027431 2. 4027430 3. 2123961

4. 2123959 5. 37904498 6. 37904499

Type here to search

Windows Taskbar: Type here to search, Start button, File Explorer, Google Chrome, Microsoft Edge, File Manager, File Explorer, 18:09, ENG, 01-11-2021, 21

Practical No: Experiment no.2 A)

Aim: Element analysis using chemsketch.

Theory: ACD/ChemSketch Freeware is a drawing package that allows you to draw chemical structures including organics, organometallics, polymers, and Markush structures. It also includes features such as calculation of molecular properties (e.g., molecular weight, density, molar refractivity etc.), 2D and 3D structure cleaning and viewing, functionality for naming structures (fewer than 50 atoms and 3 rings), and prediction of logP. ChemSketch is a comprehensive structure editor with a variety of tools and functionality that ease the communication of scientific and chemical information. It can generate names from molecular structure, molecular properties from chemical structure. It can draw molecular structures, generate structures from InChI or SMILES strings, or copy/paste from ChemDraw, Insert pre-drawn templates of amino acids, aromatics, steroids, sugars, and more. One can search the Dictionary of >170, 000 systematic, trivial, and trade names, draw reactions and complex chemical schema (including biotransformation maps), use graphical templates and tools to communicate chemistry and chemical biology concepts (e.g., chemical bond types, Lewis structures, molecular orbitals, Newman projections, peptide sequences, and more).

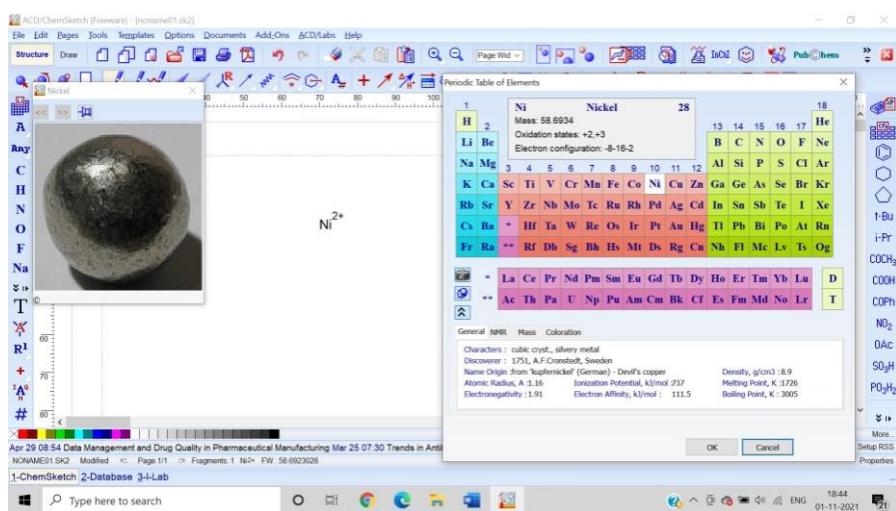
Methodology:

My disease: Glioma

My drug:- Gliadel, Carmustine

Chemsketch:

Go to Chem Sketch and in periodic table select one element:



Nickel:

Mass: 58.6934

Oxidation States: +2 , +3

EC: -8 -16 -2

Roll No: BID 19006

76

SEARCH FOR
gliadel

Treating this as a text search.

COMPOUND BEST MATCH

Carmustine; 154-93-8; 1,3-Bis(2-Chloroethyl)-1-Nitrosourea; BCNU; Carmustin; Gliadel; Nitrumon; Carmubris; ...

Compound CID: 2578

MF: C5H9Cl2N3O2 MW: 214.05g/mol

IUPAC Name: 1,3-bis(2-chloroethyl)-1-nitrosourea

Isomeric SMILES: C(C)NC(=O)N(CCC)N=O

InChIKey: DLGOEMSEDOSKAD-UHFFFAOYSA-N

InChI: InChI=1S/C5H9Cl2N3O2/c6-1-3-8-5(11)10(9-12)4-2-7/h1-4H2,(H,8,11)

Create Date: 2005-03-25

[Summary](#) [Similar Structures Search](#) [Related Records](#) [PubMed \(MeSH Keyword\)](#)

Compounds (1)	Substances (16)	Literature (144)
Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary		

Type here to search

00:57 02-11-2021 [23]

SEARCH FOR
gliadel

Treating this as a text search.

COMPOUND BEST MATCH

Carmustine; 154-93-8; 1,3-Bis(2-Chloroethyl)-1-Nitrosourea; BCNU; Carmustin; Gliadel; Nitrumon; Carmubris; ...

Compound CID: 2578

MF: C5H9Cl2N3O2 MW: 214.05g/mol

IUPAC Name: 1,3-bis(2-chloroethyl)-1-nitrosourea

Isomeric SMILES: C(C)NC(=O)N(CCC)N=O

InChIKey: DLGOEMSEDOSKAD-UHFFFAOYSA-N

InChI: InChI=1S/C5H9Cl2N3O2/c6-1-3-8-5(11)10(9-12)4-2-7/h1-4H2,(H,8,11)

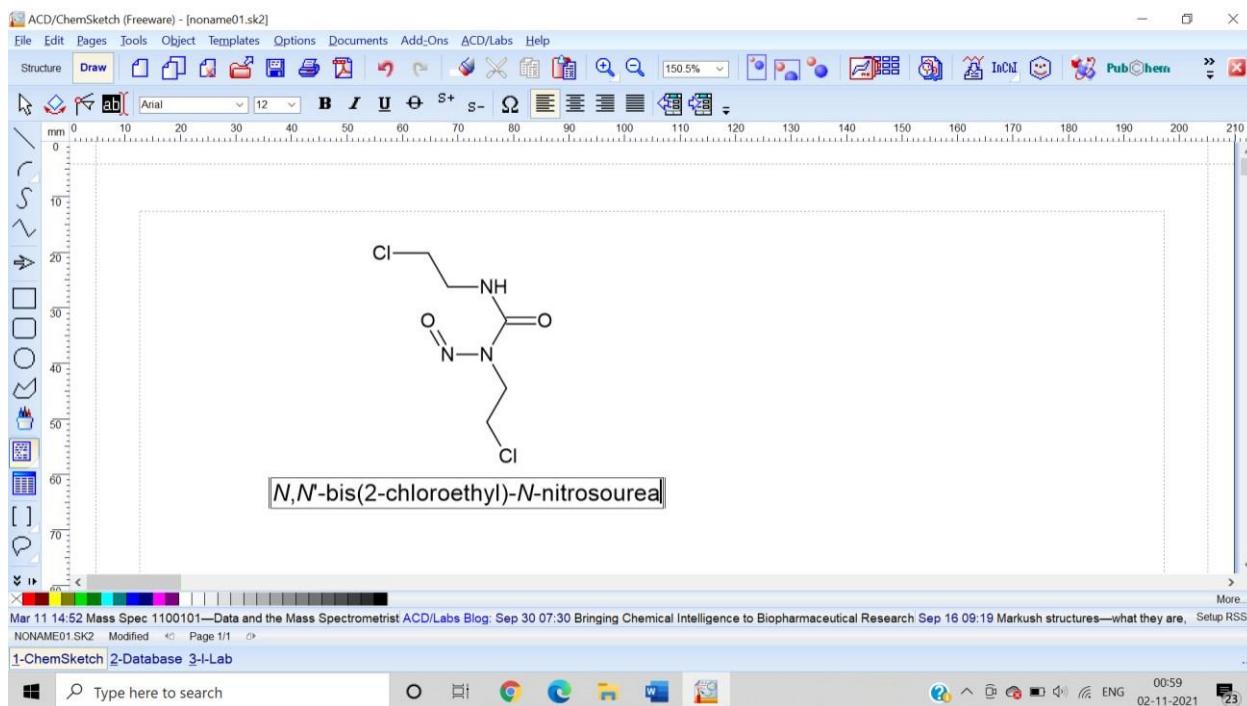
Create Date: 2005-03-25

[Summary](#) [Similar Structures Search](#) [Related Records](#) [PubMed \(MeSH Keyword\)](#)

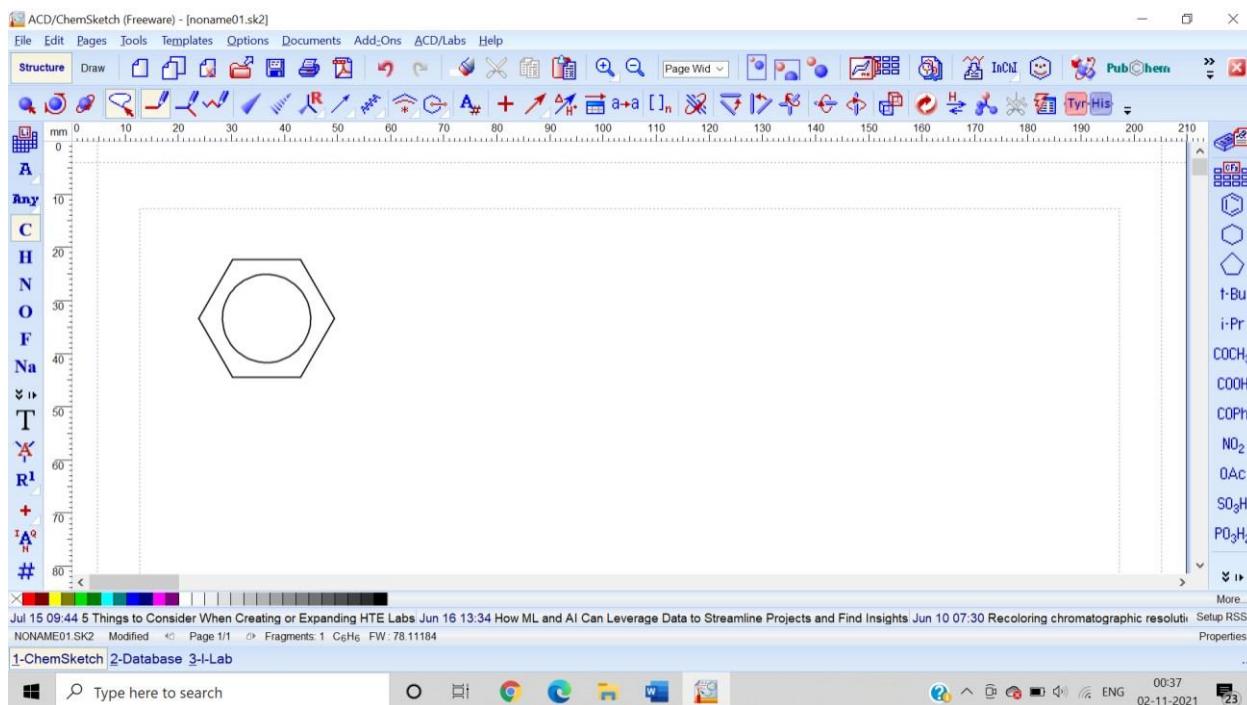
Compounds (1)	Substances (16)	Literature (144)
Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary		

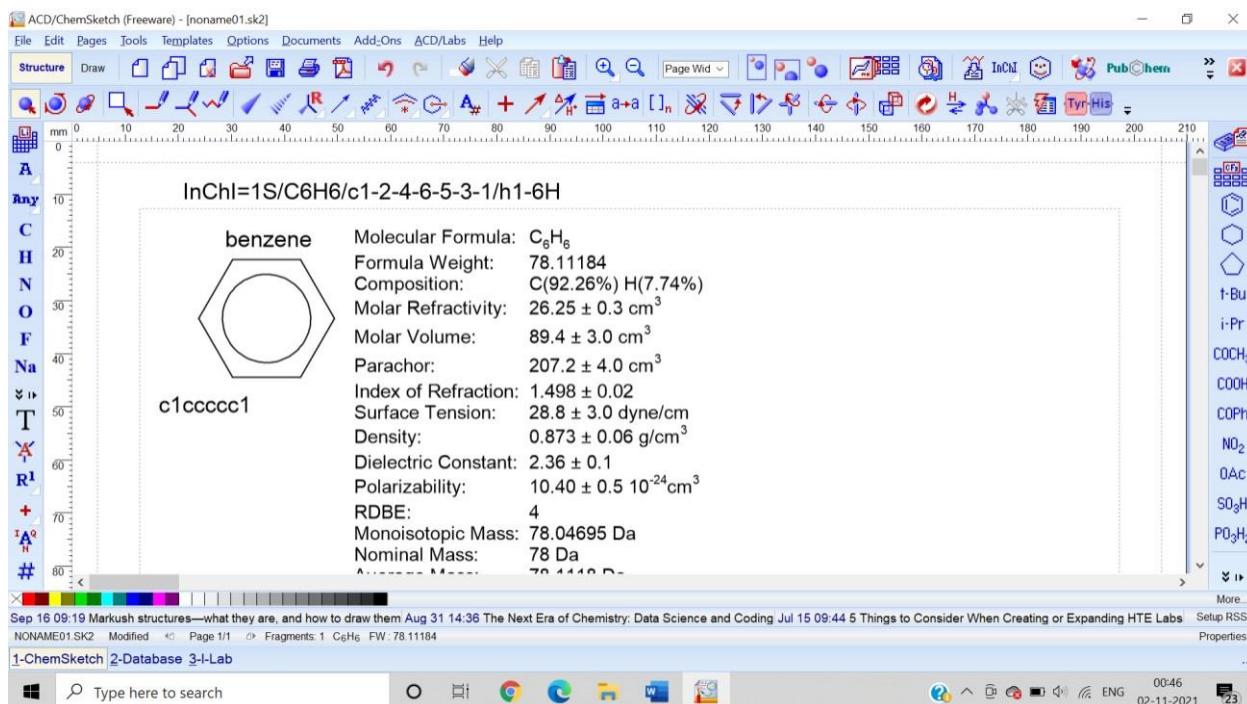
Type here to search

00:58 02-11-2021 [23]

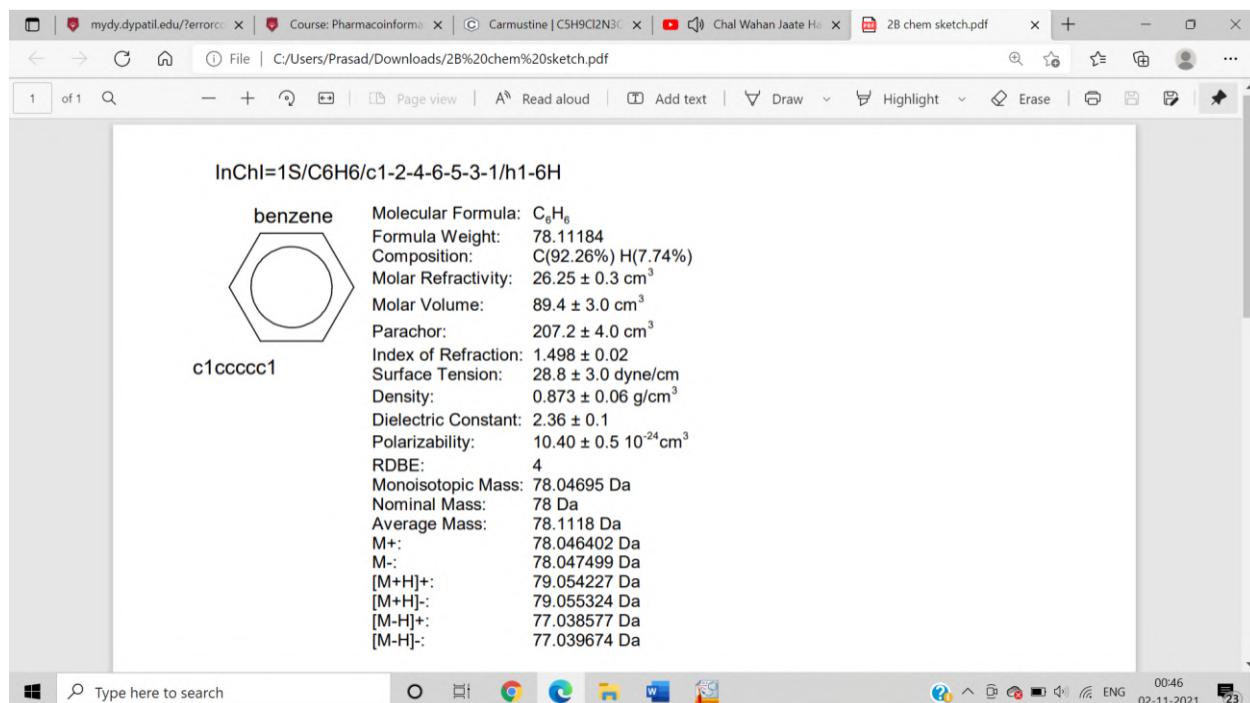


1) Aromaticity:-

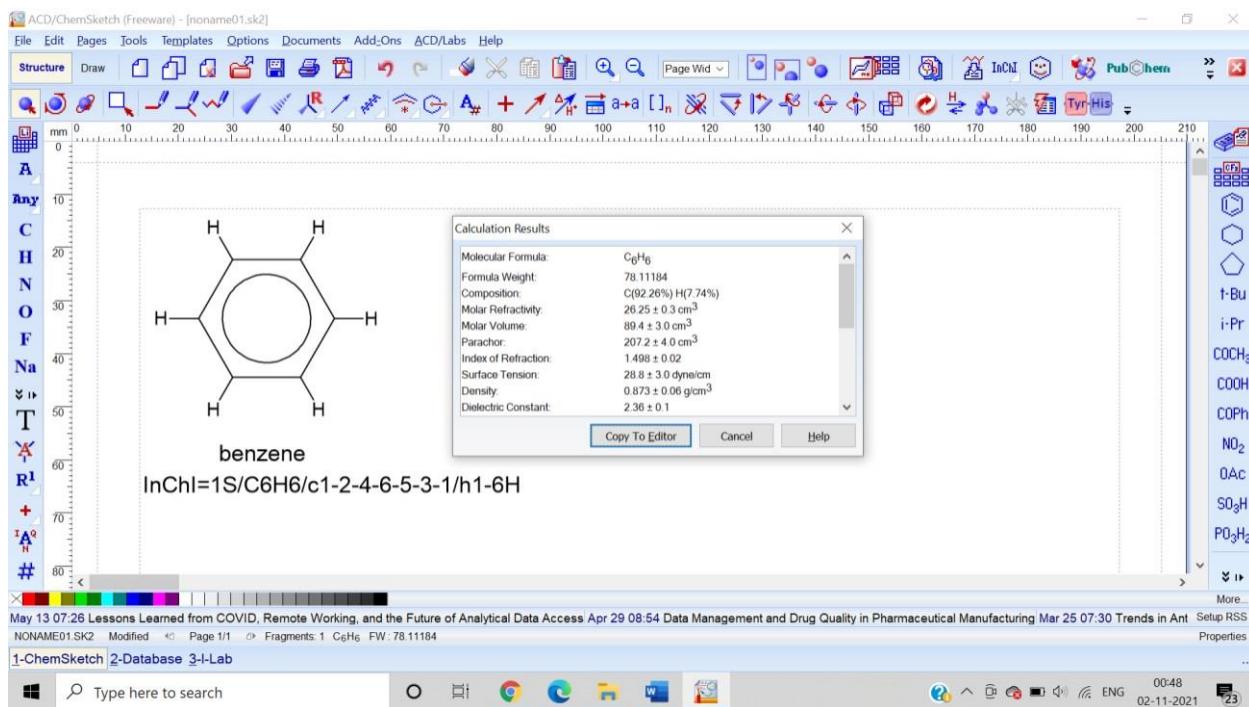
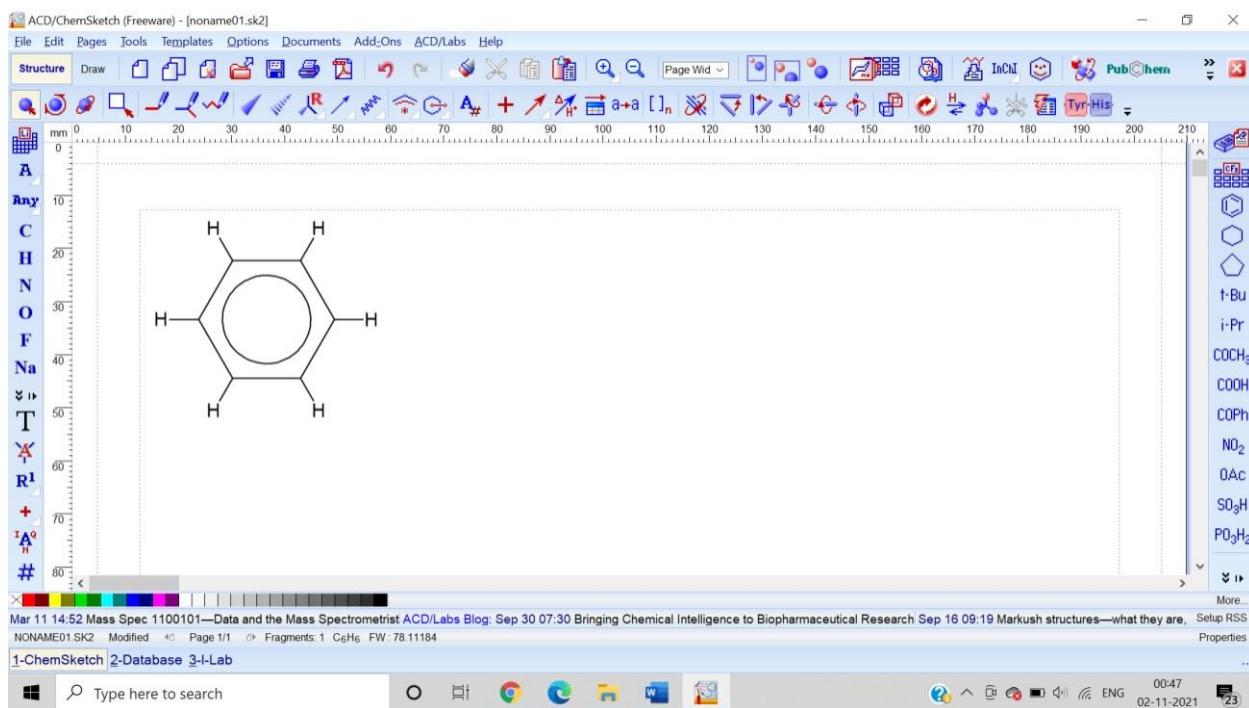




Save it in PDF formate:-

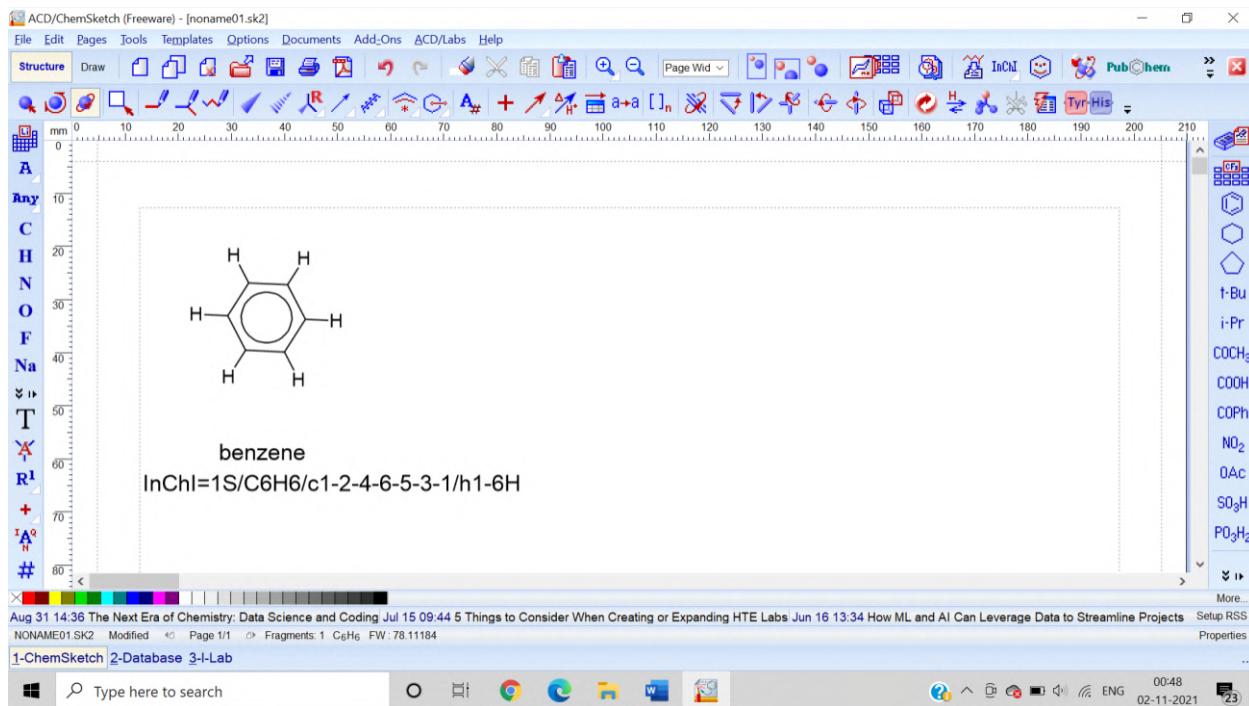


2) Explicit hydrogen

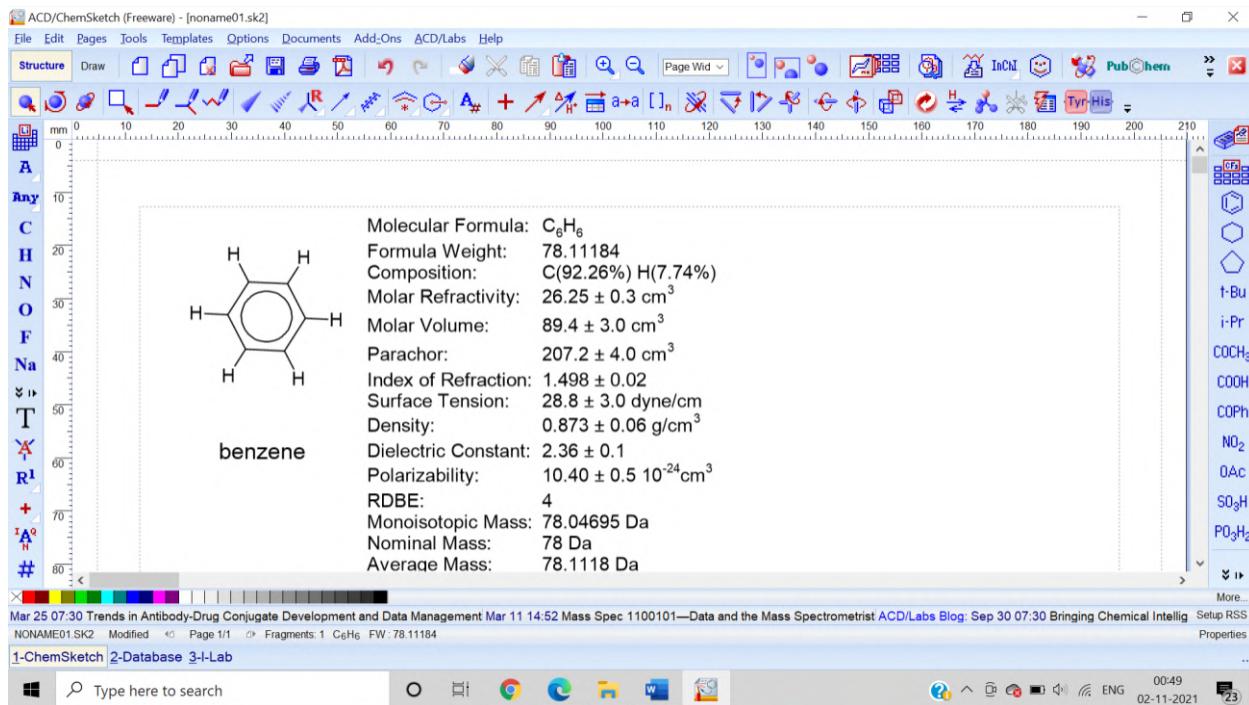


3) 3D optimization:-

Result:



Conclusion:



Practical no.2 B

Aim: Introduction to chemsketch.

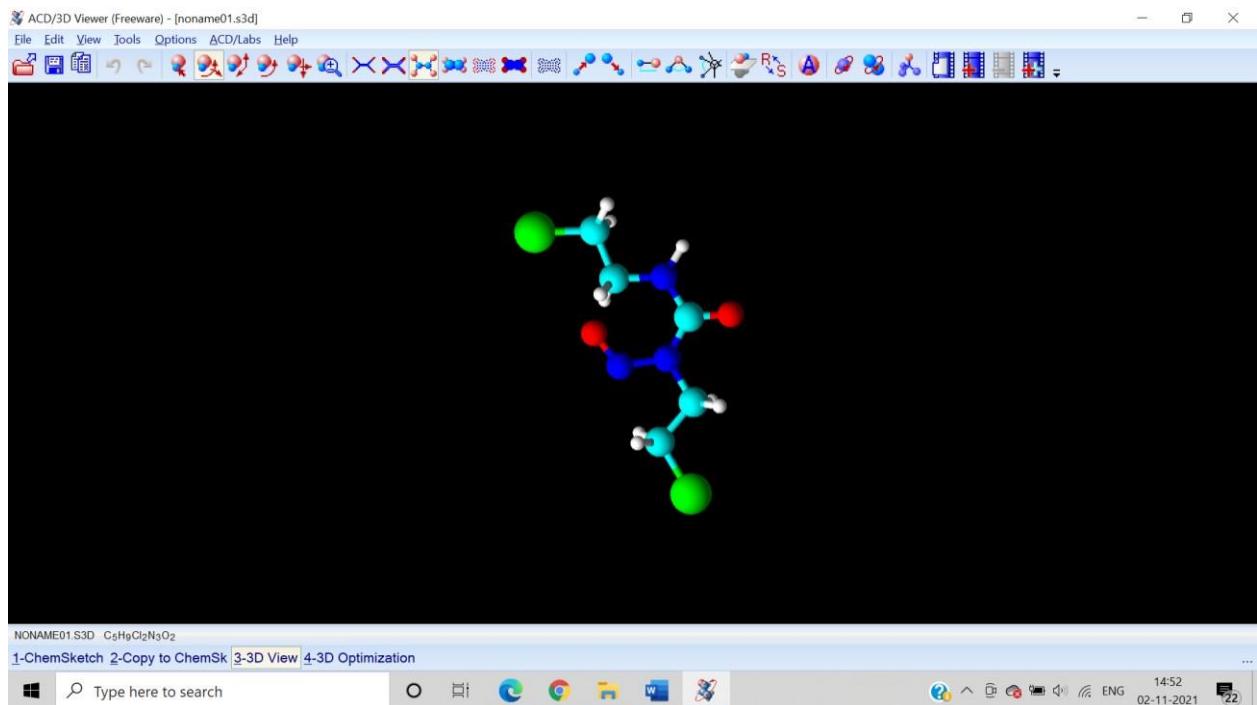
Theory:

ACD/ChemSketch Freeware is a drawing package that allows you to draw chemical structures including organics, organometallics, polymers, and Markush structures. It also includes features such as calculation of molecular properties (e.g., molecular weight, density, molar refractivity etc.), 2D and 3D structure cleaning and viewing, functionality for naming structures (fewer than 50 atoms and 3 rings), and prediction of logP. ChemSketch is a comprehensive structure editor with a variety of tools and functionality that ease the communication of scientific and chemical information. It can generate names from molecular structure, molecular properties from chemical structure. It can draw molecular structures, generate structures from InChI or SMILES strings, or copy/paste from ChemDraw, Insert pre-drawn templates of amino acids, aromatics, steroids, sugars, and more. One can search the Dictionary of >170, 000 systematic, trivial, and trade names, draw reactions and complex chemical schema (including biotransformation maps), use graphical templates and tools to communicate chemistry and chemical biology concepts (e.g., chemical bond types, Lewis structures, molecular orbitals, Newman projections, peptide sequences, and more).

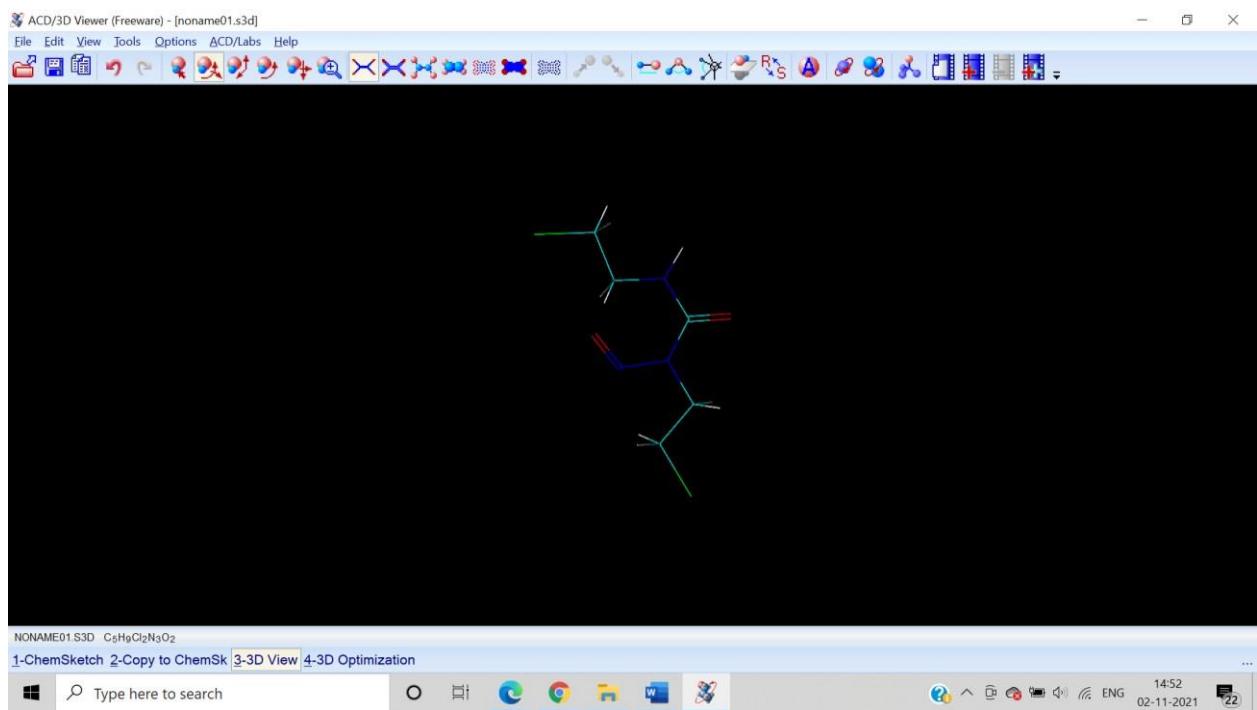
Methodology:

Open the ChemSketch software, sketch the drug molecule using the given tools.

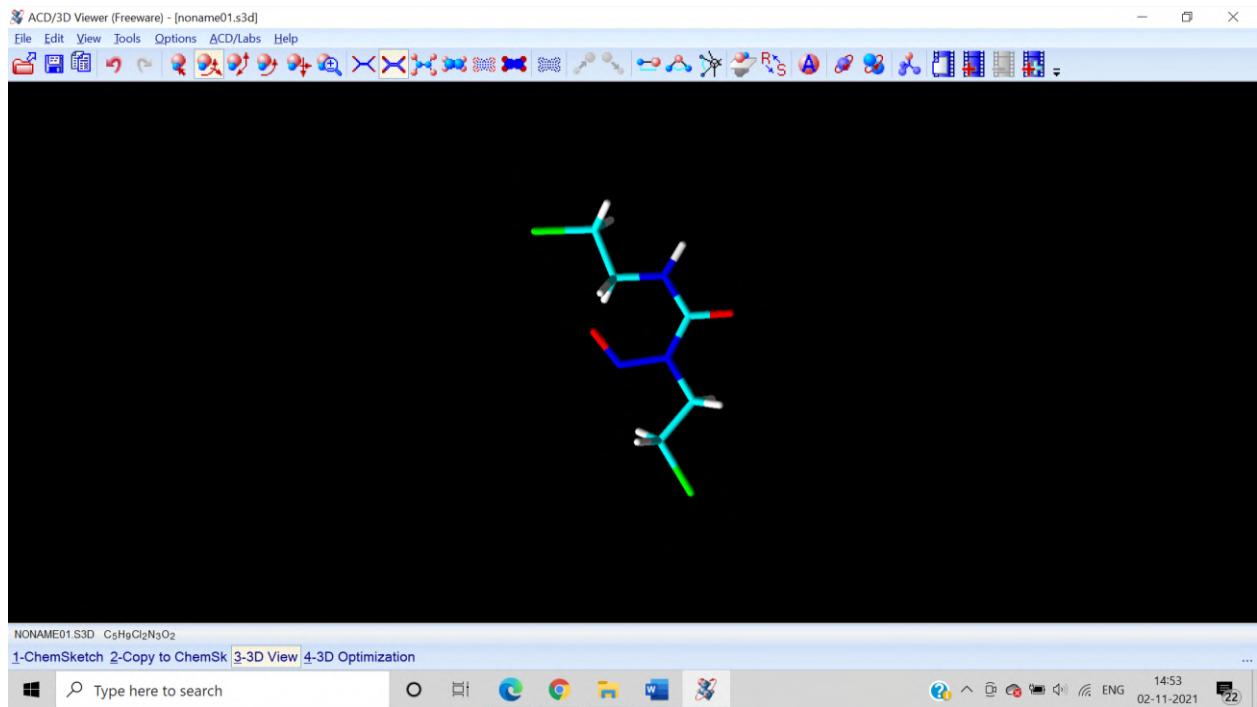
3D view:-



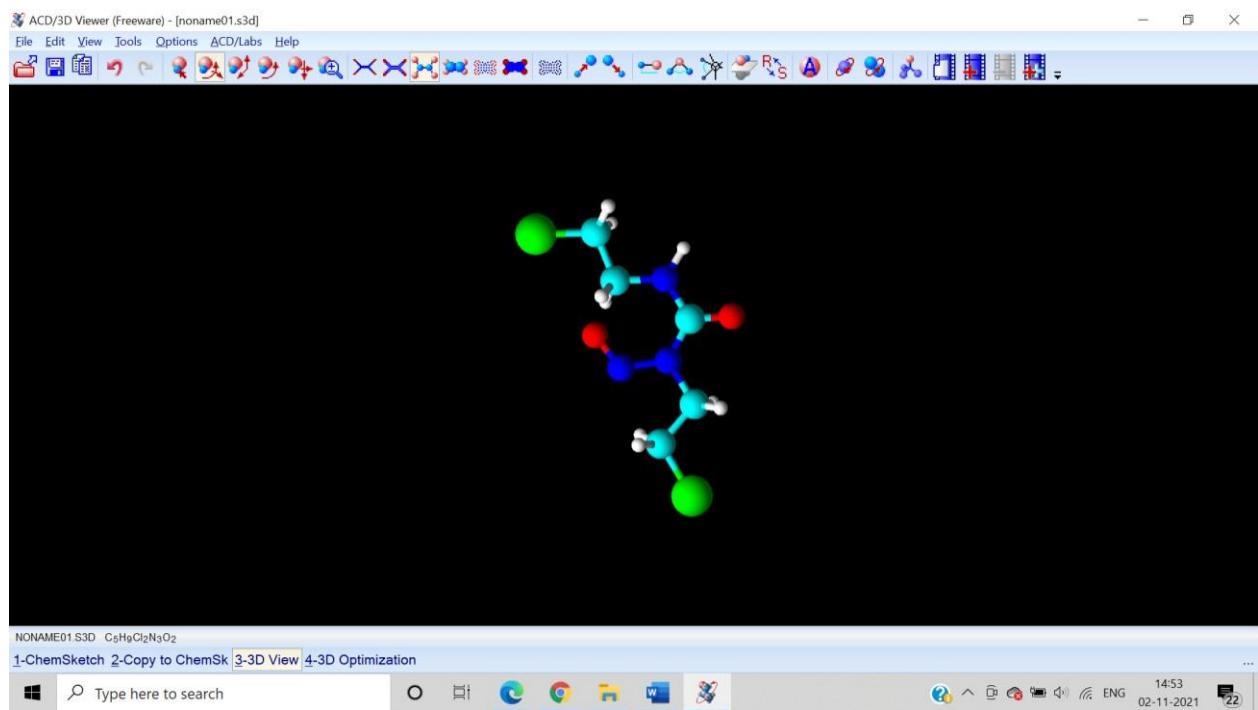
1.wireframe:-



2.sticks:-

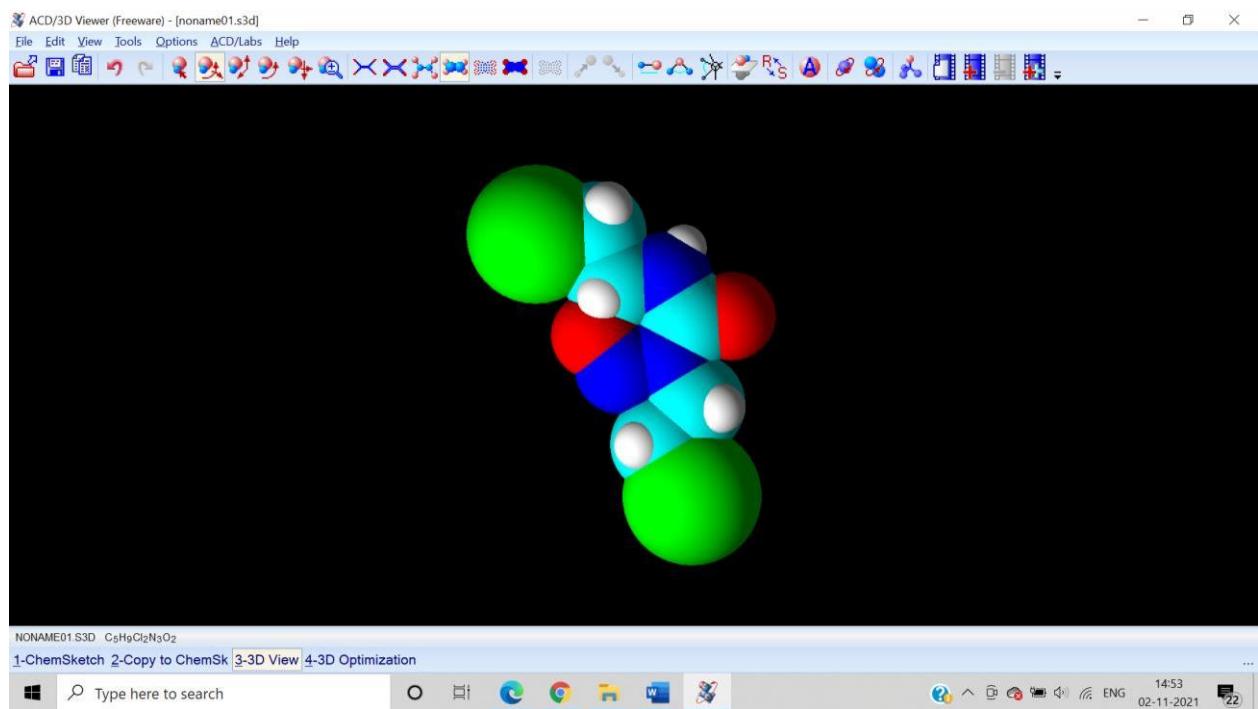


3. balls and sticks:-

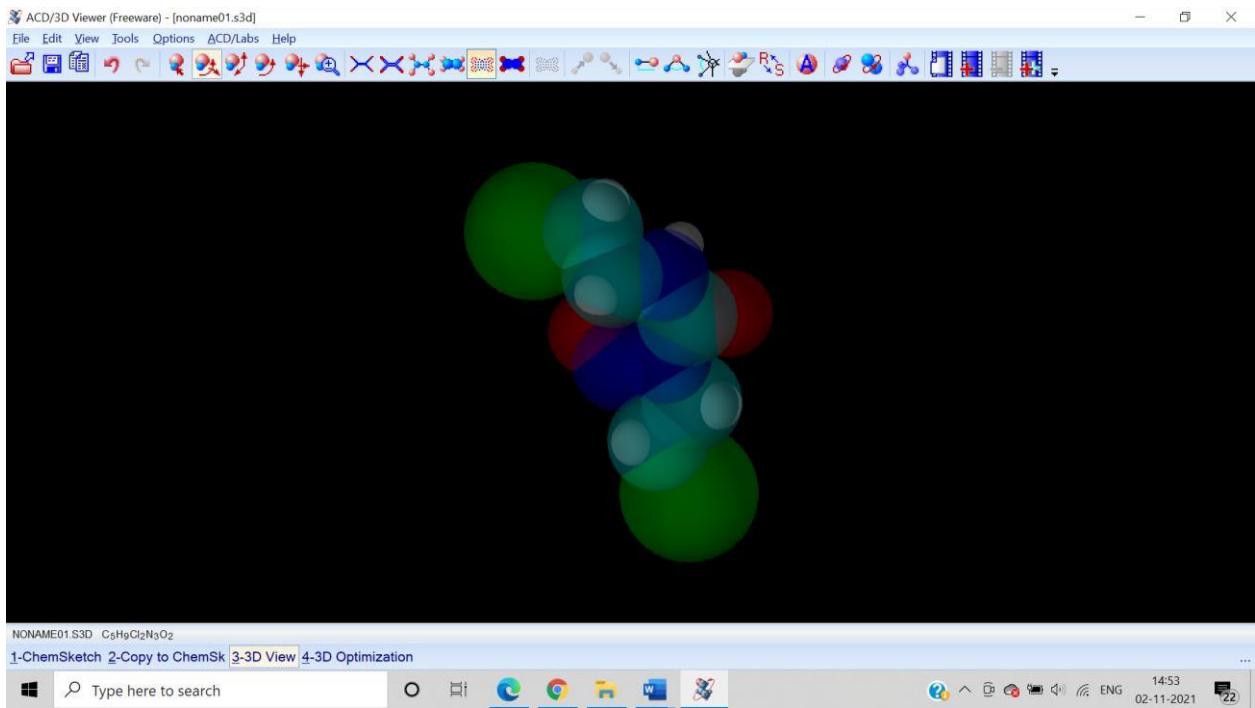


4. Spacefill:-

→

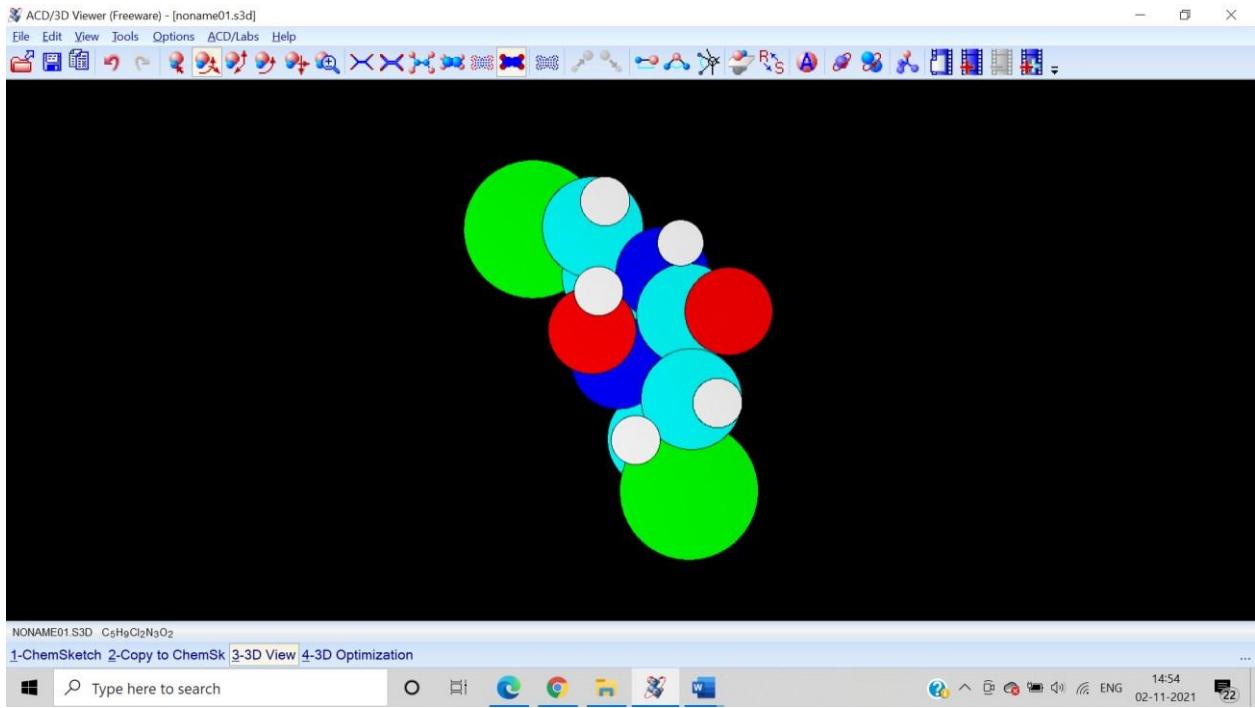


5. dots only

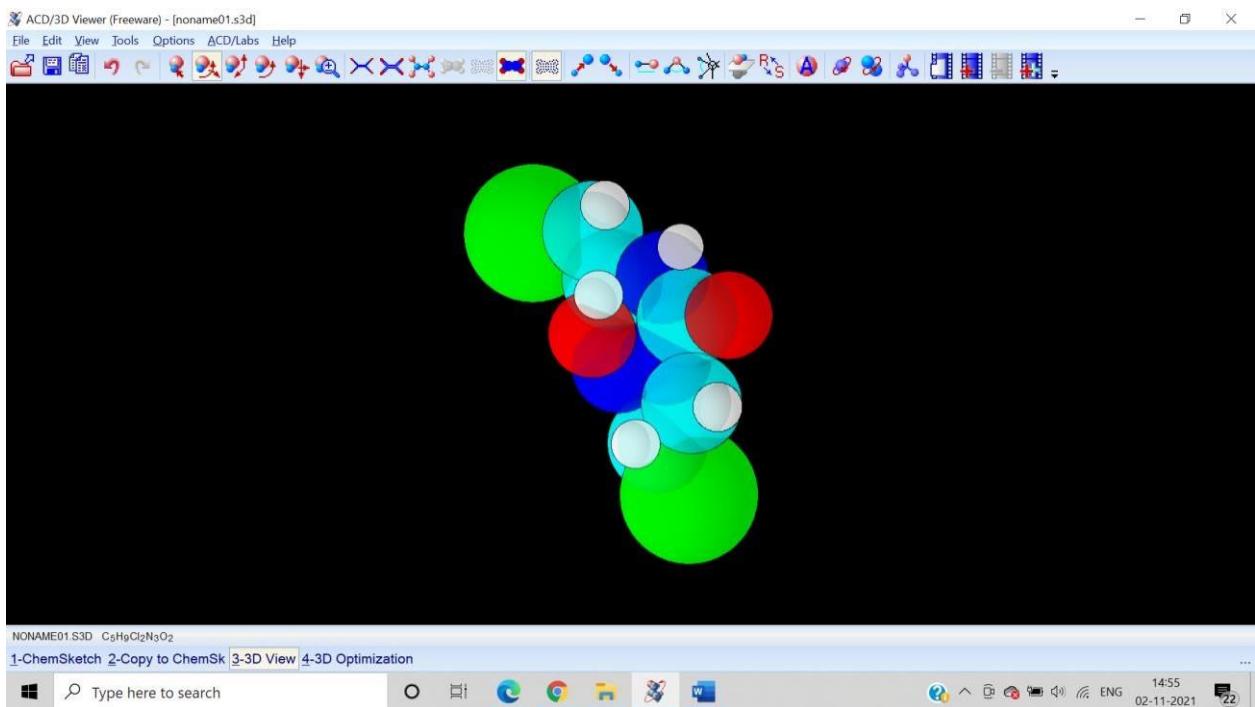


6. Disks:-



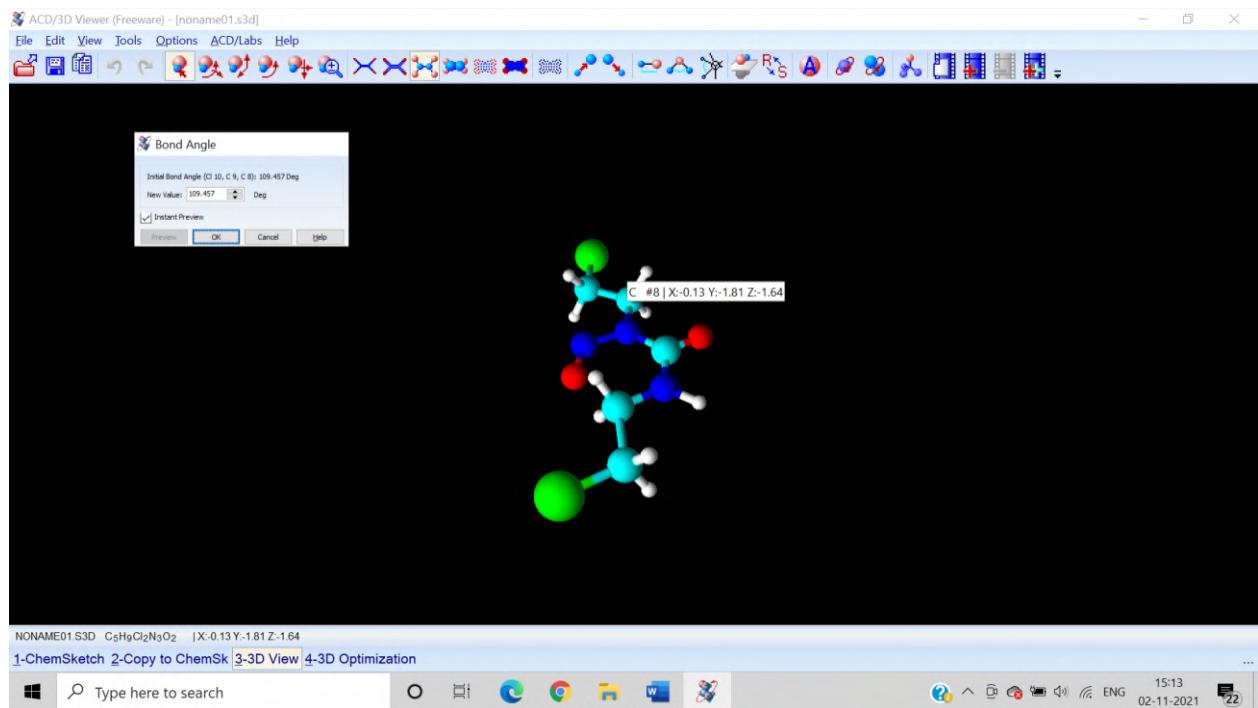


7. with dots:-

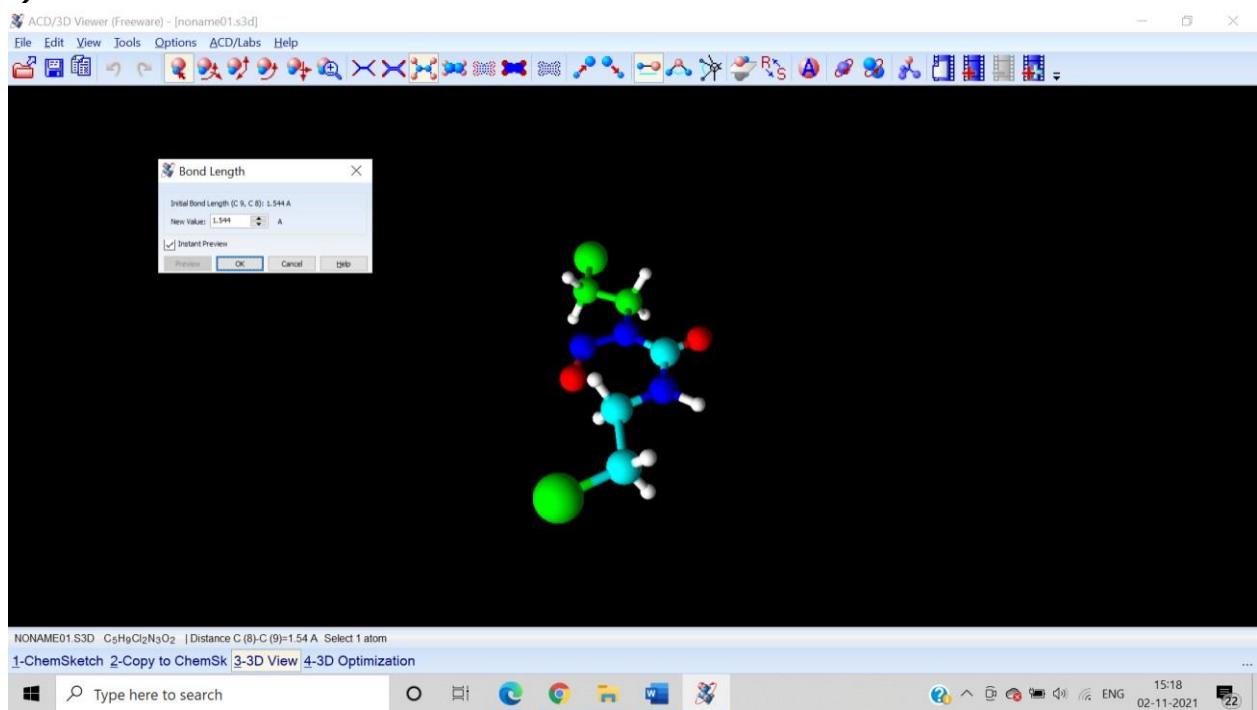


8. bond angle:-

→



9. bond length:-



PRACTICAL NO: 2C

AIM: Sketch a chemical structure using Marvin sketch.

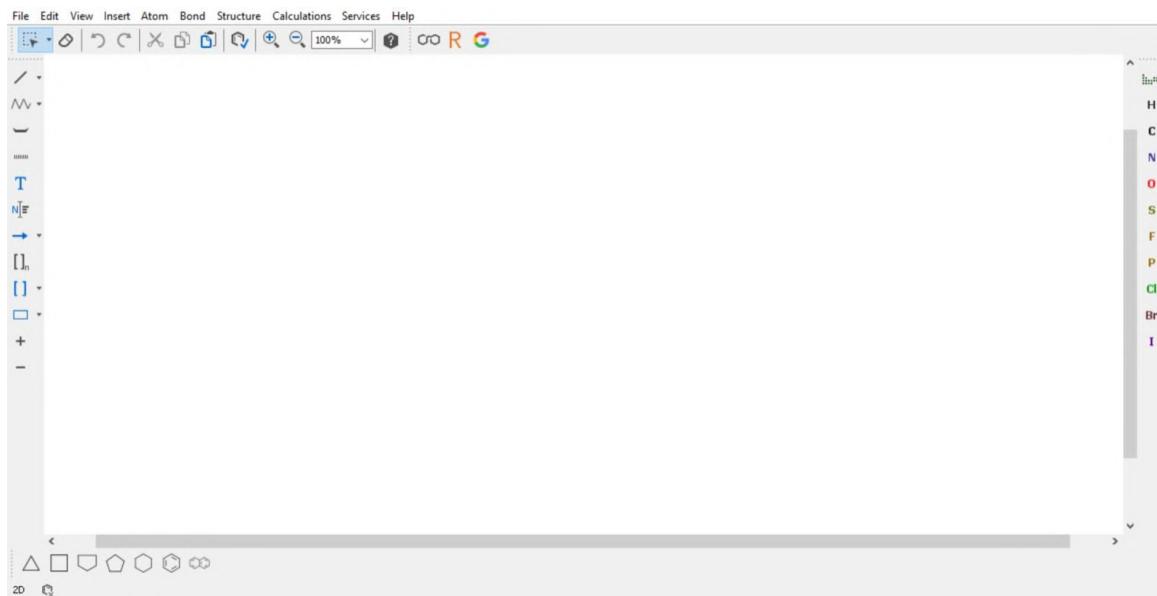
Theory:

MarvinSketch features an extensive set of functionalities to enable the fast and accurate drawing of chemical compounds, reactions, Markush structures and query molecules. Furthermore, MarvinSketch has built-in structure and valence checkers to provide guidance, and integrated property calculators to pull live results - upon your request. Not only does MarvinSketch translate chemistry into a digital environment, it also supports the widest selection of industrially acknowledged standard chemical file formats. It has built in cleaning and standardizer functions of the molecule in 2D and 3D, supports both MDL and Daylight query features in the same query and advanced query features for drawing - wide range of functional groups and templates. Additionally, it can predict pKa, calculate logD, logP and many other properties, perform Markush enumeration, see the structure source for the many file formats Marvin supports, draw reaction mechanisms, peptides, polymers, complexes, or mixtures under less than 20 seconds using graphical arrows, built-in templates and abbreviations

Methodology:

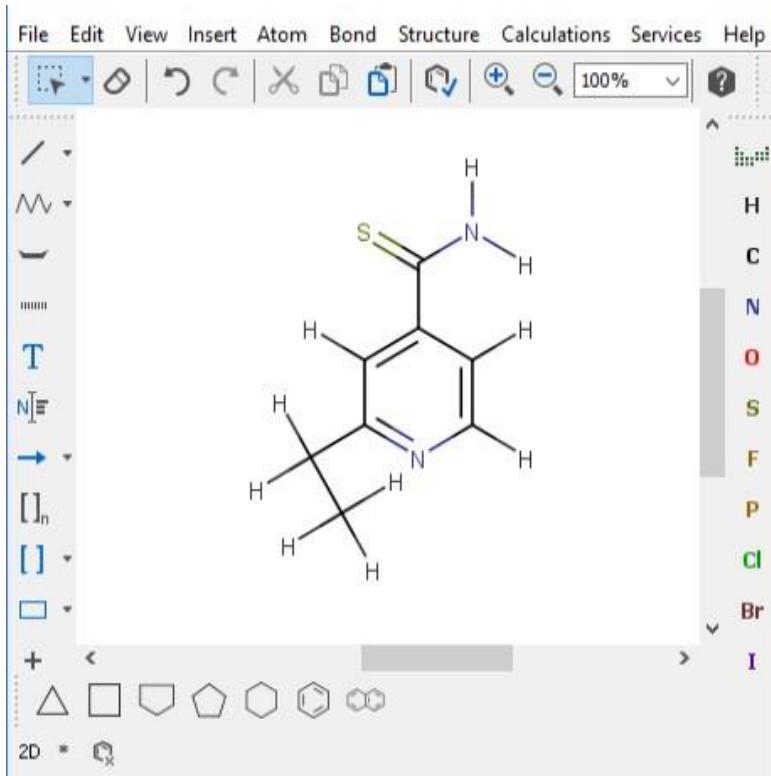
1. Structure>Clean 2D>Clean in 2D
2. Structure>Structure to name>Dialogue Box opens>Ok
3. Structure>Check Structure>New window appears displaying the following result.
4. Calculations>Elemental Analysis>Dialogue Box appears>OK
5. Calculations>Other>H Bond Donor/Acceptor>Dialogue box appears>OK
6. Calculations>Other>Huckel Analysis>Dialogue box appears>OK
7. Calculations>Other>Refractivity>Dialogue box appears>OK
8. Calculations>Geometry>Topology Analysis>Dialogue box appears>OK
9. Calculations>Geometry>Polar Surface Area (2D)>Dialogue box appears>OK
10. Calculations>Geometry>Molecular Surface Area (3D)>Dialogue box appears>OK
11. Calculations>Protonation>Pka>Dialogue box appears>OK
12. Calculations>Partitioning>logP>Dialogue box appears>OK
13. Calculations>Charge>Charge>Dialogue box appears>Ok

14. Calculations>Charge>Polarizability>Dialogue box appears>OK
15. Calculations>Conformations>Molecular Dynamics>Dialogue box appears>OK
16. Calculations>Isomers>Resonance>Dialogue box appears>OK
17. Calculations>Conformation>Conformers>Dialogue box appears>OK

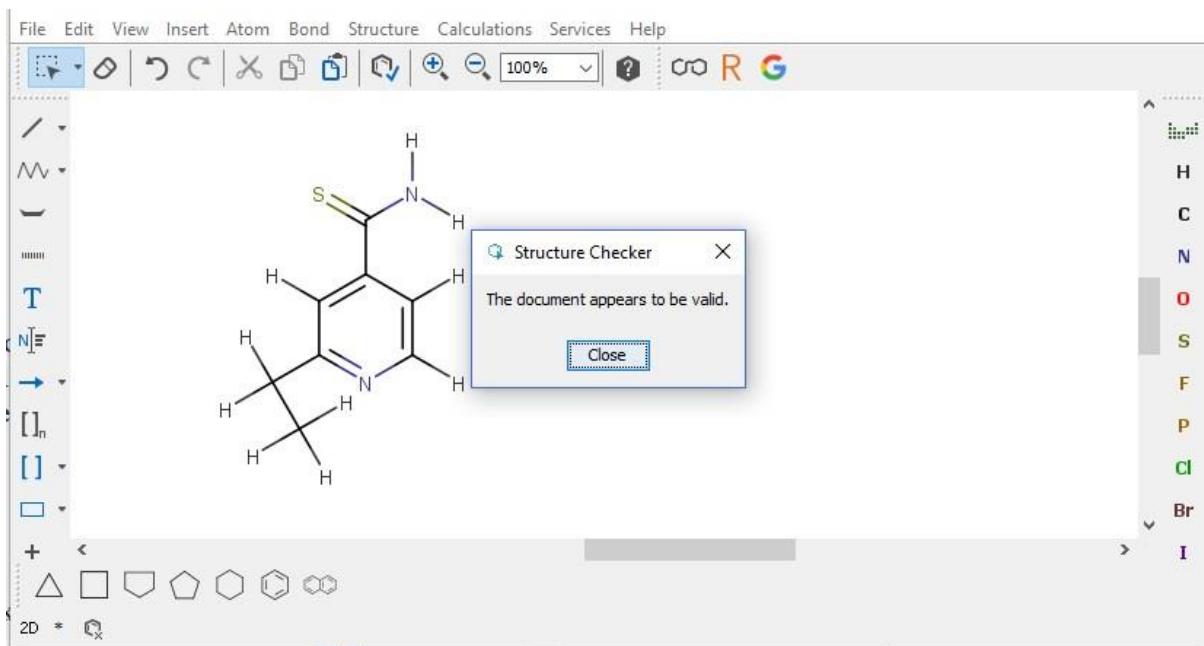


RESULT:

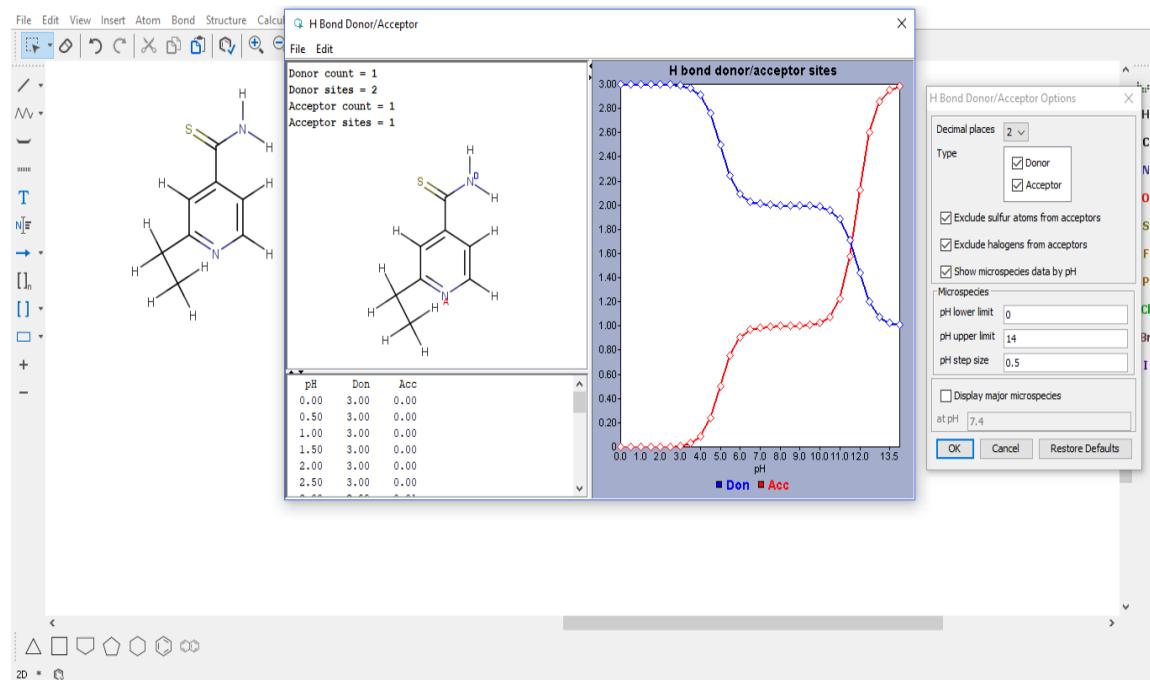
→ Clean in 2D



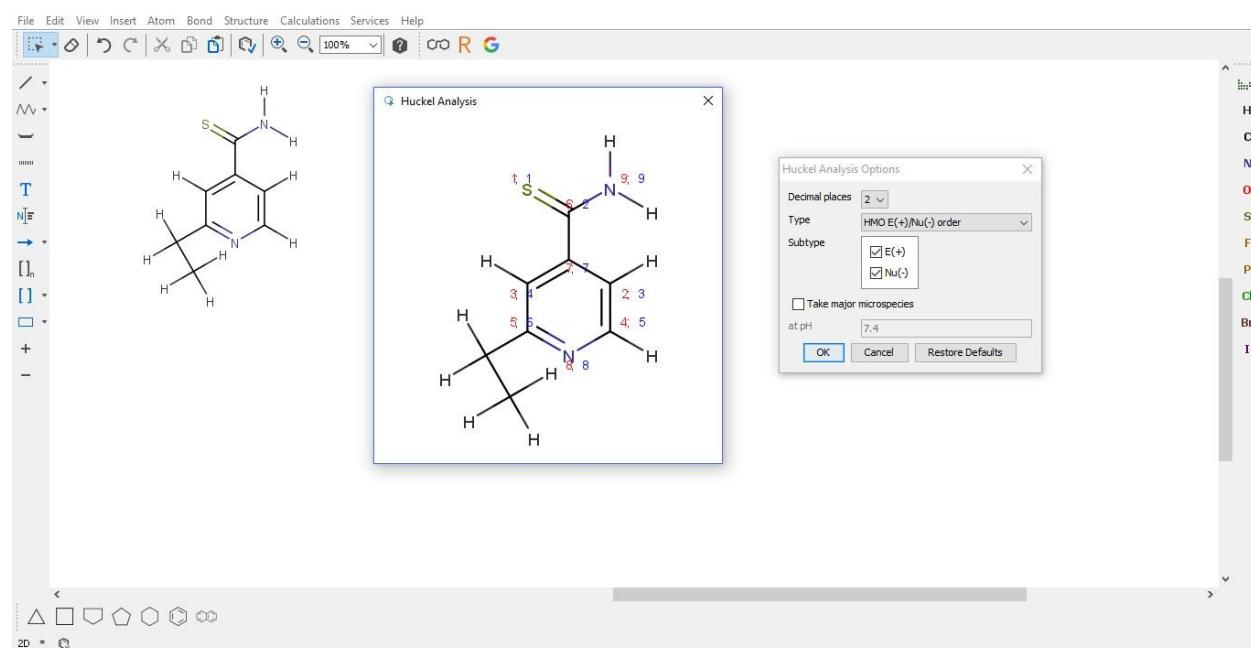
→Check Structure:



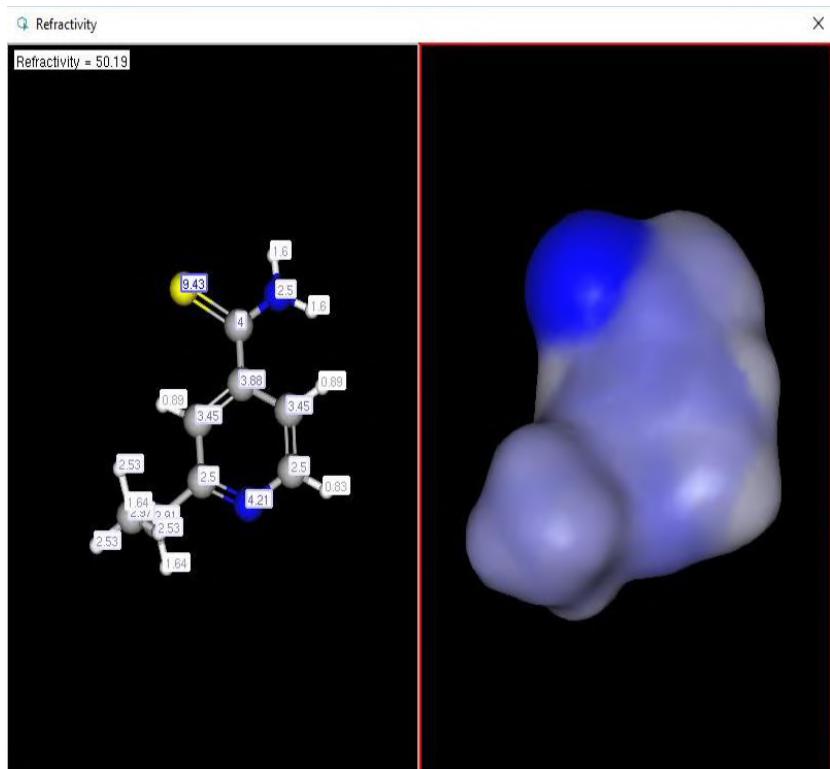
→H Bond Donor/Acceptor



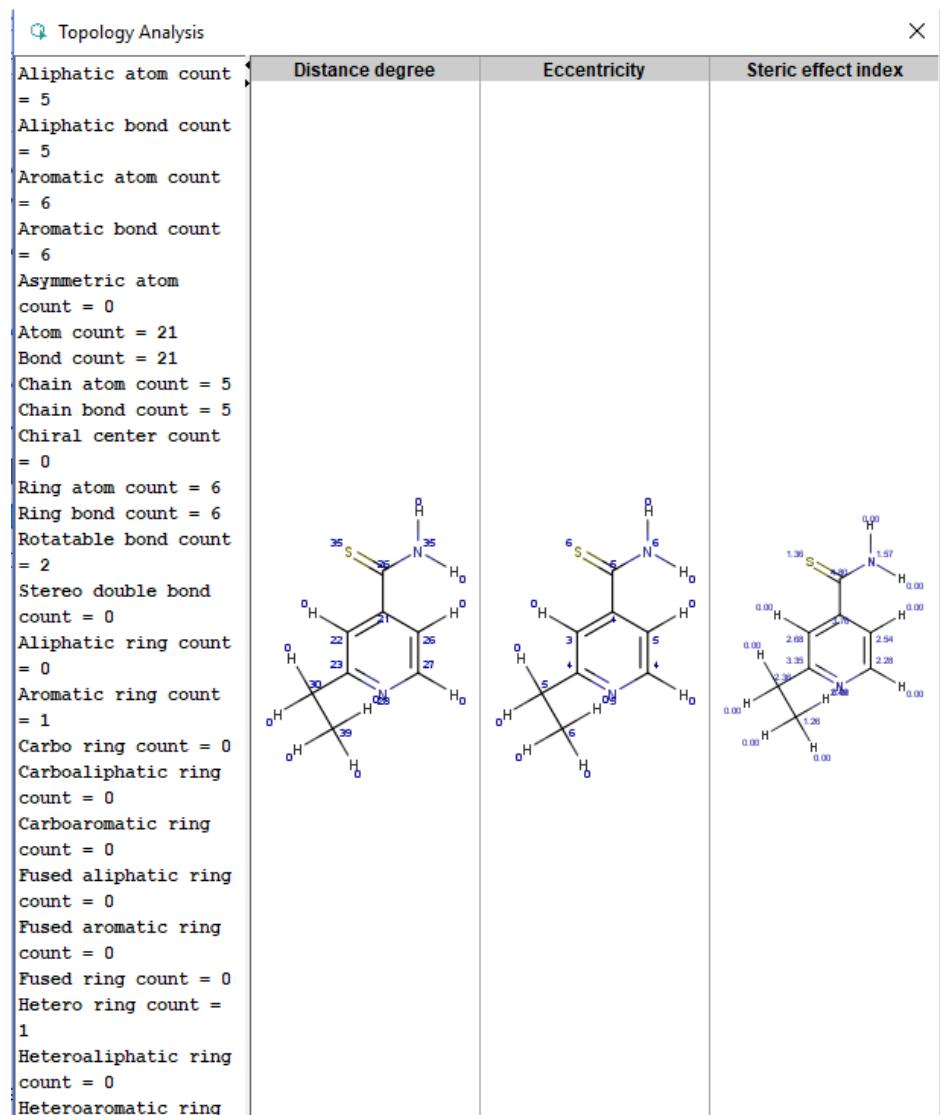
→Huckel Analysis



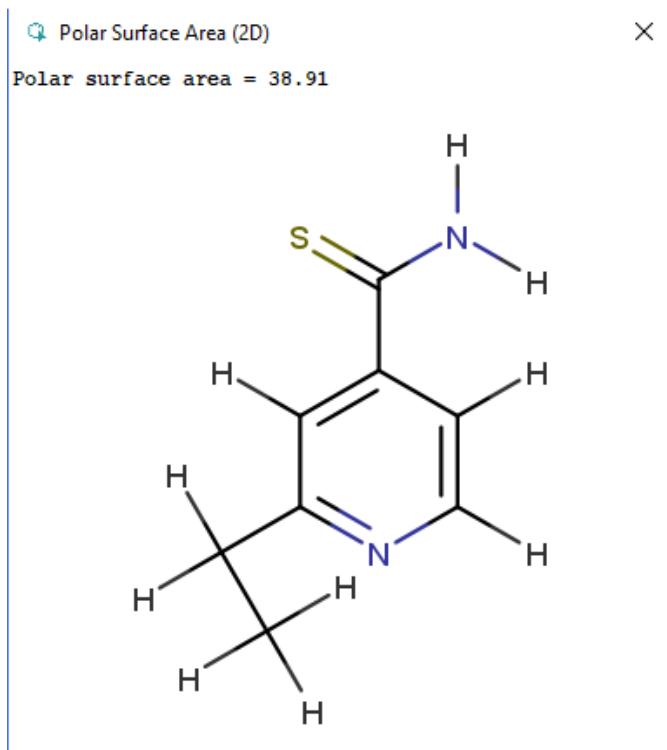
→Refractivity



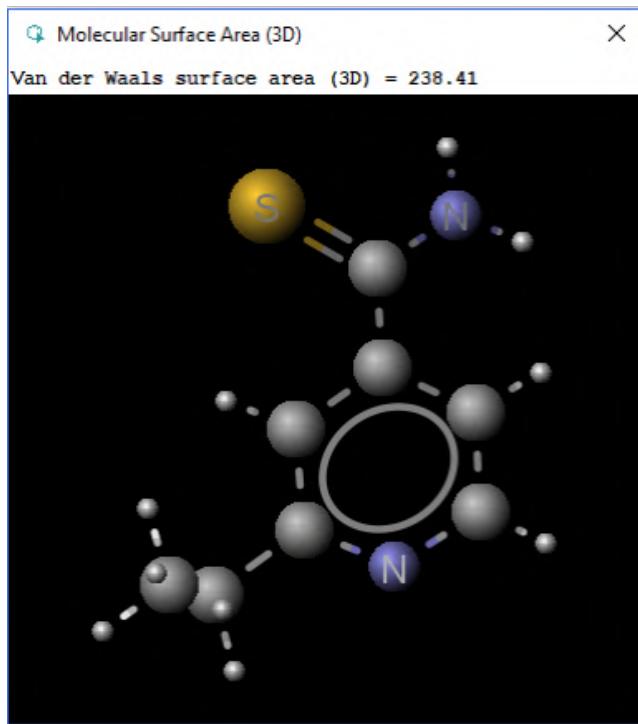
→Topology Analysis



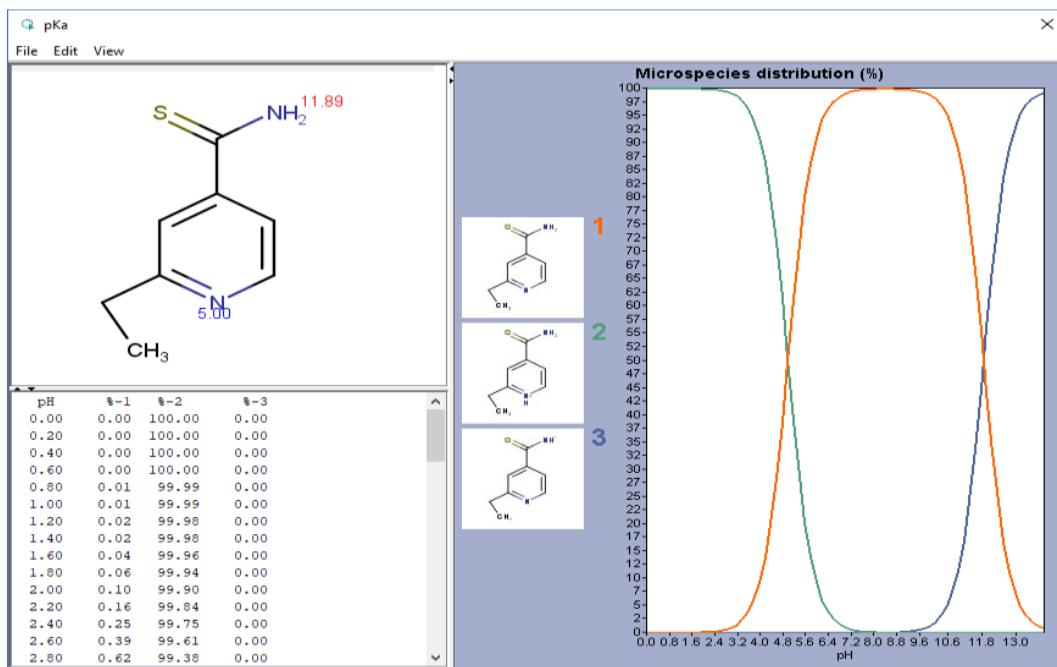
→Polar Surface Area (2D)



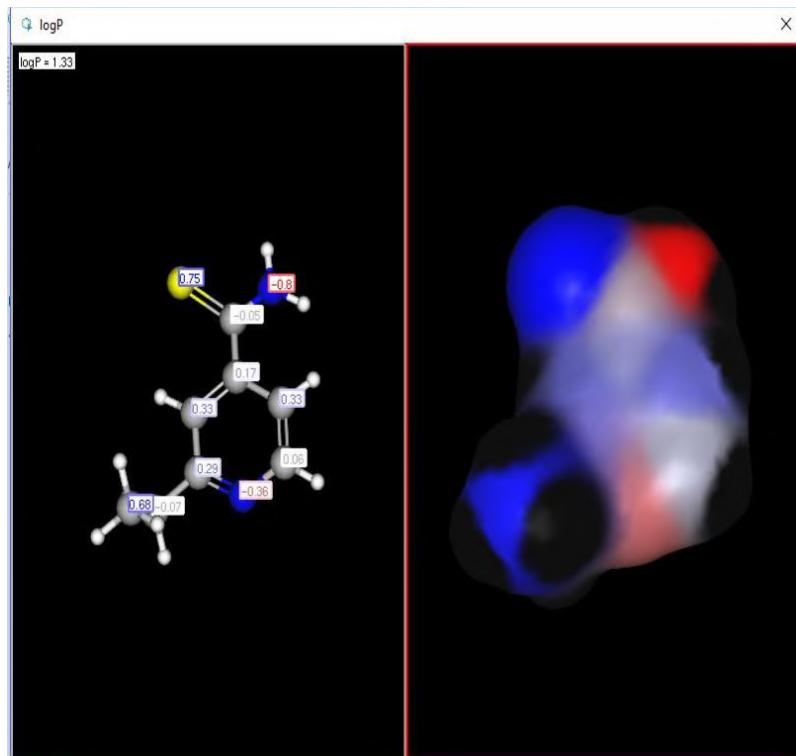
→Molecular Surface Area (3D)



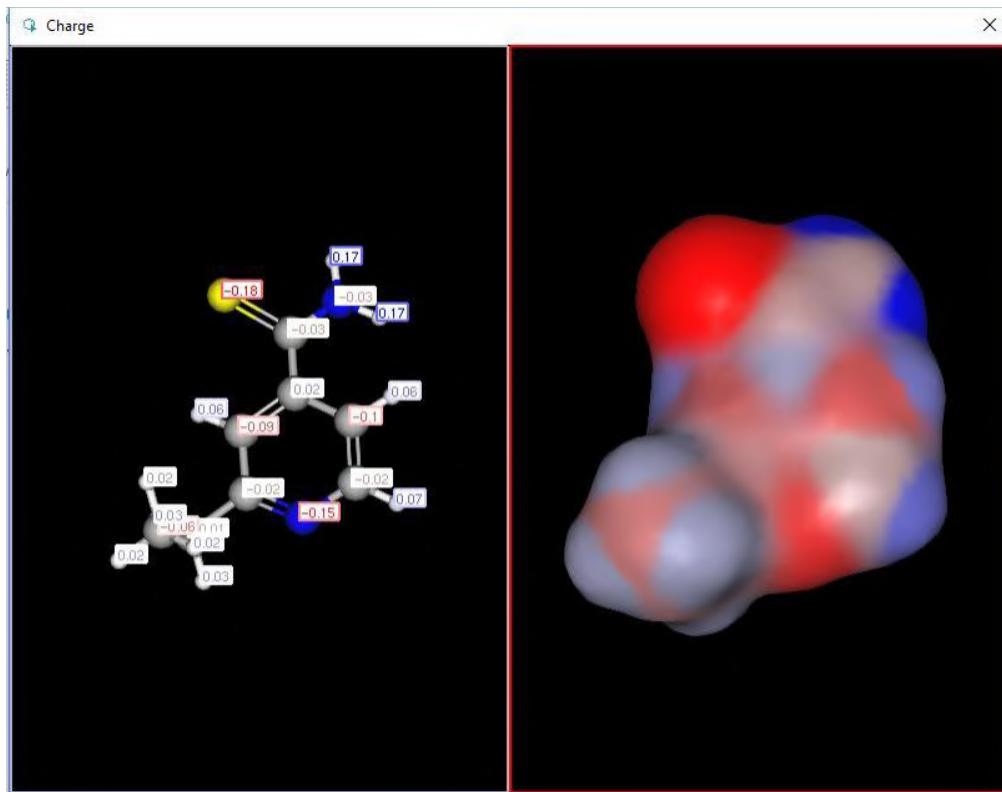
→Pka



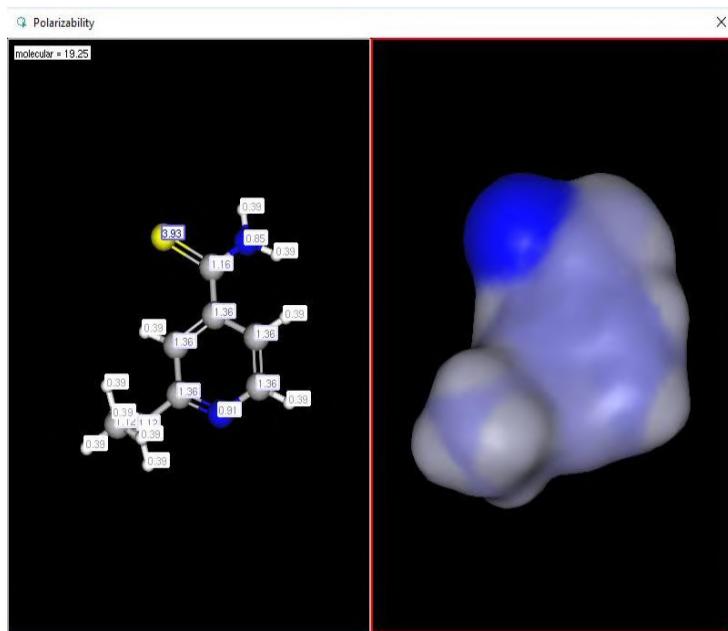
→logP



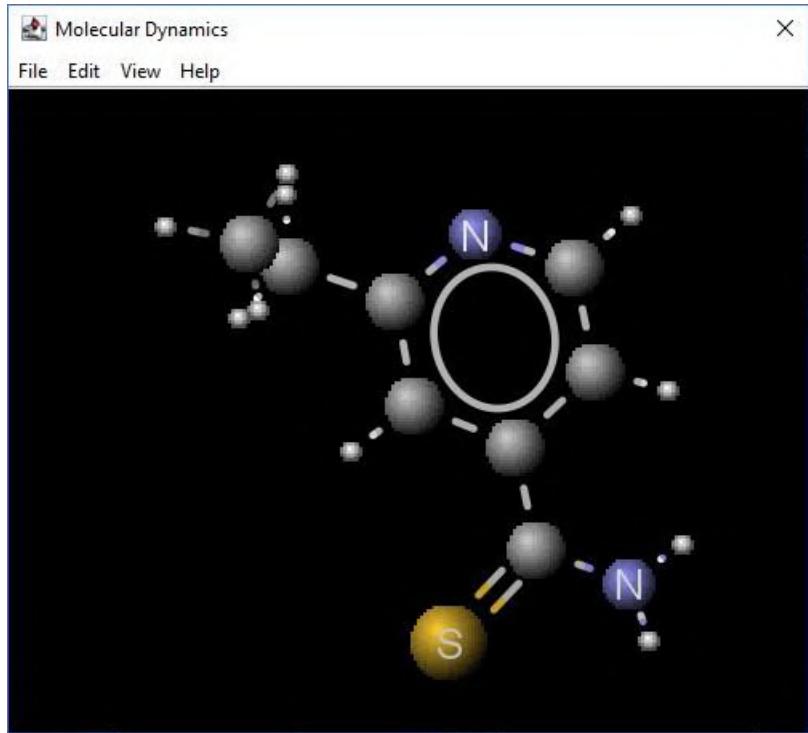
→Charge



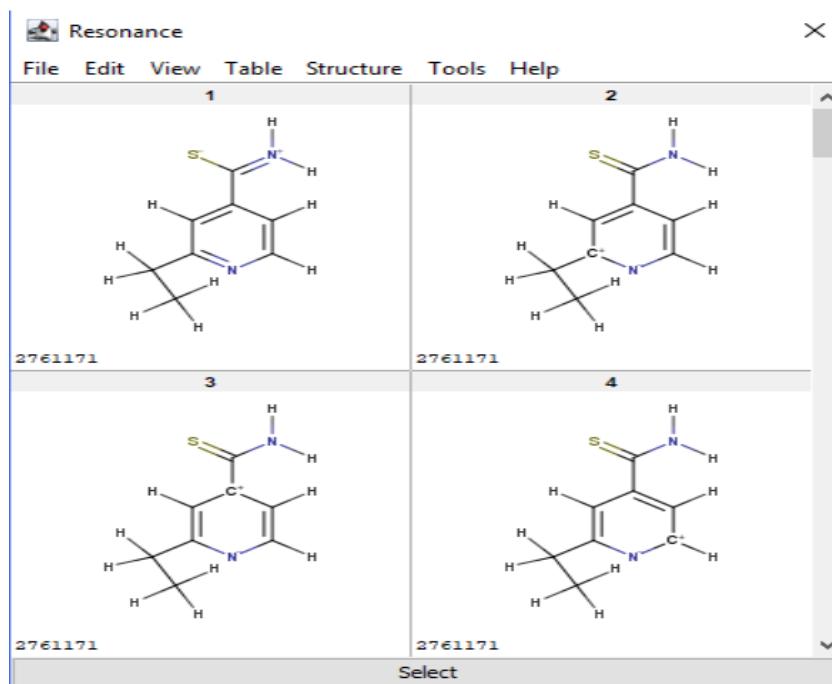
→Polarizability



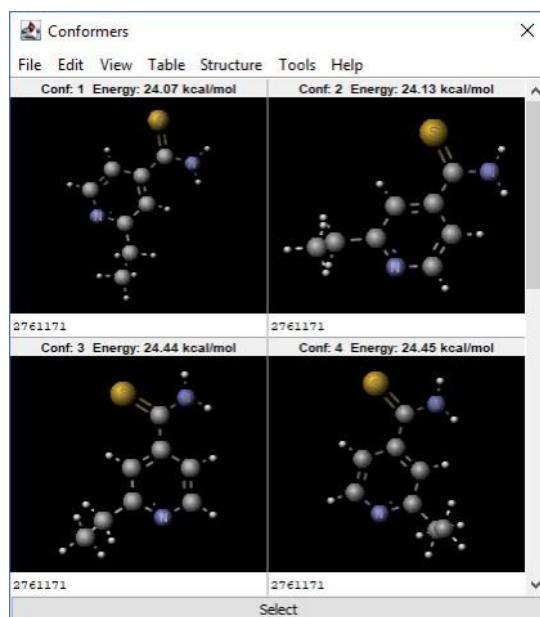
→Molecular Dynamics



→Resonance



→Conformers



CONCLUSION:

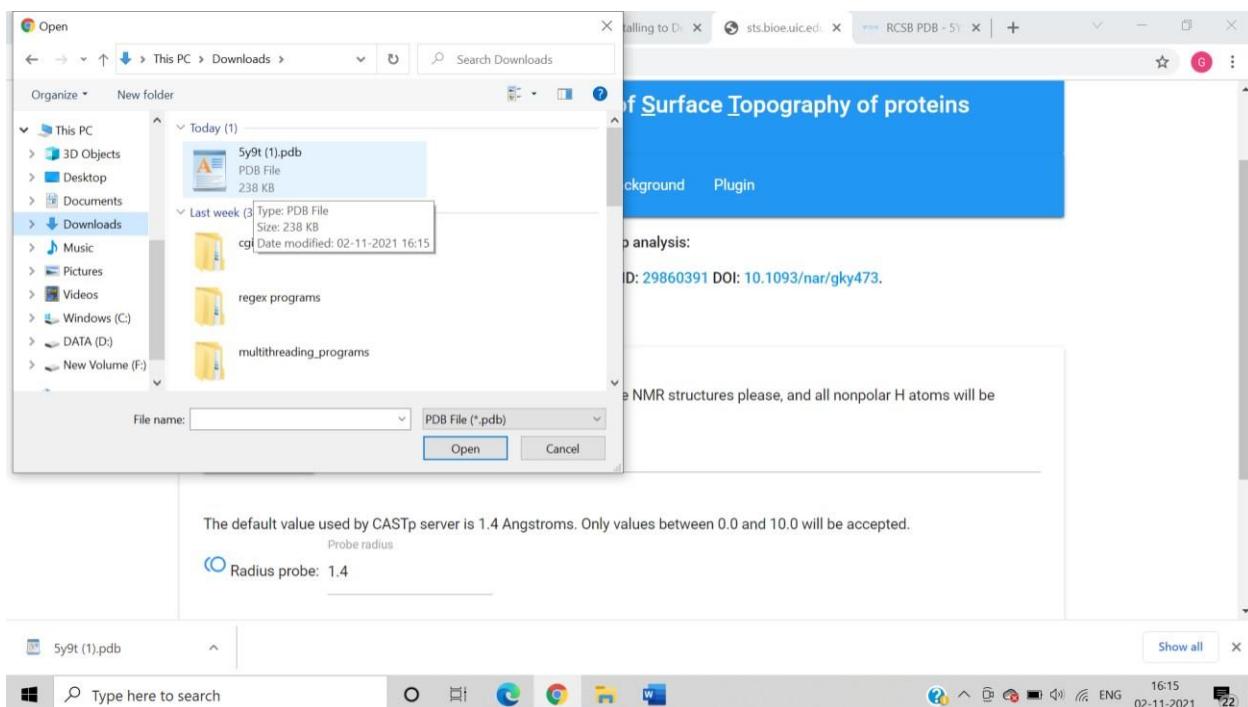
Chemical structure was done using Marvin sketch. Various features provided by Marvin sketch was used to study the structure.

Experiment no: 3A

Aim: Protein active site prediction using CastP.

Theory:- Computed Atlas of Surface Topography of proteins (Cast P) is used for locating, delineating and measuring concave surface regions on three-dimensional structures of proteins . It aims to provide comprehensive and detailed quantitative characterization of interior voids and surface pockets of proteins. The measurement includes the area and volume of pocket or void by solvent accessible surface model (Richards' surface) and by molecular surface model (Connolly's surface), all calculated analytically. It measures the size of mouth openings of individual pockets, which helps to assess the accessibility of binding sites to various ligands and substrates. It can be used to study surface features and functional regions of proteins; includes a graphical user interface, flexible interactive visualization, as well as on-the-fly calculation for user uploaded structures.

Methodology:-



Radius: 1.4

Please cite this paper if you publish or present results using CASTP analysis:

Tian et al., Nucleic Acids Res. 2018. PMID: 29860391 DOI: 10.1093/nar/gky473.

For questions and bugs, please contact uic.lianglab@gmail.com.

Upload a molecular structure in standard PDB format. No multiple NMR structures please, and all nonpolar H atoms will be ignored

5y9t (1).pdb

The default value used by CASTP server is 1.4 Angstroms. Only values between 0.0 and 10.0 will be accepted.

Probe radius
Radius probe: 1.4

Email (optional): Your email address

5y9t (1).pdb

Put your email address and submit:

For questions and bugs, please contact uic.lianglab@gmail.com.

Upload a molecular structure in standard PDB format. No multiple NMR structures please, and all nonpolar H atoms will be ignored

5y9t (1).pdb

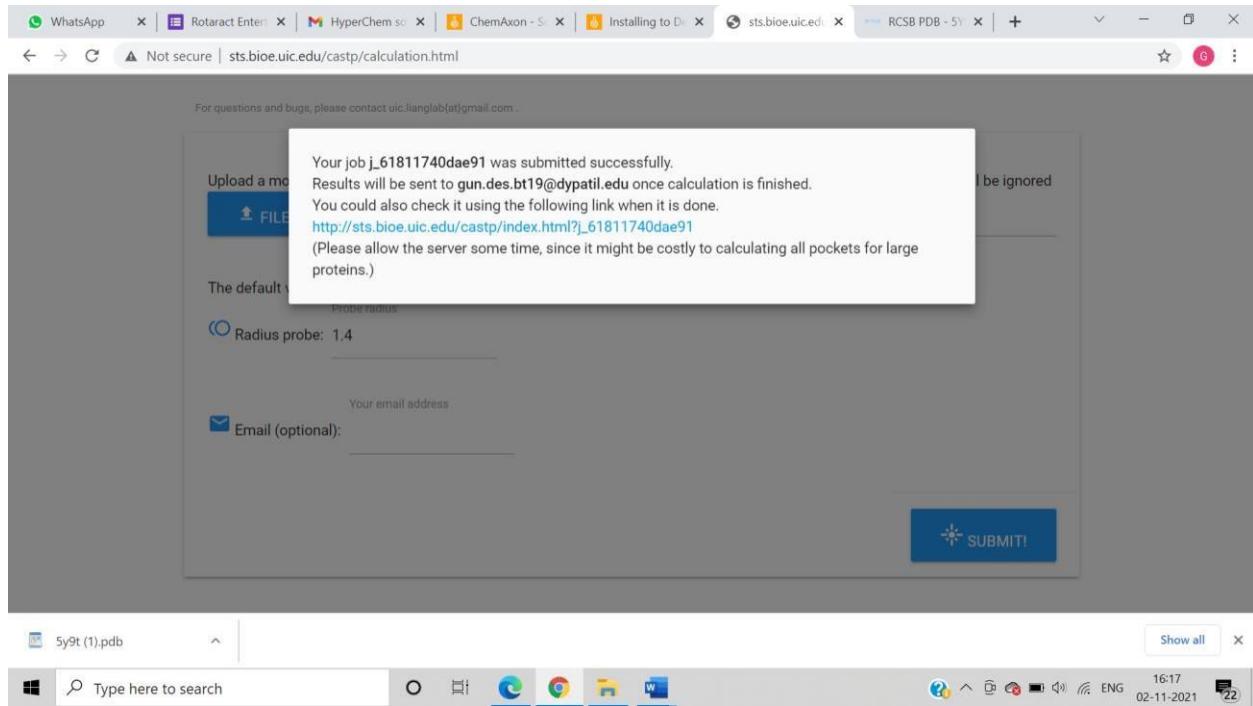
The default value used by CASTP server is 1.4 Angstroms. Only values between 0.0 and 10.0 will be accepted.

Radius probe: 1.4

Your email address
Email (optional): gun.des.bt19@dypatil.edu

SUBMIT!

5y9t (1).pdb

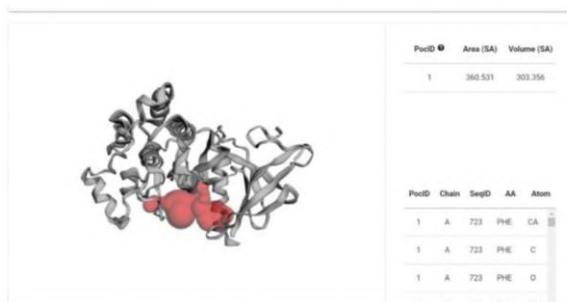


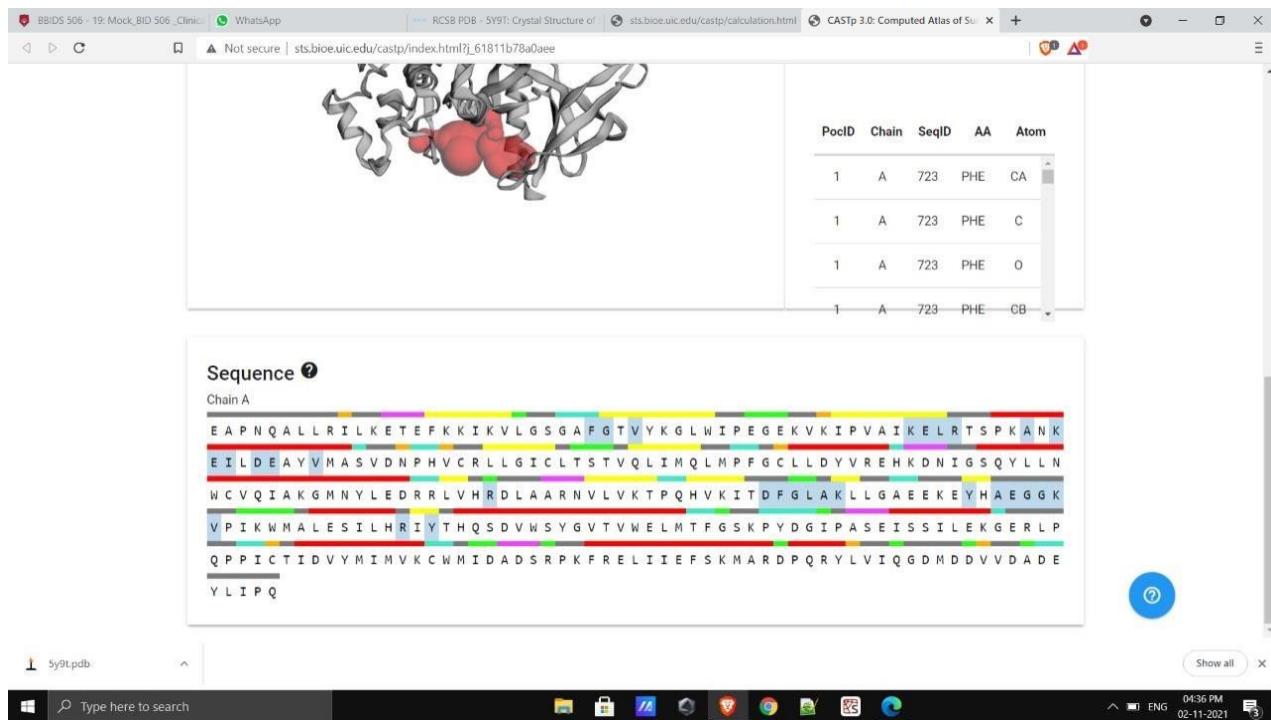
Result:-

The largest cavity is colored red:-

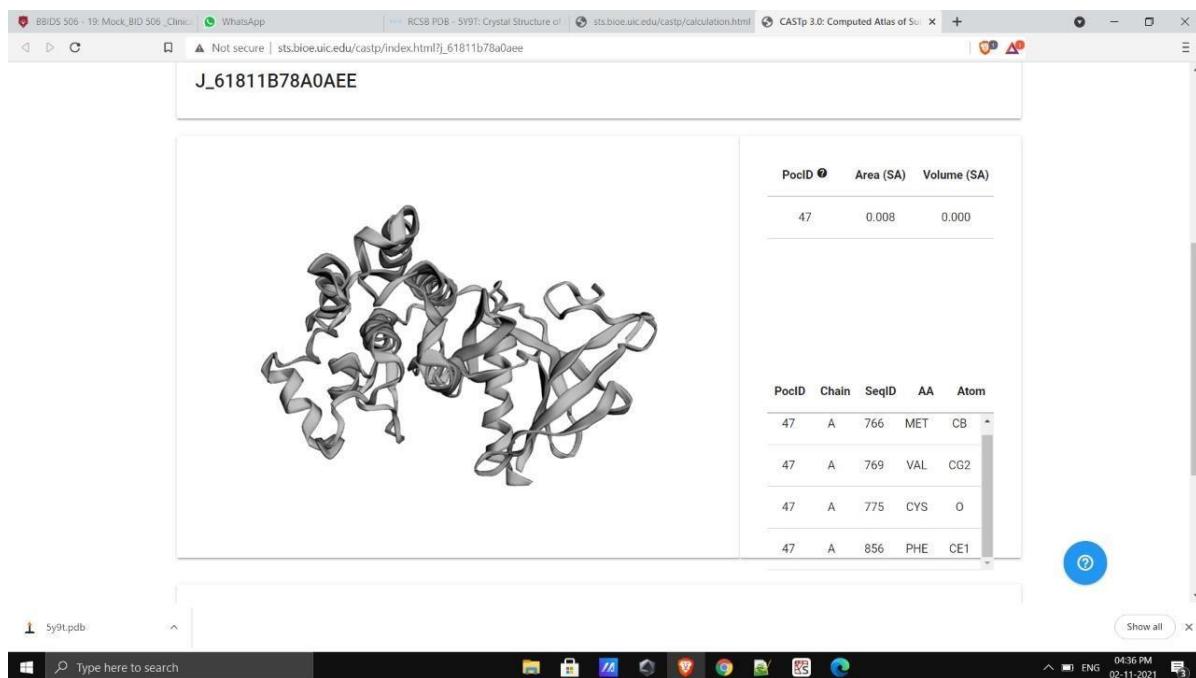
Area = 360.531

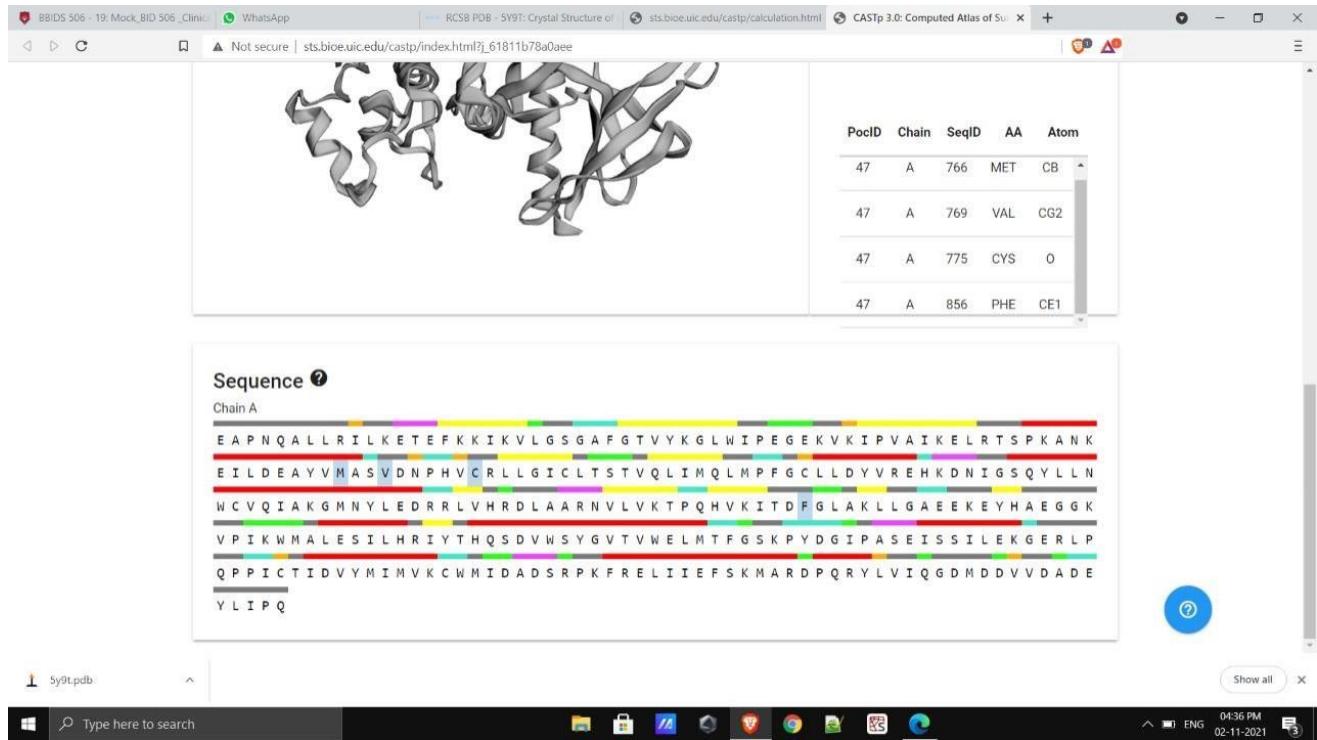
Volume = 303.356





The highlighted amino acid residues in Chain A results in the formation of cavity 1. With Surface Area of 360.531 Å and volume 303.356





Experiment no: 3B

Aim: Protein structure summary and analysis using PDBSUM

Theory: The PDBsum is a pictorial database that provides an at-a-glance overview of the contents of each 3D structure deposited in the Protein Data Bank (PDB). It shows the molecule(s) that make up the structure (*i.e.* protein chains, DNA, ligands and metal ions) and schematic diagrams of their interactions. Extensive use is made of the freely available RasMol molecular graphics program to view the molecules and their interactions in 3D.

Methodology:-

First go to drug bank and search for your disease drug:

Disease: Glioma

Drug: Carmustine, Gliadel

Given Gene name: GSR

The screenshot shows the DRUGBANK Online interface. The search bar at the top contains the query "gliadel". The main content area displays the details for "2. Glutathione reductase, mitochondrial". The data is presented in a table format:

Kind	Protein	General Function	Nadp binding
Organism	Humans	Specific Function	Maintains high levels of reduced glutathione in the cytosol.
Pharmacological action	Yes	Gene Name	GSR
Actions	Inhibitor	Uniprot ID	P00390
		Uniprot Name	Glutathione reductase, mitochondrial
		Molecular Weight	56256.565 Da

Below the table, there is a section titled "References" with one entry: "1. Akella SS, Harris C. Pyridine nucleotide flux and glutathione oxidation in the cultured rat conceptus. Reprod Toxicol. 1999 May; 9(3):331-6." The bottom of the screen shows a taskbar with various icons and the system tray indicating the date and time.

Go to RCSB PDB and in search bar search the gene name GSR:

Screenshot of the RCSB PDB search results for "GSR".

Search Query: GSR

Results:

- in Chemical ID(s)**: GSR
- in Additional Structure Keywords**: Q74D82, GsR143A, Peptidase, M48 family, Structural Genomics, PSI-2, Protein Structure Initiative, Northeast Structural Genomics Consortium, NESG, Hydrolase, Metalloprotease, Protease
- in Structure Title**: Northeast Structural Genomics target GsR13
- in Primary Citation Title**: SOLUTION STRUCTURE OF PUTATIVE UNCHARACTERIZED PROTEIN GSU1278 FROM METHANOCALDOCOCCUS JANNASCHII, NORTHEAST STRUCTURAL GENOMICS CONSORTIUM (NESG) TARGET GsR195

Advanced Search Query Builder:

- Full Text: GSR
- Structure Attribute: GSR
- Chemical Attribute: GSR
- Sequence: GSR
- Sequence Motif: GSR
- Structure Similarity: GSR

System Status: 17:20 02-11-2021

You will get 198 hits for the gene name GSR:

Screenshot of the RCSB PDB search results for "GSR" showing 198 hits.

Displaying 1 to 25 of 198 Structures

STRUCTURAL ENTITY ID	STRUCTURE DESCRIPTION	RELEASED	METHOD	ORGANISMS	MACROMOLECULE	UNIQUE LIGANDS
5F51	Structure of <i>B. abortus</i> WrbA-related protein A (apo)	2016-03-09	X-RAY DIFFRACTION 2.53 Å	<i>Brucella abortus</i> 2308	NAD(P)H dehydrogenase (quinone) (protein)	SO4
5F4B	Structure of <i>B. abortus</i> WrbA-related protein A (Wrpa)	2016-03-09	X-RAY DIFFRACTION 2.498 Å	<i>Brucella abortus</i> 2308	NAD(P)H dehydrogenase (quinone) (protein)	

Filters:

- SCIENTIFIC NAME OF SOURCE ORGANISM: *Homo sapiens* (51), *Bos taurus* (23), *Enterococcus faecalis* (8), *Pseudomonas* sp. KKS102 (7), *Xanthobacter autotrophicus* Py2 (7), *Afipia carboxidovorans* OM5 (6), *Arabidopsis thaliana* (6), *Critchidia fasciculata* (6), *Pseudomonas putida* (6), *Rhodobacter capsulatus* (6), More...
- TAXONOMY: Eukaryota (108), Bacteria (89), Archaea (2)
- EXPERIMENTAL METHOD: X-RAY DIFFRACTION (195), SOLUTION NMR (2), ELECTRON MICROSCOPY (1)

POI YMFR ENTITY TYPE: <https://www.rcsb.org/structure/5F4B>

System Status: 17:21 02-11-2021

WhatsApp | Rotaract Entertainn | Inbox (332) - gun.d | RCSB PDB: Search | GSR in All results | PROCHECK summary

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation Careers MyPDB

Protein (197)
 DNA (1)

REFINEMENT RESOLUTION (Å)
 0.5 - 1.0 (1)
 1.0 - 1.5 (13)
 1.5 - 2.0 (55)
 2.0 - 2.5 (72)
 2.5 - 3.0 (40)
 3.0 - 3.5 (13)
 3.5 - 4.0 (1)
 4.0 - 4.5 (1)

RELEASE DATE
 1985 - 1989 (1)
 1990 - 1994 (20)
 1995 - 1999 (24)
 2000 - 2004 (37)
 2005 - 2009 (40)
 2010 - 2014 (45)
 2015 - 2019 (19)
 2020 - 2024 (12)

ENZYME CLASSIFICATION NAME

https://www.rcsb.org/structure/3DK9

3D View

Macromolecule NAD(P)-reductase (quinone) (protein)
Unique Ligands CL, FMN

4GR1
THE BINDING OF THE RETRO-ANALOGUE OF GLUTATHIONE DISULFIDE TO GLUTATHIONE REDUCTASE
Schulz, G.E., Janes, W.
(1990) J Biol Chem 265: 10443-10445

Released 1991-10-15
Method X-RAY DIFFRACTION 2.4 Å
Organisms Homo sapiens
Macromolecule GLUTATHIONE REDUCTASE (protein)
Unique Ligands FAD, PO4, RGS

3DK9
Catalytic cycle of human glutathione reductase near 1 Å resolution
Berkholz, D.S., Faber, H.R., Savvides, S.N., Karplus, P.A.
(2008) J Mol Biol 382: 371-384

Released 2008-08-05
Method X-RAY DIFFRACTION 0.95 Å
Organisms Homo sapiens
Macromolecule Glutathione reductase (protein)
Unique Ligands FAD, SO4

Download File View File

Contact Us

Type here to search

17:21 02-11-2021

WhatsApp | Rotaract Entertainn | Inbox (332) - gun.d | RCSB PDB - 3DK9 | GSR in All results | PROCHECK summary

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation Careers MyPDB

PDB PROTEIN DATA BANK 183584 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Enter search terms or PDB ID(s). Advanced Search | Browse Annotations Help

Celebrating 40 YEARS OF Protein Data Bank

Structure Summary 3D View Annotations Experiment Sequence Genome Ligands Versions

Biological Assembly 1 3DK9 Display Files Download Files

3DK9 Catalytic cycle of human glutathione reductase near 1 Å resolution
DOI: 10.2210/pdb3DK9/pdb
Classification: OXIDOREDUCTASE
Organism(s): Homo sapiens
Expression System: Escherichia coli
Mutation(s): No

Deposited: 2008-06-24 Released: 2008-08-05
Deposition Author(s): Berkholz, D.S., Faber, H.R., Savvides, S.N., Karplus, P.A.

Experimental Data Snapshot wwPDB Validation 3D Report Full Report

Type here to search

17:22 02-11-2021

Go to EMBL EBI and paste the PDB code obtained:

3D9K :

S WhatsApp | Rotaract Entertainment | Not secure | ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdicode=index.html | RCSB PDB - 3DK9: Cataly | PDBsum home page

EMBL-EBI PDBsum

Databases Tools Research Training Industry About Us Help Site Index

PDBsum Pictorial database of 3D structures in the Protein Data Bank

PDBsum is a pictorial database that provides an at-a-glance overview of the contents of each 3D structure deposited in the Protein Data Bank (PDB). It shows the molecule(s) that make up the structure (ie protein chains, DNA, ligands and metal ions) and schematic diagrams of the interactions between them. [Read more...](#)

PDB code (4 chars) 3DK9 **Find**

Example: "1kfv"

Alpha Fold model (human proteins only) **Find** see

Analyses
Enter UniProt accession (or UniProt id), to find Alpha Fold model of given protein. Eg P00734 (THRB_HUMAN).

Text search **Search**
Scans all TITLE, HEADER, COMPND, SOURCE and AUTHOR records in the PDB (eg to find a given protein by name).

Search by sequence **Search**
Perform FASTA search vs all sequences in the PDB to get a list of the closest matches.

Search by **UniProt id:** **Pfam id:** **Ensembl id:**
 (eg P03023, LAC1_ECOLI, etc) (eg PF07992) (eg ENSG00000086205, ENST00000256999)
Search **Search** **Search**

Contents
PDBsum contains 208,818 entries including 1,451 superseded entries. Last update: 2 Sep 2021

In-house version
PDBsum

In-house version companies to provide their own structures below left

Related data
BC-PDB
Enzyme 3D structures organized by their numbering hierarchy
DrugPort
Structures of drugs

Type here to search

17:23 02-11-2021

1) list all the features:-

a)Top page:-

PDBsum entry 3d9k

PDBsum

Top page Protein Ligands Prot-prot Clefts Tunnels Links PDB Id 3d9k

PDB Id: 3d9k

Name: Transcription

Title: Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxy-terminal domain of RNA-polymerase II

Structure: RNA-binding protein 16 Chain a, b Fragment: ctd interacting domain of scaf8, unp residues 1-136. Synonym: RNA-binding motif protein 16. Engineered yes. Ctd-peptide: CTD-pep_2. Expression: Eukaryotic.

Source: Homo sapiens, Man Organism, taxid: 9606, Gene: rpm16, kaa1116. Expressed in: escherichia coli. Synthetic: yes. Other details: peptide derived from the conserved repeat sequence in RNA polymerase II ctd

Resolution: 2.20 Å **R-factor:** 0.193 **R-free:** 0.252

Authors: R Becker, B Loll, A Meinhart

Key ref: R Becker et al. (2008). Snapshots of the RNA Processing Factor SCAF8 Bound to Different Phosphorylated Forms of the Carboxy-terminal Domain of RNA Polymerase II. *J Biol Chem.* 283: 22059-22069. PubMed id: 18550522 DOI: 10.1074/jbc.M803540209

Date: 27-May-08 **Release date:** 10-Jun-08

PROCHECK

3Dmol

Contents

- Protein chains
- Ligands

Q9UPN6 (SCAF8_HUMAN) - SR-related and CTD-associated factor 8 from Homo sapiens

Seq Struc

Seq Struc

Seq Struc

1271 a.a. 137 a.a.*

Type here to search

17:23 ENG 02-11-2021 22

Interpretation: The top page of PDBsum for the given target protein provides an overview including information like it's structure, chains, enzyme reactions ,procheck analysis.

b)Ligands:

The screenshot shows the PDBsum entry for PDB code 3d9k. The ligand is PRO-SER-TYR-SEP-PRO-THR, identified as Phosphoserine (C₃H₈NO₆P). A validation report table provides details on atoms, missing atoms, and chiral centers for each residue.

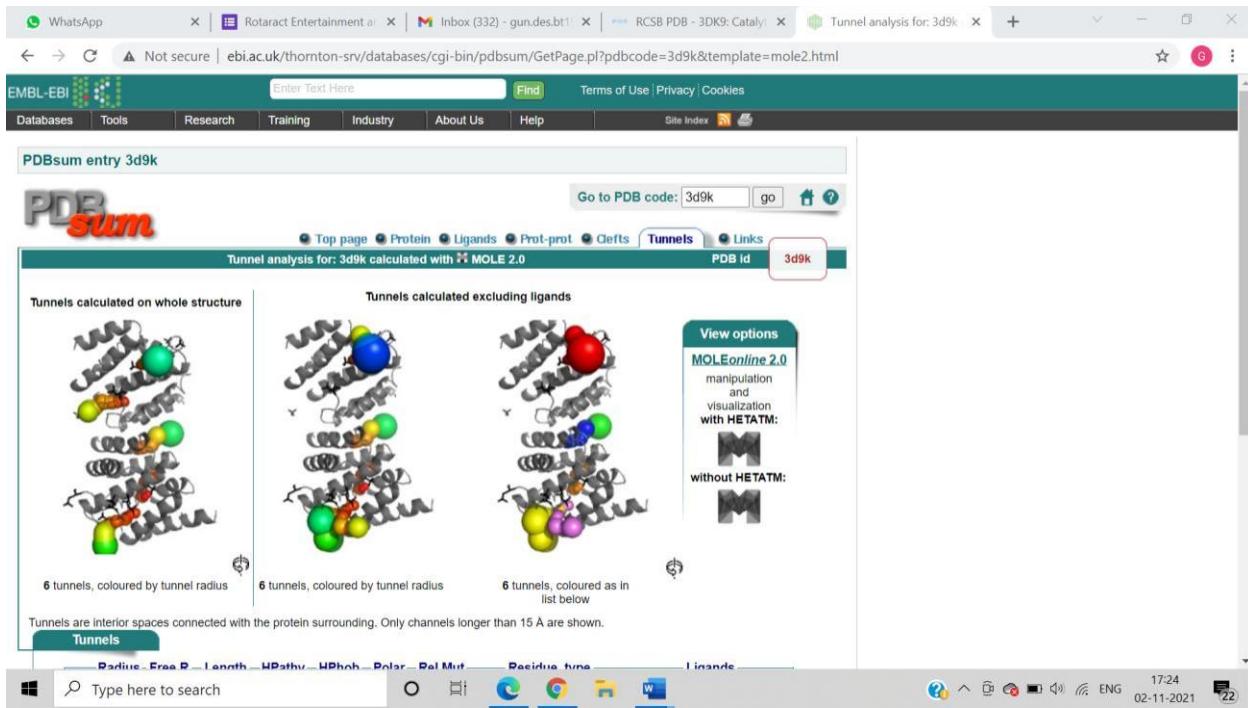
Residue	Atoms	Missing	Incorrect Chiral Centres							
	Dic.	Struc	Link	Subs	Atoms	Rings	Planar	High	C	Other
PRO -1(Y)	8	8	1	1	0	0	0	0	0	0
SER 0(Y)	7	6	0	0	1	0	-	-	-	-
TYR 1(Y)	13	13	1	1	0	0	0	0	0	0
SEP 2(Y)	11	11	1	1	0	0	0	0	0	0
PRO 3(Y)	8	8	1	1	0	0	0	0	0	0
THR 4(Y)	8	7	0	0	1	0	-	-	-	-

Advanced Analysis table:

Residue	Name Mismatches	Count
PRO -1(Y)	O: OXT	1
SER 0(Y)	SP: SGL: SGP: LGL: GLG: GLT	1

Interpretation: The Ligand tab provides the information on the ligands involved in interaction with the protein with a validation report and Ligplot at the bottom of the page.

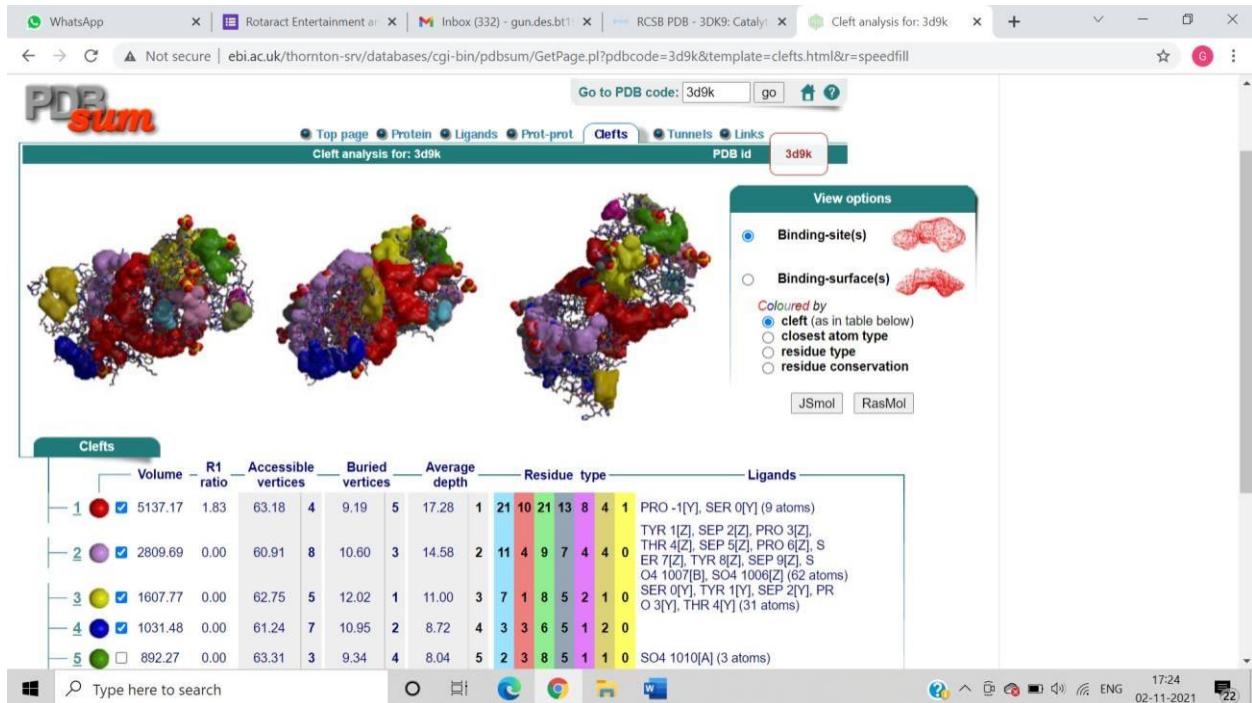
c) Tunnels:



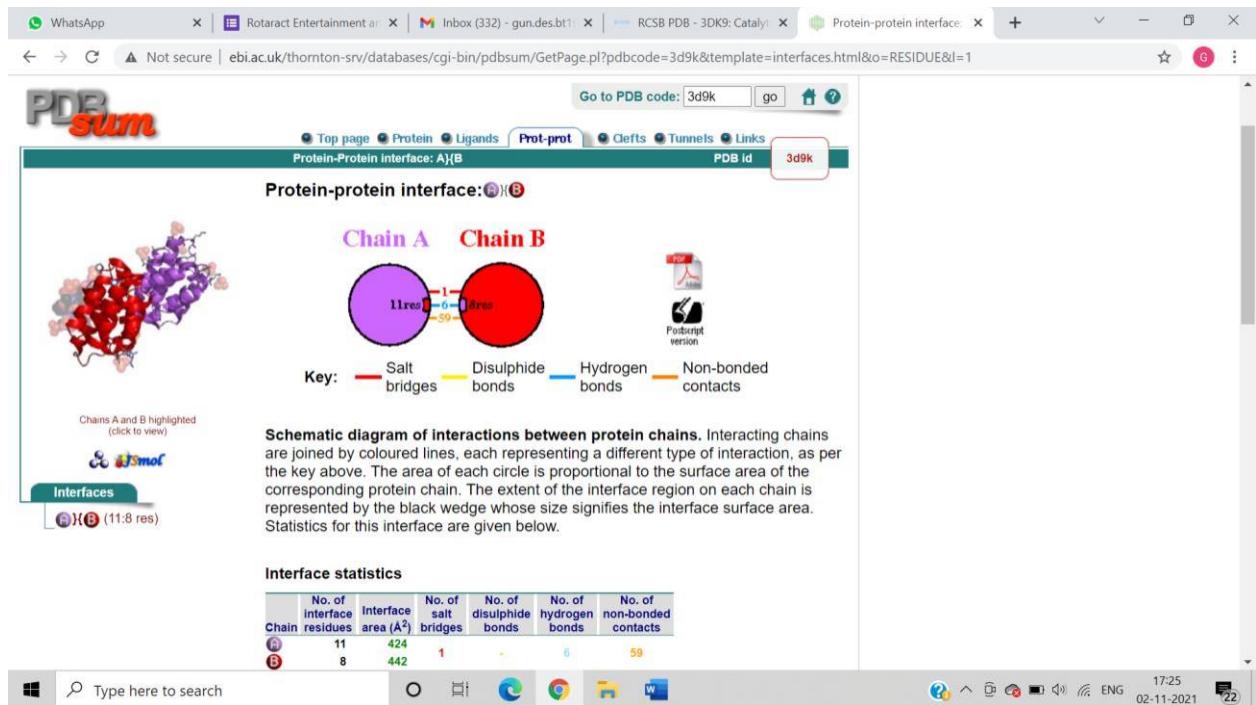
Interpretation:- Tunnels are interior spaces connected with the protein surrounding. Only channels longer than 15 Å are shown.

d)Clefs:

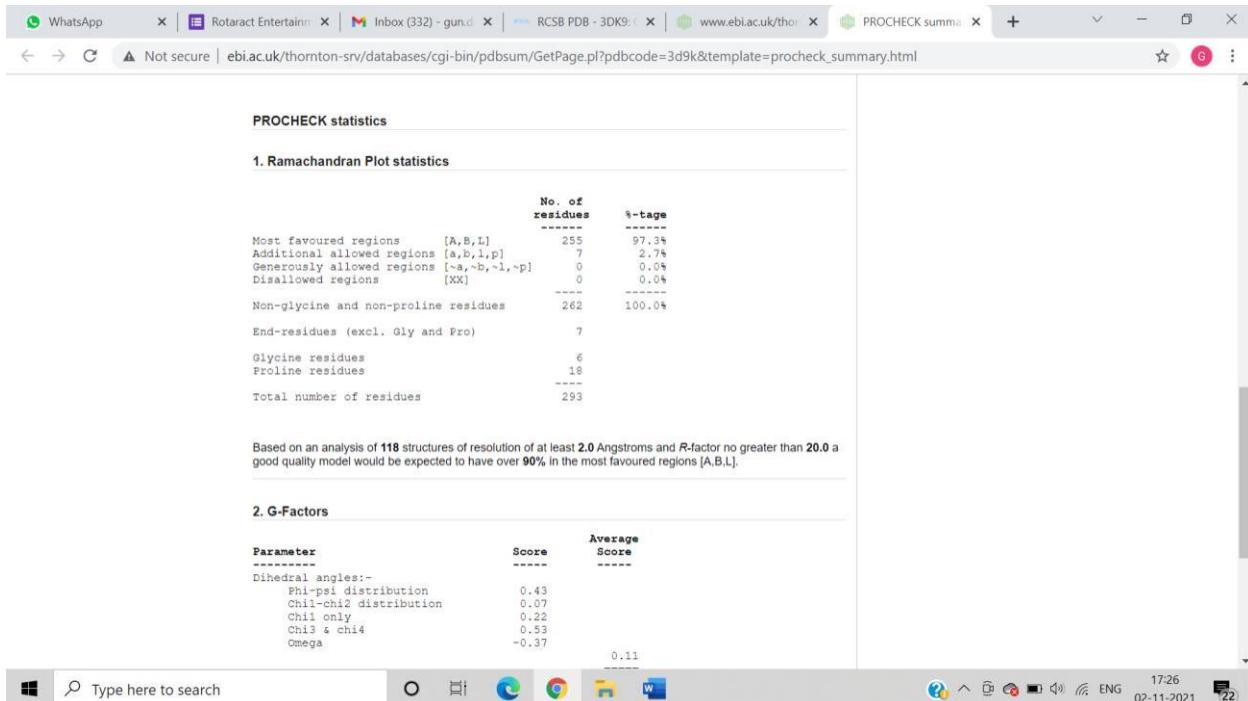
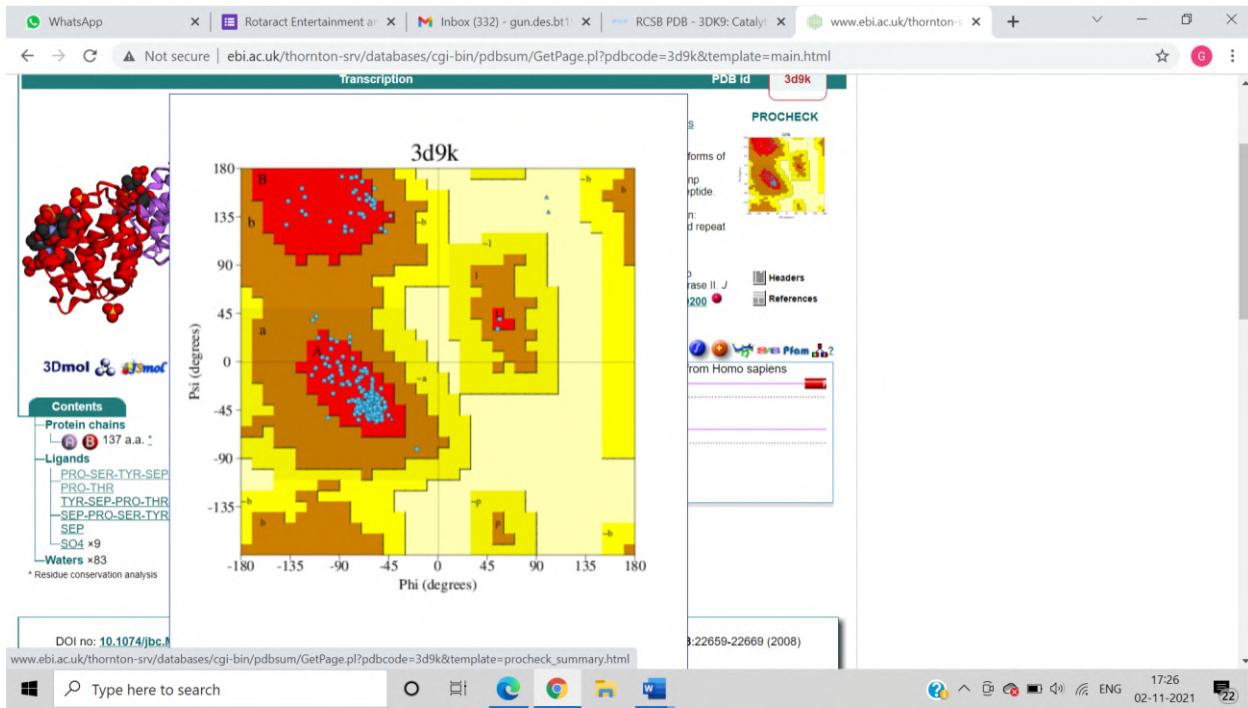
Clefs are the hydrophobic spaces present on surfaces of the protein, 10 clefs are present in the structure, the largest cleft is colored in red with a volume of 5137.17Å, consisting 21 positive residues, 10 negative residues, 21 neutral residues, 13 aliphatic residues, 8 aromatic residues, 4 residues consisting proline and glycine and 1 residues of cysteine.



e) Protein-protein interface:



2) Procheck :-



→ Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20.0 a good quality model would be expected to have over 90% in the most favoured regions [A,B,L].

WhatsApp | Rotaract Entertainn | Inbox (332) - gun.d | RCSB PDB - 3DK9 | www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbscode=3d9k&template=procheck_summary.html | PROCHECK summary

Not secure | ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbscode=3d9k&template=procheck_summary.html

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20.0 a good quality model would be expected to have over 90% in the most favoured regions [A,B,L].

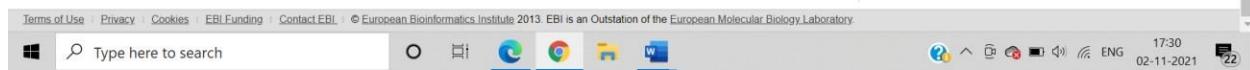
2. G-Factors

PARAMETER	Score	Average Score
Dihedral angles:-		
Phi-psi distribution	0.43	
Chi1-chi2 distribution	0.07	
Chi1 only	0.22	
Chi3 & chi4	0.53	
Omega	-0.37	0.11
Main-chain covalent forces:-		
Main-chain bond lengths	0.55	
Main-chain bond angles	0.48	0.51
OVERALL AVERAGE	0.26	

G-factors provide a measure of how unusual, or out-of-the-ordinary, a property is.

Values below -0.5* - unusual
Values below -1.0** - highly unusual

Important note: The main-chain bond-lengths and bond angles are compared with the Engh & Huber (1991) ideal values derived from small-molecule data. Therefore, structures refined using different restraints may show apparently large deviations from normality.



3)List of interactions:

List of protein-ligand interactions									
PDB code: 3d9k Ligand PRO-SER-TYR-SEP-PRO-THR									
Hydrogen bonds									
----- A T O M 1 -----> <----- A T O M 2 ----->									
Atom Atom Res Res Atom Atom Res Res									
no. name name no. Chain no. name name no. Chain Distance									
1. 160 O 21 A <-> 2326 N TYR 1 Y 2.79									
2. 174 N LYS 23 A --> 2316 O PRO -1 Y 2.95									
3. 545 OD2 ASP 67 A <-> 2337 OH TYR 1 Y 3.26									
Non-bonded contacts									
----- A T O M 1 -----> <----- A T O M 2 ----->									
Atom Atom Res Res Atom Atom Res Res									
no. name name no. Chain no. name name no. Chain Distance									
1. 154 CB PRO 20 A --- 2324 CB SER 0 Y 3.68									
2. 160 O ILE 21 A --- 2321 CA SER 0 Y 3.41									
3. 160 O ILE 21 A --- 2322 C SER 0 Y 3.52									
4. 160 O ILE 21 A --- 2324 CB SER 0 Y 3.68									
5. 160 O ILE 21 A --- 2326 N TYR 1 Y 2.79									
6. 160 O ILE 21 A --- 2327 CA TYR 1 Y 3.74									
7. 160 O ILE 21 A --- 2330 CB TYR 1 Y 3.60									
8. 166 CA SER 22 A --- 2316 O PRO -1 Y 3.80									
9. 168 C SER 22 A --- 2316 O PRO -1 Y 3.84									
10. 174 N LYS 23 A --- 2316 O PRO -1 Y 2.95									
11. 175 CA LYS 23 A --- 2316 O PRO -1 Y 3.77									
12. 178 CB LYS 23 A --- 2316 O PRO -1 Y 3.40									
13. 201 CR MET 26 A --- 2323 CD2 TYR 1 Y 3.81									

List of protein-ligand interactions									
PDB code: 3d9k Ligand PRO-SER-TYR-SEP-PRO-THR									
Hydrogen bonds									
----- A T O M 1 -----> <----- A T O M 2 ----->									
Atom Atom Res Res Atom Atom Res Res									
no. name name no. Chain no. name name no. Chain Distance									
16. 515 CB TYR 64 A --- 2332 CD1 TYR 1 Y 3.86									
17. 516 CG TYR 64 A --- 2332 CD1 TYR 1 Y 3.78									
18. 518 CD2 TYR 64 A --- 2329 O TYR 1 Y 3.84									
19. 518 CD2 TYR 64 A --- 2332 CD1 TYR 1 Y 3.84									
20. 519 CE1 TYR 64 A --- 2354 CD PRO 3 Y 3.72									
21. 520 CE2 TYR 64 A --- 2329 O TYR 1 Y 3.48									
22. 545 OD2 ASP 67 A --- 2337 OH TYR 1 Y 3.26									
23. 545 OD2 ASP 67 A --- 2349 CA PRO 3 Y 3.47									
24. 545 OD2 ASP 67 A --- 2352 CB PRO 3 Y 3.56									
25. 546 N SER 68 A --- 2337 OH TYR 1 Y 3.75									
26. 547 CA SER 68 A --- 2337 OH TYR 1 Y 3.53									
27. 550 CB SER 68 A --- 2336 CZ TYR 1 Y 3.89									
28. 550 CB SER 68 A --- 2337 OH TYR 1 Y 3.50									
29. 580 CD ARG 71 A --- 2337 OH TYR 1 Y 3.89									
30. 582 NE ARG 71 A --- 2350 C PRO 3 Y 3.87									
31. 582 NE ARG 71 A --- 2351 O PRO 3 Y 3.88									
32. 584 CZ ARG 71 A --- 2350 C PRO 3 Y 3.81									
33. 586 NH1 ARG 71 A --- 2357 C THR 4 Y 3.70									
34. 586 NH1 ARG 71 A --- 2358 O THR 4 Y 3.67									
35. 588 NH2 ARG 71 A --- 2358 C PRO 3 Y 3.78									
36. 588 NH2 ARG 71 A --- 2352 CB PRO 3 Y 3.27									
37. 588 NH2 ARG 71 A --- 2355 N THR 4 Y 3.62									
38. 966 CD1 LEU 116 A --- 2352 CB PRO 3 Y 3.75									
39. 966 CD1 LEU 116 A --- 2353 CG PRO 3 Y 3.86									
40. 1271 CB ASP 16 B --- 2317 CB PRO -1 Y 3.73									
41. 1272 CB ASP 16 B --- 2316 O PRO -1 Y 3.23									
42. 1272 CG ASP 16 B --- 2317 CB PRO -1 Y 3.77									
43. 1273 OD1 ASP 16 B --- 2316 O PRO -1 Y 3.30									
44. 1274 OD2 ASP 16 B --- 2315 C PRO -1 Y 3.77									
45. 1274 OD2 ASP 16 B --- 2316 O PRO -1 Y 3.15									
46. 1274 OD2 ASP 16 B --- 2317 CB PRO -1 Y 3.34									
47. 1284 CE2 TYR 17 B --- 2317 CB PRO -1 Y 3.81									
48. 1286 OH TYR 17 B --- 2317 CB PRO -1 Y 3.76									
49. 1286 OH TYR 17 B --- 2318 CG PRO -1 Y 3.61									
50. 1286 OH TYR 17 B --- 2319 CD PRO -1 Y 3.59									
Number of hydrogen bonds: 3									
Number of non-bonded contacts: 50									

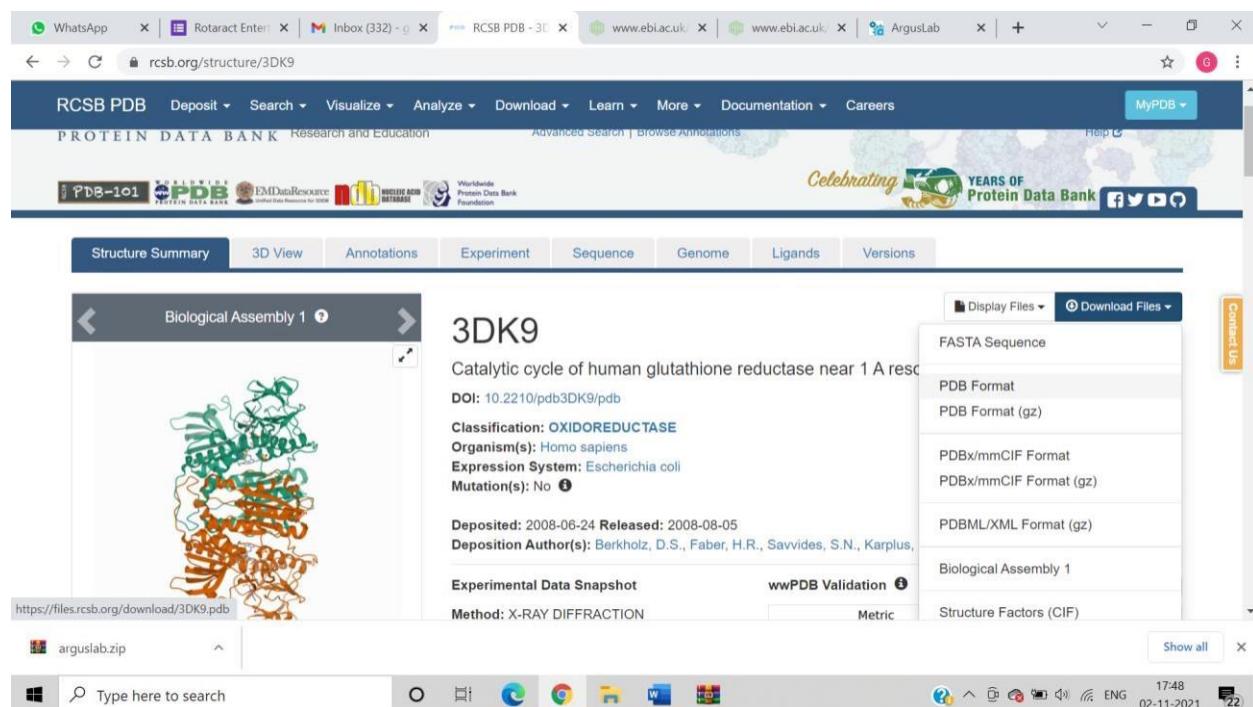
Practical no: 04

Aim: Perform molecular docking (Protein and ligand) using Argus algorithm in Argus lab

Theory: Molecular docking is a kind of bioinformatic modelling which involves the interaction of two or more molecules to give the stable adduct. Depending upon binding properties of ligand and target, it predicts the three-dimensional structure of any complex. Molecular docking generates different possible adduct structures that are ranked and grouped together using scoring function in the software. Docking simulations predict optimized docked conformer based upon total energy of the system. In spite of all potential approaches, ligand chemistry (tautomerism and ionization), receptor flexibility (single conformation of rigid receptor) and scoring function (differentiate true binding mode) still remained the challenge. Many important aspects of molecular docking in terms of its approaches, types, applications and challenges are briefly discussed in this article.

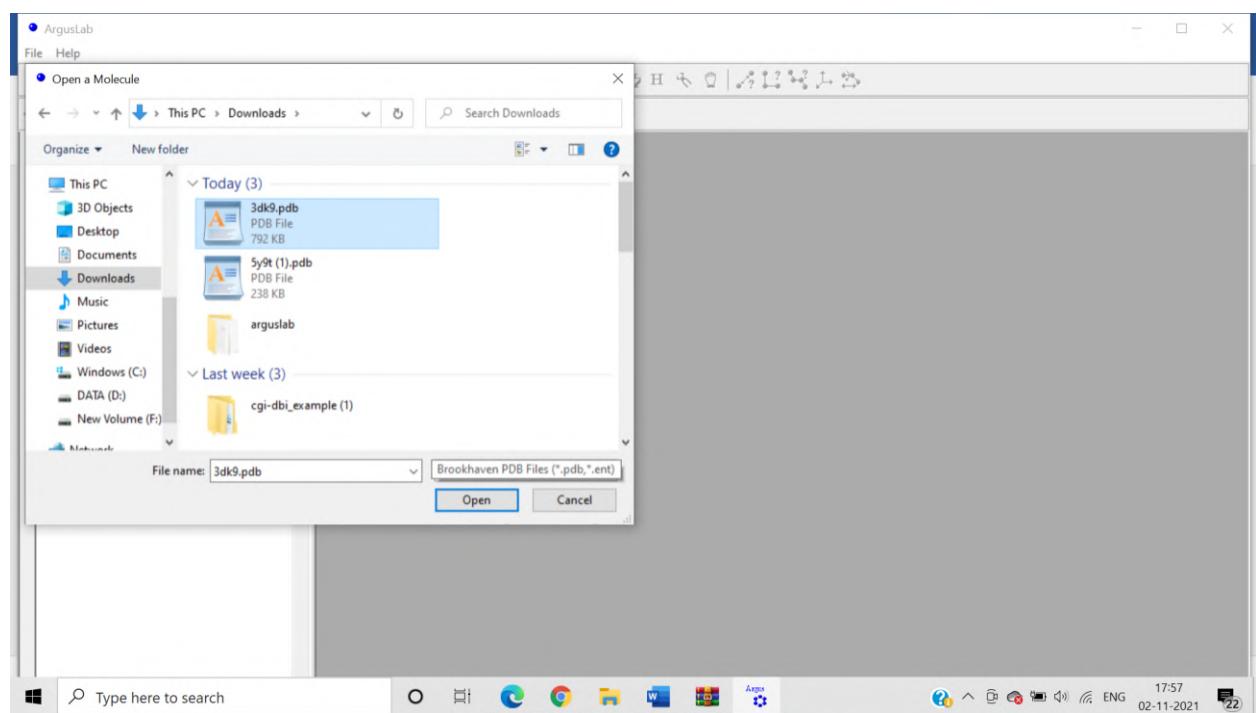
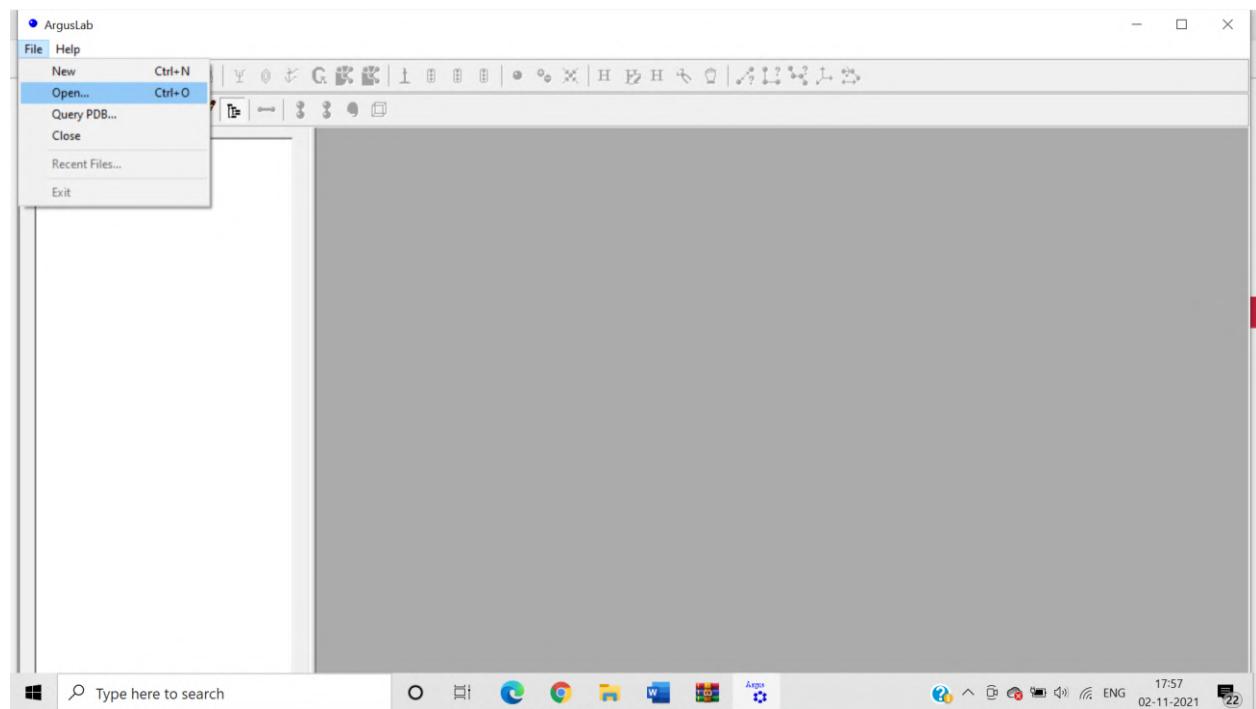
Methodology:

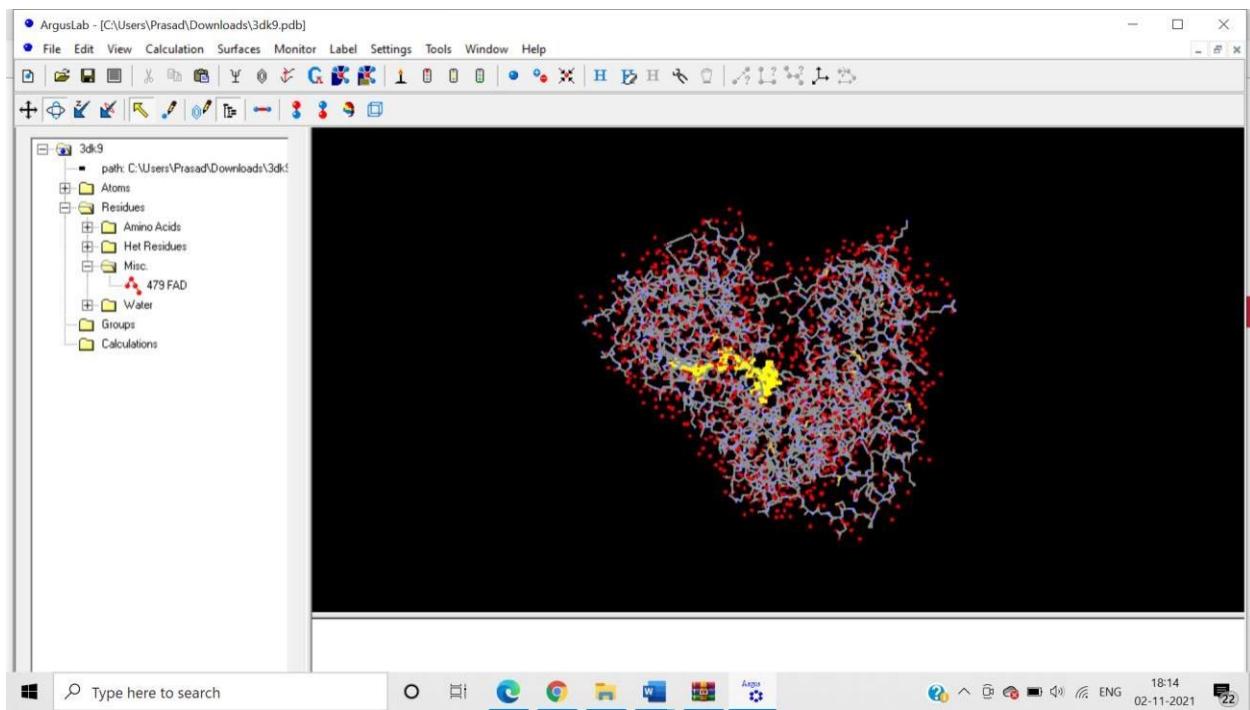
Open the 3DK9 file in RCSB PDB site and download it in pdb format:



1. Click on file and open a pdb file .
2. In the left hand side panel, click on the plus sign.
3. Click on residues.
4. Click on misc.

5. A list of ligands is shown, click on one (the selected part is highlighted).

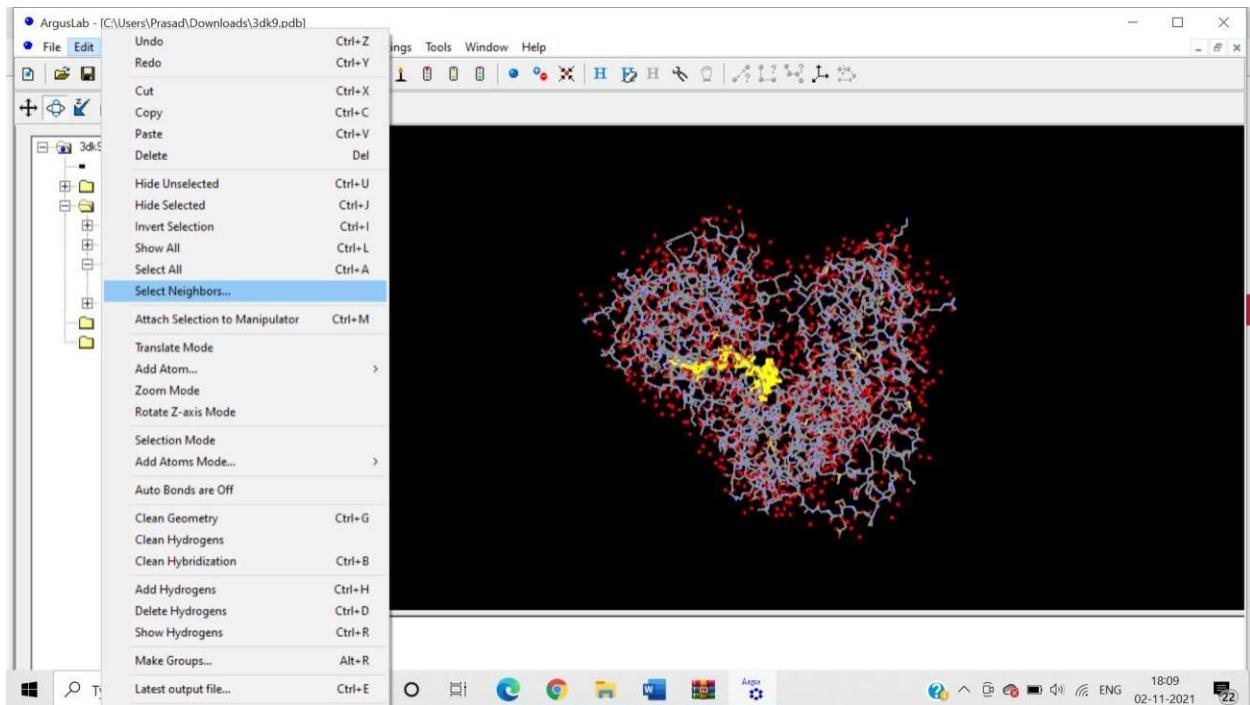




6. Click on edit.

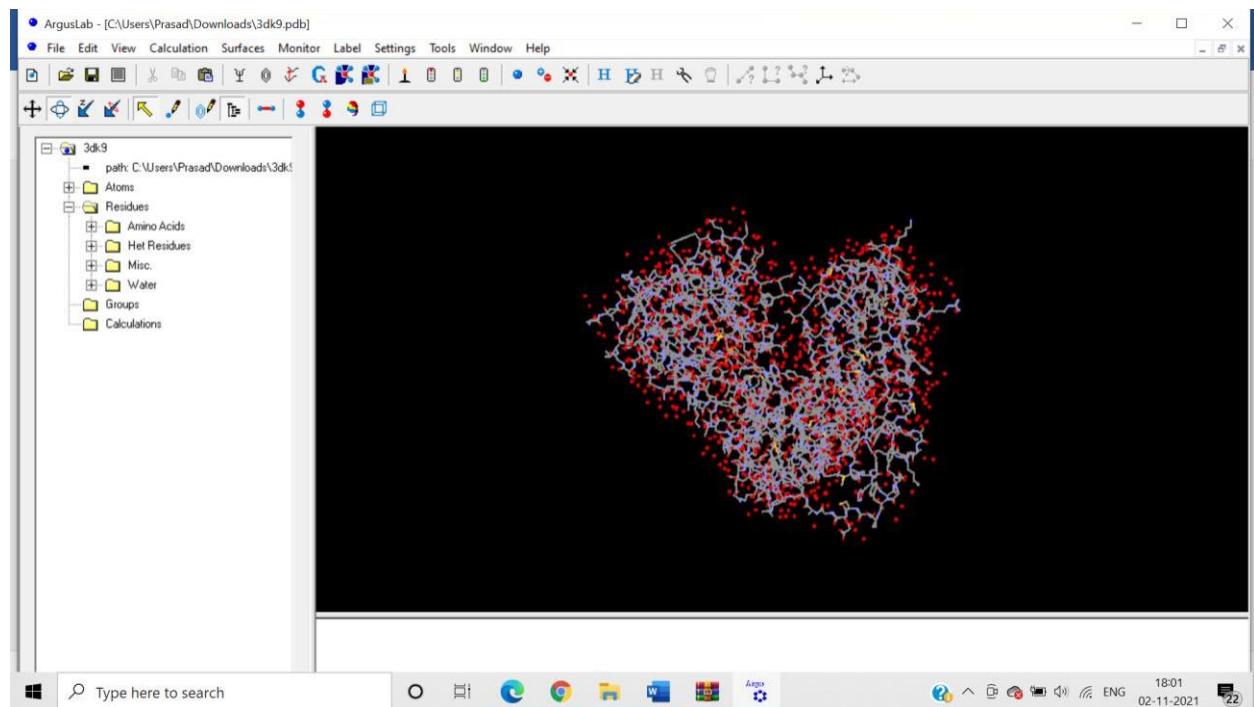
7. Click select neighbors.

8. A dialogue box will appear as shown below.



9. Click ok.

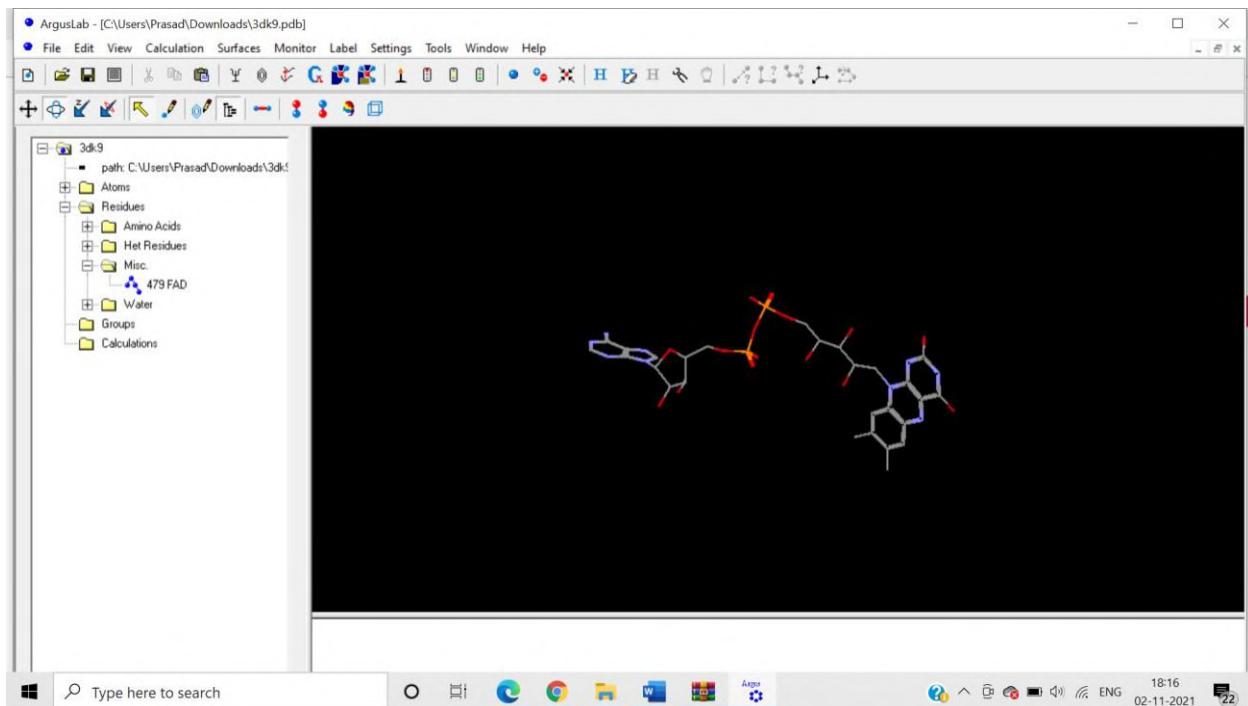
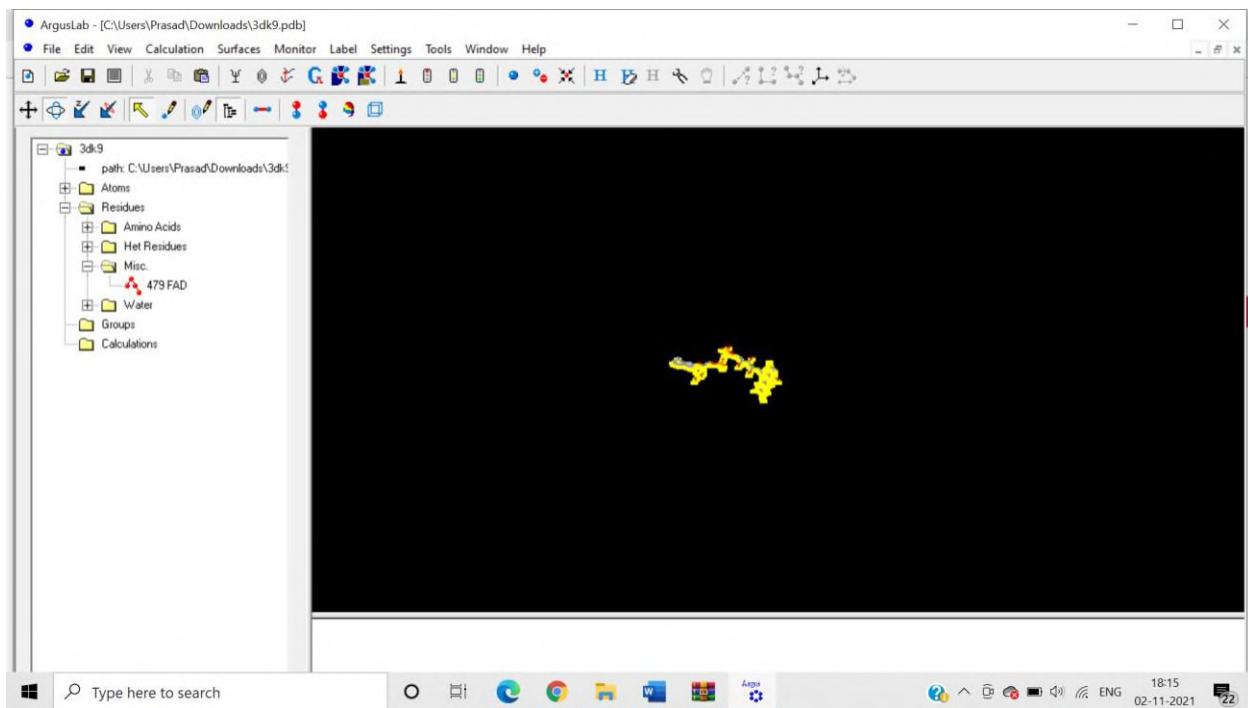
10. It will highlight the area around it (360 degrees).



11. Click on edit.

12. Click hide unselected.

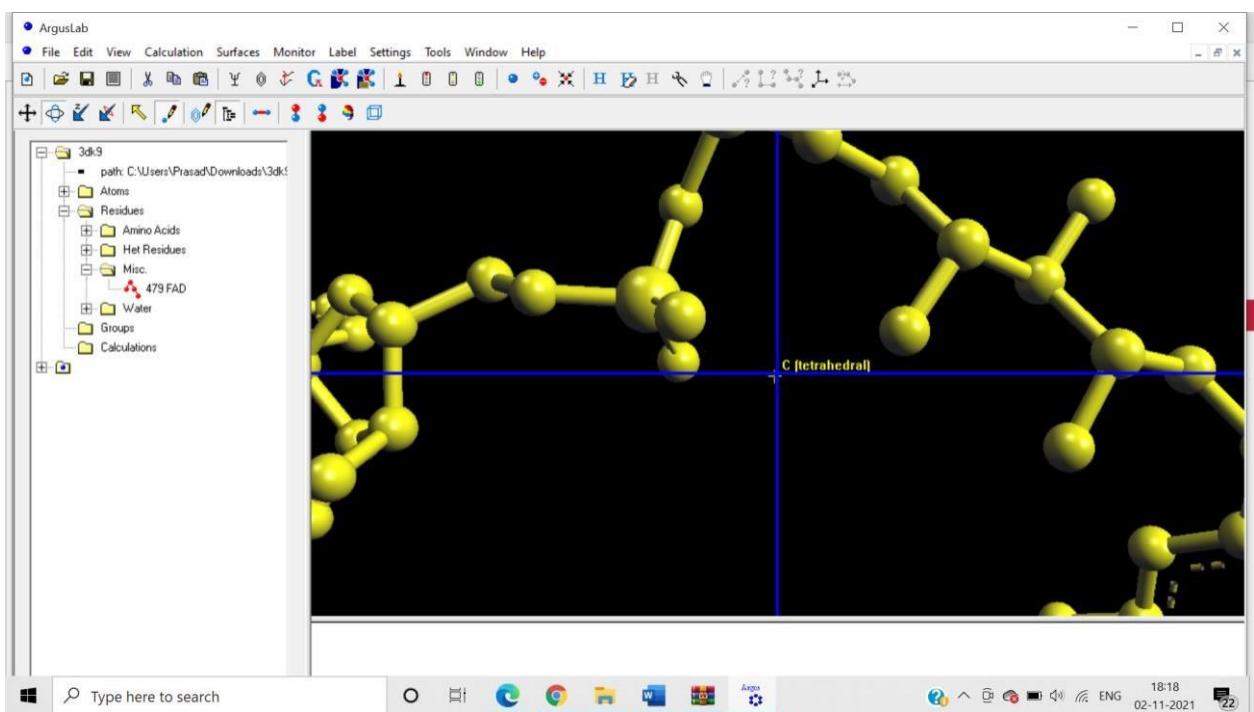
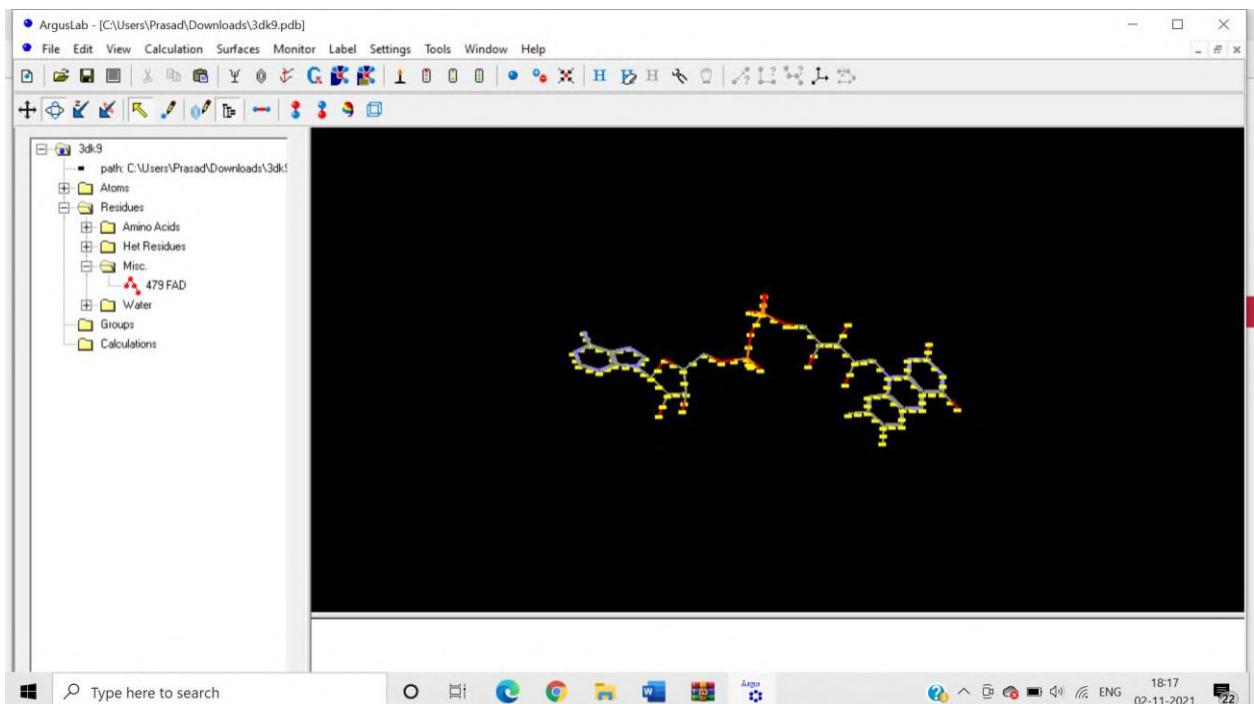
13. Click on the center the icon molecule.

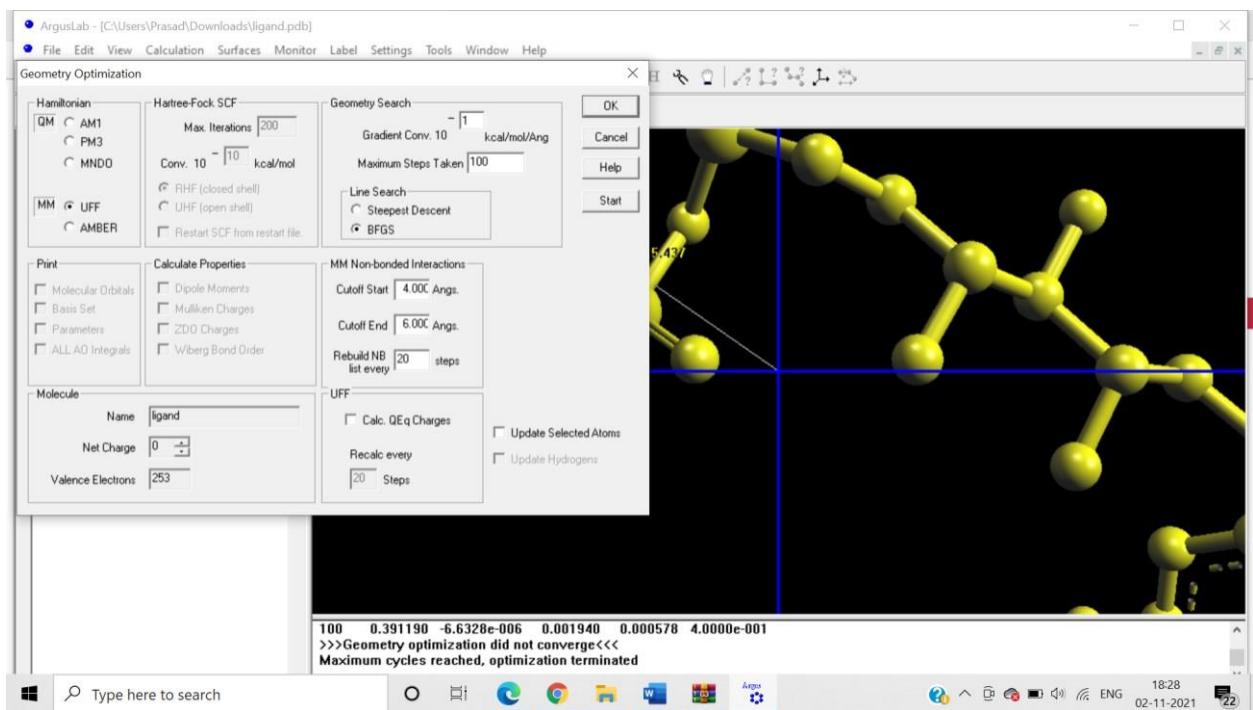
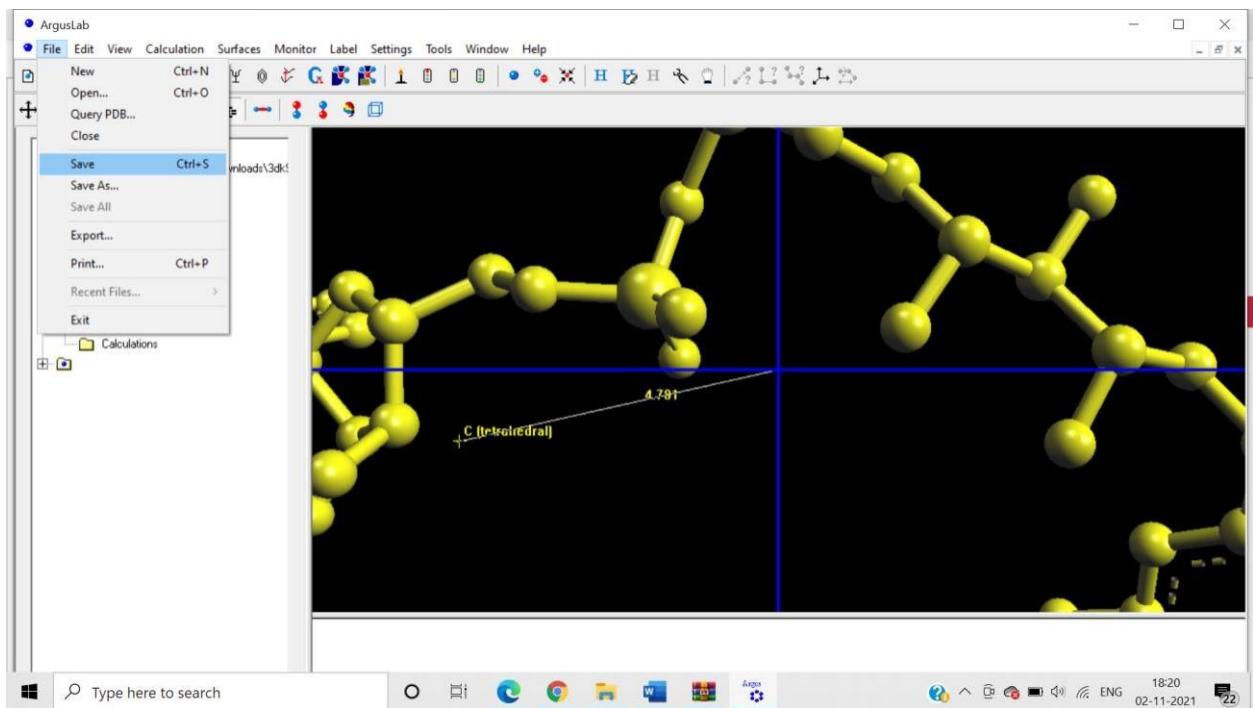


14. Copy the ligand.

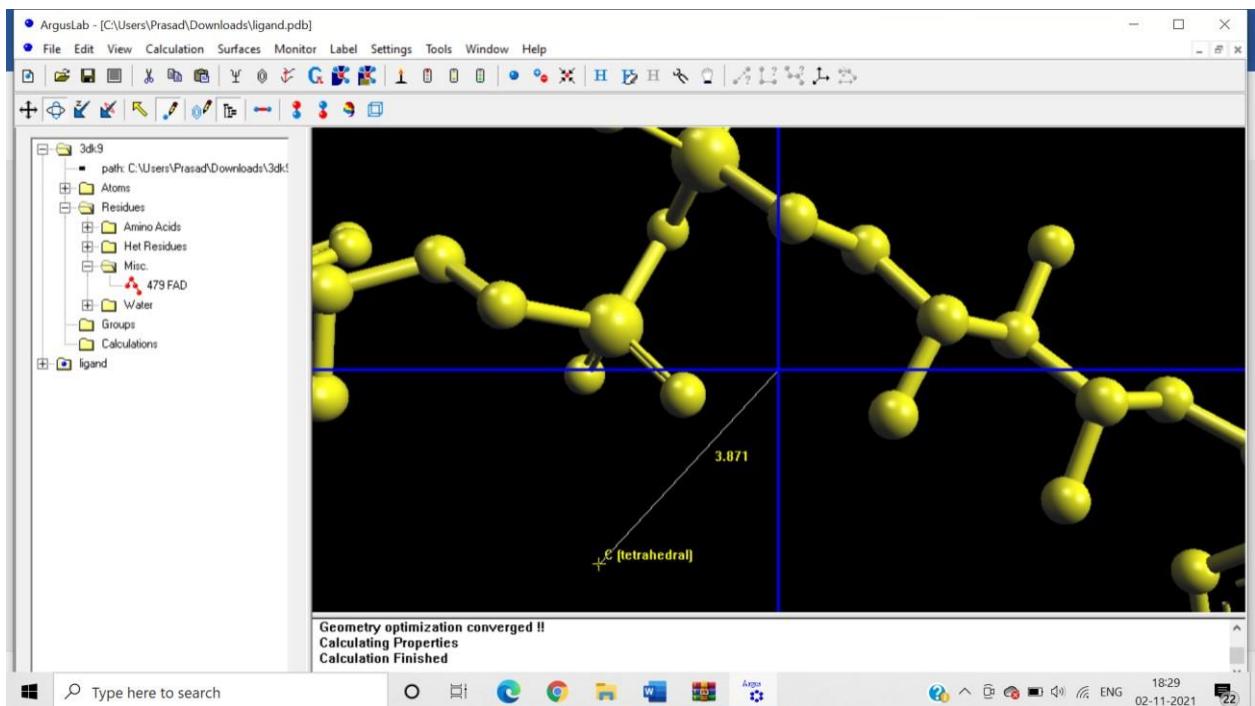
15. Create a new file and paste.

16. Click on the clean geometry icon and save it in a pdb format.





Geometry optimization converged:



This is an energy optimized structure.

Molecular Docking:

1. We have to train the system.
2. Right click on the ligand and select make a ligand group residue.
3. Group folder activated.
4. Click on group and right click on the ligand and select make a binding site group for this group.
5. Click on calculation.
6. Click on dock ligand.
7. Calculate size.
8. Start

Conclusion:- Argus Lab is used for molecular docking of ligand of pdb id. Ligand optimization was performed. After the superimposition of the optimized ligand, the old ligand is deleted. After docking we get the best optimum results of minimized energy.

Practical no: 04 B

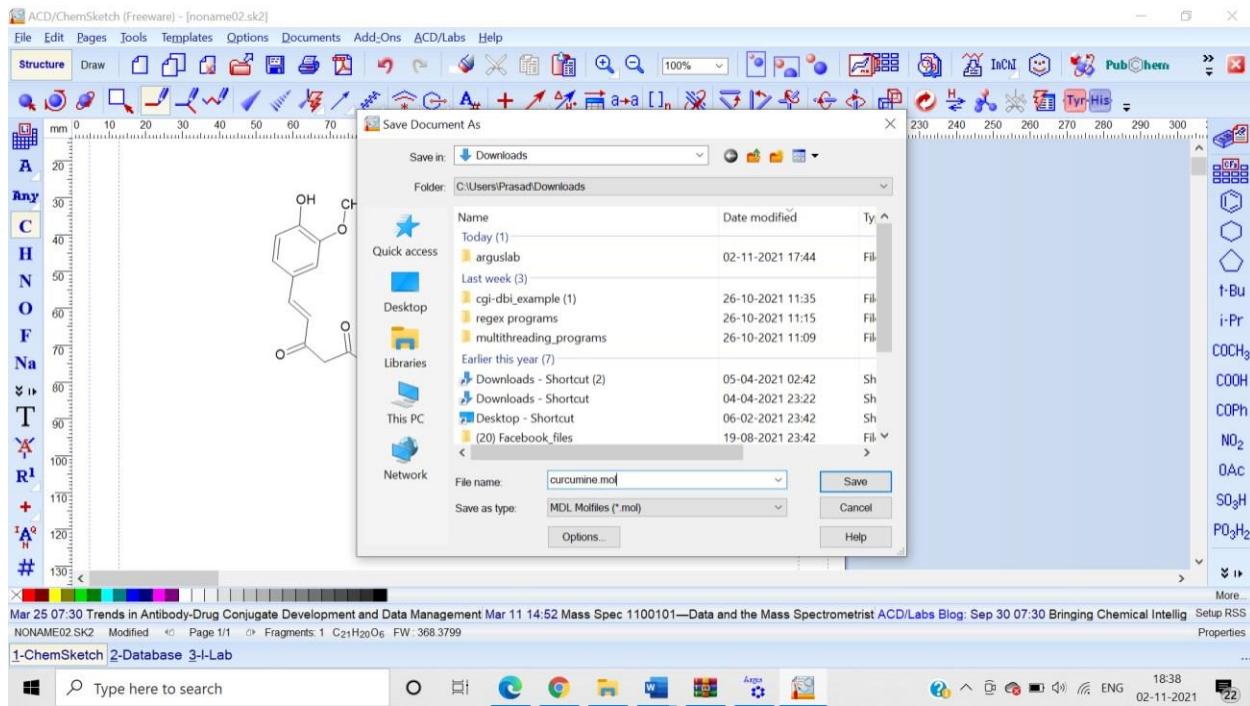
Aim: Perform molecular docking (Protein and ligand) using Genetic algorithm in Argus lab.

Theory: Molecular docking is a kind of bioinformatic modelling which involves the interaction of two or more molecules to give the stable adduct. Depending upon binding properties of ligand and target, it predicts the three-dimensional structure of any complex. Molecular docking generates different possible adduct structures that are ranked and grouped together using scoring function in the software. Docking simulations predict optimized docked conformer based upon total energy of the system. In spite of all potential approaches, ligand chemistry (tautomerism and ionization), receptor flexibility (single conformation of rigid receptor) and scoring function (differentiate true binding mode) still remained the challenge. Many important aspects of molecular docking in terms of its approaches, types, applications and challenges are briefly discussed in this article.

Methodology:

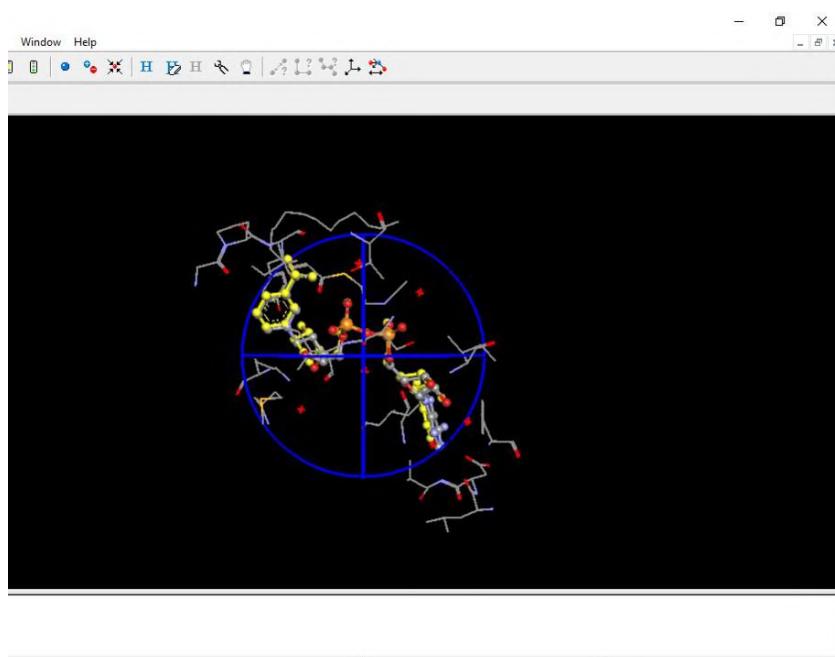
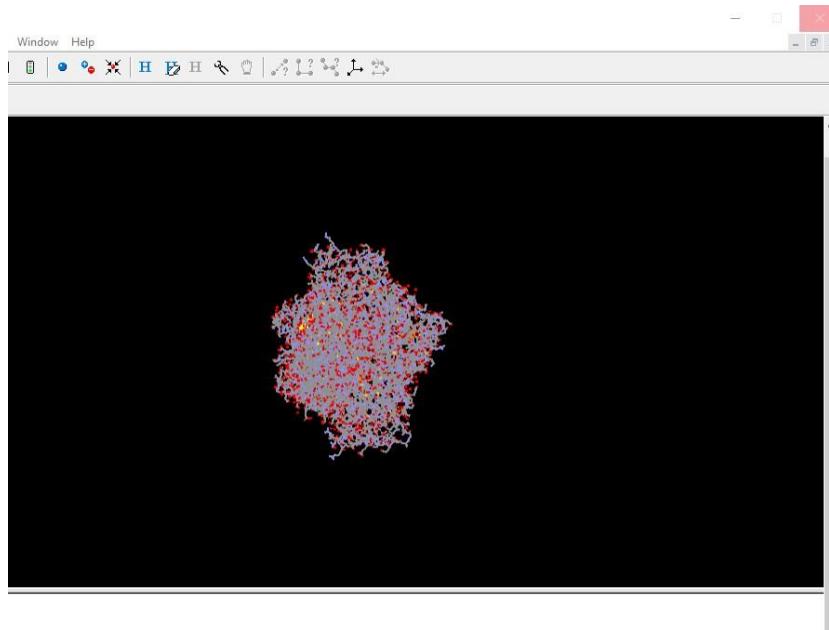
The screenshot shows a Microsoft Edge browser window with the following details:

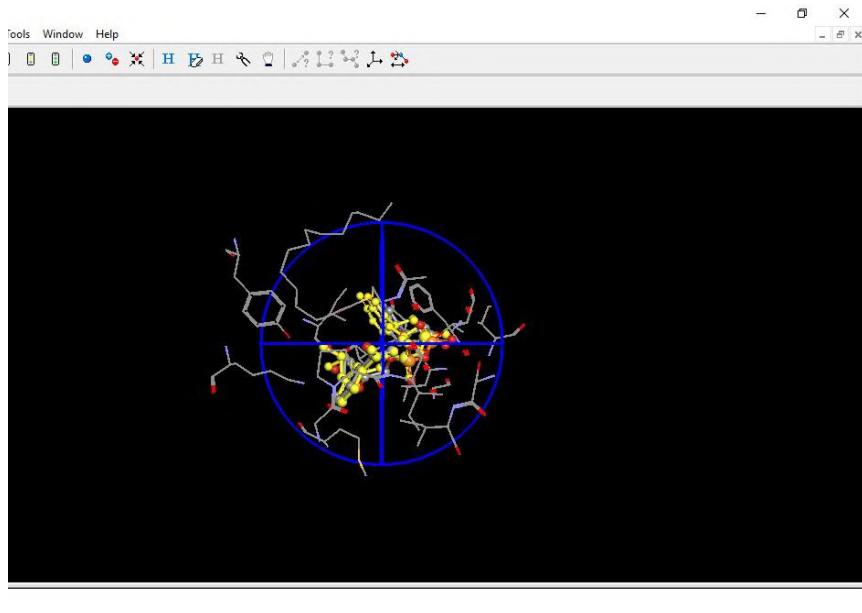
- Address Bar:** https://pubchem.ncbi.nlm.nih.gov/compound/969516
- Page Title:** COMPOUND SUMMARY
- Section Headers:** PubChem CID (969516), Structure, Chemical Safety, Molecular Formula
- Structure Section:** Shows a 2D and 3D chemical structure of Curcumin. The 2D structure is a yellow-orange molecule with a central ring system and various substituents. The 3D structure is a perspective view of the same molecule.
- Chemical Safety Section:** Contains hazard symbols for Irritant and Health Hazard, along with a link to the Laboratory Chemical Safety Summary (LCSS) Datasheet.
- Molecular Formula Section:** Shows the formula C₂₁H₂₀O₆ and the name curcumin.
- Right Sidebar:** Titled "CONTENTS", it lists sections such as Title and Summary, 1 Structures, 2 Names and Identifiers, etc., each with a dropdown arrow.
- Bottom Navigation:** Includes a search bar, taskbar icons (File, Home, Back, Forward, Stop, Refresh), and system status indicators (Windows logo, battery level, signal strength, network, volume, date/time).



Step:

1. Click on file and open a pdb file .
2. In the left hand side panel, click on the plus sign.
3. Click on residues.
4. Click on misc.
5. A list of ligands is shown, click on one (the selected part is highlighted).
6. Click on edit.
7. Click select neighbors.
8. A dialogue box will appear as shown below.
9. Click ok.
10. It will highlight the area around it (360 degrees).





11 Click on edit.

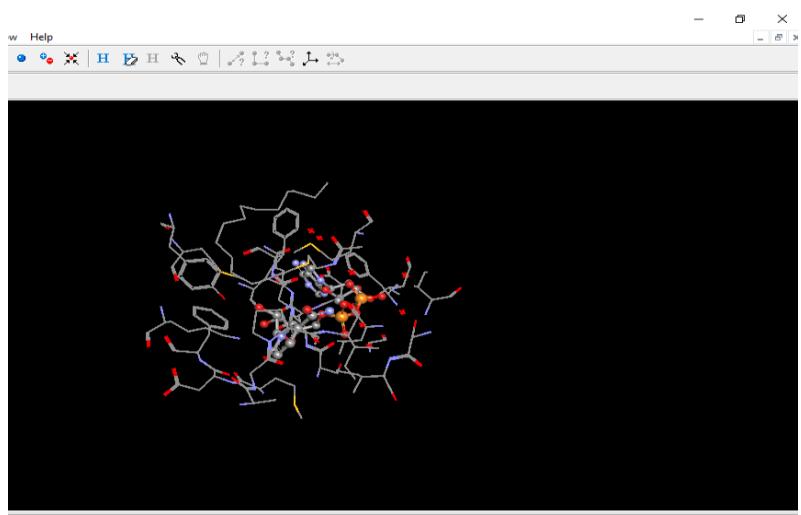
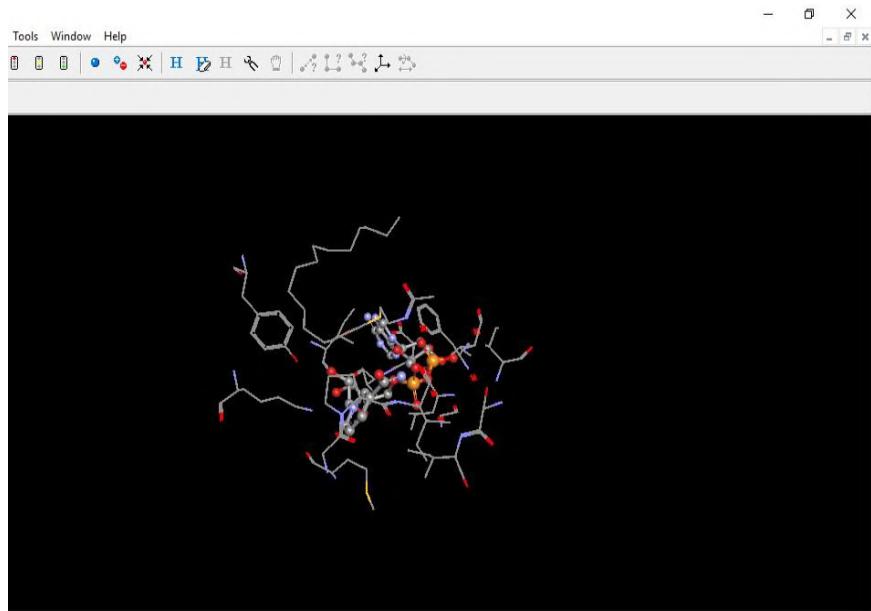
12 Click hide unselected.

13 Click on the center the icon molecule.

14 Copy the ligand.

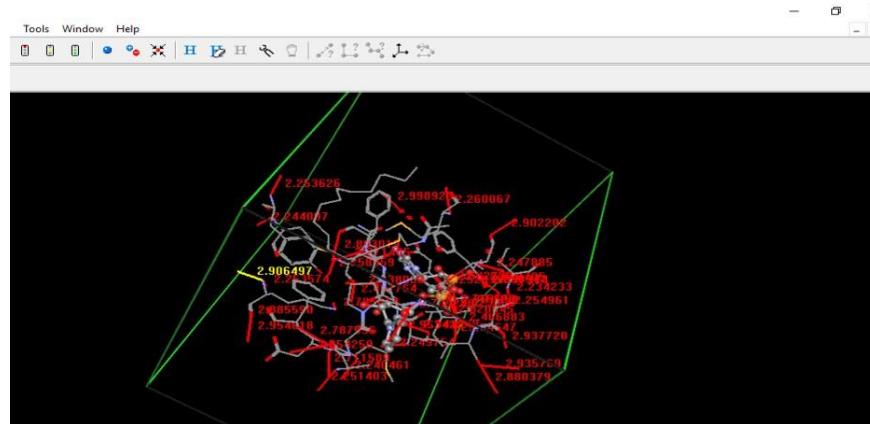
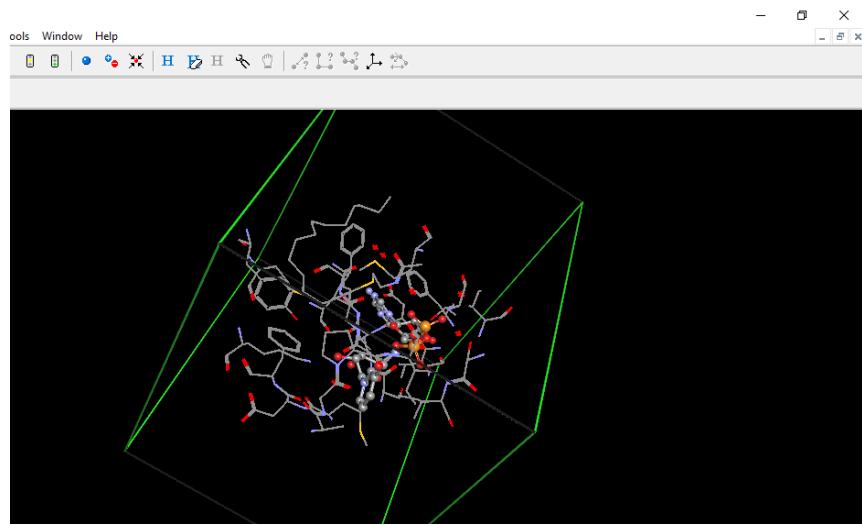
15 Create a new file and paste.

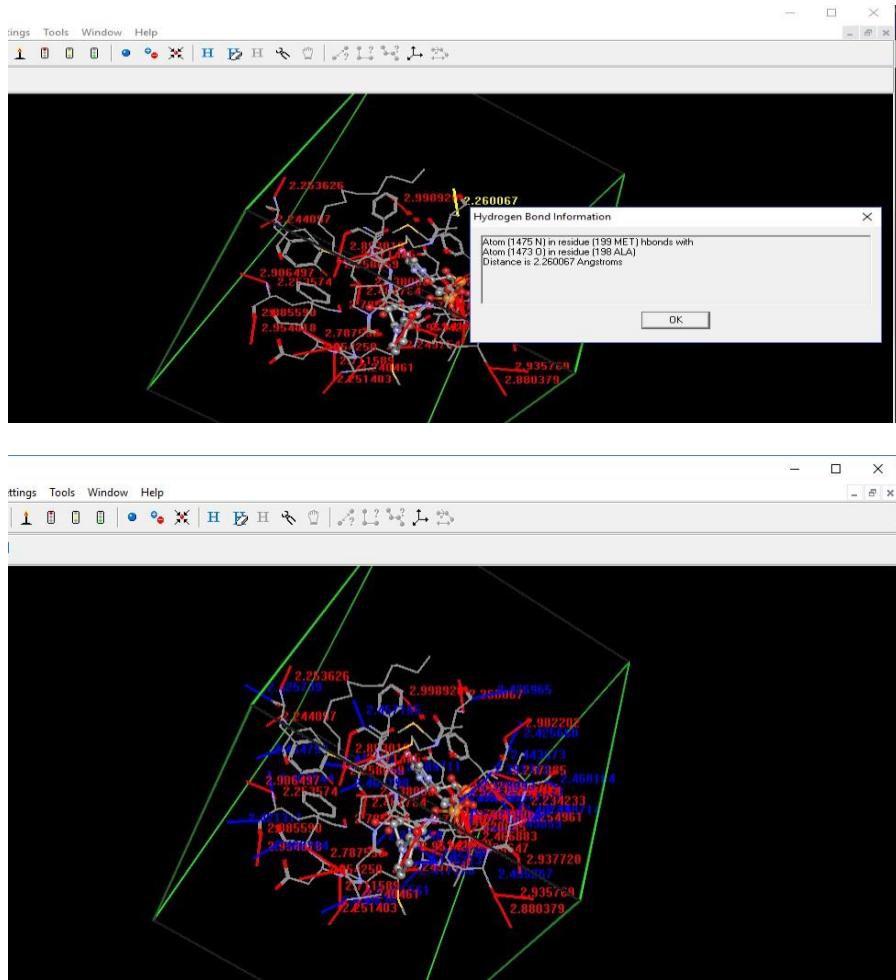
16 Click on the clean geometry icon and save it in a pdb format.



Super Imposition:

1. Select the ligand and go to settings, display and ball and cylinder high.
2. Copy paste the optimized structure onto the unoptimized structure and superimpose.
3. Delete the old ligand.





Molecular Docking:

1. We have to train the system.
2. Right click on the ligand and select make a ligand group residue.
3. Group folder activated.
4. Click on group and right click on the ligand and select make a binding site group for this group.
5. Click on calculation.
6. Click on dock ligand.
7. Calculate size.
8. Click on GADock.
9. Start.

Results :-The result is shown above,.

Conclusion:-

Argus Lab is used for molecular docking of ligand of pdb file. Ligand optimization was performed. After the superimposition of the optimized ligand, the old ligand is deleted. After docking we get the best optimum results of minimized energy.

Practical No: 05

Aim: Comparative Molecular Docking Using Argus Lab

Theory: ArgusLab is a molecular modeling, graphics, and drug design program for Windows operating systems. It's getting a little dated by now, but remains surprisingly popular. To date, there are > 20,000 downloads, ArgusLab is freely licensed. You don't need to sign anything. You can use as many copies as you need if you are teaching a class where your students might benefit from using ArgusLab. You are not allowed to redistribute ArgusLab from other websites or sources. However, you may link to this website from your own websites if you like.

First find compounds you want to do comparison between:

Disease: Glioma

Drug: Carmustine, Gliadel

So go to Pubchem to obtain structure of carmustine:

The screenshot shows the PubChem Compound Summary page for Carmustine (CID 2578). The main content area includes the PubChem CID (2578), two chemical structures (2D and 3D), hazard symbols for Acute Toxic and Health Hazard, and a link to the Laboratory Chemical Safety Summary (LCSS) Datasheet. On the right, a sidebar titled "CONTENTS" lists various sections such as Title and Summary, Structures, Names and Identifiers, Chemical and Physical Properties, Spectral Information, Related Records, Chemical Vendors, Drug and Medication Information, Pharmacology and Biochemistry, and Use and Manufacturing. At the bottom, there is a search bar, a taskbar with icons for File, Home, Google, Mail, and Word, and a system tray showing the date and time (03-11-2021, 06:58).

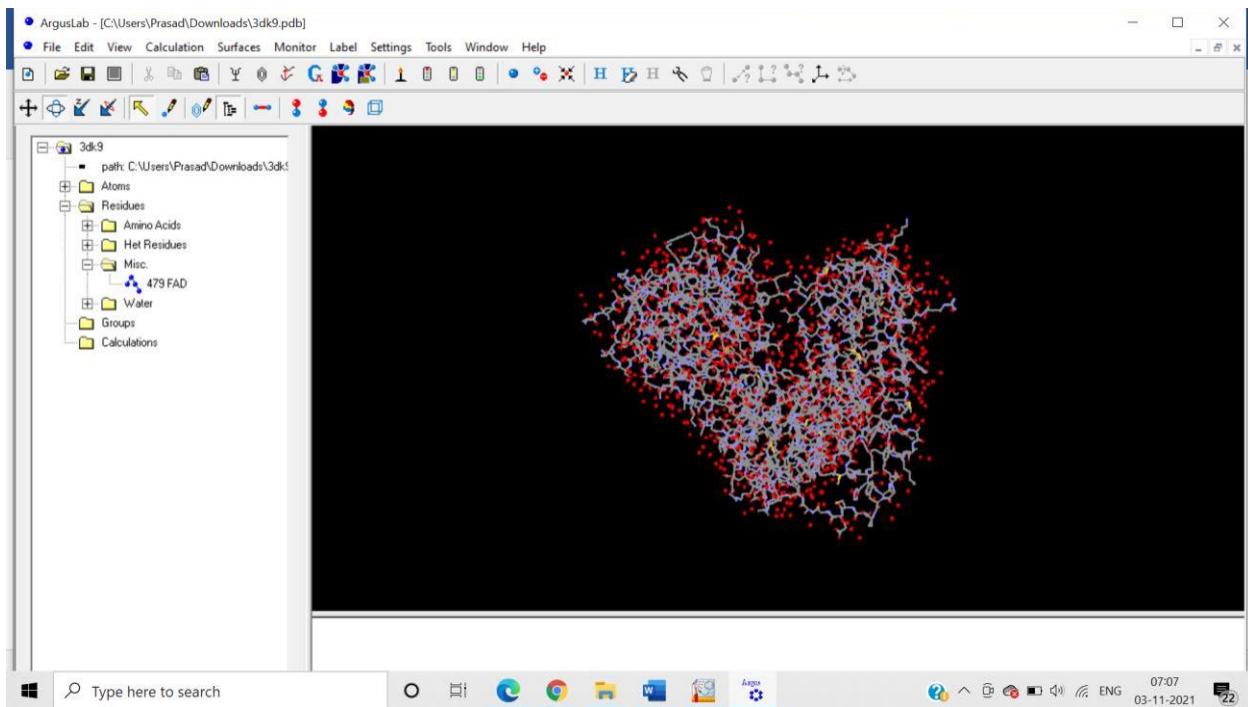
Copy the smile of Carmustine:

The screenshot shows the PubChem page for Carmustine (Compound). It displays the InChI Key (DLGOEMSEDOSKAD-UHFFFAOYSA-N) and Canonical SMILES (C(C)(C)NC(=O)N(CCCl)N=O). The page also includes sections for Molecular Formula (C₅H₉Cl₂N₃O₂) and a sidebar with various chemical properties and links.

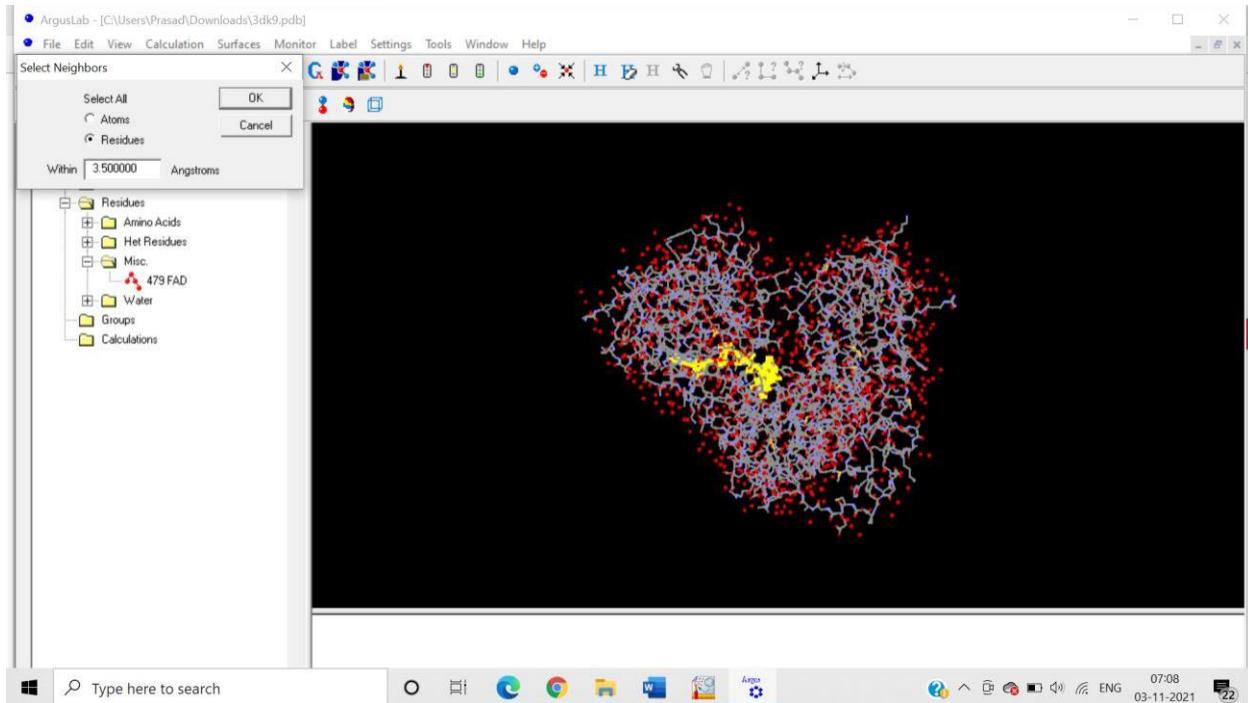
Come to chemsketch and paste It to draw the structure of the same respectively:

The screenshot shows the ACD/ChemSketch Freeware interface. A chemical structure of Carmustine (4,6-dichloro-2-methyl-N-nitro-2-nitroso-N-phenyl-s-triazine-4-amine) is drawn on the canvas. The software's toolbar and periodic table are visible, along with a status bar at the bottom.

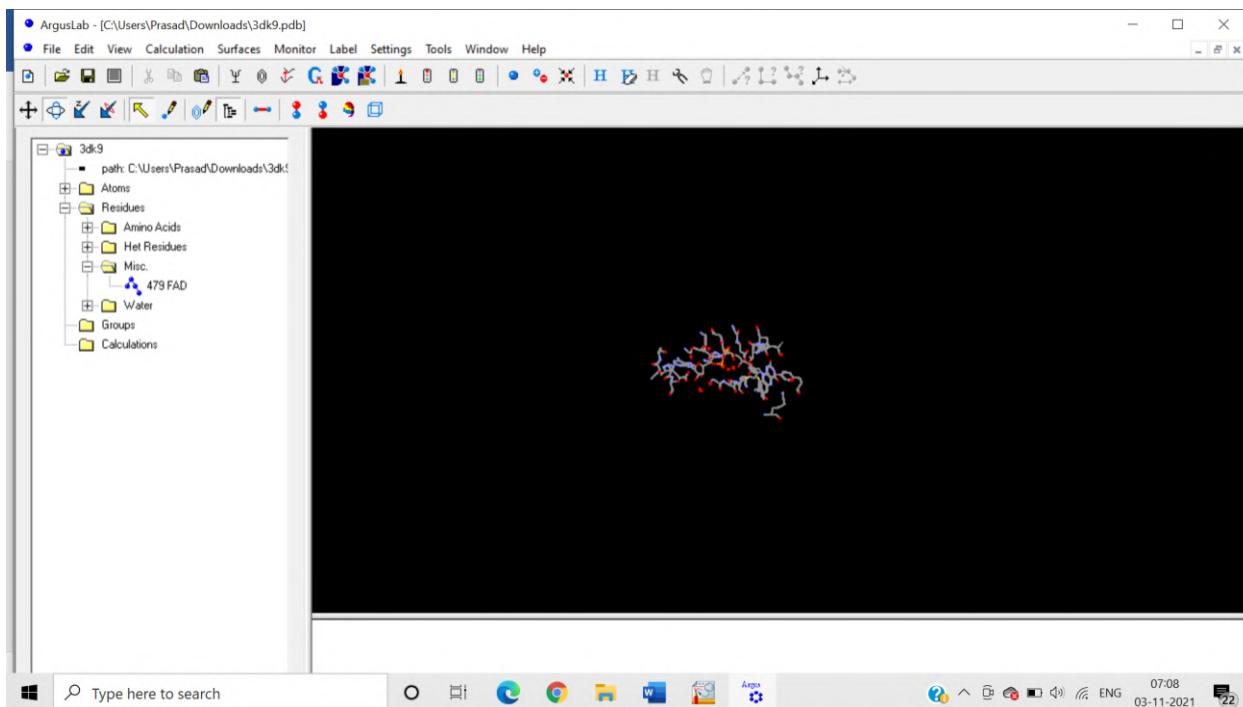
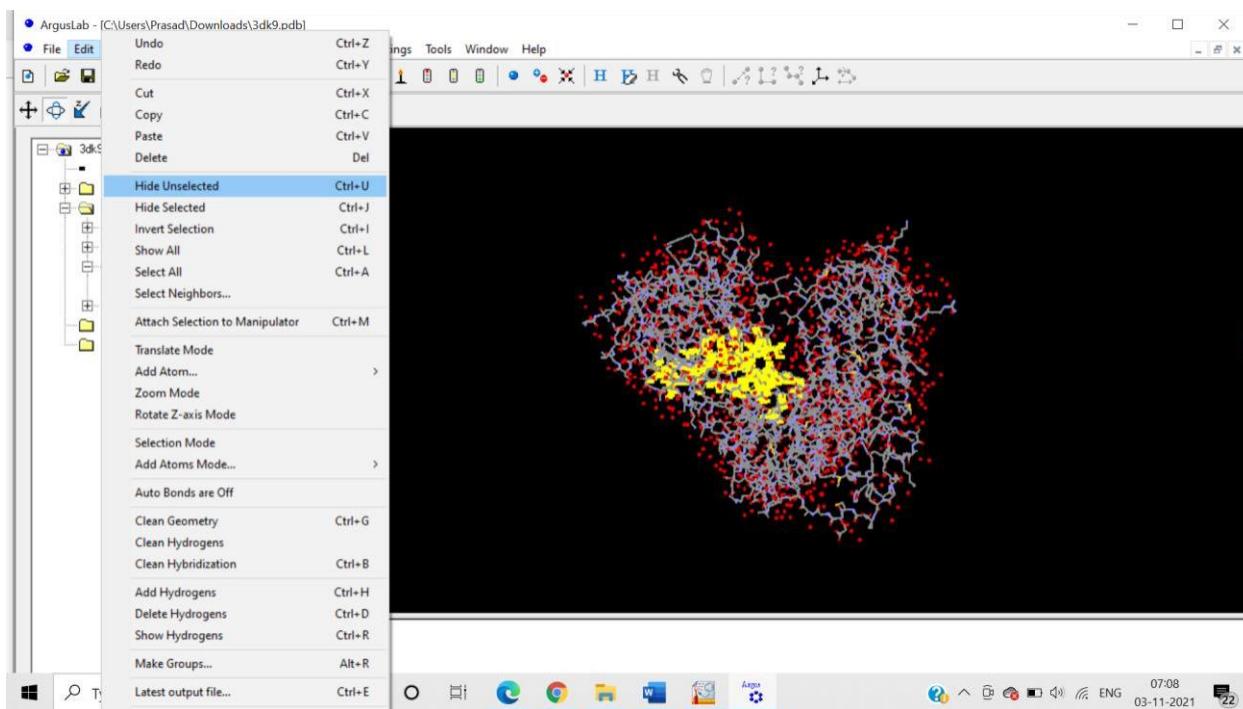
Come to Arguslab and open misc FAD structure:



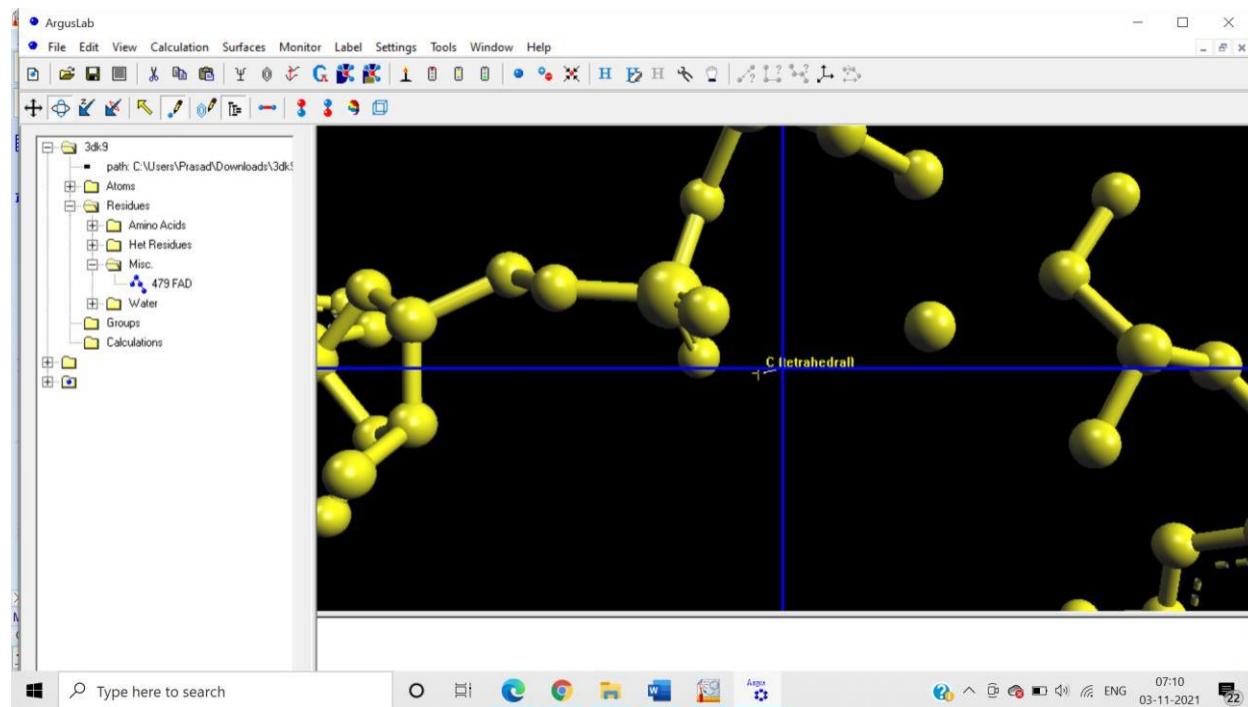
Neighbours Given: 3.500000 Angstroms



Go to hide unselected to hide the unselected structure:



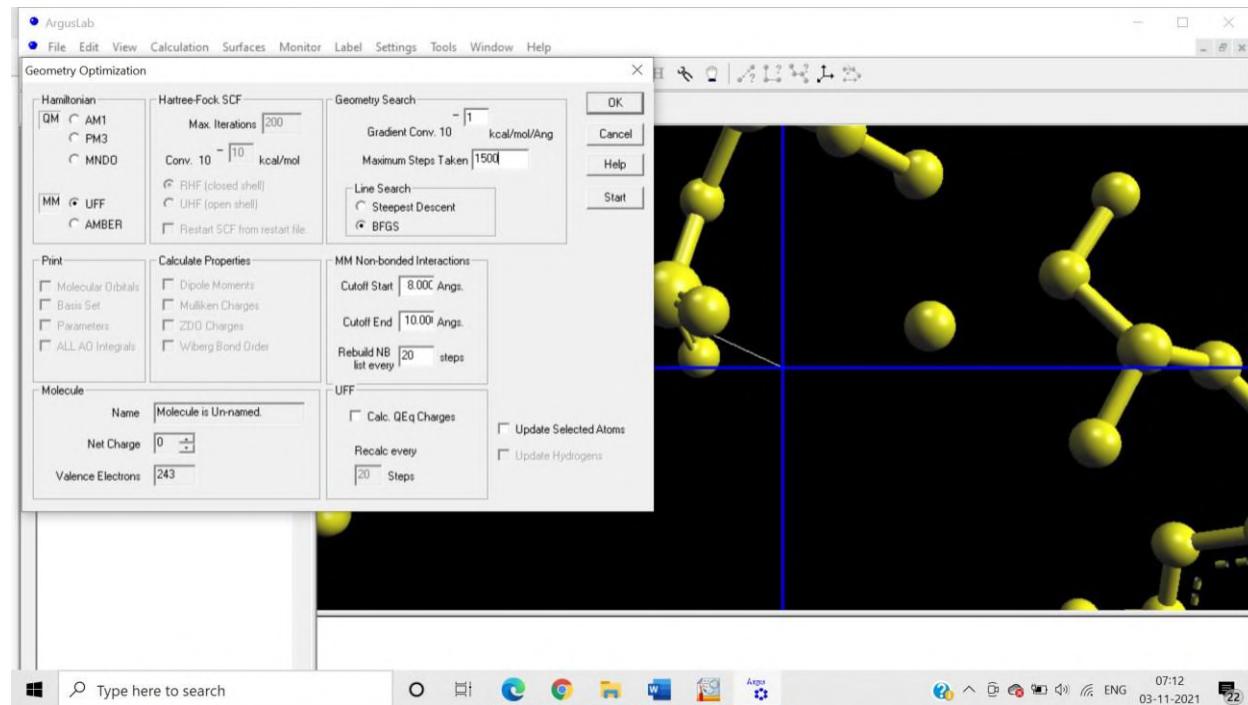
New file , Paste:



Change the settings to UFF,

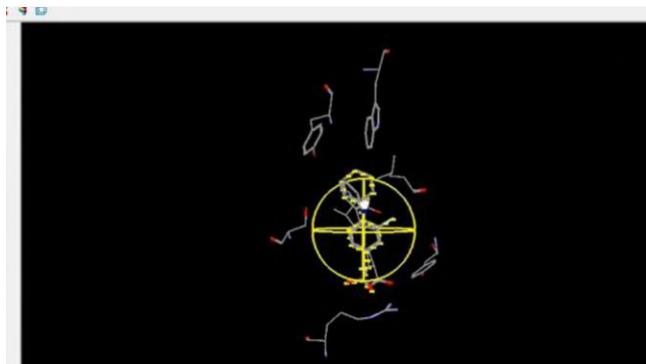
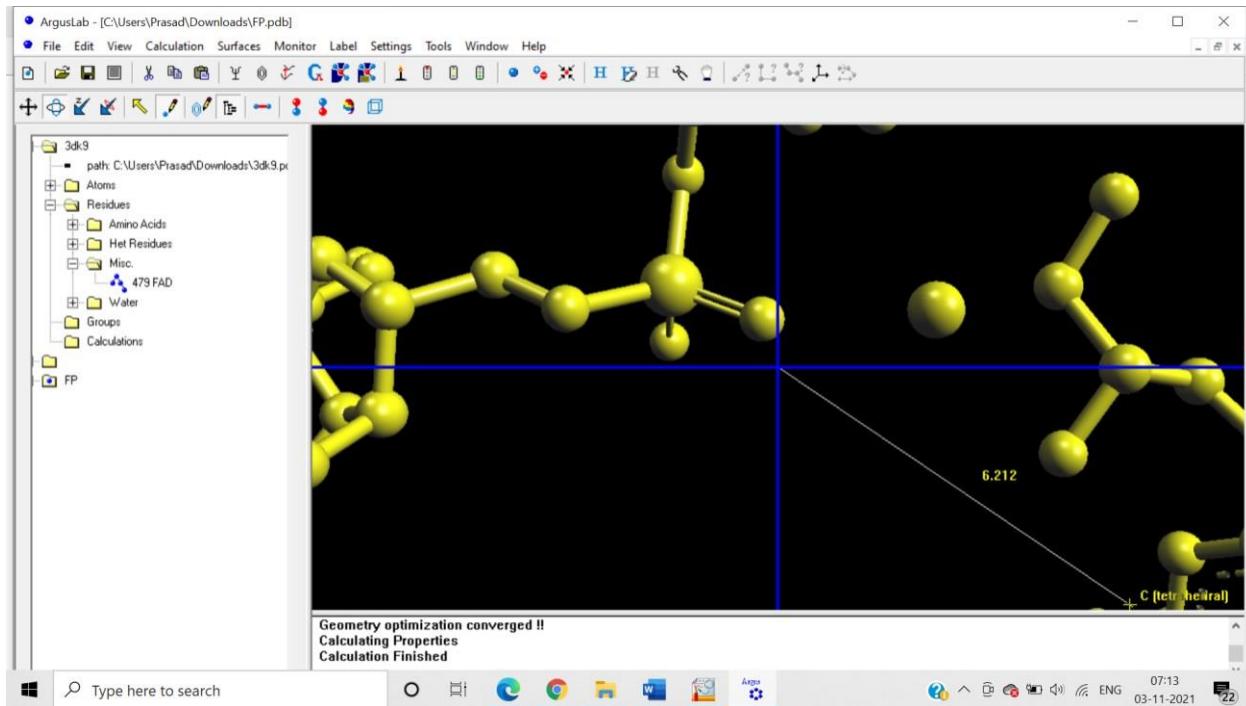
Maximum steps taken 1500

Start:

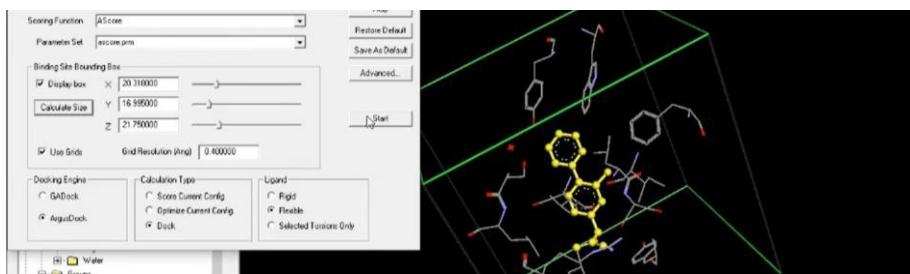


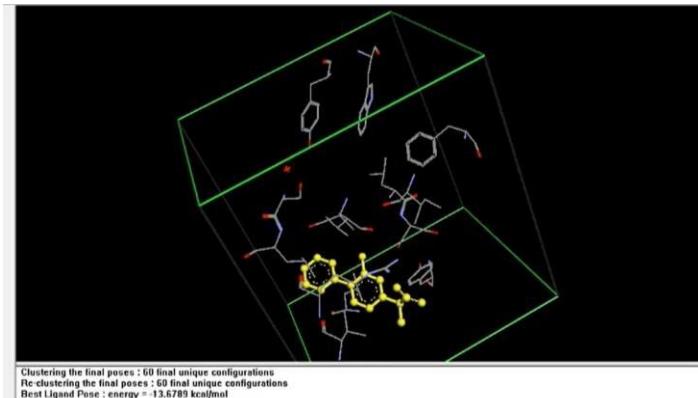
Geometry optimization converge:

Calculations finished:

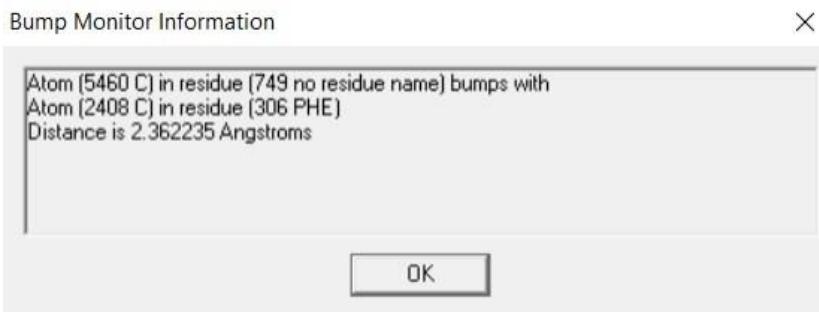
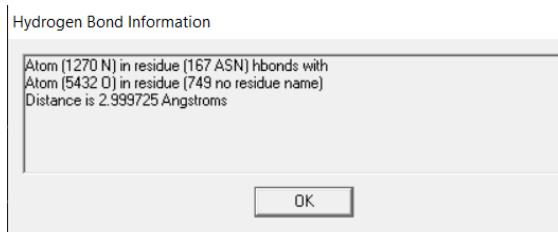


Dock ligand into binding site:





1. Report the best energy reading of the docking. Also report the hydrogen bonding and bumps.



2. Then select a different molecule, here we have used for comparative docking.
3. Optimize the molecule and repeat the entire procedure of docking using that molecule.
4. Note the readings

Result:

Hydrogen Bond Information

X

Atom (1766 N) in residue (227 TRP) hbonds with
Atom (5429 O) in residue (0 no residue name)
Distance is 2.987116 Angstroms

OK

Bump Monitor Information

X

Atom (5434 C) in residue (0 no residue name) bumps with
Atom (384 C) in residue (55 TYR)
Distance is 2.353934 Angstroms

OK

Conclusion:

Comparative docking between antihistamines drug molecule and NSAID molecule was done successfully. The interactions were compared of the compound.

Practical no: 6

Aim: Virtual Screening Using igemdock

Theory: iGEMDOCK provides biological insights by deriving the pharmacological interactions from screening compounds. The pharmacological interactions represent conserved interacting residues that often form binding pockets with specific physico-chemical properties to play the essential functions of the target protein. Experiment results show that the success rate of iGEMDOCK is 78% (root-mean-square derivations below 2.0 angstrom) on 305 protein-compound complexes. For virtual screening, pharmacological interactions derived by iGEMDOCK often involve the biological functions and enrich the hit rates on three public sets (i.e., estrogen receptor α for antagonists (ER) and agonists(ERA) and thymidine kinase (TK)); useful for understanding the ligand binding mechanisms and discovering lead compounds.

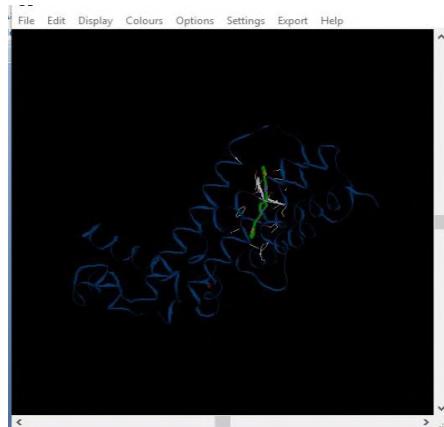
Methodology:

- 1) Setting parameters

Under docking accuracy settings → default setting → Select drug screening → START
DOCKINGPopulation Size:200 Generations:70 Number of solution:3

Results:

- 1) Display the structure



Interpretation: structure of protein complex bound with ligand is displayed

- 2) Structure with ligand



Interpretation: protein structure with ligand is displayed

3) Finished docking process

Interpretation: results are obtained after 24 minutes

4) Post-dock analysis

Interpretation: Post-dock analysis of the given compounds is obtained and presented In tabular form with values like energy, vanderwaal,H-bond,Electronegativity; molecule no-5 selatinib has highest energy

5) Interaction table



Interpretation: Interaction table for given compounds is displayed

6) View the selected docked structure

Interpretation: compounds 1 and 2 i.e.(lapatinib and selatinib) are viewed in rasmol viewer

7) Cluster analysis:

Interpretation: cluster analysis for all 6 compounds and the energy provided in interaction table were displayed , the colour coding scheme if in green describe the interaction ,whereas the area covered in black states no interaction between compound and the amino acid.

Practical no: 7

AIM: Protein docking using hex.

THEORY:

Protein Docking is not flexible but rigid. They don't make bond angles, bond deviation, etc. They make surface matches, and since they are mega structures, their degree of freedom should be restricted.

Hex is an interactive molecular graphics program for calculating and displaying feasible docking modes of pairs of protein and DNA molecules. Hex can also calculate protein-ligand docking, assuming the ligand is rigid, and it can superpose pairs of molecules using only knowledge of their 3D shapes. Hex has been available for about 12 years now, it is still the only docking and superposition program to use spherical polar Fourier (SPF) correlations to accelerate the calculations, and it's still one of the few docking programs which has built-in graphics to view the results. Also, it is the first protein docking program to be able to use modern graphics processor units (GPUs) to accelerate the calculations.

METHODOLOGY:

1. Open the receptor protein and the ligand protein on hex. Make sure they are saved in the hex example folder.
2. Keep an eye on the **R** value seen at the bottom of the screen. Once you load the ligand after the receptor, the **R** value changes.
3. Click on graphics, solid model and apply it to the receptor. Try all the models (solid model, solid surface, dot surface, harmonic surface). Make sure you apply the surfaces to either receptor or ligand, since the main purpose of changing the model is to distinguish between a receptor and the ligand.
4. Now to dock the protein, click on control, and then docking. Change Correlation type to shape + electrostatic + DARS. Keeping all the other settings on default, activate the process of docking.
5. After the process of docking, the best energy is shown in the window at the top left hand side corner.
6. In the message dialogue box, a log of all the results is shown, the first log is the best result.
7. In graphics, the animation feature will show how the docking process occurs. You can change frame and spin speed. If you notice the message box, you'll see that as the frame shifts, it logs there. If you disable the models for the receptor and ligand and activate the hydrogen bonds for the complex through the molecule option in the control button, you'll see where the ligand and the receptor forms hydrogen bonds as the ligand moves in the animation.
8. To save the result of docking, click in file and save. You can type the number of orientations you want (100 here) and the top 100 orientations will get saved. If you want

to save just 1 file for all the orientations, click on unified Multi model file. You can view the saved file on any viewer (Spdbv, biovia, discovery studio)

RESULT:

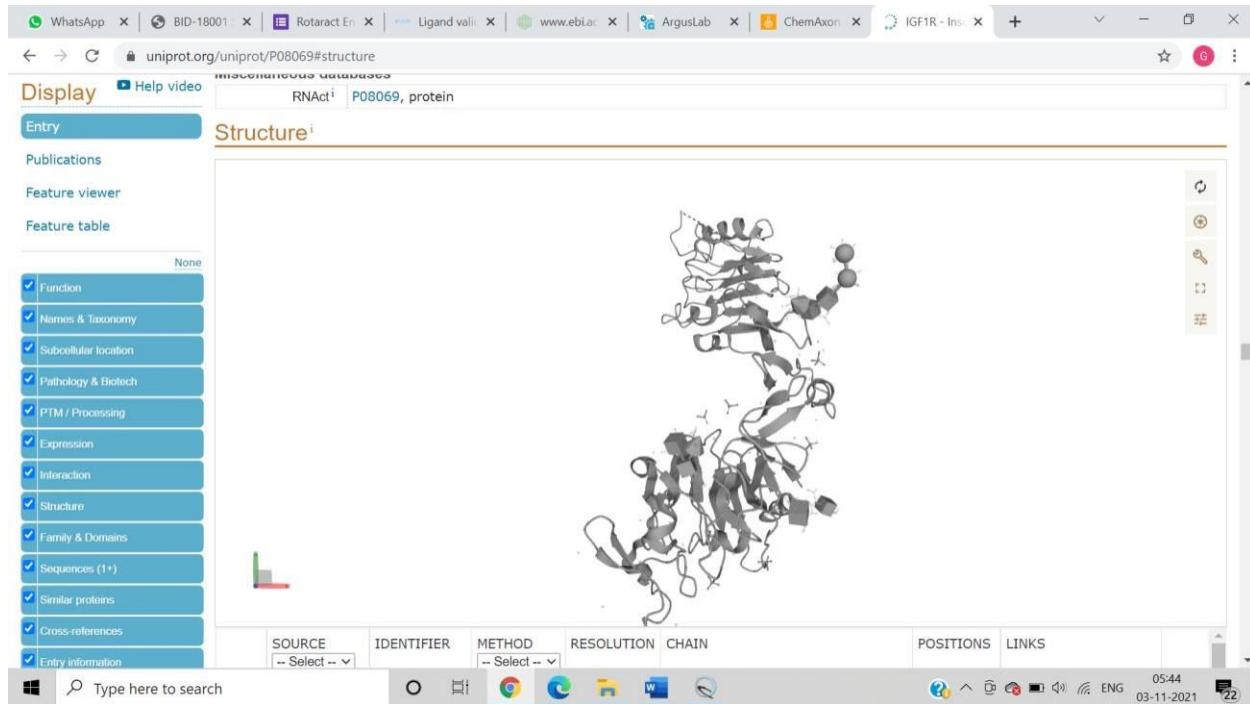
First go to UniProt and search insulin in search bar and see the number of hits.

So we got 87979 hits.

The screenshot shows the UniProtKB 2021_03 results page. The search term "insulin" is entered in the search bar. The results table displays 25 entries out of 87,979. The columns include Entry, Entry name, Protein names, Gene names, Organism, and Length. The first entry listed is P08069, IGF1R_HUMAN, Insulin-like growth factor 1 receptor, IGF1R, Homo sapiens (Human), 1,367.

Entry	Entry name	Protein names	Gene names	Organism	Length
P08069	IGF1R_HUMAN	Insulin-like growth factor 1 receptor...	IGF1R	Homo sapiens (Human)	1,367
P24062	IGF1R_RAT	Insulin-like growth factor 1 receptor...	Igf1r	Rattus norvegicus (Rat)	1,370
P09208	INSR_DROME	Insulin-like receptor	InR dinr, Dir-a, Inr-a, CG18402	Drosophila melanogaster (Fruit fly)	2,144
Q60751	IGF1R_MOUSE	Insulin-like growth	Igf1r	Mus musculus (Mouse)	1,373

Structure for P08069:



Select anyone resolution and download the pdb file:

WhatsApp | BID-18001 | Rotaract En | Ligand vali | www.ebi.ac | ArgusLab | ChemAxon | IGF1R - Ins | + - X

← → C 🔒 uniprot.org/uniprot/P08069#structure

Display Help video

Entry Publications Feature viewer Feature table

None

- Function
- Names & Taxonomy
- Subcellular location
- Pathology & Biotech
- PTM / Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequences (1+)
- Similar proteins
- Cross-references

SOURCE	IDENTIFIER	METHOD	RESOLUTION	CHAIN	POSITIONS	LINKS
PDB	3F5P	X-ray	2.90 Å	A/B/C/D/E/F/G/H/I/J/K/L/M/R/S/T	981-1286	PDB · RCSB-PDB · PDBsum
PDB	3I81	X-ray	2.08 Å	A	982-1286	PDB · RCSB-PDB · PDBsum
PDB	3LVP	X-ray	3.00 Å	A/B/C/D	951-1286	PDB · RCSB-PDB · PDBsum
PDB	3LW0	X-ray	1.79 Å	A/B/C/D	983-1286	PDB · RCSB-PDB · PDBsum
PDB	3NWS	X-ray	2.14 Å	A	982-	PDB · RCSB-PDB ·

Secondary structure 1 1367

Legend: Helix Turn Beta strand PDB Structure known for this area

Show more details

3D structure databases

Type here to search 05:45 03-11-2021 22

Download the pdb file of 3LW0:

The screenshot shows the RCSB PDB website for structure 3LW0. The main content includes:

- Structure Summary:** Shows a 3D ribbon model of the IGF-1RK protein-ligand complex.
- Experimental Data Snapshot:** Includes Method: X-RAY DIFFRACTION, Resolution: 1.79 Å, R-Value Free: 0.185, and R-Value Work: 0.157.
- wwPDB Validation:** A chart showing percentile ranks for various metrics. The data is as follows:

Metric	Percentile Ranks	Value
Rfree	1	0.195
Clashscore	1	1
Ramachandran outliers	0	0
Sidechain outliers	1.4%	1.4%
RSRZ outliers	2.6%	2.6%

Legend: Worse (red), Better (blue). Percentile relative to all X-ray structures. Percentile relative to X-ray structures of similar resolution.

Classification: TRANSFERASE

Organism(s): Homo sapiens

Expression System: Spodoptera frugiperda

Mutation(s): Yes

Download the pdb file of 3LW0:

WhatsApp | BID-18001 | Rotaract En | Ligand vali... | www.ebi.ac... | ArgusLab | ChemAxon | RCSB PDB

rcsb.org/structure/3LW0

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation Careers MyPDB

Structure Summary 3D View Annotations Experiment Sequence Genome Ligands Versions

3LW0

IGF-1RK in complex with ligand MSC1609119A-1

DOI: 10.2210/pdb3LW0/pdb

Classification: TRANSFERASE
Organism(s): Homo sapiens
Expression System: Spodoptera frugiperda
Mutation(s): Yes

Deposited: 2010-02-23 Released: 2010-09-29
Deposition Author(s): Graedler, U., Heinrich, T., Boeltcher, H., Blaukat, A., S.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION Resolution: 1.79 Å R-Value Free: 0.185 R-Value Work: 0.157

wwPDB Validation

Metric	Rfree
	Clashscore
Ramachandran outliers	Sidechain outliers
RSRZ outliers	Worse Percent
Better Percent	

Global Symmetry: Asymmetric - C1 https://files.rcsb.org/download/3LW0.pdb (monomer - A1)

3D View: Structure | Electron Density | Ligand Interaction

Type here to search

Display Files Download Files

FASTA Sequence
PDB Format
PDB Format (gz)
PDB/mmCIF Format
PDB/mmCIF Format (gz)
PDB/XML Format (gz)

Contact Us

Biological Assembly 1
Biological Assembly 2
Biological Assembly 3
Biological Assembly 4
Structure Factors (CIF)
Structure Factors (CIF - gz)

Validation Full PDF

05:46 03-11-2021 22

WhatsApp | BID-18001 | Rotaract En | Ligand vali... | www.ebi.ac... | ArgusLab | ChemAxon | InR - Insulin | RCSB PDB

uniprot.org/uniprot/P09208

UniProt UniProtKB Advanced Search

BLAST Align Retrieve/ID mapping Peptide search SPARQL Help Contact

The new UniProt website is here! Take me to UniProt BETA

UniProtKB - P09208 (INSR_DROME)

Display Help video BLAST Align Format Add to basket History Add a publication Feedback

Entry

Protein: **Insulin-like receptor**
Gene: **InR**
Organism: *Drosophila melanogaster* (Fruit fly)
Status: Reviewed - Annotation score: 5 - Experimental evidence at protein level

Function:

None

Function: Has a ligand-stimulated tyrosine-protein kinase activity. Required for cell survival. Regulates body size and organ size by altering cell number and cell size in a cell-autonomous manner. Involved in the development of the embryonic nervous system, and is necessary for axon guidance and targeting in the visual system. Also plays a role in life-span determination.

6 Publications

3lw0.pdb

Show all

Type here to search

05:48 03-11-2021 22

P06213 (INSR_HUMAN)

Receptor tyrosine kinase which mediates the pleiotropic actions of insulin. Binding of insulin leads to phosphorylation of several intracellular substrates, including, insulin receptor substrates (IRS1, 2, 3, 4), SHC, GAB1, CBL and other signaling intermediates. Each of these

phosphorylated proteins serve as docking proteins for other signaling proteins that contain Src-homology-2 domains (SH2 domain) that specifically recognize different phosphotyrosine residues, including the p85 regulatory subunit of PI3K and SHP2. Phosphorylation of IRSs proteins lead to the activation of two main signaling pathways: the PI3K-AKT/PKB pathway, which is responsible for most of the metabolic actions of insulin, and the Ras-MAPK pathway, which regulates expression of some genes and cooperates with the PI3K pathway to control cell growth and differentiation. Binding of the SH2 domains of PI3K to phosphotyrosines on IRS1 leads to the activation of PI3K and the generation of phosphatidylinositol-(3, 4, 5)-triphosphate (PIP3), a lipid second messenger, which activates several PIP3-dependent serine/threonine kinases, such as PDPK1 and subsequently AKT/PKB. The net effect of this pathway is to produce a translocation of the glucose transporter SLC2A4/GLUT4 from cytoplasmic vesicles to the cell membrane to facilitate glucose transport. Moreover, upon insulin stimulation, activated AKT/PKB is responsible for: anti-apoptotic effect of insulin by inducing phosphorylation of BAD; regulates the expression of gluconeogenic and lipogenic enzymes by controlling the activity of the winged helix or forkhead (FOX) class of transcription factors. Another pathway regulated by PI3K-AKT/PKB activation is mTORC1 signaling pathway which regulates cell growth and metabolism and integrates signals from insulin. AKT mediates insulin-stimulated protein synthesis by phosphorylating TSC2 thereby activating mTORC1 pathway. The Ras/RAF/MAP2K/MAPK pathway is mainly involved in mediating cell growth, survival and cellular differentiation of insulin. Phosphorylated IRS1 recruits GRB2/SOS complex, which triggers the activation of the Ras/RAF/MAP2K/MAPK pathway. In addition to binding insulin, the insulin receptor can bind insulin-like growth factors (IGF1 and IGFII). Isoform Short has a higher affinity for IGFII binding. When present in a hybrid receptor with IGF1R, binds IGF1. PubMed:12138094 shows that hybrid receptors composed of IGF1R and INSR isoform Long are activated with a high affinity by IGF1, with low affinity by IGF2 and not significantly activated by insulin, and that hybrid receptors composed of IGF1R and INSR isoform Short are activated by IGF1, IGF2 and insulin. In contrast, PubMed:16831875 shows that hybrid receptors composed of IGF1R and INSR isoform Long and hybrid receptors composed of IGF1R and INSR isoform Short have similar binding characteristics, both bind IGF1 and have a low affinity for insulin.

UniProtKB - P06213 (INSR_HUMAN)

Display Help video BLAST Align Format Add to basket History Add a publication Feedback

Entry

Protein Insulin receptor
Gene INSR
Organism Homo sapiens (Human)
Status Reviewed - Annotation score: 5 - Experimental evidence at protein levelⁱ

Function

Receptor tyrosine kinase which mediates the pleiotropic actions of insulin. Binding of insulin leads to phosphorylation of several intracellular substrates, including, insulin receptor substrates (IRS1, 2, 3, 4), SHC, GAB1, CBL and other signaling intermediates. Each of these phosphorylated proteins serve as docking proteins for other signaling proteins that contain Src-homology-2 domains (SH2 domain) that specifically recognize different phosphotyrosine residues, including the p85 regulatory subunit of PI3K and SHP2. Phosphorylation of IRS proteins lead to the activation of two main signaling pathways: the PI3K-AKT/PKB pathway, which is responsible for most of the metabolic actions of insulin, and the Ras-MAPK pathway, which regulates expression of some genes and cooperates with the PI3K pathway to control cell growth and differentiation. Binding of the SH2 domains of PI3K to phosphotyrosines on IRS1 leads to the activation of PI3K and the generation of phosphatidylinositol-(3, 4, 5)-triphosphate (PIP3), a lipid second messenger, which activates

3lw0.pdb Show all

Type here to search 05:49 03-11-2021

Download the pdb file of 1IR3:

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation Careers MyPDB

183584 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Enter search terms or PDB ID(s).

Advanced Search | Browse Annotations Help

PDB-100 PDB Protein Data Bank EAMDataResource Nucleic Acid Database Worldwide Protein Data Bank Foundation

Celebrating 40 YEARS OF Protein Data Bank

Structure Summary 3D View Annotations Experiment Sequence Genome Ligands Versions Contact Us

Biological Assembly 1 1IR3 Display Files Download Files

PHOSPHORYLATED INSULIN RECEPTOR TYROSINE KINASE IN COMPLEX WITH PEPTIDE SUBSTRATE AND ATP ANALOG

DOI: 10.2210/pdb1IR3/pdb

Classification: COMPLEX (TRANSFERASE/SUBSTRATE)
Organism(s): Homo sapiens
Expression System: Spodoptera frugiperda
Mutation(s): Yes

Deposited: 1997-09-22 Released: 1998-01-07

3lw0.pdb Show all

Type here to search 05:51 03-11-2021

1IR3

PHOSPHORYLATED INSULIN RECEPTOR TYROSINE KINASE IN COMPLEX WITH PEPTIDE SUBSTRATE AND ATP ANALOG

WhatsApp | BID-18001 | Rotaract En | Ligand vali | www.ebi.ac | ArgusLab | ChemAxon | Igf1r - Insu | + | - | X

← → 🔍 uniprot.org/uniprot/Q60751#structure

UniProt UniProtKB Advanced Search

BLAST Align Retrieve/ID mapping Peptide search SPARQL Help Contact

The new UniProt website is here! Take me to UniProt BETA

UniProtKB - Q60751 (IGF1R_MOUSE)

Display Help video BLAST Align Format Add to basket History Add a publication Feedback

Entry Publications Feature viewer Feature table

Protein: **Insulin-like growth factor 1 receptor**
 Gene: **Igf1r**
 Organism: **Mus musculus (Mouse)**
 Status: **Reviewed - Annotation score: 5/5 - Experimental evidence at protein level**

Function: Receptor tyrosine kinase which mediates actions of insulin-like growth factor 1 (IGF1). Binds IGF1 with high affinity and IGF2 and insulin (INS) with a lower affinity. The activated IGF1R is involved in cell growth and survival control. IGF1R is crucial for tumor transformation and survival of malignant cell. Ligand binding activates the receptor kinase, leading to receptor autoprophosphorylation, and tyrosines phosphorylation of multiple substrates, that function as signaling adapter proteins including, the insulin-receptor substrates (IRS1/2), Shc and 14-3-3 proteins. Phosphorylation of IRSs proteins lead to the

1ir3.pdb 3lw0.pdb Show all

Type here to search 05:53 03-11-2021

Q60751 (IGF1R_MOUSE)

WhatsApp | BID-18001 | Rotaract En | Ligand vali | www.ebi.ac | ArgusLab | ChemAxon | Igf1r - Insu | + | - | X

← → 🔍 uniprot.org/uniprot/Q60751#structure

Display Help video Structure

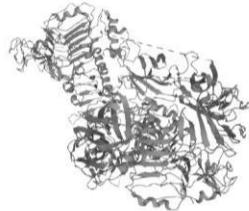
Entry Publications Feature viewer Feature table

None

Function
 Names & Taxonomy
 Subcellular location
 Pathology & Biotech
 PTM / Processing
 Expression
 Interaction
 Structure
 Family & Domains
 Sequences (1+)
 Similar proteins

1ir3.pdb 3lw0.pdb Show all

Type here to search 05:53 03-11-2021

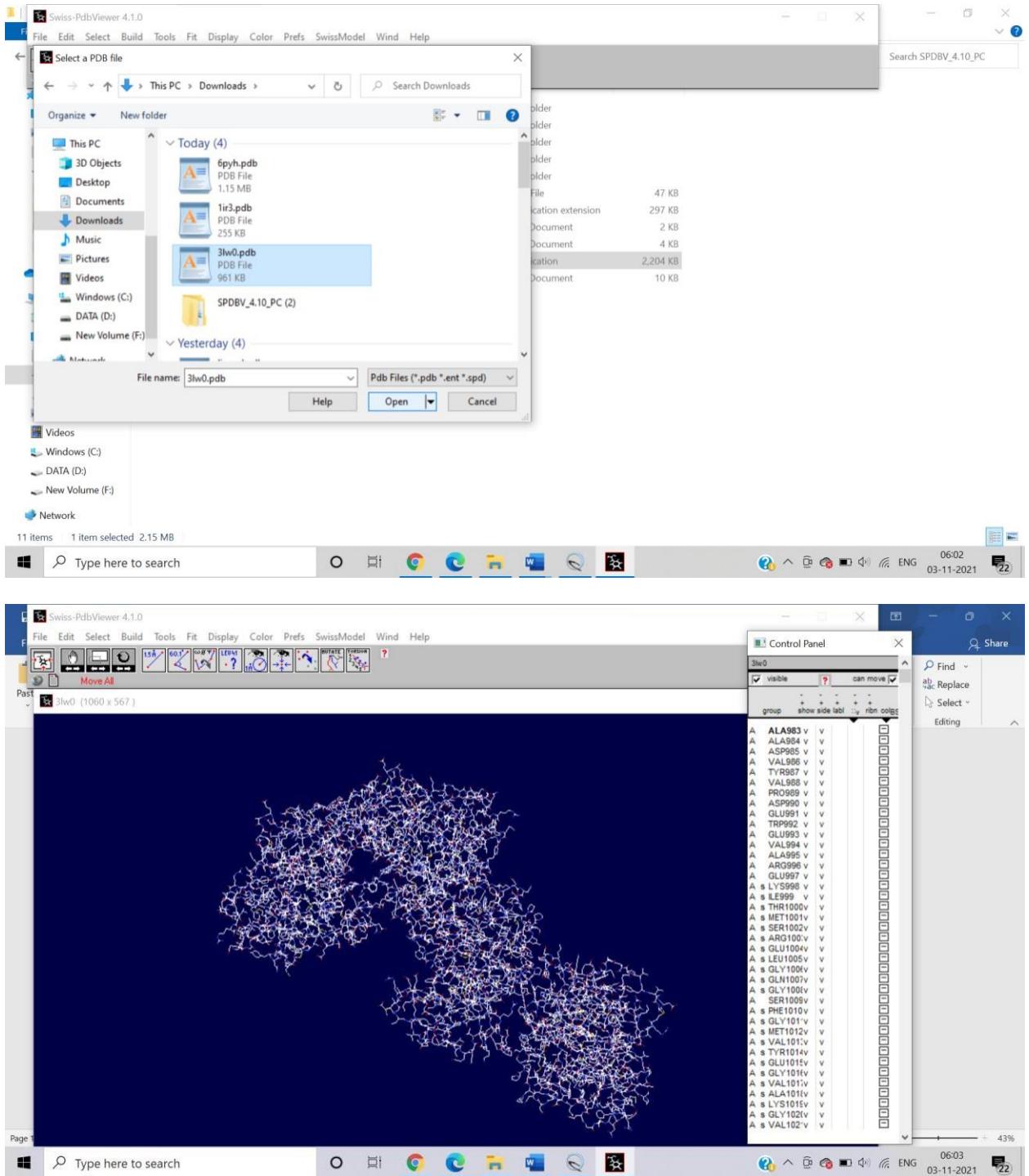


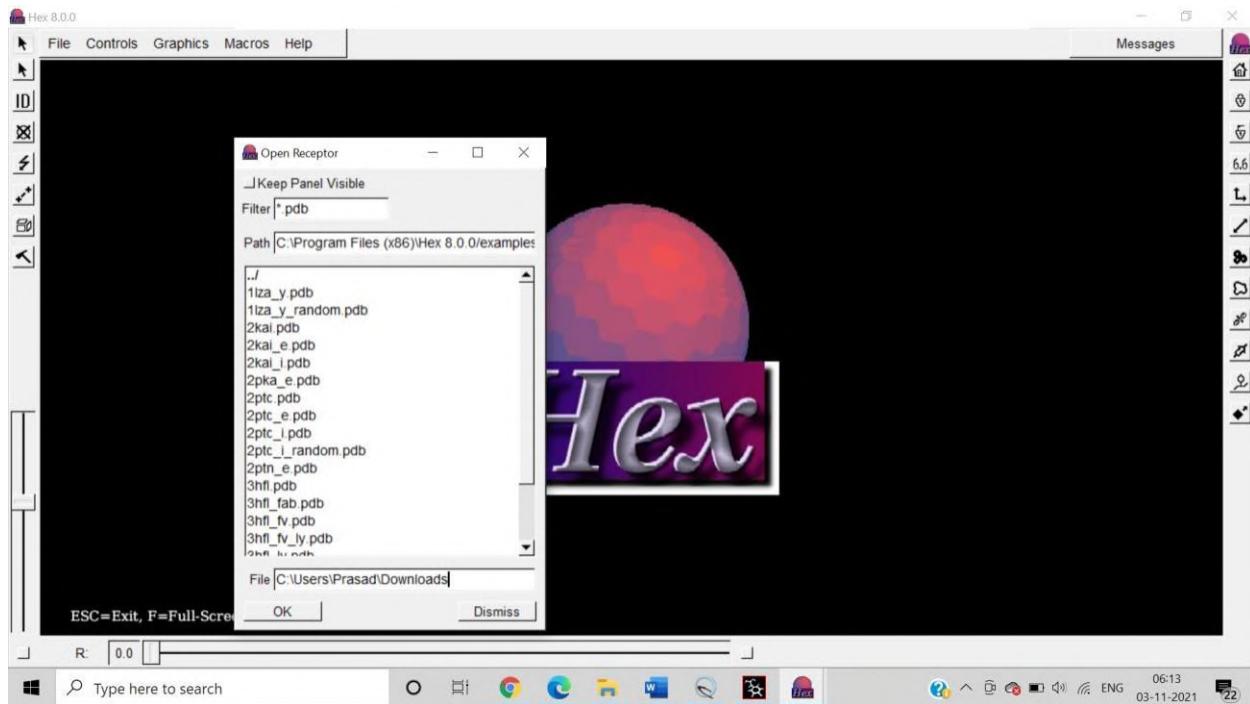
Download the pdb file of 6PYH:

The screenshot shows the RCSB PDB website interface. At the top, there is a navigation bar with links for RCSB PDB, Deposit, Search, Visualize, Analyze, Download, Learn, More, Documentation, and Careers. A search bar is present with the placeholder "Enter search terms or PDB ID(s)". Below the search bar, there is a banner for "Celebrating 50 YEARS OF Protein Data Bank". The main content area displays the structure summary for PDB ID 6PYH. It features a 3D ribbon model of the protein complex. To the left, there is a "Biological Assembly 1" section with arrows for navigation. To the right, detailed information about the structure is provided, including its classification as a SIGNALING PROTEIN/HORMONE, organisms involved (Mus musculus, Homo sapiens), expression system (Homo sapiens, Escherichia coli), and mutation status (No). Below this, deposition and release dates are listed as 2019-07-29 and 2019-10-23 respectively. At the bottom of the page, there is a search bar with the text "Type here to search" and a toolbar with various icons.

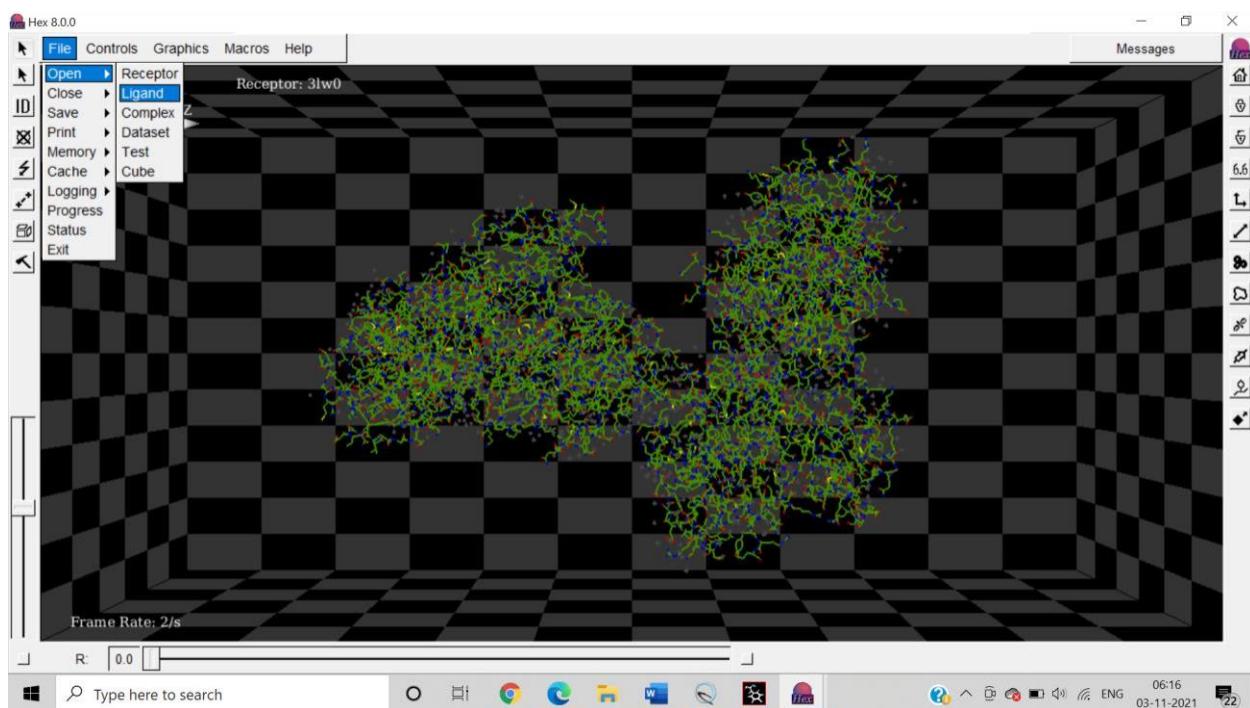
6PYH

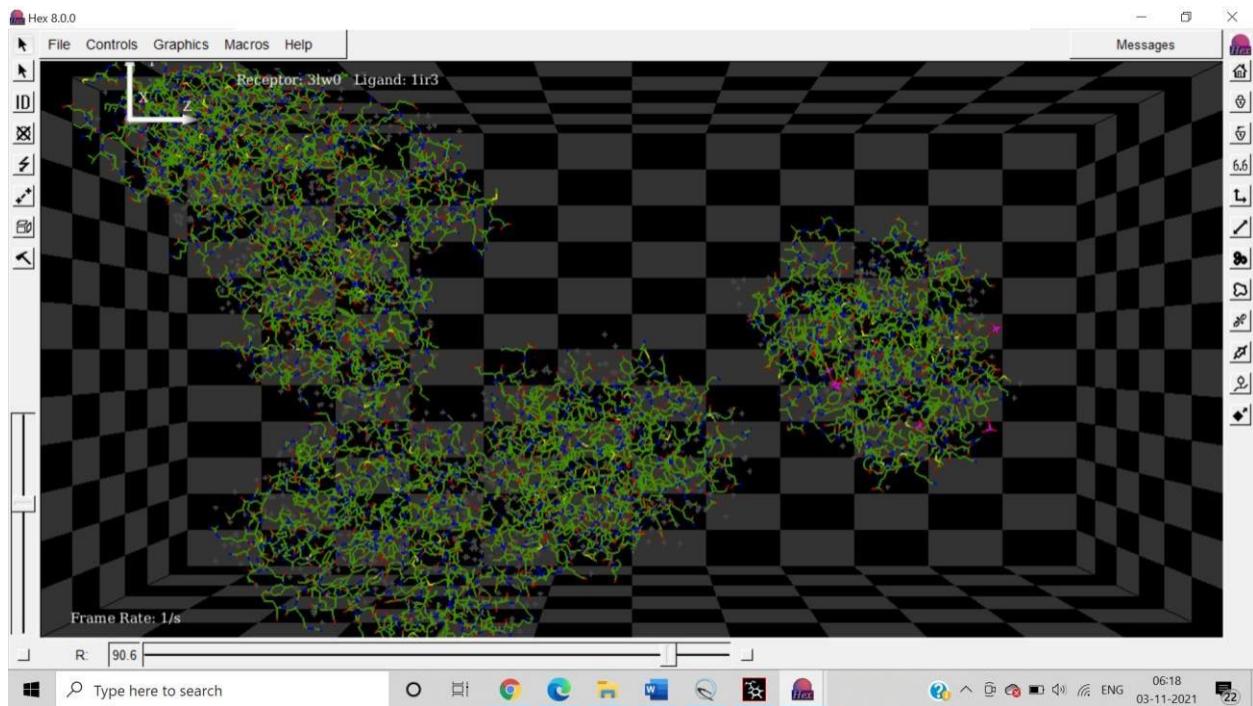
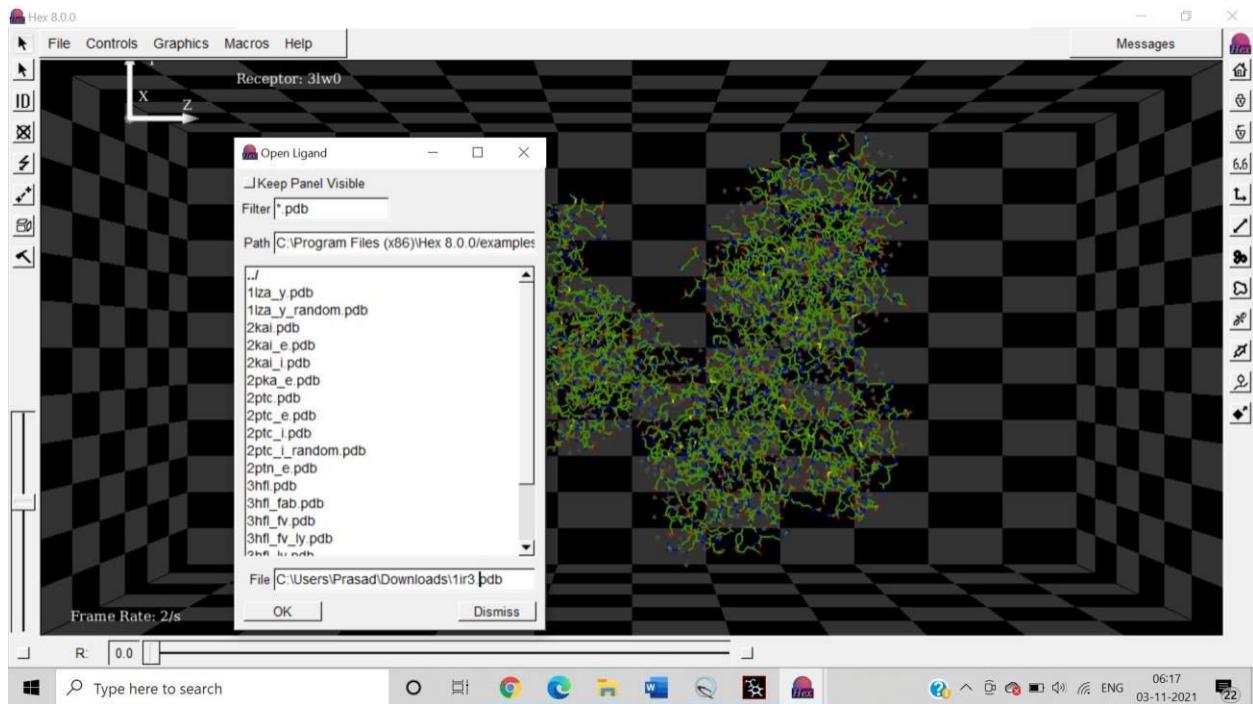
Cryo-EM structure of full-length IGF1R-IGF1 complex. Only the extracellular region of the complex is resolved.

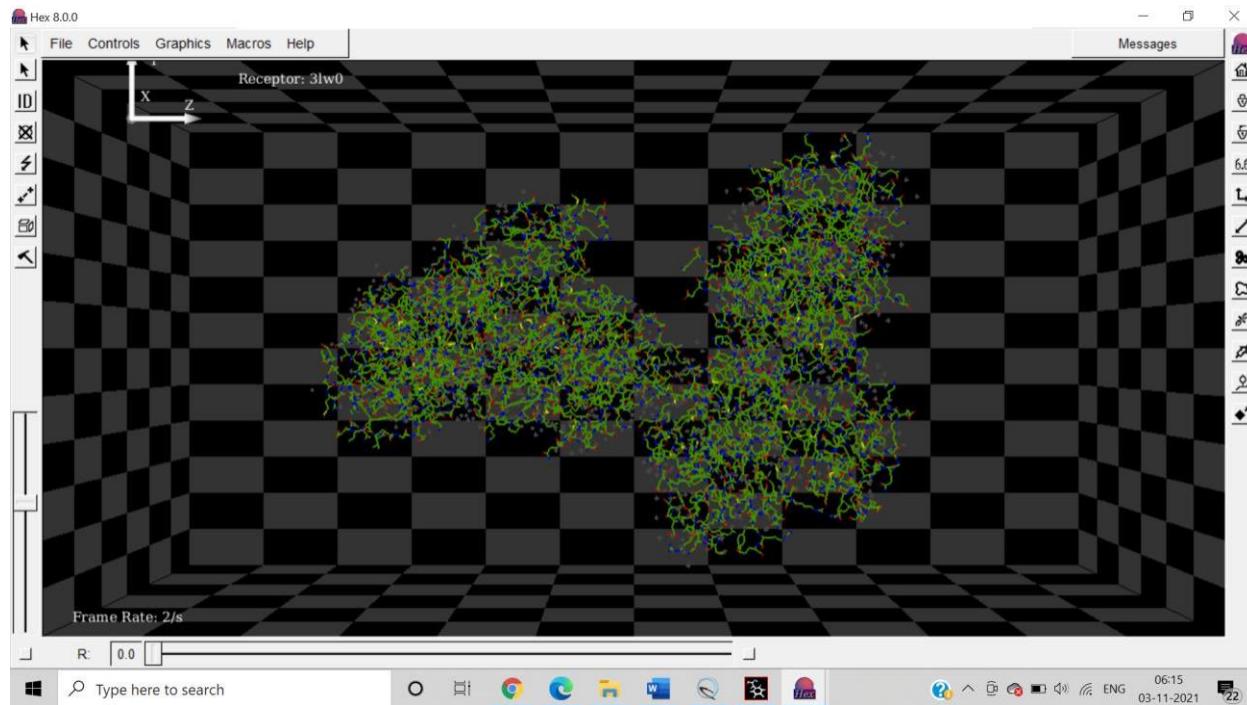
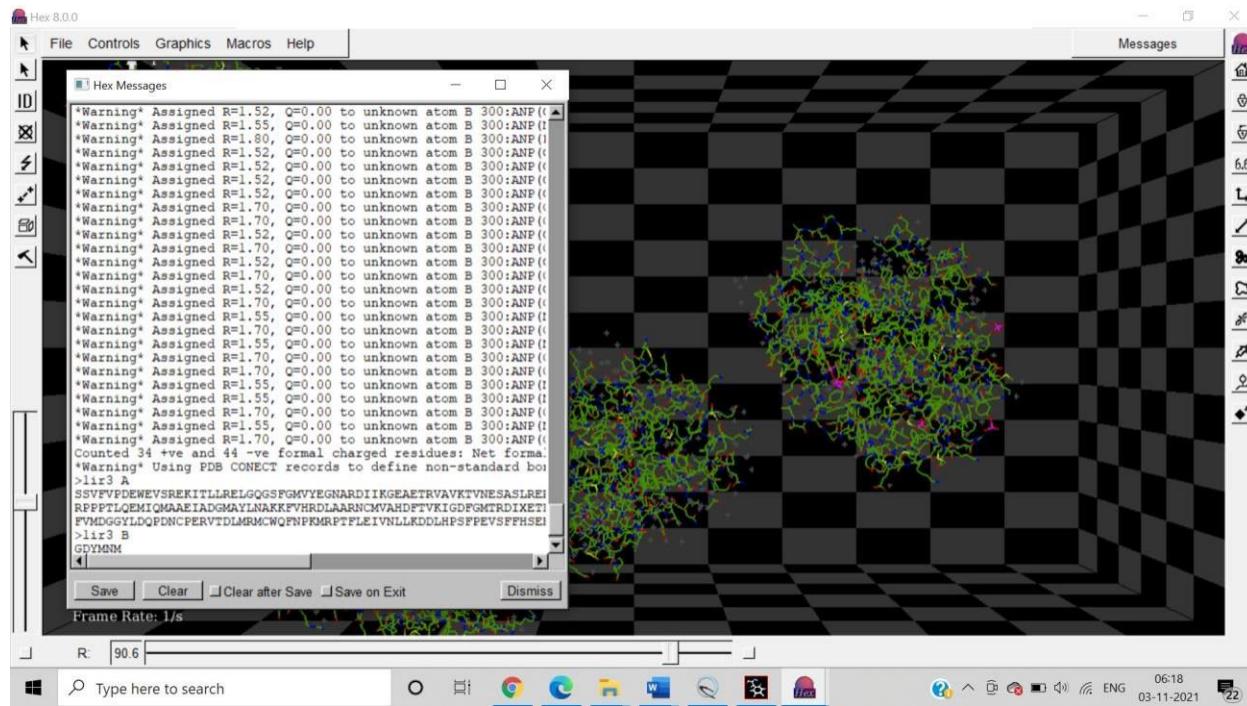


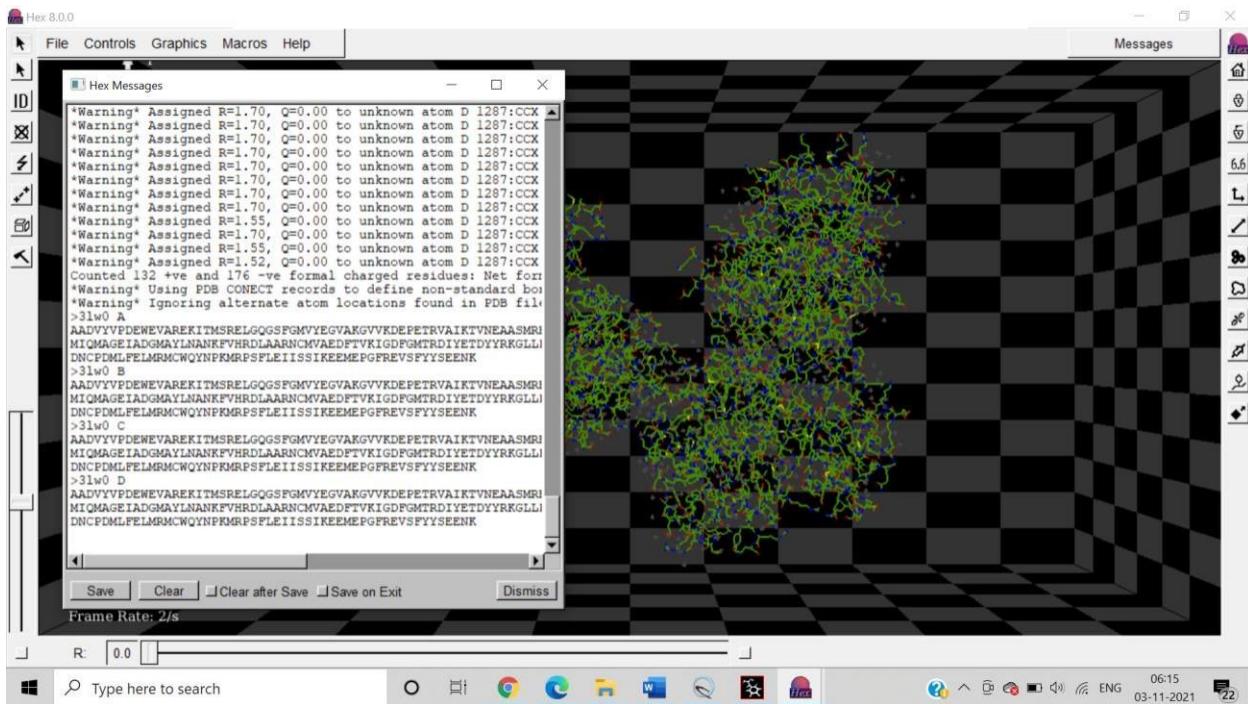


LIGAND:

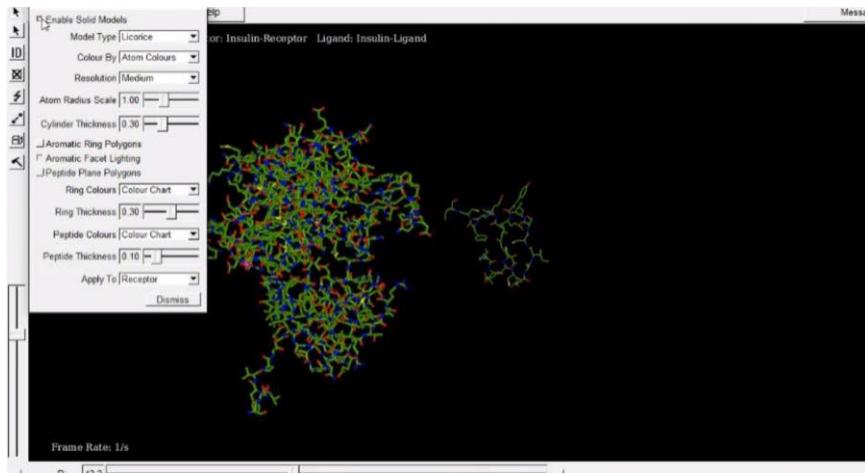




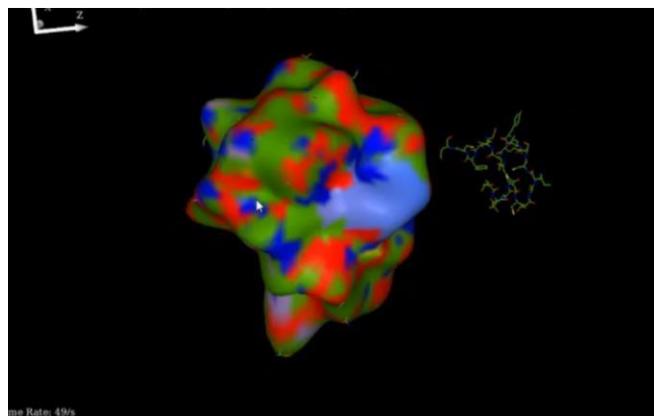




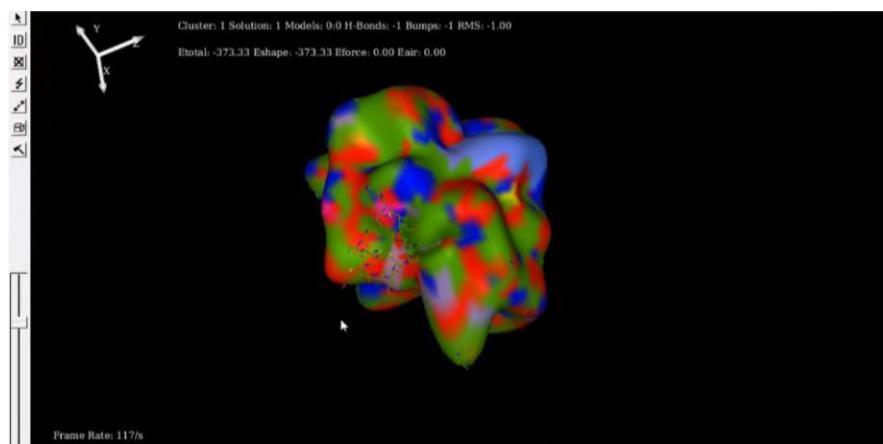
BALL AND STICK MODEL:



HARMONIC SURFACE: Ligand



After the process of protein docking.



Result:

Frame rate = 117/s

1Bumps

-1RMS -1.00

Etotal = -373.33

Eshapes= -373.33

Eforce = 0.00

CONCLUSION:

Protein docking was successfully carried out using hex.

Practical no : 8A

Aim: Perform Structural Based pharmacophores using PharmMapper.

Theory: A pharmacophore is an abstract description of molecular features that are necessary for molecular recognition of a ligand by a biological macromolecule.

Pharmacophore mapping is the definition and placement of pharmacophoric features and the alignment techniques used to overlay 3D structures.

PharmMapper:

PharmMapper Server is a freely accessed web-server designed to identify potential target candidates for the given probe small molecules (drugs, natural products, or other newly discovered compounds with binding targets unidentified) using pharmacophore mapping approach. Benefited from the highly efficient and robust mapping method, PharmMapper bears high throughput ability and can identify the potential target candidates from the database within a few hours. It is backed up by a large, in-house repertoire of pharmacophore database extracted from all the targets in TargetBank, DrugBank, BindingDB and PDTD. Over 7,000 receptor-based pharmacophore models (covering 1,627 drug targets information, 459 of which are human protein targets) are stored and accessed by PharmMapper. It finds the best mapping poses of the user uploaded molecules (in Tripos Mol2 or MDL SDF format) against all the targets in PharmTargetDB and top N potential drug targets as well as respective molecule's aligned poses are outputted.

PharmMapper Steps:

- Convert the file of standard drug using chimera in mol2 format
- Open pharmMapper
- Upload the file containing the standard drug in the upload query file section under the job submission.
- Enter your email and your drug name in description submit
- Set the parameters and yes to generate the conformers [300] and set all target option.
- Submit the job.
- Using the job ID that is generated after the submission if the job checks the status of the job under the Check job tab.
- Result link Obtained via email

PharmMapper Steps:

- Convert the file of standard drug using chimera in mol2 format
- Open pharmMapper

- Upload the file containing the standard drug in the upload query file section under the job submission.
- Enter your email and your drug name in description□ submit
- Set the parameters and yes to generate the conformers [300] and set all target option.
- Submit the job.
- Using the job ID that is generated after the submission if the job checks the status of the job under the Check job tab.
- Result link Obtained via email.

Methodology:

Step 1: Open the PharmMapper online server and upload your query file in mol2 format, give your mail ID and hit ‘continue’.

Step 2: keep the parameters at default and hit submit

Result:

Practical no : 8B :

Aim: - Perform Structural Based pharmacophores using PharmMapper.

Theory:

PharmaGist:

PharmaGist is a freely available web server for pharmacophore detection. The employed method is ligand based. It does not require the structure of the target receptor; instead, the input is a set of structures of drug-like molecules that are known to bind to the receptor. Candidate pharmacophores are computed by multiple flexible alignments of the input ligands. The main innovation of this approach is that the flexibility of the input ligand is handled explicitly and in deterministic manner within the alignment process. The method is highly efficient, where a typical run with up to 33 drug-like molecules take seconds to a few minutes on a standard PC. Another important characteristic of the method is the capability of detecting pharmacophores shared by different subsets of input molecules. This capability is a key advantage when the ligands belong to different binding modes or when the input contains outliers.

Methodology:-

Step 1: Create a zip file of 5 mol2 format molecules; enter those in the space given along with your mail ID and hit ‘submit’.

PharmaGist

Websrvr
[About] [WebServer] [FAQ] [Help / Getting Started] Contact: pdock@tau.ac.il

```
Validating input molecules:  
Filename 0: 1.mol2  
1.mol2  
run vega for 1.mol2  
/specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.linux -c /specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.config -t 1.mol2  
for 1.mol2 exit mode is 0 Filename 1: 1.mol2 OK  
Filename 1: 4.mol2  
4.mol2  
run vega for 4.mol2  
/specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.linux -c /specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.config -t 4.mol2  
for 4.mol2 exit mode is 0 Filename 2: 4.mol2 OK  
Filename 2: 2.mol2  
2.mol2  
run vega for 2.mol2  
/specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.linux -c /specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.config -t 2.mol2  
for 2.mol2 exit mode is 0 Filename 3: 2.mol2 OK  
Filename 3: 3.mol2  
3.mol2  
run vega for 3.mol2  
/specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.linux -c /specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.config -t 3.mol2  
for 3.mol2 exit mode is 0 Filename 4: 3.mol2 OK  
Filename 4: Diclofenac.mol2  
Diclofenac.mol2  
run vega for Diclofenac.mol2  
/specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.linux -c /specific/a/home/cci/psdock/websrvr/pharma/bin/pharmagist.config -t Diclofenac.mol2  
for Diclofenac.mol2 exit mode is 0 Filename 5: Diclofenac.mol2 OK
```

Running PharmaGist for 5 molecules.

Conclusion:

- Pharmacophore mapping was performed using both PharmMapper and PharmaGist servers using a complex receptor along with 6derivatives.
- PharmMapper is based on genetic

- algorithm and it uses structure based virtual screening (SBVS) and PharmaGist uses ligand based virtual screening (LBVS).
- In PharmMapper, it was concluded that the rank 299 ligand had the best fit based on the fitness score of 2.319 along with 3 hydrophobic interactions and zero positive, 1 negative, 0 donors, 2 acceptor molecules.

PharmGist Steps:

- Zip all of the derivatives of the drugs along with the standard drug molecule into a folder. Selected all the derivatives → right click → compress to zip folder. (Make sure that you do not zip the entire folder since the system won't be able to detect the derivatives).
- Open the Pharm Gist web tool and submit the all derivates by uploading the zip file into the input page.
- Set the number of output Pharmacophores as, enter the email address and submit the input.

Result: The result link can be obtained via the email which can be further used to analyses the result.

Experiment No 9:

Aim:- Toxicity analysis using SwissADME.

Theory:-This website allows you to compute physicochemical descriptors as well as to predict ADME parameters, pharmacokinetic properties, druglike nature and medicinal chemistry friendliness of one or multiple small molecules to support drug discovery.The main idea is to give free access to a number of parameters and predictive models in order to compute the physicochemistry and estimate the pharmacokinetics, druglikeness and medicinal chemistry friendliness of small molecules. All selected methods are in-house, implemented from publications or directly computed through non-commercial executables. Apart from performance and robustness, all models should be free-for-academics, easy to analyse and straightforwardly translatable to chemistry for drug design purposes. Besides, some effort were put on simple and flexible input/output as well as clear layout in the web page. This allows experts as well as nonexperts to make use of the website for their research. SwissADME was made for application in drug discovery and medicinal chemistry contexts, which stresses for a balance between accuracy and speed in order to deal with a large number of molecules. Because of the predictive nature of the data returned by SwissADME, values should be handled with due care. Extra precaution should be taken if using SwissADME for any other activities outside the scope of drug design/discovery.

Methodology:

Google→SwissADME→Paste the canonical smile [Can Enter the list of SMILES here]

Run

Result below→Screenshot→Conclusion about the properties

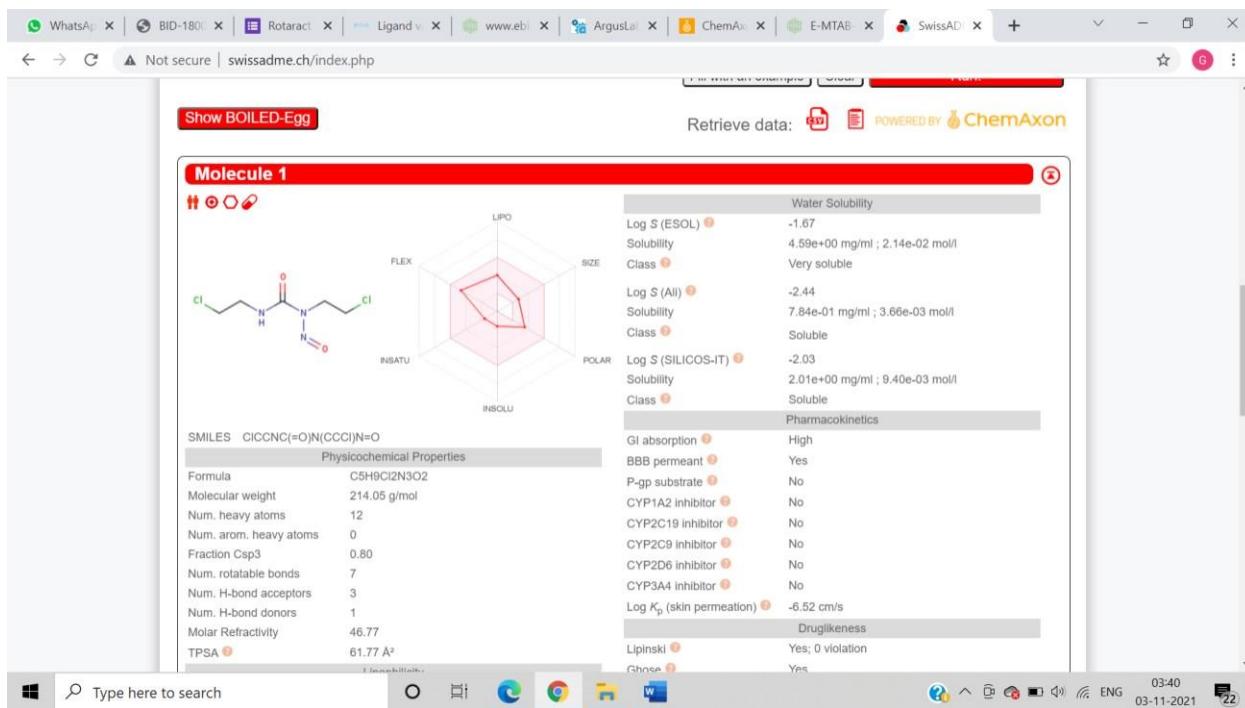
Can retrieve the data in PDF & notepad form.

First go to drug bank and retrieve your drug name:

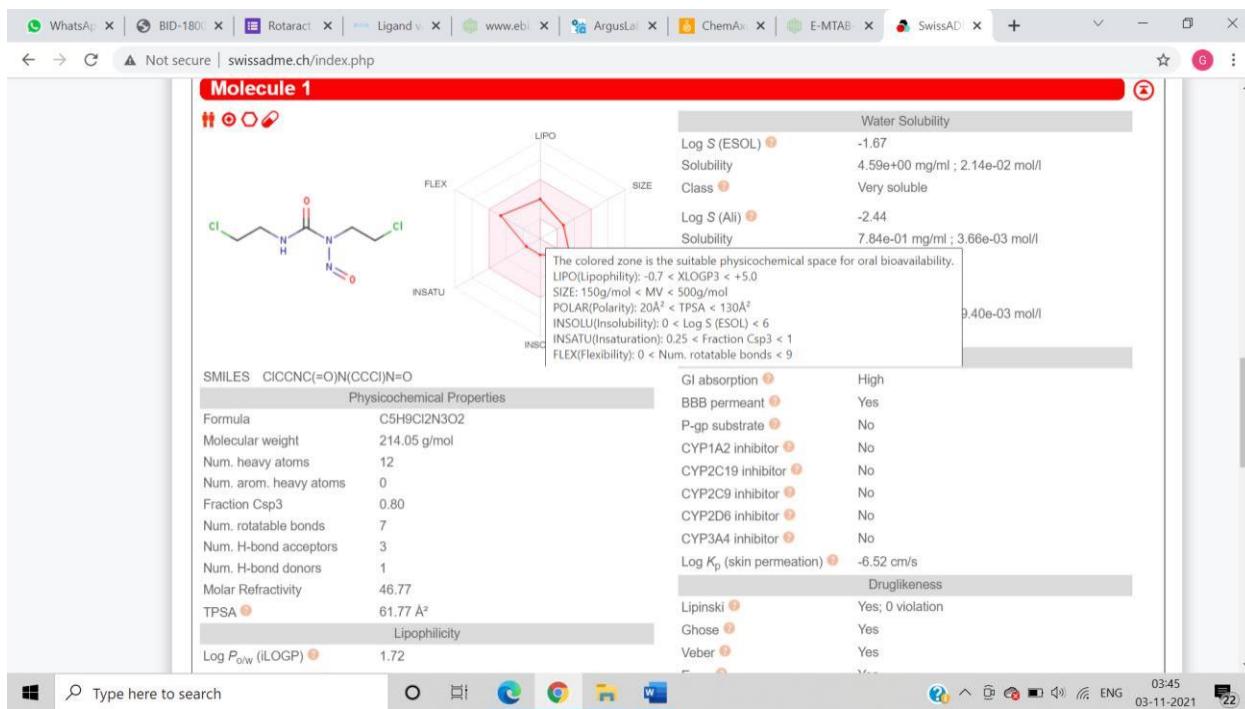
Disease: Glioma

Drug: Carmustine

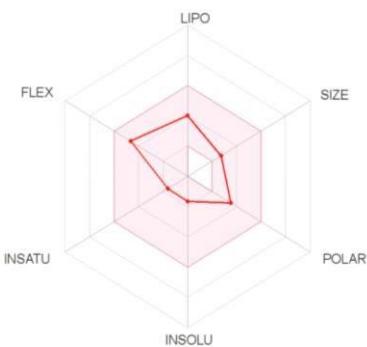
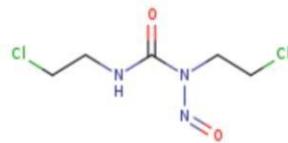
Enter the smiles in the box and run the program:



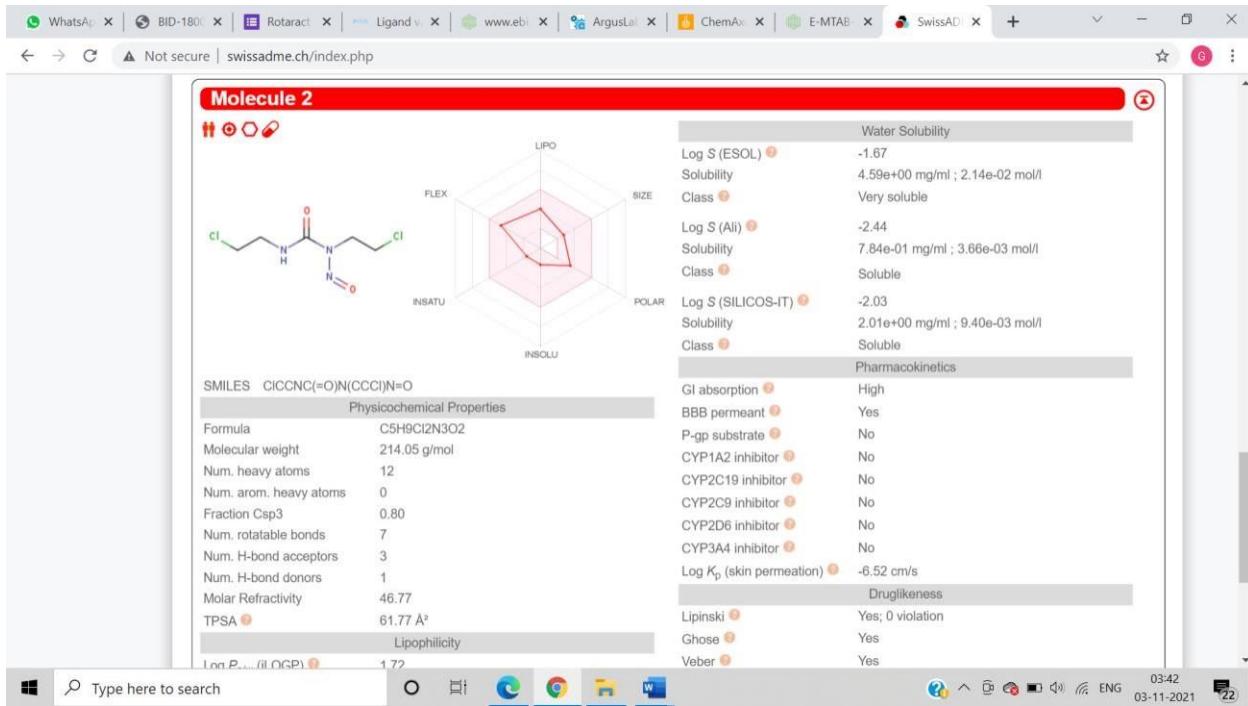
The coloured zone is the suitable physicochemical space for oral bioavailability .



Molecule 1



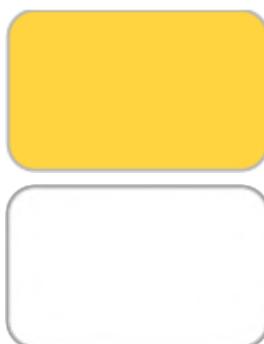
Formula	C5H9Cl2N3O2
Molecular weight	214.05 g/mol
Num. heavy atoms	12
Num. arom. heavy atoms	0
Fraction Csp3	0.80
Num. rotatable bonds	7
Num. H-bond acceptors	3
Num. H-bond donors	1
Molar Refractivity	46.77
TPSA	61.77 Å ²



Log $P_{o/w}$ (iLOGP) 1.72

The BOILED-Egg can be applied in a variety of settings, from the filtering of chemical libraries at the early steps of drug discovery, to the evaluation of drug candidates for development.

The BOILED-Egg model delivers a rapid, intuitive, easily reproducible yet statistically unprecedented robust method to predict the passive gastrointestinal absorption and brain access of small molecules useful for drug discovery and development.



BBB

HIA



PGP+

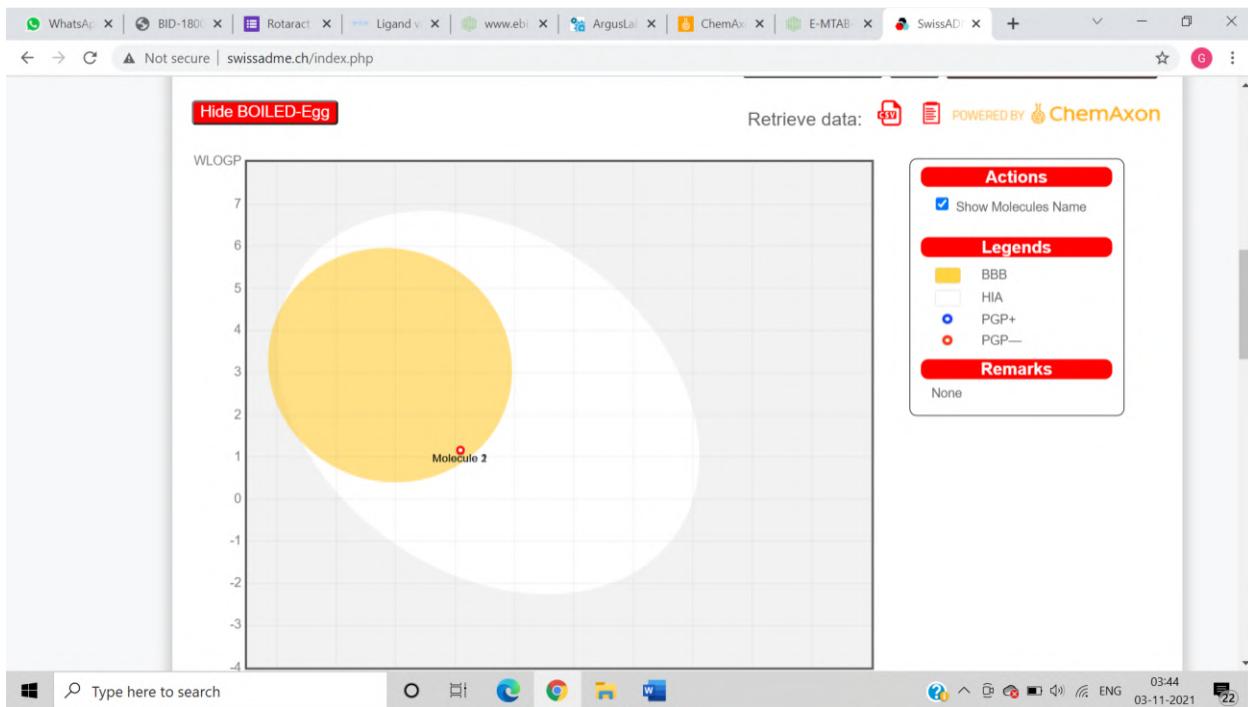
<=""

td="" style="width:

11px; height:
11px;">>

PGP—
<=""
td="" style="width:
11px; height:
11px;">>

Show molecule name:



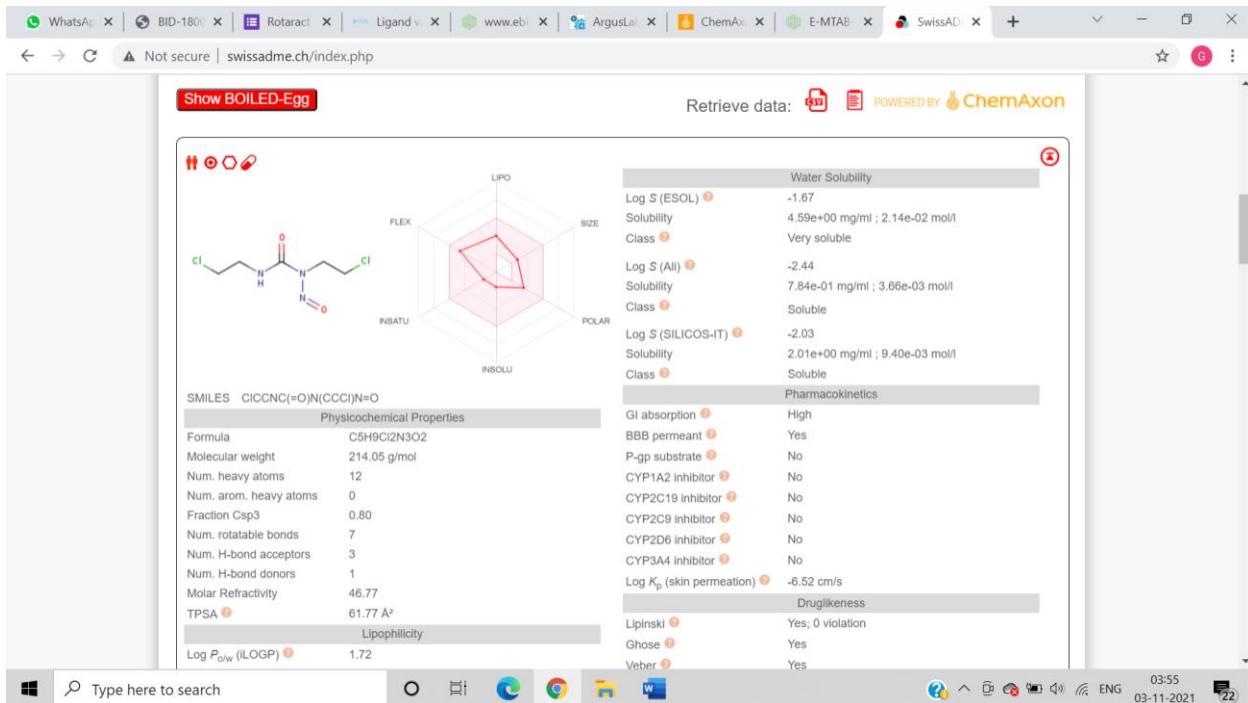
Go to Pub Chem and retrieve 10 smiles of Carmustine:

The screenshot shows the PubChem search interface. A search query "carmustine" has been entered into the search bar. The results page displays the "COMPOUND BEST MATCH" for Carmustine, with its CID (2578) and various chemical properties listed: MF: C₅H₉Cl₂N₃O₂, MW: 214.05g/mol, IUPAC Name: 1,3-bis(2-chloroethyl)-1-nitrosourea, Isomeric SMILES: C(CCl)NC(=O)N(CCCl)N=O, InChIKey: DLGOEMSEDOSKAD-UHFFFAOYSA-N, InChi: InChI=1S/C5H9Cl2N3O2/c6-1-3-8-5(11)10(9-12)4-2-7/h1-4H2,(H,8,11), Create Date: 2005-03-25. Below the main result, there are tabs for "Summary", "Similar Structures Search", "Related Records", and "PubMed (MeSH Keyword)". The browser's top navigation bar and taskbar are visible at the bottom.

Copy paste the 10 smiles and run the program:-

The screenshot shows the Marvin JS interface. On the right, a text input field contains a list of SMILES strings: C(CC1NC(=O)N(CCC1)N=O, C(CNC(=O)N(CCC1)N=O)NC(=O)N(CCC1)N=O, C(CC1NC(=O)N(CCN(C(=O)NCCCC1)N=O)N=O, C(CC1NC(=O)N(CCF)N=O, C(CC1NC(=O)N(CCC1)[NH2+] [O-], C(CF)NC(=O)N(CCC1)N=O, [2H]C([2H]) (CNC(=O)N(CC([2H])([2H])C1)N=O)C1, CC(CNC(=O)N(CC(C)C1)N=O)C1, CNC(=O)N(CCC1)N=O.C1. Below the input field are buttons for "Fill with an example", "Clear", and "Run!". The Marvin JS interface includes a toolbar on the left and a molecular viewer window on the right. The browser's top navigation bar and taskbar are visible at the bottom.

Result:



Formula C5H9Cl2N3O2

Molecular weight 214.05 g/mol

Num. heavy atoms 12

Num. arom. heavy atoms 0

Fraction Csp3 0.80

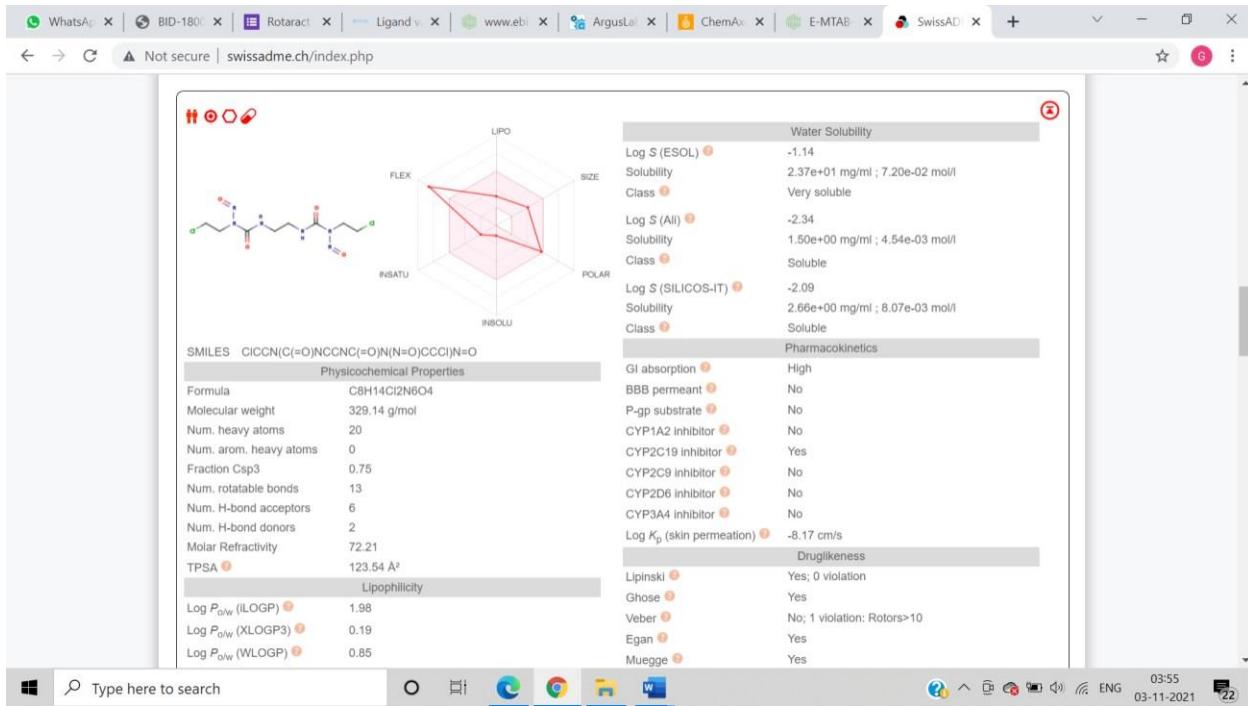
Num. rotatable bonds 7

Num. H-bond acceptors 3

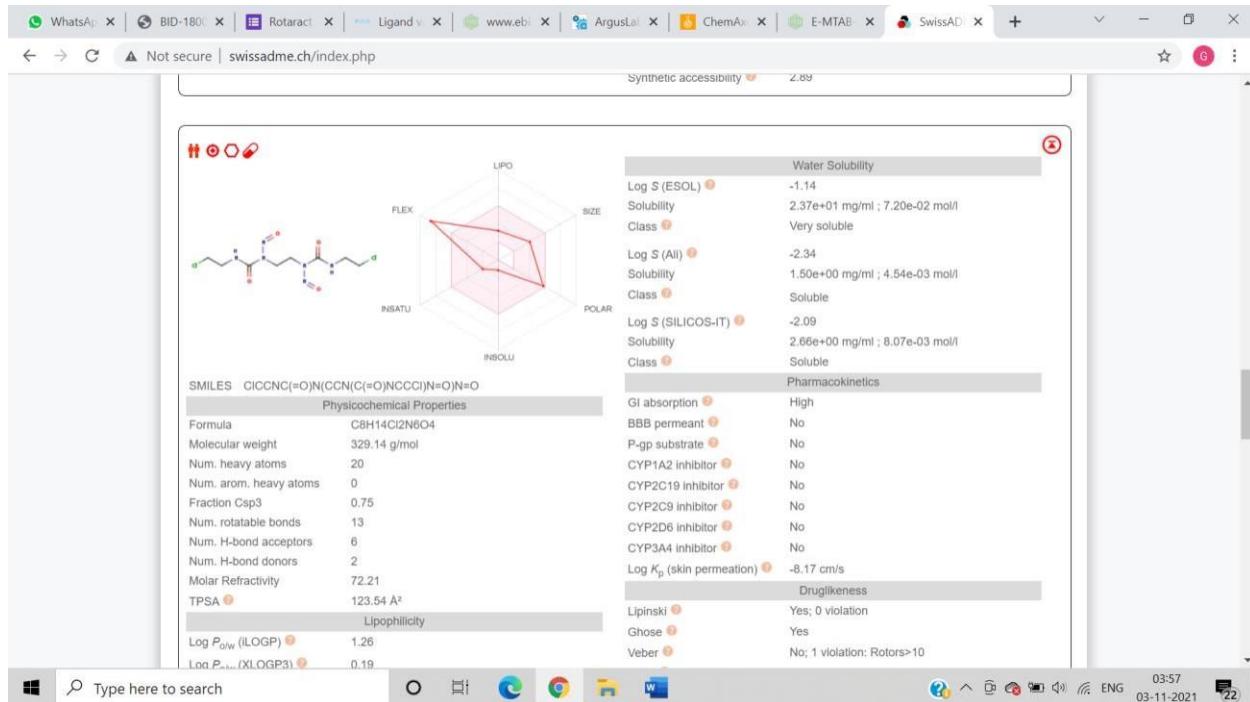
Num. H-bond donors 1

Molar Refractivity 46.77

TPSA 61.77 Å²



Formula	C8H14Cl2N6O4
Molecular weight	329.14 g/mol
Num. heavy atoms	20
Num. arom. heavy atoms	0
Fraction Csp3	0.75
Num. rotatable bonds	13
Num. H-bond acceptors	6
Num. H-bond donors	2
Molar Refractivity	72.21
TPSA	123.54 Å ²



Formula C8H14Cl2N6O4

Molecular weight 329.14 g/mol

Num. heavy atoms 20

Num. arom. heavy atoms 0

Fraction Csp3 0.75

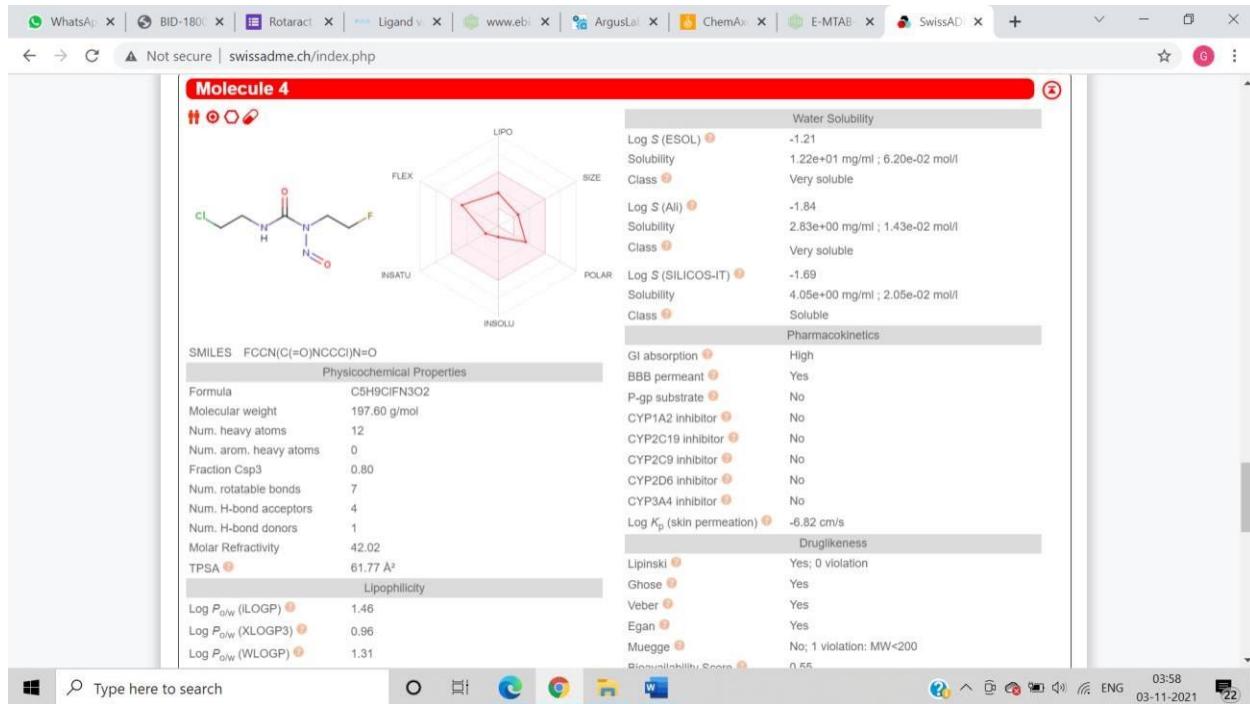
Num. rotatable bonds 13

Num. H-bond acceptors 6

Num. H-bond donors 2

Molar Refractivity 72.21

TPSA 123.54 Å²



Formula C5H9ClFN3O2

Molecular weight 197.60 g/mol

Num. heavy atoms 12

Num. arom. heavy atoms 0

Fraction Csp3 0.80

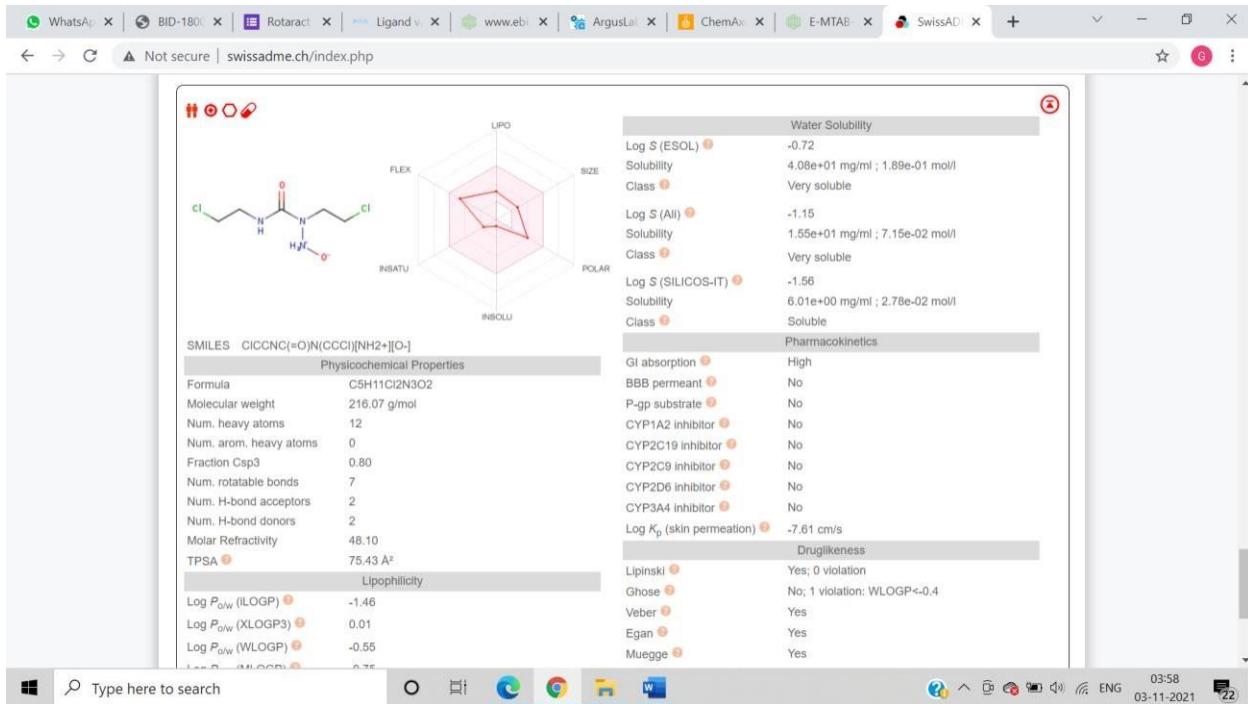
Num. rotatable bonds 7

Num. H-bond acceptors 4

Num. H-bond donors 1

Molar Refractivity 42.02

TPSA 61.77 Å²



Formula C5H11Cl2N3O2

Molecular weight 216.07 g/mol

Num. heavy atoms 12

Num. arom. heavy atoms 0

Fraction Csp3 0.80

Num. rotatable bonds 7

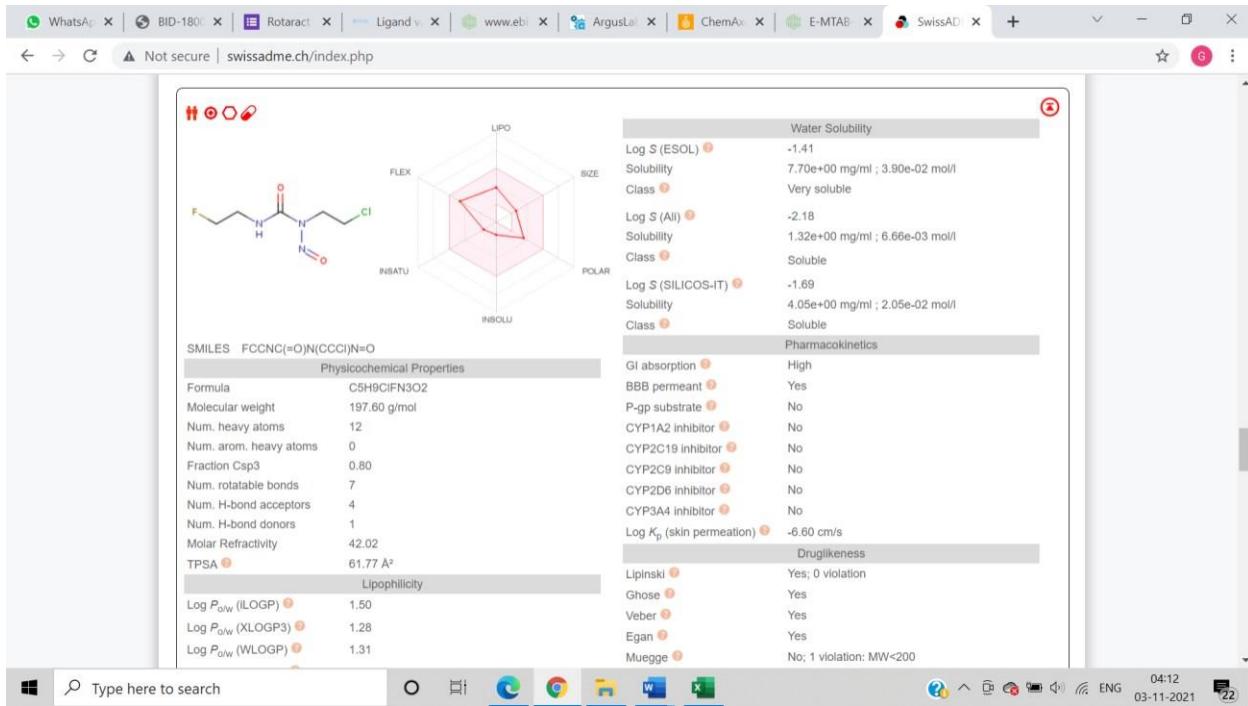
Num. H-bond acceptors 2

Num. H-bond donors 2

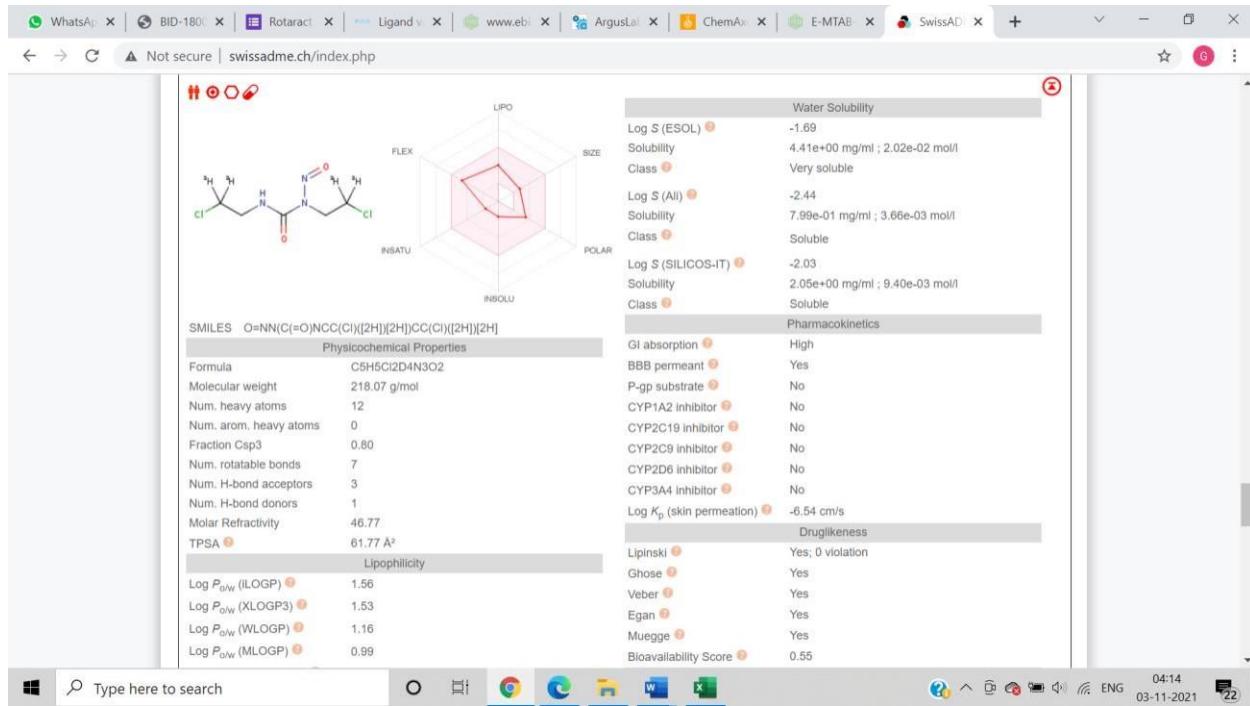
Molar Refractivity 48.10

75.43 Å²

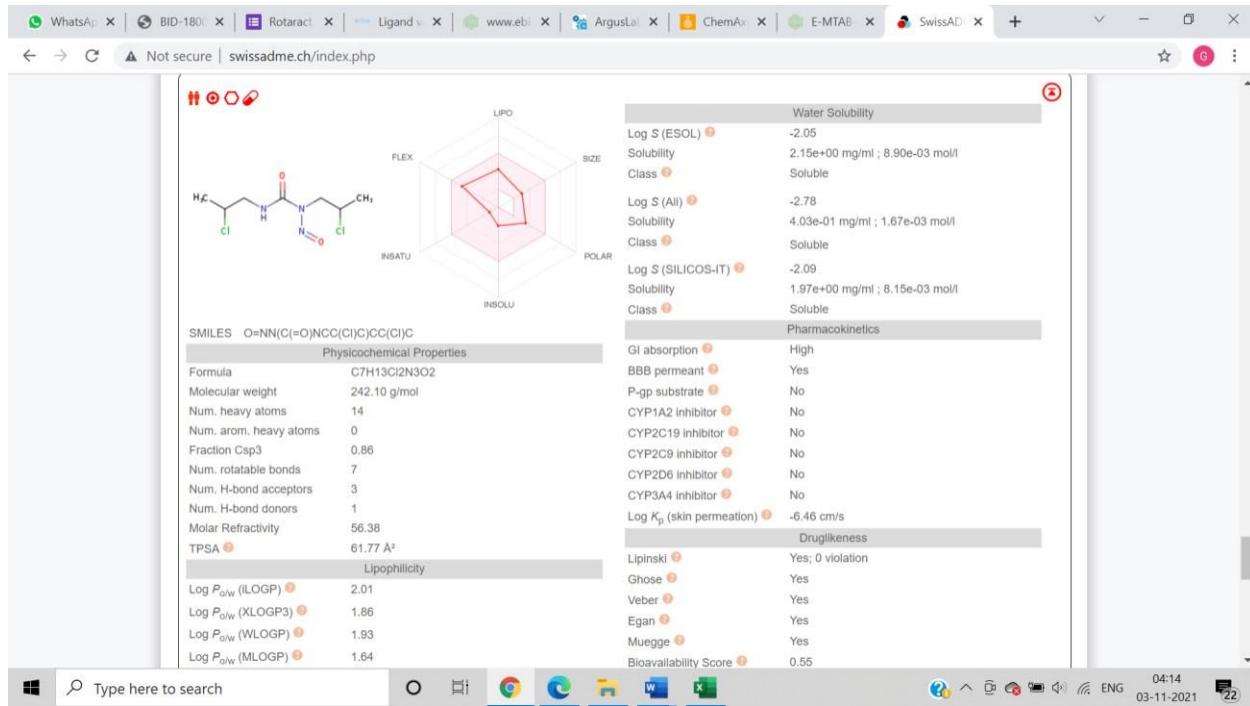
TPSA ?



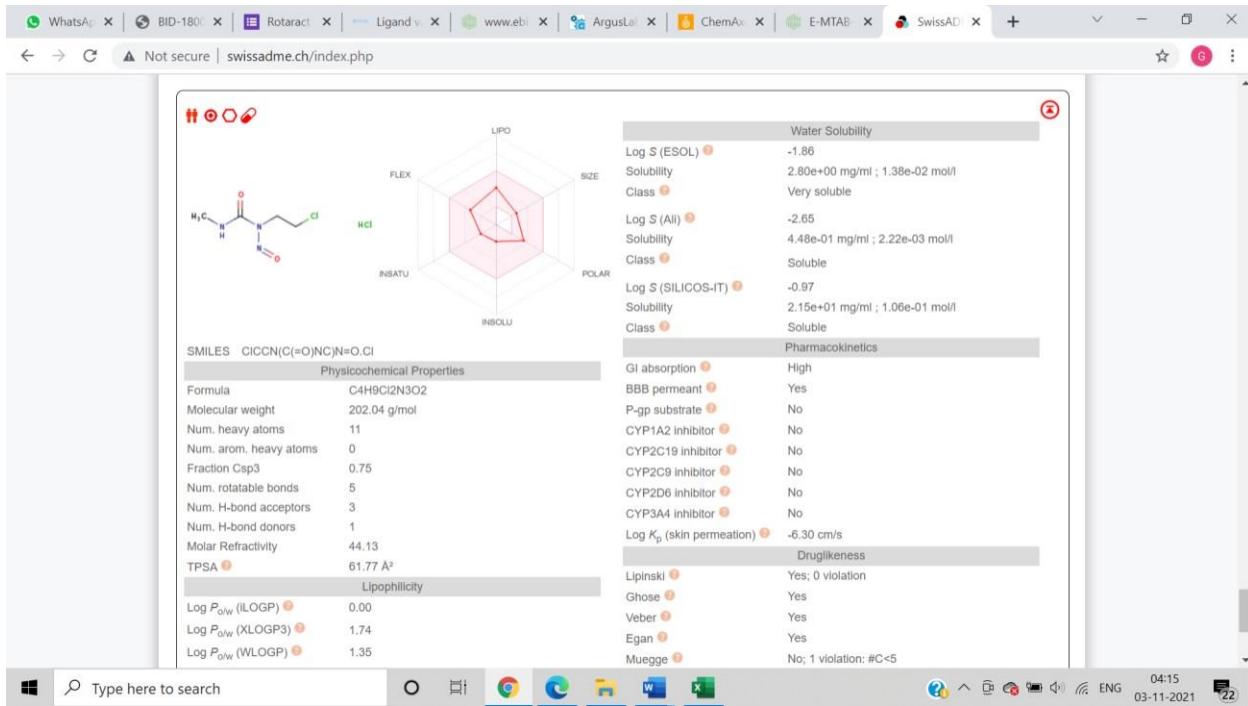
Formula	C5H9ClFN3O2
Molecular weight	197.60 g/mol
Num. heavy atoms	12
Num. arom. heavy atoms	0
Fraction Csp3	0.80
Num. rotatable bonds	7
Num. H-bond acceptors	4
Num. H-bond donors	1
Molar Refractivity	42.02
TPSA	61.77 Å ²



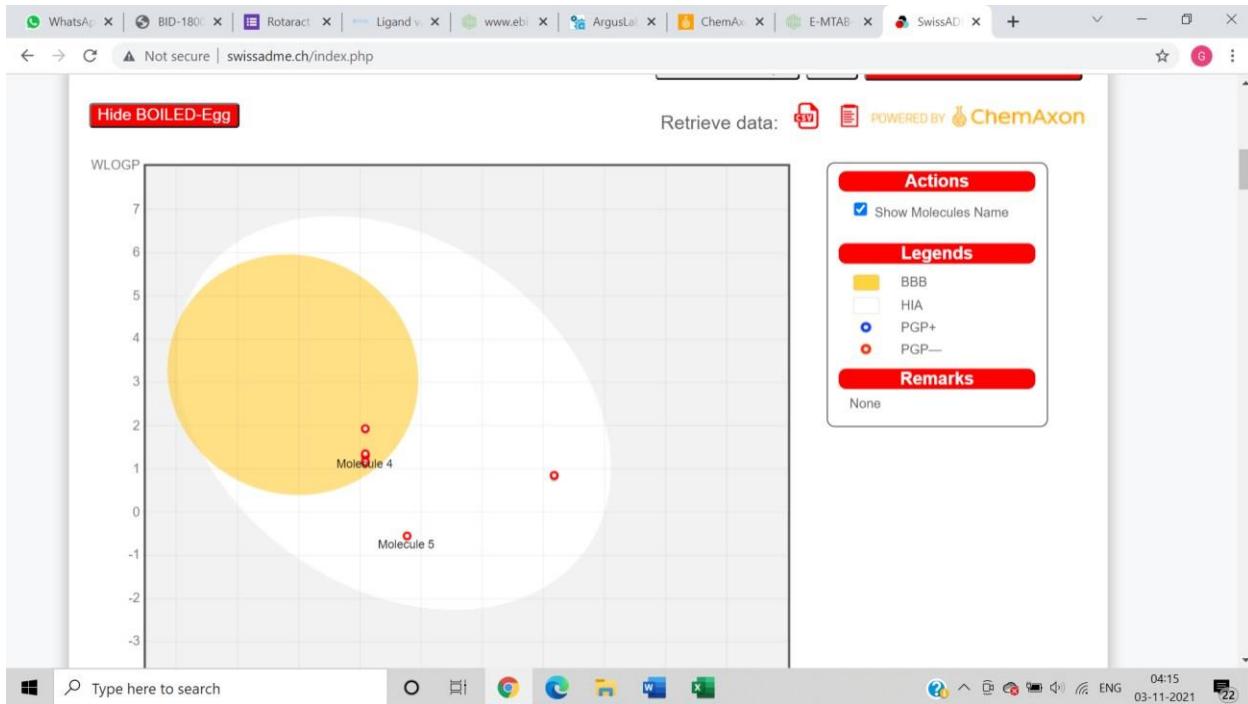
Formula	C ₅ H ₅ Cl ₂ D ₄ N ₃ O ₂
Molecular weight	218.07 g/mol
Num. heavy atoms	12
Num. arom. heavy atoms	0
Fraction Csp ₃	0.80
Num. rotatable bonds	7
Num. H-bond acceptors	3
Num. H-bond donors	1
Molar Refractivity	46.77
TPSA	61.77 Å ²



Formula	C ₇ H ₁₃ Cl ₂ N ₃ O ₂
Molecular weight	242.10 g/mol
Num. heavy atoms	14
Num. arom. heavy atoms	0
Fraction Csp ₃	0.86
Num. rotatable bonds	7
Num. H-bond acceptors	3
Num. H-bond donors	1
Molar Refractivity	56.38
TPSA	61.77 Å ²



Formula	C4H9Cl2N3O2
Molecular weight	202.04 g/mol
Num. heavy atoms	11
Num. arom. heavy atoms	0
Fraction Csp3	0.75
Num. rotatable bonds	5
Num. H-bond acceptors	3
Num. H-bond donors	1
Molar Refractivity	44.13
TPSA	61.77 Å ²



swissadme (1).csv - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

A1 Molecule

Molecule	Canonical Formula	MW	#Heavy at	#Aromatic Fraction	C	#Rotatable	#H-bond a	#H-bond c	MR	TPSA	iLOGP	XLOGP3	WLOGP	MLOGP	Silicos-IT	L Consensus	ESOL	Log SES
C1CCNC(=C)C9C12N		214.05	12	0	0.8	7	3	1	46.77	61.77	1.72	1.53	1.16	0.99	0.66	1.21	-1.67	4.
C1CCN(C(=C)C8H14Cl2)C12N		329.14	20	0	0.75	13	6	2	72.21	123.54	1.98	0.19	0.85	0.65	-0.45	0.65	-1.14	2.
C1CCNC(=C)C8H14Cl2N		329.14	20	0	0.75	13	6	2	72.21	123.54	1.26	0.19	0.85	0.65	-0.45	0.5	-1.14	2.
Molecule 4: C1CCNC(=C)C9C12N		197.6	12	0	0.8	7	4	1	42.02	61.77	1.46	0.96	1.31	0.81	0.44	1	-1.21	1.
Molecule 5: C1CCNC(=C)C9C12N		216.07	12	0	0.8	7	2	2	48.1	75.43	-1.46	0.01	-0.55	-0.75	-1.21	-0.79	-0.72	4.
FCCNC(=C)C9C12N		197.6	12	0	0.8	7	4	1	42.02	61.77	1.5	1.28	1.31	0.81	0.44	1.07	-1.41	7.
O=NN(C(=C)C9C12N)C12N		218.07	12	0	0.8	7	3	1	46.77	61.77	1.56	1.53	1.16	0.99	0.66	1.18	-1.69	4.
O=NN(C(=C)C7H13Cl2)C12N		242.1	14	0	0.86	7	3	1	56.38	61.77	2.01	1.86	1.93	1.64	0.99	1.69	-2.05	2.
C1CCNC(=C)C4H9Cl2N		202.04	11	0	0.75	5	3	1	44.13	61.77	0	1.74	1.35	0.63	-0.26	0.69	-1.86	2.

Result:

- 1 standard drug molecule and it's 10 derivative Smile was pasted in the section and results for 11 molecules were obtained , based on the results we observe several different sections that determine various properties.
- For molecule 1, from the results obtained the molecule has chemical formula of C₂₉H₂₆ClFN₄O₄S, it's molecular weight is 166.24 gms; It has 11 heavy atoms, 6 Aromatic heavy atoms, csp3 fraction is 0.25. There are 2 rotatable bonds, 1 H bonds acceptors and 1 H bond donors respectively. The molar reflectivity is 49.97 and TPSA of given molecule is 71 Å².
- WLOGP value for the molecule is 1.28, Further we observe that the molecule is very soluble in water with solubility value 2.54e+00 mg/ml ; 1.53e-02 mol/l; 3.62e-07 mol/l.
- The GI absorption rate is high for this molecule , This drug molecule cannot pass through BBB ,pgp substrate no; This drug molecule acts as an inhibitor for the enzymes given below CYP1A2 inhibitor ,CYP2C19 inhibitor, CYP2C9 inhibitor,CYP2D6 inhibitor,CYP3A4 inhibitor and does not inhibit activity of enzyme CYP1A2.
- This drug molecule violates Lipinski's rule -> No; 1 violation: MW<200

Conclusion: Molecular properties of molecules were determined using SwissADME . The selected molecules was Carmustine which showed solubility and high GI absorption. Log K_p was found. Carmustine can cross Glioma.

Experiment No 10A

Aim: Gene Expression data analysis using database and GEO2r tool.

Theory: Gene expression omnibus (GEO) is a publicly available database for storing functional genomics data. The tools at GEO help the user to query and download the GEO dataset and experiments. It is an international public repository which distributes microarray, next generation sequencing and other forms of high throughput genomic data.

It is serendipitous that the word “geo” is a prefix meaning “earth” because not only does GEO primarily host global gene expression data, GEO itself is indeed a global resource; at the time of this writing GEO contains submissions from 72 nations. There are no fees to submit data to GEO, download data, or use GEO tools. Scientists submit to GEO in order to share their data with the research community and/or as a requirement of publication or grant directives. GEO supports the Minimum Information About a Microarray Experiment (MIAME) and Minimum :

by the Functional Genomics Data Society for standardization of information about microarray and sequencing experiments that enable the data to be interpreted and replicated by the research community. The GEO database handles the majority of direct submissions from the research community and at the time of this writing holds 54,640 public studies, comprising over 1.3 million samples, derived from 2889 different organisms. An up-to-date summary of GEO data types and content is provided at <http://www.ncbi.nlm.nih.gov/geo/summary/>.

While the chief role of GEO is to serve as a public data archive, the database is not simply an online warehouse of data. GEO strives to make the data it contains accessible to the research community. Due to the complex nature of the data generated by genomic experiments most studies are analyzed by bioinformaticians and statisticians, or researchers with specialized analysis software. Researchers who lack these skills or software face a substantial challenge if they wish to analyze genomics experiments themselves. In order to make such data analysis accessible to all researchers, GEO has developed several tools for data query, visualization, and analysis that can be performed directly on the GEO website and do not require the download or manipulation of the data files.

Methodology:

Homepage of GEO:

COVID-19 Information

Gene Expression Omnibus

GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

Getting Started

- Overview
- FAQ
- About GEO DataSets
- About GEO Profiles
- About GEO2R Analysis
- How to Construct a Query
- How to Download Data

Tools

- Search for Studies at GEO DataSets
- Search for Gene Expression at GEO Profiles
- Search GEO Documentation
- Analyze a Study with GEO2R
- Studies with Genome Data Viewer Tracks
- Programmatic Access
- FTP Site
- ENCODE Data Listings and Tracks

Browse Content

Repository Browser
DataSets: 4348
Series: 162799
Platforms: 22704
Samples: 4700084

Keyword or GEO Accession Search

Disease: Glioma

Drug: Carmustine

So we have 676248 hits for Glioma:

Gene Expression Omnibus

GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

glioma

There are 68500 results for "glioma"
There are 676248 results for "glioma"

Setting Started

- Overview
- FAQ
- About GEO DataSets
- About GEO Profiles
- About GEO2R Analysis
- How to Construct a Query
- How to Download Data

Tools

- Search for Studies at GEO DataSets
- Search for Gene Expression at GEO Profiles
- Search GEO Documentation
- Analyze a Study with GEO2R
- Studies with Genome Data Viewer Tracks
- Programmatic Access
- FTP Site
- ENCODE Data Listings and Tracks

Information for Submitters

- Login to Submit
- Submission Guidelines
- Update Guidelines
- MIAME Standards
- Citing and Linking to GEO

For glioma disease we got 68500 hits

Single-cell transcriptomics revealed a critical role of SPP1/CD44-mediated crosstalk between macrophages and cancer cells in glioma.

Summary :- single-cell RNA sequencing analysis of one high-grade glioma with wide-type IDH genes, revealing the landscape of tumor and immune cells in the ecosystem of this tumor

The screenshot shows the NCBI GEO Accession Display page for study GSE185231. The study is titled "Single-cell transcriptomics revealed a critical role of SPP1/CD44-mediated crosstalk between macrophages and cancer cells in glioma". It was performed by Jia D, Liu Y, He C, Ma X, using expression profiling by high throughput sequencing. The study is public on Oct 31, 2021. The platforms used were GPL24676 Illumina NovaSeq 6000 (Homo sapiens).

Series GSE185231	
Status	Public on Oct 31, 2021
Title	Single-cell transcriptomics revealed a critical role of SPP1/CD44-mediated crosstalk between macrophages and cancer cells in glioma
Organism	Homo sapiens
Experiment type	Expression profiling by high throughput sequencing
Summary	a single-cell RNA sequencing analysis of one high-grade glioma with wide-type IDH genes, revealing the landscape of tumor and immune cells in the ecosystem of this tumor
Overall design	one sample from glioma tissue
Contributor(s)	Jia D, Liu Y, He C, Ma X
Citation missing	<i>Has this study been published? Please login to update or notify GEO.</i>
Submission date	Oct 04, 2021
Last update date	Oct 31, 2021
Contact name	Deshui Jia
E-mail(s)	deshuijia@gmail.com
Organization name	Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine
Street address	No.650,Xin SongJiang Road
City	Shanghai
ZIP/Postal code	201600
Country	China

Platforms (1) GPL24676 Illumina NovaSeq 6000 (Homo sapiens)

CANCER AND DRUG TREATMENT:

Title: Decoding Expression Dynamics of Protein and Transcriptome at the Single Cell Level through Multi-Omics Sequencing in Paired Picoliter Chambers

Get the latest research from NIH: <https://www.nih.gov/coronavirus>.
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

HOME | SEARCH | SITE MAP | GEO Publications | FAQ | MIAME | Email GEO | Not logged in | Login

Scope: Self Format: HTML Amount: Quick GEO accession: GSE186402 Go

Series GSE186402 Query DataSets for GSE186402

Status: Public on Oct 25, 2021
 Title: Decoding Expression Dynamics of Protein and Transcriptome at the Single Cell Level through Multi-Omics Sequencing in Paired Picoliter Chambers
 Organisms: Homo sapiens; Mus musculus
 Experiment type: Expression profiling by high throughput sequencing
 Summary: Simultaneous analysis of mRNAs and proteins at the single-cell level provides information about the dynamics and correlations of gene and protein expressions in individual cells, enabling comprehensive study of cellular heterogeneity and expression patterns. Here, we present a platform for high-throughput cellular indexing of mRNAs and proteins, named multi-Paired-seq, with high cell utilization, accurate molecular measurement and low cost. Based on hydrodynamic differential flow resistance, multi-Paired-seq largely improves cell utilization (>95%). Combined with pump/valve structure, cell-free antibodies and mRNAs can be removed completely for highly accurate detection. The picoliter reaction chambers allow higher detection sensitivity and lower sequencing cost. By using multi-Paired-seq, more elaborate classifications of breast cancers are identified according to multimodal measurements, and the expression correlations between mRNAs and proteins under altered conditions are quantified. Multi-Paired-seq provides multi-modal measurements at the single cell level, which offers a new tool for cell biology, developmental biology, drug discovery and precision medicine.

Obtain the GEO accession number: GSE173886

Use GEO2R to compare two or more groups of Samples in order to identify genes that are differentially expressed across experimental conditions. Results are presented as a table of genes ordered by significance. [Full Instructions](#)

GEO accession: GSE173886 Set Expression data from C42B prostate cancer cell line treated with enzalutamide and/or sphingosine kinase inhibitor PF-543 or ABC294640

Group	Accession	Title	Source name
-	GSM5262362	ABC294640_0hr	C42B cells treated with 1t
-	GSM5262363	DMSO_0hr_rep1	C42B cells treated with 0
-	GSM5262364	DMSO_0hr_rep2	C42B cells treated with 0
-	GSM5262365	DMSO_0hr_rep3	C42B cells treated with 0
-	GSM5262366	enzalutamide_0hr_rep1	C42B cells treated with 1t
-	GSM5262367	enzalutamide_0hr_rep2	C42B cells treated with 1t
-	GSM5262368	enzalutamide-ABC294640_0hr_rep1	C42B cells treated with 1t
-	GSM5262369	enzalutamide-ABC294640_0hr_rep2	C42B cells treated with 1t
-	GSM5262370	enzalutamide-ABC294640_0hr_rep3	C42B cells treated with 1t
-	GSM5262371	enzalutamide-PF543_0hr_rep1	C42B cells treated with 1t
-	GSM5262372	enzalutamide-PF543_0hr_rep2	C42B cells treated with 1t
-	GSM5262373	enzalutamide-PF543_0hr_rep3	C42B cells treated with 1t
-	GSM5262374	PF543_0hr_rep1	C42B cells treated with 1t
-	GSM5262375	PF543_0hr_rep2	C42B cells treated with 1t

GEO2R Options Profile graph R script

Version info: R 3.2.3, Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8
Data plots for selected GEO samples
library(GEOquery)

R script:-

```
# Version info: R 3.2.3, Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8
#####
# Data plots for selected GEO samples
library(GEOquery)
library(limma)
library(umap)

# load series and platform data from GEO

gset <- getGEO("GSE173886", GSEMatrix =TRUE, getGPL=FALSE)
if (length(gset) > 1) idx <- grep("GPL23159", attr(gset, "names")) else idx <- 1
gset <- gset[[idx]]

ex <- exprs(gset)
# log2 transform
qx <- as.numeric(quantile(ex, c(0., 0.25, 0.5, 0.75, 0.99, 1.0), na.rm=T))
LogC <- (qx[5] > 100) ||
  (qx[6]-qx[1] > 50 && qx[2] > 0)
if (LogC) { ex[which(ex <= 0)] <- NaN
  ex <- log2(ex) }

# box-and-whisker plot
par(mar=c(7,4,2,1))
title <- paste ("GSE173886", "/", annotation(gset), sep ="")
boxplot(ex, boxwex=0.7, notch=T, main=title, outline=FALSE, las=2)

# expression value distribution plot
par(mar=c(4,4,2,1))
title <- paste ("GSE173886", "/", annotation(gset), " value distribution", sep ="")
plotDensities(ex, main=title, legend=F)

# mean-variance trend
ex <- na.omit(ex) # eliminate rows with NAs
plotSA(lmFit(ex), main="Mean variance trend, GSE173886")

# UMAP plot (multi-dimensional scaling)
ex <- ex[!duplicated(ex), ] # remove duplicates
ump <- umap(t(ex), n_neighbors = 7, random_state = 123)
plot(ump$layout, main="UMAP plot, nbrs=7", xlab="", ylab="", pch=20, cex=1.5)
library("maptools") # point labels without overlaps
pointLabel(ump$layout, labels = rownames(ump$layout), method="SANN", cex=0.6)
```

```

# Version info: R 3.2.3, BiocLite 2.88.0, GEOQuery 2.40.0, limma 3.26.8
#####
# Data plots for selected GEO samples
# https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE173886
library(limma)
library(gplots)
library(grid)

# load series and platform data from GEO
gset <- getGEO("GSE173886", GSEMatrix=TRUE, getGPL=FALSE)
gset$genes$ID[gset$genes$ID %in% grep("GSM523199", gset$genes$name)] <- 1
idx <- grep("GSM523199", gset$genes$name)
gset$genes$ID[idx] <- 1

ex <- exprs(gset)
# log2 transform
ex <- as.data.frame(t(quantile(ex, c(0., 0.25, 0.5, 0.75, 0.99, 1.0), na.rm=T)))
log2_ex <- log2(ex)
log2_ex[is.na(log2_ex)] <- 0
if (log2_ex[is.na(log2_ex)] < 0) log2_ex[is.na(log2_ex)] <- -Inf
ex <- log2(ex)

# box-and-whisker plot
par(mar=c(4,2,1))
boxplot(ex[,1], main="GSE173886", ylab="annotation(gset)", sep="")
boxplot(ex, boxexe=0.7, notch=T, main=title, outline=FALSE, las=2)
# expression value distribution plot
par(mar=c(4,2,1))
boxplot(ex[,1], main="GSE173886", ylab="annotation(gset)", sep="")
plotDensity(ex, main=title, legend=FALSE)
# mean-variance trend
ex <- na.omit(ex) # eliminate rows with NAs
trend(ex, main="mean-variance trend", GSE173886)
# MDS plot (multi-dimensional scaling)
ex <- ex[!(duplicated(ex)), ] # remove duplicates
ump <- umap(t(ex), n_neighbors = 7, random_state = 123)
library("gridtools") # point labels without overlap
label(ump$layout, labels = rownames(ump$layout), method="3ANN", cex=0.6)
points(ump$layout, col="red", pch=19, cex=0.6)
# PCA plot
par(mar=c(4,2,1))
PCA(ex, main="PCA", ylab="annotation(gset)", sep="")
# TMAF plot (Multi-dimensional scaling)
ex <- ex[!(duplicated(ex)), ] # remove duplicates
tmaf(ex, main="TMAF", ylab="annotation(gset)", sep="")
# mean-variance trend
ex <- na.omit(ex) # eliminate rows with NAs
trend(ex, main="mean-variance trend", GSE173886)
# MDS plot (multi-dimensional scaling)
ex <- ex[!(duplicated(ex)), ] # remove duplicates
ump <- umap(t(ex), n_neighbors = 7, random_state = 123)
library("gridtools") # point labels without overlap
label(ump$layout, labels = rownames(ump$layout), method="3ANN", cex=0.6)

```

Train it for differentiation between DMSO and Enzalutamide:

2 samples of DMSO and 5 samples of Enzalutamide altered.

Samples		Define groups	
Group	Sample ID	Group	Title
enzalutamide	282362	enzalutamide	ABC294640_0hr
enzalutamide	282363	enzalutamide	DMSO_0hr_rep1
enzalutamide	282364	enzalutamide	DMSO_0hr_rep2
enzalutamide	GSM5282365	enzalutamide	DMSO_0hr_rep3
enzalutamide	GSM5282366	enzalutamide	enzalutamide_0hr_rep1
DMSO	GSM5282367	enzalutamide	enzalutamide_0hr_rep2
DMSO	GSM5282368	enzalutamide	enzalutamide-ABC294640_0hr_rep1
	GSM5282369	enzalutamide	enzalutamide-ABC294640_0hr_rep2
	GSM5282370	enzalutamide	enzalutamide-ABC294640_0hr_rep3
	GSM5282371	enzalutamide	enzalutamide-PF543_0hr_rep1
	GSM5282372	enzalutamide	enzalutamide-PF543_0hr_rep2
	GSM5282373	enzalutamide	enzalutamide-PF543_0hr_rep3
	GSM5282374	PF543	PF543_0hr_rep1
	GSM5282375	PF543	PF543_0hr_rep2

The screenshot shows the GEO2R web application interface. At the top, a navigation bar includes links for Dashboard, Course: Ph, Inbox (605), Carmustine, Kites "Dil K, Curcumin, GEO2R - GI, GEO Access, and a search bar with the URL https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE173886.

The main content area displays a table of samples under the heading "Samples". The table has columns for "Group", "Sample ID", and "Title". The "Group" column shows three groups: DMSO, DMSO, and DMSO. The "Title" column lists various sample names such as ABC294640_6hr, DMSO_6hr_rep1, DMSO_6hr_rep2, etc. A dropdown menu titled "Define groups" is open, showing a list of samples: DMSO (3 samples), enzalutamide, and GSM5282364. The "enzalutamide" entry is currently selected.

The screenshot shows the results page for the GEO2R analysis of GSE173886. The top navigation bar and search bar are identical to the previous screenshot.

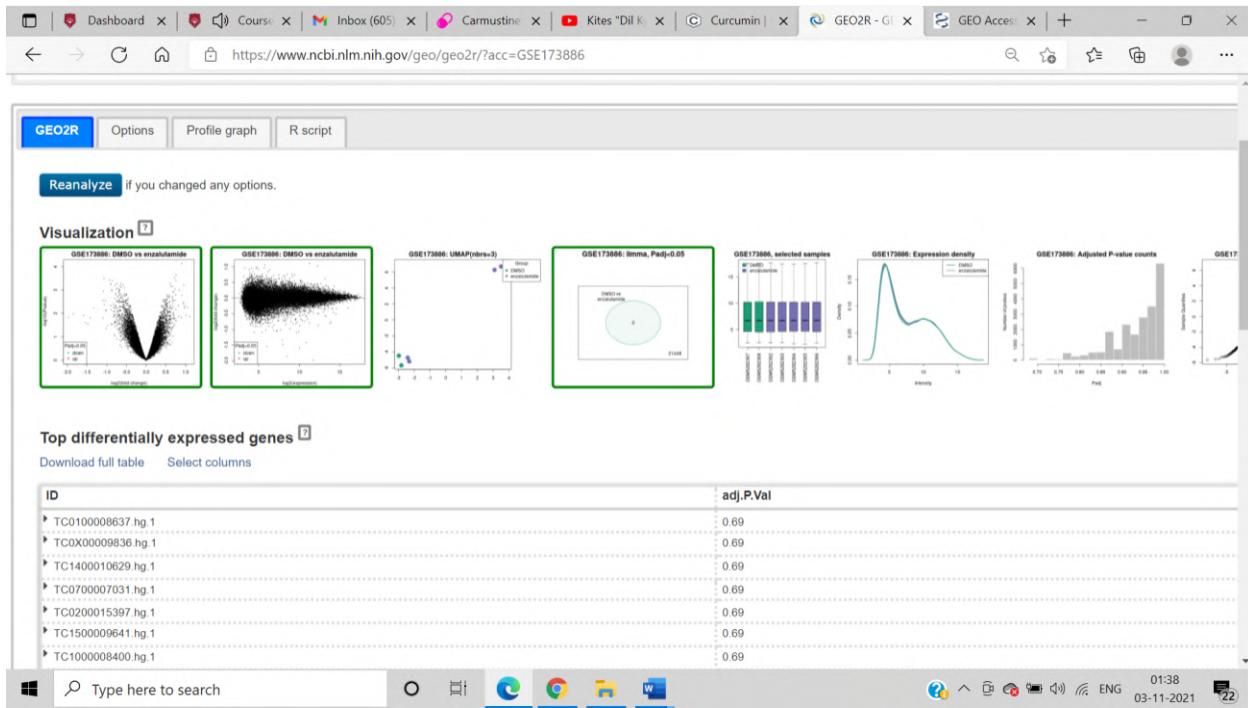
The main content area displays a table of samples with the same structure as the previous screenshot. The "Title" column now includes additional entries for enzalutamide-ABC294640_6hr_rep1 through rep3, and enzalutamide-PF543_6hr_rep1 through rep3. The status bar at the bottom right indicates the date and time: 03-11-2021 01:32.

The screenshot shows the options page for the GEO2R analysis of GSE173886. The top navigation bar and search bar are identical to the previous screenshots.

The main content area contains a sidebar with various options. It includes fields for "Significance level cut-off" (set to 0.05) and "Volcano and MA plot contrasts" (set to 0 selected). There is also a checkbox for "DMSO vs enzalutamide".

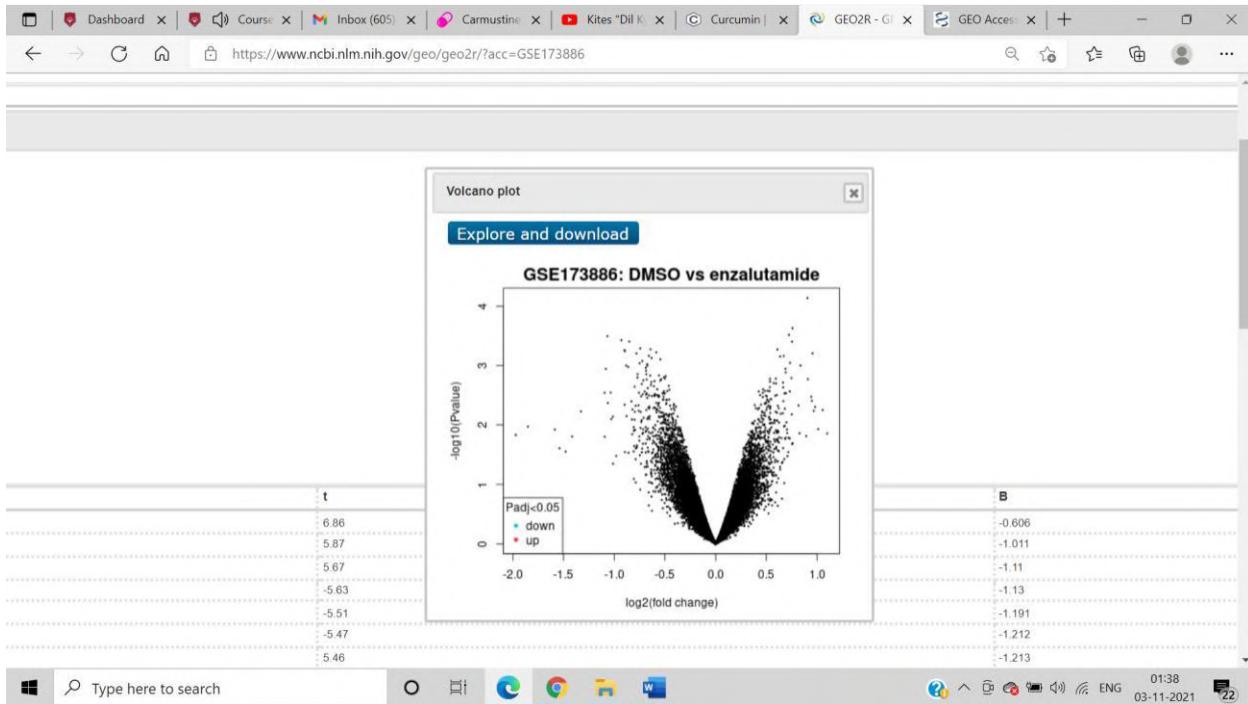
The screenshot shows the Windows taskbar. The address bar displays the URL https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE173886#options. The taskbar also shows icons for File Explorer, Edge browser, Google Chrome, File Explorer, and Microsoft Word.

Following is the analysis for DMSO and Enzalutamide:



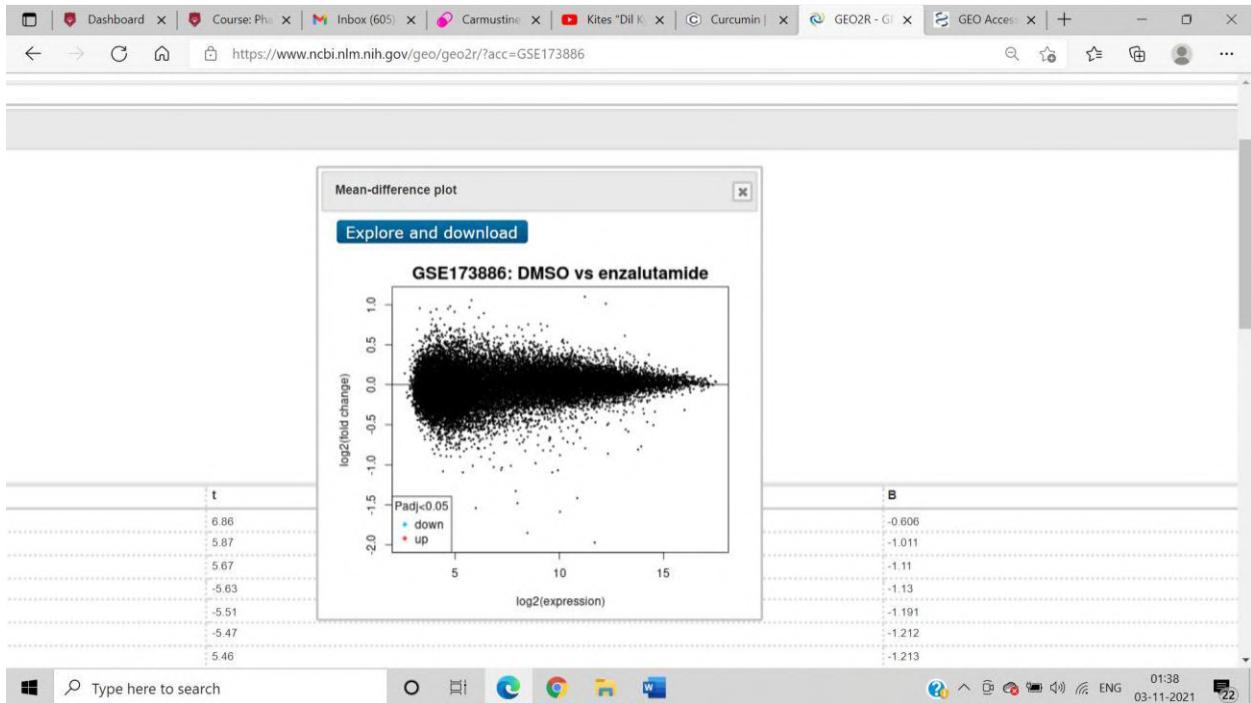
Volcano Plot:

GSE173886 of DMSO vs Enzalutamide. It's between the fold change and Pvalue graph chart.



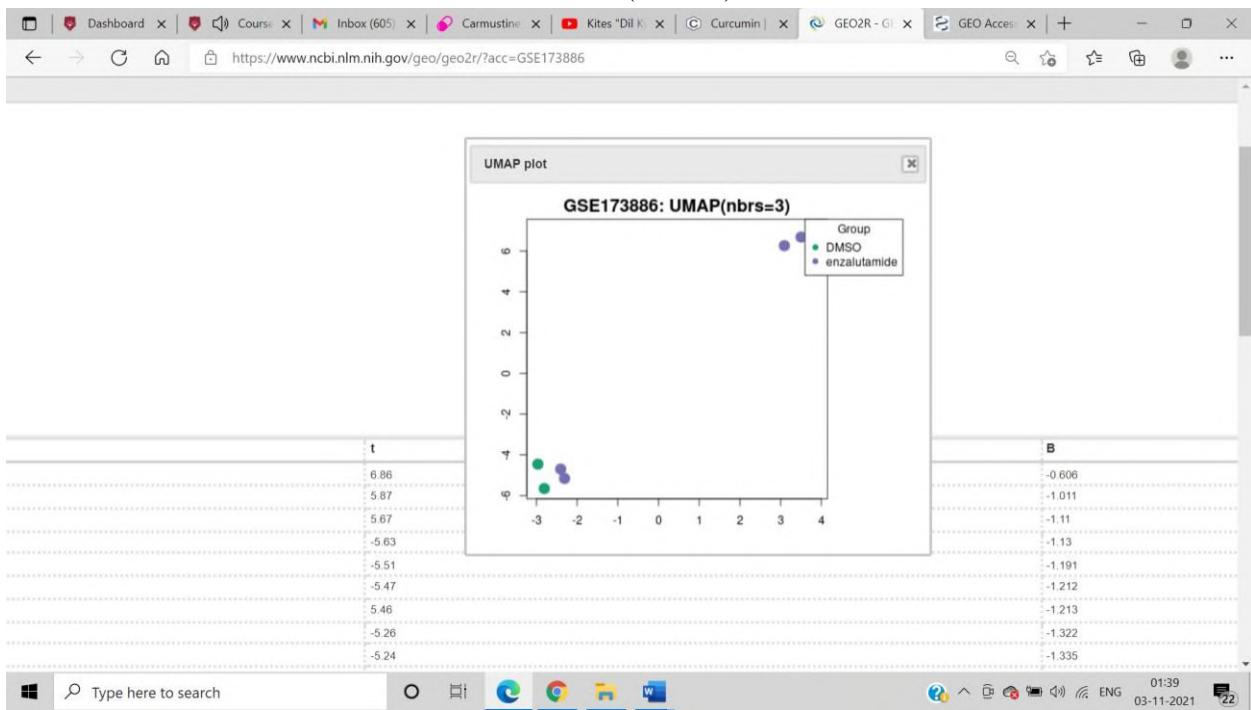
Mean-difference plot:-

GSE173886 of DMSO vs Enzalutamide. Its between the fold change and Pvalue graph chart



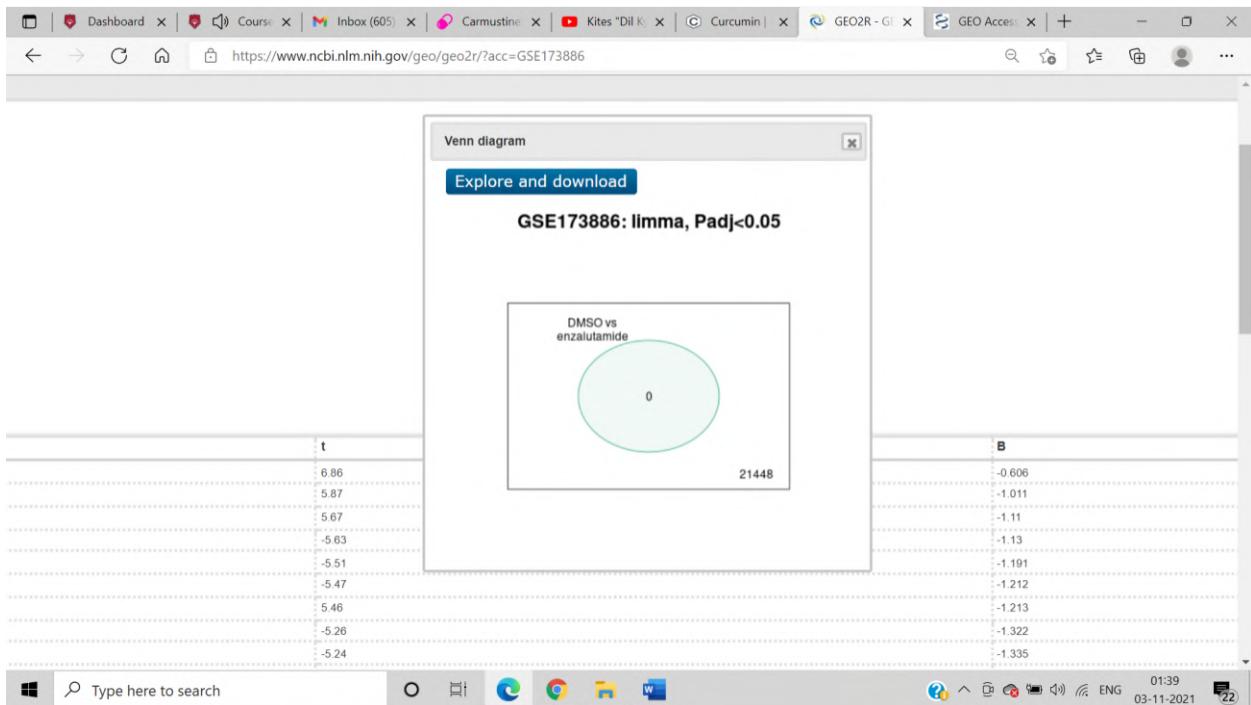
UMAP plot:-

GSE173886 of DMSO vs Enzalutamide. UMAP(nbrs=3)

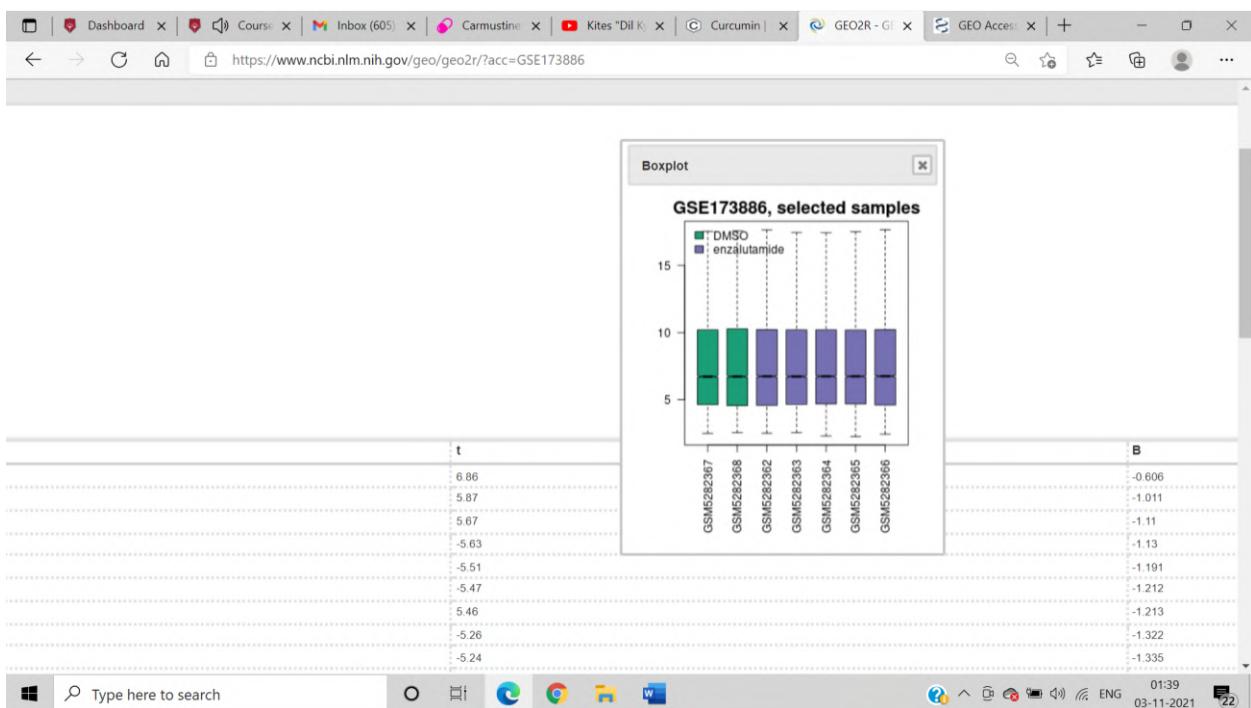


Venn Diagram:

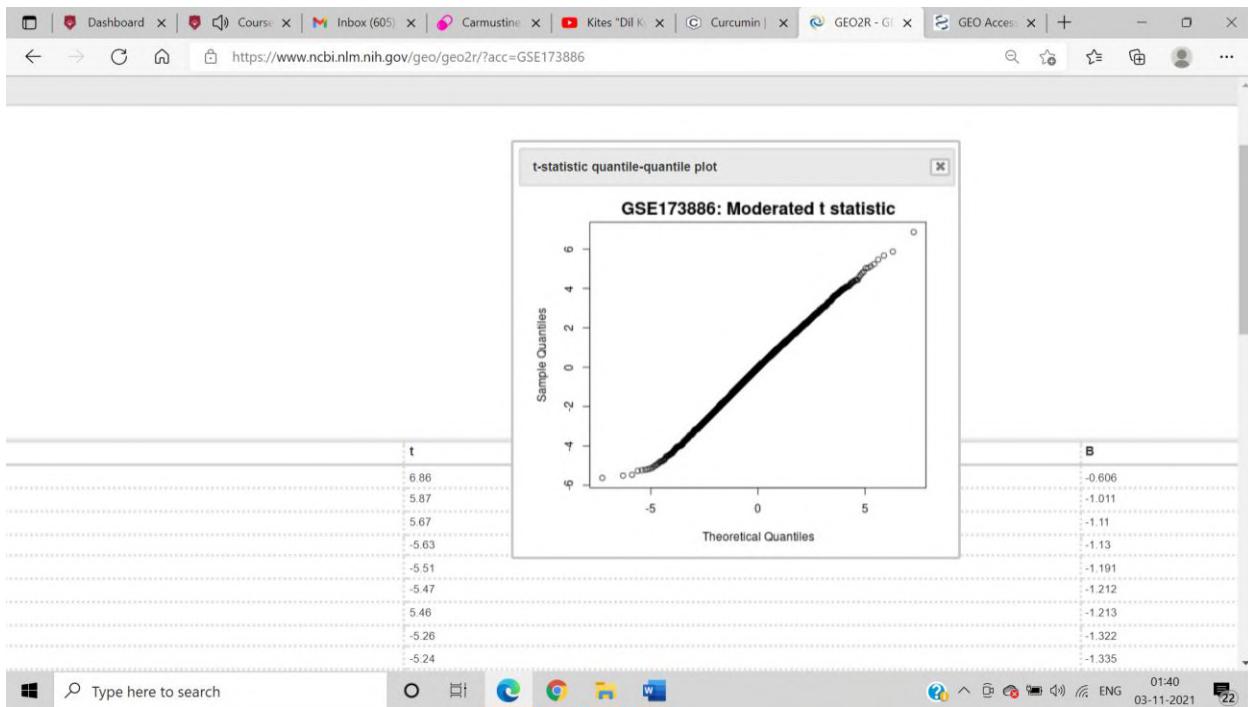
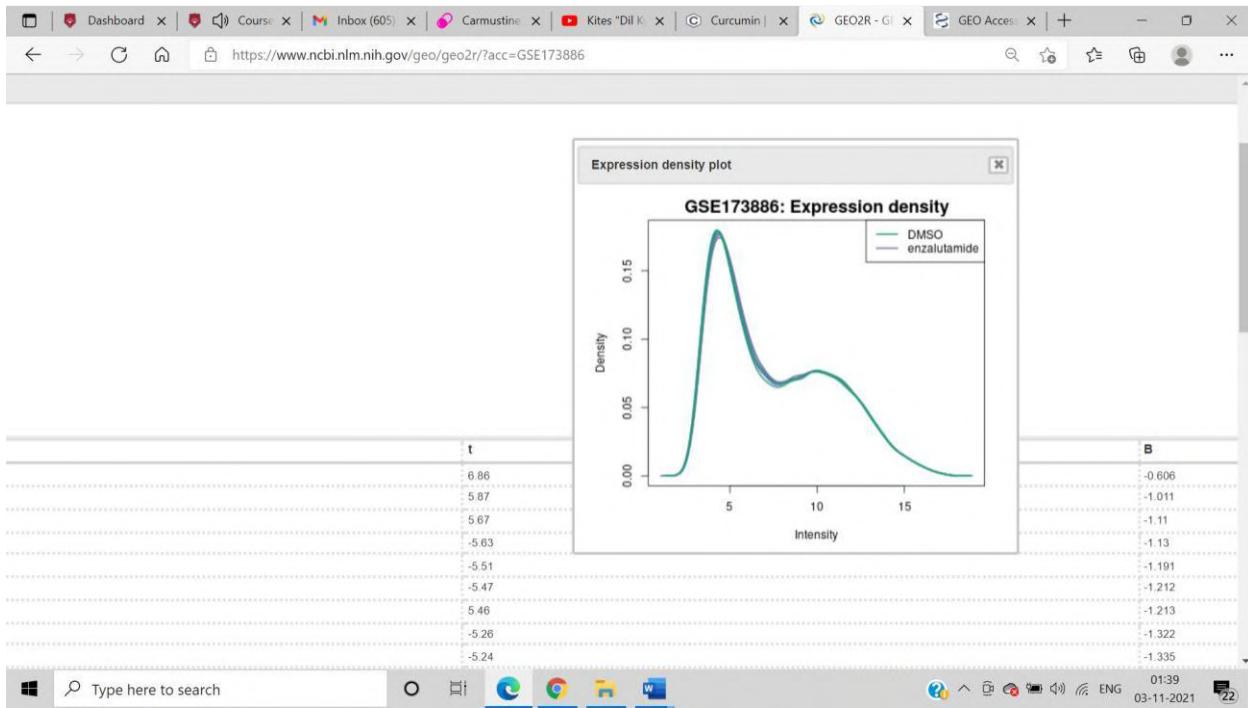
GSE173886: Ilmma, Padj<0.05

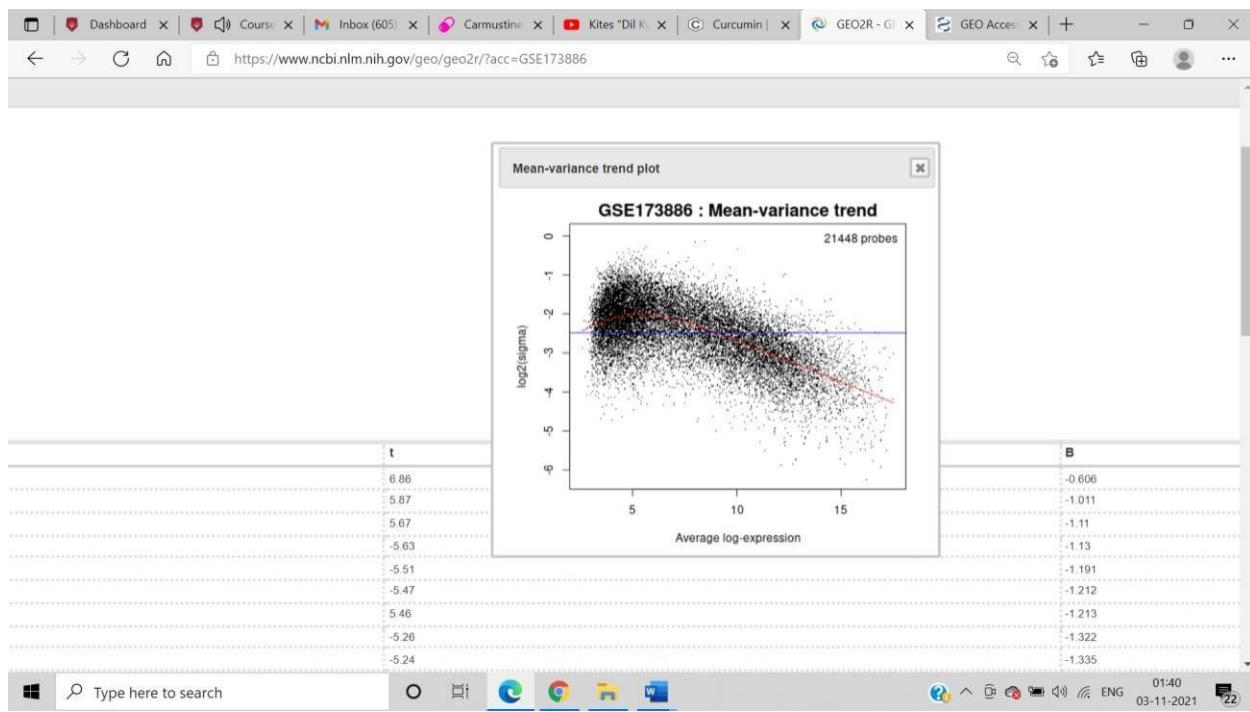


GSE173886 :- Box plot for the selected samples:



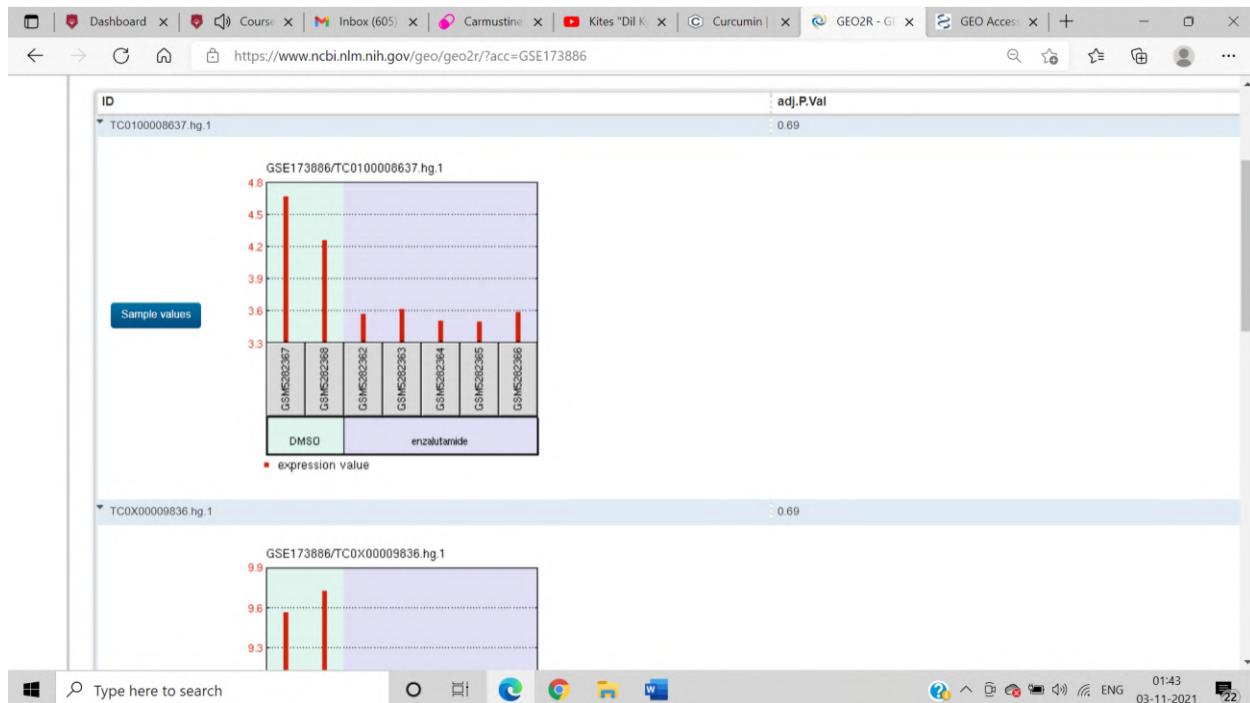
Expression density plot :

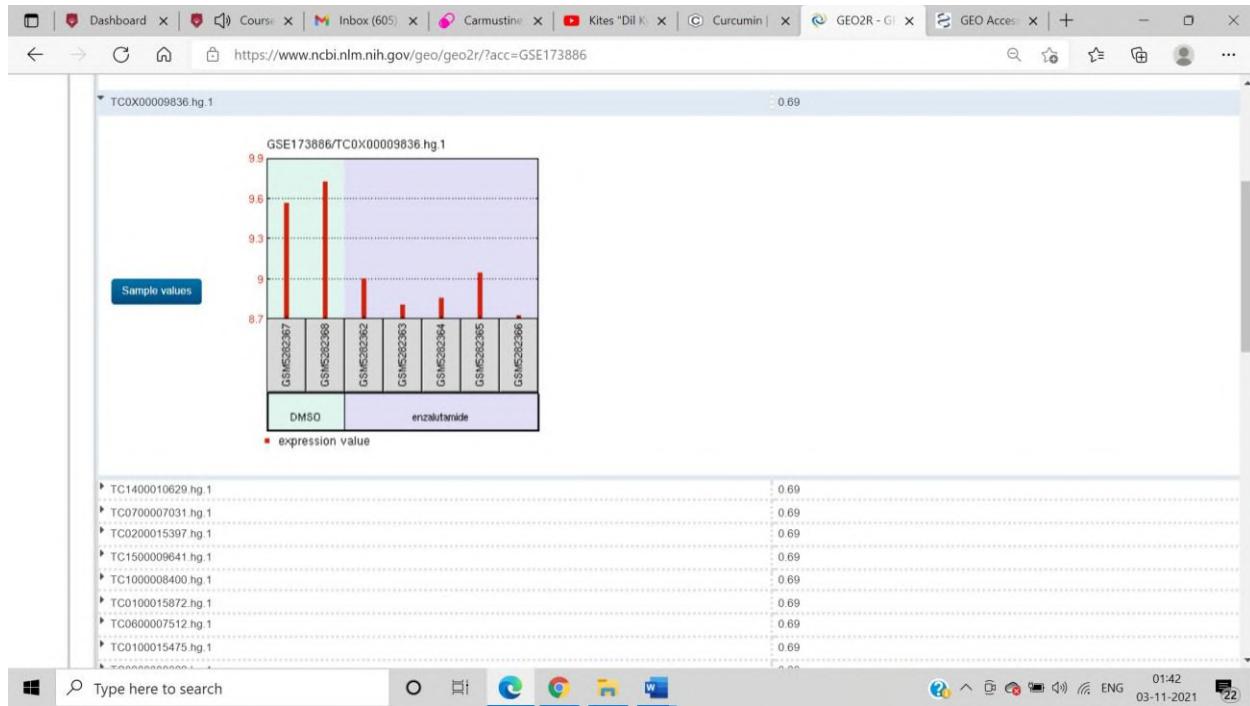




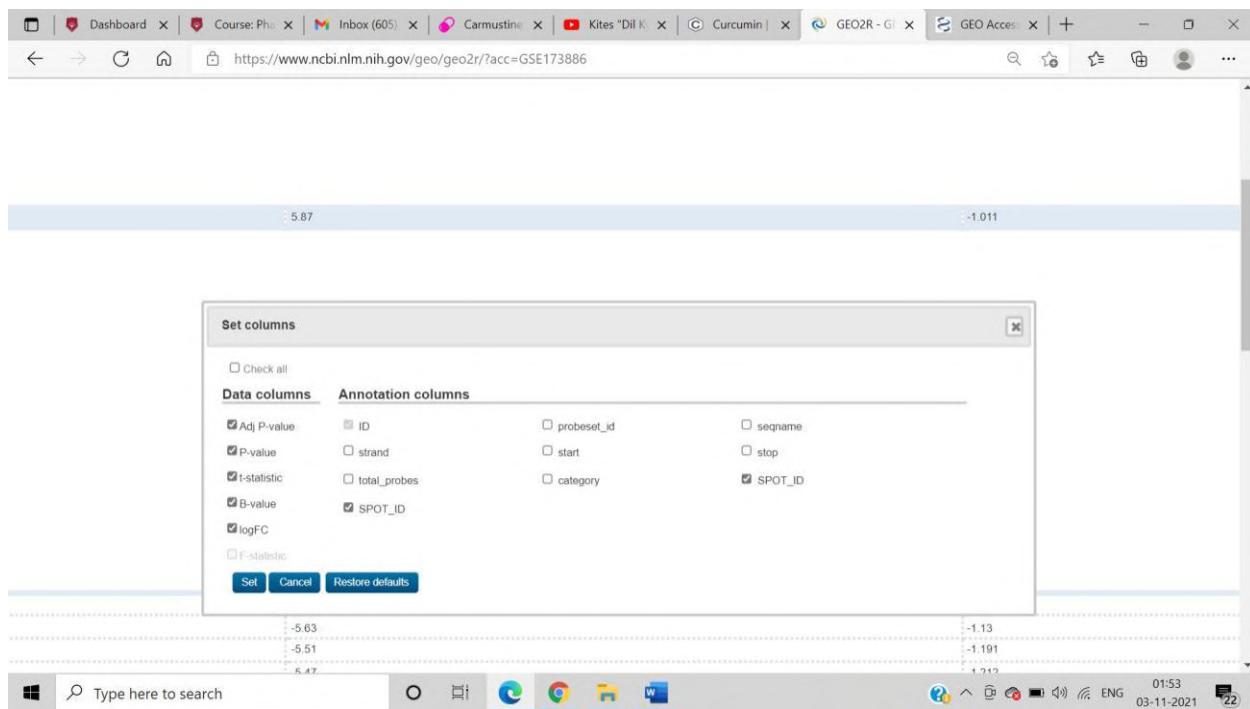
ID no:TC0100008637.hg.1

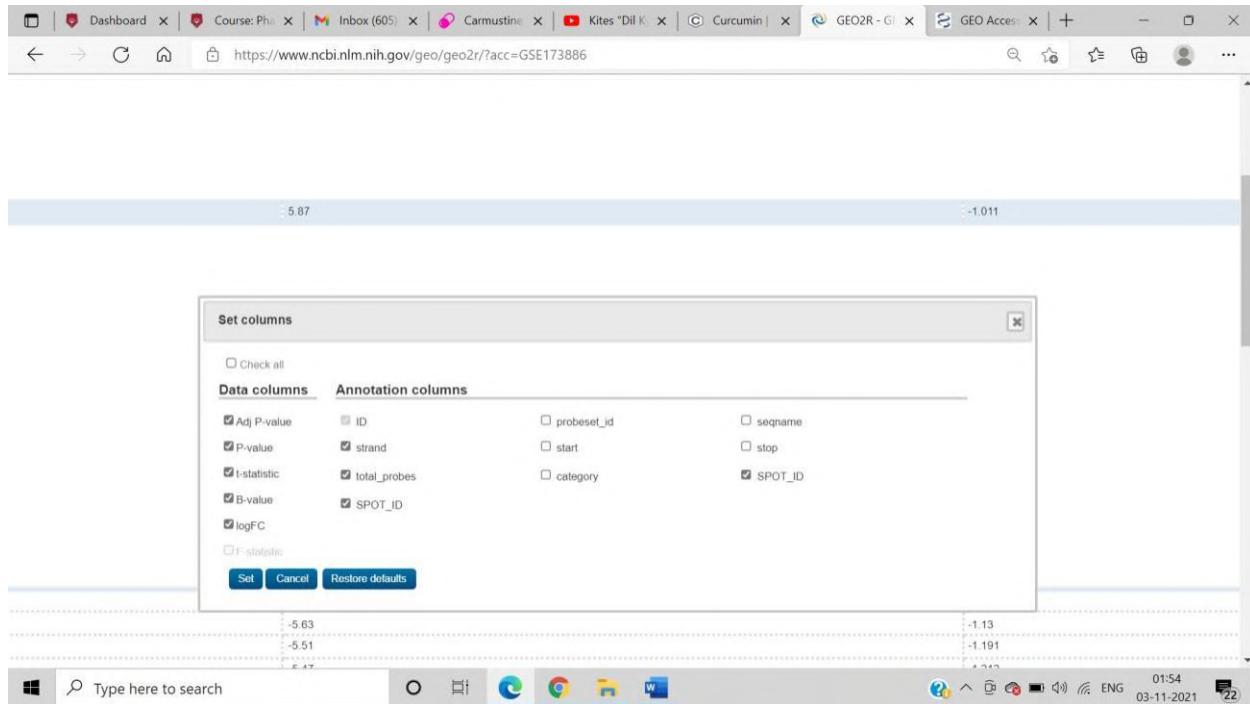
Red=expression value





You can edit the columns:





Result:

GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

Conclusion:

In this case, we chose CANCER AND DRUG TREATMENT as the key-term and got many hits. GEO has a great service of showing the Cluster AnalysisGEO provide user-friendly mechanisms that allow users to locate, review and download gene expression profiles of interest. It displays the data in graphical as well as in the tabular form.

Experiment no: 10B

Aim: ArrayExpress

Theory: ArrayExpress is one of the major public repositories for functional genomics datasets. Most of the data is genome-wide gene expression data, measured on microarray or next-generation sequencing (NGS) platforms. A range of DNA assays are also hosted by ArrayExpress, such as ChIP-seq or genotyping.

The main object in ArrayExpress is the experiment. An experiment usually groups several assays belonging to one study or publication. Each experiment contains metadata describing the biological specimen and experimental procedures, as well as resulting data files. All data and files in ArrayExpress are provided by the user, either submitted directly or imported from other databases, such as Gene Expression Omnibus (GEO) at NCBI. Directly submitted datasets are manually curated to promote compliance with the MIAME and MINSEQE guidelines. These minimal information standards support the sharing and reuse of scientific data.

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression, provide the scientific community with freely available information on the abundance and localisation of RNA (and proteins) across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others.

Methodology:

1) Go to the homepage of ArrayExpress and search your disease

Disease: Glioma

Drug: Carmustine, gliadel

The screenshot shows the ArrayExpress website homepage. At the top, there is a navigation bar with links to EMBL-EBI, Services, Research, Training, and About us. The EMBL-EBI logo and Hinxtion logo are also present. A search bar at the top right contains the query "Glioma". Below the search bar, there is a link to "advanced search" and examples of search terms: E-MEXP-31, cancer, p53, Geuvadis. The main content area features a large title "ArrayExpress – functional genomics data" and a sub-section titled "Data Content" which states "Updated today at 02:00" and lists statistics: 74924 experiments, 2567549 assays, and 61.27 TB of archived data. There is also a "Browse ArrayExpress" button. On the left, there is a "Latest News" section with a single item: "1 October 2020 - ArrayExpress is moving to BioStudies". A banner at the bottom of the page requires users to accept cookies and provides links to Privacy Notice and Terms of Use.

2) So we got 942 hits of experiments for glioma disease:-

The screenshot shows the search results page for "Glioma" on ArrayExpress. The search bar at the top contains the query "Glioma". The main content area is titled "Search results for Glioma" and displays a table of 942 experiments. The table has columns for Page (1-25 of 942), Accession, Title, Type, Organism, Assays, Released, Processed, Raw, and Atlas. The first few rows of the table are:

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Atlas
E-MTAB-9860	mRNA-Seq of zebrafish embryos (96hpf) exposed to different concentrations of the organochloride Methoxychlor below acute toxicity levels against untreated control groups	RNA-seq of coding RNA	Danio rerio	12	15/10/2021			-
E-MTAB-9859	mRNA-Seq of zebrafish embryos (96hpf) exposed to different concentrations of Imidacloprid below acute toxicity levels against untreated control groups	RNA-seq of coding RNA	Danio rerio	12	15/10/2021			-
E-MTAB-9855	mRNA-Seq of zebrafish embryos (96hpf) exposed to different concentrations of Carbaryl below acute toxicity levels against untreated control groups	RNA-seq of coding RNA	Danio rerio	9	15/10/2021			-
E-MTAB-9854	mRNA-Seq of zebrafish embryos (96hpf) exposed to different concentrations of the organochloride Methoxychlor below acute toxicity levels against untreated control groups	RNA-seq of coding RNA	Danio rerio	9	15/10/2021			-

A banner at the bottom of the page requires users to accept cookies and provides links to Privacy Notice and Terms of Use.

Results:

- 1) E-MTAB-9860 - mRNA-Seq of zebrafish embryos (96hpf) exposed to different concentrations of the organochloride Methoxychlor below acute toxicity levels against untreated control groups

The screenshot shows a web browser window with the URL [ebi.ac.uk/arrayexpress/experiments/E-MTAB-9860/?query=Glioma](https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9860/?query=Glioma). The page is titled "ArrayExpress" and displays information for experiment E-MTAB-9860, which is a mRNA-Seq study of zebrafish embryos (Danio rerio) exposed to Methoxychlor at various concentrations. Key details include 12 samples and 8 protocols. A banner at the bottom states: "This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our [Privacy Notice](#) and [Terms of Use](#)." There is also a link to "I agree, dismiss this banner". The browser interface shows tabs for WhatsApp, BID-180, Rotaract, Ligand, www.ebi, ArgusLab, ChemAx, and E-MTAB, along with standard navigation and search bars.

Description:-

The aim of this mRNA expression profiling experiment was to screen for ecotoxicogenomic fingerprints in zebrafish (*Danio rerio*) embryos as aquatic vertebrate non-target model exposed to sub lethal concentrations of synthetic organochloride insecticide Methoxychlor (CAS 72-43-5). It is structurally highly similar to its precursor molecule DDT (Dichlorodiphenyltrichloroethane). Like its precursor, today Methoxychlor is banned from the use as pesticide in the United States and the European Union due to its acute toxicity, high bioaccumulation potential and endocrine disruption activity. The Insecticide Resistance Action Committee (IRAC) classified Methoxychlor after its mode of action (MoA) in the target organism as a sodium channel modulator (Group 3B). Our goal is to identify toxicogenomic profiles with predictive character and potential molecular key events (KE) explaining upstream adverse effects in aquatic non-target organisms for this substance of particularly high environmental concern. This will provide useful information to refine and improve existing adverse outcome pathways (AOP). Furthermore, integrating the obtained profiles for this and other tested chemicals in a collective database will enable us in the future to derive predictions about the ecotoxicological hazard for chemicals with unknown apical effects, based on similarly altered transcriptomic and proteomic profiles. In a modified version of the zebrafish embryo toxicity test (OECD 236), 15 fertilized

eggs were exposed to two different sub lethal concentrations of Methoxychlor for 96 hours under semi-static conditions. Each test comprised of a low exposure (LE, 20 µg/L), mid exposure (ME, 60 µg/L), high exposure (HE, 180 µg/L) and negative control (NC) group and was performed in triplicates. At 96 hours post fertilization (hpf), 10 larvae were randomly picked for each sample and pooled for RNA and protein extraction with NucleoSpin RNA/Protein kit (Macherey-Nagel). RNA quality was assessed with a 2100 Bioanalyzer system (Agilent) before coding RNA was purified (PolyA selection with TruSeq RNA Library Prep Kit v2) and sequenced on an Illumina HiSeq 4000 System (Illumina) in 50 bp single read mode, producing roughly 30 million reads per sample. Adapter sequences were removed with trimmomatic and sequences were aligned to the D.rerio reference genome GRCz11 with STAR. Counting of feature mapped reads was performed through featureCounts. Library gene count tables were then merged to a single count matrix as input for differential gene expression analysis with DESeq2

Links:

So below files you can get direct links to GEO or ENA-ERP links as shown below:

This service is part of the ELIXIR infrastructure
ArrayExpress is an ELIXIR Core Data Resource [Learn more](#)

This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our [Privacy Notice](#) and [Terms of Use](#).

I agree, dismiss this banner

Result: Using arrayexpression atlas we obtain results for baseline expression which provide us with graphical results for the experiments and classify whether the data is obtained using different features and functions of the site.

Conclusion:

ArrayExpress is a database that stores data from high-throughput functional genomics experiments. It helps in retrieving the information of various samples that are related to the studies of gene expression analysis. ArrayExpress is one of the repositories recommended by major scientific journals to archive functional genomics data from microarray and sequencing platforms to support reproducible research.

For high-throughput sequencing based experiments the raw data is brokered to the European Nucleotide Archive, while the experiment descriptions and processed data are archived in ArrayExpress.

Experiment No: 11A

AIM: MOLECULAR DYNAMICS USING HYPERCHEM: Study of small molecule.

THEORY:

Hypercube, Inc. is a scientific software company, incorporated in 1985, specializing in molecular modeling software. HyperChem is, noted for its ease of use, extensive functionality, and modest price. Our most important platform is Microsoft Windows; HyperChem, on a PC under Windows, has the largest number of installations of any full-featured molecular modeling program.

Hypercube, Inc. produces two versions of the core HyperChem product: HyperChem 8.0 and HyperChem for MAC. HyperChem 8.0 includes the Chemist's Developer Kit, an advanced customization tool; HyperNMR, for a priori simulation of NMR spectra; and HyperChem Data, a chemical database program with over 10,000 molecules included.

Hypercube, Inc. has set new standards for ease of use and molecular modeling power on PC-based systems. The goal is to bring molecular modeling to all chemists and chemistry students.

METHODOLOGY:

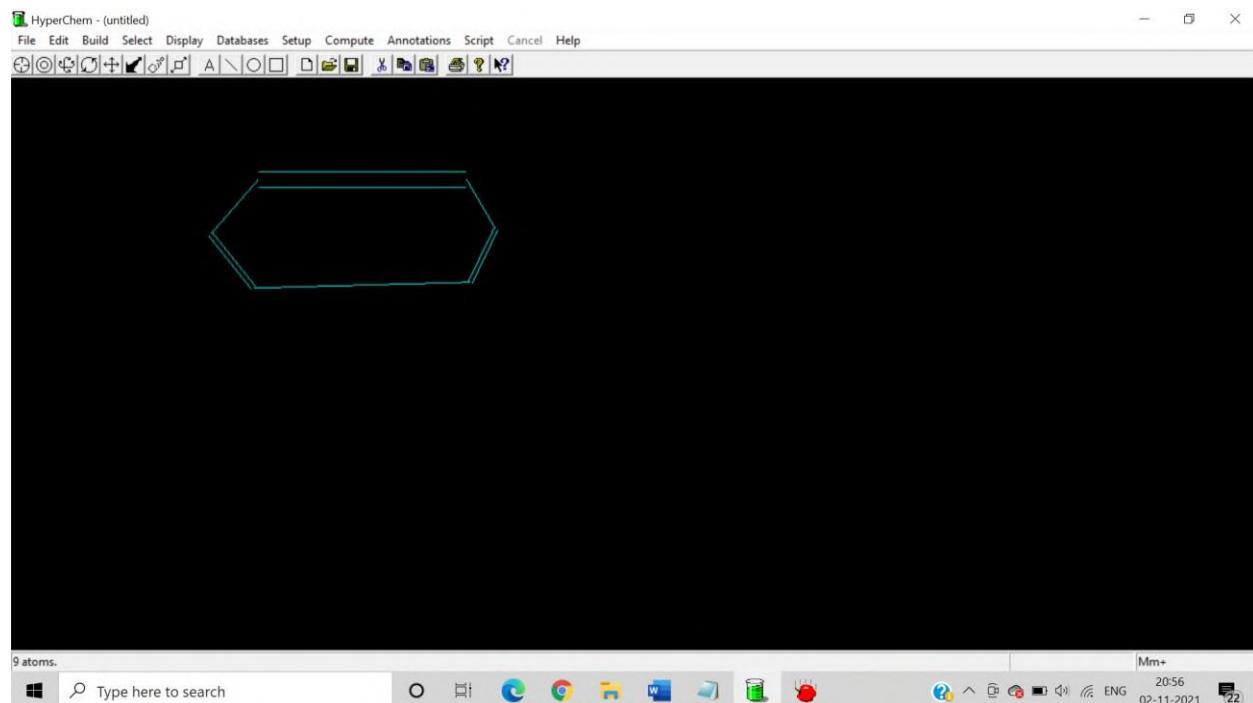
1. Start log through files option. It'll keep a track of every step. To create a structure, click on draw twice to open the periodic table. Click on the molecule you want to use to create a structure.
2. Click on setup and molecular mechanics and click on ok on the dialogue box(MM+).
3. Click on compute and single point to find the energy and gradient of the molecule.
4. Click on compute and geometry optimization to get an optimal conformation of the molecule. In the dialogue box that appears, change maximum cycles to 500 to get the best results, converging of the molecule is important.
5. If the molecule is too big for the window, click on edit and zoom to change the magnification, to fit it to window.
6. To show double bonds, left click on the bond. To delete a bond, right click.
7. Since the molecule has been changed, geometry optimization to get the optimal conformation is necessary.
8. To add a side chain to the molecule, open the periodic table and click on the molecule you want to add. Perform geometry optimization again, with the maximum cycles as 1500 or just change the number and save the molecule (mol1).
9. For the same small molecule, open periodic box from setup and click on ok. The molecule will be surrounded by water molecules.
10. Perform geometry optimization to find the optimal conformation. Change number of maximum cycles.
11. Click on compute and then molecular dynamics. Click on average and click on EKIN, EPOT, ETOT, TEMP and add it to average and graph column, since we want the result

for both. Click on ok and proceed. The system will take some time to calculate and show the average and graph.

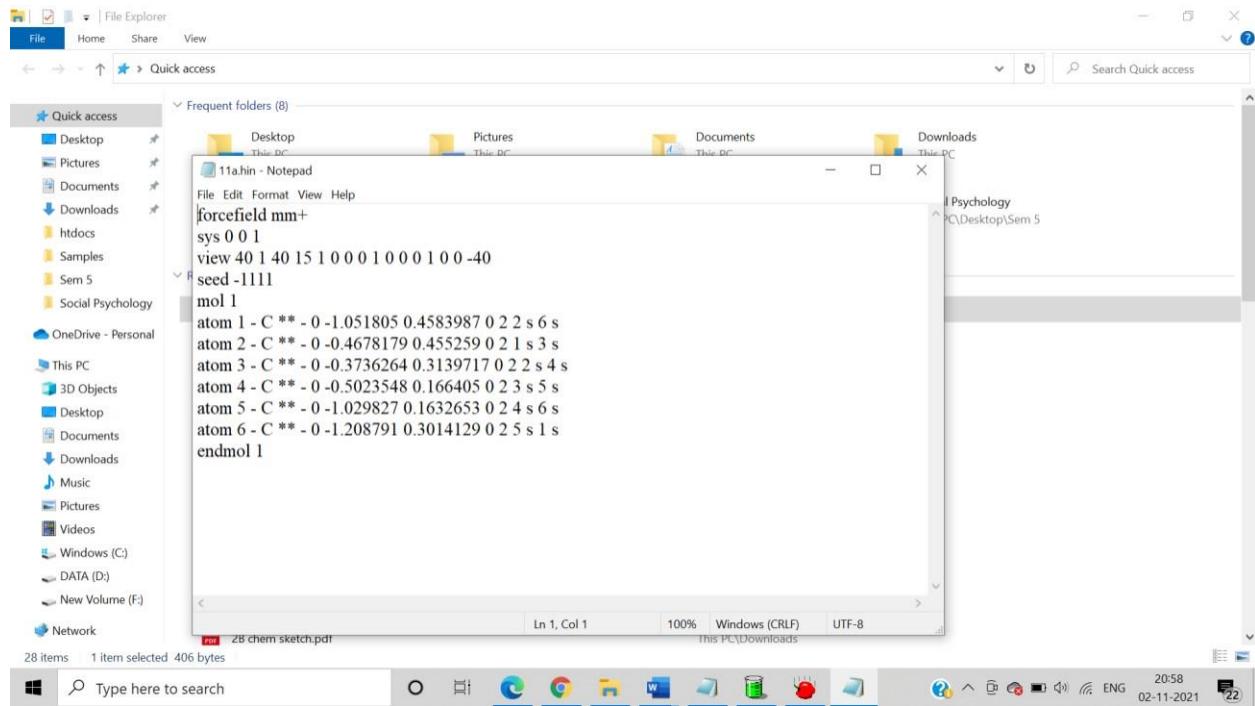
12. Open the excel file to see the log of change in kinetic energy, potential energy, total energy and temperature.

RESULT:

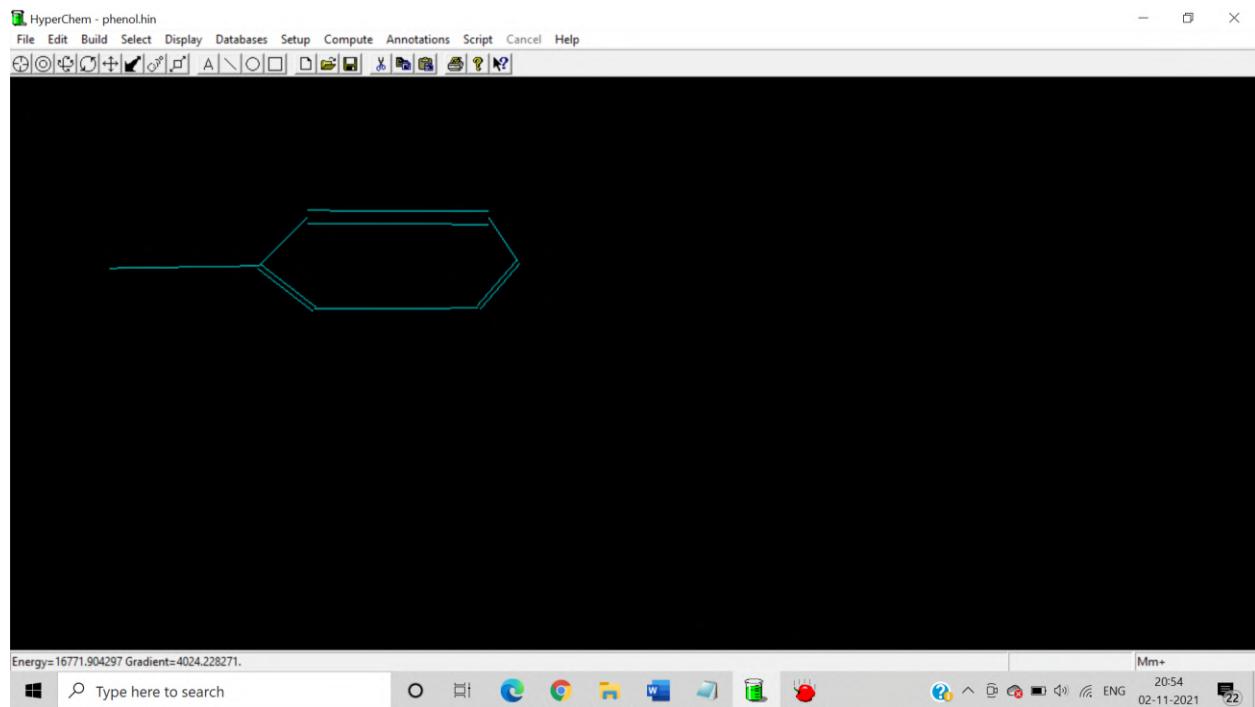
→ You will see 9 atom created.



Automatically a note file is created as you save the file.



Given is the energy and gradient before optimazitaion:



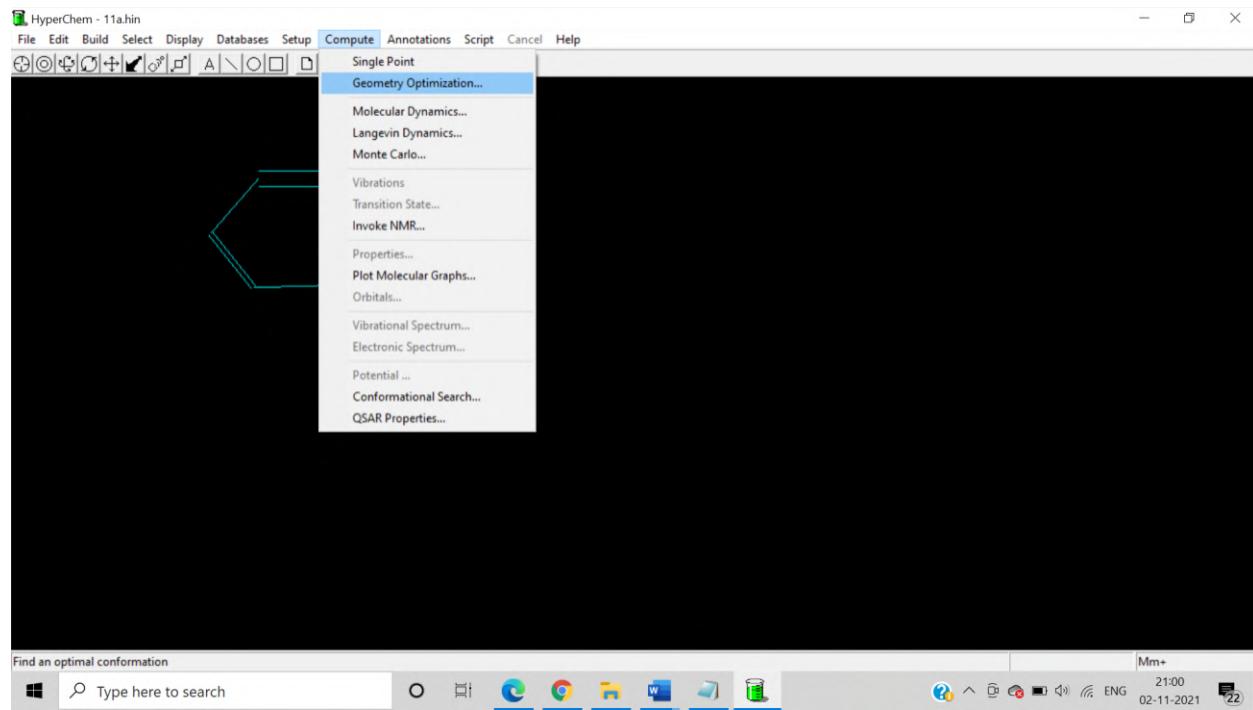
→ Energy = 16771.904297

Gradient = 4024.228271

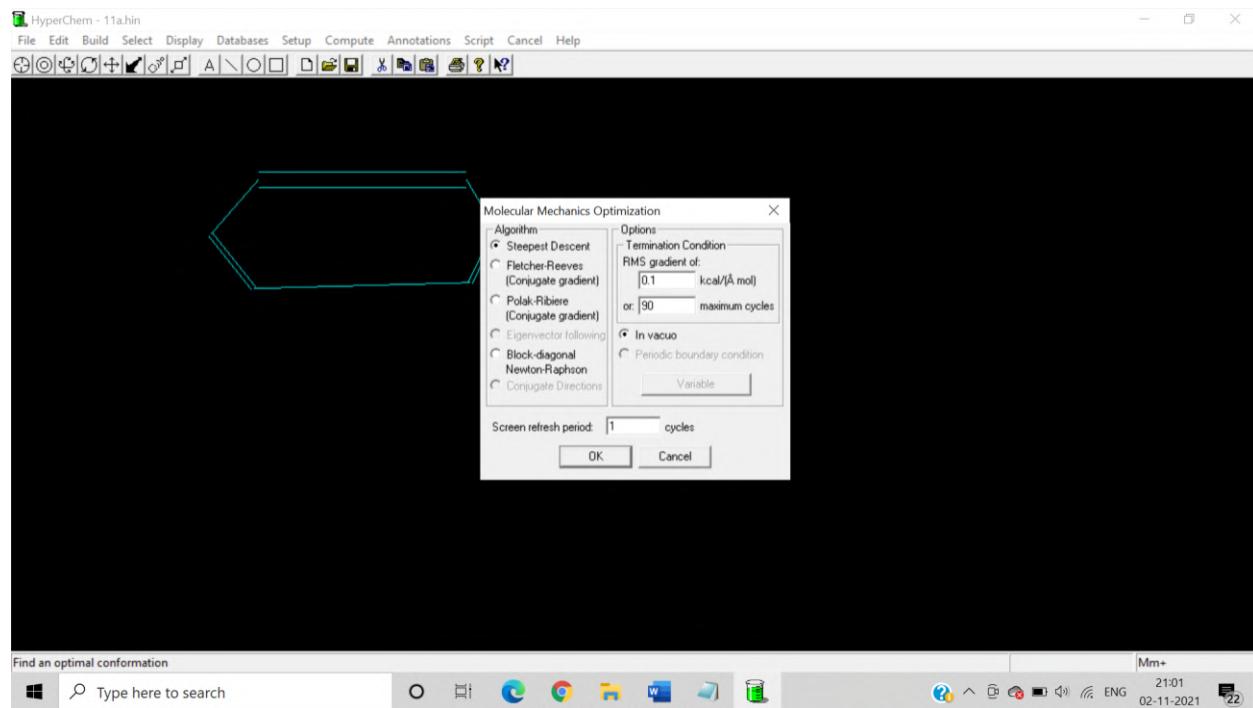
Roll No: BID 19006

204

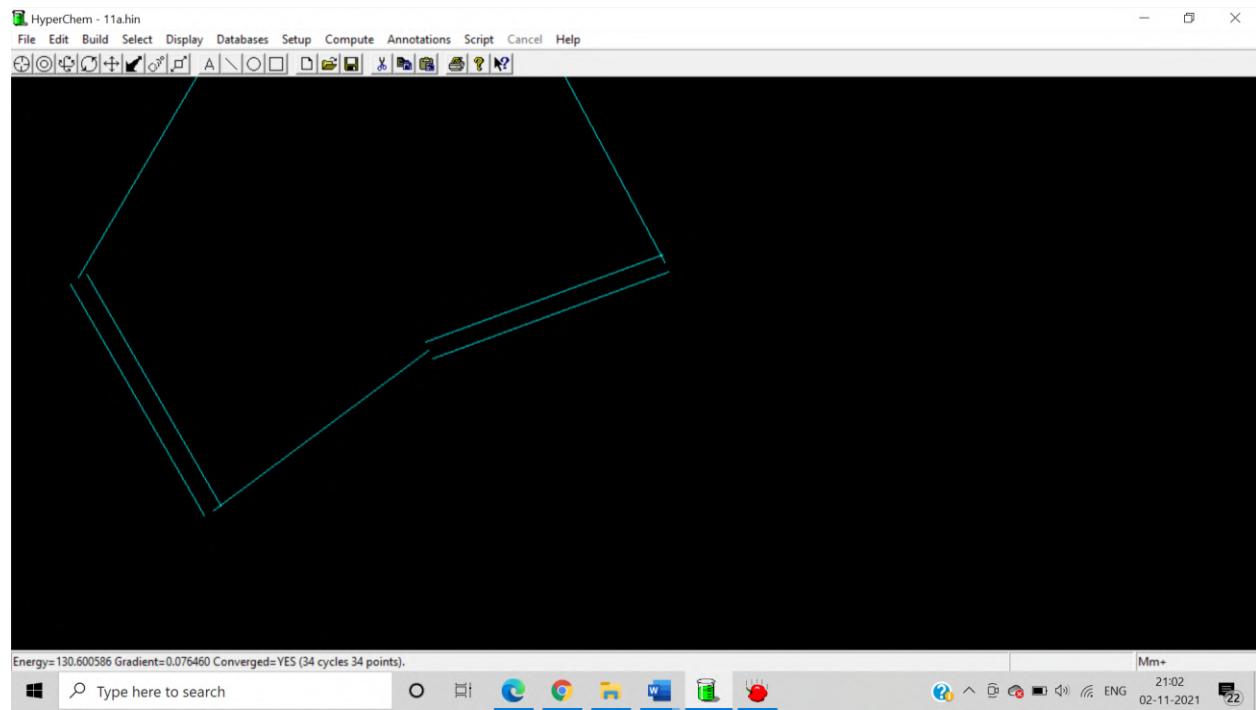
Geometry Optimization:



Go for steepest descent:



Given is the energy and gradient after geometry optimization:

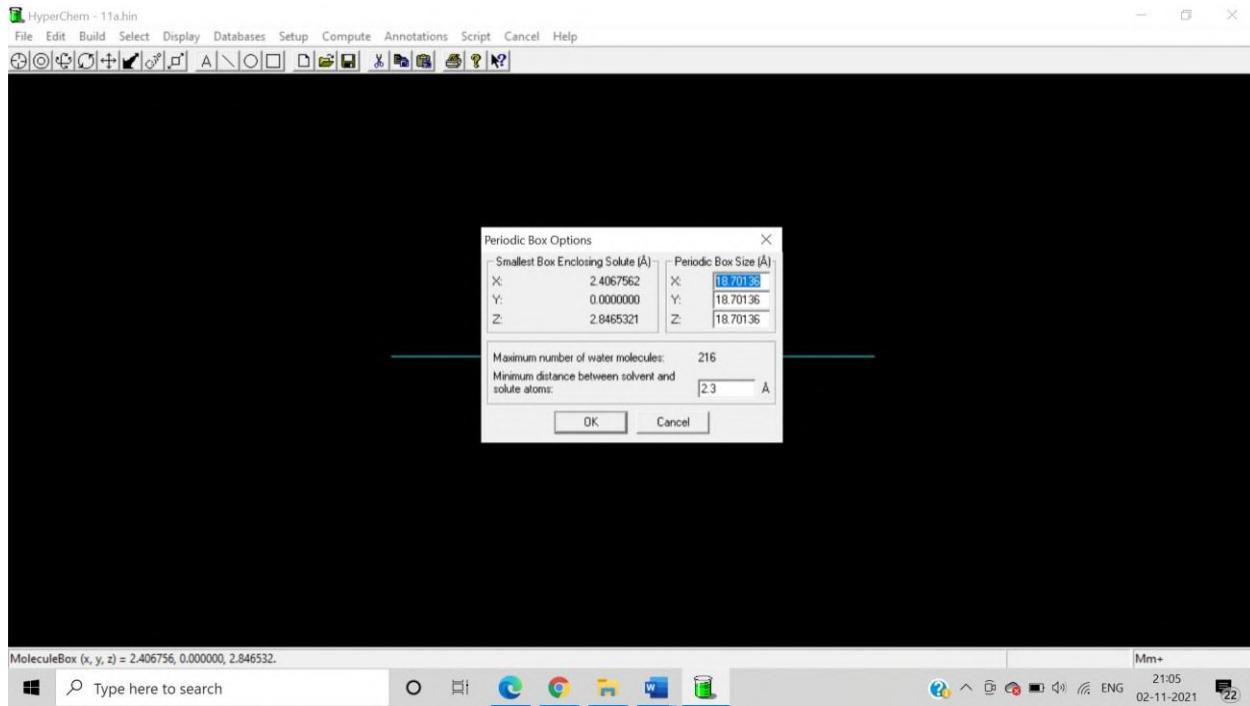
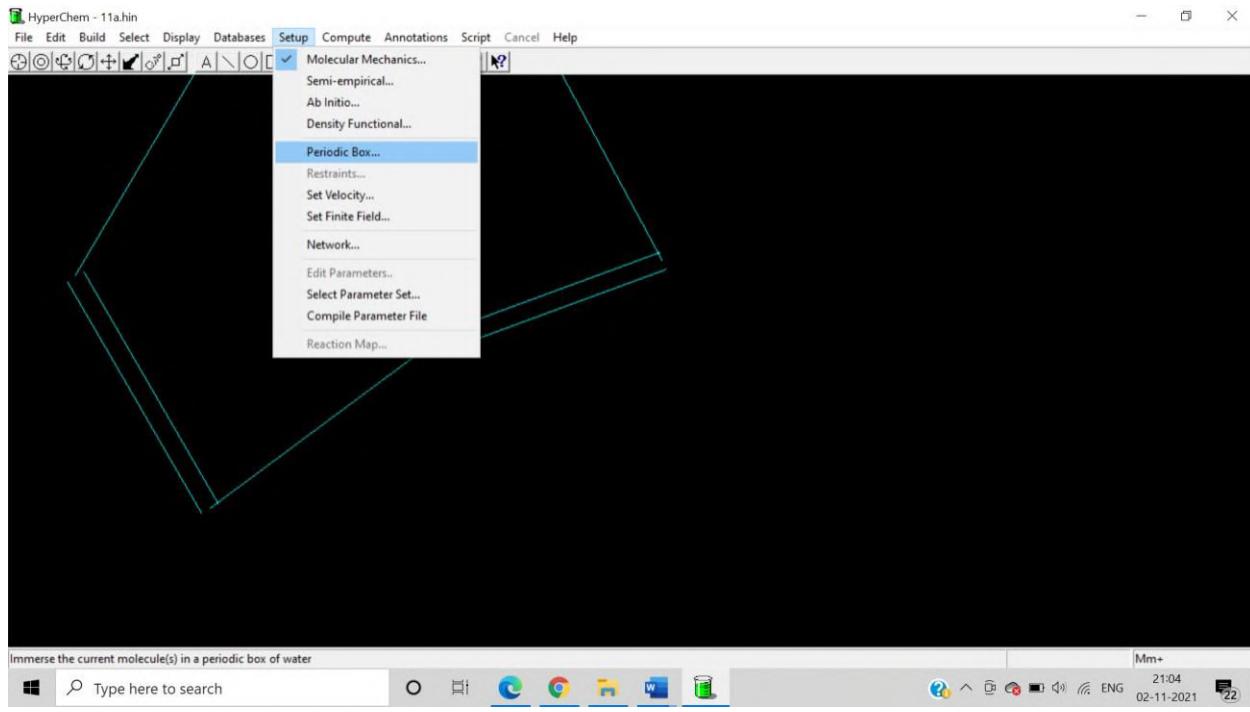


Energy: 130.600586

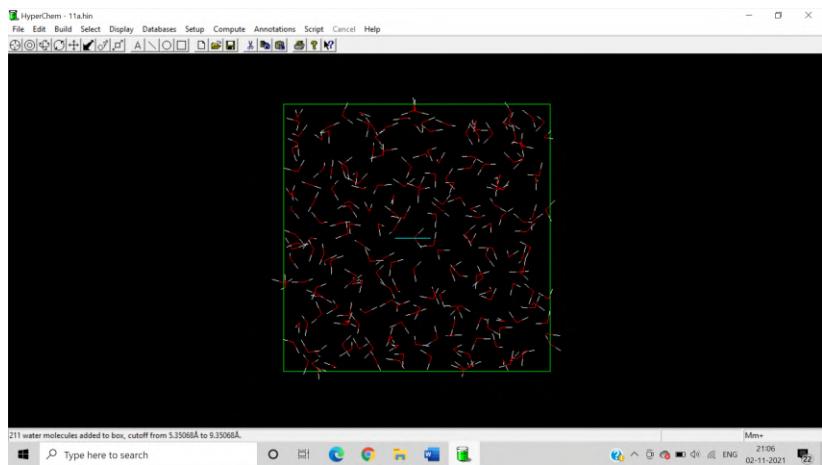
Gradient: 0.076460

Converged: Yes

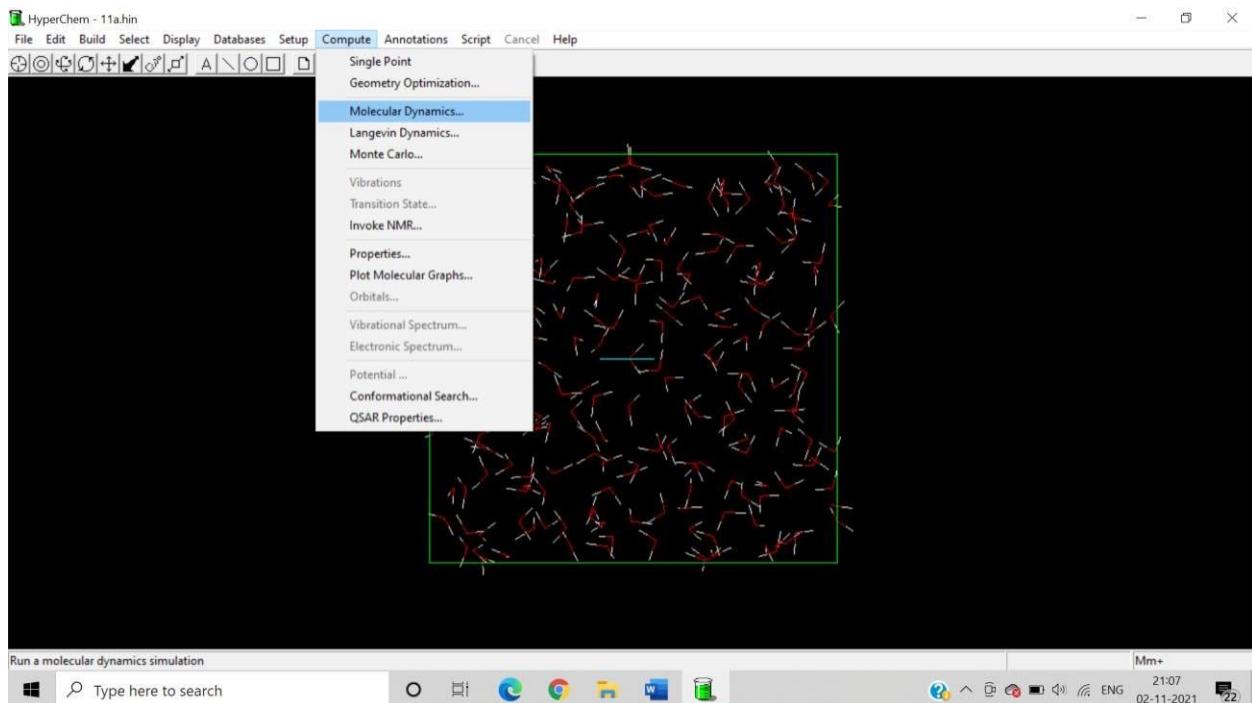
Setup: Periodic Box:

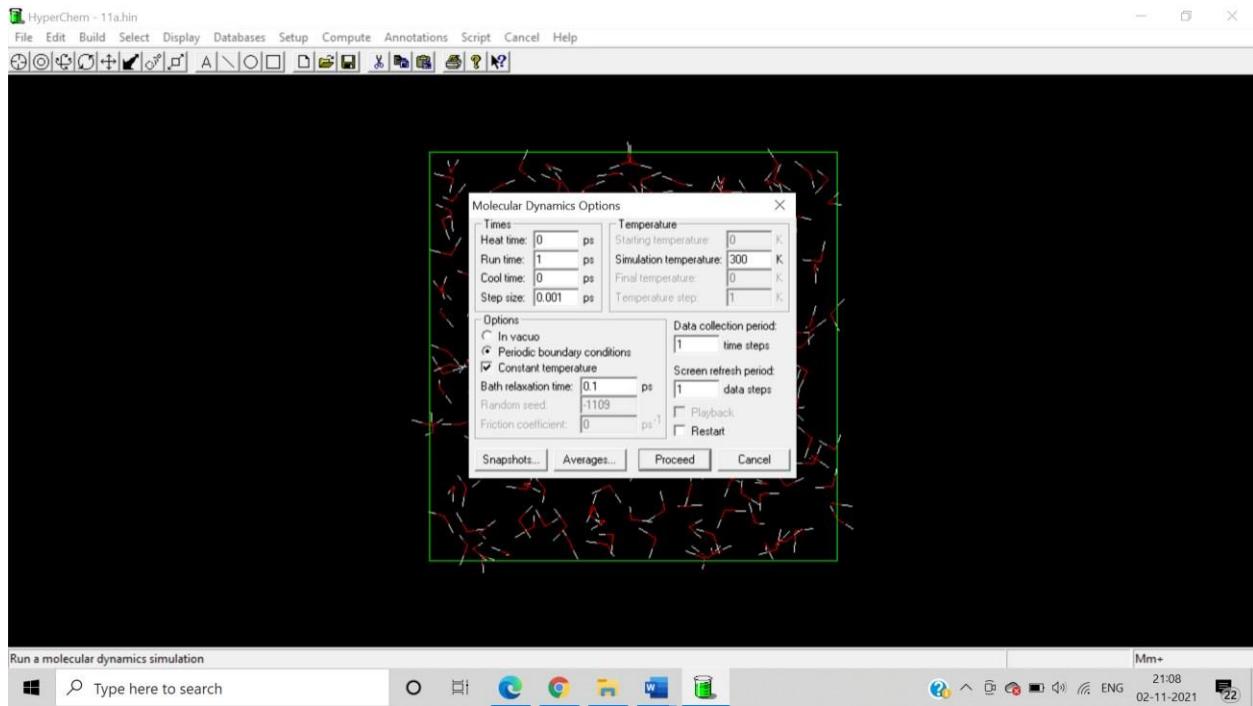


You will get 211 water molecules added to the box and cutoff from 5.35068 Armstrong to 9.35068 Armstrong



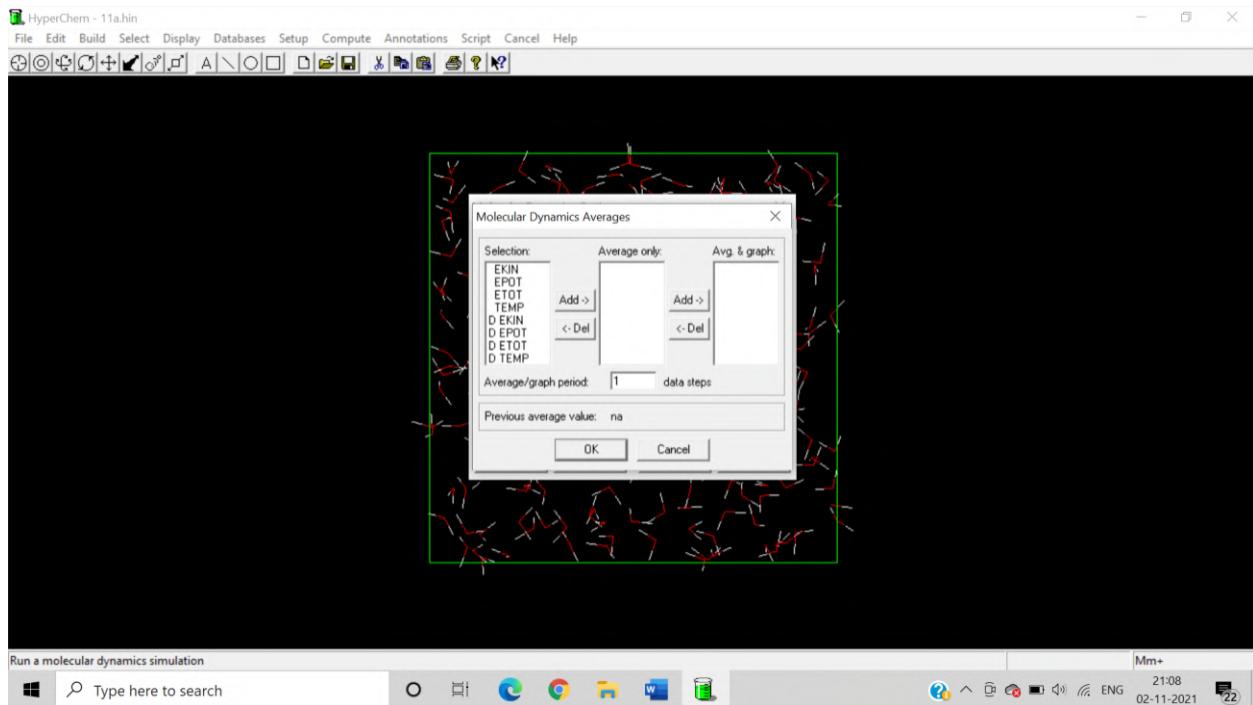
Molecular Dynamics:





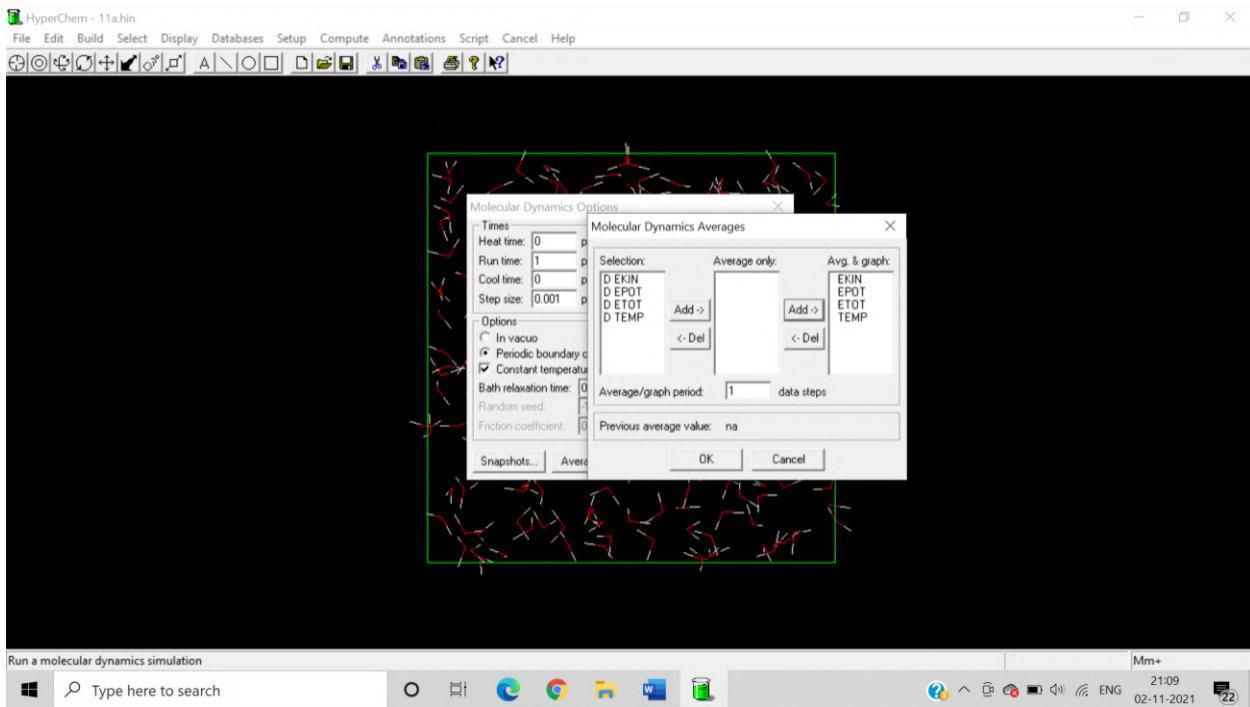
Average:

Copy the EKIN EPOT ETOT and TEMP till Avg and graph section to get the correct result:-



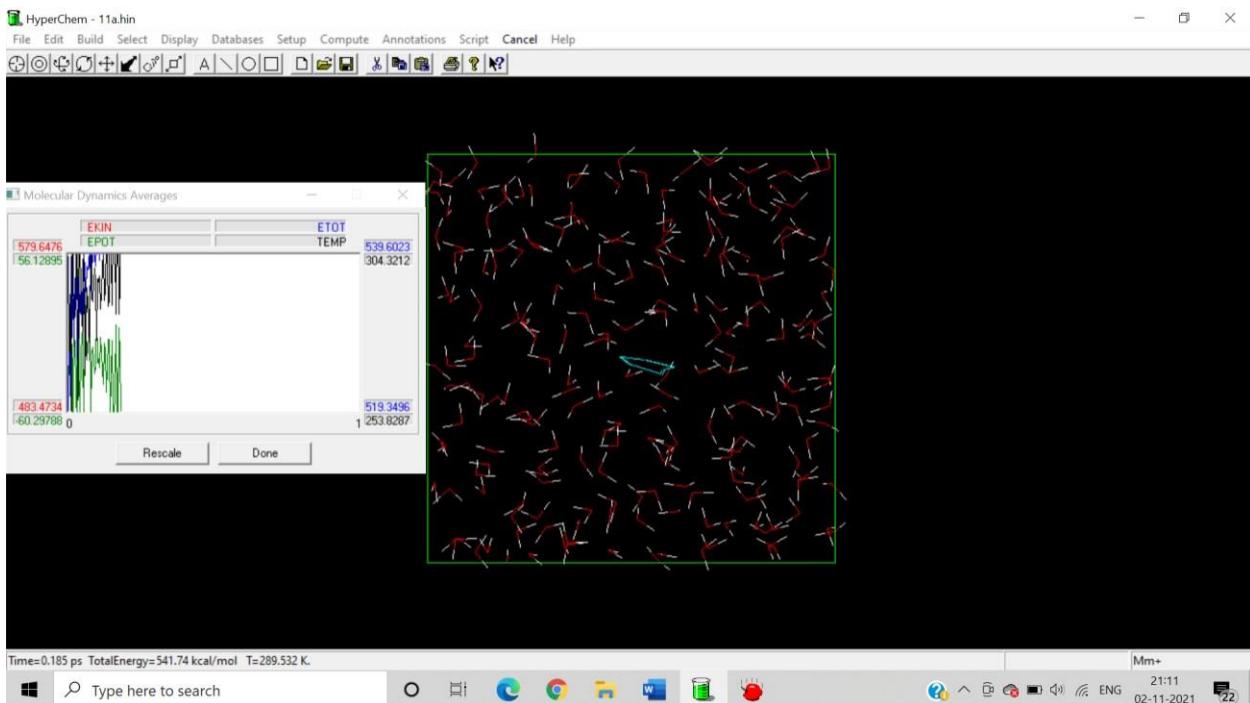
Average and Graph:

Copy the EKIN EPOT ETOT and TEMP till Avg and graph section

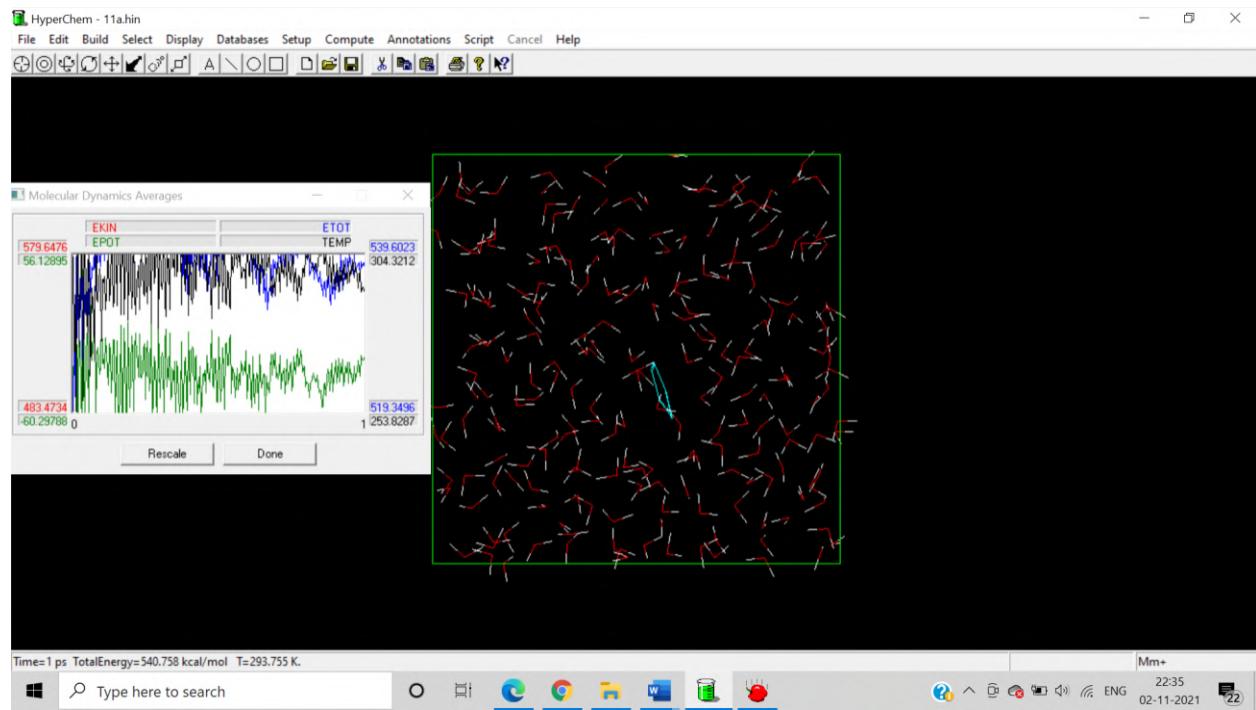


Graph:

Wait for 5 mins until the whole graph is obtained:



Result:



So at the start the potential energy was 483.4734 and initial kinetic energy was -60.29788.

The final energy is 56.12895 and 579.6476.

The screenshot shows an Excel spreadsheet titled '11a.csv - Excel'. The table contains data from a molecular dynamics simulation, with columns labeled 'time', 'EKIN', 'EPOT', 'ETOT', and 'TEMP'. The data starts at time 0 with values 579.6476, 56.12895, 539.6023, and 304.3212 respectively, and continues for 18 time steps up to 0.016.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	time	EKIN	EPOT	ETOT	TEMP														
2	0	579.6476	-60.29788	519.3497	304.3212														
3	0.001	568.2559	-42.3682	525.8877	298.3404														
4	0.002	520.5812	7.958268	528.5394	273.3107														
5	0.003	485.3008	54.0306	539.3314	254.7881														
6	0.004	501.6285	31.75614	533.3846	263.3603														
7	0.005	544.8705	-16.9133	527.9572	286.0628														
8	0.006	555.9994	-34.7506	521.2488	291.9056														
9	0.007	517.6115	14.39759	532.009	271.7515														
10	0.008	477.9618	61.22085	539.1826	250.935														
11	0.009	483.4734	56.12895	539.6023	253.8287														
12	0.01	530.9285	3.188899	534.1174	278.7431														
13	0.011	582.7742	-52.804	529.9683	305.9627														
14	0.012	604.8557	-75.952	528.9037	317.5557														
15	0.013	586.7139	-57.1296	529.5842	308.031														
16	0.014	553.0337	-11.515	541.5187	290.3486														
17	0.015	551.324	-12.3423	538.9817	289.451														
18	0.016	586.0345	-52.4728	533.5617	307.6743														

Result:

Go to 2D charts and using chart design draft a graph appropriately for the same:

Where:

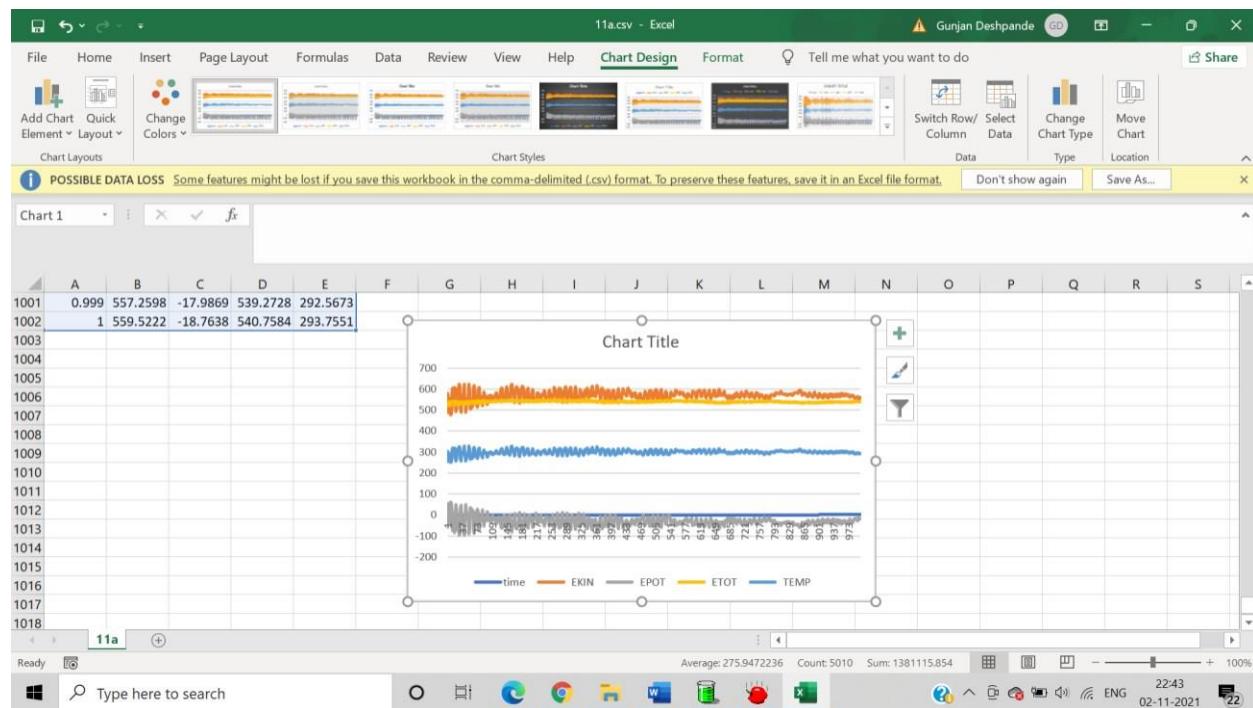
Dark Blue = time

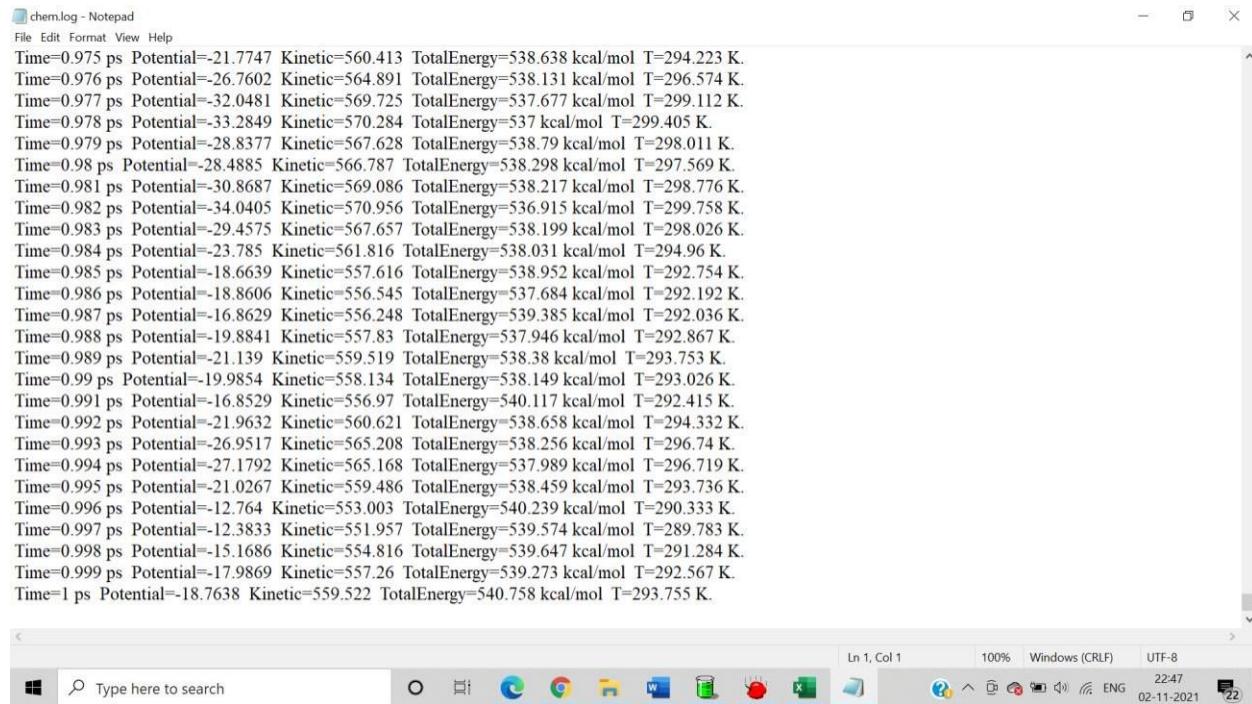
Orange= EKIN

Grey = EPOT

Yellow = ETOT

Light blue =TEMP





```
chem.log - Notepad
File Edit Format View Help
Time=0.975 ps Potential=-21.7747 Kinetic=560.413 TotalEnergy=538.638 kcal/mol T=294.223 K.
Time=0.976 ps Potential=-26.7602 Kinetic=564.891 TotalEnergy=538.131 kcal/mol T=296.574 K.
Time=0.977 ps Potential=-32.0481 Kinetic=569.725 TotalEnergy=537.677 kcal/mol T=299.112 K.
Time=0.978 ps Potential=-33.2849 Kinetic=570.284 TotalEnergy=537 kcal/mol T=299.405 K.
Time=0.979 ps Potential=-28.8377 Kinetic=567.628 TotalEnergy=538.79 kcal/mol T=298.011 K.
Time=0.98 ps Potential=-28.4885 Kinetic=566.787 TotalEnergy=538.298 kcal/mol T=297.569 K.
Time=0.981 ps Potential=-30.8687 Kinetic=569.086 TotalEnergy=538.217 kcal/mol T=298.776 K.
Time=0.982 ps Potential=-34.0405 Kinetic=570.956 TotalEnergy=536.915 kcal/mol T=299.758 K.
Time=0.983 ps Potential=-29.4575 Kinetic=567.657 TotalEnergy=538.199 kcal/mol T=298.026 K.
Time=0.984 ps Potential=-23.785 Kinetic=561.816 TotalEnergy=538.031 kcal/mol T=294.96 K.
Time=0.985 ps Potential=-18.6639 Kinetic=557.616 TotalEnergy=538.952 kcal/mol T=292.754 K.
Time=0.986 ps Potential=-18.8606 Kinetic=556.545 TotalEnergy=537.684 kcal/mol T=292.192 K.
Time=0.987 ps Potential=-16.8629 Kinetic=556.248 TotalEnergy=539.385 kcal/mol T=292.036 K.
Time=0.988 ps Potential=-19.8841 Kinetic=557.83 TotalEnergy=537.946 kcal/mol T=292.867 K.
Time=0.989 ps Potential=-21.139 Kinetic=559.519 TotalEnergy=538.38 kcal/mol T=293.753 K.
Time=0.99 ps Potential=-19.9854 Kinetic=558.134 TotalEnergy=538.149 kcal/mol T=293.026 K.
Time=0.991 ps Potential=-16.8529 Kinetic=556.97 TotalEnergy=540.117 kcal/mol T=292.415 K.
Time=0.992 ps Potential=-21.9632 Kinetic=560.621 TotalEnergy=538.658 kcal/mol T=294.332 K.
Time=0.993 ps Potential=-26.9517 Kinetic=565.208 TotalEnergy=538.256 kcal/mol T=296.74 K.
Time=0.994 ps Potential=-27.1792 Kinetic=565.168 TotalEnergy=537.989 kcal/mol T=296.719 K.
Time=0.995 ps Potential=-21.0267 Kinetic=559.486 TotalEnergy=538.459 kcal/mol T=293.736 K.
Time=0.996 ps Potential=-12.764 Kinetic=553.003 TotalEnergy=540.239 kcal/mol T=290.333 K.
Time=0.997 ps Potential=-12.3833 Kinetic=551.957 TotalEnergy=539.574 kcal/mol T=289.783 K.
Time=0.998 ps Potential=-15.1686 Kinetic=554.816 TotalEnergy=539.647 kcal/mol T=291.284 K.
Time=0.999 ps Potential=-17.9869 Kinetic=557.26 TotalEnergy=539.273 kcal/mol T=292.567 K.
Time=1 ps Potential=-18.7638 Kinetic=559.522 TotalEnergy=540.758 kcal/mol T=293.755 K.
```

Conclusion: Time=1 ps Potential=-18.7638 Kinetic=559.522 TotalEnergy=540.758 kcal/mol T=293.755 K.

Practical No: 11B

AIM: Study of peptide wild and mutant.

THEORY:

Hypercube, Inc. is a scientific software company, incorporated in 1985, specializing in molecular modeling software. HyperChem is, noted for its ease of use, extensive functionality, and modest price. Our most important platform is Microsoft Windows; HyperChem, on a PC under Windows, has the largest number of installations of any full-featured molecular modeling program.

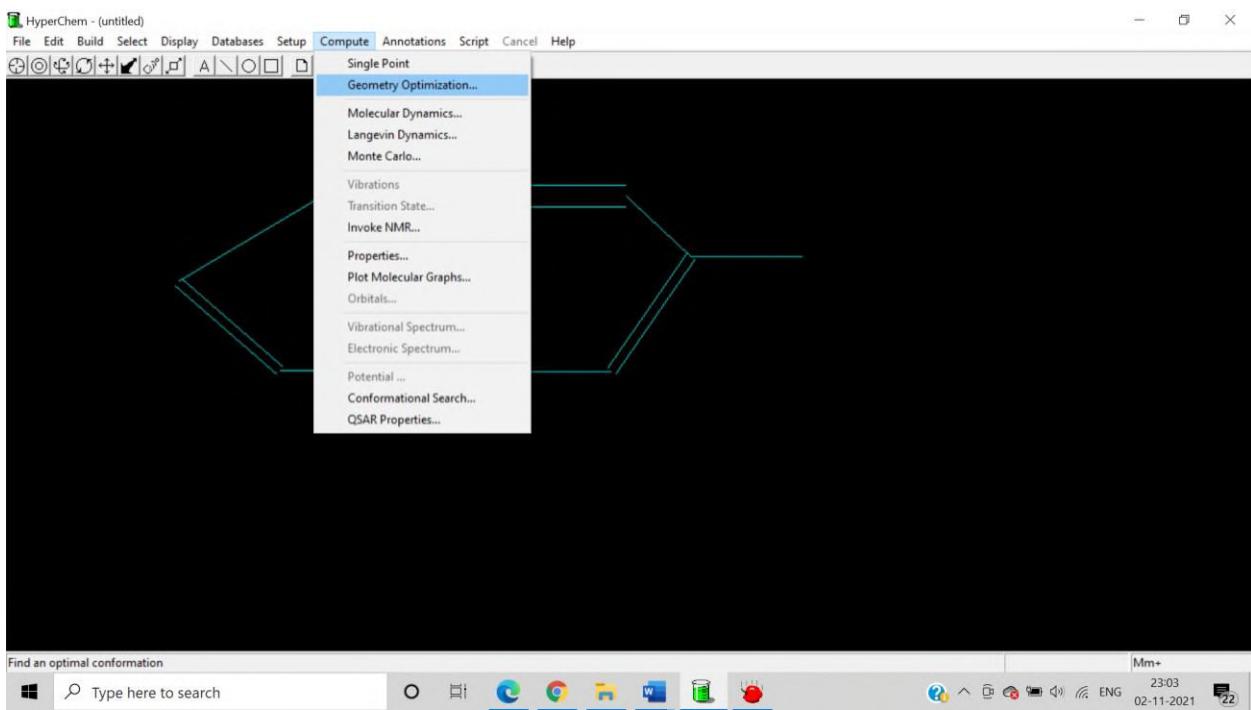
Hypercube, Inc. produces two versions of the core HyperChem product: HyperChem 8.0 and HyperChem for MAC. HyperChem 8.0 includes the Chemist's Developer Kit, an advanced customization tool; HyperNMR, for a priori simulation of NMR spectra; and HyperChem Data, a chemical database program with over 10,000 molecules included.

Hypercube, Inc. has set new standards for ease of use and molecular modeling power on PC-based systems. The goal is to bring molecular modeling to all chemists and chemistry students.

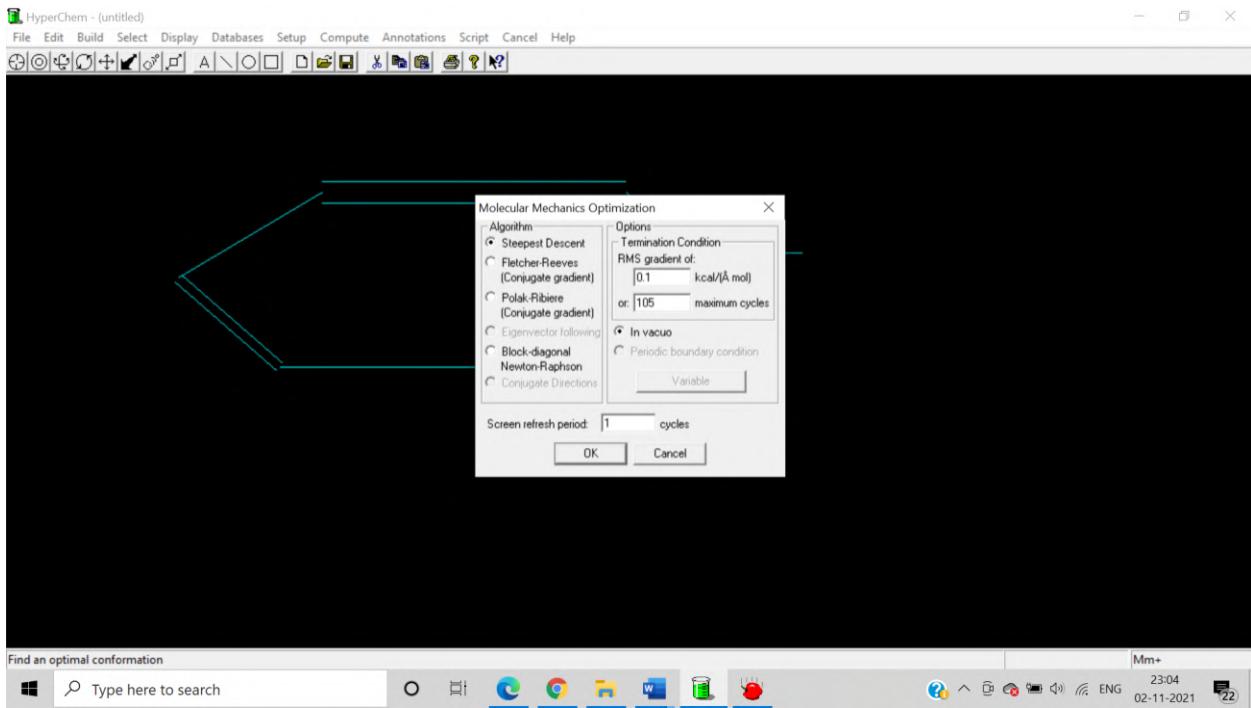
METHODOLOGY

1. Start log. Click on select and residues. Click on display and labels and click on name in the residue box. This will help us distinguish one residue from the other as it will show the name.
2. Click on databases and amino acids. Switch to alpha helix and click on amino acids to create a peptide (ala, gly, ser, thr, leu, ile).
3. Click on single point through compute to find the energy and gradient of the peptide created.
4. Select an amino acid on which mutation has to be performed (ser). Go to databases and mutate (phe).
5. Deselect everything by right clicking on the blank, black space. Perform geometry optimization and change the maximum cycle to 2000.
6. Open the text document to see the log.

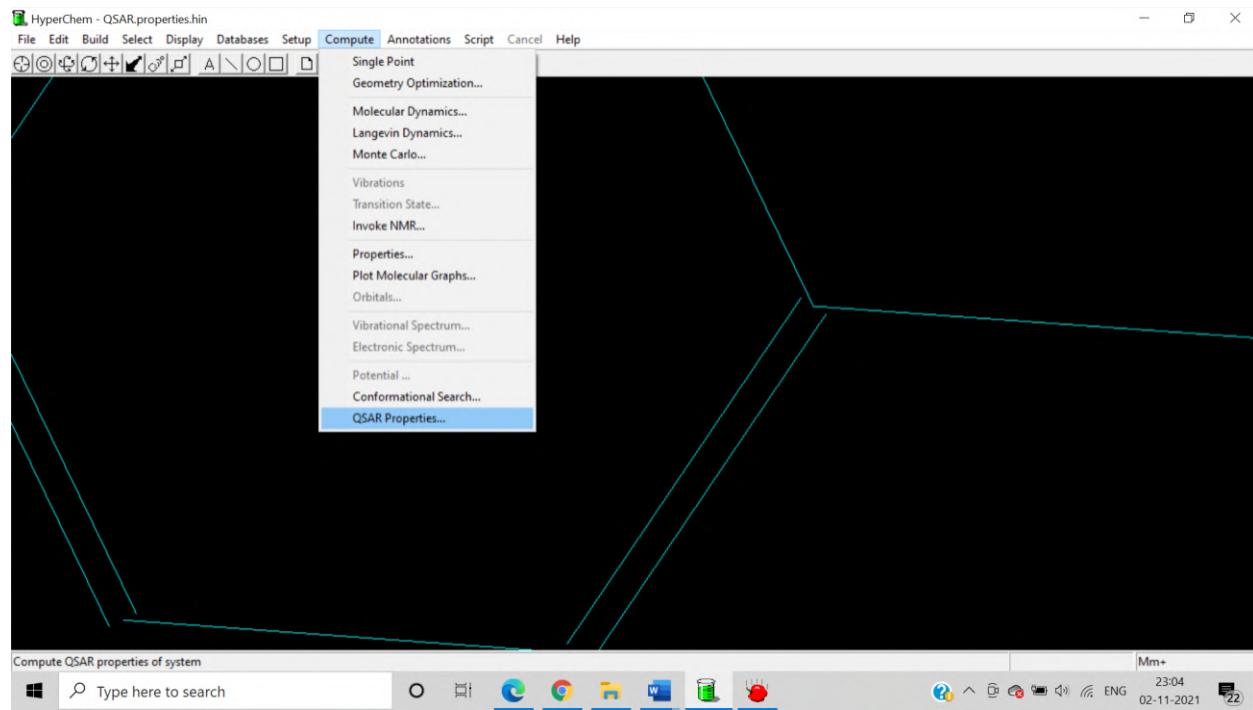
First do the Geometry Optimization: of the structure :-



Steepest Descent:-

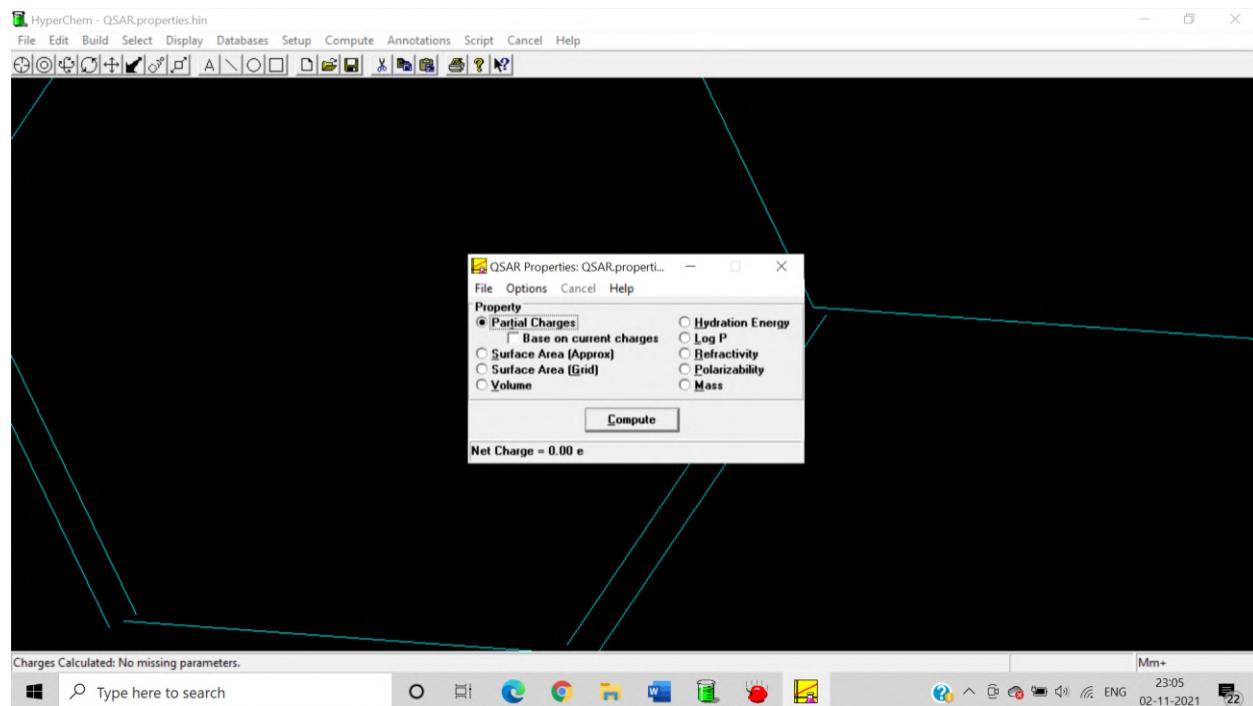


Compute:- To QSAR Properties:-



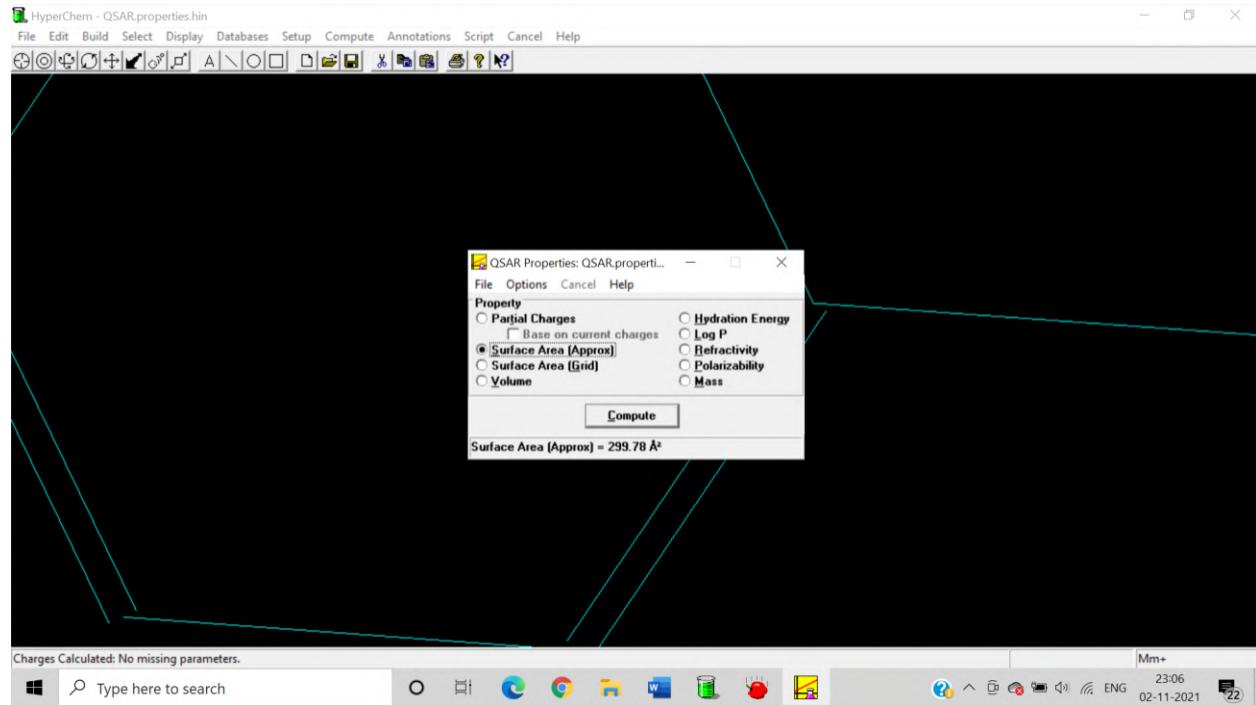
Now compute the Partial Charges:-

Net charge = 0.00e



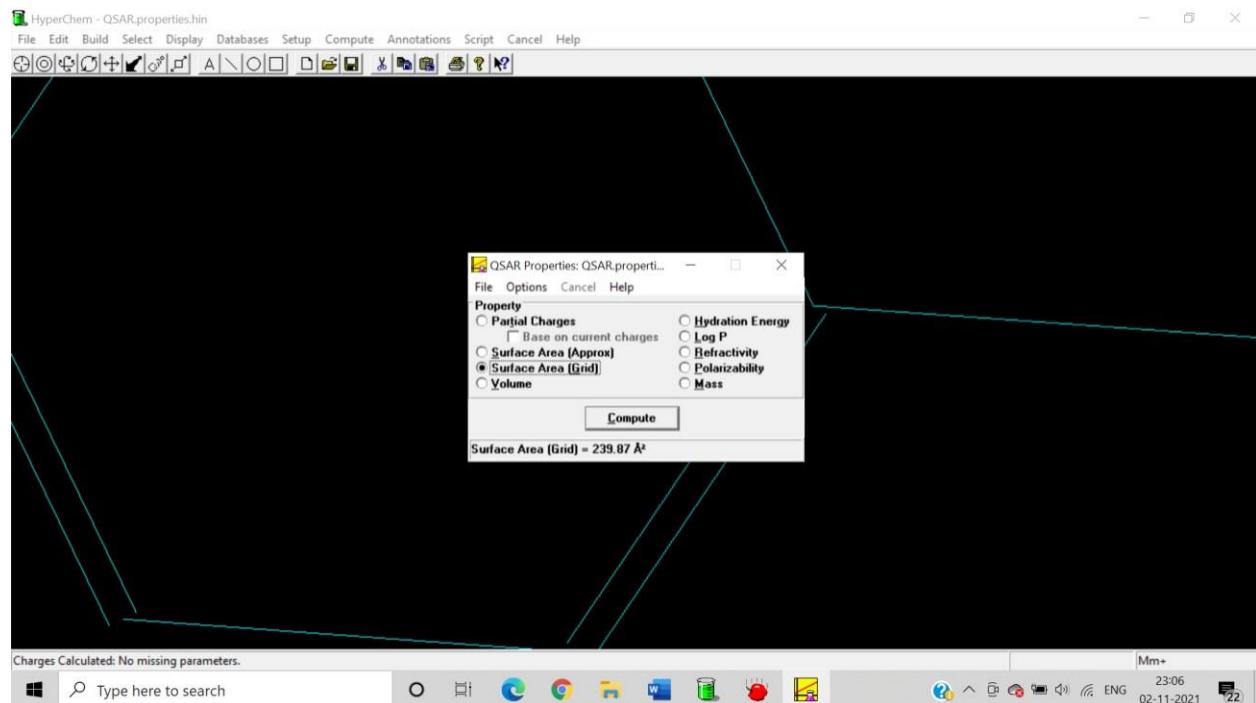
Now compute the Surface Area (Approx) –

Surface Area (Approx)=299.78 A³



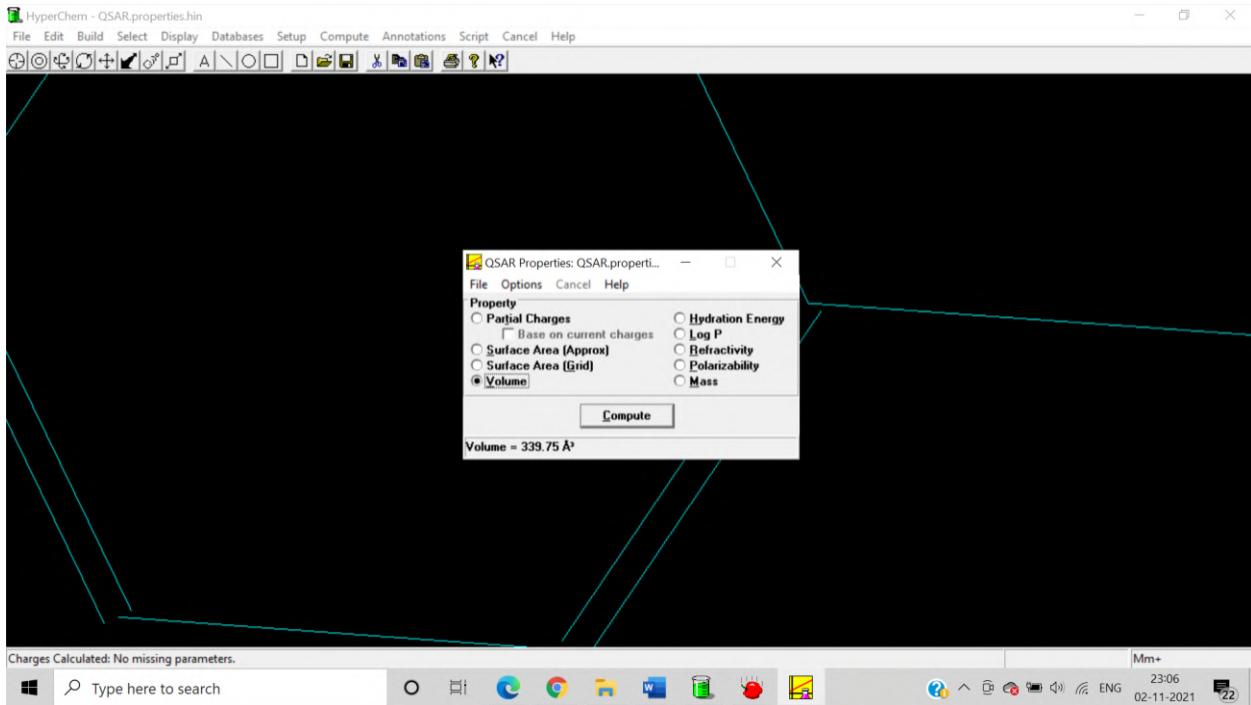
Now compute the Surface Area (Grid)-

Surface Area (Grid)-239.87A³

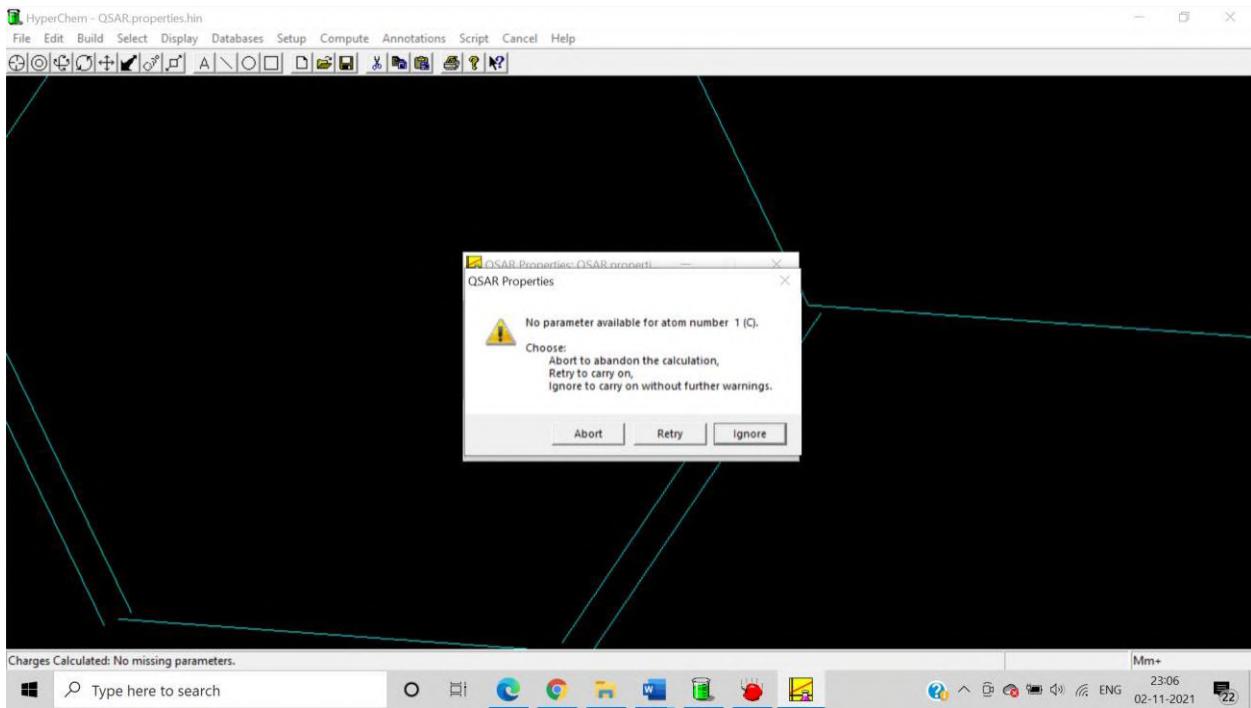


Now compute the Volume:-

Volume = 339.75 \AA^3

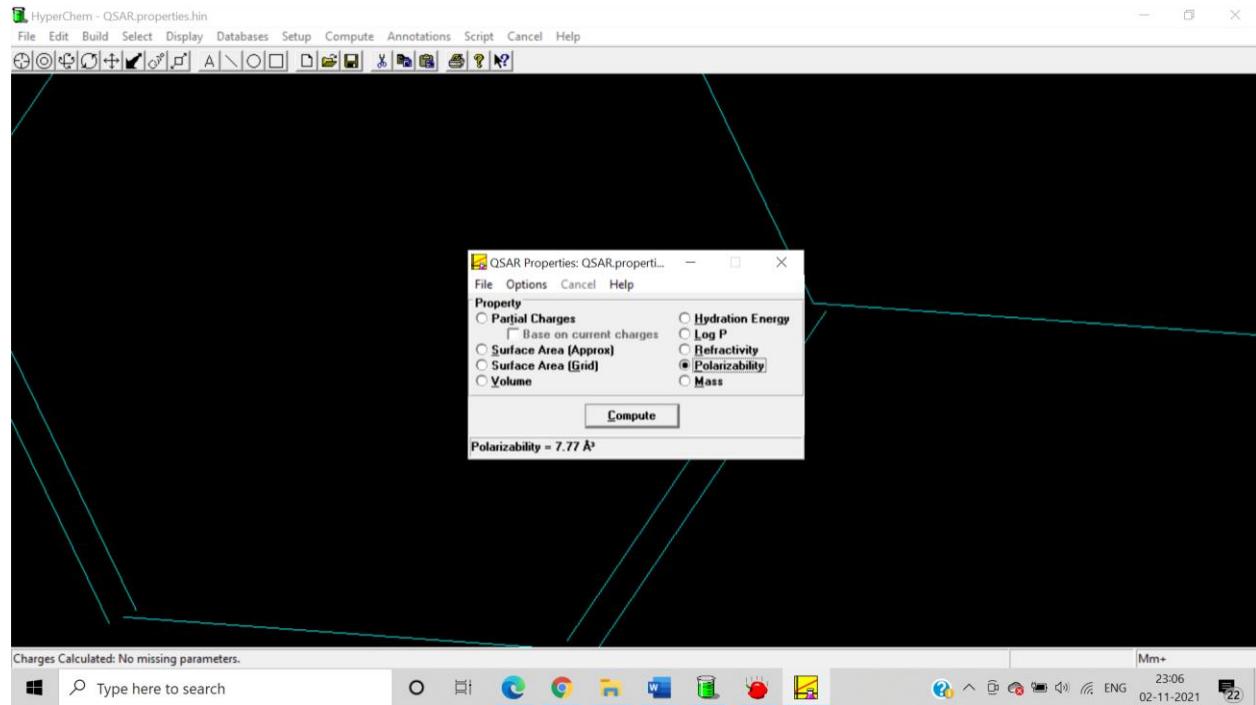


Now compute the Hydration energy:-



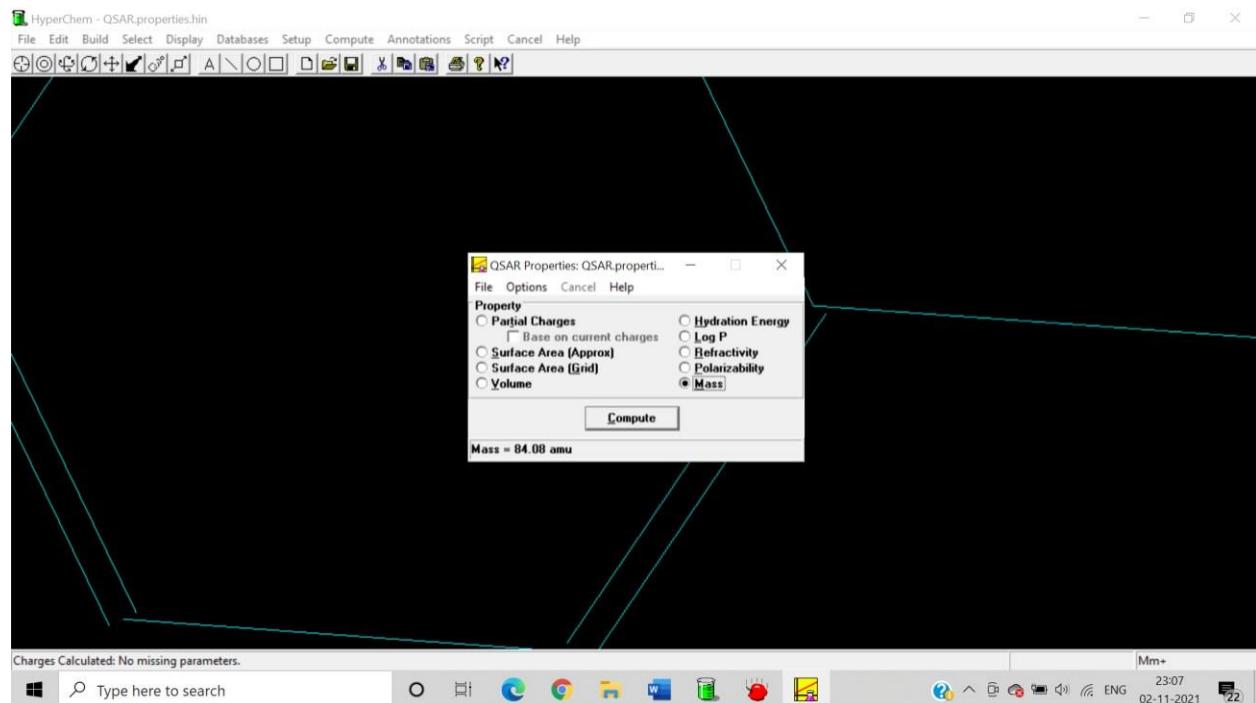
Now compute the Polarizability:-

Polarizability=7.77A³

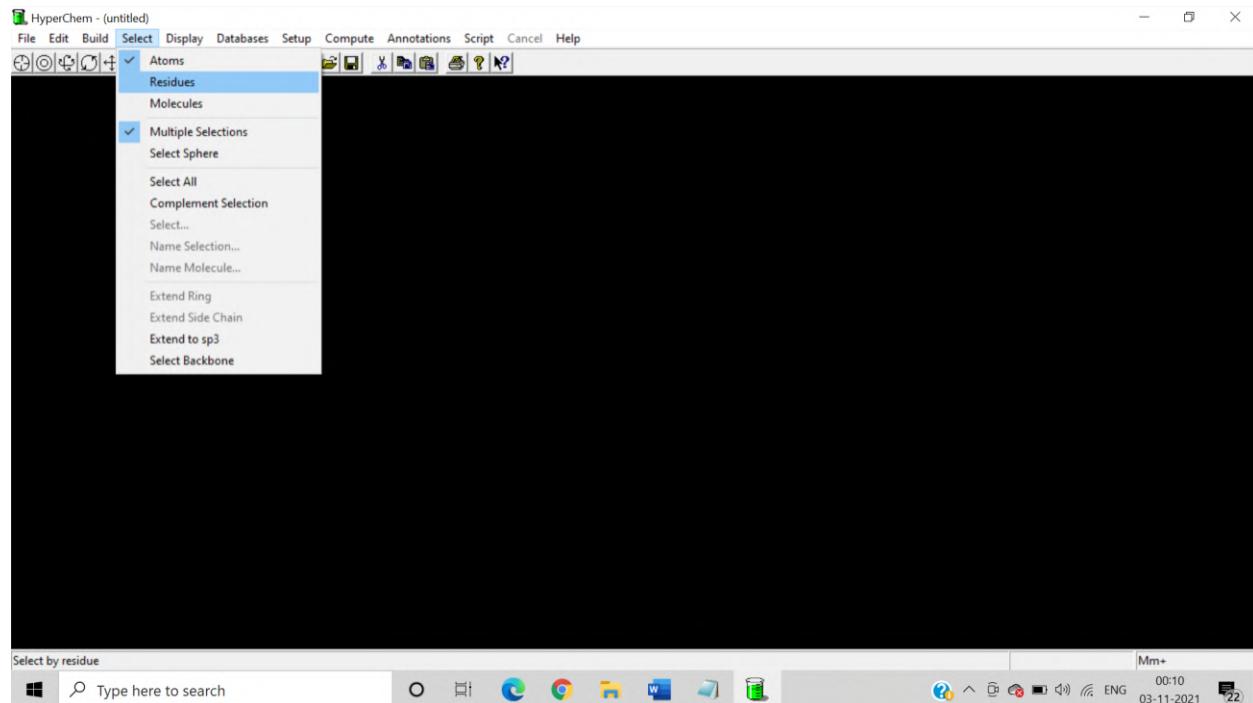


Now compute the mass:-

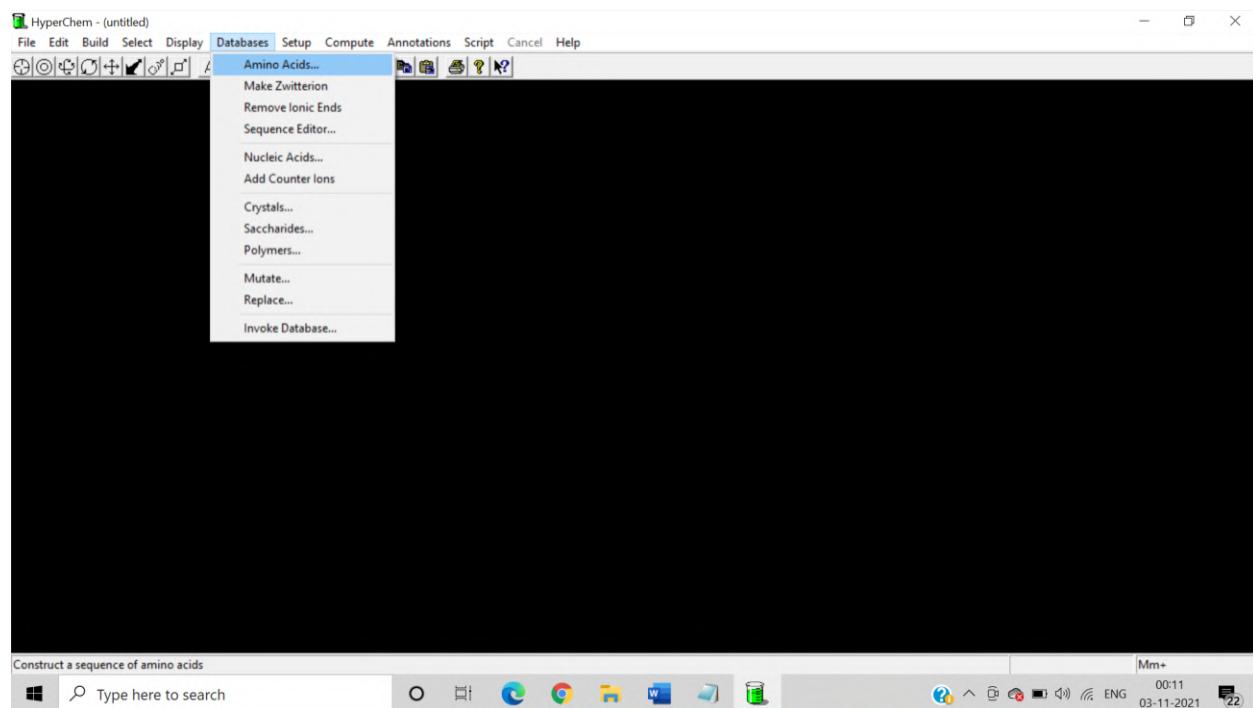
Mass=84.08 amu

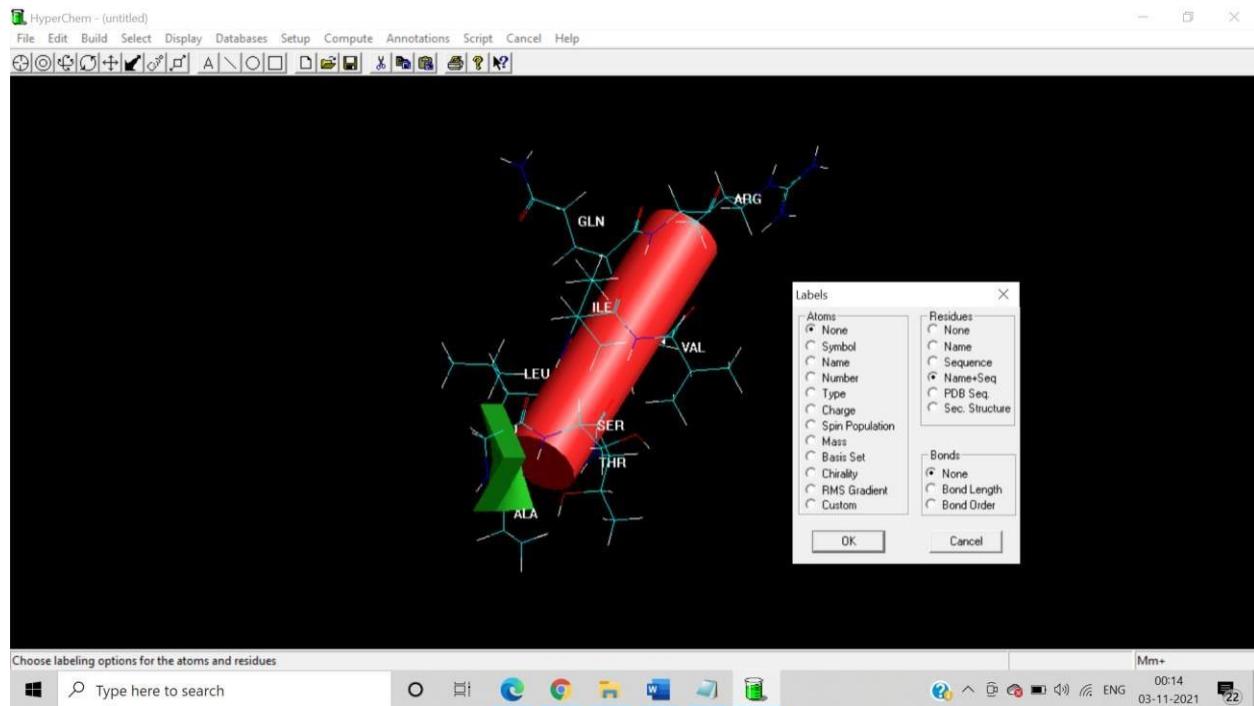
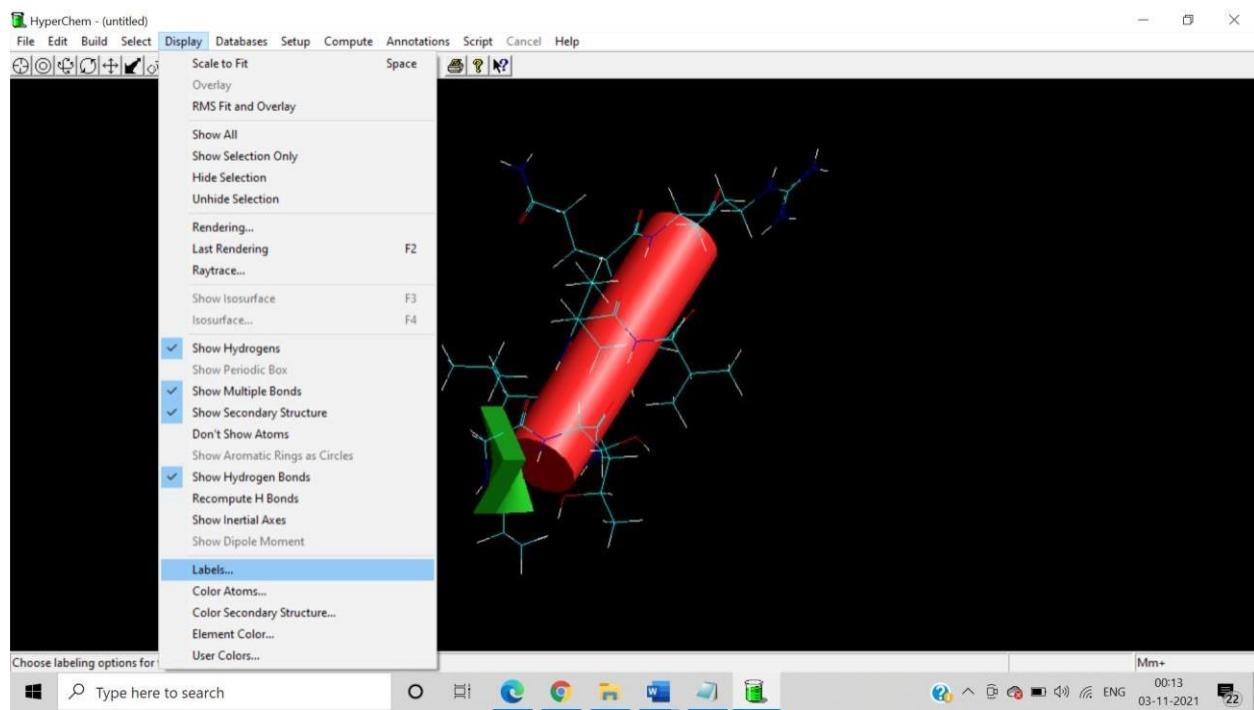


Working for PEPTIDE-



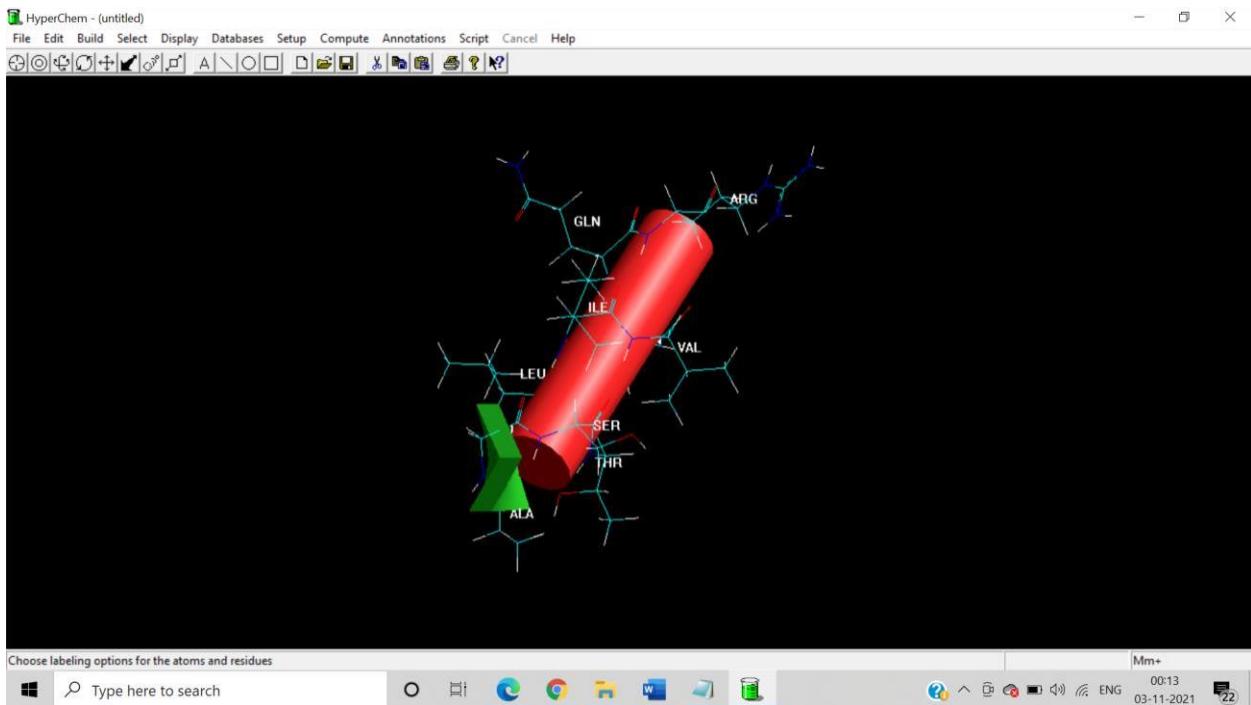
First construct a sequence of amino acids:



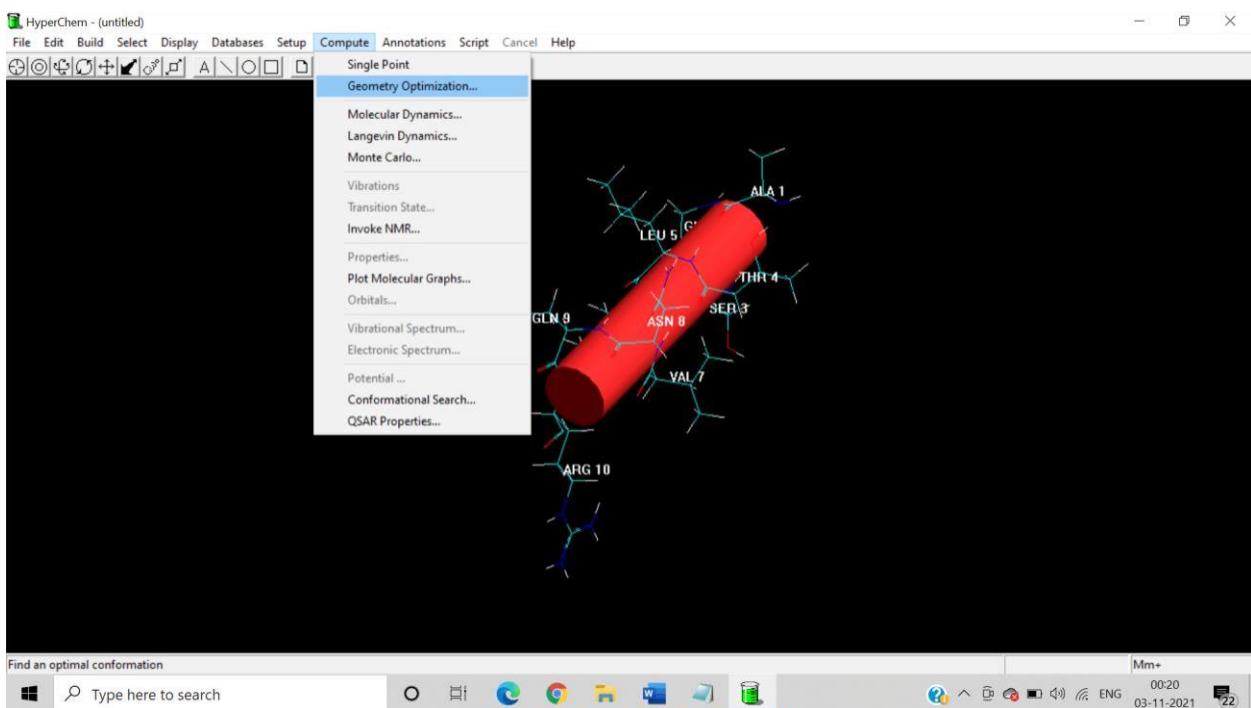


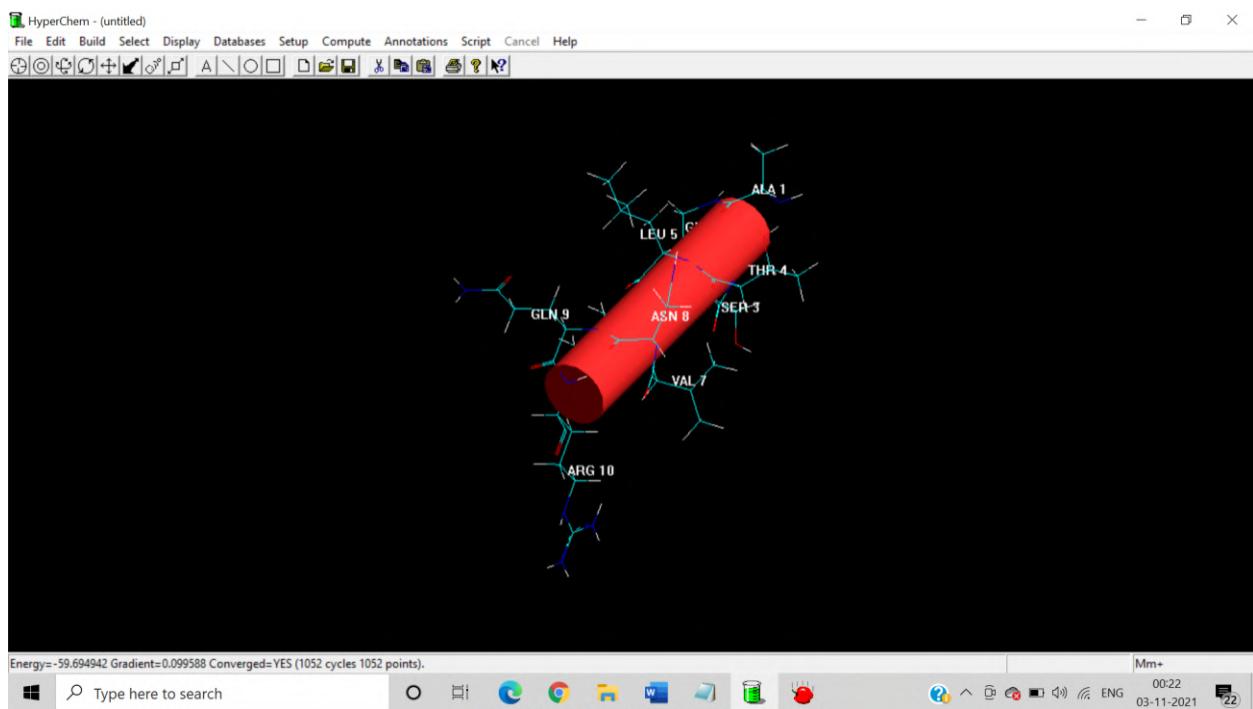
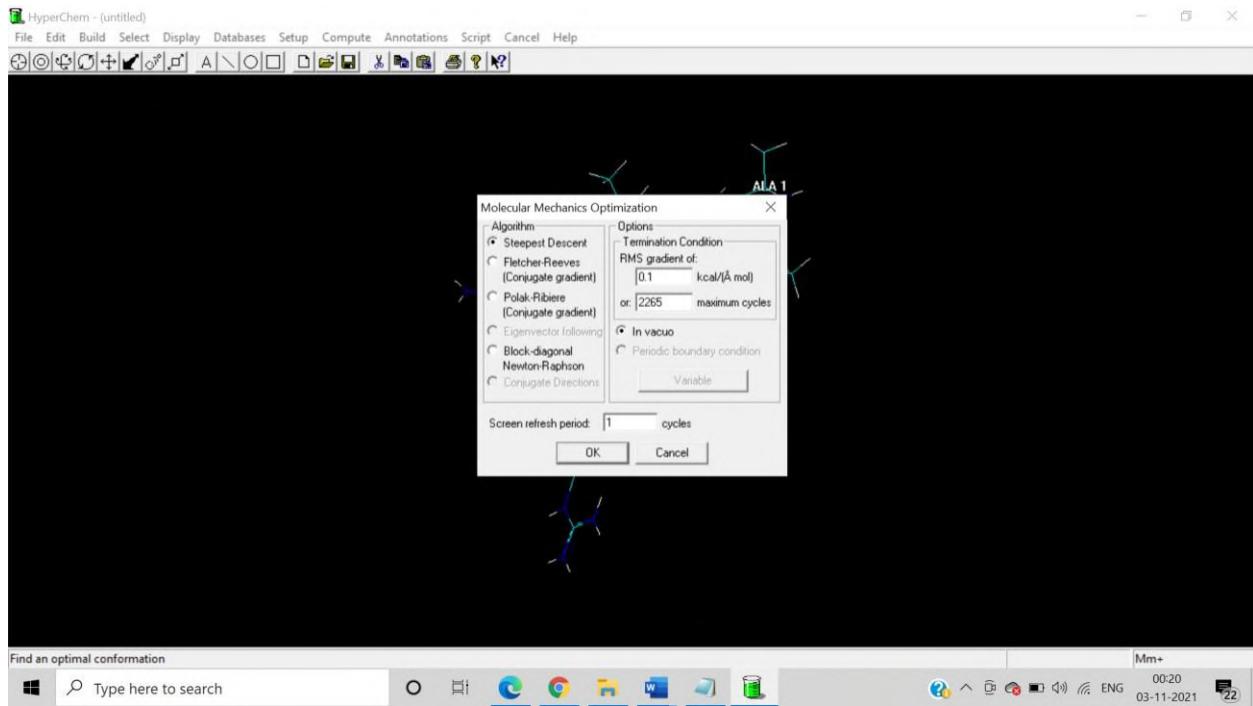
Result:

Label and give number to all 10 aa :



Geometry Optimization of all 10 aa:





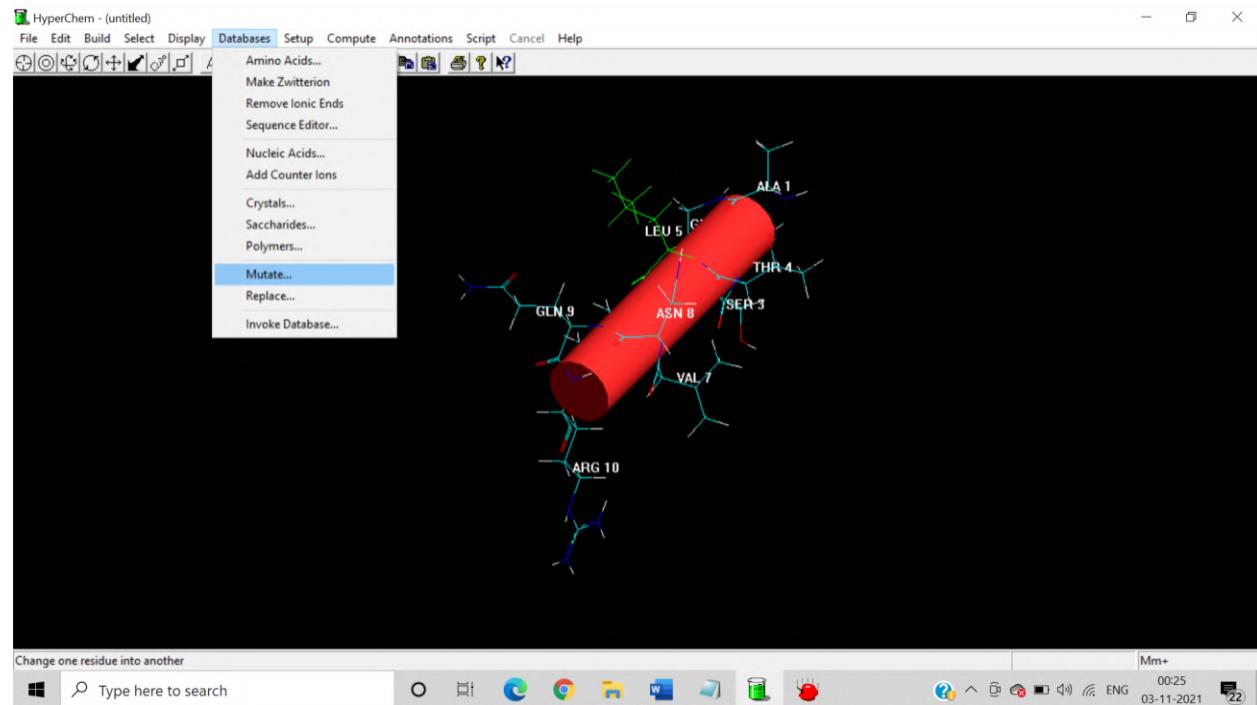
Energy :- 59.69494942

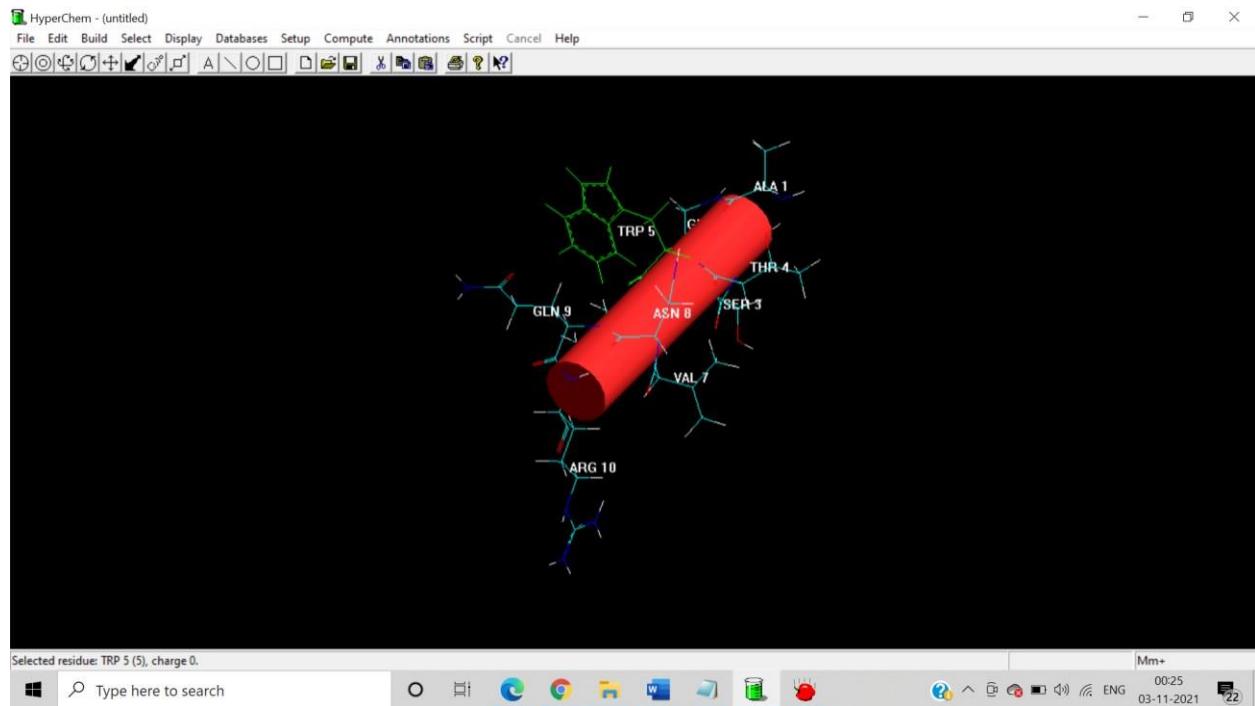
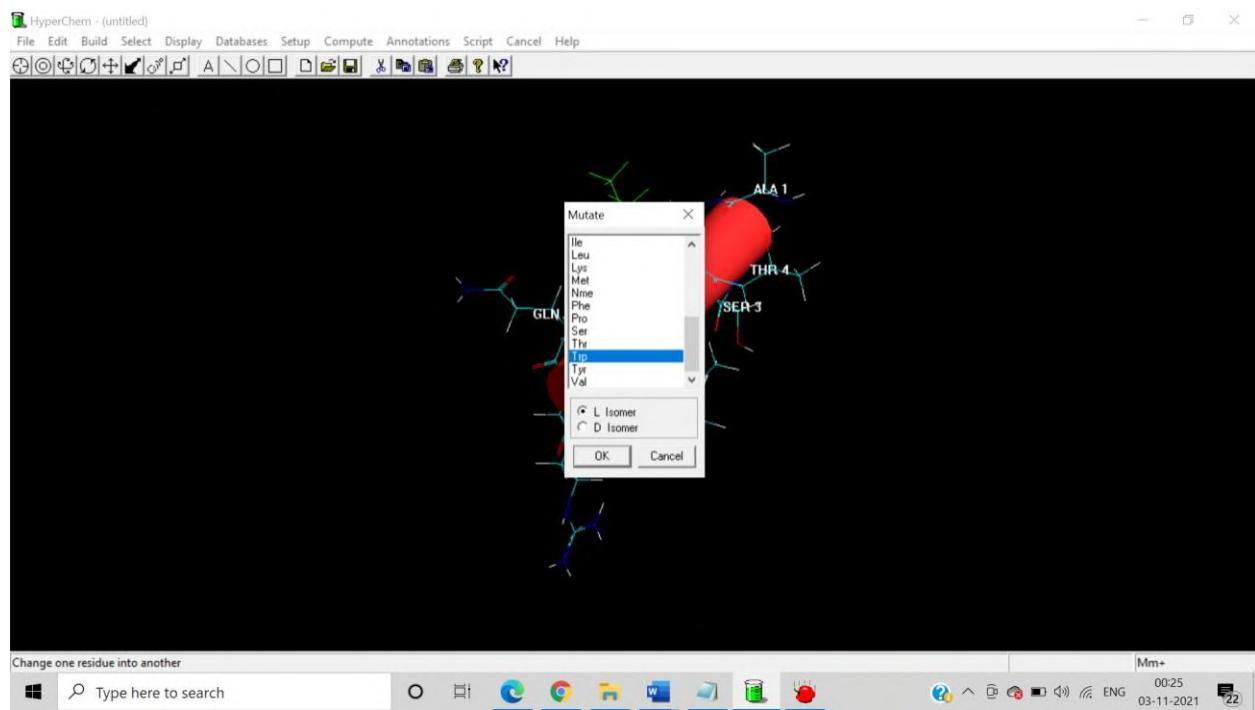
Graident:- 0.099588

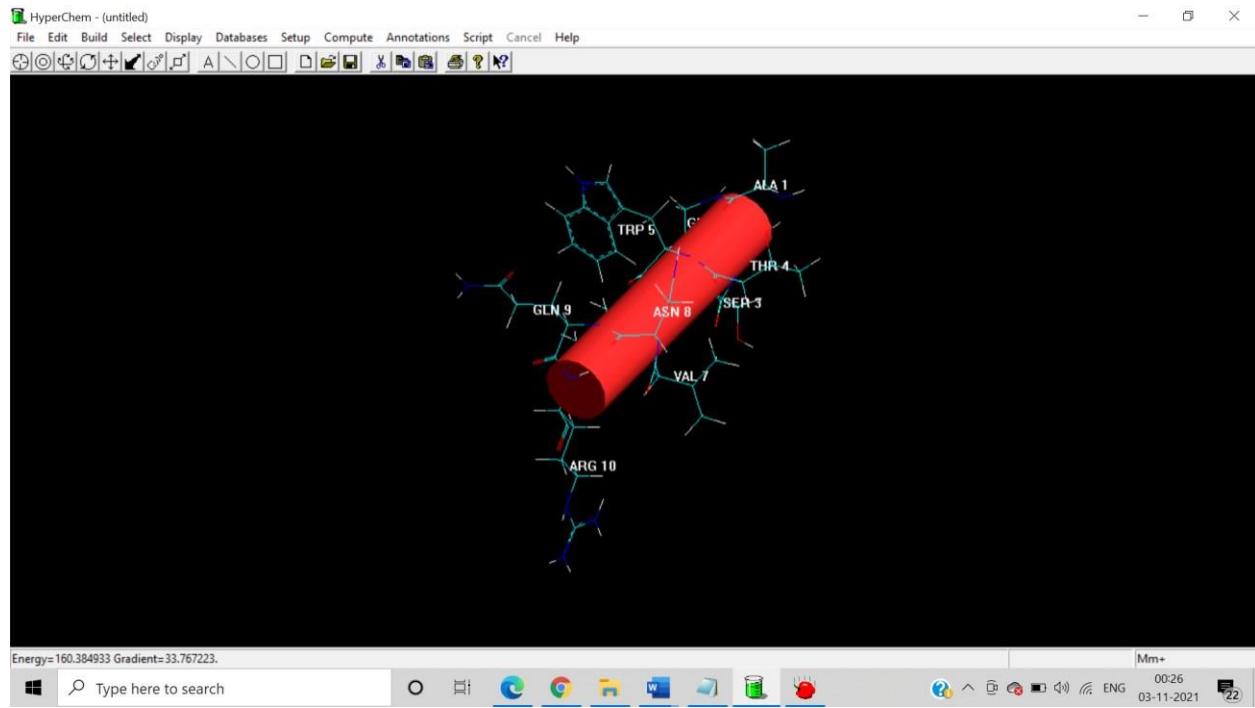
Converged: Yes

Cycles:-1052

Points:-1052



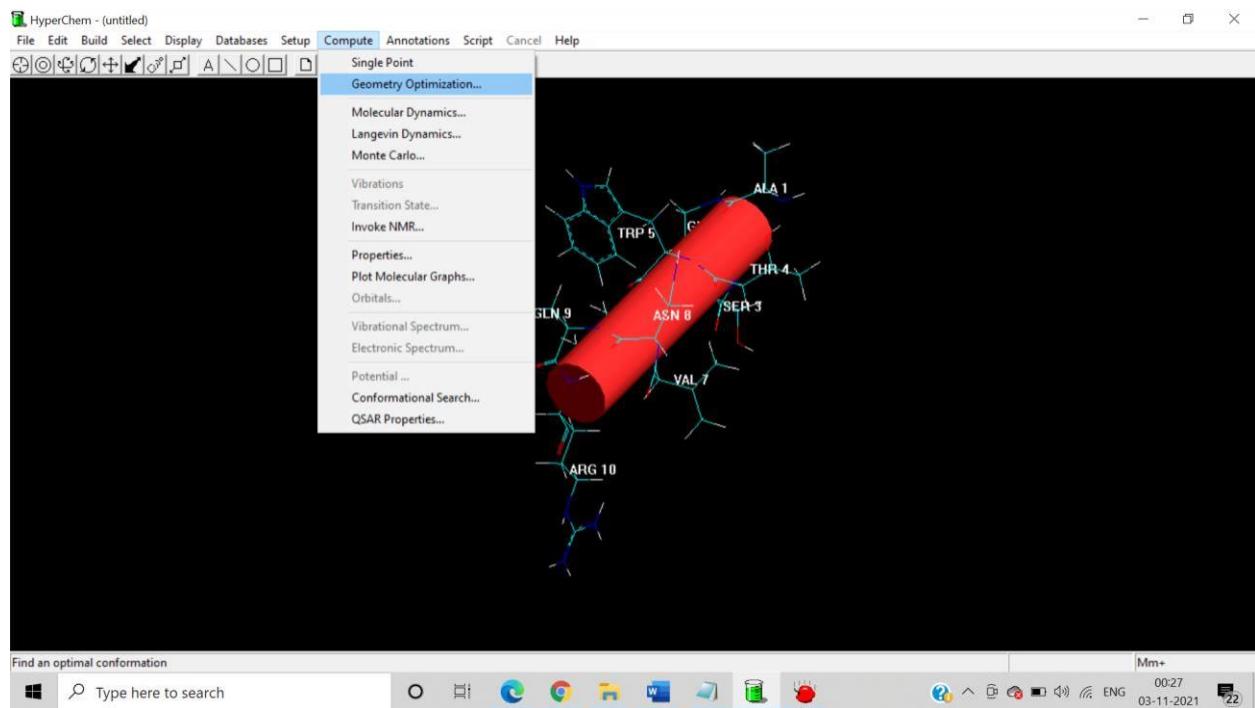




Now the single point energy is: 160.384933

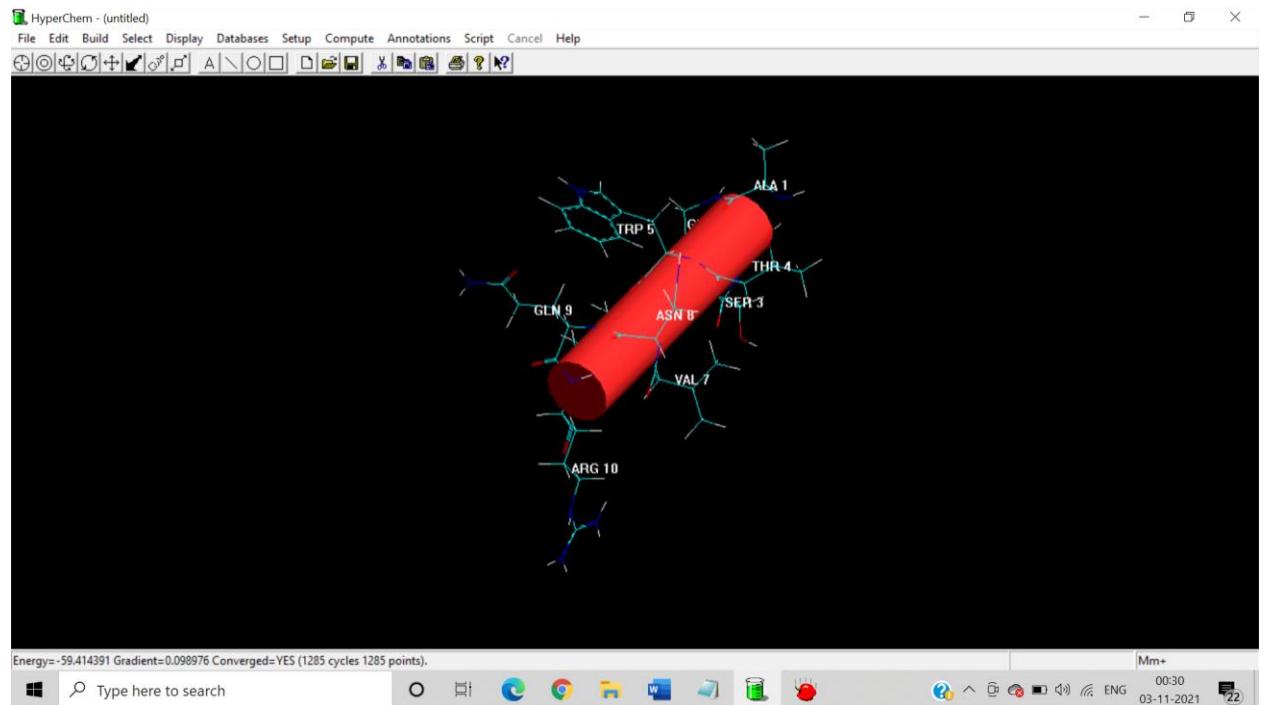
Gradient:-33.767223

Now go for Geometry Optimization:



Result:

After GO and Mutation:



Conclusion:

The final energy = 59.414391

Gradient:- 0.098976

Converged: Yes