

DATE: 12/08/2024

PRACTICAL: 1

High Performance Thin Layer Chromatography (HPTLC)

AIM:

To perform qualitative and quantitative analysis of Atorvastatin using High Performance Thin Layer Chromatography (HPTLC).

THEORY:

High Performance Thin Layer Chromatography (HPTLC) is an enhanced or automated form of thin layer chromatography (TLC). Enhancement has been made to the TLC technique to automate the different steps to increase resolution and to allow more accurate quantitation. It is a quality assessment tool used for the evaluation of pharmaceutical and botanical materials.

The process involves a suitable adsorbent material like silica, bonded silica phases or alumina which is coated in the form of a thin film over a support material like aluminum or glass. This forms the stationary phase. Mobile phase is usually a single organic solvent or a mixture of two or more solvents depending upon the nature of the stationary phase and sample. It allows the analysis of a broad number of compounds both efficiently and cost effectively. Same analysis can be viewed using different wavelengths of light thereby providing a more complete profile of the plant or drug.

HPTLC remains one of the most flexible, reliable, and cost-efficient separation technique ideally suited for the analysis of botanicals and herbal drugs. Used with standardized procedures, it guarantees reproducible results, a vital element in the routine identification of complex fingerprints of plant extracts and pharmaceutical products.

Derivatization plays an important role in enabling the detection of separated compounds that are colorless and cannot be visualized with UV radiation or fluorescence. A suitable reagent is applied to the TLC plate, which reacts with the sample compounds and transforms them into detectable derivatives.

HPTLC consists of following instrument components:

1. LINOMAT-V Applicator:

With the Linomat-V applicator, samples are sprayed onto the chromatographic layer in the form of narrow bands with the help of N₂ Gas. This technique allows larger volumes to be applied than by contact transfer (spotting). During the spraying, the solvent of the sample evaporates almost entirely concentrating the sample into a narrow band of selectable length. Starting zones sprayed on as narrow bands ensure the highest resolution attainable with any given Thin-Layer Chromatographic system. For qualitative and quantitative HPTLC as well as for demanding preparative separation, spray-on application as bands is a necessity.



Fig 1: HPTLC Applicator CAMAG

2. TLC SCANNER- III:

The CAMAG TLC SCANNER III is used for the densitometric evaluation of objects from the fields of planar chromatography. Densitometry uses a light beam of selectable length to scan the tracks of the chromatogram. For quantifying the separated compounds, the signal from a chromatogram fraction is compared to the signal from the sample free plate background. It has the following features:

1. Fully automatic scan
2. Ultra-fast scan
3. Spectrum development
4. Beam / sample alignment check
5. 190 –800 nm range due to 3 lamps provided in the scanner (D2, W, Hg)
6. Spot check
7. Lamp use tracked



Fig 2: CAMAG TLC Scanner

3. CAMAG HPTLC Software: visionCATS

visionCATS stands for ease of use and intuitive simplicity. The software organizes the workflow of HPTLC, controls the involved CAMAG® HPTLC PRO Modules and CAMAG®

instruments, and manages data. The easy to navigate user interface effectively guides the user through the chromatographic process – from definition of samples and substances to reporting of results. Simply select one of the default methods and start working: fill in the sample table, select a developing solvent and the derivatization reagent.

If necessary, modify detection parameters. Then visionCATS will guide you. Creating your own method is easy as well: just select the desired steps. The sample-oriented approach allows for creating virtual plates from tracks originating from different plates, e.g. for batch-to-batch comparison or long-term stability testing. With visionCATS relevant samples can be located easier and faster than ever: a powerful search tool within the file explorer that includes extended preview functionalities enables the user to easily search for text and date, samples, methods, and analysis files. Custom search filters can also be created.

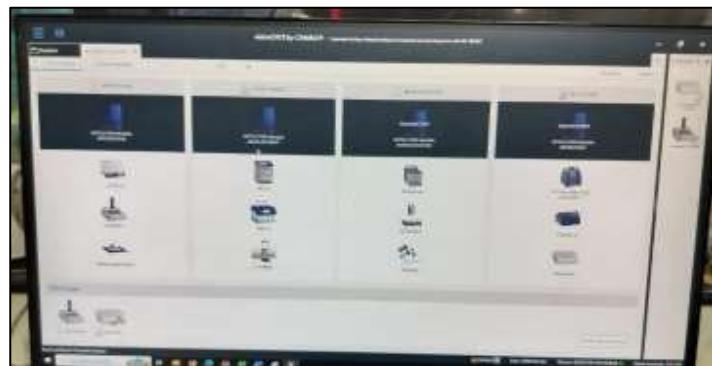


Fig 3: HPTLC visionCATS software

4. UV Cabinet

The UV Cabinet 4 consists of a UV Lamp 4 and the Viewing Box 4. The front of the Viewing Box 4 is closed with a roller shutter. Eyes are protected by the UV filter in the viewing window. The distance between UV lamp and object is optimally selected for homogeneous illumination and observation of the entire TLC plate through the viewing window.



Fig 4: UV Cabinet

Gas used in HPTLC:

1. Nitrogen Gas

When samples are applied to the TLC plate, nitrogen gas helps in minimizing the evaporation of solvents from the sample spot. This ensures that the sample remains stable and does not change its concentration before the chromatographic separation begins. Nitrogen gas can help maintain

a consistent flow and pressure during the sample application process, which is crucial for accurate and reproducible results. This ensures that the sample is applied evenly and precisely onto the TLC plate.

By controlling the application environment and preventing solvent evaporation, nitrogen gas helps in achieving sharper and more defined spots on the TLC plate, which improves the overall quality of the chromatographic separation.



Fig 5: Nitrogen gas (yellow)

PRINCIPLE:

HPTLC is a liquid-solid chromatographic technique which works on the principle that different compounds will have different solubilities and different degrees of adsorption to the two phases between which they are to be partitioned.

The stationary phase used is silica 60 F254 while the mobile phase varies according to the sample to be analyzed. The solute particles get adsorbed to the polar sites on the stationary phase due to the attractive forces between them. As the solvent rises on the plate by capillary action up through the adsorbent, differential adsorption occurs between the compounds of the mixture dissolved in the solvent and stationary adsorbent phase. The more affinity of a given component of the mixture to the stationary phase, less time it will spend in mobile phase and the more slowly it will migrate up the plate. The resolution depends on the relative movement of solute particles in mobile phase which is referred to as Retention factor.

$$\text{Retention factor (R}_f\text{)} = \frac{\text{Distance travelled by solute from point of application}}{\text{Distance travelled by solvent from point of application}}$$

REQUIREMENTS:

1. Apparatus: Test tubes, beakers, pipettes -10ml and 1 ml, Twin trough chamber.
2. Chemicals: Toluene, Ethyl acetate, Methanol and Formic acid (10:9:1:1)
3. Standard: - Stock solution of API (10 mg in 10 ml)
4. Sample: Atorvastatin
5. HPTLC syringe
6. TLC plate: Silica gel 60 F254 (20cm x 10cm)
7. Instruments: Linomat V applicator, Visualizer, scanner III, Vortex, Weighing balance.

- Miscellaneous: Tissue paper and rubber bulb.

PROCEDURE:

- Sample preparation:
 - Atorvastatin sample from Atorva tablet
 - 1000 ppm Atorvastatin sample – 1ml stock solution of sample + 9ml of methanol as diluent.
- TLC Plate and Mobile phase Preparation:
 - Cut the TLC plate – Silica 60 F254: 20cm x 10 cm
 - Preparation of mobile phase – Toluene, Ethyl acetate, Methanol and Formic acid (10:9:1:1) mix the given ratio in a stoppered test tube.
 - Vortex the mobile phase.
- Saturation of chamber:
 - Take the CAMAG chamber (Twin trough chamber).
 - Cut the saturation paper to fit in it perfectly.
 - Pour the mobile phase onto the saturation paper.
 - Divide it equally into both the troughs.
 - Keep it for saturation for 20 min, undisturbed.
- Setting software program and Application and scanning of sample:
 - Switch on all the switches of instrument. Switch on the computer, Linomat –V applicator and knob of nitrogen gas cylinder.
 - Prepare a file using ‘Vision cats’ software on computer and load the samples using applicator.
 - Run the plate in saturated mobile phase till the solvent reaches up to 3/4th of the plate. The plate is then removed and let the solvent on the plate be evaporated.
 - Scan the developed plate using CAMAG scanner -III at desire wavelength.
 - Identify the bands and calculate Rf values of respective bands using formula

CHROMATOGRAPHIC CONDITIONS:

- Stationary phase: Aluminum sheet coated with silica gel 60 F254
- Mobile phase: Toluene, Ethyl acetate, Methanol and Formic acid (10:9:1:1)
- Plate dimension: 20 x 10 cm
- Development chamber: CAMAG Twin trough chamber
- Dimension of development chamber: 10x10 cm
- Application volume of sample: 5.0µl
- Saturation time: 20 min
- Development time: Till the solvent front reaches 80% of the plate height.
- Sample applicator: CAMAG LINOMAT- V
- Software: Vision CATS
- Detector: CAMAG TLC SCANNER-III
- Scanning wavelength: 254nm

OBSERVATION TABLE:

Track	Vial ID	Description	Volume	Position	Type
-------	---------	-------------	--------	----------	------

1	1	ATV 20ppm	0.2 μ L	N/A	Reference
2	2	ATV 40ppm	0.4 μ L	N/A	Reference
3	3	ATV 60ppm	0.6 μ L	N/A	Reference
4	4	ATV 80ppm	0.8 μ L	N/A	Reference
5	5	ATV 100ppm	1.0 μ L	N/A	Reference
6	6	ATV 150ppm	1.5 μ L	N/A	Reference
7	7	ATV 200ppm	2.0 μ L	N/A	Reference
8	8	ATV 250ppm	2.5 μ L	N/A	Reference
9	9	SPL	5.0 μ L	N/A	Sample
10	10	SPL	5.0 μ L	N/A	Sample

OBSERVATIONS:

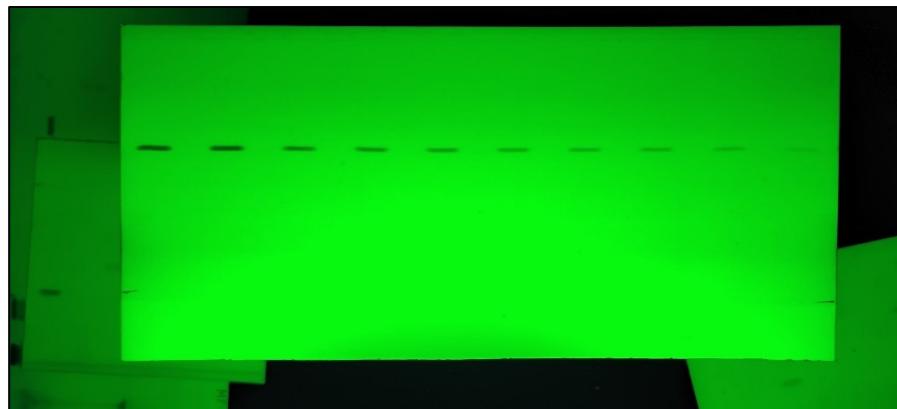


Fig 6: Visualization of bands under UV cabinet

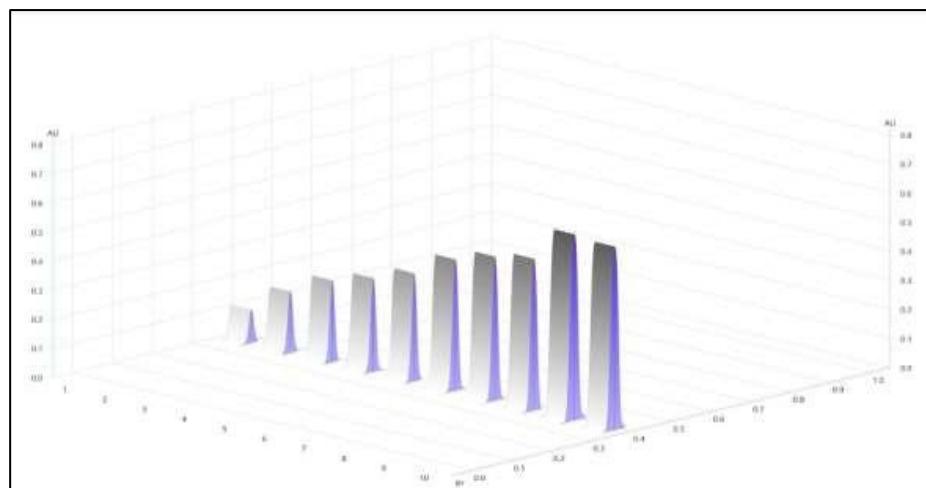


Fig 7: 3D visualization of chromatogram

Track 1

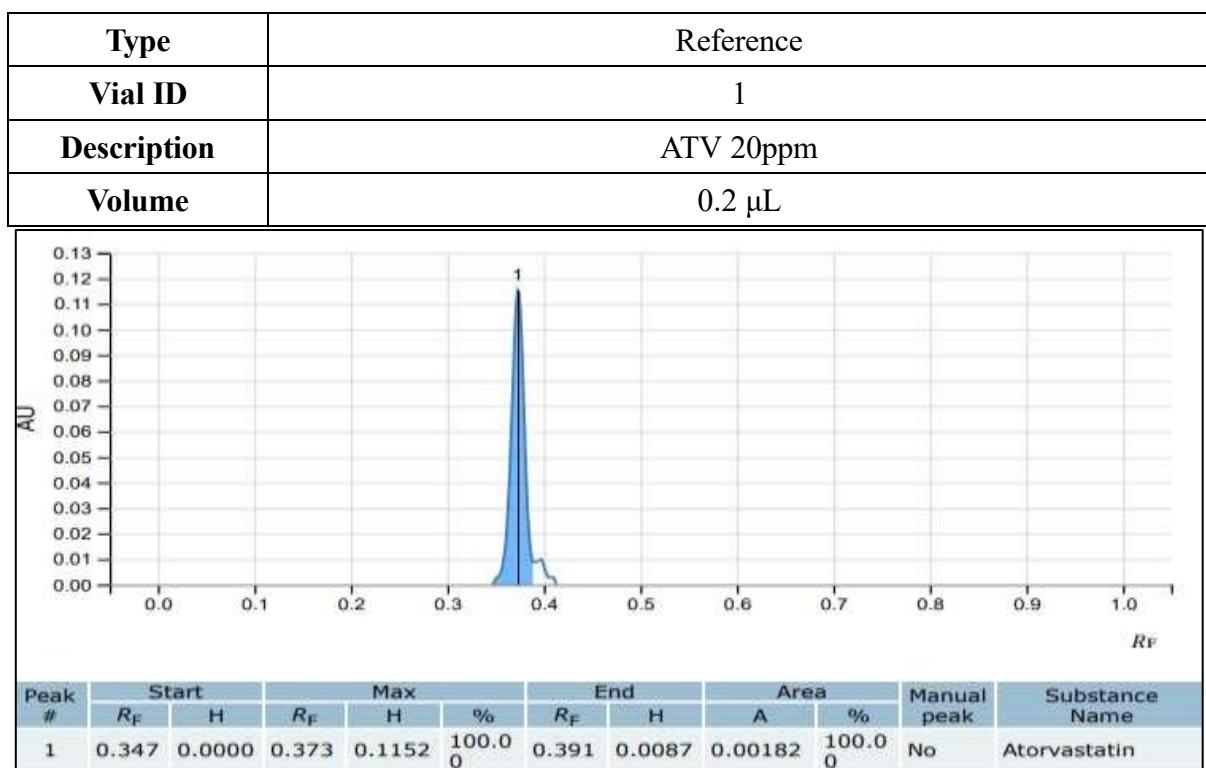


Fig 8: Track 1 chromatogram peak and Rf values

Track 2

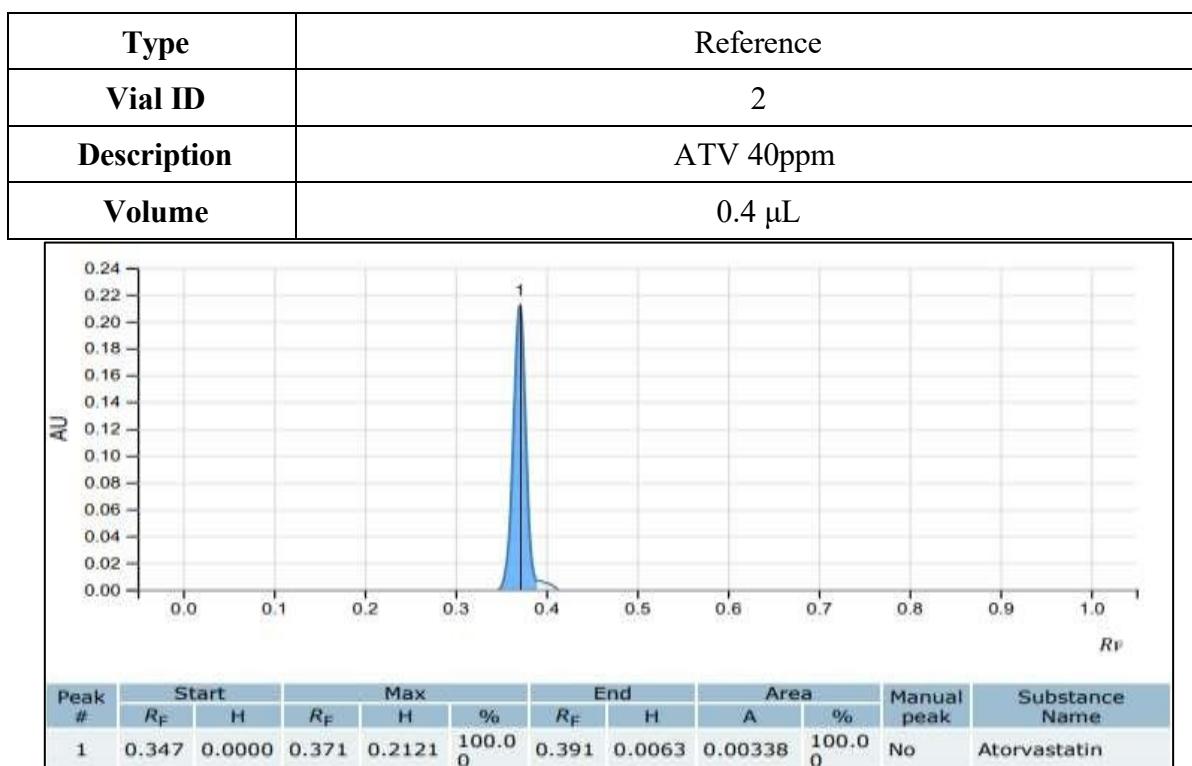


Fig 9: Track 2 chromatogram peak and R_f value

Track 3

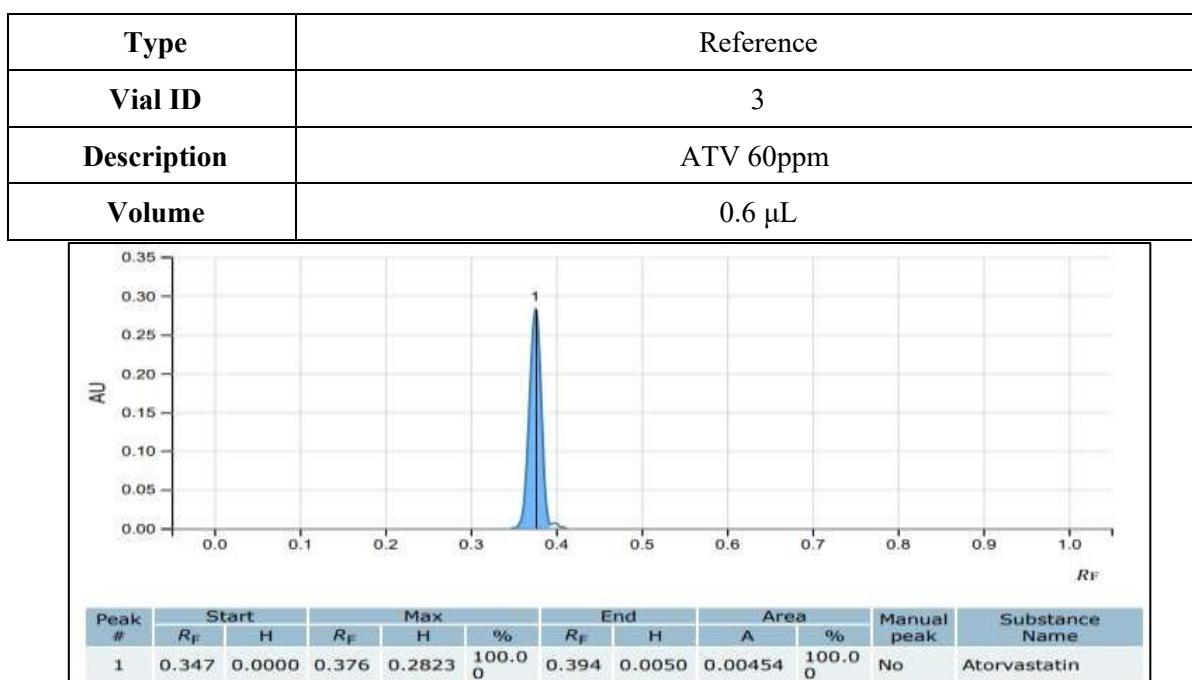


Fig 10: Track 3 chromatogram peak and R_f values

Track 4

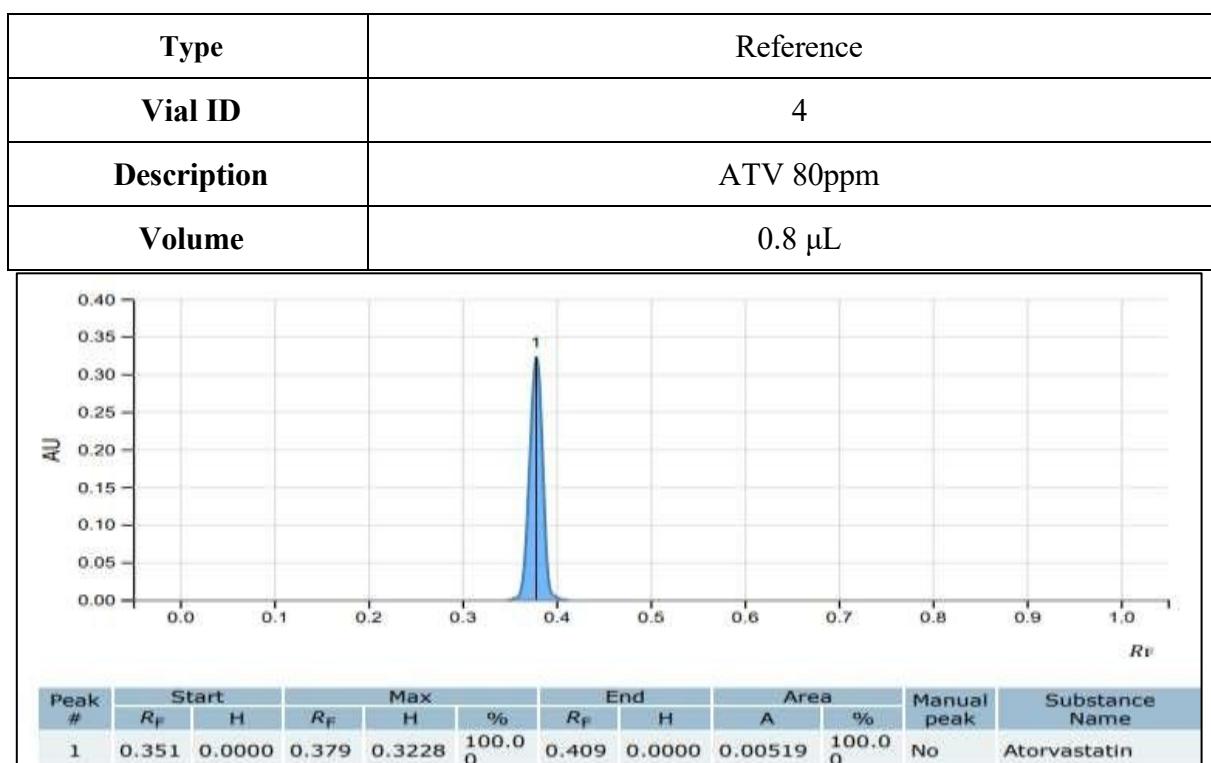


Fig 11: Track 4 chromatogram peak and Rf values

Track 5

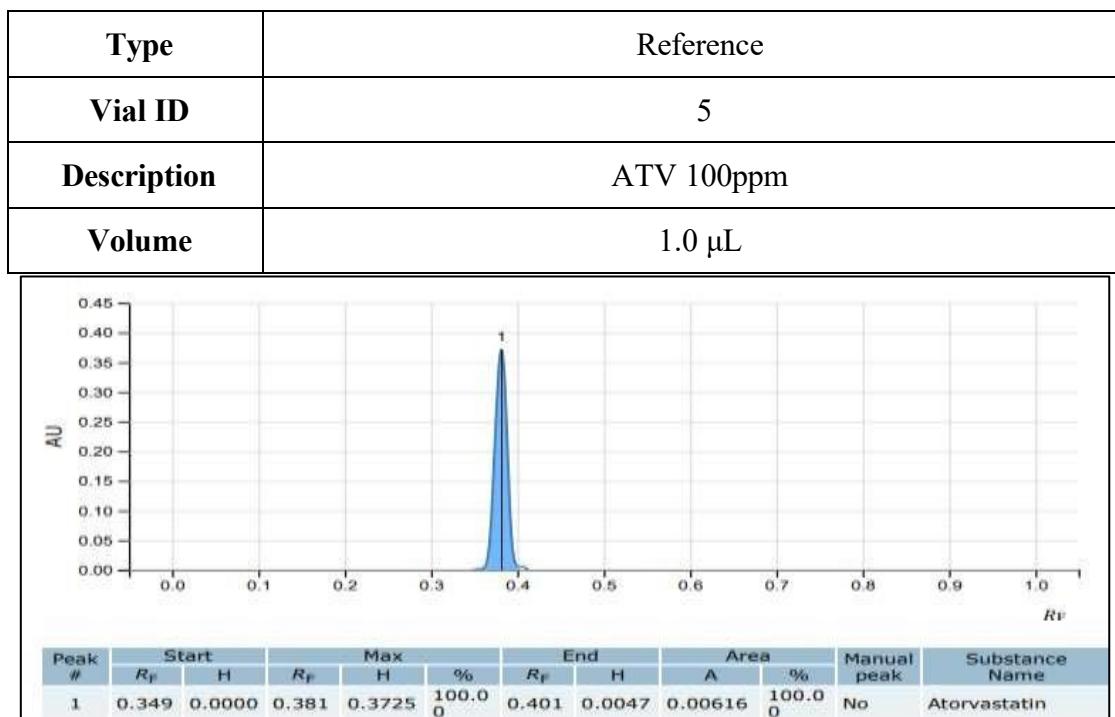


Fig 13: Track 5 chromatogram peak and Rf values

Track 6

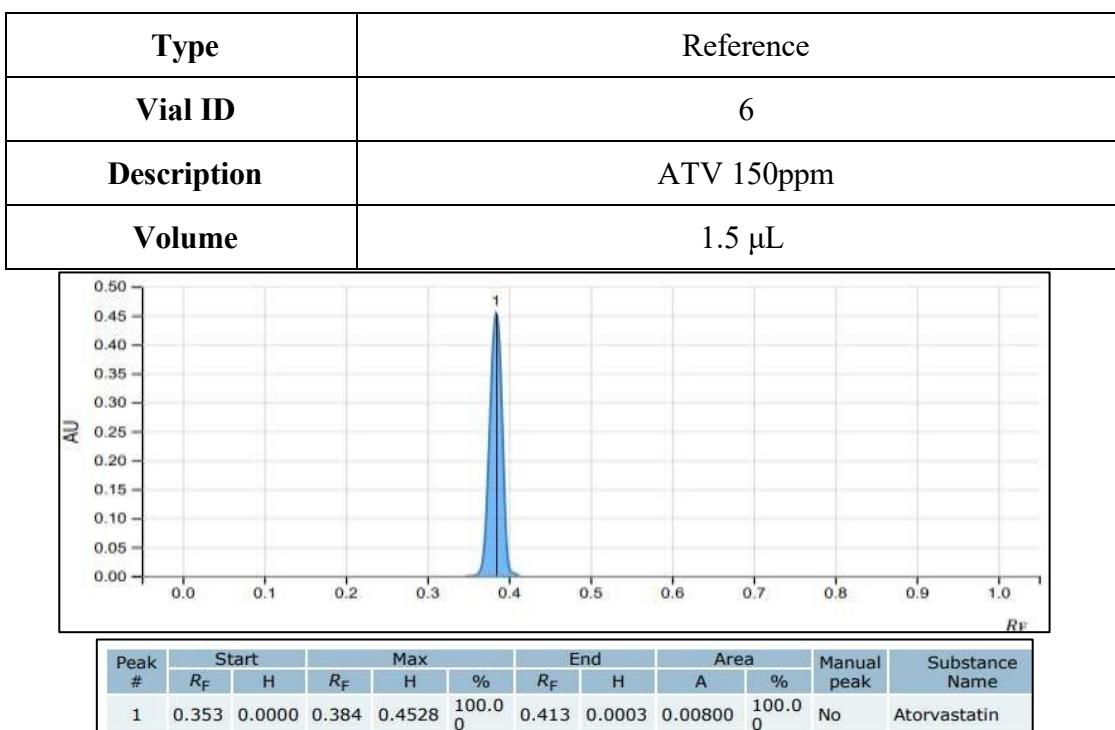


Fig 14: Track 6 chromatogram peak and Rf values

Track 7

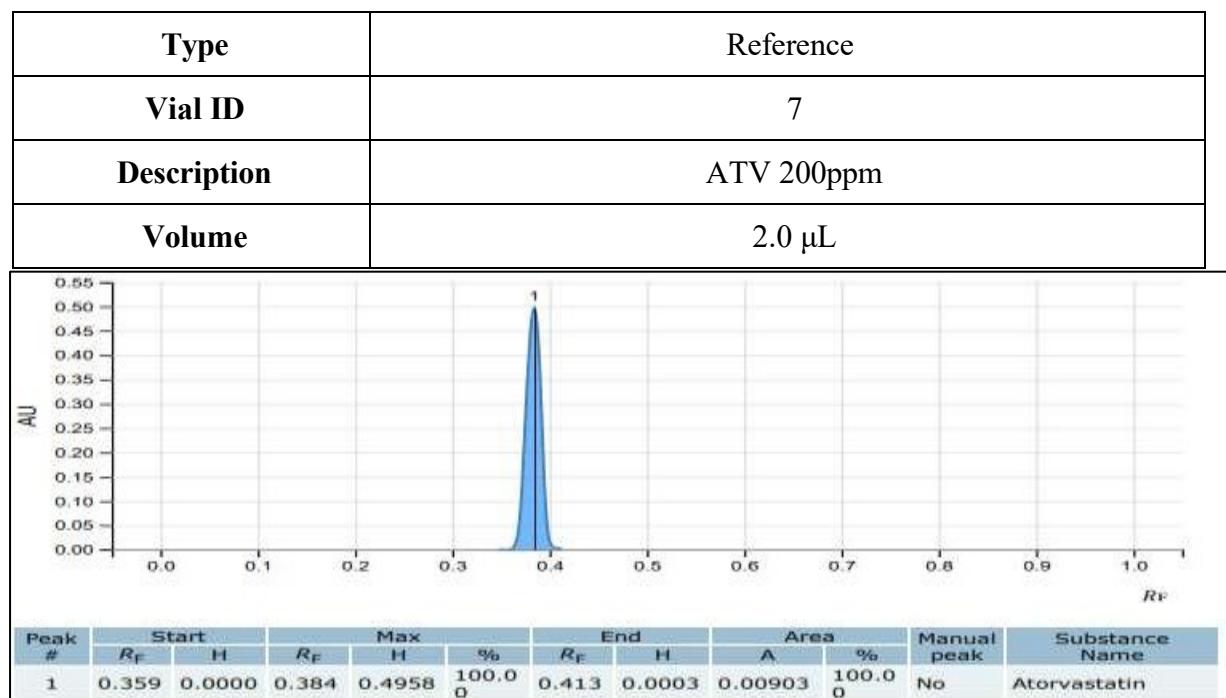


Fig 15: Track 7 chromatogram peak and Rf values

Track 8

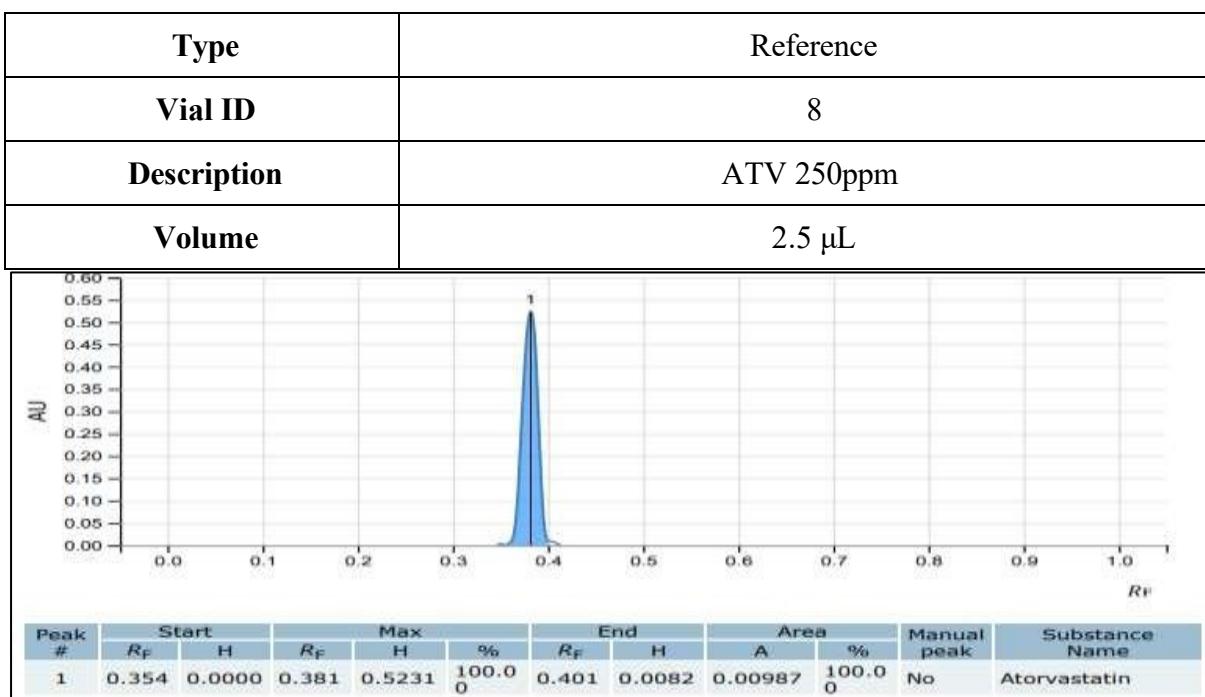


Fig 16: Track 8 chromatogram peak and Rf values

Track 9

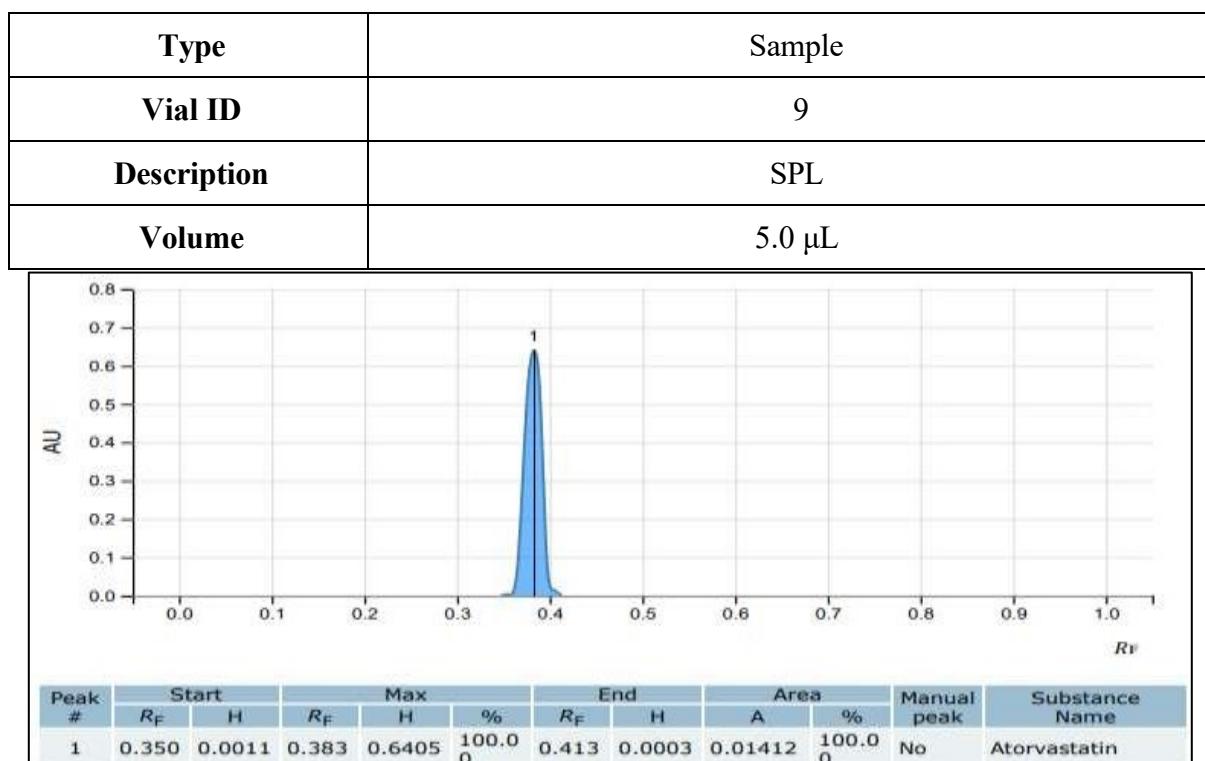


Fig 17: Track 9 chromatogram peak and Rf values

Track 10

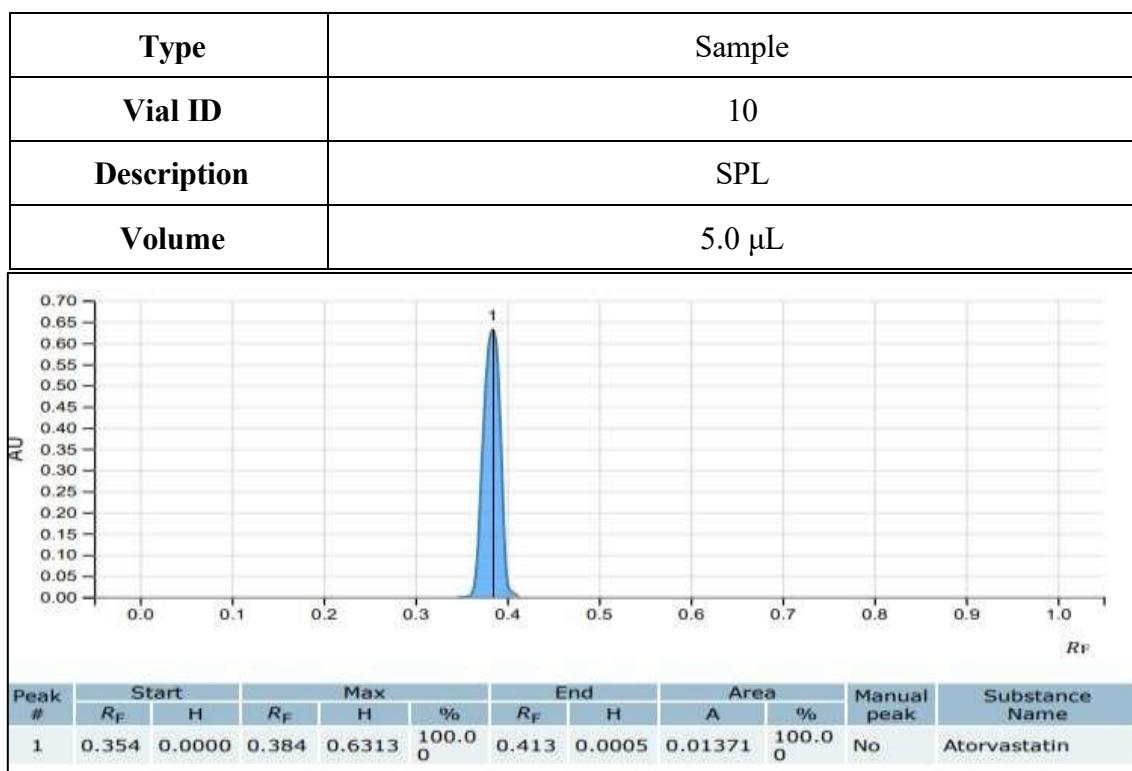


Fig 18: Track 10 chromatogram peak and Rf values

Results:						
Atorvastatin		(2 sample assignments) @ 254 nm				
Sample '10'		45.12 mg/ml	(CV unavailable)	(1 applications)	451.2 mg in 10.000 mg	
Volume: 5.0 μL		45.12 mg/ml	(CV unavailable)	(1 replicas)		
Track 10		45.12 mg/ml	225.6 μg			
Sample '9'		46.46 mg/ml	(CV unavailable)	(1 applications)	464.6 mg in 10.000 mg	
Volume: 5.0 μL		46.46 mg/ml	(CV unavailable)	(1 replicas)		
Track 9		46.46 mg/ml	232.3 μg			

Fig 19: Results for Atorvastatin sample

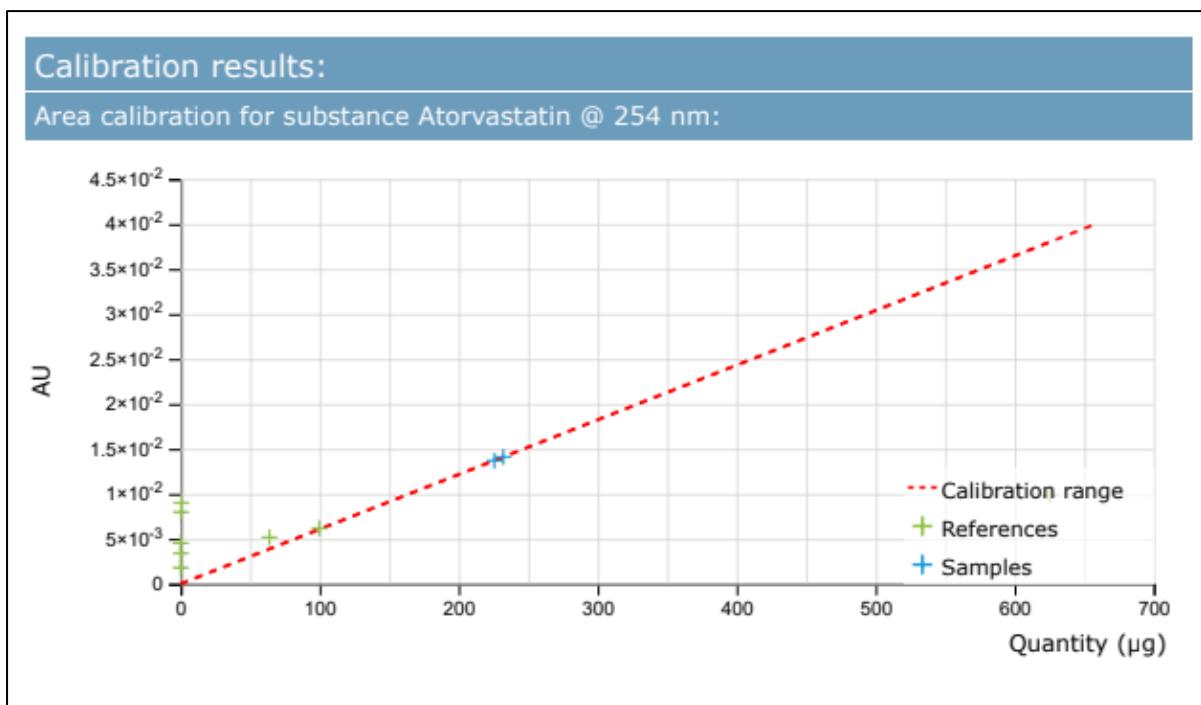


Fig 20: Area calibration graph for substance Atrovatatin at 254nm

RESULTS:

HPTLC analysis was conducted on a sample of Atorvastatin, where the chromatograms of both the standard and the samples were examined based on their respective concentrations. The compounds separated into distinct bands on an aluminum plate coated with silica gel 60 F₂₅₄. The Active Pharmaceutical Ingredient (API) was quantified by measuring both the peak area and height. The retention factor for each compound was calculated to evaluate the separation efficiency and to confirm the identity of the components.

CONCLUSION:

Qualitative and Quantitative analysis was done by using High Performance Thin Layer Chromatography (HPTLC) for the given sample of Atorvastatin.

REFERENCES:

1. Daniel, D. (n.d.) HPTLC
2. https://vikramuniv.ac.in/files/wp-content/uploads/M._Sc._IV_SEM-Paper_VUnit_IV_HPTLC-Dr_Darshana_Mehta.pdf
3. Attimarad M, Ahmed KK, Aldhubaib BE, Harsha S. High-performance thin layer chromatography: A powerful analytical technique in pharmaceutical drug discovery. Pharm Methods. 2011 Apr;2(2):71-5. doi: 10.4103/2229-4708.84436. PMID: 23781433; PMCID: PMC3658041.
4. HPTLC. (n.d.). CAMAG. <https://www.camag.com/products/hptlc>

DATE: 21/08/2024

PRACTICAL: 2
High Performance Liquid Chromatography (HPLC)

AIM:

To separate Diclofenac sodium from its combination formulation (Dicloran A) using HPLC.

INTRODUCTION:

Chromatography is an analytical technique where a sample mixture is separated into its individual component. High Performance Liquid Chromatography (HPLC) is one such mode of chromatography. It involves Mass- transfer between stationary and mobile phase.

HPLC utilizes a liquid mobile phase and solid / liquid stationary phase to separate the components of a mixture. The sample is first dissolved in a solvent and forced to flow through a chromatographic column under high pressure. In the column, the mixtures separate into its components. The column is packed with the stationary phase composed of irregularly or spherically shaped particles.

The amount of resolution is important and is dependent on the extent of interaction between the solute components and both the phases. As a result, HPLC acquires a high degree of versatility not found in other chromatographic system and it could easily separate a wide variety of chemical mixture.

INSTRUMENTATION:

HPLC instrumentation includes a pump, injector, column, detector, and data acquisition system. The heart of the system is column, where the separation occurs since the stationary phase is composed of micrometre size porous particles. A high-pressure pump is required to move the mobile phase through the column.

Eventually each component elutes from the column as a narrow band and is recorded by the recorder. Detection of the eluting compound is important. The response of detector is displayed on a chart recorder or computer screen and is known as chromatogram.

1. Stationary Phase:

Modern HPLC stationary phases are made up of small rigid porous particles with high surface area. Particle size 3 to 10 micrometres. Particle size distribution as narrow as possible depending on the type of ligand attached to the surface. The adsorbent could be normal phase (-OH, -NH₂) or reversed phase and even anion or cation exchanger

2. Mobile Phase:

In HPLC, the types of band composition of eluent are one of the variables influencing the separation of the particles should be purity, detector compatibility, solubility of sample, low viscosity, chemical inertness.

The HPLC instrument includes the following components:

1. Pump:

The role of the pump is to force a liquid (called the mobile phase) through the liquid chromatography at a specific flow rate, expressed in millimeters per min (ml / min).

Normal flow rate in HPLC is in the 1 to 2 ml / min range.

Typical pumps can reach pressure in the range of 6000 – 9000 psi (400 to 600 bar).

During the chromatographic experiment, a pump can deliver a constant mobile phase composition or an interest mobile phase composition (gradient).

2. Injector:

The injector serves to introduce the liquid sample into the flow stream of the mobile phase. Typical sample volume is 5 to 20 microliters.

The injector must also be able to withstand the high pressure of the liquid system.

An auto sampler is the automatic version for when the user has many samples to analyze or when manual injection is not practical.

3. Column:

Consider the “heart of the chromatograph” – the column’s stationary phase separates the sample components of interest using physical and chemical parameters.

The small particles inside the column are what cause the high pressure at normal flow rates. The pump must push hard to move the mobile phase through the column and this resistance cause a high pressure within the chromatograph.

4. Detector:

The detector can detect the individual analyte that come out (elute) from the column.

A detector serves to measure the amount of these molecules so that the chemist can quantitatively analyze the sample components.

The detector provides an output to a recorder or computer that result in the liquid chromatogram (i.e. the graph of the detector response).

5. Data Acquisition System:

The computer not only control the all the modules of HPLC but it takes the signal from the detector and uses it to the signal to determine the time of elution (retention time) of the sample components and the amount of sample (quantitative analysis).

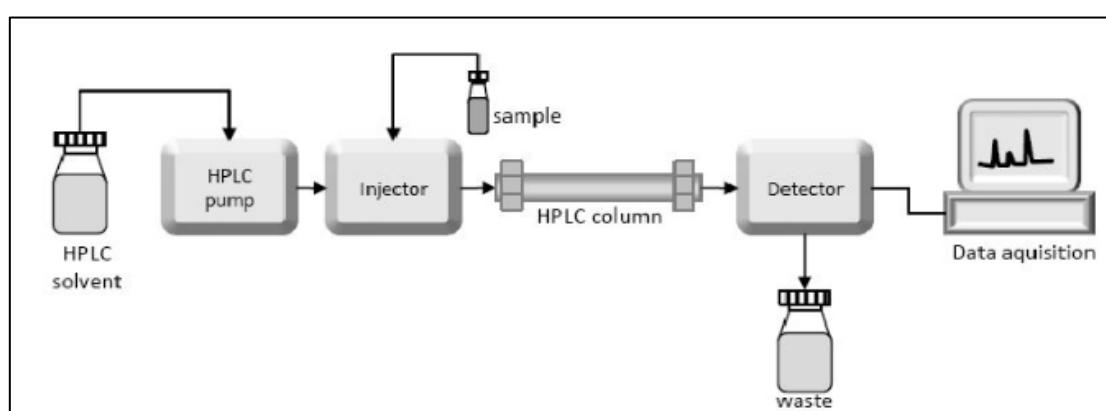


Fig 1: Instrumentation for HPLC

PRINCIPLE:

Reversed phase HPLC is the most popular mode of chromatography. In reverse phase chromatography, the stationary phase is relatively non-polar than the mobile phase. The

stationary phase is the liquid coated on inert solid support which is mostly silica. Hence the basic principle of reversed phase HPLC is partitioning between the stationary phase and mobile phase which flows through the column. In reverse phase, retention time is longer for molecules which are more non-polar while polar molecules elute more readily. The retention time can be increased by adding more water to the mobile phase. This makes mobile phase more hydrophilic relative to the more hydrophobic stationary phase. Similarly, retention time can be decreased by adding more organic solvent to the mobile phase.

APPLICATIONS OF HPLC:

1. Pharmaceutical Industry:

Drug development and quality control: HPLC is used to analyze the purity, potency, and stability of pharmaceutical products.

Pharmacokinetics: It helps in monitoring how drugs are absorbed, distributed, metabolized, and excreted in the body.

Impurity profiling: Detects trace impurities in drugs to ensure safety and compliance with regulations.

2. Clinical Diagnostics:

Biomarker analysis: HPLC is used to measure biomarkers in blood, urine, and other biological fluids.

Therapeutic drug monitoring: It helps monitor drug levels in patients, ensuring optimal dosing.

Disease diagnosis: Identifies abnormal levels of compounds associated with certain diseases (e.g., vitamins, amino acids).

3. Food and Beverage Industry:

Quality control: Ensures the safety and consistency of food products by detecting contaminants, preservatives, and additives.

Nutritional analysis: Quantifies nutrients like vitamins, sugars, amino acids, and fatty acids in food products.

Pesticide residue analysis: Monitors pesticide levels to ensure food safety.

4. Environmental Monitoring:

Pollutant detection: HPLC is used to detect trace levels of pollutants, such as pesticides, herbicides, and organic compounds, in water, soil, and air.

Water quality testing: Analyses drinking water, wastewater, and other water sources for harmful chemicals.

Waste management: Assesses chemical components in industrial effluents and waste products.

5. Biotechnology and Life Sciences:

Protein and peptide analysis: Used for protein purification and analysis of protein structure, function, and interactions.

DNA/RNA analysis: HPLC separates nucleotides and oligonucleotides for genetic and molecular studies.

Metabolomics: Identifies and quantifies small molecules (metabolites) in biological samples to understand metabolic pathways.

6. Forensic Science:

Toxicology: HPLC identifies drugs, poisons, and toxins in biological samples from crime scenes.

Drug abuse testing: Detects illegal substances and performance-enhancing drugs in athletes.

Explosives and arson investigation: Analyses trace evidence from explosive materials or fire accelerants.

7. **Chemical Industry:**

Purity assessment: Analyses the purity of raw materials, chemicals, and solvents.

Product development: Supports the creation of new chemical compounds and materials by analyzing reaction products.

Process monitoring: Helps optimize chemical manufacturing processes by monitoring reaction intermediates and byproducts.

Diclofenac:

Diclofenac sodium is non-steroidal anti-inflammatory drug (NSAID) which is used to treat minor aches and used as an analgesic to reduced pain. It is available as sodium and potassium salts. It is available as a generic drug in several formulations.

Action of DFS:

It works by blocking the action of cyclooxygenase which is involved in production of prostaglandin. This prostaglandin produced in response to injury or certain diseases and would otherwise go on to cause pain.

Uses of diclofenac sodium:

1. Treatment of pain, inflammation disorders.
2. It is used in treatment of various type of arthritis.
3. It is also used in treatment of chronic disorder and acute non-bacterial inflammation of anterior part of eye.
4. It is used in pain management in case of kidney and gall stone and also in case of active migraines.

Side effects:

1. DFS may cause side effects. Common side effects with DFS are stomach pain, constipation, diarrhea, heart burn or indigestion, headache, nausea, etc.
2. Contradiction: Hypersensitivity against Diclofenac inflammatory intestinal disorders such as ulcerative colitis, severe renal insufficiency

Formulation and Combination Formulation:

Dicloran– Single Formulation. It contains only Diclofenac sodium as an API. Dicloran A – Combination Formulation. It contains Diclofenac sodium and paracetamol.

REQUIREMENTS:

1. Apparatus:
 - a. Standard volumetric flask (10ml, 25ml),
 - b. Pipettes (1ml, 5ml),
 - c. 250ml Schott bottles,
 - d. 50ml beakers,
 - e. Hamilton syringe (100 μ l).
2. Chemicals:
 - a. Methanol (HPLC grade),
 - b. Milli Q water
3. Miscellaneous:
 - a. Mortar and Pestle
 - b. Syringe filter
 - c. Filtering Assembly
 - d. Dicloran and Dicloran A tablets
 - e. Standard: Diclofenac sodium (DFS)
4. Instruments:
 - a. Sonicator,
 - b. Shimadzu Prominence High Performance Liquid Chromatography
 - c. (HPLC) Gradient System.

PROCEDURE:

A. Preparation of standard solution of DFS

- a. 25 mg of standard DFS powder was weighed and dissolved in a minimum amount of methanol, and the volume was made up to 25 ml in a 25 ml standard volumetric flask using methanol to prepare a 1000 ppm stock solution.
- b. 1 ml of the above solution was pipetted out into a 10 ml standard volumetric flask, and the volume was made up to 10 ml using methanol to prepare a 100-ppm solution.

B. Preparation of sample

- a. Weigh tablet of Dicloran A.
- b. Find out the average weight (so that the average weight corresponds to 50mg of DFS)
- c. Crush the tablet using mortar and pestle.
- d. Weigh the powder which gives 25mg of Diclofenac and dissolve it in minimum amount of methanol.
- e. Make the volume up to 25ml using methanol. Filter the solution through Whatman filter paper no. 41, this gives 1000ppm stock solution.
- f. Using this prepare 100ppm, 10ppm solution of Dicloran A.

C. Preparation of Mobile Phase

- a. Water: Methanol (20:80).
- b. Measured 20 ml of water and mixed it with 80 ml of methanol in a clean container.
- c. Thoroughly mixed the solution to ensure a homogenous Water: Methanol (20:80) mobile phase.

STANDARD OPERATING PROCEDURE FOR THE OPERATION OF SHIMADZU PROMINENCE HPLC SYSTEM:

1. Operating procedure:

- A. Switch on the mains for UPS power supply.
 - B. Then press the ‘Test’ button on the Smart-UPS RT 2000.
 - C. Switch on the mains for the Shimadzu Gradient HPLC system.
 - D. Switch on the mains for the computer. Switch on the CPU.
 - E. Switch on the grey button with ‘Power’ written on them on the HPLC system. Each one to switch on the LC-20 AD. Pump and SPD-M20 A Detector.
 - F. Now a red light will start flickering on the instrument and the LED display for pump will get activated.
 - G. After the instrument stabilizes the red-light goes off and a green light indicates that system is ready.
2. To load the mobile phase reservoirs in the mobile phase reservoir tray
 - A. After switching on the instrument, keep the required mobile phase components contained in the Schott bottles on the mobile phase reservoir tray. [NOTE: Filter the mobile phase through the 0.2-micron membrane filter and sonicate for half an hour]
 - B. Dip the desired tubing with solvent inlet filters (A, B, C or D) in the Schott bottles containing mobile phase and close the mouth of the bottles using aluminium foil to prevent the entry of foreign matter in the mobile phase components.
 - C. To stop purging, again press ‘Purge’ button below LED display of the pump and rotate the grey knob on the pump in clockwise direction in 180 degrees (i.e. back to the original position)
 3. To connect the instrument to the software:
 - A. After purging the system, double click on the ‘LC solution’ icon the desktop of the computer.
 - B. A window opens in which click on the ‘Operation’ mode.
 - C. Click on HPLC-1 icon which leads to another window where a password will be asked. Don’t enter any password, just click on ‘Ok’.
 - D. After these 2 beep sounds (one from the pump and the other from the detector), indicates that the instrument is online software. Now the system is online.
 - E. A new window opens which shows ‘LC Real Time analyses and LC: Connected PDA: Connected’ is highlighted in green in data acquisition mode.
 4. To start an analysis on the system
 - A. To start with any new analysis first creates a method file.
 - B. To create a method file, go to uppermost toolbar and single click on ‘File’ and select ‘New method file’.
 - C. Now in the ‘Instrument parameter’s view’ there are two options ‘Normal’ and ‘Advanced’
 - Normal:
 1. First click on Normal.
 2. In this window there are 2 options ‘Simple settings’ and ‘LC time program’.
 3. Click on ‘Simple settings’ and feed the parameters.
 4. Give LC stop time and click on ‘Apply to all acquisition times’.
 5. In the option ‘Pump’ select the HPLC mode i.e. Isocratic or Low-Pressure gradient mode.
 6. Give ‘Pump a flow’ for Isocratic mode or ‘Total Pump a flow’ for Low Pressure gradient mode.
 7. If only 2 ports are in use i.e. A and B then give solvent B conc.

8. In case of use of all the 4 ports; give solvent B, C, and D concentrations accordingly.
 9. Now click on ‘Advanced’.
- Advanced:
 1. In Advanced mode there are 5 different options.
 2. First click on ‘Data acquisition’ check the LC stop time.
 3. For carrying out gradient analysis click on LC time program and give appropriate gradient program.
 4. Click on ‘Pump’ option and enter the pressure limit (Pump A) –max (350kgf/cm²) for C- 18 column.
 5. Click on PDA and enter Start and Stop wavelength [in between 190-800 (nm)] as per need of analysis.
 6. After feeding all the above information, save the method file.
 7. For saving the method file go to uppermost toolbar, single click on ‘File’ and select the option ‘Save Method File as’.
 8. Give proper path and name to Method File and save it in the desired folder.
 9. After saving method file, click on the ‘Download’ icon on the upper right side of the ‘instrument parameters view’ window, to set those conditions in the instrument.
 10. By processing ‘Pump’ button on the LC -20 AD Pump or clicking ‘Instrument ON’ icon in the toolbar, start the mobile phase flushing.
 11. By clicking ‘Plot’ icon in the chromatogram window, software starts plotting the chromatogram.
 12. Let the mobile phase flush the column for appropriate time so as to stabilize the baseline in the chromatogram window.
5. To start a single new analysis:
 - A. First step is plotting of chromatogram for the mobile phase by clicking ‘Stop’ icon to the right-hand side upper corner of the chromatogram window.
 - B. Now click on the ‘Single Start’ icon in the vertical acquisition bar at the extreme left-hand side of the chromatogram window.
 - C. This opens a new window for ‘Single Run’ as follows:
 - Fill in all the details such as Sample name, Sample ID etc.
 - Select a method file by clicking a folder icon on the left-hand side of ‘Method File’.
 - New window of ‘Select Method File’ opens, now choose accurate path and open the desired method file.
 - Similarly select the appropriate folder to save the data file using the folder icon for ‘Data File’.
 - After feeding all the data in ‘Single Run’ window click ‘Ok’. This opens a new small window which says click ‘Start’ or inject the sample.
6. To inject the sample in the manual injection:
 - A. Now rinse the sample injection loop using such as methanol i.e. take 50ml of methanol in Hamilton ml syringe, remove the trapped air bubble if any and inject the entire solvent in the loop via manual sample injection port but do not bring the knob from ‘Load’ position to ‘Inject’ position (otherwise the solvent goes to the column).

- B. After rinsing the port, now aspirate the sample (little more than 20ml as sample capacity is 20ml) in the Hamilton ml syringe remove the trapped air bubbles, now insert the needle in the manual sample injection part and inject the whole amount of sample in it. Then quickly bring the knob from ‘Load’ position to ‘Inject’ position, so that the sample now goes to the column.
 - C. After this; software automatically starts the plotting of a chromatogram for the injected sample.
 - D. To stop the run in between click on ‘Stop’ icon at the upper right-hand side of chromatogram window.
7. Post run analysis
- A. Go to ‘LC solution’ window click on ‘Operation Mode’ and single click on ‘Post Run’.
 - B. ‘LC pasture analysis’ window opens.
 - C. To open desired data file, single click on ‘File’ in the upper most toolbar.
 - D. Now click on ‘Open’, select appropriate path and open desired data file.
 - E. A new window opens which shows a chromatogram view of selected data.
 - F. To view a peak table for selected data file, single click on ‘view’ in uppermost toolbar and select a peak table. A peak table view opens.
8. To create a report format file
- A. Single click on ‘File’ in the uppermost toolbar in the post run window, click on ‘new’ and then select ‘Report Format File’
 - B. The Report Format file opens.
 - C. Toolbar icons: Addition of a text box; LC/PDA Chromatogram; LC/PDA peak table and Sample information.
 - D. For adding every component in the report; single click on their respective icons in the toolbar and create spaces for them in the blank report format below by dragging separate boxes for each one of them.
 - E. After creating a desired report format file go to uppermost toolbar, select ‘File’ click on
 - F. ‘Save Report Format File as.
 - G. Give proper path and save it in the desired folder.
9. To switch off the system
- A. After the analysis is complete always flush the column with HPLC grade Methanol:RO water (50:50) for 30-45mins and finally store the column with 100% HPLC grade Methanol for 30mins before switching off the system.
 - B. After final Methanol flushing, turn off the mobile phase flow by pressing ‘Pump’ button on LC-20 AD pump and stop plotting the chromatogram.
 - C. Then exit from the existing windows by cancelling them and finally cancel and close the ‘LC solution’ window, the software gets closed.
 - D. Refresh the desktop and shutdown the computer and then switch off the mains of the computer.
 - E. Now to switch off the instrument, press the grey button for power on the LC- 20 AD pump and SPD-M20 A detectors and switch them off, the LED display goes off.
 - F. Then switch off the mains for the instrument.
 - G. After switching off mains for computer and instrument; switch off mains for UPS and then press the button next to ‘Test’ button on the smart UPS RT 2000.
 - H. All the green light on the UPS goes off and whole HPLC system with UPS gets shut down.

Chromatographic conditions:

Sr. No	Parameters	Specification
1	Column	LCGC Quadisil BDS C18 4.6x250mm 5μ
2	Mobile Phase	Water:Methanol(20:80)
3	Flow rate	0.8mL/min
4	Detection wavelength	272nm
5	Run time	10 minutes
6	Injector volume	20μl
7	Standard concentration	100 ppm
8	Sample concentration	100 ppm

CHROMATOGRAMS:

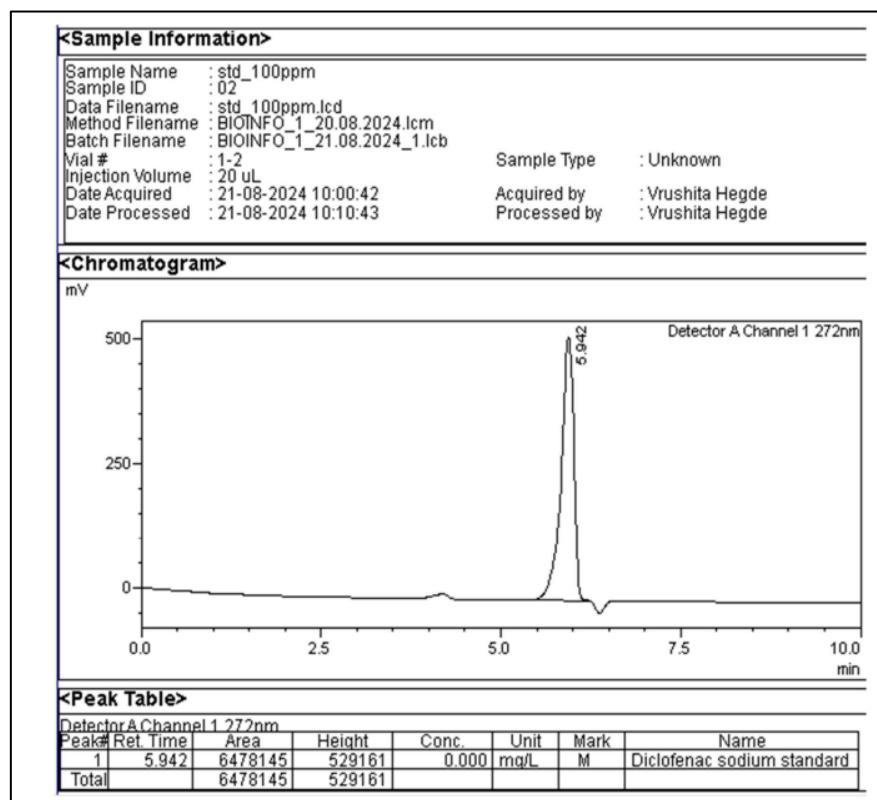


Fig 2: Chromatogram for standard

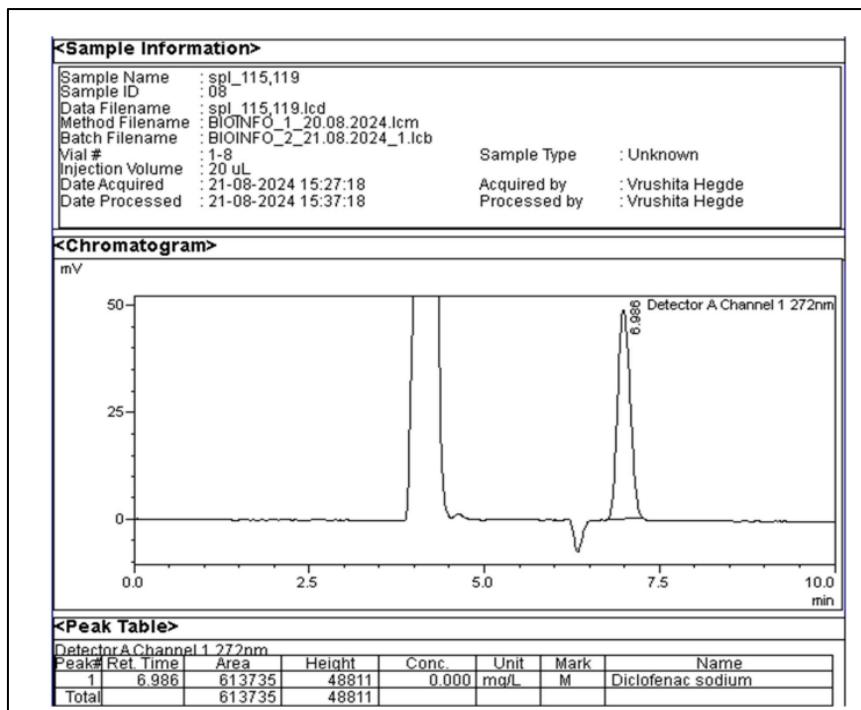


Fig 3: Chromatogram for Sample

CALCULATIONS:

Amount of DFS per tablet

$$\begin{aligned}
 &= \frac{\text{AUC Sample} \times \text{Conc. of Standard} \times \text{Weight of Standard} \times \text{Weight of Tablet}}{\text{AUC standard} \times \text{Conc. of Sample}} \\
 &= \frac{613735 \times 10 \times 100}{6478145 \times 10 \times 100} \times 465 \\
 &= 44.0537 \text{ mg/tablet}
 \end{aligned}$$

$$\begin{aligned}
 \text{Extraction efficiency} &= \frac{\text{Amount present in tablet}}{50} \times 100 = \frac{44.0537}{50} \times 100 \\
 &= 88.10\%
 \end{aligned}$$

OBSERVATIONS:

Sample	Concentration (ppm)	Rt (min)	Area under Curve (in nm ²)	Amount (in mg)
Standard	10 ppm	5.942	6478145	50
Dicloran A	10 ppm	6.986	613735	44.0537

RESULTS:

1. The retention time of standard is 5.942 mins and which approximately matches with Dicloran tablet sample with retention time 6.986.
2. The label claim of DFS sample tablet was calculated. It was found to be 44.0537 mg/tablet.
3. The extraction efficiency was found to be 88.10%

CONCLUSION:

1. The HPLC analysis of modern drug Dicloran A was successfully performed.
 2. Amount present of DFS was less than the label claim.
 3. The extraction efficiency was found to be 88.10% which matches with the acceptance criteria (80%-120%).
 4. Hence, the extraction method is compatible with the analyte of interest.
-

PRACTICAL: 3

Gas Chromatography

AIM:

To calculate the percentage of ethanol from given sample by performing Gas chromatography.

THEORY:

Chromatography is the general name given to the methods by which two or more compounds in a mixture are physically separated by distributing between two phases: a stationary phase which can be a solid or liquid supported on a solid and a mobile phase, either a gas or a liquid which flows continuously around the stationary phase.

The separation of the individual components results from the relative difference in affinity for the stationary phase.

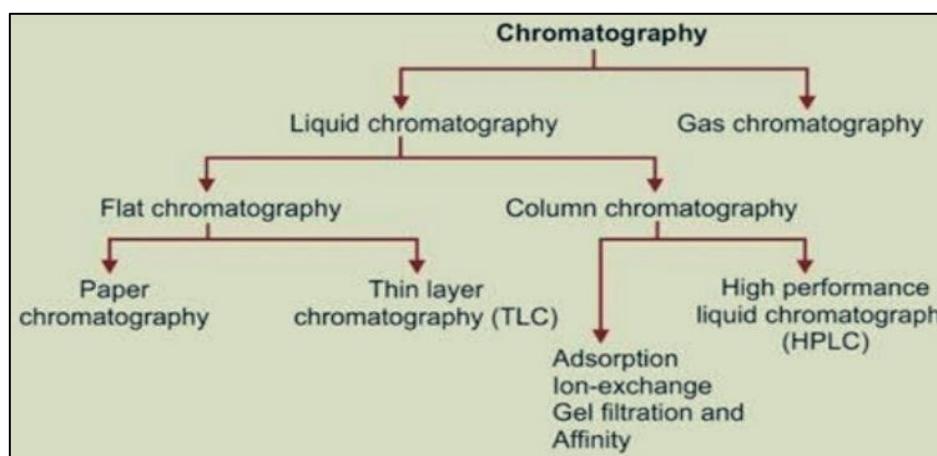


Fig 1: Types of chromatography

Gas Chromatography:

Gas chromatography (GC) is a powerful analytical technique used for separating and analyzing compounds that can be vaporized without decomposition. It differs from other forms of chromatography in that the mobile phase is a gas and the components are separated as vapours. It is widely employed in various scientific disciplines including chemistry, biochemistry, pharmaceuticals, and environmental science.

It is thus used to separate and detect small molecular weight compounds in the gas phase. The sample is either a gas or a liquid that is vaporized in the injection port. The mobile phase for gas chromatography is a carrier gas, typically helium because of its low molecular weight and being chemically inert. The pressure is applied and the mobile phase moves the analyte through the column. The separation is accomplished using a column coated with a stationary phase.

Components of Gas chromatography:

- 1. Sample Injector:** Various types of samples can be introduced into the column through the injector port

2. **Column/Stationary Phase:** This is where the separation of sample compounds takes place. The column is a coiled tube made of metal or glass material that can withstand high temperatures. Columns vary in both length and diameter.
3. **Mobile Phase:** The carrier gas is introduced from a gas cylinder into the gas chromatograph. It moves through the column at a constant flow rate and exits at the detector outlet. Unlike some other analytical methods, the mobile phase in GC does not interact with chemicals and only serves to carry them. Because of this, the carrier gas must be inert. Examples include helium, nitrogen, and argon.
4. **Oven:** It basically provides heat to the column through which the materials get vaporized.
5. **Detector:** This device at the end of the column senses each compound as it elutes. The data recorded by the detector is transmitted to a computer that produces a two-dimensional plot called a chromatogram.

PRINCIPLE:

The principle of gas chromatography is based on partitioning, the components of the sample will partition (i.e. distribute) between the two phases: the stationary phase and the mobile phase. Compounds that have a greater affinity for the stationary phase spend more time in the column and thus elute later and have a longer retention time (R_t) than samples that have a higher affinity for the mobile phase.

The sample mixture is vaporized and injected into the GC system. Inside the GC column, the sample interacts with the stationary phase, which coats the inside of the column. This stationary phase may be polar or non-polar, depending on the type of compounds being analyzed. As the carrier gas (typically helium or nitrogen) flows through the column, compounds in the sample partition between the mobile phase (carrier gas) and the stationary phase.

Compounds that interact more strongly with the stationary phase take longer to elute (come out) of the column, resulting in longer retention times. Different compounds in the mixture have different affinities for the stationary phase based on their chemical properties such as polarity, size, and interaction with the stationary phase. This differential interaction leads to separation of the compounds as they travel through the column. After separation the compounds are detected by the detectors and signals are generated in the form of peaks which are analysed by the softwares and the peaks are integrated.

Temperature is very critical factor in GC as temperature affects not only retention but also relative retention in gas chromatography and therefore, when we change temperature, we also change the selectivity of the separation.

REQUIREMENTS:

1. Chemicals:
 - a. Sample: Sanitizer
 - b. Standard: Ethanol
2. Instrument:
 - a. GC Model: GC-2014 SHIMADZU
3. Software:
 - a. GC Solution
4. Gases:

- a. Carrier gas: Nitrogen / Helium
- b. Air or Hydrogen
- 5. Injection syringe
- 6. Computer with good internet speed

INSTRUMENTAL PARAMETERS:

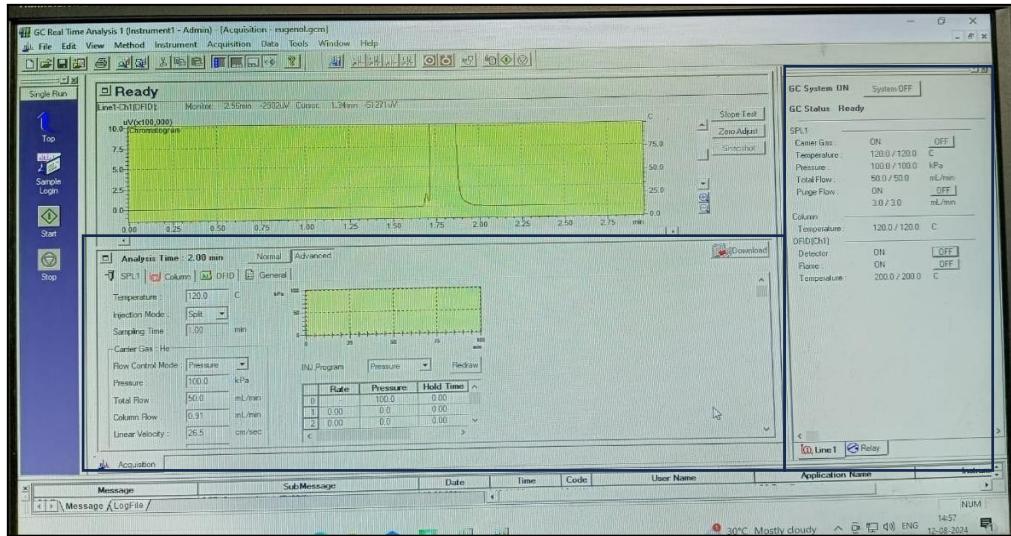


Fig 2: Conditions for the Instrument

METHODOLOGY:

1. Prepare the required sample.
2. Set the instrumental parameters according the sample, DOWNLOAD the method and switch the SYSTEM ON.
3. Wait for the System to be READY.
4. Enter the “SAMPLE ID” in the software.
5. Rinse the given GC Syringe with the sample thrice.
6. Aspirate 0.2 μ l of the sample into the syringe.
7. Press on the “START” button.
8. Integrate the chromatogram. Calculate the percent purity.

OBSERVATIONS:

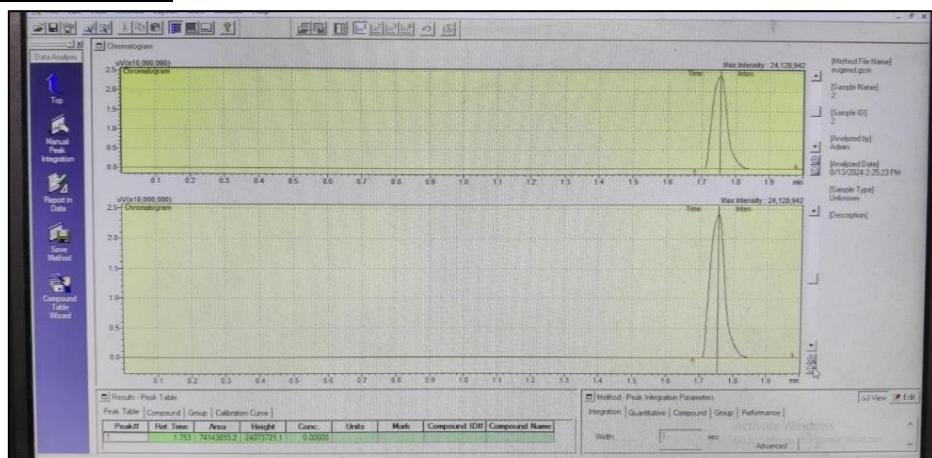


Fig 3: Chromatogram obtained for standard

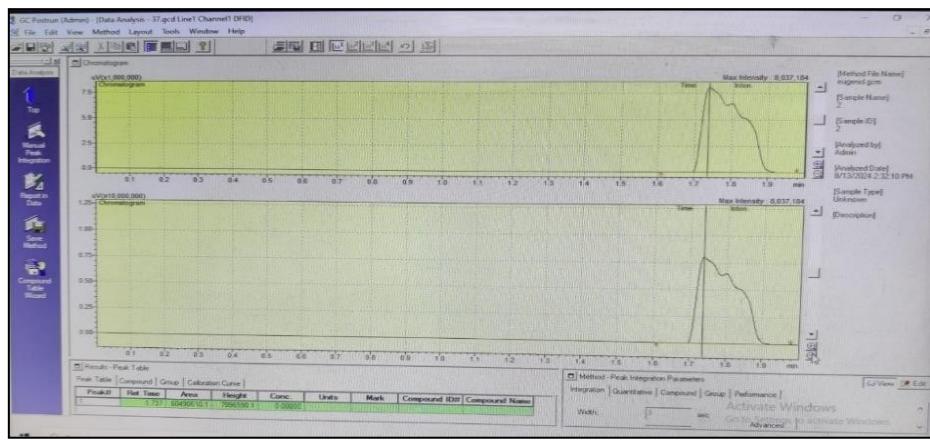


Fig 4: Chromatogram obtained for sample



Fig 5: Overlaying of standard and sample chromatogram



Fig 6: GC instrument (GC-2014 SHIMADZU)

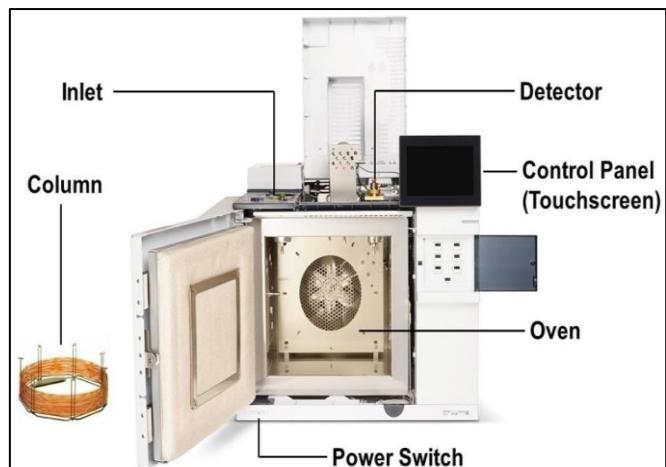


Fig 7: Parts of GC instrument

CALCULATIONS:

$$\begin{aligned}
 \text{Percentage of ethanol} &= \frac{\text{Area under curve for sample}}{\text{Area under curve for standard}} \times 100 \\
 &= \frac{60490610.1}{74143855.2} \times 100 \\
 &= 0.815854 \times 100 \\
 &= 81.5854 \%
 \end{aligned}$$

RESULTS:

The gas chromatography analysis was performed on the sanitizer sample to determine the percentage of ethanol. This was done by calculating the area under the curve for both the sample and a standard. The percentage of ethanol in the sample was found to be 81.5854%.

CONCLUSION:

Quantitative analysis of ethanol in the given sanitizer sample was done by performing gas chromatography.

REFERENCES:

1. Skoog, D. A.; Holler, F. J.; Crouch, S. R. Principles of Instrumental Analysis. Sixth Edition, Thomson Brooks/Cole, USA, 2007.
2. Jwaili M. Pharmaceutical applications of gas chromatography. Open J. Appl. Sci. 2019;9(9):683- 690. doi:10.4236/ojapps.2019.99055
3. Scott R P W. Introduction to Analytical Gas Chromatography, Revised and Expanded. CRC press, 2017.
4. McNair H M, Miller J M, Snow N H. Basic gas chromatography. John Wiley & Sons, 2019.

PRACTICAL: 4
Fourier Transform Infrared Spectroscopy (FTIR)

AIM:

To perform the FTIR analysis of Diclofenac Sodium.

THEORY:

Fourier Transform Infrared Spectroscopy, also known as FTIR Analysis or FTIR Spectroscopy, is an analytical technique used to identify organic, polymeric, and, in some cases, inorganic materials. The FTIR analysis method uses infrared light to scan test samples and observe chemical properties. The FTIR instrument sends infrared radiation of about 10,000 to 100 cm⁻¹ through a sample, with some radiation absorbed and some passed through. The absorbed radiation is converted into rotational and/or vibrational energy by the sample molecules. The resulting signal at the detector presents as a spectrum, typically from 4000 cm⁻¹ to 400cm⁻¹, representing a molecular fingerprint of the sample. Each molecule or chemical structure will produce a unique spectral fingerprint, making FTIR analysis a great tool for chemical identification. Fourier Transform Infrared Spectroscopy (FTIR) is a powerful analytical technique that measures the infrared region of the electromagnetic spectrum to determine the molecular composition and structure of materials. The basic principle behind FTIR is that different molecular bonds within a sample absorb specific frequencies of infrared light, causing the molecules to vibrate. By analyzing the absorption pattern, or spectrum, scientists can identify the types of bonds present and gain insights into the sample's composition.

FTIR analysis involves exposing a sample to infrared radiation and measuring the light that passes through or is absorbed by the material. The resulting spectrum is a graph displaying the different frequencies of infrared light absorbed by the sample, with the vertical axis showing the amount of light absorbed and the horizontal axis representing the frequency or wavelength of the light. This spectrum serves as a unique "fingerprint" for the material, allowing for identification by searching against reference databases.

FTIR spectroscopy has several advantages over traditional dispersive infrared spectrometers, including faster data collection, higher sensitivity and resolution, and the ability to analyze samples in various states (solid, liquid, or gas). The technique has found widespread applications in fields such as chemistry, materials science, pharmaceuticals, and quality control, enabling rapid and non-destructive analysis of samples as small as 20 microns. By leveraging the principles of Fourier transform mathematics and the unique vibrational properties of molecules, FTIR spectroscopy has become an essential tool for modern analytical chemistry and materials characterization.

FTIR consists of following instrument components:

Interferometer:

1. The source: Nearest glow bulb: IR rays are emitted from glowing black body source. This beam lanes through one aperture which controls the amount of energy presented to the sample.
2. Sample holder: The beam enters the sample compartment where it is transmitted through a reflected surface of a sample depending upon the time of analysis.

3. Detector: The beam finally passes through the detector for final measurement. The detectors used are specially designed to measure the signal. Detectors in FTIR spectroscopy convert infrared radiation into electrical signals for analysis, enabling quantification of molecular absorption. The detector used in FTIR is Deuterated Triglycine Sulfate (DTGS)
4. Computer: Computers are essential in FTIR spectroscopy for controlling the spectrometer and processing data, converting raw signals into interpretable spectra. They also facilitate spectral analysis, peak identification, and automated reporting, enhancing efficiency and accuracy in various applications.

The interferometer, specifically the Michelson interferometer, is a vital component of Fourier Transform Infrared Spectroscopy (FTIR) that enables the simultaneous collection of spectral data across a wide range of wavelengths. It works by splitting an incoming beam of infrared light into two paths using a beam splitter; one beam reflects off a fixed mirror while the other reflects off a movable mirror. As the movable mirror changes position, it creates an optical path difference between the two beams. When these beams recombine at the beam splitter, they produce an interference pattern that varies with the position of the movable mirror, resulting in an interferogram—a complex signal representing light intensity as a function of mirror position. This raw data is then processed using Fast Fourier Transform (FFT) to convert it into a frequency spectrum that displays absorbance versus wavenumber, allowing for the identification and quantification of materials based on their unique spectral fingerprints. The Michelson interferometer's design ensures high precision and resolution, making it essential for various applications in chemistry, biology, and materials science.



Fig 1: Interferometer

Pellet Die:

In the setup of Fourier Transform Infrared Spectroscopy (FTIR), the upper jacket, lower jacket, anvil, and plunger are essential components used primarily for sample preparation and analysis, particularly for solid samples. The upper jacket and lower jacket form part of a heating or cooling system that maintains the temperature of the sample during analysis. These jackets encase the sample holder and can be adjusted to create a controlled environment, which is crucial for obtaining accurate spectral data, especially when analysing temperature-sensitive materials. The anvil serves as a pressing mechanism that applies uniform pressure to the sample, ensuring it is evenly compressed against the infrared-transparent window, which is essential for effective light transmission and accurate measurements. The plunger works in

conjunction with the anvil to apply force to the sample, allowing for precise adjustments in pressure. Before subjecting the sample to hydraulic pressure, this process is integral in forming a pellet of the sample, typically using potassium bromide (KBr) as a matrix. Together, these components facilitate optimal sample preparation and measurement conditions, enabling FTIR to deliver reliable and reproducible results across various applications.

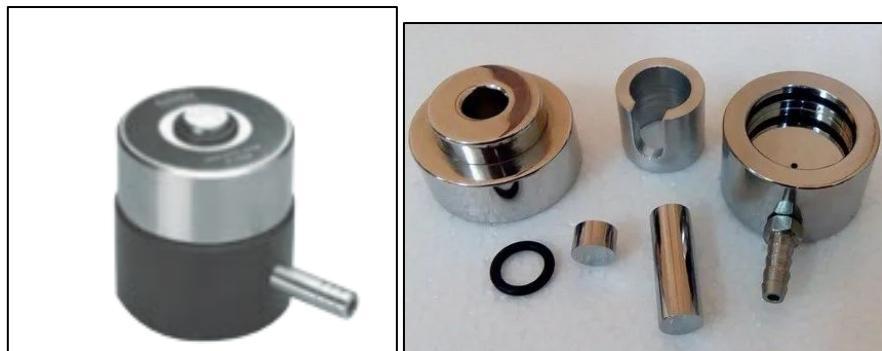


Fig 2: Pellet Die Setup

Hydraulic Pressure:

Hydraulic pressure is essential in preparing samples for Fourier Transform Infrared (FTIR) spectroscopy. It allows for the creation of uniform, transparent KBr pellets containing the sample, which is crucial for obtaining accurate and reproducible spectral data. The hydraulic press applies controlled, evenly distributed force to compress the sample-KBr mixture into a solid disc, ensuring minimal variability in thickness and density. This standardized preparation process is quick, versatile, and user-friendly, making it an integral part of the FTIR analysis workflow. The resulting pellets interact with infrared light in the spectrometer, providing valuable insights into the molecular structure of the sample.



Fig 3: Hydraulic Press

Shimadzu FTIR Spectrophotometer Software: IRsolution

IRsolution is a sophisticated software designed for use with Shimadzu FTIR spectrometers, providing a comprehensive platform for data acquisition, analysis, and reporting in infrared spectroscopy. It features advanced search capabilities, including spectral, peak, text, and combination searches, supported by a robust library containing approximately 12,000 spectra from various categories such as pharmaceuticals and contaminants. The software automates routine tasks like data manipulation and reporting, significantly reducing analysis time and allowing operators with minimal FTIR experience to perform complex analyses easily. Additionally, the "Easy Macro" function enables users to create macros for repetitive tasks, enhancing efficiency. With its intuitive interface and powerful algorithms for contaminant identification and pass/fail judgments based on pharmacopoeia standards, IRsolution streamlines the workflow of FTIR analysis while ensuring high accuracy and reliability in results.

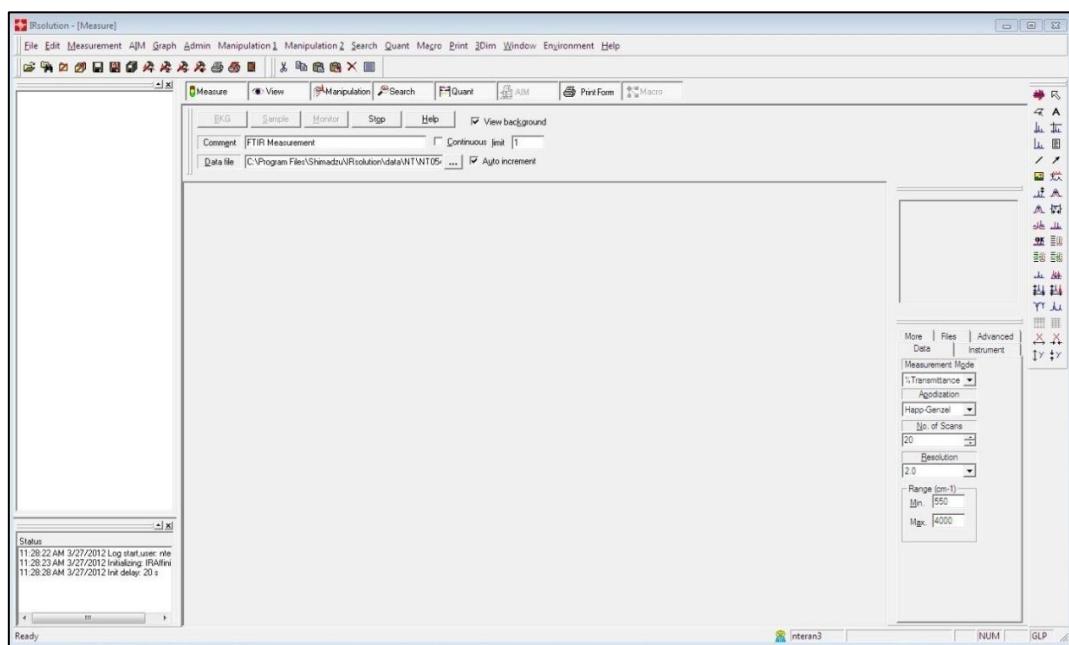


Fig 4: IRsolution Software

PRINCIPLE:

The principle of Fourier Transform Infrared (FTIR) spectroscopy is based on the absorption of infrared radiation by molecules, which leads to changes in their vibrational energy levels. When infrared light interacts with a molecule, it can only absorb energy and vibrate if the molecule undergoes a change in dipole moment and if the energy of the infrared photon matches the energy gap between vibrational levels. In an FTIR spectrometer, the infrared radiation is first split by a beam splitter into two optical paths—one reflecting off a fixed mirror and the other off a moving mirror—creating an optical path difference. When the beams recombine at the beam splitter, they produce an interference signal, known as an interferogram, which is measured by a detector. This interferogram contains information about the sample's absorption characteristics, and by applying a Fourier transform, it is converted from the time domain to the frequency domain, resulting in an infrared spectrum that reveals the characteristic absorption peaks of the sample, enabling qualitative and quantitative analysis of its molecular structure.

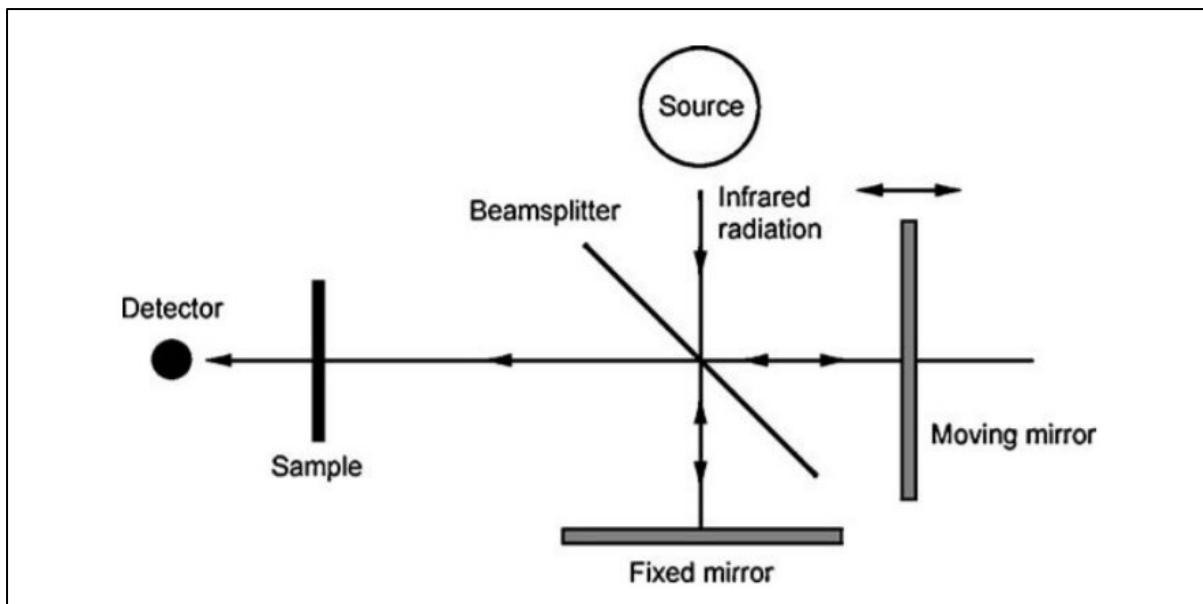


Fig 5: Basic schematic of Michelson interferometer

REQUIREMENTS:

1. Sample: Diclofenac Sodium
2. Chemical: Chloroform, KBr
3. Apparatus: Mortar and pestle, Pellet Die, Hydraulic pressure
4. Instruments: Microwave, Shimadzu FTIR Spectrophotometer model no. A213750-03405
5. Miscellaneous: Tissue Paper, Metal spatula

PROCEDURE:

1. Potassium bromide (KBr) and Diclofenac Sodium (DFS) sample in a ratio of 4:1 using a mortar and pestle to form a smooth, homogeneous powder.
2. Clean the pellet die components (upper jacket, lower jacket, and anvil) using chloroform and tissue paper to ensure a contaminant-free surface.
3. Assemble the upper and lower jacket of the pellet die and shift the anvil into position.
4. Fill the sample compartment of the anvil evenly with the ground KBr-DFS mixture using a spatula.
5. Insert the plungers into the die, with the flat surface facing downwards.
6. Place the assembled pellet die in a hydraulic press and apply a pressure of 10 tons for 10 minutes to compress the sample into a thin, transparent pellet.
7. Dismantle the pellet die and carefully scrape the formed pellet using a spatula to remove it from the die.
8. Transfer the pellet to the sample holder in the FTIR spectrometer and record the infrared spectrum.

RESULTS:

1. In the FTIR analysis of the Diclofenac Sodium (DFS) solid sample, a total of 209 peaks were observed in the recorded spectrum. For each peak, detailed information was collected, including the following parameters:
 - a. Intensity: The height of each peak, indicating the strength of absorption.
 - b. Corrected Intensity (Corr. Int): The intensity values adjusted for baseline variations and noise.
 - c. Base Height (Base H): The vertical distance from the baseline to the peak's highest point.
 - d. Base Width (Base L): The horizontal width of the peak at its base, providing insight into peak sharpness.
 - e. Area: The total area under each peak, representing the integrated absorbance over the wavelength range.
 - f. Corrected Area (Corr. Area): The area values adjusted for baseline corrections to ensure accurate quantification.
2. This comprehensive dataset allows for a thorough analysis of the molecular characteristics of Diclofenac Sodium, facilitating identification and quantification of functional groups present in the sample.

CONCLUSION:

The FTIR spectroscopy analysis of Diclofenac Sodium (DFS) successfully identified its molecular characteristics through the observation of 209 peaks in the spectrum. This study confirmed the presence of specific functional groups associated with the compound, demonstrating the effectiveness of FTIR as an analytical tool.

REFERENCES:

1. Berna, F. (2017). Fourier Transform Infrared Spectroscopy (FTIR). In A. S. Gilbert (Ed.), *Encyclopedia of Geoarchaeology* (pp. 1-5). Springer. https://doi.org/10.1007/978-1-4020-4409-0_15
2. Fadlelmoula, A., Pinho, D., Carvalho, V. H., Catarino, S. O., & Minas, G. (2022). Fourier Transform Infrared (FTIR) Spectroscopy to Analyse Human Blood over the Last 20 Years: A Review towards Lab-on-a-Chip Devices. *Micromachines*, 13(2), 187.
3. <https://doi.org/10.3390/mi13020187>
4. Hamamatsu Photonics. (n.d.). FTIR (Fourier Transform Infrared Spectroscopy). Retrieved from <https://www.hamamatsu.com/eu/en/applications/analytical-equipment/fir.html>
5. Jasco Inc. (n.d.). Sampling techniques for FTIR spectroscopy. Retrieved from
6. <https://jascoinc.com/learning-center/theory/spectroscopy/fundamentals-ftir-spectroscopy/sampling/>
7. Libretexts. (2023, April 10). How an FTIR spectrometer operates. Chemistry LibreTexts. https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_%28Physical_and_Theoretical_Chemistry%29/Spectroscopy/Vibrational_Spectroscopy/Infrared_Spectroscopy/How_an_FTIR_Spectrometer_Operates

Name: Ms. Prarthi Hrishit Kothari

Class: M. Sc. Bioinformatics (Part II)

Roll Number: 115

Course: M. Sc. Bioinformatics

Department: Department of Bioinformatics

Paper: Mandatory Paper I

**Paper Name
and Code:** Biophysical techniques, IPR,
Cheminformatics &
Combinatorial Chemistry
(GNKPSBIMJ1P503)

**Academic
Year:** 2024-25



SGCP's

Guru Nanak Khalsa College

of Arts, Science & Commerce (Autonomous)

DEPARTMENT OF BIOINFORMATICS

CERTIFICATE

This is to certify that Ms. Prarthi Hrishit Kothari (Roll No: 115) of M. Sc. Bioinformatics (Part II) has satisfactorily completed the practical for Mandatory Paper I: Biophysical techniques, IPR, Cheminformatics & Combinatorial Chemistry (GNKPSBIMJ1P503) for Semester III course prescribed by the University of Mumbai during the academic year 2024-2025.

Teacher-in-Charge
(Signature)

Head of Department
(Signature)

External Examiner
(Signature)

INDEX

SR. NO.	WEBLEM	DATE	PAGE NO.	SIGN
1	Importance of 3D structures and Methods of Generation from 1D and 2D Representations.	25/07/2024	1	
1(A)	To retrieve 2D/3D structure for ‘Adenosine’ (PubChem ID: 60961) using PubChem database and apply filters as follows: 1. To refine search using substructure, exact and similar structure approach. 2. To screen structure based on chemical property approaches (Physiochemical properties).	25/07/2024	5	
2	Introduction to Chemical Structures (1D, 2D & 3D) Drawing using BIOVIA Draw Software and File Conversion using Open Babel Tool	26/07/2024	14	
2(A)	To perform drawing/editing/manipulation of 2D and 3D structure using BIOVIA Draw Software for query ‘Adenosine’ (PubChem ID: 60961).	26/07/2024	26	
2(B)	To understand and convert various structural file format along with storage and retrieval methods using Open Babel Tool.	26/07/2024	36	
3	Introduction to Balloon Software	08/08/2024	42	
3(A)	To get conformers for the query ‘Imatinib’ (PubChem ID: 5291) using Balloon software.	08/08/2024	46	
4	Introduction to Pharmacophore mapping with ZINCPharmer mapping and Swiss PDB viewer	29/08/2024	54	
4(A)	To generate and analyze pharmacophore map for the query of ‘Acyclovir’ (PubChem ID: 135398513) using ZINCPharmer Software.	29/08/2024	59	

WEBLEM 1

Importance Of 3D Structures And Methods Of Generations From 1D And 2D Representations

INTRODUCTION:

Structural bioinformatics is a subfield of bioinformatics concerned with the analysis and prediction of biological macromolecules such as proteins, RNA, and DNA in three dimensions. It works with both empirically solved structures and computer models to make generalizations about macromolecular 3D structures, such as comparisons of overall folds and local motifs, principles of molecular folding, evolution, binding interactions, and structure/function correlations. Structural bioinformatics is a subset of computational structural biology, and the term structural has the same meaning as structural biology. The major goal of structural bioinformatics is to develop new ways for analyzing biological macromolecular data in order to address biological problems and discover new information.

Structural bioinformatics comprises data resources, algorithms, and tools for investigating, analyzing, predicting, and interpreting bio-macromolecular structures. Structural bioinformatics focuses on relationships between structures based on their spatial coordinates. As a result, established fields of bioinformatics can better evaluate the primary structure. The visualization of protein structures is a critical issue in structural bioinformatics. It enables users to view static or dynamic representations of molecules, as well as the detection of interactions that can be used to infer molecular mechanisms. Several key reasons why studying structural bioinformatics is important are Understanding structure-function relationship, protein-protein structures, drug design and discovery, etc. Structural bioinformatics serves as a bridge between experimental data and computational models. By combining results from techniques like X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy with computational analyses, researchers can gain a comprehensive understanding of biomolecular structures. This integration enhances the accuracy of structural predictions and provides a more robust framework for biological interpretation.

There are various types of structures like:

- 1. 1D (one dimensional):** One-dimensional (1D) chemical structures refer to molecular representations that focus on the linear arrangement of atoms and their connectivity without emphasizing three-dimensional spatial geometry. These representations are essential in various fields of chemistry and materials science. 1D chemical structures provide a foundational method for characterizing molecules, focusing on their atomic composition and bonding characteristics, which can be utilized in various scientific and industrial applications.



Fig 1: 1D structure of Aspirin

- 2. 2D (two dimensional):** 2D chemical structures represent molecules in a flat, two-dimensional format. They typically depict atoms as symbols (e.g., C for carbon, O for oxygen) and use lines to represent bonds between them. Common forms of 2D representations include Lewis structures, which show all atoms and bonds explicitly, and bond-line notation, which simplifies structures by omitting hydrogen atoms and representing carbon atoms at the vertices of lines. A common type of 2D representation is the structural formula, which shows how atoms are connected within the molecule. It can depict single, double, and triple bonds, and may also indicate the presence of functional groups.

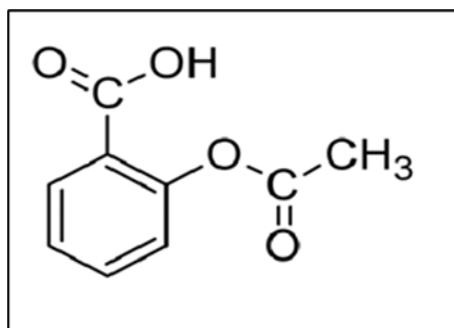


Fig 2: 2D structure of Aspirin

- 3. 3D (three dimensional):** 3D structures provide the exact coordinates of each atom in the molecule, typically in Cartesian (x, y, z) format. This allows for a complete description of the molecular geometry. The distances between bonded atoms are specified, providing information about the strength and nature of the chemical bonds. The angles between bonds are also defined, which is crucial for understanding the overall shape of the molecule and predicting its reactivity and interactions. 3D structures can be visualized using various molecular models, such as ball-and-stick models, where atoms are represented as spheres and bonds as sticks, or space-filling models, where atoms are depicted as spheres with sizes proportional to their atomic radii.

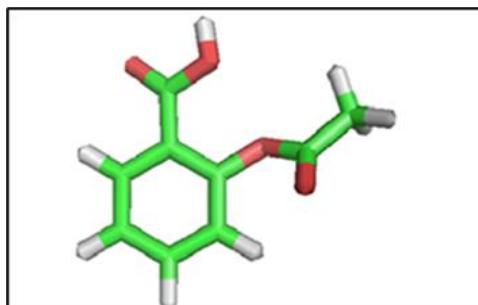


Fig 3: 3D structure of Aspirin

There are several types of studies that can be conducted using 3D chemical structures like:

1. Structure-Function Relationship Studies

Investigating how the 3D arrangement of atoms affects the binding interactions between ligands and their target proteins, which is essential for understanding drug efficacy and specificity.

2. Drug Discovery and Development

Utilizing 3D structures to design and optimize new drug candidates by predicting how structural modifications affect binding affinity and selectivity.

3. Computational and Experimental Techniques

Employing techniques like X-ray crystallography and NMR spectroscopy to elucidate the 3D structures of biomolecules, which is crucial for understanding their functions and mechanisms.

The reason behind preferring 3D structures over 1D and 2D is that 2D representations provide valuable information about connectivity and functional groups while 3D structures are more informative due to their ability to capture spatial information critical for understanding molecular interactions, predicting properties, guiding drug discovery, and exploring complex biological systems.

Applications:

1. The 3D structure of a protein can be revealed through X-ray crystallography.
2. To build an electron density map, extrapolation techniques such as Multiwavelength anomalous dispersion can be employed, which uses the location of selenium atoms as a reference to infer the rest of the structure. The electron density map is used to build a standard ball-and-stick model.
3. **NMR spectroscopy data analysis:** Nuclear magnetic resonance spectroscopy experiments provide 2D (or higher) data, with each peak corresponding to a chemical group inside the sample. To convert spectra into 3D structures, optimization methods are applied.
4. **Structure-based drug design:** Using 3D structures to guide the search for novel drug leads through molecular docking and machine learning models.
5. **Quantitative structure-activity relationship (QSAR) studies:** Establishing relationships between 3D descriptors (e.g., pharmacophores, surface properties) and biological activity.
6. **Drug discovery:** 2D Graphs are used primarily for initial screenings and structure-activity relationship (SAR) studies. They are beneficial for generating molecular fingerprints and performing virtual screening of compound libraries. 3D Structures are essential for structure-based drug design, including molecular docking studies that simulate how small molecules bind to target proteins. They enable more accurate predictions of binding affinities and interactions, which are critical for identifying potential drug candidates.

PubChem Database:

PubChem is a public chemical information resource, developed and maintained by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), an institute within the U.S. National Institutes of Health (NIH). It collects chemical substance descriptions and their biological activities from more than 500 data sources and disseminates these data to the public free of charge. Since the launch in 2004 as a component of the NIH Molecular Libraries Roadmap Initiatives, PubChem has been a key information resource for biomedical research communities in many areas such as cheminformatics, chemical biology, medicinal chemistry, and drug discovery.

PubChem contains various types of chemical information, including 2-D and 3-D structures, chemical and physical properties, bioactivity data, pharmacology, toxicology, drug target, metabolism, safety and handling, relevant patents and scientific papers, etc. While the majority of PubChem's records are about small molecules, it also contains information on a broad range of chemical entities, including siRNAs, miRNAs, carbohydrates, lipids, peptides, chemically modified macromolecules, and many others. These data are provided by various contributors, including government agencies, university labs, pharmaceutical companies, chemical vendors, publishers and a number of chemical biology resources. The data in PubChem are organised into three interlinked databases: Substance (as of writing more than 286 million substance descriptions), Compound (over 111 million unique chemical structures) and BioAssay (1.2 million biological assays covering more than 10,000 target protein sequences). Most of the chemical databases discussed in this paper also contribute their data to PubChem.

REFERENCES:

1. Ambrose, G. O. (2021, December 24). Overview of Structural Bioinformatics and its Applications. Longdom. <https://www.longdom.org/open-access/overview-of-structural-bioinformatics-and-its-applications-89200.html>
 2. Berini, P. (2014). Amplification and Lasing with Surface Plasmon Polaritons. In Handbook of surface science (pp. 309–328). <https://doi.org/10.1016/b978-0-444-59526-3.00010-0>
 3. Structural bioinformatics. (n.d.). 3wMedia Eindhoven. <https://www.carimmaastricht.nl/research/facilities/structural-bioinformatics>
 4. Evaluation of aspirin structure from 1D to 2D to 3D. (n.d.). ResearchGate. https://www.researchgate.net/figure/Evaluation-of-aspirin-structure-from-1D-to-2D-to-3D_fig1_272411549
-

WEBLEM 1(A)
PubChem
(URL: <https://pubchem.ncbi.nlm.nih.gov/>)

AIM:

To retrieve 2D / 3D structure for ‘Adenosine’ (PubChem CID: 60961) using PubChem database and apply the filters as follows:

1. To refine the search using exact, substructure and similar structure approach
2. To screen structures based on chemical property approaches (Physicochemical properties)

INTRODUCTION:

PubChem database

PubChem is a public repository for information on chemical substances and their biological activities, launched in 2004 as a component of the Molecular Libraries Roadmap Initiatives of the US National Institutes of Health (NIH). For the past 11 years, PubChem has grown to a sizable system, serving as a chemical information resource for the scientific research community.

PubChem consists of three inter-linked databases, Substance, Compound and BioAssay. The Substance database contains chemical information deposited by individual data contributors to PubChem, and the Compound database stores unique chemical structures extracted from the Substance database. Biological activity data of chemical substances tested in assay experiments are contained in the BioAssay database.

PubChem contains one of the largest corpus of publicly available chemical information. It has more than 157 million depositor-provided chemical substance descriptions, 60 million unique chemical structures and 1 million biological assay descriptions, covering about 10 thousand unique protein target sequences. PubChem has rapidly grown to a key chemical information resource that serves scientific communities in many areas such as cheminformatics, chemical biology, medicinal chemistry and drug discovery. PubChem contains a substantial amount of literature-derived bioactivity data of chemical substances, manually extracted from tens of thousands of scientific articles by data contributors such as ChEMBL and BindingDB. In addition, through integration with data from DrugBank, the Hazardous Substances Data Bank and other databases, the annotation of chemical records includes pharmacology, drug target information, toxicology, safety and handling information.

Adenosine

Adenosine is an important organic compound and a nucleoside that plays a crucial role in various physiological processes. It is composed of two main components: the nitrogenous base adenine and the sugar ribose. These components are linked by a β -N9-glycosidic bond, forming the structure of adenosine. The chemical formula of adenosine is $C_{10}H_{13}N_5O_4$, and its structure can be represented as follows:

1. **Adenine:** A purine base that features a fused double-ring structure containing nitrogen atoms.
2. **Ribose:** A five-carbon sugar that is part of the nucleoside structure.

Adenosine is a key player in cellular energy transfer, primarily through its derivatives: adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP). ATP, in particular, is often referred to as the "molecular unit of currency" for energy transfer in cells, as it captures and releases energy necessary for various cellular processes, including muscle contraction and nerve impulse propagation.

In addition to its role in energy metabolism, adenosine also functions as a signaling molecule in the body, influencing processes such as sleep regulation, immune response, and cardiovascular function. It operates through four known receptor subtypes (A1, A2A, A2B, and A3), which are G-protein-coupled receptors that mediate various physiological effects.

METHODOLOGY:

1. Open homepage of PubChem database (URL: <https://pubchem.ncbi.nlm.nih.gov/>)
2. Search for the query 'Adenosine' (PubChem CID: 60961).
3. Exact search for query 'Adenosine' was obtained with compound CID: 60961.
4. Sections like Compounds, Substances, Genes, Proteins, Pathways, Cell Lines, Bioassays, Literature and Patents were obtained.
5. Open one entry (PubChem CID: 60961) from the obtained entries.
6. Use a variety of filters from the table of contents, such as Structure, Names and Identifiers, Chemical and Physical Properties, Chemical Vendors, Literature and Patents to get the desired results.

OBSERVATIONS:

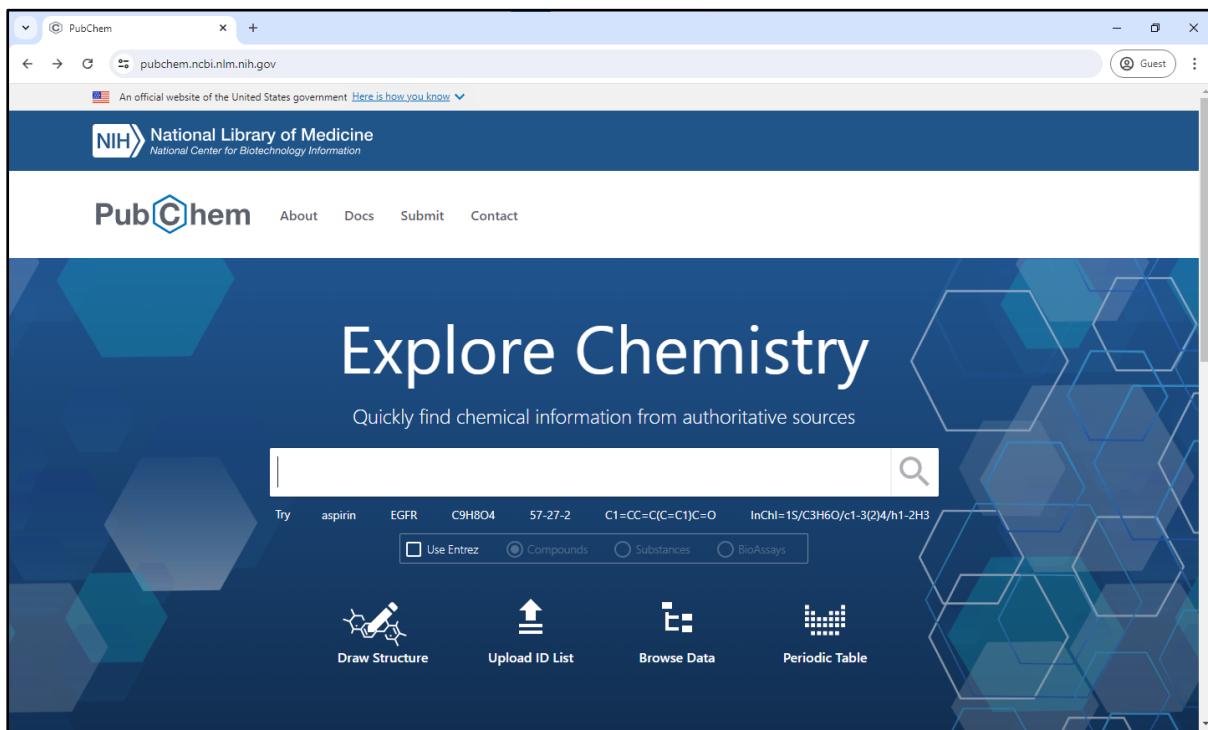


Fig 1: Homepage of PubChem database

A screenshot of the PubChem result page for the query "adenosine". The search bar at the top contains "adenosine". The "BEST MATCH" section shows a detailed card for Adenosine (PubChem CID: 60961). The card includes the chemical structure, compound ID, molecular formula (C10H13N5O4), MW (267.24g/mol), IUPAC Name (2-(6-aminopurin-9-yl)-5-(hydroxymethyl)oxolane-3,4-diol), Isomeric SMILES, InChIKey, InChI, and Create Date (2004-09-16). Below the card are links for Summary, Similar Structures Search, Related Records, and PubMed (MeSH Keyword). At the bottom, there are links for Compounds (4,263), Substances (32,288), Genes (949), Proteins (2,340), Pathways (980), Cell Lines (1), BioAssays (22,930), and Literature (283,366).

Fig 2.a: Result page for query 'Adenosine' (PubChem CID: 60961)

Compounds (4,263)

Substances (32,288) **Genes (949)** **Proteins (2,340)** **Pathways (980)** **Cell Lines (1)** **BioAssays (22,930)** **Literature (283,366)**

Patents (11,860)

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...](#)

4,263 results [Filters](#) SORT BY [Relevance](#)

[Download](#)

[Search in Entrez](#)

ACTIONS ON RESULTS WITH ID TYPE:
Compounds
[Push to Entrez](#)
[Save for Later](#)
[Linked Data Sets](#)

adenosine; 58-61-7; Adenocard; Adenoscan; Adenine riboside; ...
Compound CID: 60961
MF: C₁₀H₁₃N₅O₄ MW: 267.24g/mol
IUPAC Name: (2R,3R,4S,5R)-2-(6-aminopurin-9-yl)-5-(hydroxymethyl)oxolane-3,4-diol
Isomeric SMILES: C1=NC(=C2C(=N1)N(C=N2)[C@H]3[C@@H]([C@H]([C@H](O3)CO)O)O)N
InChIKey: OIRDTCQYFTABQQQ-KQYNXXCUSA-N
InChI: InChI=1S/C10H13N5O4/c11-8-5-9(13-2-12-8)15(3-14-5)10-7(18)6(17)4(1-16)19-10/h2-4,6-7,10,16-18H,1H2,(H2,11,12,13) t4-,6-,7-,10-/m1/s1
Create Date: 2004-09-16

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

5'-adenylic acid; adenosine 5'-monophosphate; Adenosine monophosphate; adenosine phosphate; 61-19-8; ...
Compound CID: 6083
MF: C₁₀H₁₃N₅O₇P MW: 347.22g/mol

Fig 2.b: Result page for query ‘Adenosine’ (PubChem CID: 60961)

PubChem [CID60961 structure](#)

Identity (1) **Similarity (>1,000)** **Substructure (>1,000)** **Superstructure (>1,000)** **3D Similarity (>1,000)** [Settings](#)

Fingerprint Tanimoto-based 2-dimensional similarity search.

Percentage of the database searched: 5%. [Search All](#)

1,000 results (incomplete) [Filters](#) SORT BY [Relevance](#)

[Download](#)

ACTIONS ON RESULTS WITH ID TYPE:
Compounds
[Push to Entrez](#)
[Save for Later](#)
[Linked Data Sets](#)

5'-adenylic acid; adenosine 5'-monophosphate; Adenosine monophosphate; adenosine phosphate; 61-19-8; ...
Compound CID: 6083
MF: C₁₀H₁₃N₅O₇P MW: 347.22g/mol
IUPAC Name: [(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyxolan-2-yl]methyl dihydrogen phosphate
Isomeric SMILES: C1=NC(=C2C(=N1)N(C=N2)[C@H]3[C@@H]([C@H]([C@H](O3)COP(=O)(O)O)O)N
InChIKey: UDMBCSSLTHHNCD-KQYNXXCUSA-N
InChI: InChI=1S/C10H14N5O7P/c11-8-5-9(13-2-12-8)15(3-14-5)10-7(17)6(16)4(22-10)1-21-23(18,19)20/h2-4,6-7,10,16-17H,1H2,(H2,11,12,13) t2-,4-,6-,7-,10-/m1/s1
Create Date: 2004-09-16

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

adenosine; 58-61-7; Adenocard; Adenoscan; Adenine riboside; ...
Compound CID: 60961
MF: C₁₀H₁₃N₅O₄ MW: 267.24g/mol

Fig 3: Result page for similar structure search

PubChem CID: 60961

Structure: 2D, 3D, Crystal

Chemical Safety: Irritant

Molecular Formula: C₁₀H₁₃N₅O₄

Synonyms: adenosine, 58-61-7

CONTENTS:

- 1 Structures
- 2 Biologic Description
- 3 Names and Identifiers
- 4 Chemical and Physical Properties
- 5 Spectral Information
- 6 Related Records
- 7 Chemical Vendors
- 8 Drug and Medication Information
- 9 Food Additives and Ingredients
- 10 Pharmacology and Biochemistry
- 11 Use and Manufacturing
- 12 Identification
- 13 Safety and Hazards
- 14 Toxicity
- 15 Associated Disorders and Diseases
- 16 Literature
- 17 Patents
- 18 Interactions and Pathways
- 19 Biological Test Results
- 20 Taxonomy
- 21 Classification
- 22 Information Sources

Fig 4: Result page after selecting an entry ‘Adenosine’ (PubChem CID: 60961)

PubChem CID: 60961

Structure: 2D, 3D, Crystal

Chemical Safety: Irritant

Molecular Formula: C₁₀H₁₃N₅O₄

Synonyms: adenosine, 58-61-7, Adenocard, Adenoscan, Adenine riboside

Molecular Weight: 267.24 g/mol

CONTENTS:

- 1 Structures
- 2 Biologic Description
- 3 Names and Identifiers
- 4 Chemical and Physical Properties
- 5 Spectral Information
- 6 Related Records
- 7 Chemical Vendors
- 8 Drug and Medication Information
- 9 Food Additives and Ingredients
- 10 Pharmacology and Biochemistry
- 11 Use and Manufacturing
- 12 Identification
- 13 Safety and Hazards
- 14 Toxicity
- 15 Associated Disorders and Diseases
- 16 Literature
- 17 Patents
- 18 Interactions and Pathways
- 19 Biological Test Results
- 20 Taxonomy
- 21 Classification
- 22 Information Sources

Fig 5: Select the category from Table of Contents

Fig 5a: Selecting and viewing information for the category – ‘Structure’

Fig 5b: Selecting and viewing information for the category – ‘Names and Identifiers’

4 Chemical and Physical Properties

4.1 Computed Properties

Property Name	Property Value	Reference
Molecular Weight	267.24 g/mol	Computed by PubChem 2.2 (PubChem release 2021.10.14)
XLogP3	-1.1	Computed by XLogP3 3.0 (PubChem release 2021.10.14)
Hydrogen Bond Donor Count	4	Computed by Cactus 3.4.8.18 (PubChem release 2021.10.14)
Hydrogen Bond Acceptor Count	8	Computed by Cactus 3.4.8.18 (PubChem release 2021.10.14)
Rotatable Bond Count	2	Computed by Cactus 3.4.8.18 (PubChem release 2021.10.14)
Exact Mass	267.09675391 g/mol	Computed by PubChem 2.2 (PubChem release 2021.10.14)
Monoisotopic Mass	267.09675391 g/mol	Computed by PubChem 2.2 (PubChem release 2021.10.14)
Topological Polar Surface Area	140 Å ²	Computed by Cactus 3.4.8.18 (PubChem release 2021.10.14)
Heavy Atom Count	19	Computed by PubChem
Formal Charge	0	Computed by PubChem
Complexity	335	Computed by Cactus 3.4.8.18 (PubChem release 2021.10.14)
Isotope Atom Count	0	Computed by PubChem
Defined Atom Stereocenter Count	4	Computed by PubChem
Undefined Atom Stereocenter Count	0	Computed by PubChem
Defined Bond Stereocenter Count	0	Computed by PubChem

Fig 5c: Selecting and viewing information for the category – ‘Chemical and Physical Properties’

7 Chemical Vendors

64 vendors

AKos Consulting & Solutions PubChem SID: 152028013	Purchasable Chemical: AKOS015888594
Sigma-Aldrich PubChem SID: 329747369 PubChem SID: 329748939 PubChem SID: 329770589	Purchasable Chemical: O1B900_SIGMA Purchasable Chemical: 1012123_USP Purchasable Chemical: A0230200_SIAL
Alfa Chemistry PubChem SID: 347758307 PubChem SID: 486439324 PubChem SID: 486477164	Purchasable Chemical: 5536-17-4 Purchasable Chemical: ACM0325 Purchasable Chemical: ACM1547824
AstaTech, Inc. PubChem SID: 45872624	Purchasable Chemical: AC7861
Enamine PubChem SID: 473968224 PubChem SID: 475480180	Purchasable Chemical: EN300-100931 Purchasable Chemical: Z1341543331
Ambeed PubChem SID: 376152558	Purchasable Chemical: A393785
J&H Chemical Co.,Ltd PubChem SID: 469423731 PubChem SID: 46905098	Purchasable Chemical: JH704931 Purchasable Chemical: JH865663
BenchChem PubChem SID: 445633186	Purchasable Chemical: B011128

Fig 5d: Selecting and viewing information for the category – ‘Chemical Vendors’

PubChem Adenosine (Compound)

16 Literature

16.1 Consolidated References

90,958 Items [Download](#)

Search SORT BY Publication Date - Most Recent

First analytical confirmation of drug-induced crystal nephropathy in felines caused by GS-441524, the active metabolite of Remdesivir
Publication Name: Journal of Pharmaceutical and Biomedical Analysis
Publication Date: 2024-09-01
PMID: 38823223 DOI: 10.1016/j.jpba.2024.116248

Detecting 2'-5'-adenosine linked nucleic acids via acylation of secondary hydroxy functionality
Publication Name: Bioorganic & Medicinal Chemistry Letters
Publication Date: 2024-09-01
PMID: 38857849 DOI: 10.1016/j.bmcl.2024.129847

Intracoronary adenosine compared with adrenaline and verapamil in the treatment of no-reflow phenomenon following primary PCI in STEMI patients
Publication Name: International Journal of Cardiology
Publication Date: 2024-09-01
PMID: 38844092 DOI: 10.1016/j.ijcard.2024.132228

Compression Promotes the Osteogenic Differentiation of Human Periodontal Ligament Stem Cells by Regulating METTL14-mediated lGFI
Publication Name: Current Stem Cell Research & Therapy
Publication Date: 2024-09-01
PMID: 38279741 DOI: 10.2174/011574888x244047231012103752

Small molecules that regulate the N6-methyladenosine RNA modification as potential anti-cancer agents

[Cite](#) [Download](#)

CONTENTS

- 11 Use and manufacturing
- 12 Identification
- 13 Safety and Hazards
- 14 Toxicity
- 15 Associated Disorders and Diseases
- 16 Literature**
 - 16.1 Consolidated References
 - 16.2 NLM Curated PubMed Citations
 - 16.3 Springer Nature References
 - 16.4 Thieme References
 - 16.5 Wiley References
 - 16.6 Nature Journal References
 - 16.7 Chemical Co-Occurrences in Literature
 - 16.8 Chemical-Gene Co-Occurrences in Literature
 - 16.9 Chemical-Disease Co-Occurrences in Literature
- 17 Patents
- 18 Interactions and Pathways
- 19 Biological Test Results
- 20 Taxonomy
- 21 Classification
- 22 Information Sources

Fig 5e: Selecting and viewing information for the category – ‘Literature’

PubChem Adenosine (Compound)

17 Patents

US5731296 [DrugBank](#)

17.1 Depositor-Supplied Patent Identifiers

158,317 Items [Download](#)

Search SORT BY Priority Date - Most Recent

A kind of quality control method of diquafofol tetrasodium intermediate 5'-uridine disodium
Publication Number: CN-115144509-B
Priority Date: 2022-09-05 Grant Date: 2022-11-29

A kind of method for preparing 2'-deoxy-2'-fluoro- β -D-arabinose adenosine analog
Publication Number: CN-115109814-A
Priority Date: 2022-08-26

A method for preparing 2'-deoxy-2'-fluoro- β -D-arabinoadenosine analogs
Publication Number: CN-115109814-B
Priority Date: 2022-08-26 Grant Date: 2022-11-29

A kind of extraction method of platelet exosome
Publication Number: CN-115029312-A
Priority Date: 2022-08-11

A kind of extraction method of platelet exosome
Publication Number: CN-115029312-B
Priority Date: 2022-08-11 Grant Date: 2022-11-18

[Cite](#) [Download](#)

CONTENTS

- 8 Drug and Medication Information
- 9 Food Additives and Ingredients
- 10 Pharmacology and Biochemistry
- 11 Use and Manufacturing
- 12 Identification
- 13 Safety and Hazards
- 14 Toxicity
- 15 Associated Disorders and Diseases
- 16 Literature**
- 17 Patents**
 - 17.1 Depositor-Supplied Patent Identifiers
 - 17.2 WIPO PATENTSCOPE
 - 17.3 Chemical Co-Occurrences in Patents
 - 17.4 Chemical-Disease Co-Occurrences in Patents
 - 17.5 Chemical-Gene Co-Occurrences in Patents
- 18 Interactions and Pathways
- 19 Biological Test Results
- 20 Taxonomy
- 21 Classification
- 22 Information Sources

Fig 5f: Selecting and viewing information for the category – ‘Patents’

RESULTS:

PubChem database was explored and the 2D/3D structure, similar structure, exact structure and substructure for the query ‘Adenosine’ (PubChem CID: 60961) was retrieved and studied. One best match was found which is exact search followed by 4263 hits with different sections like Compounds, Substances, Genes, Proteins, Pathways, Cell Lines, Bioassays, Literature and Patents. More than 1000 substructures and similar structures were obtained. Various categories were selected from the table of contents like Structure, Names and Identifiers, Chemical and Physical Properties, Chemical Vendors, Literature and Patents.

CONCLUSION:

The exact structure, substructures and similar structures were studied and observed based on their physiochemical properties by exploring PubChem database for query ‘Adenosine’ (PubChem CID: 60961).

REFERENCES:

1. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. Nucleic Acids Res. 2016 Jan 4;44(D1):D1202-13. doi: 10.1093/nar/gkv951. Epub 2015 Sep 22. PMID: 26400175; PMCID: PMC4702940.
 2. Team, E. W. (n.d.). adenosine (CHEBI:16335). <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI%3A16335>
 3. Mahler, G. S. (1998). Adenosine. In *Analytical profiles of drug substances and excipients* (pp. 1–37). [https://doi.org/10.1016/s0099-5428\(08\)60751-0](https://doi.org/10.1016/s0099-5428(08)60751-0)
-

WEBLEM 2

Introduction to Chemical Structures (1D, 2D and 3D) Drawing using BIOVIA Draw Software and File Conversion using OpenBabel Tool

INTRODUCTION:

1. BIOVIA Draw Software

BIOVIA Draw is a sophisticated software application designed for scientists to efficiently draw and edit complex biological structures, chemical reactions, and molecules. Developed by Dassault Systèmes, this tool is essential for scientists and researchers in various fields, enabling efficient collaborative searching, viewing, communicating, and archiving of scientific information. This tool enhances the collaborative searching, viewing, and archiving of scientific information, making it a vital resource in various research and development settings. Available in both 32-bit and 64-bit versions, BIOVIA Draw ensures compatibility with a wide range of systems, catering to diverse user needs and preferences.

Key Features of BIOVIA Draw

1. BIOVIA Draw stands out for its efficient chemical structure drawing capabilities. It allows scientists to generate structures using common chemical string notations like SMILES, IUPAC Name, InChI, and HELM through the "Structure Resolver" and "Generate Structure from Text" options.
2. The all-purpose drawing tool enables continuous drawing of bonds, pulling out rings, and adding atoms, making the process intuitive and user-friendly.
3. With drag-and-drop functionality, users can easily reuse commonly used structures and chemical abbreviations by placing them on the toolbar for quick access.
4. Right-click options provide easy access to atom, bond, and fragment properties, as well as query options, enhancing the user's control over the drawing process.
5. The software supports multiple undo/redo actions, allowing users to quickly retrace their steps and correct mistakes.
6. BIOVIA Draw also includes robust annotation tools, enabling the clear and detailed annotation of reaction schemes with text, color, and various arrow styles.

Integration and Customization

BIOVIA Draw is designed for flexible integration with various enterprise software applications. It supports integration with custom Java® and .NET applications, as well as other BIOVIA products such as BIOVIA Pipeline Pilot, BIOVIA Insight, Insight for Excel, BIOVIA Registration, BIOVIA DiscoveryGate®, BIOVIA Workbook, BIOVIA Notebook, BIOVIA Isentris®, and BIOVIA CISPro.

Users can modify the chemical drawing look-and-feel according to their organization's needs by configuring the application's XML settings.

Additionally, custom add-ins can be created to enhance the drawing experience, and the software can be integrated with existing desktop and web applications for queries and browsing.

BIOVIA Draw for Scientists

BIOVIA Draw offers scientists a quick and efficient way to draw structures and queries. The software includes a structure converter that can convert structures to IUPAC names and vice versa, including support for Enhanced Stereochemistry. It also supports conversions between structures and canonical SMILES, InChI names, and InChI keys. Users can create structures using the "Structure Resolver" and search common structure identifiers like CAS numbers, names, MDL Numbers, SMILES, and InChI via web services provided by online chemical databases such as NCI/CADD, PubChem, and BIOVIA DiscoveryGate® Available Chemicals Directory. This feature allows users to retrieve corresponding structures without having to draw them manually.

BIOVIA Draw supports the creation and editing of haptic bonds, polymers, formulations, and mixtures (Sgroups). It also includes tools for creating and editing Rgroup (Markush) queries with built-in Rgroup query logic. Additionally, the software enables users to create and edit 3D queries, supporting all 3D query features, including 3D rotation. Users can create SDFFiles from molfiles or canvas fragments, open SDFFiles to view contents, check availability against BIOVIA DiscoveryGate® Available Chemicals Directory, calculate properties, and search within the open file. The software also provides options to export or report on the results and calculations.

BIOVIA Draw offers various "Copy as" and "Paste as" options, including Molfile, Sketch String, IUPAC, SMILES, InChI (key or string), NEMA, Chime, HELM, Sequence, Bitmap, Metafile, and Text. Users can customize symbols, including composite symbols, and choose from a large library of protecting group templates. The software has improved chemical recognition of axial, tetrahedral, and geometric stereogenic centers. Users can create and edit ISIS-compatible sketches and edit legacy sketches with improved cleaning. The software also supports opening Chemdraw CDX files.

BIOVIA Draw allows users to calculate properties on structures as they draw, including AlogP, Polar Surface Area, Hydrogen and Stereo counts, composition, weights, and formulas. The software can connect to the BIOVIA Pipeline Pilot Chemistry Collection for further analysis and data cleaning via a variety of prebuilt workflows and functions, such as Clean, Calculators, and Manipulators. This connectivity enhances the utility of BIOVIA Draw, making it a powerful tool for scientific research and analysis.

BIOVIA Draw for Scientists Working with Biologics

BIOVIA Draw supports scientists working with biologics by enabling the drawing, registration, searching, and reporting on chemically modified peptide or nucleotide sequences. Users can create 1- and 3-letter peptide, DNA, or RNA sequences using a dedicated Sequence tool. This tool allows for the creation of nucleotides with any combination of natural or custom ribose, base, and phosphate building blocks. It also supports drawing crossing bonds, disulfide bridges, and attaching side-chain protecting groups.

The software can import and convert text, XHELM, HELM, FASTA, Swiss-Prot, PDB, and EMBL files into chemically significant sequences. Users can expand residues in a sequence to their full structure to illustrate chemical modifications. Additionally, the HELM Editor can be launched to generate and edit HELM string notations. BIOVIA Draw connects and synchronizes with the BIOVIA Centralized Library of templates and monomers for sequences, facilitating efficient and accurate work with biologics.

BIOVIA Draw for Developers

BIOVIA Draw enables developers to integrate structure drawing and display capabilities into their applications while customizing the software to fit organizational workflows. The updated integration with BIOVIA Pipeline Pilot allows developers and administrators to enhance BIOVIA Draw's functionality using various Pipeline Pilot Collections, such as Chemistry and ADMET. These collections support different types of outputs, including calculated values, molecules, SD files, charts, and tables. Users can run example or custom protocols on their structures from within BIOVIA Draw and select the necessary parameters to obtain the required answers.

Developers can use BIOVIA Draw with web applications, including ASP.NET, MVC, Java EE, and JavaScript, through the new DrawBridge executable. The software supports Microsoft Windows 10 and Office 2016 and 2019 (both 32 and 64-bit versions). Additionally, BIOVIA Draw's look-and-feel can be easily extended with custom add-ins, including ACD/Labs calculators, ACD/Name, bioavailability (Rule of 5), Rgroup/Sgroup/Stereochemistry/Reaction Enumerators, OSRA, TLC creation, and various predictive spectra and lookup capabilities.

Advantages

1. Provides comprehensive tools for drawing and editing complex molecules, chemical reactions, and biological sequences.
2. Enhances collaboration by facilitating the searching, viewing, communication, and archiving of scientific information.
3. Manages complex biological entities, including peptides, oligonucleotides, and oligosaccharides.
4. Includes a biological sequence editor supporting custom residues and linkers.
5. Offers fast and flexible editing with a range of add-ins for various molecular calculations.
6. Allows developers to create and integrate custom add-ins via a documented API.

7. Supports 2D and 3D representations, suitable for life sciences and materials science applications.
8. Customizes protein and nucleic acid sequences and synchronizes with a centralized library.
9. Imports and exports various sequence formats and integrates with other BIOVIA tools.
10. Extends functionality through add-ins from the SWYM community and customizable UI via XML configuration.

BIOVIA Draw is a versatile and powerful tool that significantly enhances the efficiency and accuracy of chemical and biological drawing and editing tasks. Its comprehensive feature set, seamless integration capabilities, and support for a wide range of formats and databases make it an indispensable tool for scientists. By streamlining the drawing process and facilitating collaboration, BIOVIA Draw helps researchers focus on innovation and discovery, pushing the boundaries of scientific knowledge.

INSTALLATION STEPS:

Draw and edit complex molecules

Students, teachers and researchers in an academic settings can—at no charge—download BIOVIA Draw for rapidly drawing chemical structures and chemically intelligent queries.

[Register to download](#) the installer.

Read: BIOVIA Draw Datasheet

DOCUMENT
BIOVIA Draw Datasheet
PDF Document - 6.3Mb

First Name*

Last Name*

Company*

India

City*

Department*

Industry*

I agree to Dassault Systèmes' Terms of Use and Privacy Policy and I am informed that my personal data may be shared with trusted stakeholders for business purposes, including Dassault Systèmes and SOLIDWORKS business partners and events sponsors.

Submit

Fig 1: Fill the Registration Details and Click on ‘Submit’

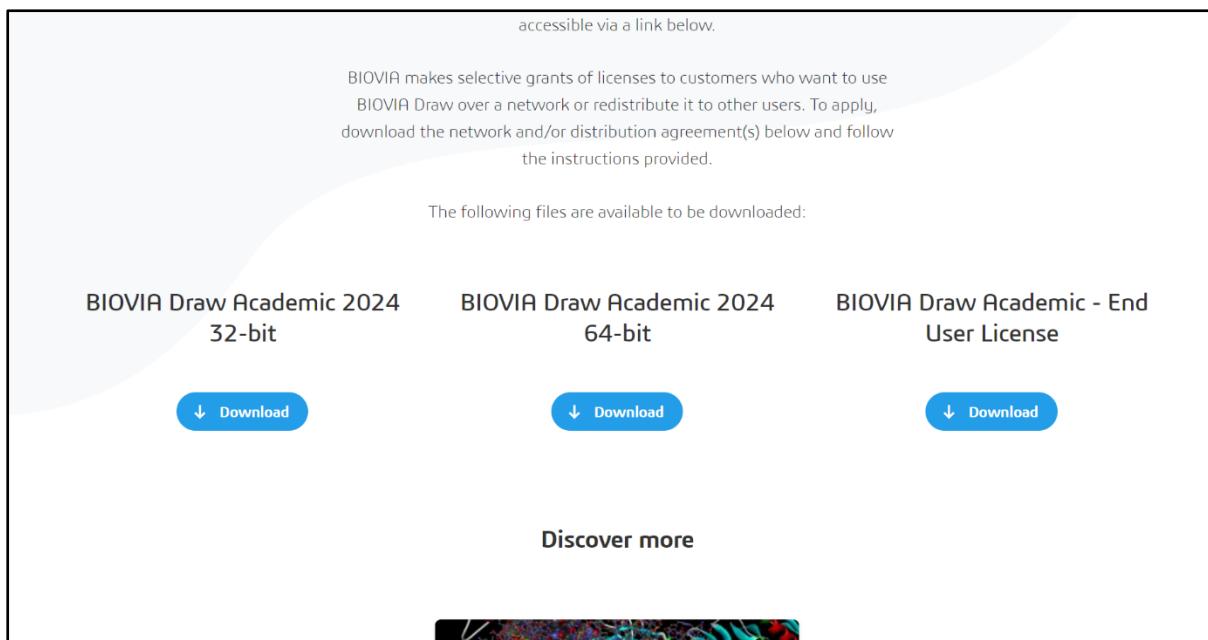


Fig 2: Download the BIOVIA Draw Academic 2024 64-bit

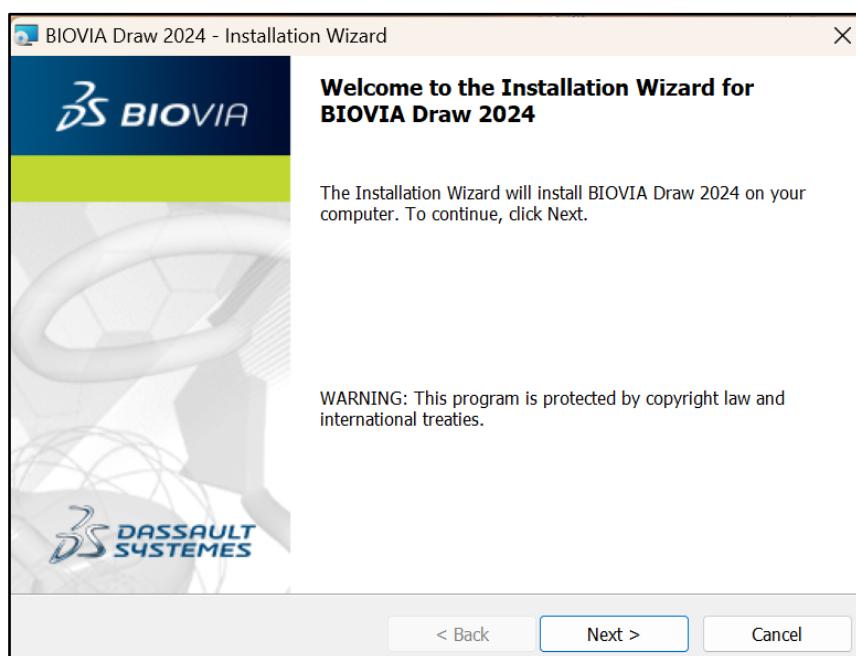


Fig 3: Click on 'Next'

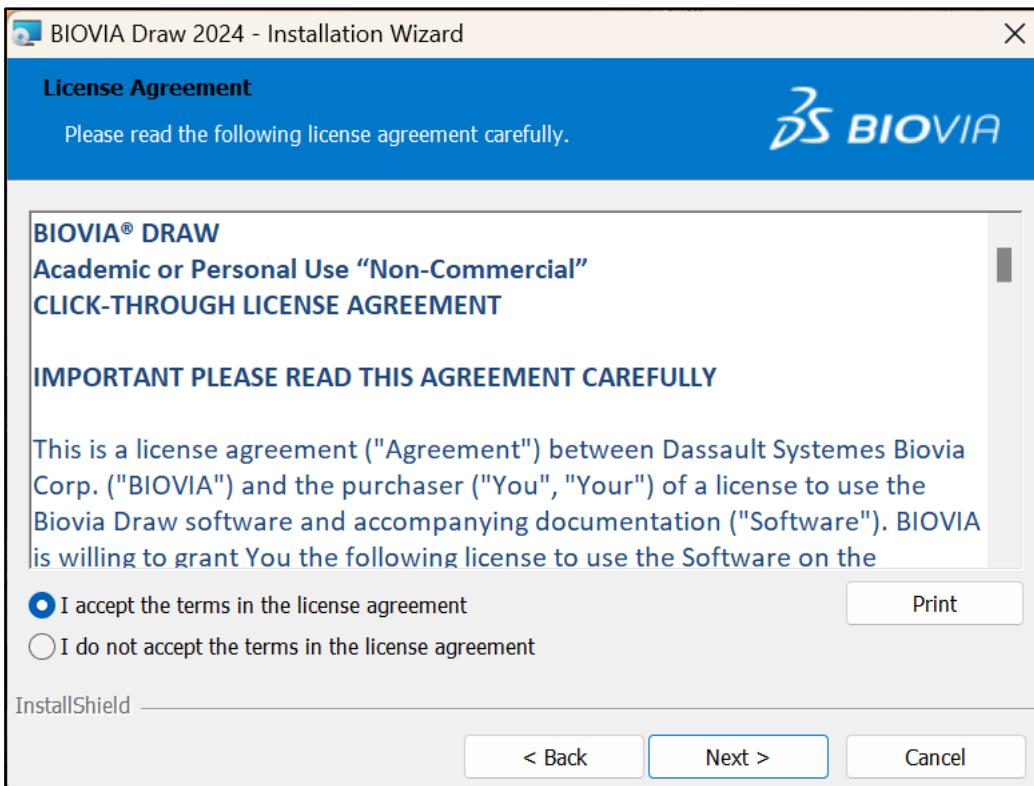


Fig 4: Select the option ‘I accept the terms in the license agreement’ and click on ‘Next’

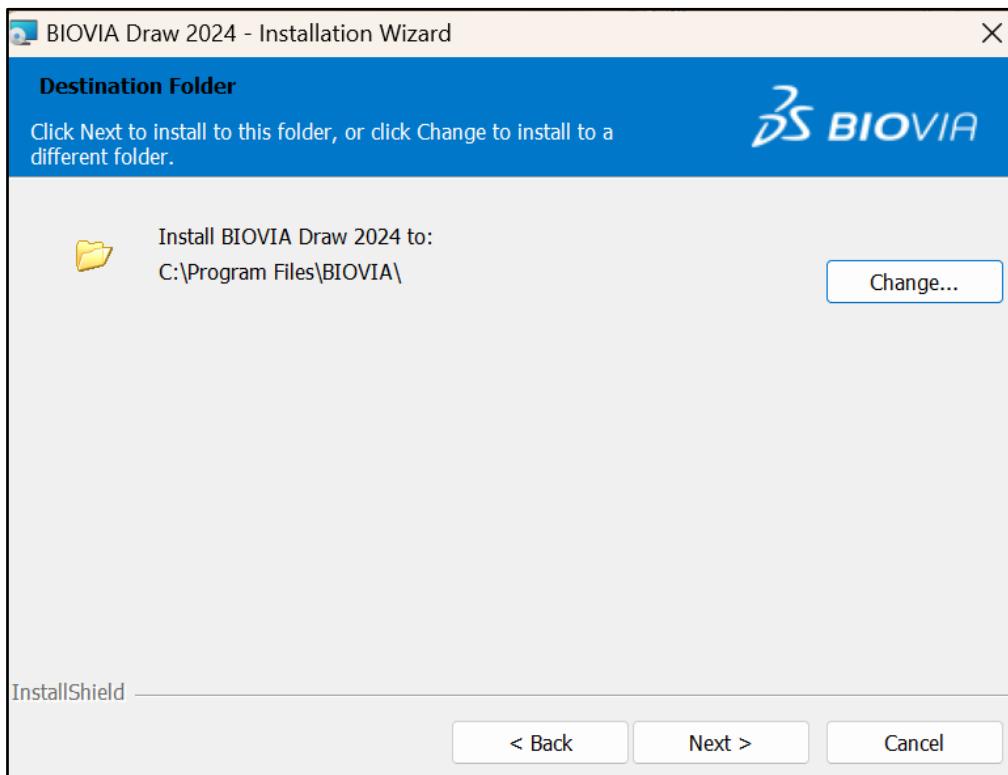


Fig 5: Select the folder to save the software and click on ‘Next’

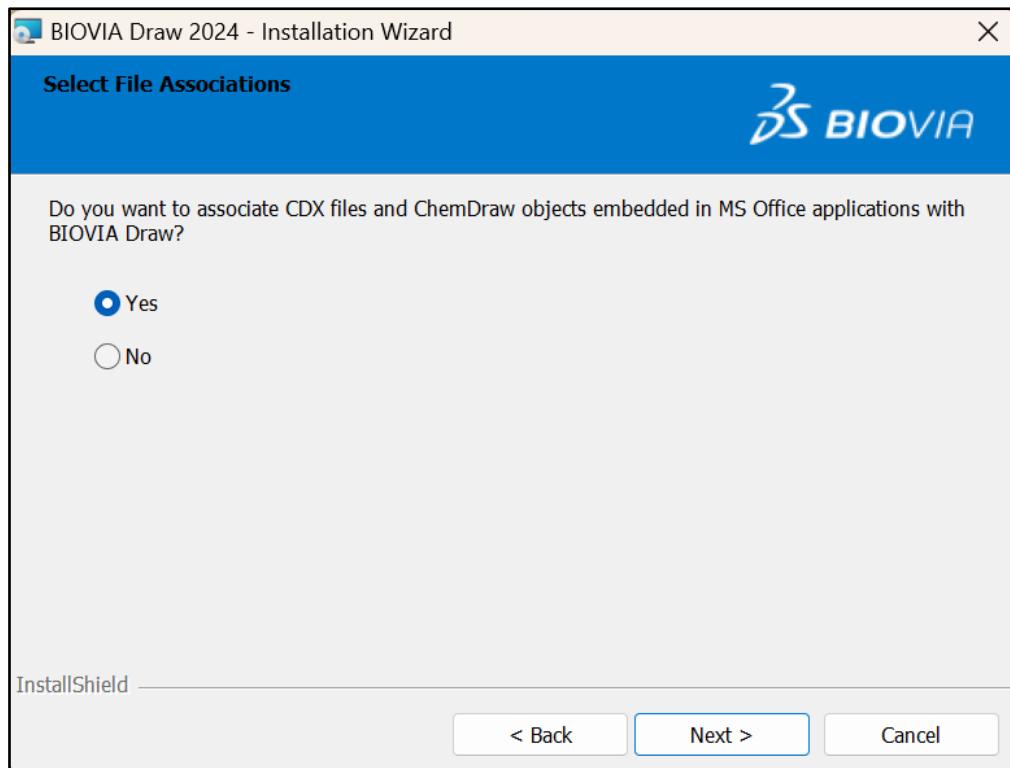


Fig 6: Select ‘Yes’ option and click on ‘Next’

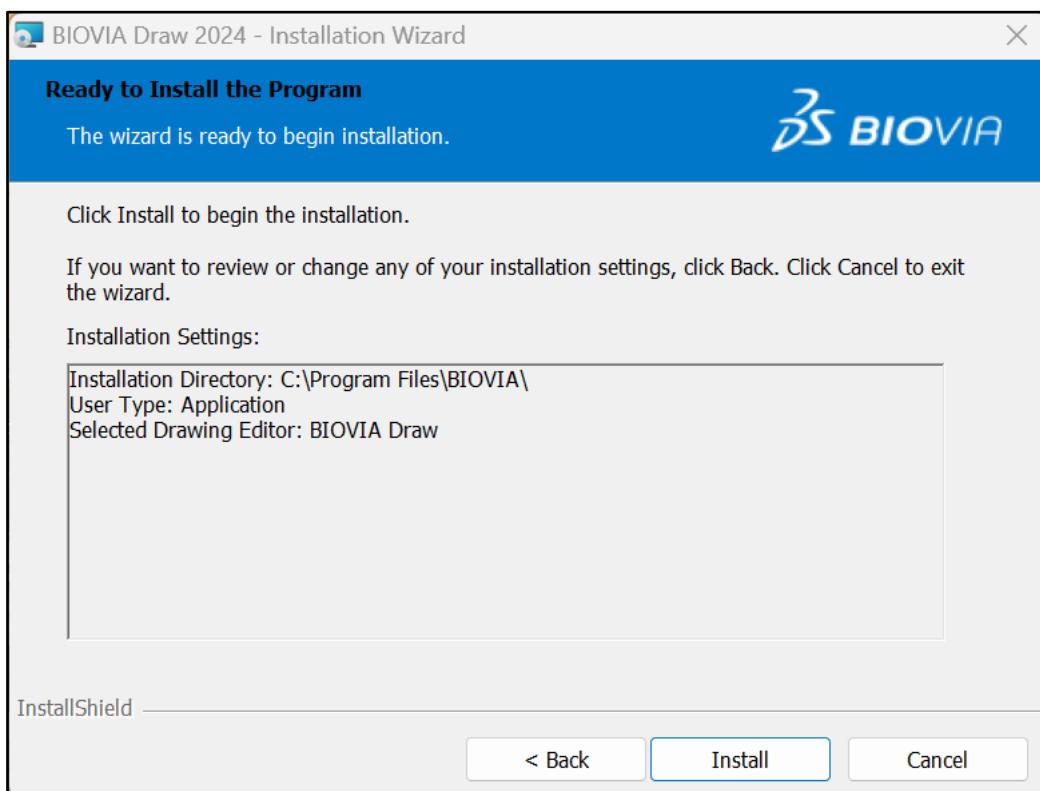


Fig 7: Click on ‘Install’

2. Open Babel Tool

Open Babel is a chemical toolbox designed to speak the many languages of chemical data. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas.

Open Babel supports 111 chemical file formats in total. It can read 82 formats and write 85 formats. These encompass common formats used in cheminformatics (SMILES, InChI, MOL, MOL2), input and output files from a variety of computational chemistry packages (GAMESS, Gaussian, MOPAC), crystallographic file formats (CIF, ShelX), reaction formats (MDL RXN), file formats used by molecular dynamics and docking packages (AutoDock, Amber), formats used by 2D drawing packages (ChemDraw), 3D viewers (Chem3D, Molden) and chemical kinetics and thermodynamics (ChemKin, Thermo). Formats are implemented as "plugins" in Open Babel, which makes it easy for users to contribute new file formats.

File Formats

- 1. MOL file format:** An MDL Molfile is a file format for holding information about the atoms, bonds, connectivity and coordinates of a molecule. The molfile consists of some header information, the Connection Table (CT) containing atom info, then bond connections and types, followed by sections for more complex information.
- 2. .sdf file format:** The structural data file (SDF) belongs to a family of chemical-data file formats. It is based on the MOL-file format, both developed by Molecular Design Limited (MDL), now named BIOVIA (belongs Dassault Systems). The SDF contains information on the chemical structure and associated data of compounds in plain text. The original MOL-file only encoded a single molecule, whereas files in SDF format can encode single or multiple molecules that are then delimited by lines consisting of four dollar signs (\$\$\$\$). SDF files are formatted ASCII files that store information about the positions of the individual atoms (either in 2D or 3D space) that make up the molecule. The data on connectivity and hybridization state are also encoded, although their use is less frequent and often inconsistent.
- 3. SMILES file format:** SMILES (Simplified Molecular Input Line Entry System) is a chemical notation that allows a user to represent a chemical structure in a way that can be used by the computer. SMILES is an easily learned and flexible notation. The SMILES notation requires that you learn a handful of rules. You do not need to worry about ambiguous representations because the software will automatically reorder your entry into a unique SMILES string when necessary
- 4. SMARTS file format:** SMARTS is a language that allows you to specify substructures using rules that are straightforward extensions of SMILES. SMARTS provides a number of primitive symbols describing atomic properties beyond those used in SMILES (atomic symbol, charge, and isotopic specifications).

5. **XYZ Cartesian coordinate system:** The XYZ file format is a simple text-based format used to represent molecular structures in three-dimensional Cartesian coordinates. Each file begins with a line indicating the number of atoms, followed by a comment line. Subsequent lines list each atom's symbol and its X, Y, Z coordinates in angstroms.

INSTALLATION STEPS:

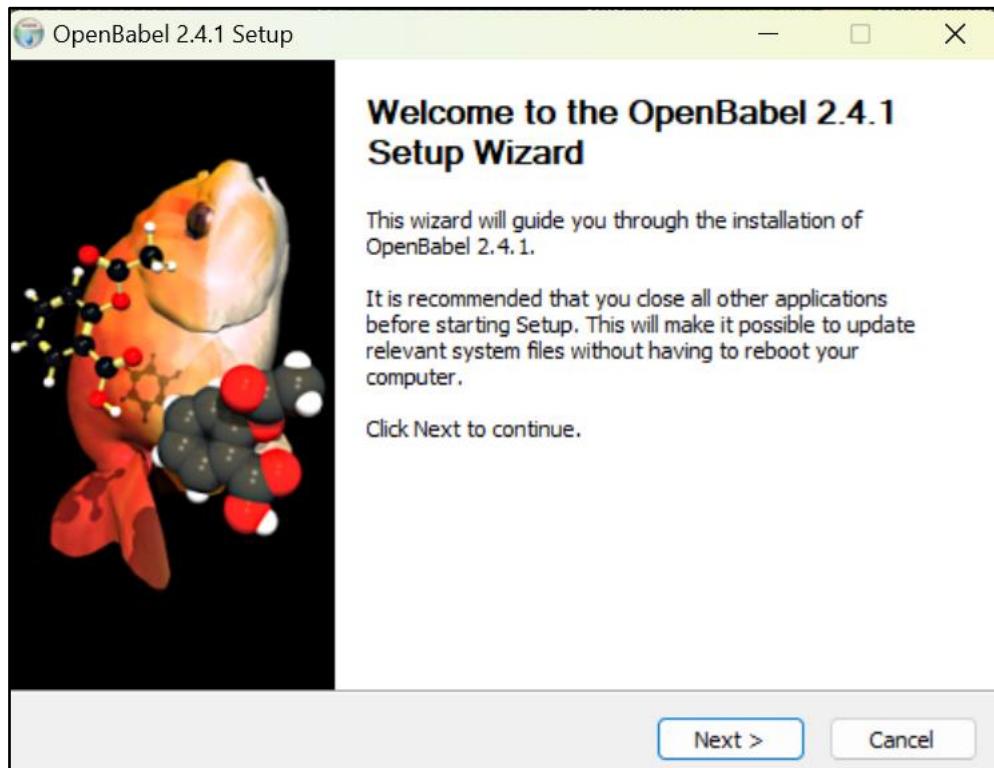


Fig 1: Click on ‘Next’

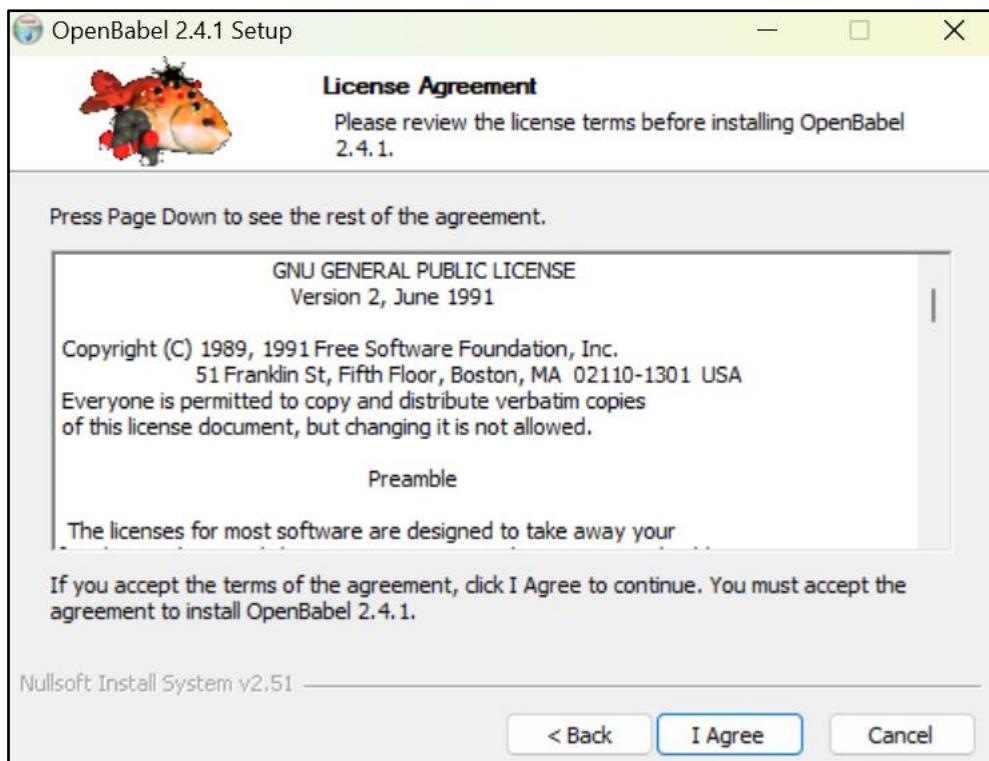


Fig 2: Click on ‘I Agree’

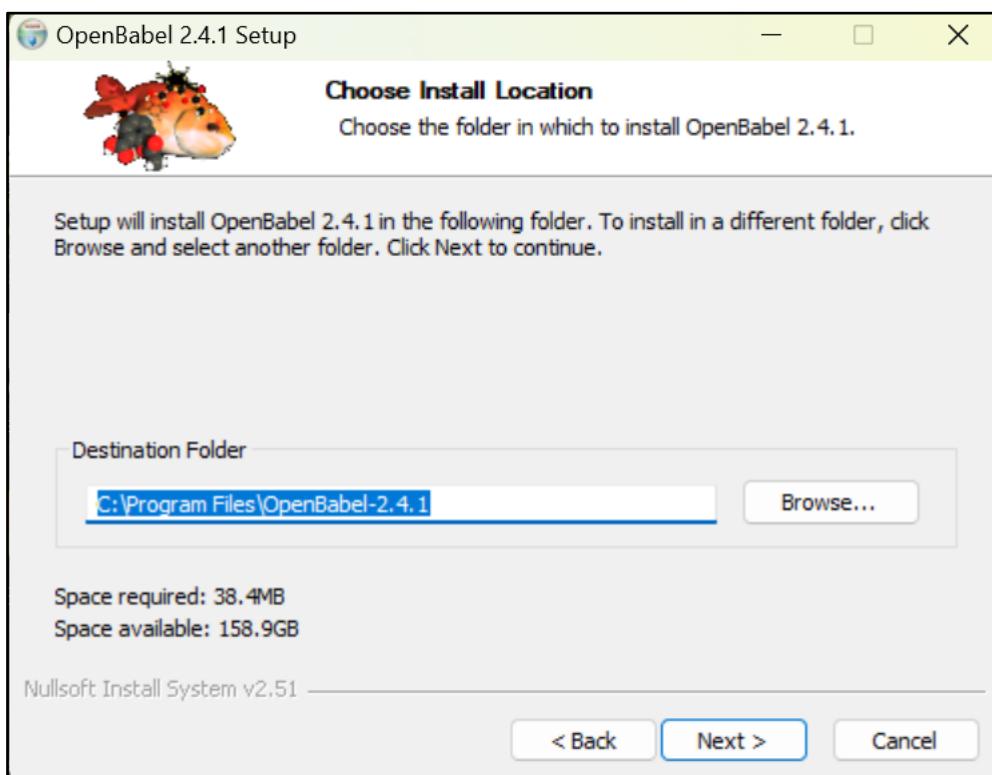


Fig 3: Select the path to save the tool and click on ‘Next’

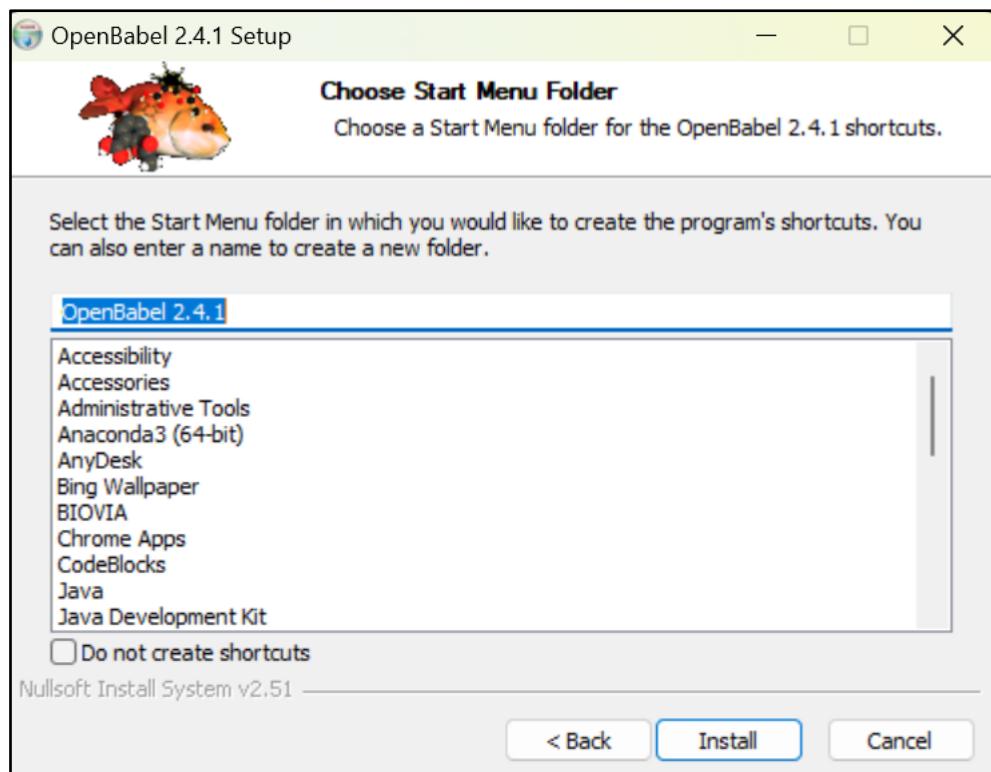


Fig 4: Click on ‘Install’

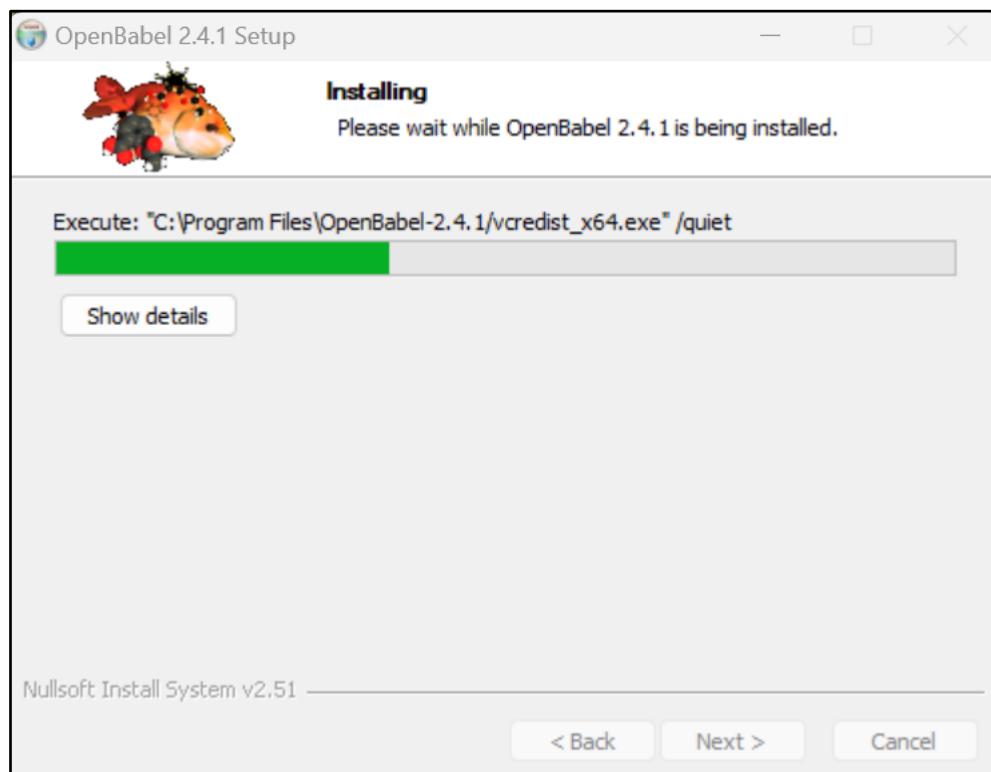


Fig 5: Open Babel Installing

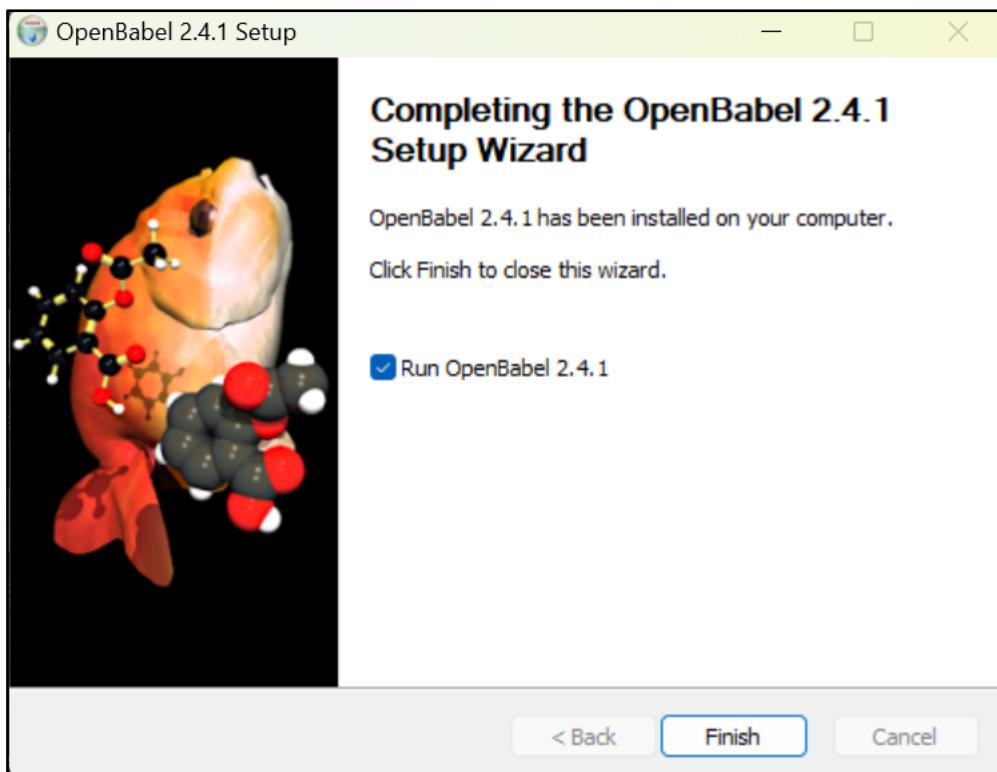


Fig 6: Click on ‘Finish’ and Installation process finishes

REFERENCES:

1. Dassault Systèmes. (2020). BIOVIA Draw Datasheet. In BIOVIA Draw [Datasheet]. <https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/BIOVIA/PDF/biovia-draw-ds.pdf>
2. BIOVIA Draw 2024 SP1 | What’s new? (2024, July 4). [Video]. Dassault Systèmes. <https://www.3ds.com/products/biovia/draw>
3. O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. Journal of cheminformatics, 3, 33. <https://doi.org/10.1186/1758-2946-3-33>
4. User Guide — Open Babel openbabel-3-1-1 documentation. (n.d.). <https://openbabel.org/docs/index.html>
5. OPENBABEL - Chemical file format converter. (n.d.). <https://www.cheminfo.org/Chemistry/Cheminformatics/FormatConverter/index.html>

WEBLEM 2(A)
BIOVIA Draw Software
(URL: <https://discover.3ds.com/biovia-draw-academic>)

AIM:

To perform drawing / editing / manipulation of 2D and 3D structure using BIOVIA Draw Software for the query ‘Adenosine’ (PubChem CID: 60961).

INTRODUCTION:

BIOVIA Draw enables scientists to draw and edit complex molecules, chemical reactions and biological sequences with ease, facilitating the collaborative searching, viewing, communicating, and archiving of scientific information. BIOVIA Draw offers scientist unique capabilities for managing complex biological entities including the ability to register and retrieve peptides, oligonucleotides, and oligosaccharides. Scientists have access to many features including a biological sequence editor that allows the definition of custom residues and linkers, Markush structure tools, and haptic and hydrogen bond tools.

BIOVIA Draw is a powerful software tool designed for scientists to facilitate the drawing and editing of complex molecules, chemical reactions, and biological sequences. Developed by Dassault Systèmes, it enhances collaborative searching, viewing, communicating, and archiving of scientific information.

As an enterprise software application, BIOVIA Draw offers flexible integration with custom Java® and .NET applications as well as integration with BIOVIA Pipeline Pilot, BIOVIA Insight and Insight for Excel, BIOVIA Registration, BIOVIA DiscoveryGate®, BIOVIA Workbook, BIOVIA Notebook, BIOVIA Isentris® and BIOVIA CISPro. Modify the chemical drawing look-and-feel according to the organization’s needs by configuring the applications XML. Create custom add-ins to enhance the scientist’s drawing experience. Integrate with existing desktop applications. Leverage Web applications for query and browsing.

Key Features

1. Structure converter converts structure-to-IUPAC name and IUPAC name-to-structure (now including Enhanced Stereochemistry); structure-to-canonical SMILES and SMILES-to-structure; structure-to-InChI name and InChI name to-structure; and structure-to-InChI key.
2. Create structures using the “Structure Resolver,” search common structure identifiers like CAS number, name, MDL Number, SMILES, InChI and more via webservice provided by the following online chemical databases; NCI/CADD (which currently has 96M unique structures), PubChem (96.5M unique structures), BIOVIA DiscoveryGate® Available Chemicals Directory (12.7M unique structures), or configure others to retrieve the corresponding structure without having to draw it.

3. Create and edit haptic bonds, polymers, formulations and mixtures (Sgroups).
4. Create and edit Rgroup (Markush) queries including built-in Rgroup query logic.
5. Create and edit 3D queries, all 3D query features supported including 3D rotate.
6. Create SDFfiles from molfiles or canvas fragments.
7. Open SDFfiles and view contents, check availability against BIOVIA DiscoveryGate® Available Chemicals Directory, calculate properties, search within the open file, export or report on the results and calculations.
8. Copy as and Paste as options
 - Molfile, Sketch String
 - IUPAC
 - SMILES, InCHi (key or string), NEMA, Chime
 - HELM, Sequence
 - Bitmap, Metafile, Text
9. Customize symbols including composite symbols • Choose from large library of protecting group templates.
10. Take advantage of improved chemical recognition of axial, tetrahedral and geometric stereogenic centers.
11. Create and edit ISIS-compatible sketches, edit legacy sketches with improved cleaning of sketches.
12. Calculate properties on structures as you draw (AlogP, Polar Surface Area, Hydrogen and Stereo counts, composition, weights and formula)
13. Connect to the BIOVIA Pipeline Pilot Chemistry Collection to further analyze and clean data via a variety of prebuilt workflows and functions, including Clean, Calculators, Manipulators and more.

Overall, BIOVIA Draw is a comprehensive tool that supports a wide range of scientific disciplines, making it an essential resource for researchers in chemistry and biology.

Adenosine

Adenosine is an important organic compound and a nucleoside that plays a crucial role in various physiological processes. It is composed of two main components: the nitrogenous base adenine and the sugar ribose. These components are linked by a β -N9-glycosidic bond, forming the structure of adenosine. The chemical formula of adenosine is $C_{10}H_{13}N_5O_4$, and its structure can be represented as follows:

1. **Adenine:** A purine base that features a fused double-ring structure containing nitrogen atoms.
2. **Ribose:** A five-carbon sugar that is part of the nucleoside structure.

Adenosine is a key player in cellular energy transfer, primarily through its derivatives: adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP). ATP, in particular, is often referred to as the "molecular unit of currency" for energy transfer in cells, as it captures and releases energy necessary for various cellular processes, including muscle contraction and nerve impulse propagation.

In addition to its role in energy metabolism, adenosine also functions as a signaling molecule in the body, influencing processes such as sleep regulation, immune response, and cardiovascular function. It operates through four known receptor subtypes (A1, A2A, A2B, and A3), which are G-protein-coupled receptors that mediate various physiological effects.

METHODOLOGY:

To use BIOVIA Draw software effectively, follow these steps:

1. Download the installer from the official BIOVIA website or through an academic resource if you qualify for the free version.
2. Run the installer and follow the on-screen instructions to complete the installation.
3. Open BIOVIA Draw from your applications menu or desktop shortcut.
4. Explore the main workspace, which includes a drawing canvas, toolbars, and menus.
5. Take note of the drawing tools available, such as bond drawing, atom placement, and templates for common structures.
6. Use the all-purpose drawing tool to continuously draw bonds. Click to place atoms and drag to create bonds.
7. Select the atom tool to add different types of atoms to your structure.
8. Drag commonly used structures or fragments from the library onto the canvas for quick access.
9. Add text, colours, and arrows to annotate your chemical structures or reaction schemes for clarity.
10. Save your work in the desired format (e.g: .molfiles, .SDF) for further use or sharing.
11. You can also export structures for inclusion in Microsoft Office documents or presentations.
12. Explore features like the Structure Resolver or Generate structure from text or Generate text from Structure to convert chemical names to structures or vice versa. Use the built-in query options to search for structures using identifiers like CAS numbers or SMILES.
13. By following these steps, you can effectively utilize BIOVIA Draw for your chemical drawing and editing needs, enhancing your research and presentations.

OBSERVATIONS:

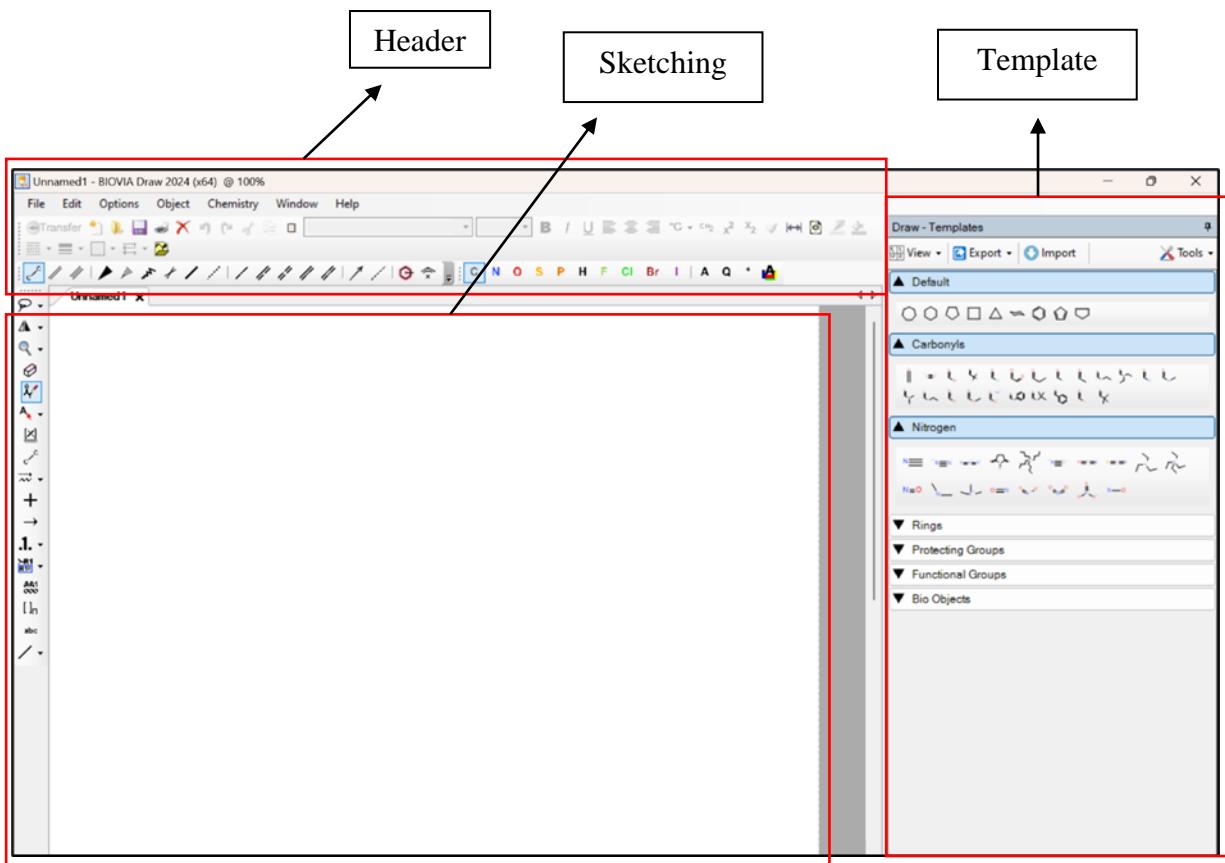


Fig 1: Homepage of BIOVIA Draw Software

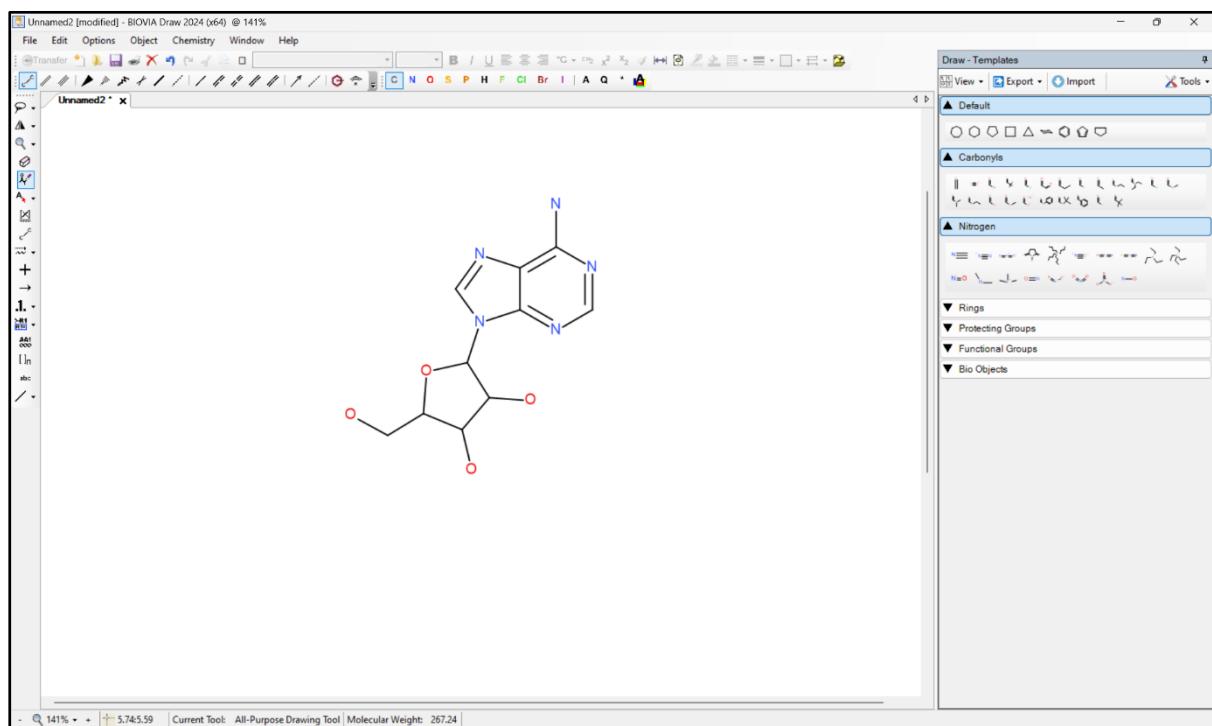


Fig 2: Structure of Adenosine

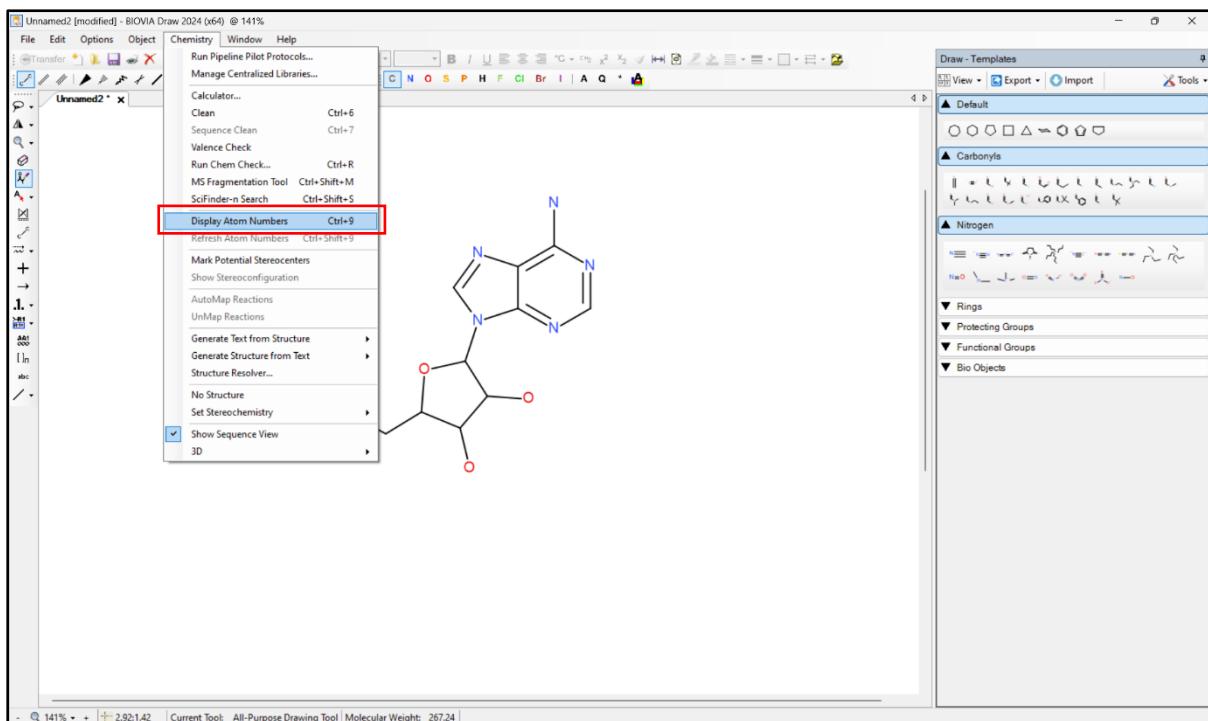


Fig 3: Option applied – ‘Display Atom Numbers’

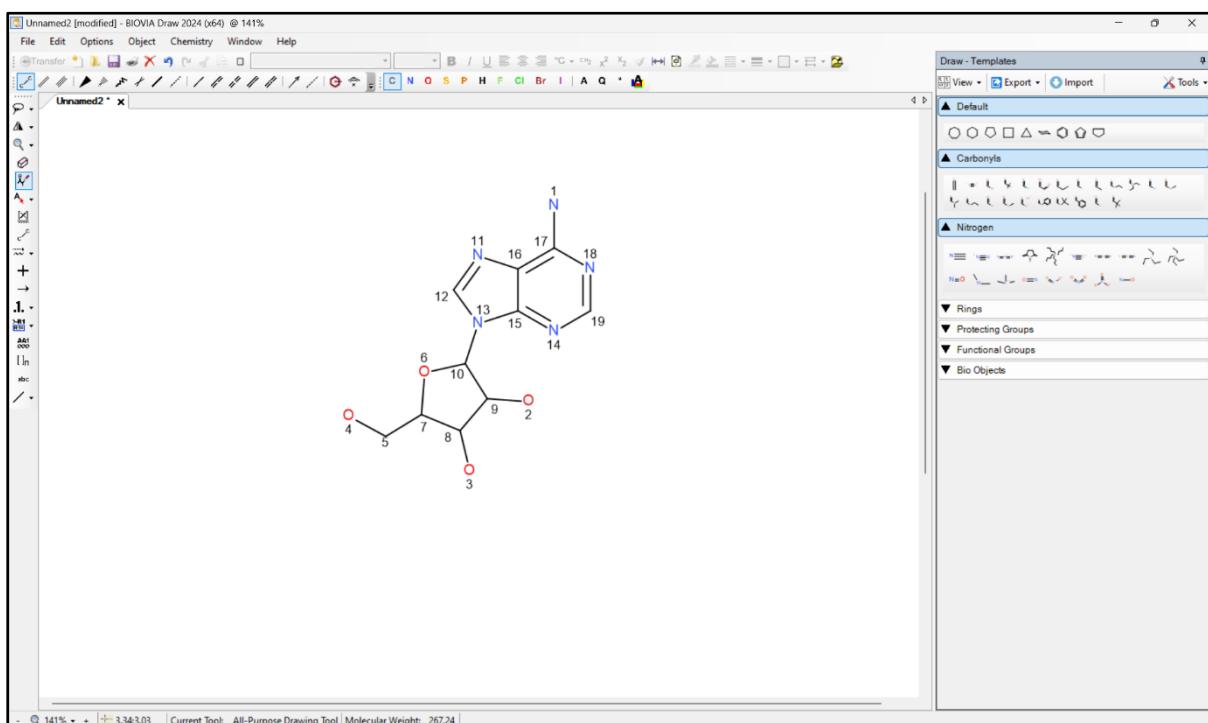


Fig 3a: Results for option applied – ‘Display Atom Numbers’

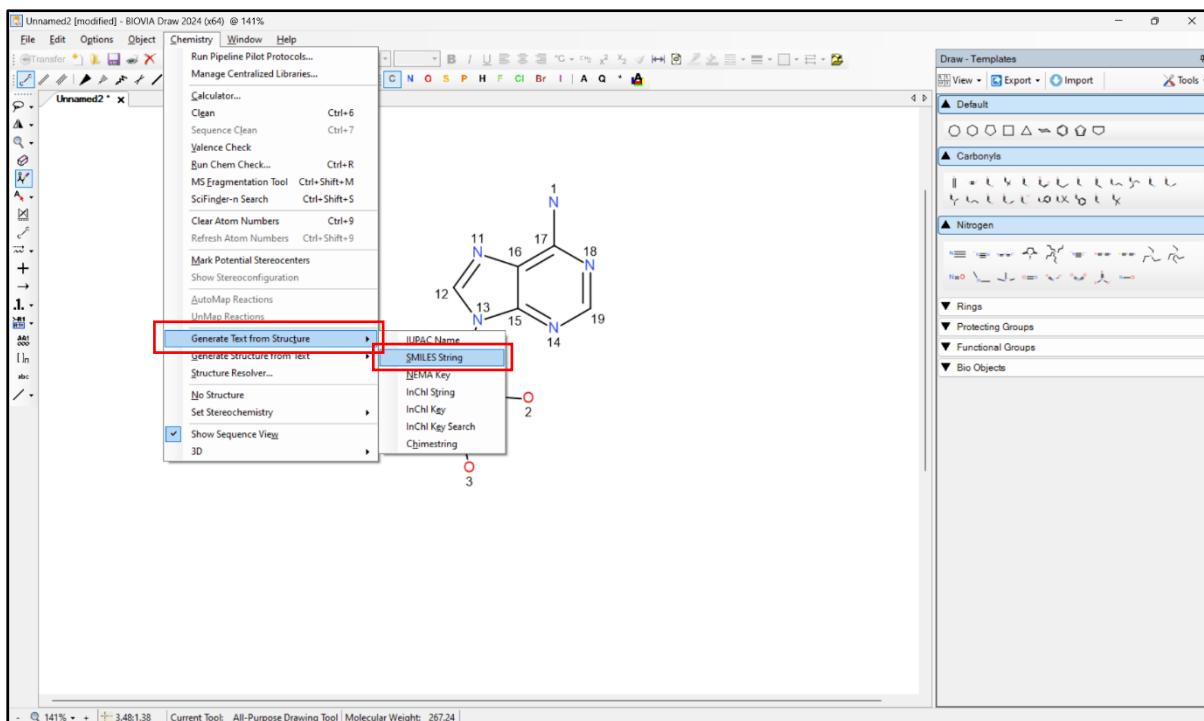


Fig 4: Option applied – ‘Generate Text from Structure – SMILES String’

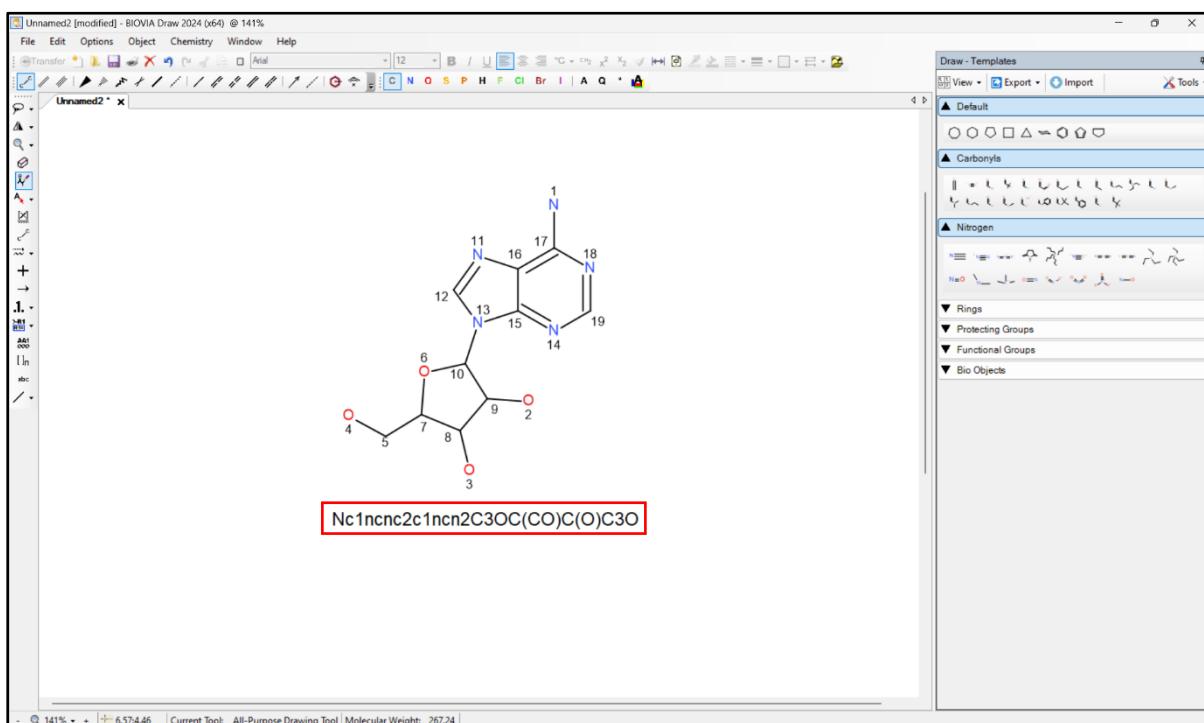


Fig 4a: Results obtained for option applied – ‘Generate Text from Structure – SMILES String’

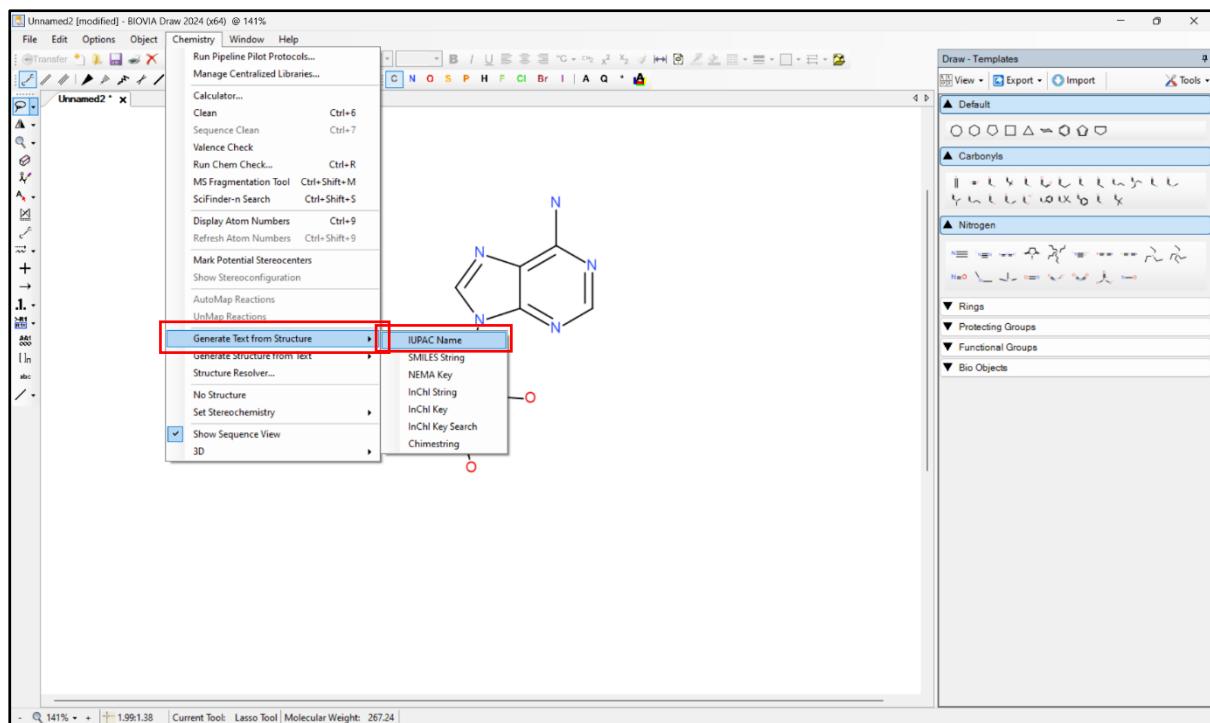


Fig 5: Option applied – ‘Generate Text from Structure – IUPAC Name’

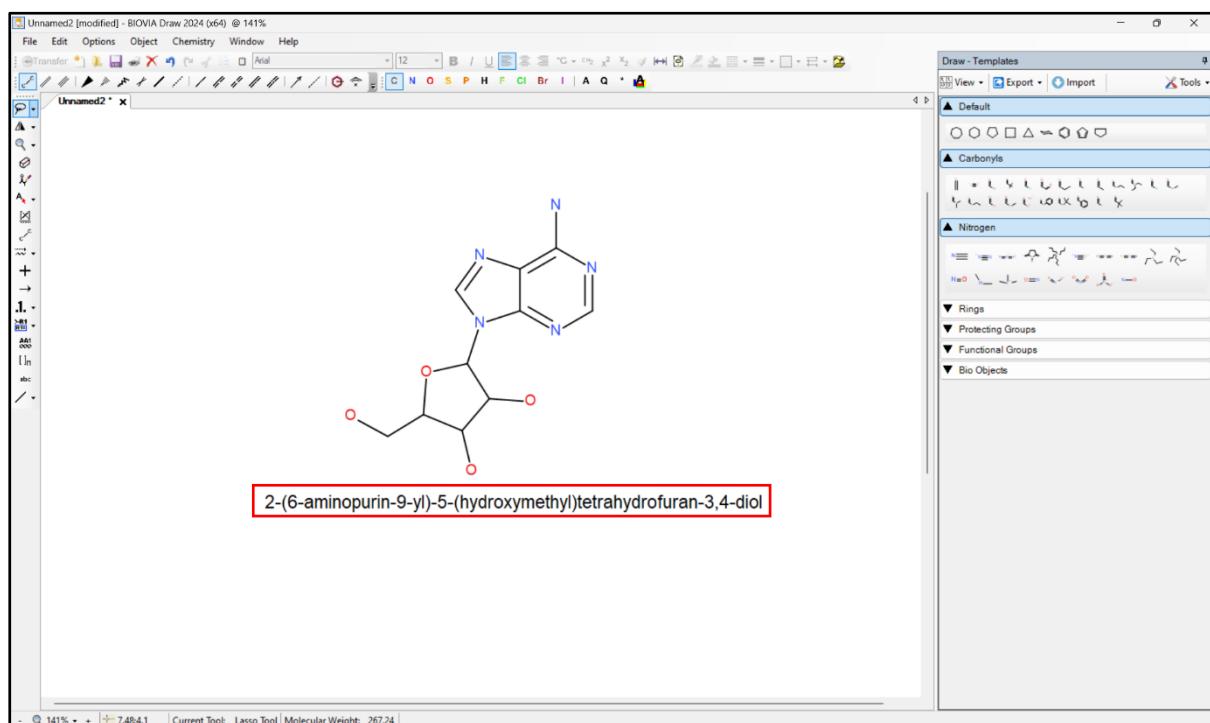


Fig 5a: Results obtained for option applied – ‘Generate Text from Structure – IUPAC Name’

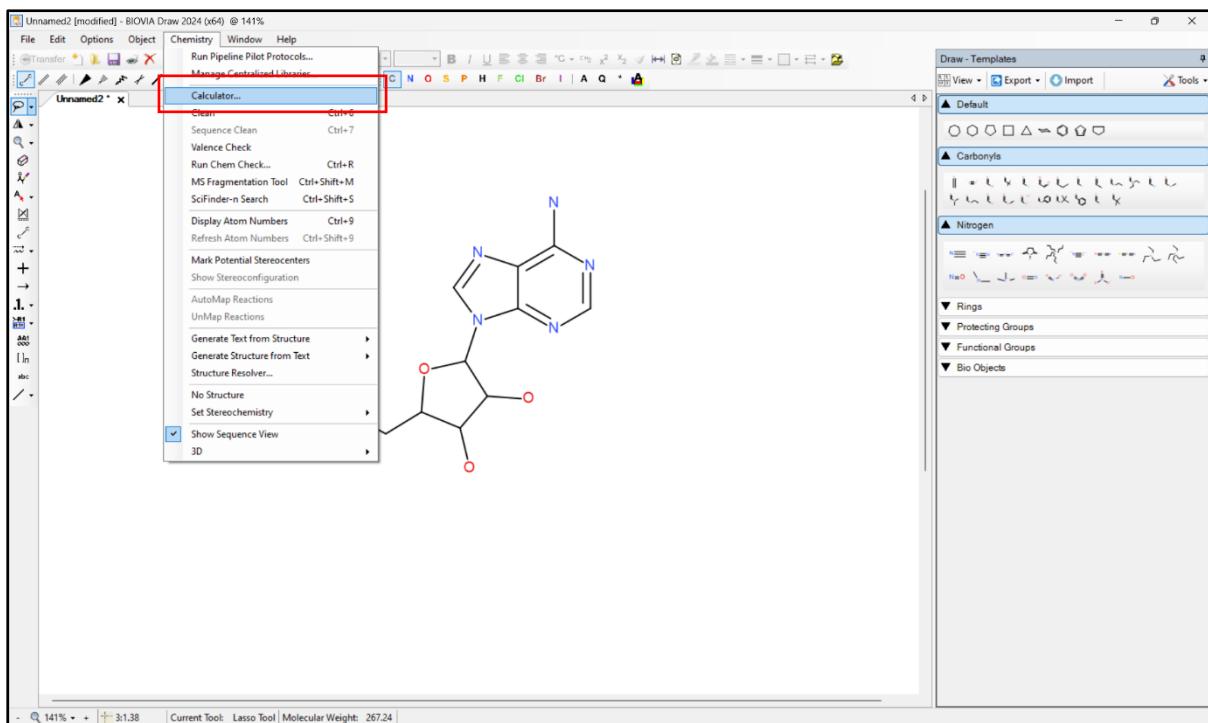


Fig 6: Option applied – ‘Calculator’

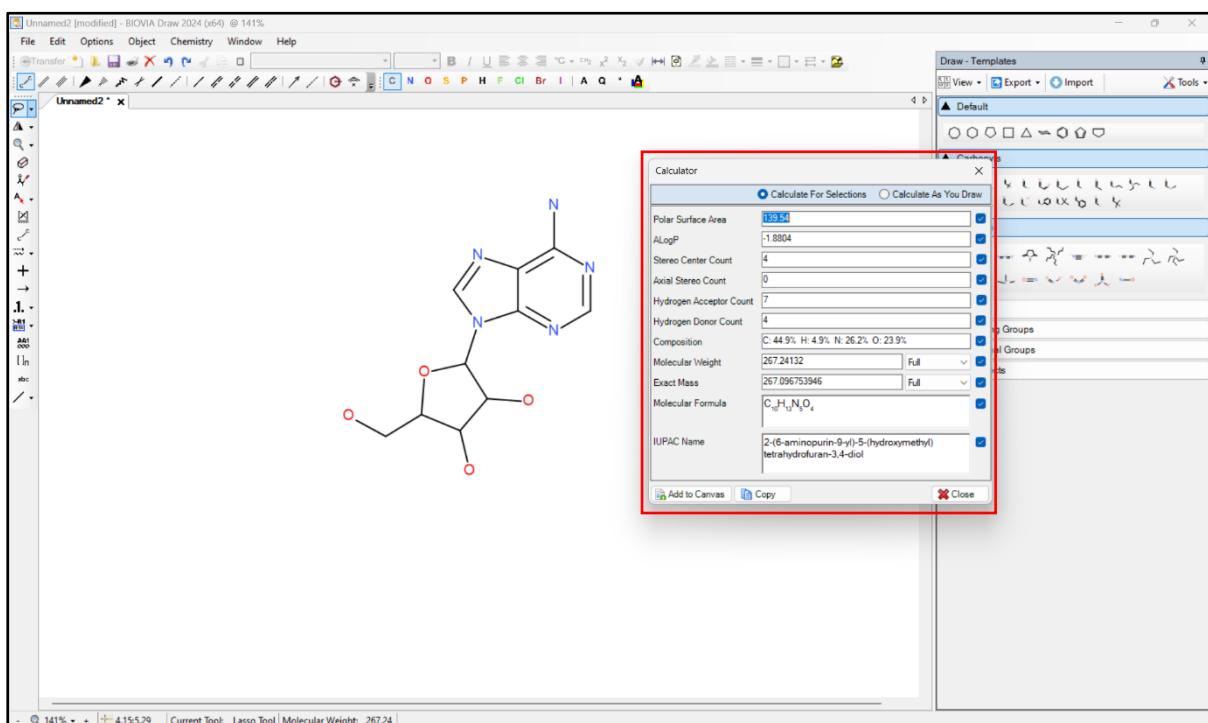


Fig 6a: Dialog box obtained for option applied – ‘Calculator’

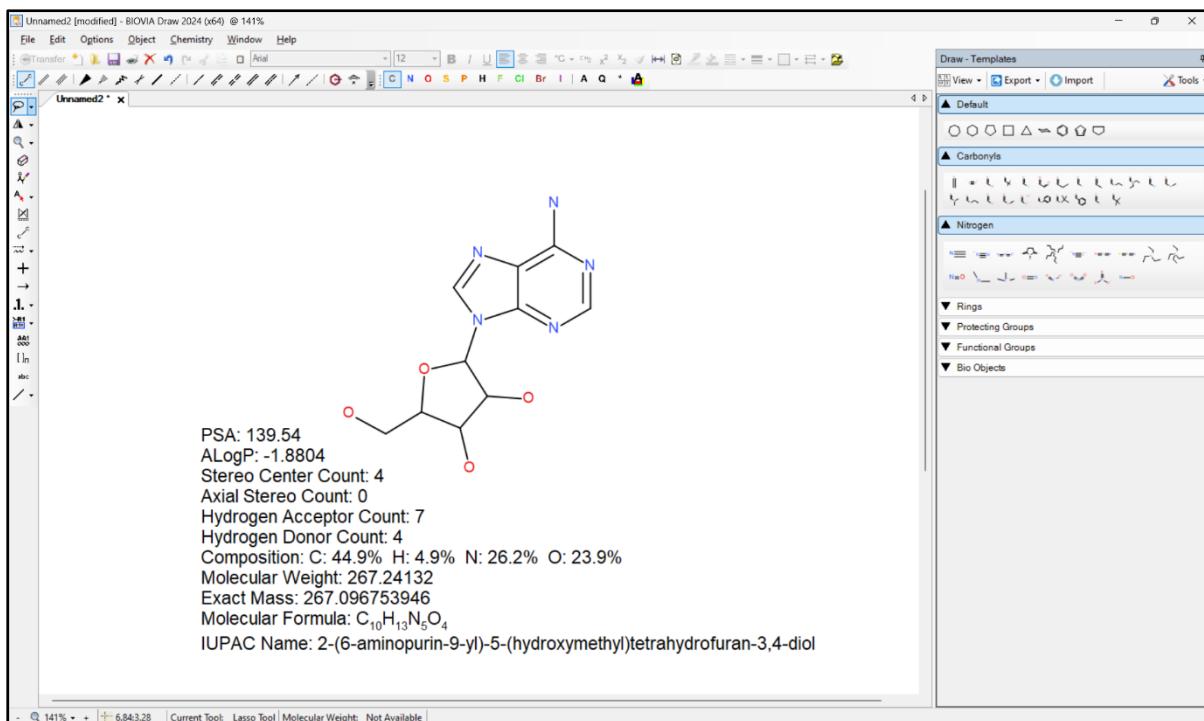


Fig 6b: Dialog box obtained for option applied – ‘Calculator’

RESULTS:

The structure of ‘Adenosine’ was drawn using the BIOVIA DRAW Software. Following information was retrieved for the structure:

1. **SMILES String:** C1=NC(=C2C(=N1)N(C=N2)C3C(C(C(O3)CO)O)O)N
2. **IUPAC Name:** 2-(6-aminopurin-9-yl)-5-(hydroxymethyl)tetrahydrofuran-3, 4-diol

3. **Physicochemical Properties:**
 - PSA: 139.54
 - ALogP: -1.8804
 - Stereo Center Count: 4
 - Axial Stereo Count: 0
 - Hydrogen Acceptor Count: 7
 - Hydrogen Donor Count: 4
 - Composition: C: 44.9% H: 4.9% N: 26.2% O: 23.9%
 - Molecular Weight: 267.24132
 - Exact Mass: 267.096753946
 - Molecular Formula: C₁₀H₁₃N₅O₄

CONCLUSION:

BIOVIA Draw is a comprehensive chemical drawing and editing tool that enables the creation and manipulation of complex biological compounds, molecules, and chemical reactions, offering structure and query drawing capabilities, biological sequence handling features, and the ability to integrate and customize the application according to organizational needs. Its structure and query drawing tools allow users to quickly sketch molecules, reactions, and sequences with precision, generate structures from common chemical string notations, and create structures with R-groups for queries or enumerations. BIOVIA Draw also provides the ability to draw, register, search, and report on chemically modified peptide or nucleotide sequences, and can be seamlessly integrated with custom Java and .NET applications as well as various BIOVIA products, allowing organizations to incorporate structure drawing and display into their workflows and applications while modifying the look-and-feel to suit their specific requirements.

REFERENCES:

1. Dassault Systèmes. (2020). BIOVIA Draw Datasheet. In BIOVIA Draw [Datasheet]. <https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/BIOVIA/PDF/biovia-draw-ds.pdf>
 2. BIOVIA Draw 2024 SP1 | What's new? (2024, July 4). [Video]. Dassault Systèmes. <https://www.3ds.com/products/biovia/draw>
 3. Admin, A. (2021, July 11). BIOVIA Draw 2020 – Chemical Structure Drawing Tool – My Biosoftware – Bioinformatics Softwares blog. <https://mybiosoftware.com/accelrys-draw-chemical-structure-drawing-tool.html>
 4. Team, E. W. (n.d.). adenosine (CHEBI:16335). <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI%3A16335>
 5. Mahler, G. S. (1998). Adenosine. In *Analytical profiles of drug substances and excipients* (pp. 1–37). [https://doi.org/10.1016/s0099-5428\(08\)60751-0](https://doi.org/10.1016/s0099-5428(08)60751-0)
-

WEBLEM 2(B)
Open Babel
(URL: <https://sourceforge.net/projects/openbabel/>)

AIM:

To understand and convert various structural file formats along with storage and retrieval methods using Open Babel tool.

INTRODUCTION:

Open Babel Tool

Open Babel is a chemical toolbox designed to speak the many languages of chemical data. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas.

Open Babel supports 111 chemical file formats in total. It can read 82 formats and write 85 formats. These encompass common formats used in cheminformatics (SMILES, InChI, MOL, MOL2), input and output files from a variety of computational chemistry packages (GAMESS, Gaussian, MOPAC), crystallographic file formats (CIF, ShelX), reaction formats (MDL RXN), file formats used by molecular dynamics and docking packages (AutoDock, Amber), formats used by 2D drawing packages (ChemDraw), 3D viewers (Chem3D, Molden) and chemical kinetics and thermodynamics (ChemKin, Thermo). Formats are implemented as "plugins" in Open Babel, which makes it easy for users to contribute new file formats.

There are various file formats which are supported by Open Babel tool, some of them are:

- 1. MOL file format:** An MDL Molfile is a file format for holding information about the atoms, bonds, connectivity and coordinates of a molecule. The molfile consists of some header information, the Connection Table (CT) containing atom info, then bond connections and types, followed by sections for more complex information.
- 2. sdf file format:** The structural data file (SDF) belongs to a family of chemical-data file formats. It is based on the MOL-file format, both developed by Molecular Design Limited (MDL), now named BIOVIA (belongs Dassault Systems).The SDF contains information on the chemical structure and associated data of compounds in plain text. The original MOL-file only encoded a single molecule, whereas files in SDF format can encode single or multiple molecules that are then delimited by lines consisting of four dollar signs (\$\$\$\$). SDF files are formatted ASCII files that store information about the positions of the individual atoms (either in 2D or 3D space) that make up the molecule. The data on connectivity and hybridization state are also encoded, although their use is less frequent and often inconsistent.

- 3. SMILES file format:** SMILES (Simplified Molecular Input Line Entry System) is a chemical notation that allows a user to represent a chemical structure in a way that can be used by the computer. SMILES is an easily learned and flexible notation. The SMILES notation requires that you learn a handful of rules. You do not need to worry about ambiguous representations because the software will automatically reorder your entry into a unique SMILES string when necessary.
- 4. SMARTS file format:** SMARTS is a language that allows you to specify substructures using rules that are straightforward extensions of SMILES. SMARTS provides a number of primitive symbols describing atomic properties beyond those used in SMILES (atomic symbol, charge, and isotopic specifications).
- 5. XYZ Cartesian coordinate system:** The XYZ file format is a simple text-based format used to represent molecular structures in three-dimensional Cartesian coordinates. Each file begins with a line indicating the number of atoms, followed by a comment line. Subsequent lines list each atom's symbol and its X, Y, Z coordinates in angstroms.

Adenosine

Adenosine is an important organic compound and a nucleoside that plays a crucial role in various physiological processes. It is composed of two main components: the nitrogenous base adenine and the sugar ribose. These components are linked by a β -N9-glycosidic bond, forming the structure of adenosine. The chemical formula of adenosine is $C_{10}H_{13}N_5O_4$, and its structure can be represented as follows:

- 1. Adenine:** A purine base that features a fused double-ring structure containing nitrogen atoms.
- 2. Ribose:** A five-carbon sugar that is part of the nucleoside structure.

Adenosine is a key player in cellular energy transfer, primarily through its derivatives: adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP). ATP, in particular, is often referred to as the "molecular unit of currency" for energy transfer in cells, as it captures and releases energy necessary for various cellular processes, including muscle contraction and nerve impulse propagation.

In addition to its role in energy metabolism, adenosine also functions as a signaling molecule in the body, influencing processes such as sleep regulation, immune response, and cardiovascular function. It operates through four known receptor subtypes (A1, A2A, A2B, and A3), which are G-protein-coupled receptors that mediate various physiological effects.

METHODOLOGY:

1. Download the open babel tool from the website source forge.
(URL: <https://sourceforge.net/projects/openbabel/>)
2. Upload the mol file obtained from the BIOVIA software in the input section of the open babel tool.
3. Select the desired file formats from the output section.
4. Select various parameters from the properties panel.
5. Click on "CONVERT" for the conversion.

OBSERVATIONS:

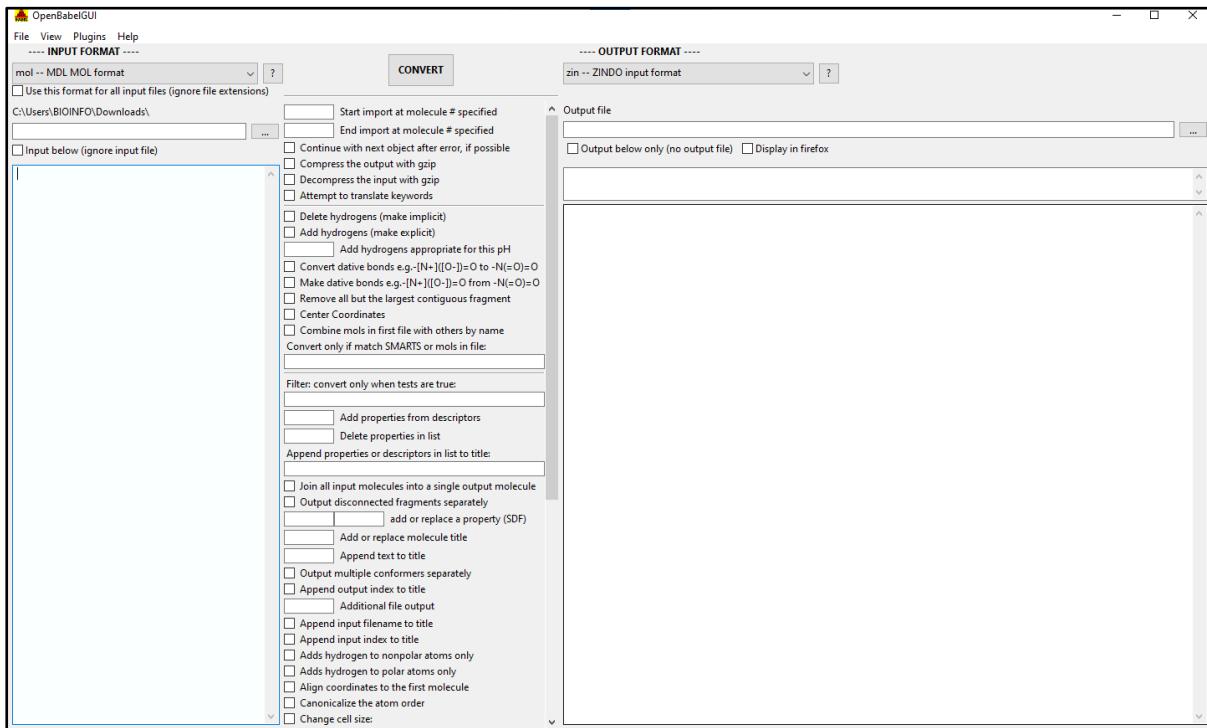


Fig 1: Homepage of Open Babel tool

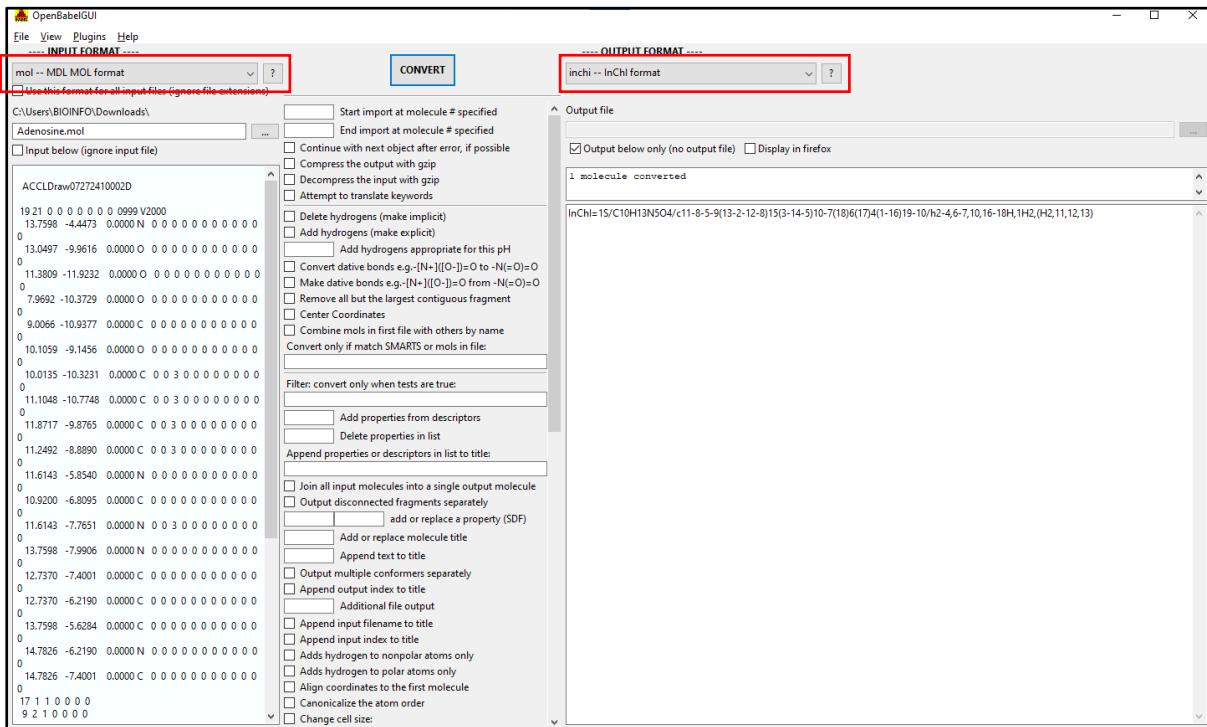


Fig 2: Conversion of mol --MDL MOL format to inchi --InChI format

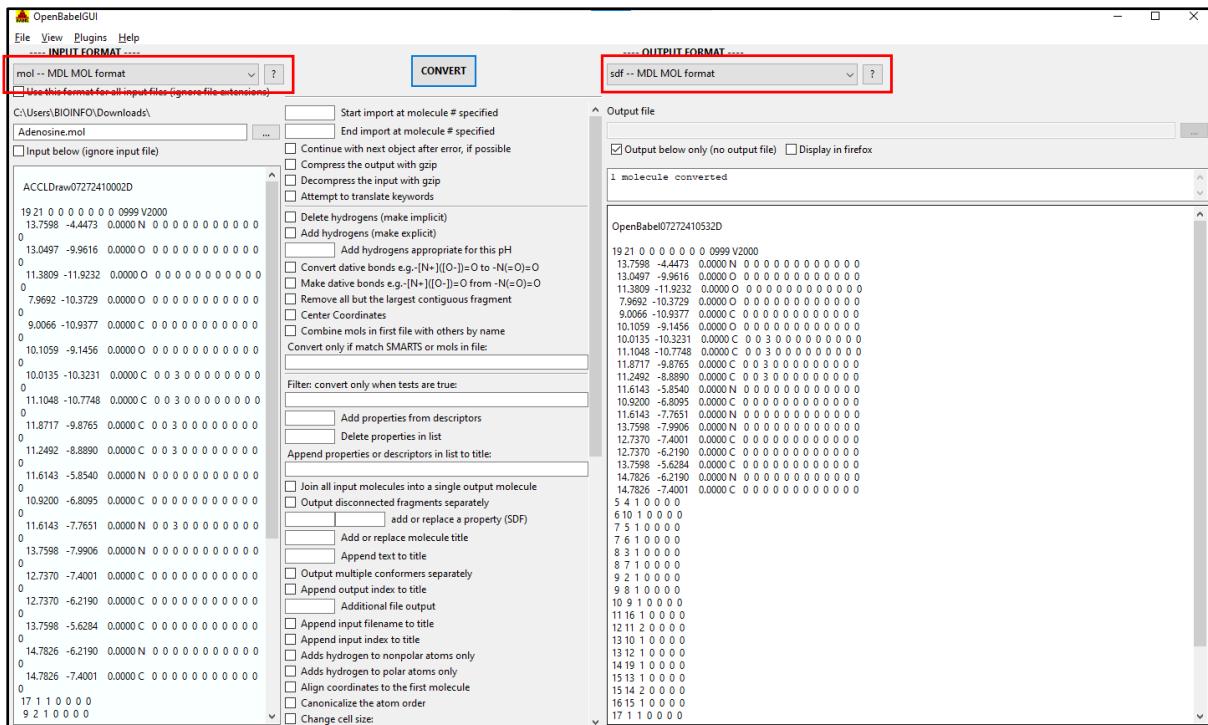


Fig 3: Conversion of mol --MDL MOL format to sdf --MDL MOL format

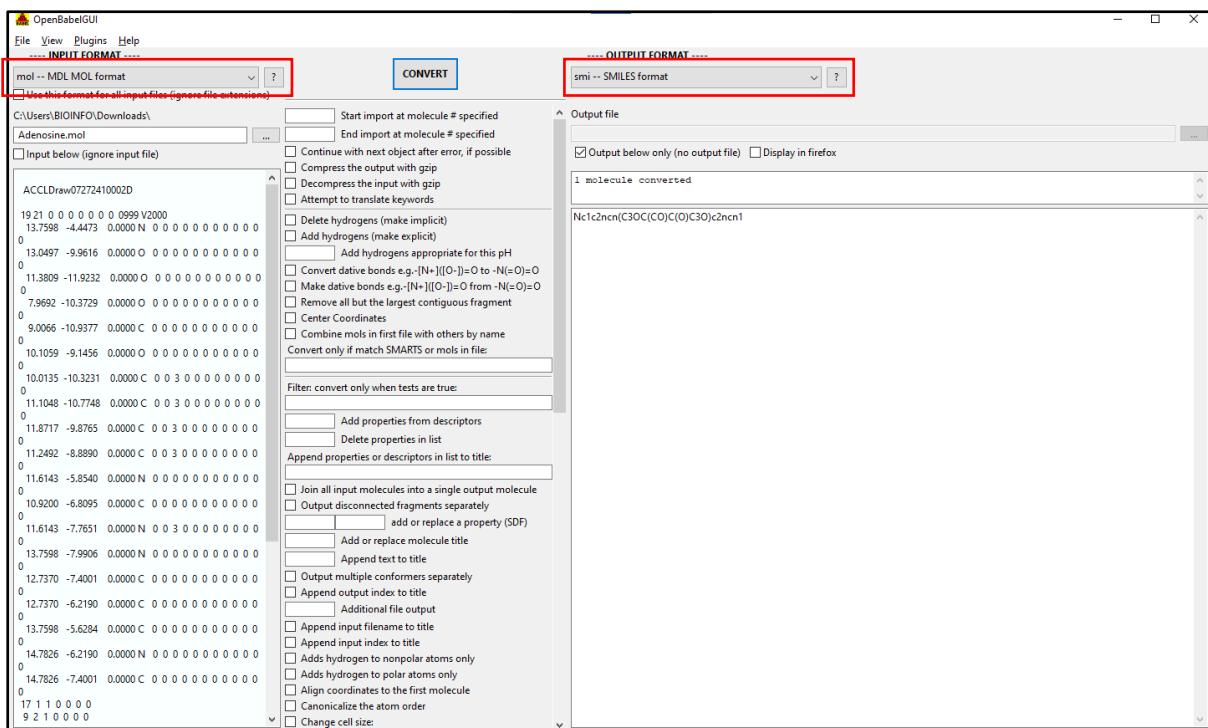


Fig 4: Conversion of mol --MDL MOL format to smi --SMILES format

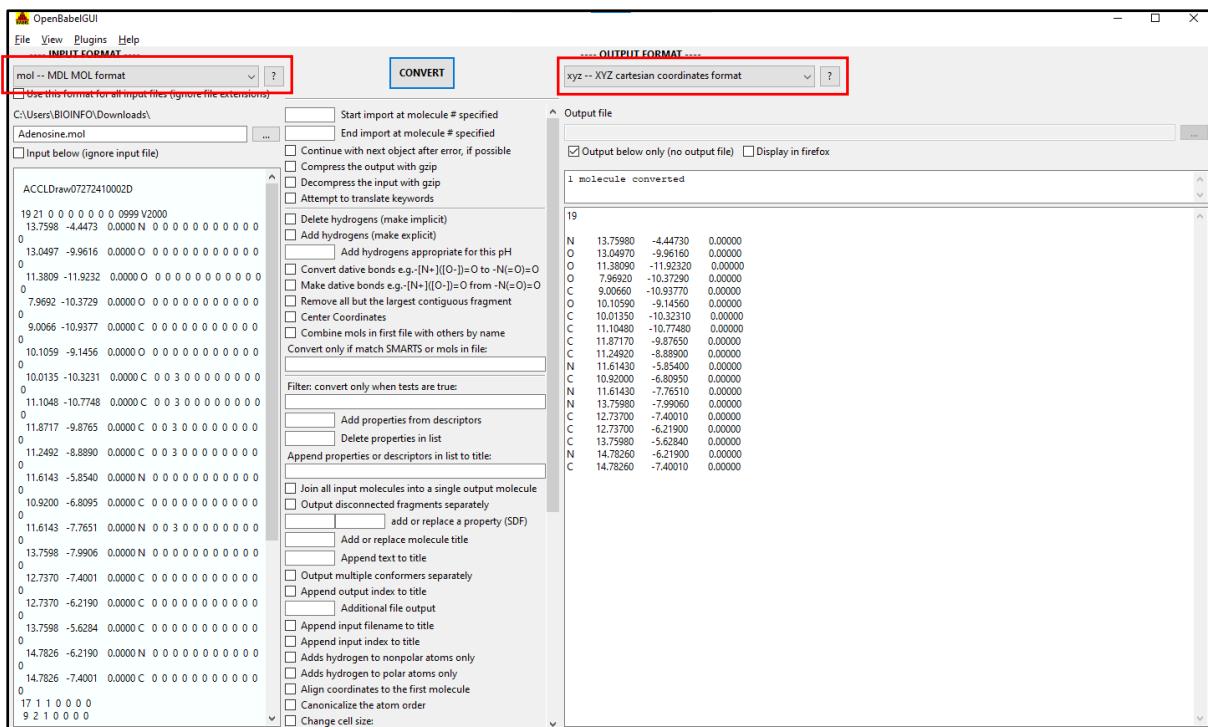


Fig 5: Conversion of mol --MDL MOL format to xyz --XYZ cartesian coordinates format

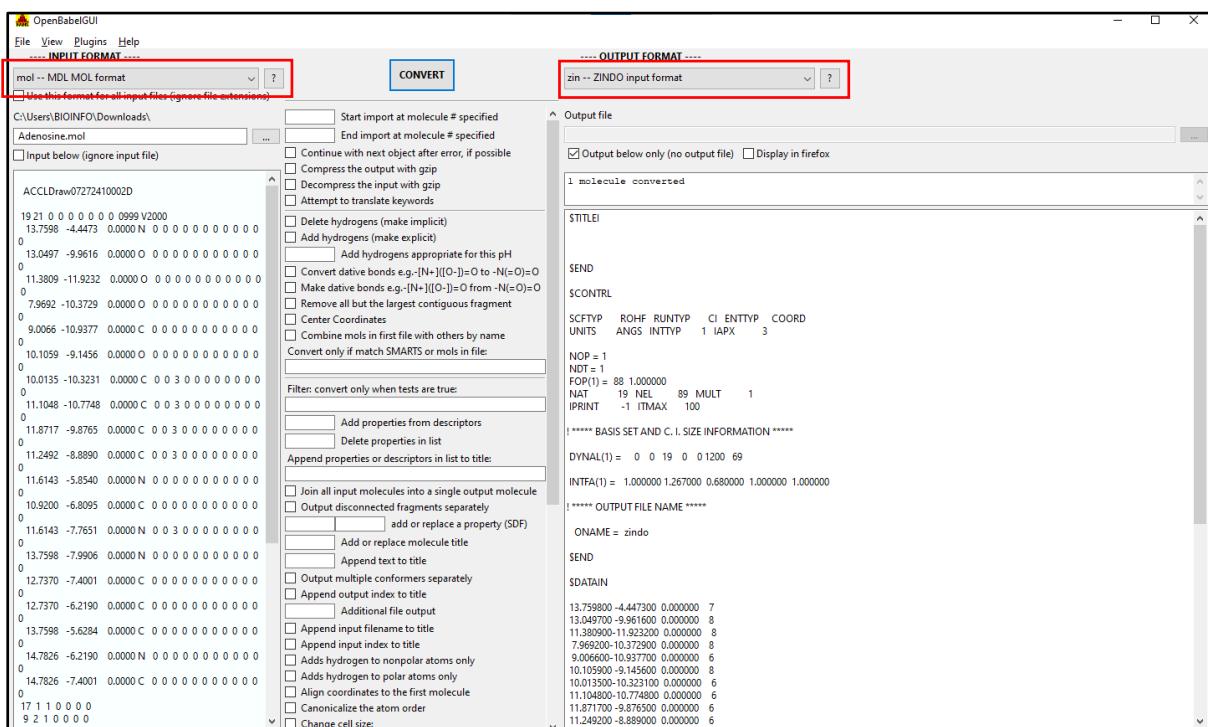


Fig 6: Conversion of mol --MDL MOL format to zin --ZINDO input format

RESULTS:

Open Babel is a tool used for conversion of file formats. It was used to convert the mol --MDL MOL format from BIOVIA software into different formats. The file was successfully changed into several formats, including inchi --InChI format, sdf --MDL MOL format, smi --SMILES format, --XYZ cartesian coordinates format and zin --ZINDO input format. Each format is useful for different purposes. InChI format is for chemical identifiers, SDF format is for structural data, SMILES format is for a simplified way to write molecules, XYZ format is for coordinates and ZINDO format is for semiempirical quantum-mechanics program ZINDO. This makes it easier for users to use the data in different software and for various analyses.

CONCLUSION:

Open babel tool was explored and various structural file formats were converted into different other file formats. File formats like inchi --InChI format, sdf --MDL MOL format, smi --SMILES format, --XYZ cartesian coordinates format and zin --ZINDO input format were studied and explored.

REFERENCES:

1. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of cheminformatics*, 3, 33. <https://doi.org/10.1186/1758-2946-3-33>
 2. User Guide — Open Babel openbabel-3-1-1 documentation. (n.d.). <https://openbabel.org/docs/index.html>
 3. OPENBABEL - Chemical file format converter. (n.d.). <https://www.cheminfo.org/Chemistry/Cheminformatics/FormatConverter/index.html>
 4. Team, E. W. (n.d.). adenosine (CHEBI:16335). <https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI%3A16335>
 5. Mahler, G. S. (1998). Adenosine. In *Analytical profiles of drug substances and excipients* (pp. 1–37). [https://doi.org/10.1016/s0099-5428\(08\)60751-0](https://doi.org/10.1016/s0099-5428(08)60751-0)
-

WEBLEM 3

Introduction to Balloon Software

(URL: https://users.abo.fi/mivainio/balloon/change_log.htm)

INTRODUCTION:

BALLOON is a free command-line program for 3D molecular model generation and conformational analysis developed by Mikko Vainio. It runs on any platform/operating system for which a binary executable has provided. BALLOON creates 3D atomic coordinates from molecular connectivity via distance geometry and conformer ensembles using multi-objective genetic algorithm. It uses MMFF94-type non-polarizable charges and polarizable (conformation-dependent) electronegativity equalization charges (SFKEEM and EEM) for performing geometry optimization. This is a simple video tutorial to build a 3D small molecule by drawing its 2D chemical structure using JSME Molecule Editor and the BALLOON tool.

Features

1. Supported input file types are –

- MDL/Symyx/Accelrys SD file format, versions V2000 and V3000.
- SMILES string representation. SMILES can be input from the command-line, in a configuration file ('input-file=<smiles>'), or in a text file with one SMILES (followed by an optional molecule name) per line.
- MOL2 file format.
- Visipoint binary file format (VBF). This is an experimental format being currently developed. An informal specification is in preparation, please ask if you are interested in the details.
- XMol XYZ file format. The parser expects hydrogen atoms to be present and guesses bond orders based on geometry. Formal atomic charges are assigned to atoms for which no charge is listed in the XYZ file.
- A very crude implementation of the PDB and Gaussian output log formats. These may be useful for the extraction of ligand structures from PDB files (dump the protein before attempting this!) and quantum mechanically optimized structures from Gaussian logs. Bonding and hybridisation states of atoms are guessed based on their local geometry.

2. Supported output file types are MDL/Symyx/Accelrys SD file format, versions V2000 and V3000 (latter used when the size of the structure so requires or if forced using option 'output-format=sdf:v3'), MOL2, and VBF formats.

3. Assigns partial charges to atoms (MOL2 output format). Available charges are MMFF94-type non-polarizable charges and polarizable (conformation-dependent) electronegativity equalization charges (SFKEEM and EEM). The polarizable models are not intended to be used while performing geometry optimization, only when assigning partial charges - see options 'onlycharge' and 'chargemodel'.

4. Molecular superposition using a Gaussian shape-density and partial charge weighted shape-density representation of the structures. (New in version 0.6.0). This is an experimental feature that, unfortunately, doesn't seem to work very well.

Usage examples

Balloon lists available options when called with '-h' or '--help'. In order to perform energy minimization using the MMFF94-like force field, Balloon needs to know the location of the force field parameter file (named 'MMFF94.mff' in the distribution). The location of the file is given using the '--forcefield' ('-f' for short) switch or via the environment variable BALLOON_FORCEFIELD. If the parameter file is not found, Balloon uses distance geometry only. The conformer sampling genetic algorithm always requires the force field parameters.

5. Generate an ensemble of conformations

```
balloon -f /path/to/MMFF94.mff --nconfs 20 --nGenerations 300 --rebuildGeometry in.sdf  
out.sdf
```

The option 'rebuildGeometry' causes conformations to be created from scratch. Otherwise, if the input structure has 3D coordinates, Balloon will use it as the initial conformation that the genetic algorithm alters to create new conformations. If valid 3D structures are used as input, the option 'rebuildGeometry' should be omitted. Note that although the initial ensemble size is 20 conformations, the final ensemble may contain more or less depending on the flexibility of the structure.

6. Generate one random conformation from SMILES

This example uses only distance geometry and no energy minimization:

```
balloon --nconfs 1 --noGA "C[n]1cnc2N(C)C(=O)N(C)C(=O)c12" caffeine.sdf
```

Now the same as above, but with energy minimization (just provide the force field parameter file):

```
Balloon -f /path/to/MMFF94.mff -nconfs 1 -noGa "C[n]1cnc2N(C)C(=O)c12"  
caffine_min.sdf
```

7. Assign MMFF94-like partial atomic charges

The output is in MOL2 format, geometry untouched:

```
balloon -f /path/to/MMFF94.mff --onlycharge in.sdf out.mol2
```

INSTALLATION:

The screenshot shows the homepage of the Balloon software. The title "Balloon" is at the top left. Below it is a horizontal navigation bar with links: "about", "news", "features", "download", "change log", and "faq". The "download" link is underlined, indicating it is the active page. The main content area starts with the text "Latest version 1.8.2 (March 13 2022)". Under "About", there is a detailed description of the software's purpose and capabilities, mentioning SMILES, SDF, and MOL2 formats, and its portability across Linux, Mac OS X, and Microsoft Windows. There are also sections for "Contact information" (listing Mikko Vainio's email) and "Bibliographic references" (citing a 2007 paper by Mikko J. Vainio and Mark S. Johnson).

Fig 1: Homepage of Balloon Software

This screenshot is similar to Fig 1, showing the homepage of the Balloon software. The "download" link in the navigation bar is highlighted with a red rectangle. Below the navigation bar, the text "BALLOON End User License Agreement" is displayed in bold blue capital letters. A large block of text follows, detailing the End User License Agreement (EULA) terms, emphasizing that it is a legal agreement between the licensee and the copyright owner(s) of the software.

Fig 1a: Click on the download section

BE BOUND BY ITS TERMS.

Current version is 1.8.2 (March 13 2022)

Please select your platform:

Linux 2.6.18, 64-bit x86_64
 Linux 2.6.18, 64-bit x86_64
Windows, 64-bit x86_64
 Mac OS X 10.15, 64-bit x86_64

System Requirements

BALLOON is a command-line program and will most likely run on any platform that can run any of the operating systems for which a binary executable is provided.

Related material

Other downloads related to BALLOON:
 Reference structures used in the publication ▾

[Download](#)

Fig 1b: Download the software for your required platform

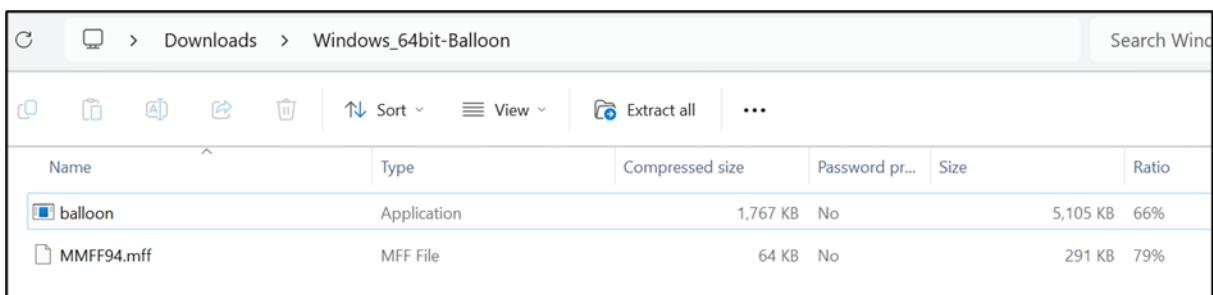


Fig 1c: Extract the zip files by clicking on 'Extract all'

REFERENCES:

1. Anders, S., Pyl, P. T., & Huber, W. (2014). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. <https://doi.org/10.1093/bioinformatics/btu638>
2. Ohashi, N., & Kohno, T. (2020). Analgesic Effect of Acetaminophen: A Review of Known and Novel Mechanisms of Action. *Frontiers in pharmacology*, 11, 580289. <https://doi.org/10.3389/fphar.2020.580289>

WEBLEM 3(A)
Balloon Software

(URL: https://users.abo.fi/mivainio/balloon/change_log.htm)

AIM:

To get the conformers for the query ‘Imatinib’ (PubChem ID: 5291) using Balloon software.

INTRODUCTION:

- Purpose:** Balloon is a free command-line program for generating 3D molecular models and performing conformational analysis.
- Features:** It creates 3D atomic coordinates from molecular connectivity using distance geometry and conformer ensembles generated by a multi-objective genetic algorithm.
- Input/Output:** Balloon accepts SMILES, SDF or MOL2 format as input and outputs SDF or MOL2 files. It handles flexibility of aliphatic rings and stereochemistry about double bonds and tetrahedral chiral atoms.
- Availability:** Balloon runs on Linux, Mac OS X, and Microsoft Windows platforms and is freely available

Imatinib

Imatinib, marketed under the brand names Gleevec and Glivec, is a targeted therapy medication primarily used in the treatment of certain cancers, particularly chronic myeloid leukemia (CML) and gastrointestinal stromal tumors (GIST). It is classified as a tyrosine kinase inhibitor (TKI), which means it works by blocking specific proteins (tyrosine kinases) that promote cancer cell growth and proliferation.

Mechanism of Action

Imatinib specifically inhibits the BCR-ABL tyrosine kinase, an abnormal protein produced by the Philadelphia chromosome mutation commonly found in CML. By binding to the ATP pocket of this protein, imatinib prevents its activation and subsequent signaling that leads to cancer cell multiplication. Additionally, imatinib also targets other tyrosine kinases such as c-KIT and PDGFRA, which are involved in various cancers and disorders.

Clinical Use

Imatinib is indicated for:

- Chronic Myeloid Leukemia (CML):** Particularly effective in patients with the Philadelphia chromosome.
- Gastrointestinal Stromal Tumors (GIST):** Used when tumors express activating mutations of c-KIT.
- Acute Lymphoblastic Leukemia (ALL):** Specifically for Philadelphia chromosome-positive cases.

- 4. Dermatofibrosarcoma Protuberans:** When surgical removal is not feasible.
- 5. Hypereosinophilic Syndrome and Systemic Mastocytosis.**

Administration and Pharmacokinetics

Imatinib is administered orally, with a high bioavailability of approximately 98%. It is typically taken once or twice daily, depending on the specific condition being treated. The drug is metabolized mainly in the liver and has a half-life of about 18 hours.

Side Effects

Common side effects of imatinib include:

- 1. Nausea and vomiting**
- 2. Diarrhea**
- 3. Fatigue**
- 4. Muscle cramps**
- 5. Rash**

Serious side effects can include liver dysfunction, heart problems, and severe allergic reactions. Monitoring during treatment is essential to manage these potential adverse effects effectively.

METHODOLOGY:

1. Visit the website https://users.abo.fi/mivainio/balloon/change_log.html and download the software by following the steps.
2. Proceed to PubChem and enter the query ‘Imatinib’ and click on search.
3. Select the best match molecule for the query ‘Imatinib’ (PubChem ID: 5291) and click on download.
4. Download the 2D structure or 3D conformer of the molecule in SDF format and save the file in the software’s folder itself.
5. Start the software by pressing the ‘shift + right key’ and open in Command prompt or PowerShell window.
6. Copy the commands for acquiring the SMILES and conformers for the query from Balloon.
7. View the SDF file of the conformer and its lowest energy.

OBSERVATIONS:

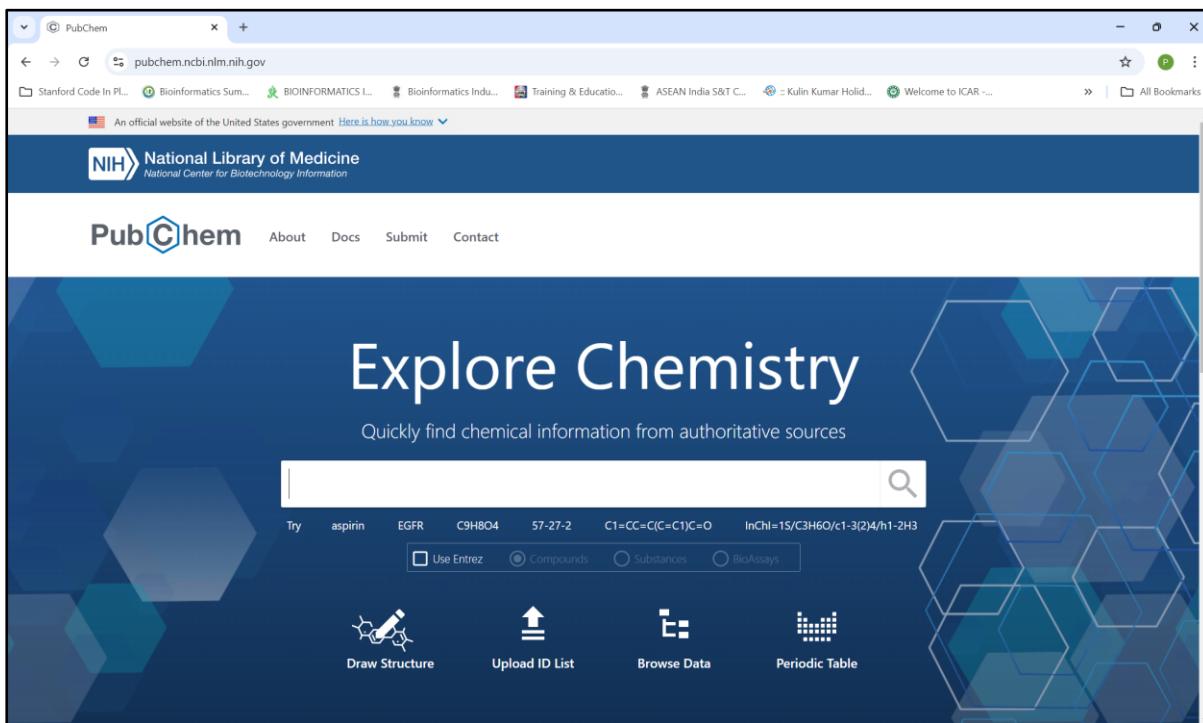


Fig 1: Homepage of PubChem

A screenshot of a web browser displaying the PubChem compound summary page for Imatinib (PubChem CID 5291). The URL in the address bar is pubchem.ncbi.nlm.nih.gov/compound/5291. The page has a white header with the NIH logo and the PubChem logo. On the left, there is a "COMPOUND SUMMARY" section for Imatinib. This section includes: "PubChem CID" (5291), "Structure" (2D and 3D chemical structures), "Chemical Safety" (two hazard symbols: Irritant and Health Hazard), "Molecular Formula" (C₂₉H₃₁N₇O), and "Synonyms" (Imatinib). On the right, there is a "CONTENTS" sidebar with a table of contents listing 16 sections, including "Title and Summary" (which is highlighted in light blue), "1 Structures", "2 Names and Identifiers", "3 Chemical and Physical Properties", etc. At the top right of the main content area, there are two buttons: "Cite" and "Download". The "Download" button is highlighted with a red rectangle.

Fig 2: Click on 'Download'

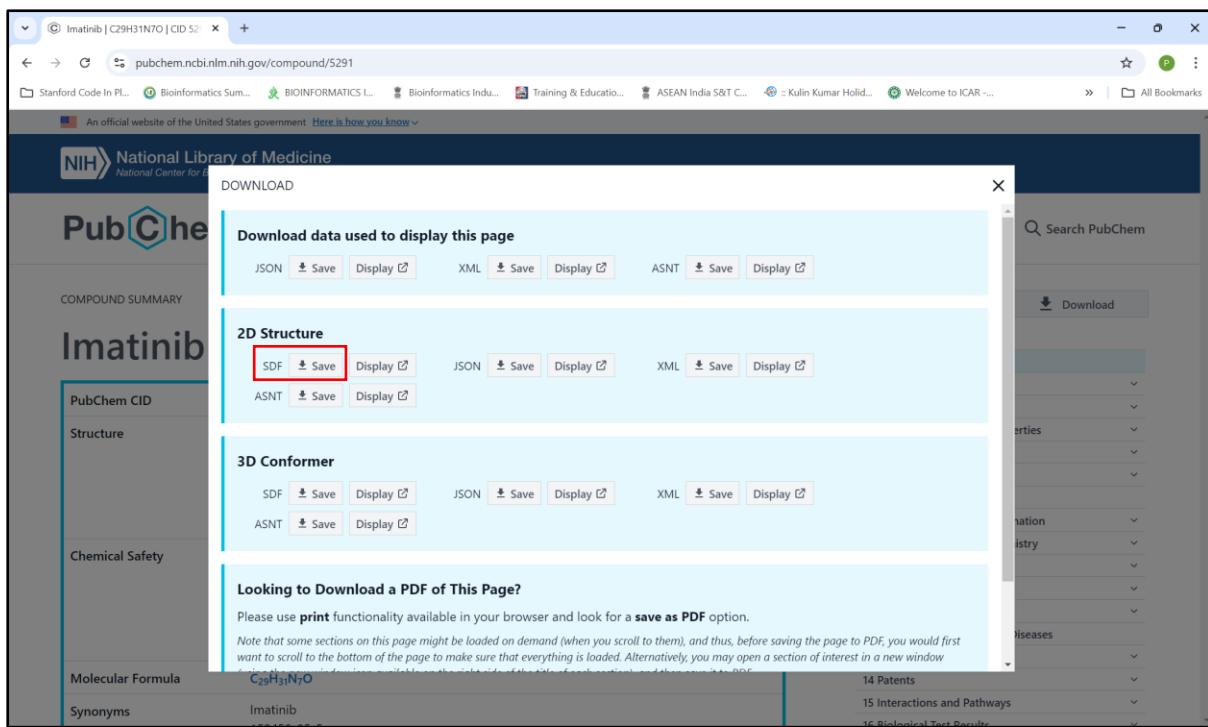


Fig 2a: Download the 2D structure of the query ‘Imatinib’ (PubChem ID: 5291) in SDF format

option `rebuildGeometry` should be omitted. Note that although the initial ensemble size is 20 conformations, the final ensemble may contain more or less depending on the flexibility of the structure.

- **Generate one random conformation from SMILES**

This example uses only distance geometry and no energy minimization:

```
balloon --nconfs 1 --noGA "C[n]1cnc2N(C)C(=O)N(C)C(=O)c12"
caffeine.sdf
```

Now the same as above, but with energy minimization (just provide the force field parameter file):

```
balloon -f /path/to/MMFF94.mff --nconfs 1 --noGA
"C[n]1cnc2N(C)C(=O)N(C)C(=O)c12" caffeine_min.sdf
```

Fig 3: Copy the commands to generate conformers from SMILES using the Balloon software

```

Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\prarthi kothari\Downloads\Windows_64bit-Balloon> .\balloon.exe --nconfs 1 --noGA "CC1=C(C=C(C=C1)NC(=O)C2=CC=C(C=C2)CN3CCN(CC3)C)NC4=NC=CC(=N4)C5=CN=CC=C5" Imatinib.sdf

Balloon version 1.8.2.6bf18fc2e8b2f84a6e59f2fea4dbadf39e35491b 64-bit build Mar 13 2022 15:55:48
Described in http://dx.doi.org/10.1021/ci6005646
Copyright (C) 2006-2021 Mikko J. Vainio and J. Santeri Puranen.
Copyright (C) 2010 Visipoint Ltd. www.visipoint.fi
All rights reserved.
THIS SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND.

1 molecule processed in 00:00:00.710098 wall,
00:00:00.609375 user,
00:00:00 system,
00:00:00.609375 total CPU time.
PS C:\Users\prarthi kothari\Downloads\Windows_64bit-Balloon>

```

Fig 4: Paste the first command and the canonical SMILES format of the query ‘Imatinib’ (PubChem ID: 5291)

```

PS C:\Users\prarthi kothari\Downloads\Windows_64bit-Balloon> .\balloon.exe -f MMFF94.mff --nconfs 20 --nGeneration 5 --i
nput-file Structure2D_COMPOUND_CID_5291.sdf --output-file Conformers_Imatinib.sdf --rebuildGeometry

Balloon version 1.8.2.6bf18fc2e8b2f84a6e59f2fea4dbadf39e35491b 64-bit build Mar 13 2022 15:55:48
Described in http://dx.doi.org/10.1021/ci6005646
Copyright (C) 2006-2021 Mikko J. Vainio and J. Santeri Puranen.
Copyright (C) 2010 Visipoint Ltd. www.visipoint.fi
All rights reserved.
THIS SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND.

1 molecule processed in 00:00:11.125360 wall,
00:00:10.515625 user,
00:00:00 system,
00:00:10.515625 total CPU time.
PS C:\Users\prarthi kothari\Downloads\Windows_64bit-Balloon>

```

Fig 5: Paste the second command to get the conformers for the query ‘Imatinib’ (PubChem ID: 5291)

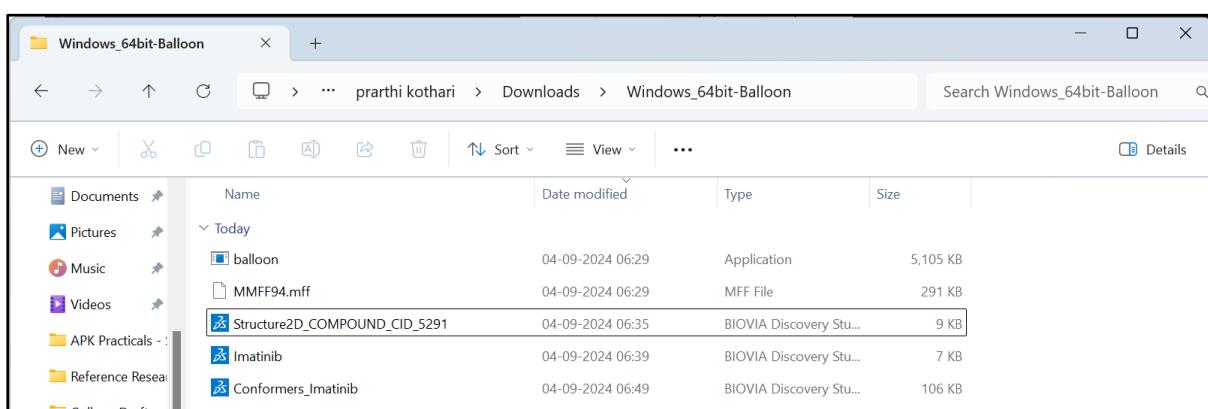


Fig 6: the SDF files generated will be saved in the folder along with the query file and the software

Structure2D_COMPOUND_CID_525	
File	Edit
§291	
-OEChem-09032420582D	
68	72
7.1962	-0.5000
2.8660	-3.0000
2.8660	-5.0000
6.3301	1.0000
8.0622	4.0000
6.3301	4.0000
7.1962	5.5000
2.8660	4.0000
2.0000	-3.5000
3.7320	-3.5000
2.0000	-4.5000
3.7320	-4.5000
2.8660	-2.0000
2.8660	-6.0000
3.7320	-1.5000
3.7320	-0.5000
4.5981	-2.0000
4.5981	0.0000
5.4641	-1.5000
5.4641	-8.5000
6.3301	-8.0000
7.1962	3.0000
8.0622	3.0000
7.1962	3.5000
8.9282	2.5000
8.0622	1.0000
8.9282	1.5000
9.7942	3.0000
7.1962	4.5000
5.4641	4.5000
4.5981	4.0000
5.4641	5.5000
6.3301	6.0000
4.5981	3.0000
3.7320	4.5000
3.7320	2.5000
2.8660	3.0000
1.7880	-2.9174
1.3894	-3.6877
4.3426	-3.6877
3.9441	-2.9174
1.3894	-4.3923

Fig 7: 2D Structure of ‘Imatinib’ (PubChem ID: 5291) viewed in Notepad

Fig 8: Imatinib SDF file opened and viewed in Notepad

Fig 9: 3D Conformers of ‘Imatinib’ (PubChem ID: 5291) viewed in Notepad

Structure2D_COMPOUND_CID_5291.sdf Conformers_Imatinib.sdf

File Edit View

```
> <PUBCHEM_IUPAC_INCHI>
InChI=1S/C29H31N7O/c1-21-5-10-25(18-27(21)34-29-31-13-11-26(33-29)24-4-3-12-30-19-24)32-28(37)23-8-6-22(7-9-23)20-36-16-14-35(2)15-17-36/h3-13,18-19H,14-17,20H2,1-2H3,(H,32,37)
(H,31,33,34)

> <PUBCHEM_IUPAC_INCHIKEY>
KTUFNOKKBWMGRW-UHFFFAOYSA-N

> <PUBCHEM_IUPAC_NAME>
4-[(4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide

> <PUBCHEM_IUPAC_NAME_MARKUP>
4-[(4-methylpiperazin-1-yl)methyl]-<I>N</I>-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide

> <PUBCHEM_IUPAC_OPENEYE_NAME>
4-[(4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[[4-(3-pyridyl)pyrimidin-2-yl]amino]phenyl]benzamide

> <PUBCHEM_IUPAC_SYSTEMATIC_NAME>
4-[(4-methylpiperazin-1-yl)methyl]-N-[4-methyl-3-[(4-pyridin-3-ylpyrimidin-2-yl)amino]phenyl]benzamide

> <PUBCHEM_IUPAC_TRADITIONAL_NAME>
4-[(4-methylpiperazono)methyl]-N-[4-methyl-3-[[4-(3-pyridyl)pyrimidin-2-yl]amino]phenyl]benzamide

> <PUBCHEM_MOLECULAR_FORMULA>
C29H31N7O

> <PUBCHEM_MOLECULAR_WEIGHT>
493.6

> <PUBCHEM_MONOISOTOPIC_WEIGHT>
493.25998864

> <PUBCHEM_OPENEYE_CAN_SMILES>
CC1=C(C=C(C1)NC(=O)C2=CC=C(C=C2)CN3CCN(CC3)C)NC4=NC=CC(=N4)C5=CN=CC=C5

> <PUBCHEM_OPENEYE_ISO_SMILES>
CC1=C(C=C(C1)NC(=O)C2=CC=C(C=C2)CN3CCN(CC3)C)NC4=NC=CC(=N4)C5=CN=CC=C5

> <PUBCHEM_TOTAL_CHARGE>
0

> <PUBCHEM_XLOGP3_AA>
3.5

> <energy>
80.155374645437803

****
```

Fig 10: Conformer having the lowest energy of 80.155374645437803

RESULTS:

The Balloon software was downloaded through the website and the files were extracted from the zip files. The 2D conformations of the compound ‘Imatinib’ (PubChem ID: 5291) were downloaded from PubChem. The conformers were run on Command prompt for the query ‘Imatinib’ (PubChem ID: 5291) and 12 conformers were generated with the energies – 80.155374645437803, 81.121658905099963, 82.380720654829673, 88.428029960169937, 88.561380311804115, 89.024456798179727, 89.570855196822606, 90.977543177392846, 93.421321406231101, 93.684635828982579, 93.823190322942452 and 94.778222045701739. The conformer with the lowest energy 80.155374645437803 was selected.

CONCLUSION:

The application Balloon was explored and the 12 conformers for the query ‘Imatinib’ (PubChem ID: 5291) were generated. The conformer with the lowest energy i.e., 80.155374645437803 was selected for further downstream process.

REFERENCES:

1. Ohashi, N., & Kohno, T. (2020). Analgesic Effect of Acetaminophen: A Review of Known and Novel Mechanisms of Action. *Frontiers in pharmacology*, 11, 580289. <https://doi.org/10.3389/fphar.2020.580289>
 2. Balloon https://users.abo.fi/mivainio/balloon/change_log.htm
 3. Iqbal, N., & Iqbal, N. (2014). Imatinib: a breakthrough of targeted therapy in cancer. *Cancer research and practice*, 2014, 357027. <https://doi.org/10.1155/2014/357027>
 4. Imatinib: Uses, Interactions, Mechanism of Action | DrugBank Online. (n.d.). DrugBank. <https://go.drugbank.com/drugs/DB00619>
-

WEBLEM 4

Introduction to Pharmacophore Mapping with ZINCPharmer Mapping and Swiss PDB Viewer

INTRODUCTION:

When the 3D structure of a protein target is not yet characterized, and/or when a set of ligands (with or without known binding affinity) is available, pharmacophore models can be developed and utilized as search queries for virtual screening of databases. These pharmacophore models can vary from sub-structural type pharmacophores to feature-based pharmacophores, where pharmacophoric points are represented by chemical features such as hydrogen bond acceptors/donors, hydrophobic regions, acidic or basic features, etc. Additionally, when there is a need to consider the 3D structure, virtual screening must address increased complexity in terms of molecular functionality and flexibility, necessitating more advanced tools for data analysis. To incorporate this into virtual screening, a chemical function-based approach is typically the most versatile. The uniqueness of these pharmacophores lies in their broad definition, which represents different interaction types between organic molecules and proteins. The usefulness of such models as queries for 3D database searches has been recently reviewed. These pharmacophores can be generated from ligand sets or from the structure of an active site. Following the virtual screening and filtering process, a robust method for ranking the resulting hits based on their potential bioactivity is essential.

There are two main approaches to developing a pharmacophore: direct and indirect methods. Direct methods use information from both the ligand and the receptor, while indirect methods rely only on a collection of ligands that have been shown to interact with a particular receptor. Indirect methods are particularly useful when the structure of the receptor is unknown, which is often the case since crystal structures are available for less than 10% of drug targets. However, direct methods are becoming increasingly important due to the growing number of known protein structures resulting from the Structural Genomics project. Once a pharmacophore model is identified, it serves as a valuable tool for discovering and developing new lead compounds.

Steps in Identifying a Pharmacophore

Generally, all algorithms for pharmacophore identification follow these six steps:

1. Input
2. Conformational Search
3. Feature extraction
4. Structure Representation
5. Pattern Identification
6. Scoring

ZINCPharmer:

ZINCPharmer is free pharmacophore search software for screening the purchasable subset of the ZINC database (updates occur monthly). ZINCPharmer can import LigandScout and MOE pharmacophore definitions, as well as identify pharmacophore features directly from structure. An interaction pharmacophore can be derived from a group of known active ligands by finding a consensus pharmacophore that all these ligands can adopt. These methods don't need a ligand-bound structure but can be computationally intensive if the input set includes many flexible ligands. PharmaGist is a free web server capable of identifying a consensus pharmacophore for a set of up to 32 ligands within minutes. On the other hand, structure-based methods require a ligand-bound structure and identify a potential pharmacophore by examining the interaction site. ZINCPharmer allows users to generate an initial pharmacophore hypothesis directly from structures available in the PDB (Protein Data Bank) and also supports importing pharmacophore definitions created using more computationally demanding techniques available in third-party tools.

Given a library of explicit compound conformations, conformers that match a 3D pharmacophore can be identified using either fingerprint-based or alignment-based methods. Fingerprints are effective for similarity metrics but can provide less precise results because they discretize the pharmacophore representation. The EDULISS online database offers fingerprint-based screening for a single-conformer library of a few million compounds, though the query fingerprint needs to be manually created from pairwise distance constraints. Alignment-based methods, while more accurate and interpretable, require more computational resources. Screening a library of fewer than a million conformers might take minutes to hours. However, since there are far fewer protein targets compared to potential ligands, alignment-based pharmacophore screening can be very effective for reverse screening to identify matching protein targets rather than ligands. PharmMapper, for instance, takes a single ligand and screens a database of over 7,000 receptors to find potential targets.

Swiss PDB Viewer:

Swiss-PdbViewer (also known as DeepView) is a user-friendly application that enables the simultaneous analysis of multiple proteins. It allows for the superimposition of proteins to deduce structural alignments and compare active sites or other relevant regions. The software provides an intuitive graphical and menu interface that simplifies the extraction of information on amino acid mutations, hydrogen bonds, angles, and distances between atoms.

Developed by Nicolas Guex since 1994, Swiss-PdbViewer was originally closely integrated with SWISS-MODEL, an automated homology modeling server created at the Structural Bioinformatics Group within the Swiss Institute of Bioinformatics (SIB) at the Biozentrum in Basel.

INSTALLATION STEPS:

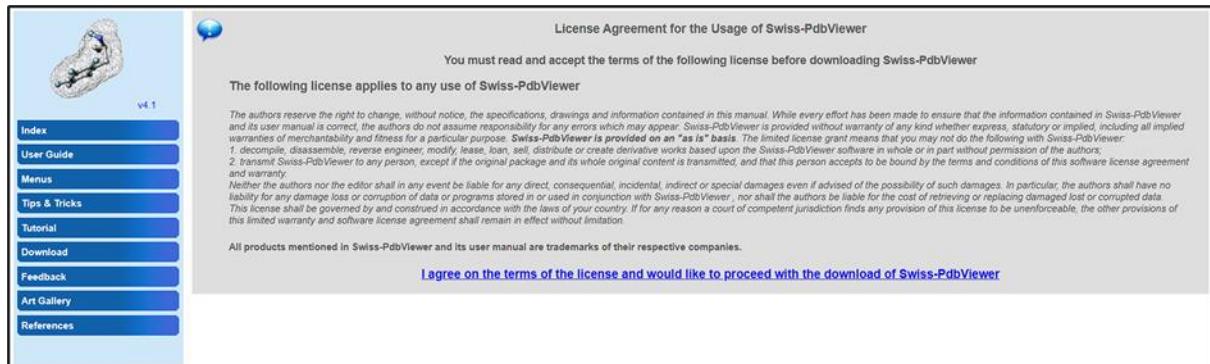


Fig 1: Installation page for Swiss PDB Viewer: Click on ‘I Agree’

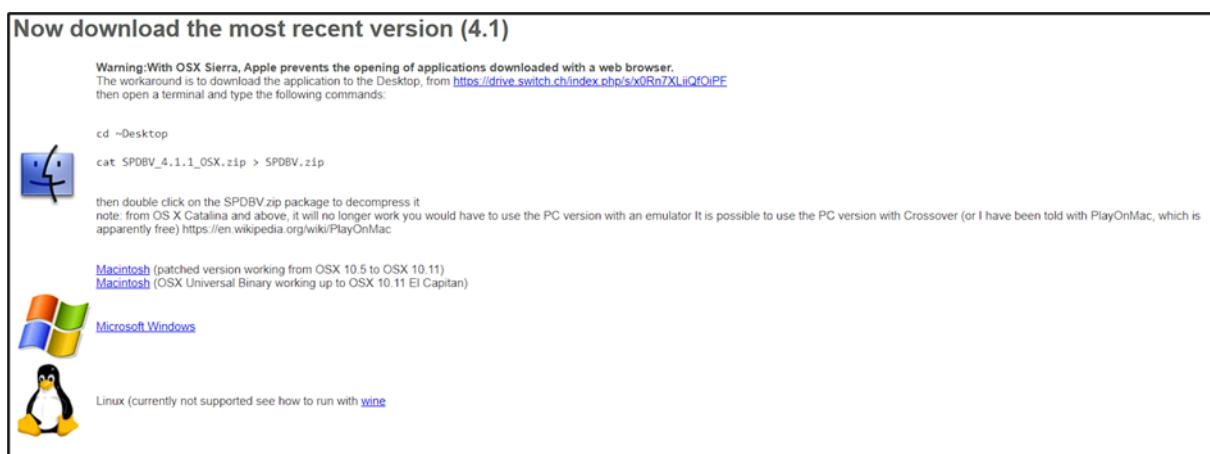


Fig 2: Click on Microsoft Windows

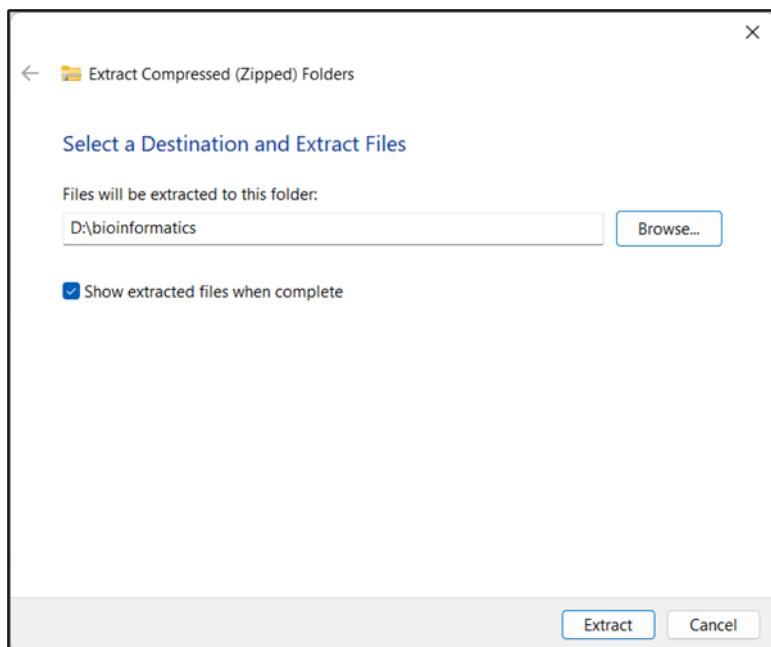


Fig 3: Extract files into your folder

stuff	31-08-2024 22:43	File folder
download	31-08-2024 22:42	File folder
scripts	31-08-2024 22:43	File folder
temp	25-07-2012 17:24	File folder
usrstuff	31-08-2024 22:43	File folder
1crn.pdb	31-08-2024 22:42	BIOVIA Discovery ... 47 KB
libcurl.dll	31-08-2024 22:42	Application extens... 297 KB
LICENSE.txt	31-08-2024 22:42	Text Document 2 KB
README.txt	31-08-2024 22:42	Text Document 4 KB
spdbv.exe	31-08-2024 22:42	Application 2,204 KB
VERSION_NEWS.txt	31-08-2024 22:42	Text Document 10 KB

Fig 4: Open the application spdbv.exe

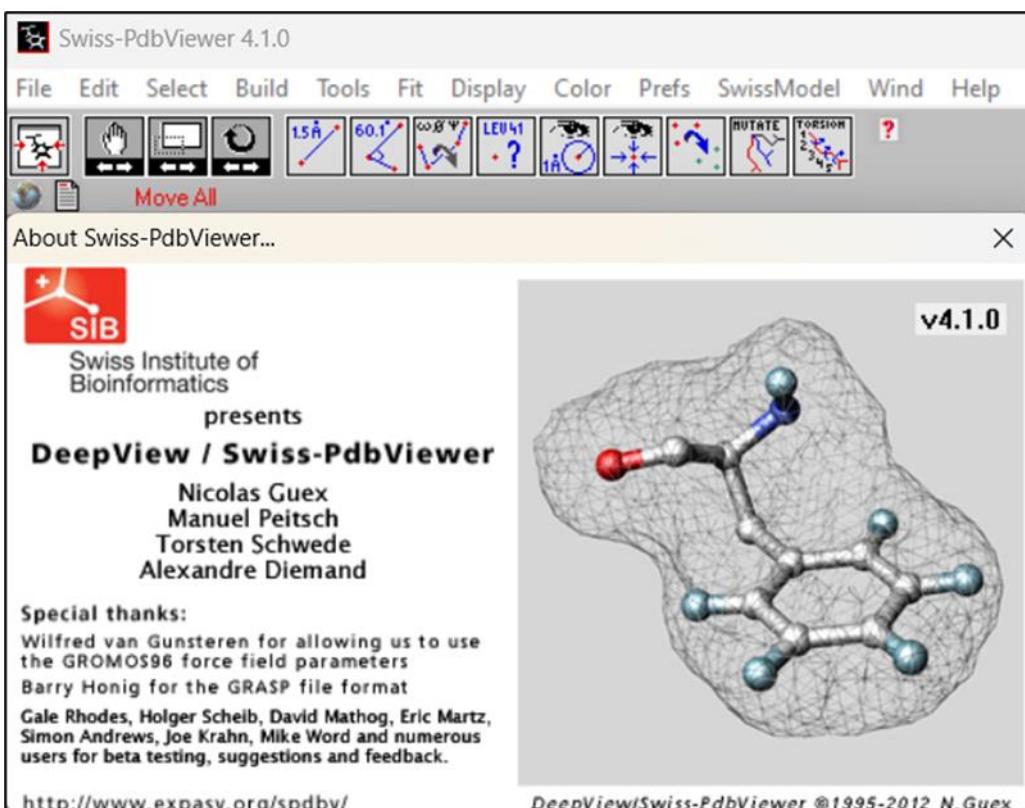


Fig 5: Homepage for Swiss PDB Viewer

REFERENCES:

1. Wolber, G., & Sippl, W. (2015). Pharmacophore identification and Pseudo-Receptor modeling. In Elsevier eBooks (pp. 489–510). <https://doi.org/10.1016/b978-0-12-417205-0.00021-3>
 2. Giordano, D., Biancaniello, C., Argenio, M. A., & Facchiano, A. (2022). Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceutics* (Basel, Switzerland), 15(5), 646. <https://doi.org/10.3390/ph15050646>
 3. Lee, J. Y., Krieger, J. M., Li, H., & Bahar, I. (2020). Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations. *Protein science: a publication of the Protein Society*, 29(1), 76–86. <https://doi.org/10.1002/pro.3732>
-

WEBLEM 4(A)
ZINCPharmer Software and Swiss PDB Viewer
(URL: <http://zincpharmer.csb.pitt.edu/>)

AIM:

To generate and analyze the pharmacophore map for the query of 'Acyclovir' (PubChem ID: 135398513) using ZINCPharmer software.

INTRODUCTION:

ZINCPharmer

ZINCPharmer is an advanced pharmacophore search tool designed for efficiently screening purchasable compounds from the extensive ZINC database. It leverages the open-source Pharmer technology, optimized for rapid searches through large datasets of fixed conformers, typically completing queries in under a minute. The software supports the integration of pharmacophore definitions from popular tools like LigandScout and MOE, and can also automatically derive pharmacophore features from molecular structures, enhancing its utility for researchers in medicinal chemistry and drug discovery.

ZINCPharmer is particularly valuable for targeting protein-protein interactions, with features like the PocketQuery function enabling the design of tailored pharmacophores. The platform offers extensive user support, including interactive examples, a comprehensive guide, and video tutorials, making it accessible to both novice and experienced users. Funded by the National Institute of General Medical Sciences, ZINCPharmer is continuously updated to include the latest purchasable compounds, ensuring its relevance in ongoing research. Researchers are encouraged to cite the associated 'Nucleic Acids Research' publication to acknowledge the software's development. Overall, ZINCPharmer is a critical tool for advancing drug discovery, providing a robust, scalable, and user-friendly platform for identifying potential drug candidates.

Swiss-PDBViewer

Swiss-PDBViewer, also known as DeepView, is a robust and user-friendly software developed by Nicolas Guex since 1994 for the analysis and manipulation of protein structures. Widely used in structural biology, this tool facilitates the detailed examination of three-dimensional protein structures, offering functionalities such as the simultaneous superimposition of multiple protein structures to deduce structural alignments and compare active sites. This capability is essential for identifying structural similarities and differences, providing insights into protein functions and interactions.

The software includes comprehensive tools for measuring amino acid mutations, hydrogen bonds, angles, and atomic distances, all accessible through an intuitive graphical user interface. Originally developed in conjunction with SWISS-MODEL, an automated homology modelling server, Swiss-PDBViewer continues to aid in protein modelling, enabling users to thread protein sequences onto 3D templates and evaluate their structural compatibility.

A notable feature of Swiss-PDBViewer is its ability to generate ray-traced quality images using POV-Ray, useful for creating high-quality visual representations of protein structures. Available across multiple operating systems, including Macintosh, Windows, Linux, and Irix, and supported by the SBGrid Consortium, Swiss-PDBViewer remains a vital tool in structural biology, biochemistry, and drug discovery, facilitating advanced protein analysis and visualization.

Acyclovir

Acyclovir, also known as aciclovir, is an antiviral medication primarily used to treat infections caused by certain types of viruses, particularly those in the herpes virus family. It is effective against herpes simplex virus (HSV) infections, varicella-zoster virus (which causes chickenpox and shingles), and can also be used in the prevention of cytomegalovirus infections in immunocompromised patients.

Acyclovir is a deoxyribonucleoside analog that mimics the structure of guanosine. Once inside the body, it is converted into its active form by viral thymidine kinase, leading to the formation of acyclovir monophosphate. This compound is then further phosphorylated by host cell enzymes to acyclovir triphosphate, which interferes with viral DNA synthesis, effectively slowing down the growth and spread of the virus.

METHODOLOGY:

1. Download 3D structure of protein ‘Crystal structure of purine nucleoside phosphorylase from mycobacterium tuberculosis in complex with acyclovir’ (PDB ID: 3IX2) from Protein Data Bank (PDB) (<https://www.rcsb.org/>).
2. Download 3D Structure of ‘Acyclovir’ molecule (PubChem ID: 135398513) in SDF Format from NCBI’s PubChem (<https://pubchem.ncbi.nlm.nih.gov/>).
3. Delete the all heteroatoms and connectivity data of atoms by removing all lines containing the ‘HETATM’ and ‘CONECT’, respectively.
4. Download Swiss-PDBViewer (aka DeepView) from its official site (<https://spdbv.unil.ch/>) by accepting its terms of license for the suitable OS. Once downloaded, decompress and extract the file contents.
5. Open the extracted file and launch the application file named as ‘spdbv’ to launch the Swiss-PDBViewer (SPDBV) software.
6. To load downloaded protein in SPDBV, from toolbar select ‘File’ > ‘Open PDB File...’ and select the protein molecule downloaded from PDB.
7. For Energy Minimization, select all the molecules ‘Select’ > ‘All...’ or ‘Ctrl + A’ and then ‘Tools’ > ‘Energy Minimization...’ or ‘Ctrl + N’. A progress bar will appear during the

process of energy minimization and once minimization is completed, a list of all the amino acids in protein with their bonds, angles, torsion, etc. along with their energy will be seen.

8. Save the energy minimized structure by ‘File’ > ‘Save’ > ‘Current Layer’ or ‘Ctrl + S’ and name the energy minimized structure accordingly (e.g.: 3IX2-em.pdb) and close SPDBV.
9. Open ZINCPharmer and go to ‘Search ZINC’ from the homepage.
10. To load protein structure, click on the ‘Load Receptor’ button at the bottom of the screen and selected the energy minimized PDB file (e.g.: 3IX2-em.pdb)
11. To load the chemical entity downloaded from PubChem, click on ‘Load Features’ and select the SDF file of the chemical entity (efavirenz) downloaded from PubChem.
12. Once the chemical structure is mapped onto the protein structure a table will be displayed in the ‘Pharmacophore’ Tab, from the ‘Enabled’ column check on all the Pharmacophore Classes.
13. If filters are needed, necessary filters can be applied from ‘Filters’ Tab. Here, 2 filters are applied:
 14. Molecular Weight: <=500 kDa & >= 100 kDa
 15. Rotatable Bonds: <= 8 & >=2
16. The ‘Viewer’ Tab can be used to customize how the protein and chemical entity looks in the view panel.
17. Once done, click on submit query to retrieve the list of chemical compounds that can be mapped onto the protein.
18. ‘Save Results’ option can be used to download the SDF file of all the compounds obtained using ZINCPharmer
19. Click on the ZINC ID from name column of the ‘Results’ Table to analyze the chemical entity.

OBSERVATIONS:

The screenshot shows the RCSB PDB homepage. In the search bar, the query 'acyclovir' has been entered. The search results page displays various protein structures related to acyclovir, including its chemical structure and biological activity. A sidebar on the left provides links to 'Deposit', 'Search', 'Visualize', 'Analyze', 'Download', and 'Learn'. The main content area features a 'Welcome' message and sections for 'Explore NEW Features' and 'PDB-101 Training Resources'. A 'Carbon Capture Mechanisms' section is also visible.

Fig 1: Homepage of Protein Data Bank (PDB) with the query ‘Acyclovir’

The screenshot shows the detailed view for the protein structure with PDB ID 3IX2. The page includes a 3D ribbon model of the protein assembly. Key information displayed includes the title 'CRYSTAL STRUCTURE OF PURINE NUCLEOSIDE PHOSPHORYLASE FROM MYCOBACTERIUM TUBERCULOSIS IN COMPLEX WITH ACYCLOVIR', the classification as a 'TRANSFERASE', and the organism as 'Mycobacterium tuberculosis variant bovis AF2122'. The page also lists experimental data such as the method (X-RAY DIFFRACTION), resolution (2.10 Å), and R-factors. On the right, there is a sidebar for file download options, with 'PDB Format' highlighted by a red box. A validation report is shown at the bottom right.

Fig 2: Download the protein with the PDB ID: 3IX2 in PDB format

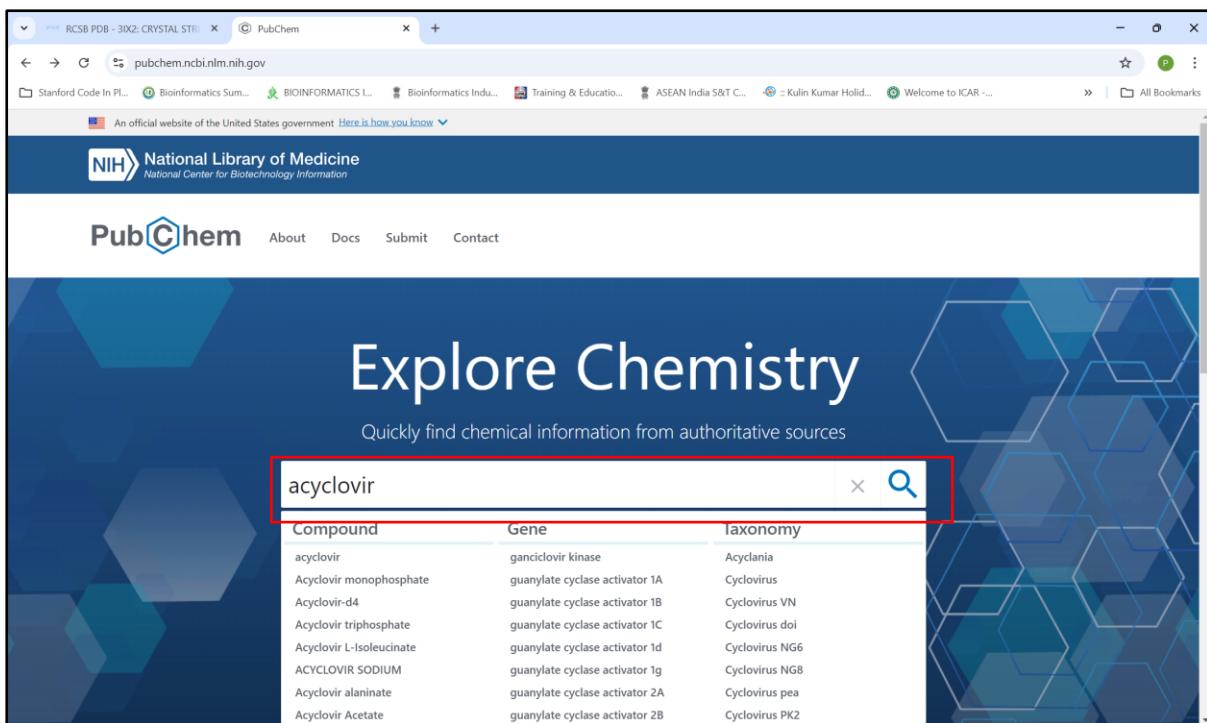


Fig 3: Homepage of PubChem with the query ‘Acyclovir’

The screenshot shows the search results for 'acyclovir'. A red box highlights the 'BEST MATCH' section. It displays the chemical structure of acyclovir, its CID (13539853), molecular formula (C₈H₁₁N₅O₃), MW (225.2g/mol), IUPAC Name (2-amino-9-(2-hydroxyethoxymethyl)-1H-purin-6-one), Isomeric SMILES (C1=NC2=C(N1COCCO)N=C(NC2=O)N), InChIKey (MKUXAQIEYXACX-UHFFFAOYSA-N), InChI (InChI=1S/C8H11N5O3/c9-8-11-6-5(7(15)12-8)10-3-13(6)4-16-2-1-14/h3,14H,1-2,4H2,(H3,9,11,12,15)), and Create Date (2019-01-10). Below this section are tabs for Summary, Similar Structures Search, Related Records, and PubMed (MeSH Keyword). At the bottom, there are tabs for Compounds, Substances, BioAssays, Literature, and Patents.

Fig 4: Search results obtained for the query ‘Acyclovir’

NIH National Library of Medicine
National Center for Biotechnology Information

PubChem About Docs Submit Contact

Search PubChem

COMPOUND SUMMARY

Acyclovir

PubChem CID 135398513

Structure

Chemical Safety

Molecular Formula C8H11N5O3

Synonyms acyclovir

CONTENTS

- 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
- 10 Identification
- 11 Safety and Hazards
- 12 Toxicity
- 13 Associated Disorders and Diseases
- 14 Literature
- 15 Patents
- 16 Interactions and Databases

Fig 5: ‘Acyclovir’ (PubChem ID: 135398513)

NIH National Library of Medicine

PubChem

COMPOUND SUMMARY

Acyclovir

PubChem CID

Structure

Chemical Safety

Molecular Formula C8H11N5O3

Synonyms acyclovir

DOWNLOAD

Download data used to display this page

2D Structure

3D Conformer

Looking to Download a PDF of This Page?

Please use print functionality available in your browser and look for a save as PDF option.

Note that some sections on this page might be loaded on demand (when you scroll to them), and thus, before saving the page to PDF, you would first want to scroll to the bottom of the page to make sure that everything is loaded. Alternatively, you may open a section of interest in a new window and save it as PDF.

14 Literature

15 Patents

16 Interactions and Databases

Fig 6: Download the 3D structure of ‘Acyclovir’ (PubChem ID: 135398513) in SDF format

```

3ix2.pdb
File Edit View
HEADER TRANSFERASE          03-SEP-09 3IX2
TITLE CRYSTAL STRUCTURE OF PURINE NUCLEOSIDE PHOSPHORYLASE FROM
TITLE 2 MYCOBACTERIUM TUBERCULOSIS IN COMPLEX WITH ACYCLOVIR
CAVEAT 3IX2 ALA C 43 HAS WRONG CHIRALITY AT ATOM CA THR C 255 HAS WRONG
CAVEAT 2 3IX2 CHIRALITY AT ATOM CA THR C 255 HAS WRONG CHIRALITY AT ATOM
CAVEAT 3 3IX2 CB LEU C 261 HAS WRONG CHIRALITY AT ATOM CA
COMPND MOL_ID: 1;
COMPND 2 MOLECULE: PURINE NUCLEOSIDE PHOSPHORYLASE;
COMPND 3 CHAIN: A, B, C;
COMPND 4 FRAGMENT: PURINE NUCLEOSIDE PHOSPHORYLASE;
COMPND 5 SYNONYM: PNP,PU-NPASE,INOSINE PHOSPHORYLASE,INOSINE-GUANOSINE
COMPND 6 PHOSPHORYLASE;
COMPND 7 EC: 2.4.2.1;
COMPND 8 ENGINEERED: YES
SOURCE MOL_ID: 1;
SOURCE 2 ORGANISM_SCIENTIFIC: MYCOBACTERIUM TUBERCULOSIS VARIANT BOVIS
SOURCE 3 AF2122/97;
SOURCE 4 ORGANISM_TAXID: 233413;
SOURCE 5 STRAIN: ATCC BAA-935 / AF2122/97;
SOURCE 6 GENE: PUNA, DEOD, BQ2927_MB3335;
SOURCE 7 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE 8 EXPRESSION_SYSTEM_TAXID: 516
KEYWDS MYCOBACTERIUM TUBERCULOSIS, PURINE NUCLEOSIDE PHOSPHORYLASE,
KEYWDS 2 ACYCLOVIR TRANSFERASE
EXPOPTA X-RAY DIFFRACTION
AUTH W.F.DE AZEVEDO JR., L.A.BASSO,D.S.SANTOS
REVDAT 3 20-SEP-23 3IX2 1 REMARK
REVDAT 2 01-MAR-23 3IX2 1 SOURCE
REVDAT 1 23-JUL-21 3IX2 0
JRNL AUTH R.A.CACERES,L.F.TIMMERS,R.G.DUCATI,D.O.DA SILVA,L.A.BASSO,
JRNL AUTH 2 W.F.DE AZEVEDO JR., D.S.SANTOS
JRNL TITL CRYSTAL STRUCTURE AND MOLECULAR DYNAMICS STUDIES OF PURINE
JRNL TITL 2 NUCLEOSIDE PHOSPHORYLASE FROM MYCOBACTERIUM TUBERCULOSIS
JRNL TITL 3 ASSOCIATED WITH ACYCLOVIR.
JRNL REF BIOCHIMIE V. 94 155 2012
JRNL REF ISSN 0308-9084
JRNL PMID 22833138
JRNL DOI 10.1016/J.BIOCHI.2011.10.003
REMARK 2
REMARK 2 RESOLUTION. 2.10 ANGSTROMS.
REMARK 3
REMARK 3 REFINEMENT.
REMARK 3 PROGRAM : REFMAC 5.2.0019
REMARK 3 AUTHORS : MURSHUDOV,SKUBAK,LEBEDEV,PANNU,STEINER,
REMARK 3 : NICHOLLS,WINN,LONG,VAGIN
REMARK 3
Ln 1 Col 1 5,52,339 characters

```

Fig 7: Protein (PDB ID: 3IX2) file viewed in Notepad

```

3ix2.pdb
File Edit View
CONECT 5711 5709 5712
CONECT 5712 5702 5705 5711
CONECT 5713 5714 5715 5716 5717
CONECT 5714 5713
CONECT 5715 5713
CONECT 5716 5713
CONECT 5717 5713
CONECT 5718 5719 5720
CONECT 5719 5718
CONECT 5720 5721 5721
CONECT 5721 5720 5722
CONECT 5722 5721 5723
CONECT 5723 5722 5724 5733
CONECT 5724 5723 5725
CONECT 5725 5724 5726
CONECT 5726 5725 5727 5733
CONECT 5727 5726 5728 5729
CONECT 5728 5727
CONECT 5729 5727 5730
CONECT 5730 5729 5731 5732
CONECT 5731 5730
CONECT 5732 5730 5733
CONECT 5733 5723 5726 5732
CONECT 5734 5735 5736 5737 5738
CONECT 5735 5734
CONECT 5736 5734
CONECT 5737 5734
CONECT 5738 5734
CONECT 5739 5740 5741
CONECT 5740 5739
CONECT 5741 5739 5742
CONECT 5742 5741 5743
CONECT 5743 5742 5744
CONECT 5744 5743 5745 5754
CONECT 5745 5744 5746
CONECT 5746 5745 5747
CONECT 5747 5746 5748 5754
CONECT 5748 5747 5749 5750
CONECT 5749 5748
CONECT 5750 5748 5751
CONECT 5751 5750 5752 5753
CONECT 5752 5751
CONECT 5753 5751 5754
CONECT 5754 5744 5747 5753
MASTER 426 0 6 31 30 0 0 6 6139 3 63 63
END

```

Fig 8: Protein (PDB ID: 3IX2) file end with HETATM (heteroatoms) and CONECT (connection) details

Ln 6303, Col 81 5,10,758 characters

100% Unix (LF) UTF-8

Fig 9: Protein (PDB ID: 3IX2) processed file end without HETATM (heteroatoms) and CONECT (connection) details

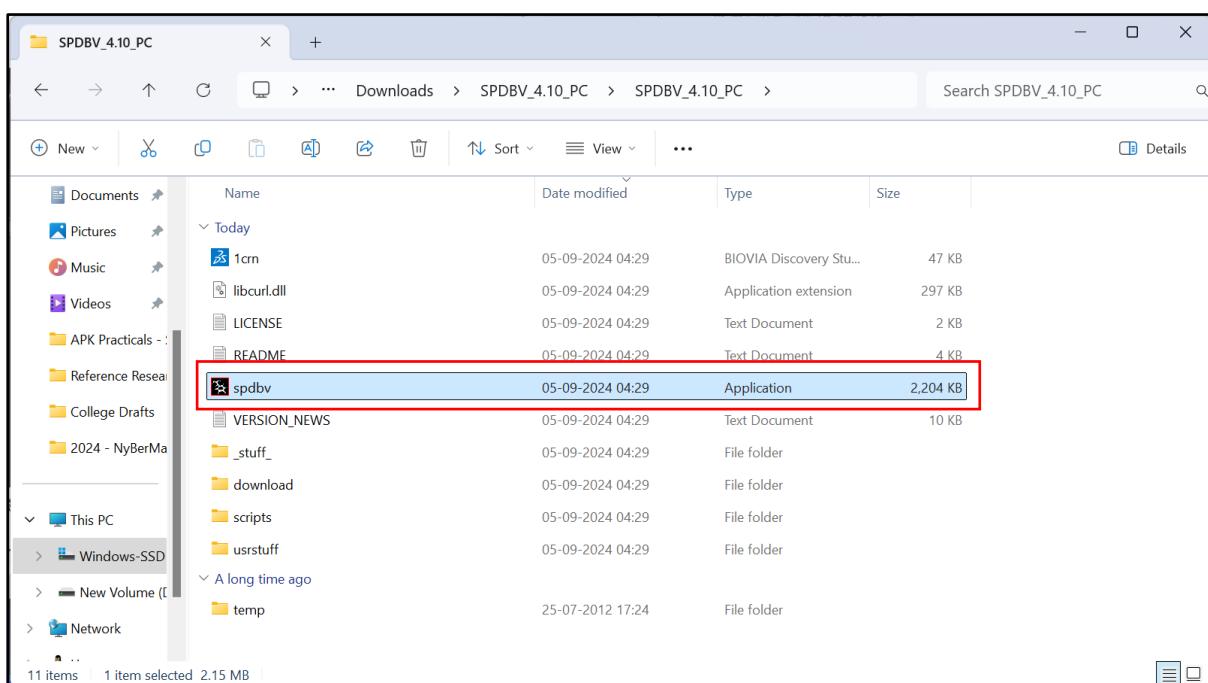


Fig 10: Launching SPDBV

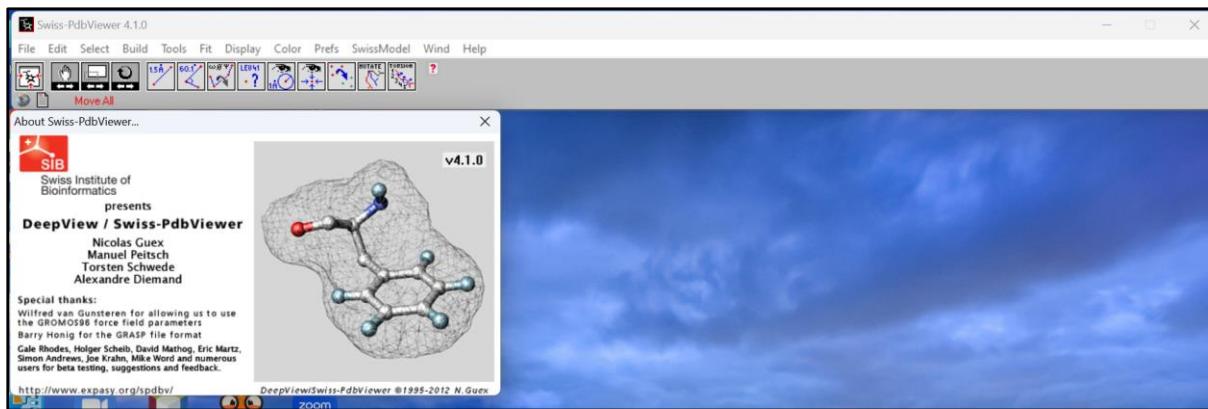


Fig 11: SPDBV Default Interface

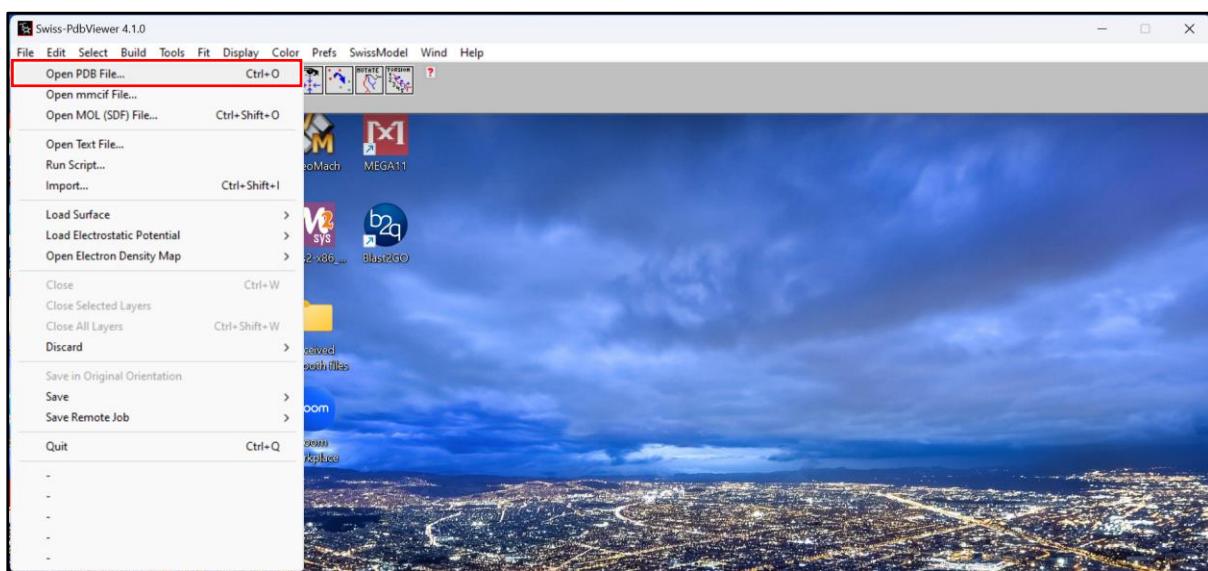


Fig 12: Loading PDB file in SPDBV

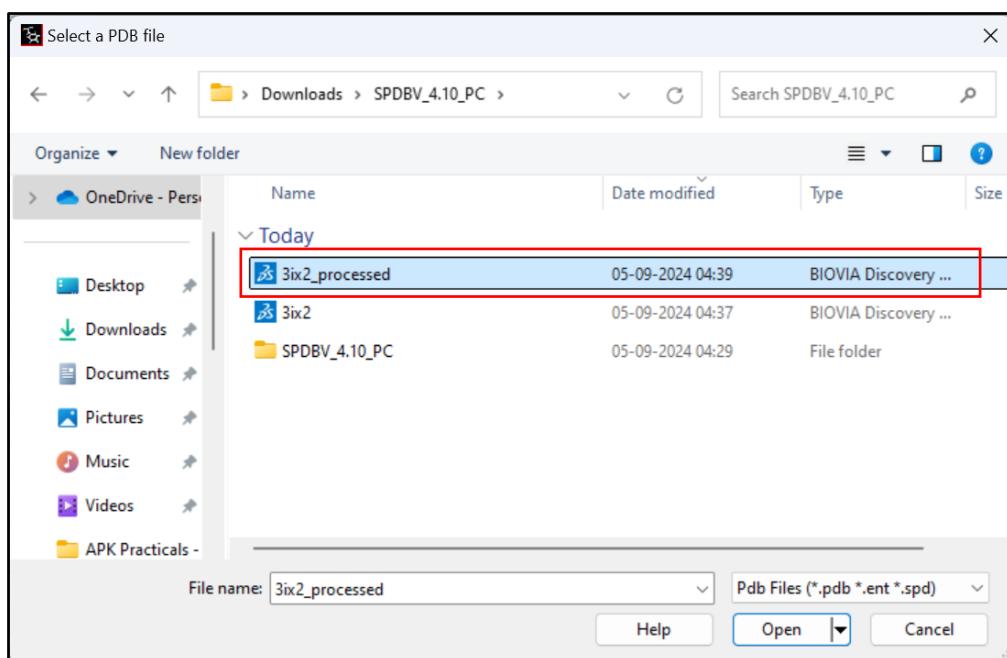


Fig 13: Selecting the processed PDB file (without HETATM and CONECT)

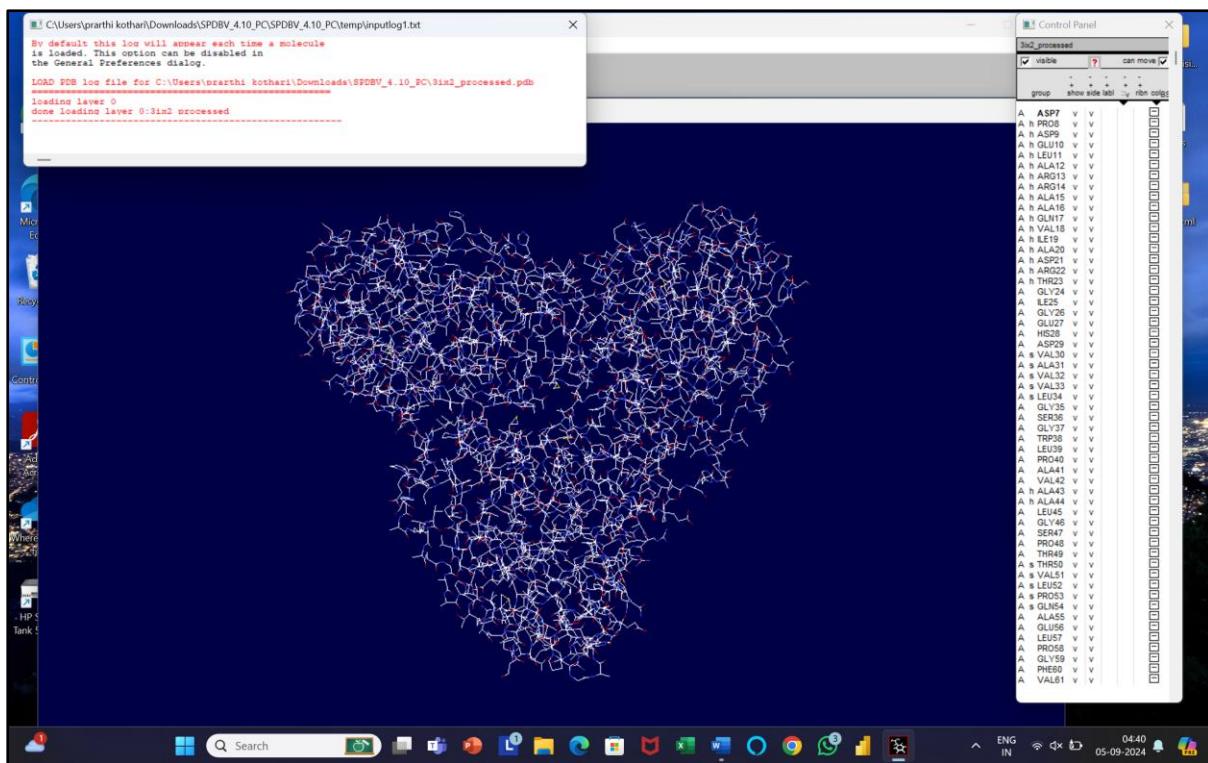


Fig 14: Processed PDB file loaded in SPDBV

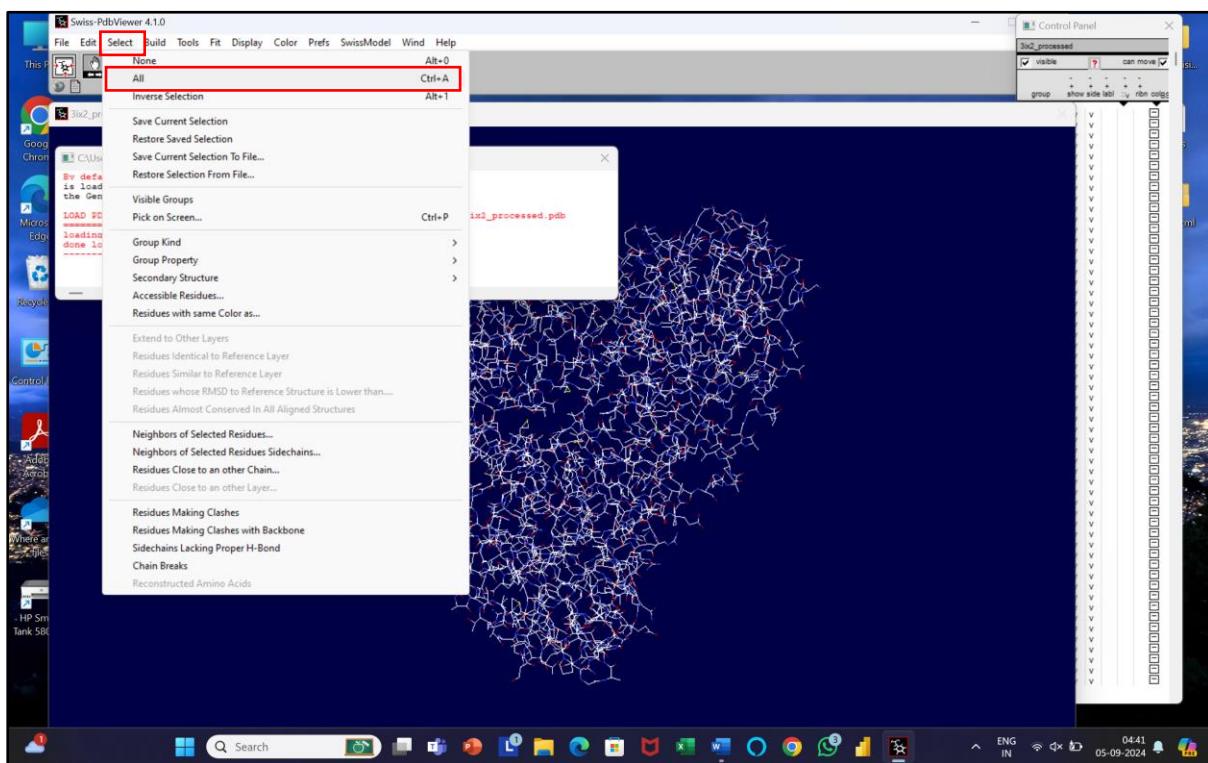


Fig 15: Selecting all atoms for energy minimization in SPDBV

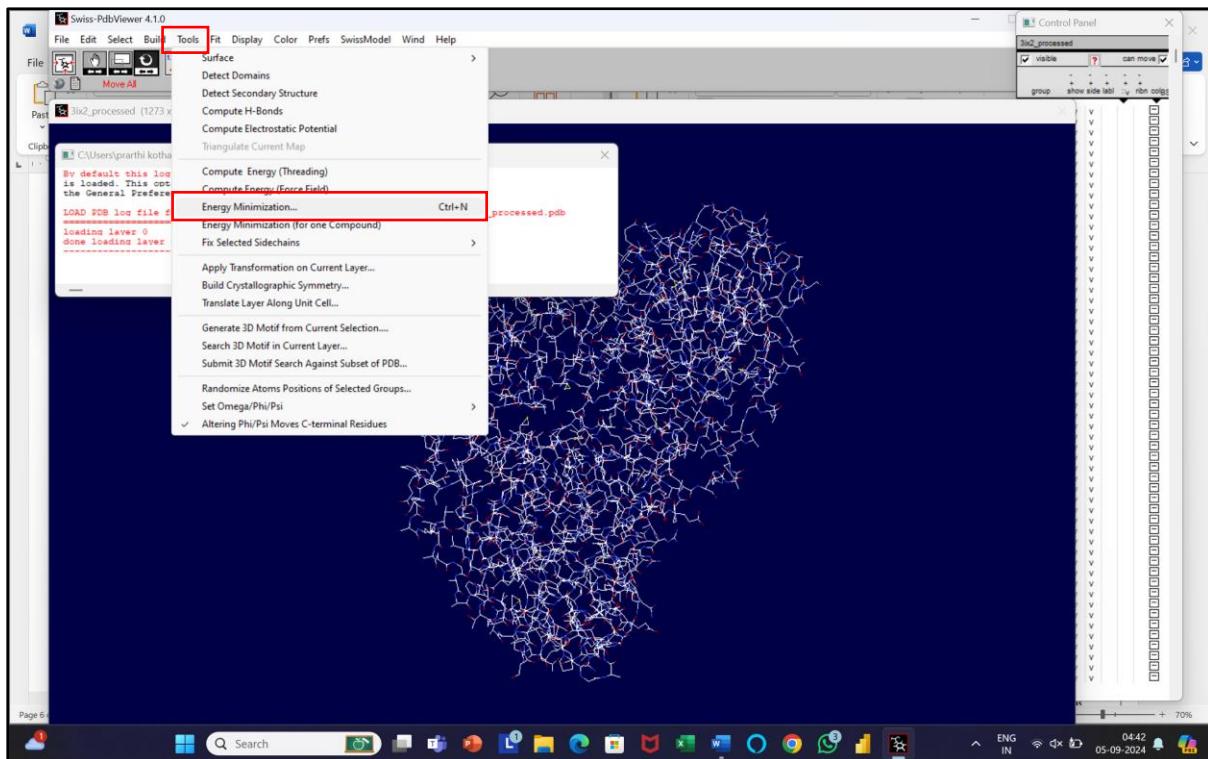


Fig 16: Minimizing energy of the protein (PDB ID: 3IX2)

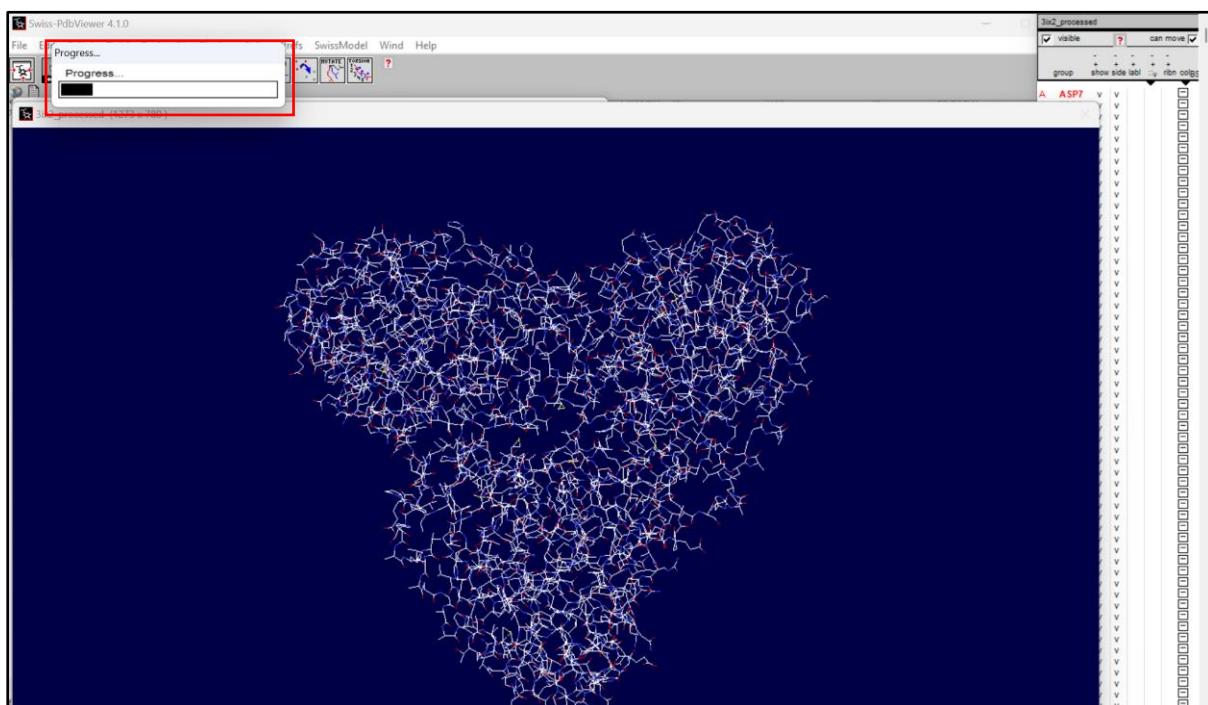


Fig 17: Energy Minimization of protein processing

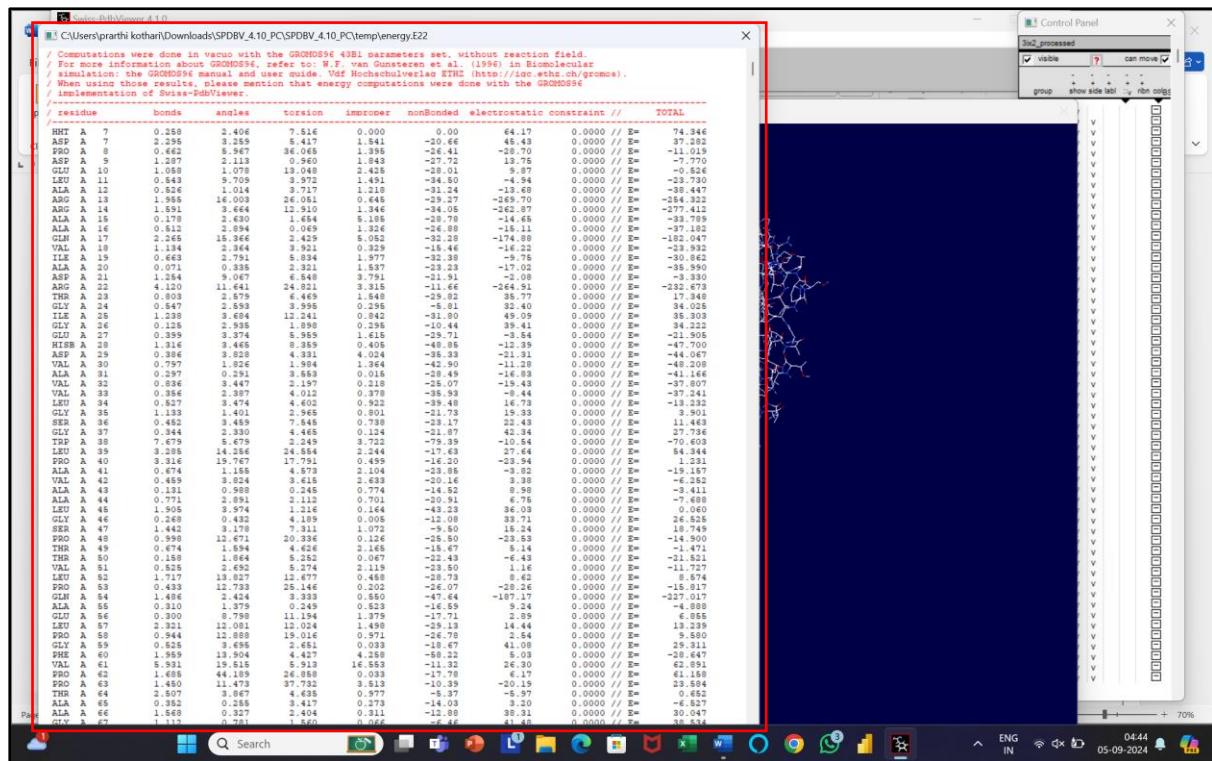


Fig 18: Details of energy minimization of each amino acid of the protein 3IX2

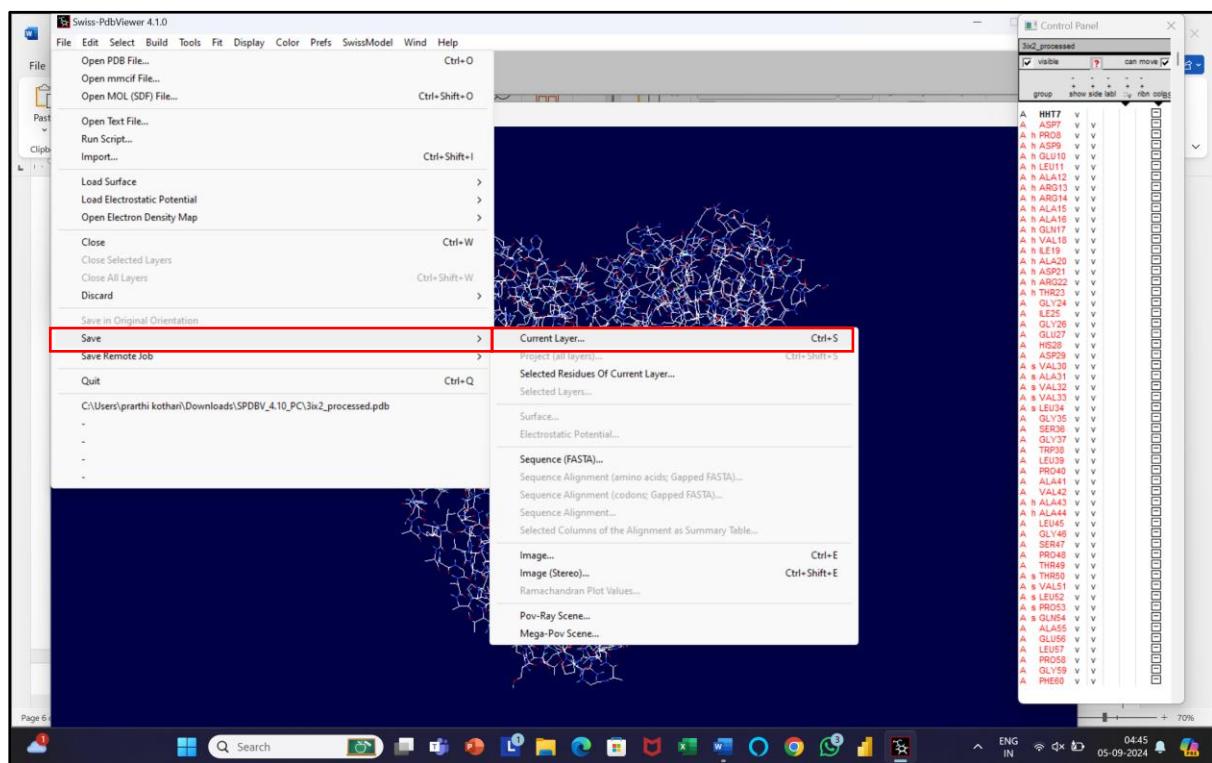


Fig 19: Saving the energy minimized protein

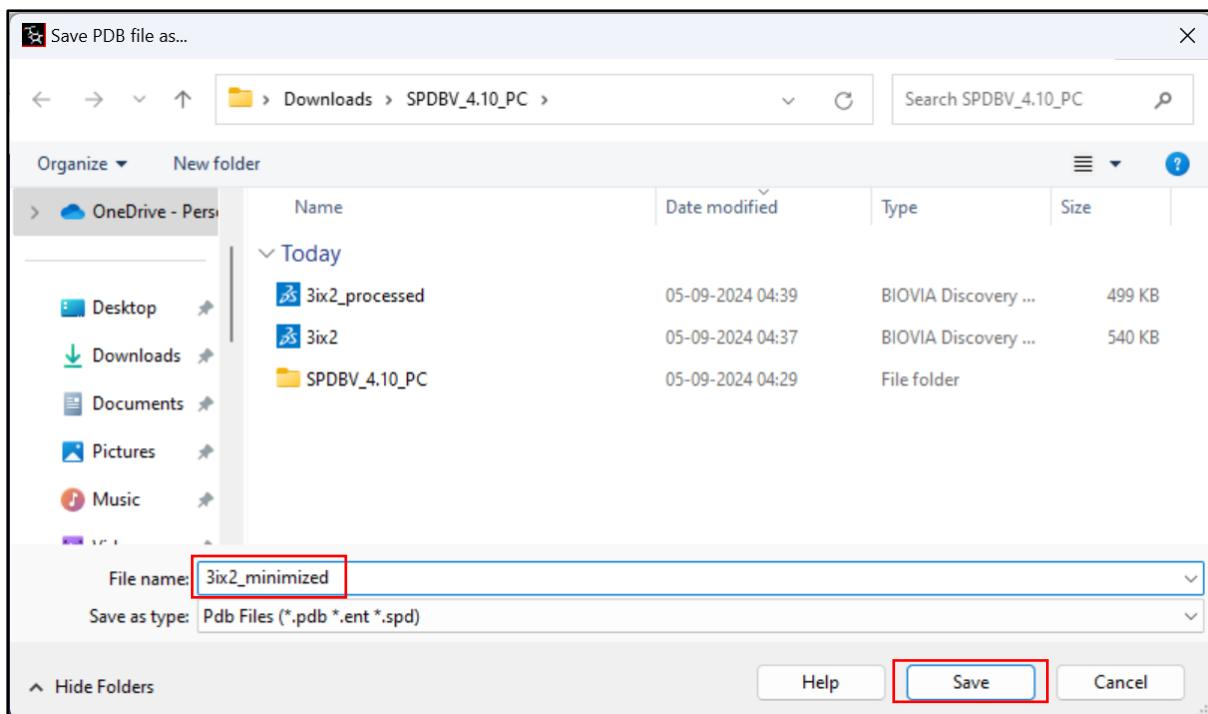


Fig 20: Saving the energy minimized protein in the desired folder

Fig 21: Homepage of ZINCPharmer

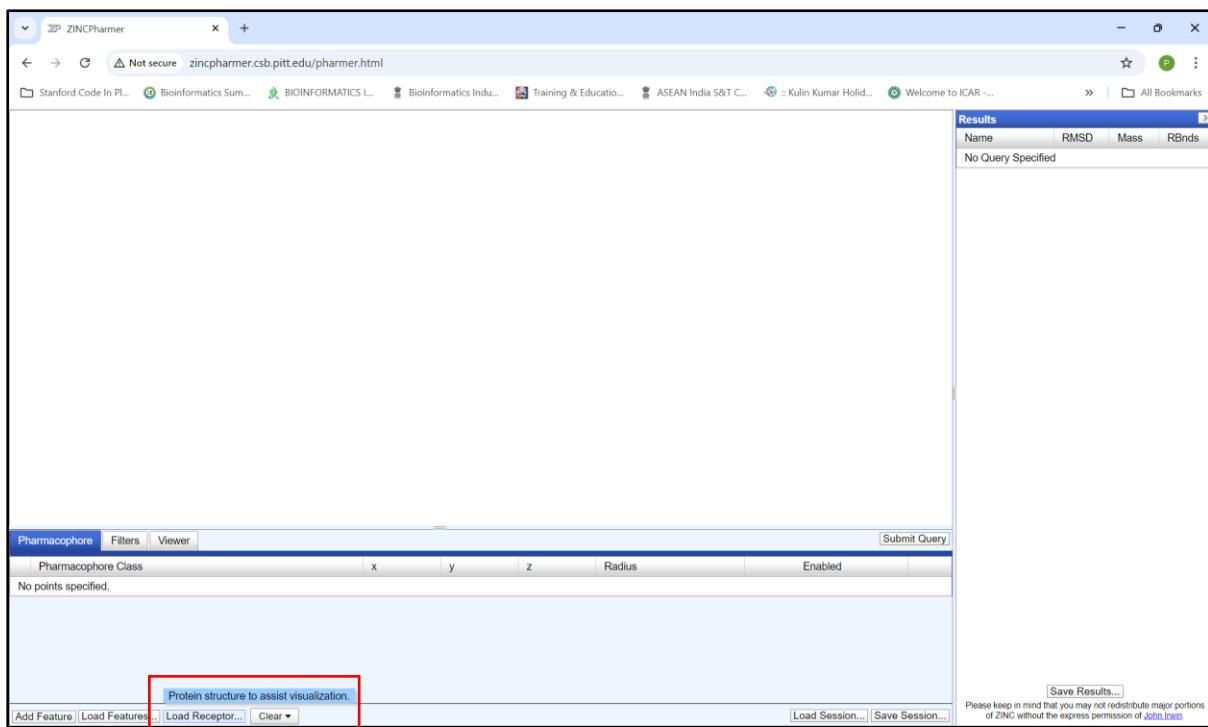


Fig 22: ZINCPharmer software page and loading energy minimized protein in ZINCPharmer

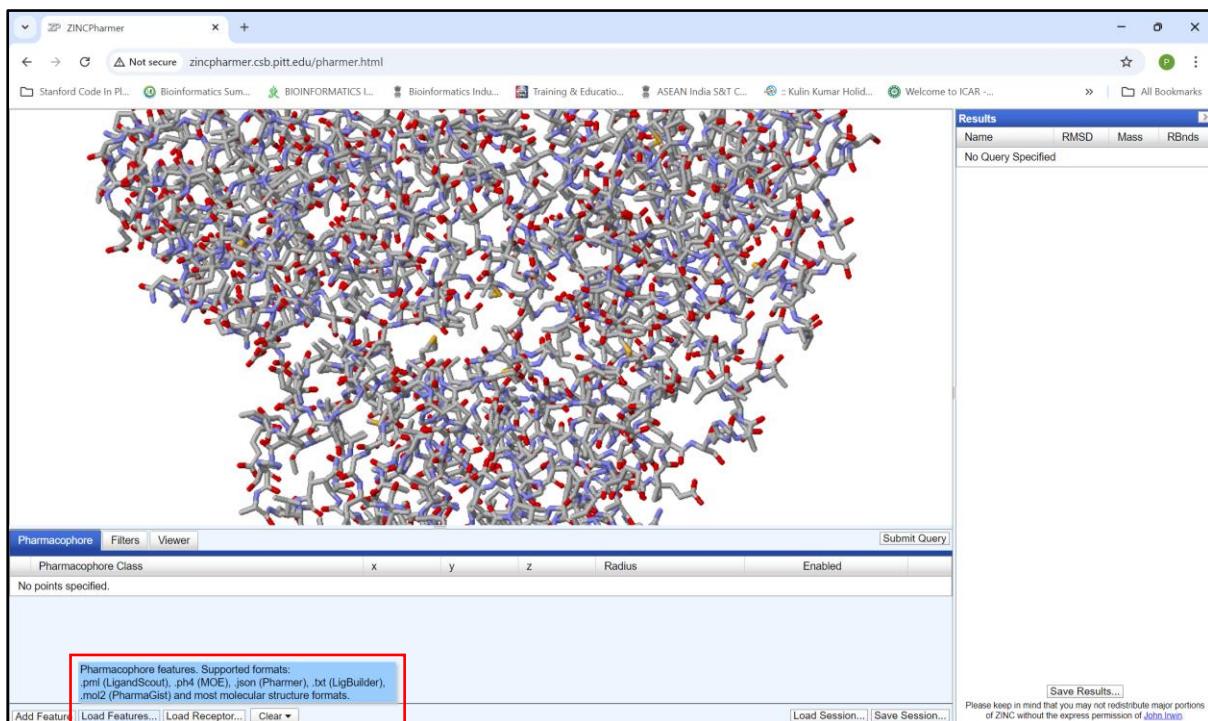


Fig 23: Visualizing energy minimized protein and loading ‘Acyclovir’ (PubChem ID: 135398513) in ZINCPharmer

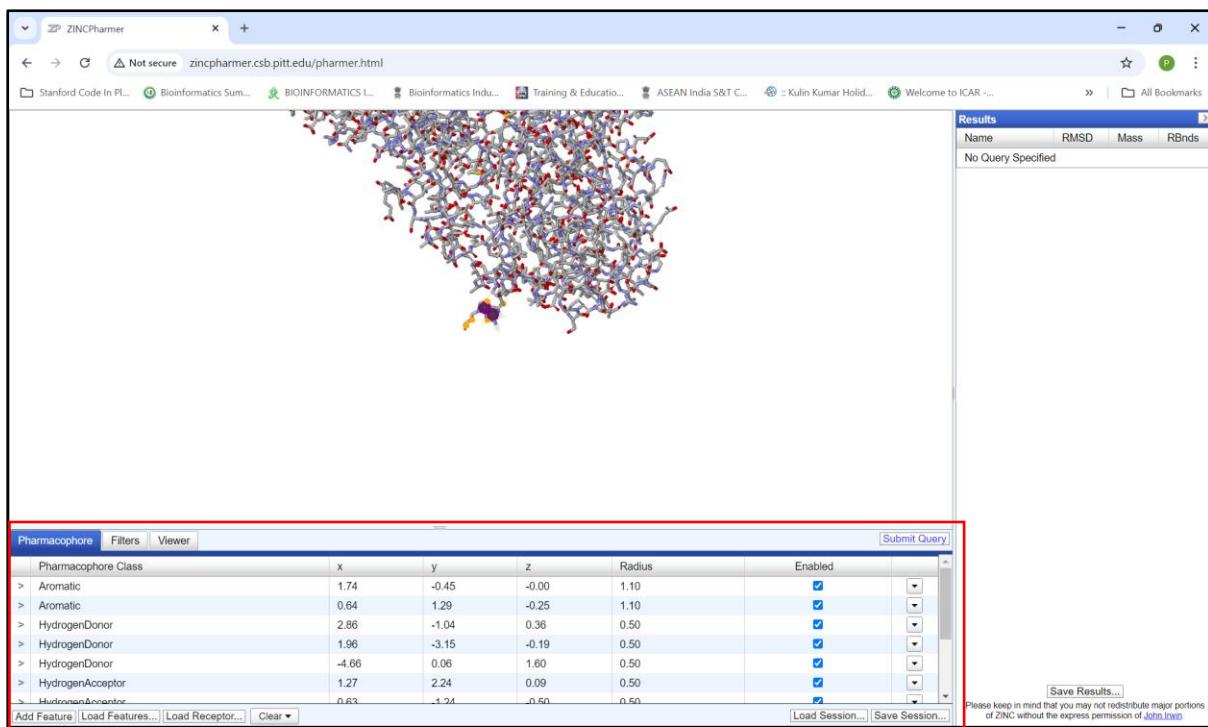


Fig 24: Setting Pharmacophore parameters in ‘Pharmacophore’ tab

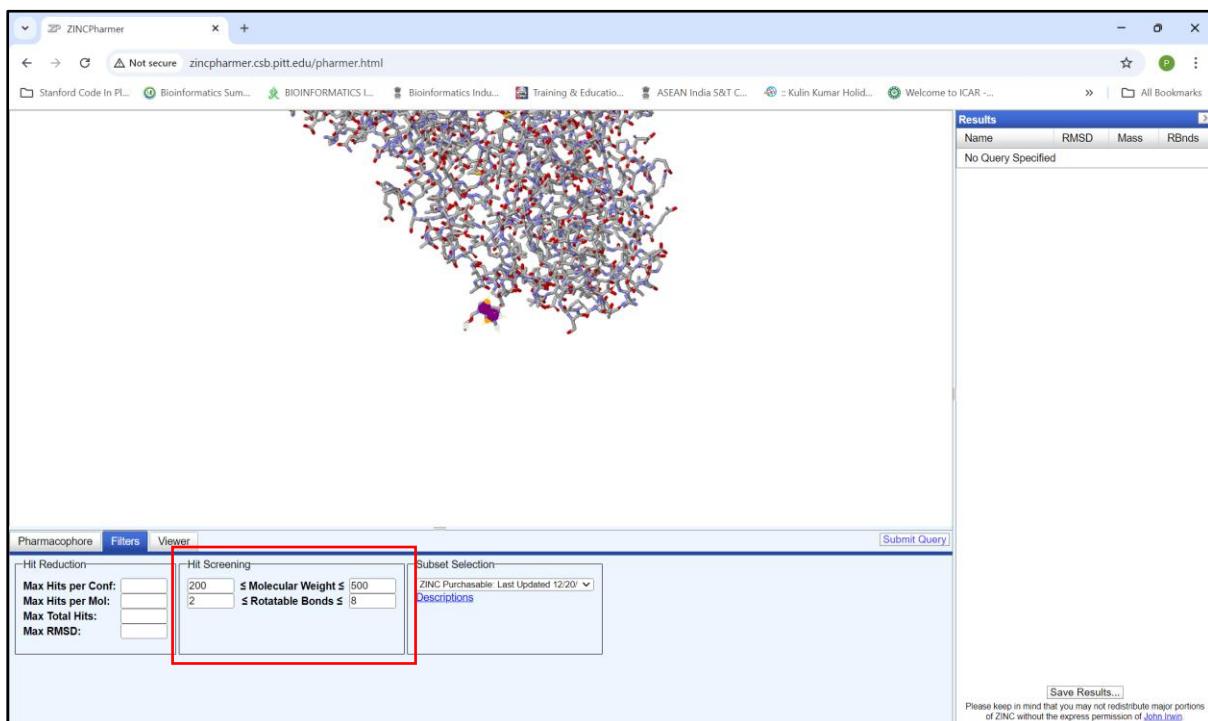


Fig 25: Setting filters to get the desirable pharmacophores in ‘Filter’ tab

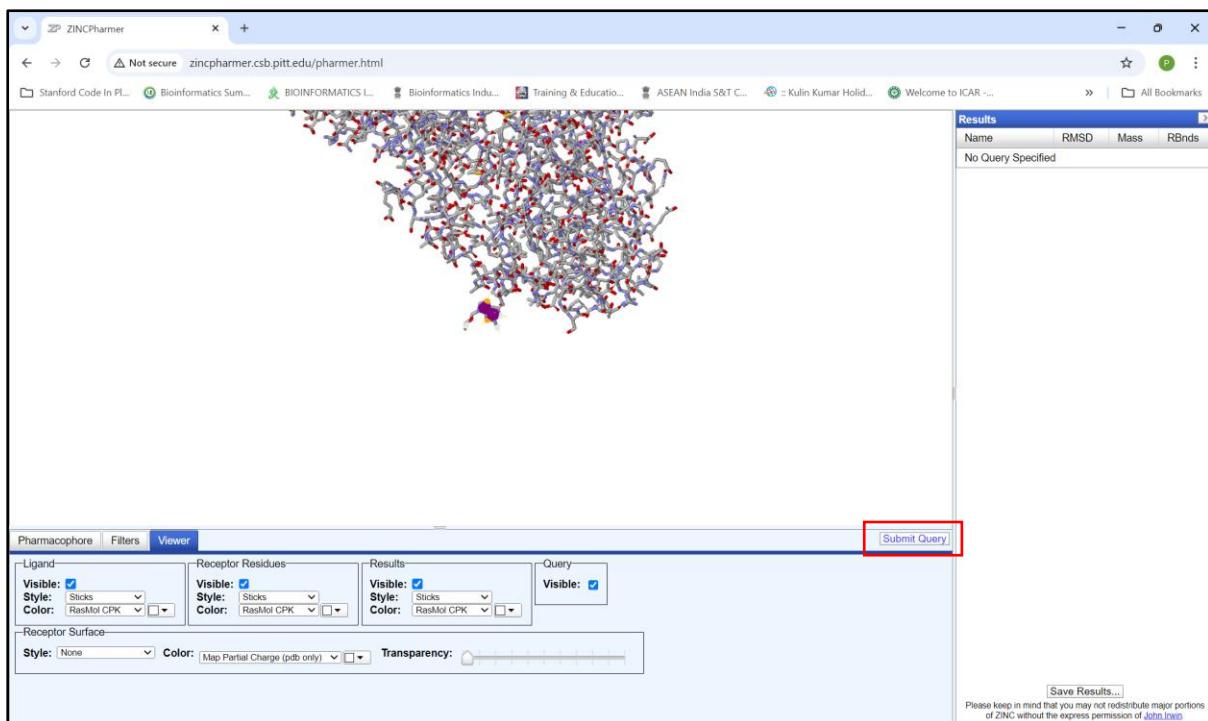


Fig 26: Customizing the molecules in the visualization panel using ‘Viewer’ tab options

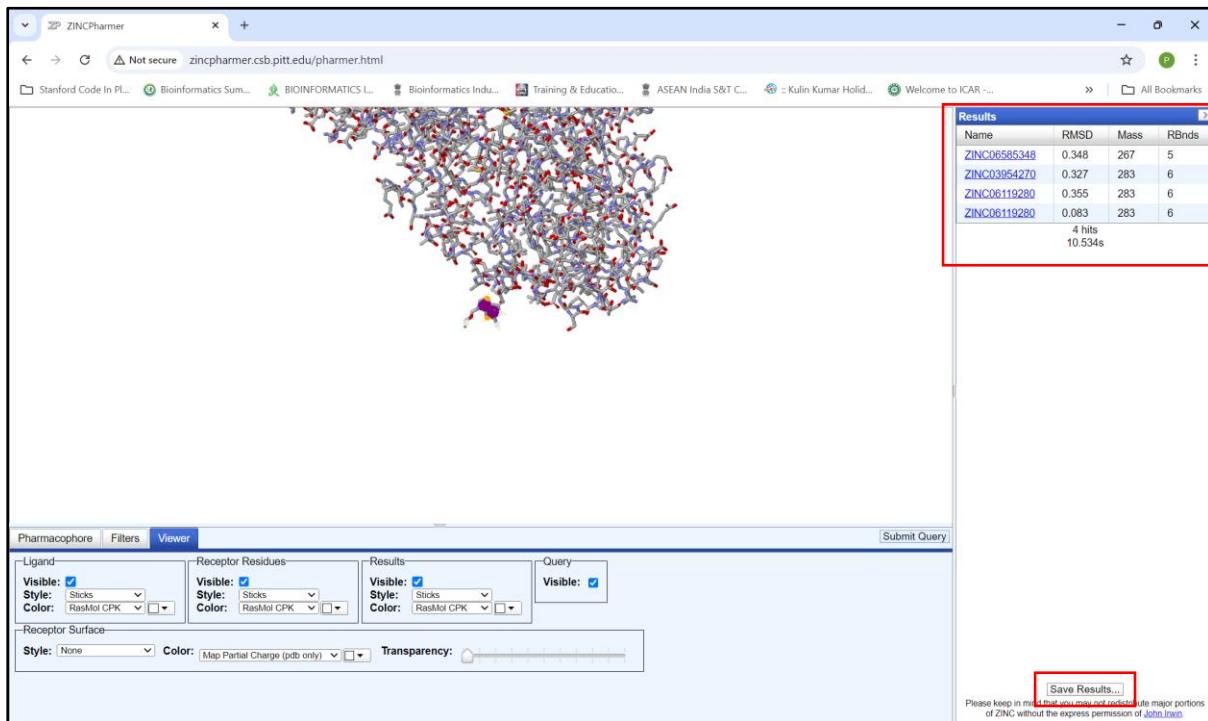


Fig 27: Results of the pharmacophore mapping and ‘Save Results’ option to download all pharmacophore results

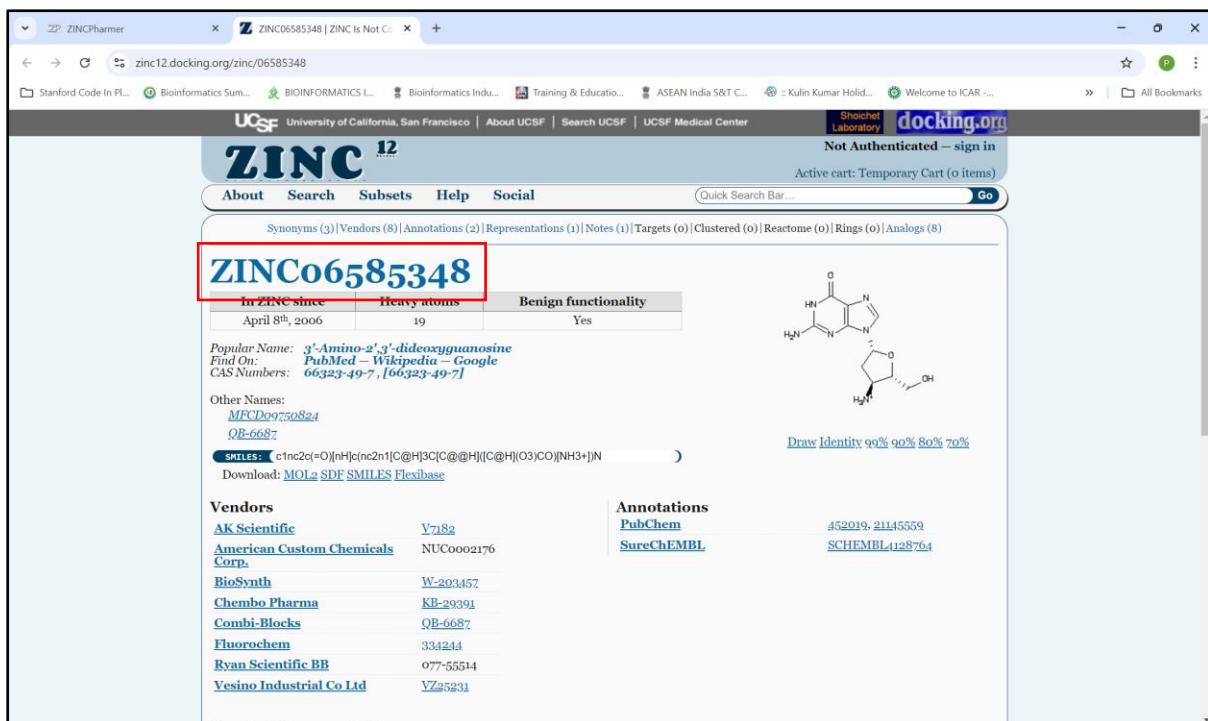


Fig 28: Chemical entity data of the first ZINC ID from the results

RESULTS:

The ZINCPPhamer software was used to generate and analyze pharmacophore map of ‘Acyclovir’ (PubChem ID: 135398513). It found 4 hits for the ‘Acyclovir’ in 10.534 seconds. The results were displayed as table with columns containing, Name (ZINC ID), RMSD Value, Mass and Rotatable Bonds of the chemical entities.

The ZINC IDs shows the data related to the specific chemical entity, such as, Entry Date, Popular Name, links to PubMed, PubChem, Wikipedia and Google, Chemical Structure, Vendors, Physical Data, Analogs, etc.

CONCLUSION:

The pharmacophore map for the query of ‘Acyclovir’ (PubChem ID: 135398513) was generated and analyzed using ZINCPPhamer Software. 4 hits were found for the query of ‘Acyclovir’ (PubChem ID: 135398513) in 10.34 seconds, along with its Name (ZINC ID), RMSD Value, Mass and Rotatable Bonds.

REFERENCES:

- Koes, D. R., & Camacho, C. J. (2012). ZINCPPhamer: pharmacophore search of the ZINC database. Nucleic Acids Research, 40(W1), W409–W414. <https://doi.org/10.1093/nar/gks37>
- Gnex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis, 18(15), 2714–2723. <https://doi.org/10.1002/elps.115018150>

3. Acyclovir: Uses, Interactions, Mechanism of Action | DrugBank Online. (n.d.).
DrugBank. <https://go.drugbank.com/drugs/DB00787>
 4. PubChem. (n.d.). Acyclovir. PubChem.
<https://pubchem.ncbi.nlm.nih.gov/compound/Acyclovir>
-