

NAME : Ch.Sri Chaitanya
 Reg.No. 192122057 Dept. Biotechnology
 SUBJECT : Protein Engineering

INDEX

SL. NO.	DATE	TITLE	PAGE NO.	FACULTY SIGN
01.	27/4/23	Qualitative -Analysis of given protein sample.	1-6	✓
02.	28/4/23	Quantitative Analysis of protein	7-8	✓
03.	05/05/23	Swiss model	9-10	✓
04.	08/05/2023	Homology modelling - predicting the protein structure	11-12	✓
05.	10/05/2023	Determination of PI of protein Using expasy protparam	13-14	✓
06.	16/05/23	Exploring genescript using Amino acid sequence .	15-18	✓
07.	15/05/23	Exploring genescript using nucleotide Sequence .	15-16	✓
08.	17/05/23	Swiss DOCK .	19-20	✓
09.	22/05/23	Thin layer chromatography	21-23	✓
10.	18/05/23	Validation of protein structure using Ramachandran Plot (Swiss 6.0)	24-25	✓

Compiled
 Date 15/6/23

15/06/23
 External

27/4/23

01. Qualitative Analysis of the given protein sample

Page No. 1

Ninhydrin Test

- i) Aim: To show the qualitative analysis of the given protein sample.

Principle: Amino acids contain a free amino group and a free carboxylic group that react together with ninhydrin to produce a coloured product. When an amino group is attached to the first (or) alpha carbon on the amino acid carbon chain, the amino group's nitrogen atom is part of a blue, product purple. Amino acids that have secondary amino group react with ninhydrin and the condensation product is yellow.

Experiment: Take 1-2 ml of the sample in a test tube and add 2 drops of ninhydrin's reagent.

Observation/Result:

protein is positive

protein is present in the given sample.

27/4/23

ii) Biuret test:

Aim: To identify the given sample qualitatively

Principle: The biuret test for proteins positively identifies proteins in solution with a deep violet colour. Biuret, $\text{H}_2\text{NCONHCCONH}_2$, reacts with copper (II) ions in a basic solution to form a deep violet complex. The peptide linkages in protein resemble those in biuret and also form deep violet complexes with basic copper (II).

Experiment:

Take a small quantity of the dispersion of the sample in a test tube and add 1 ml of NaOH solution into it. Now add 4-5 drops of 1% ²⁰⁰⁵ ^{INCU}SO₄ solution and warm the mixture for about 5 minutes.

Result:

- * No colour change is observed.
- * It indicates that no protein is present in the sample.

✓
27/4/28

Xanthoproteic test:

Aim: To identify the given sample qualitatively

Principle: Some amino acids contain aromatic groups that are derivatives of benzene. These aromatic groups can undergo reactions that are characteristic of benzene and benzene derivatives. The nitration of a benzene ring with nitric acid. The amino acids tyrosine and tryptophan contain activated benzene rings and readily undergo nitration. The amino acid phenylalanine also contains a benzene ring, but the ring is not activated and, therefore does not readily undergo nitration. This nitration reaction, when used to identify the presence of an activated benzene ring, is commonly known as the Xanthoproteic test, because the product is yellow.

Experiment:

Take about 2ml of the sample in a test tube and add few drops of conc. HNO_3 into it and heat the test tube.

Result:

No color change was observed. It indicates the absence of protein in the sample.

✓
27/11/22

iv) Millon's Test:

Aim: To identify the given sample qualitatively

Principle: Millon's Test is a test specific for tyrosine, the only amino acid containing a phenol group, a hydroxyl group attached to a benzene ring. In millon's test, the phenol group of tyrosine is first nitrated by nitric acid in the test solution. Then the nitrated tyrosine complexes mercury (I) and mercury (II) ions in the solution to form a red precipitate or a red solution, both positive results. Protein that contain tyrosine will, therefore yield as a positive result. However, some proteins containing tyrosine initially forms a white precipitate that turns red when heated, while others form a red solution immediately. Both results are considered positive.

Experiment:

* take 1 ml of the sample in a test tube and add 2 drops of millon's reagent.

Inference / Result:

protein is present in the given sample

v) Hopkins - cole Test:

Aim: To identify the given sample qualitatively.

Principle: The Hopkins - cole test is specific for tryptophan, the only amino acid containing an indole group. The indole ring reacts with glyoxylic acid in the presence of a strong acid to form a violet cyclic product. The Hopkins' cole reagent only reacts with proteins containing tryptophan. The protein solution is hydrolyzed by the concentrated sulphuric acid at the solution interface. Once the tryptophan is free, it reacts with the glyoxylic acid to form the violet product.

Experiment: In a test tube, 2ml of light exposed glacial acetic acid and 2ml of the sample liquid are taken. To this, concentrated H_2SO_4 is added along the sides of the test tube held at a slanting position. Two distinct layers of liquid are to be formed without mixing. The test tube should be observed for the formation of a colored ring at the interface of two layers.

Interference / Result:

Protein is present in the given sample

✓
Q14/2
21/4/22

(vi)

Nitroprusside Test:

Aim: to show the given sample qualitatively.

Principle: The nitroprusside test is specific for cysteine, the only amino acid containing a sulphhydryl group ($-SH$). The group reacts with nitroprusside in alkaline solution to yield a red complex.

Experiment:

2ml of the amino acid solution or the sample is added to a test tube. To the test tube, 0.5ml of the freshly prepared sodium nitroprusside is added and mixed thoroughly. Now, 0.5ml of the concentrated sodium hydroxide is added. If a positive result is not observed immediately, the test tube can be heated in a water bath for 2 minutes. The test tube is observed for the color change or formation of coloured complex.

Result/Inference:

- ★ colour changed from pale yellow to brown in colour.
- ★ the colour change indicates presence of protein in given sample.

Q1
27/4/23

02. Quantitative analysis of protein

Page No. 7

Aim: To estimate the amount of protein present in given sample.

Principle: The Bradford protein assay is used to measure the concentration of total protein in a sample. The principle of this assay is that the binding of protein molecules to coomassie dye under acidic conditions results in a colour change from brown to blue.

Procedure:

- 1) Take 5 test tubes
- 2) Name that as Blank (B), standards (s_1, s_2, s_3) Test (T) for which the amount need to quantify respectively.
- 3) Take the protein solution as shown in table
- 4) Mix all test tube and allowed to stand for 15-30 mins before the absorbance is read at 540nm.
- 5) Take absorbance . Draw the graph by taking concentration at x-axis and absorbance at Y-axis
- 6) Draw the line using the standard value.
- 7) The test absorbance were made to cut in the Standard to know the unknown concentration of the test.
- 8) The major advantages of this technique is that there is no interference from materials that adsorb at lower wavelength, and the technique is less sensitive is that there is no interference from materials that adsorb at lower wavelengths, and the technique is less sensitive to protein type because it utilizes adsorption involving peptide bonds that are common to all proteins.

Rather than specific side groups. However, it has a relatively low sensitivity compared to other UV-visible methods.

Observation:

	B	S_1 ml	S_2 ml	S_3 ml	T
Vol in ml	0	1	2	3	1
ml	3	2	1	0	2

0.07 0.093 0.025 0.1003 0.115

0.115 OD corresponds to 1.4 ml of vol

0.1g corresponds to 50 ml

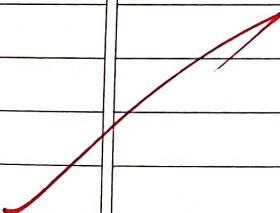
$$50 \text{ ml} \rightarrow 0.1 \text{ g}$$

$$1.4 \text{ ml} \rightarrow 0.1 / 50 \times 1.4 = 0.002$$

$$2.2 \text{ ml} \rightarrow \left[\frac{0.1}{50} \times 2.2 \right]$$

$$\underline{x \text{ axis}} = 1 \text{ unit} = 1 \text{ ml}$$

$$\underline{y \text{ axis}} = 0.03 \text{ nm} = 1 \text{ unit}.$$



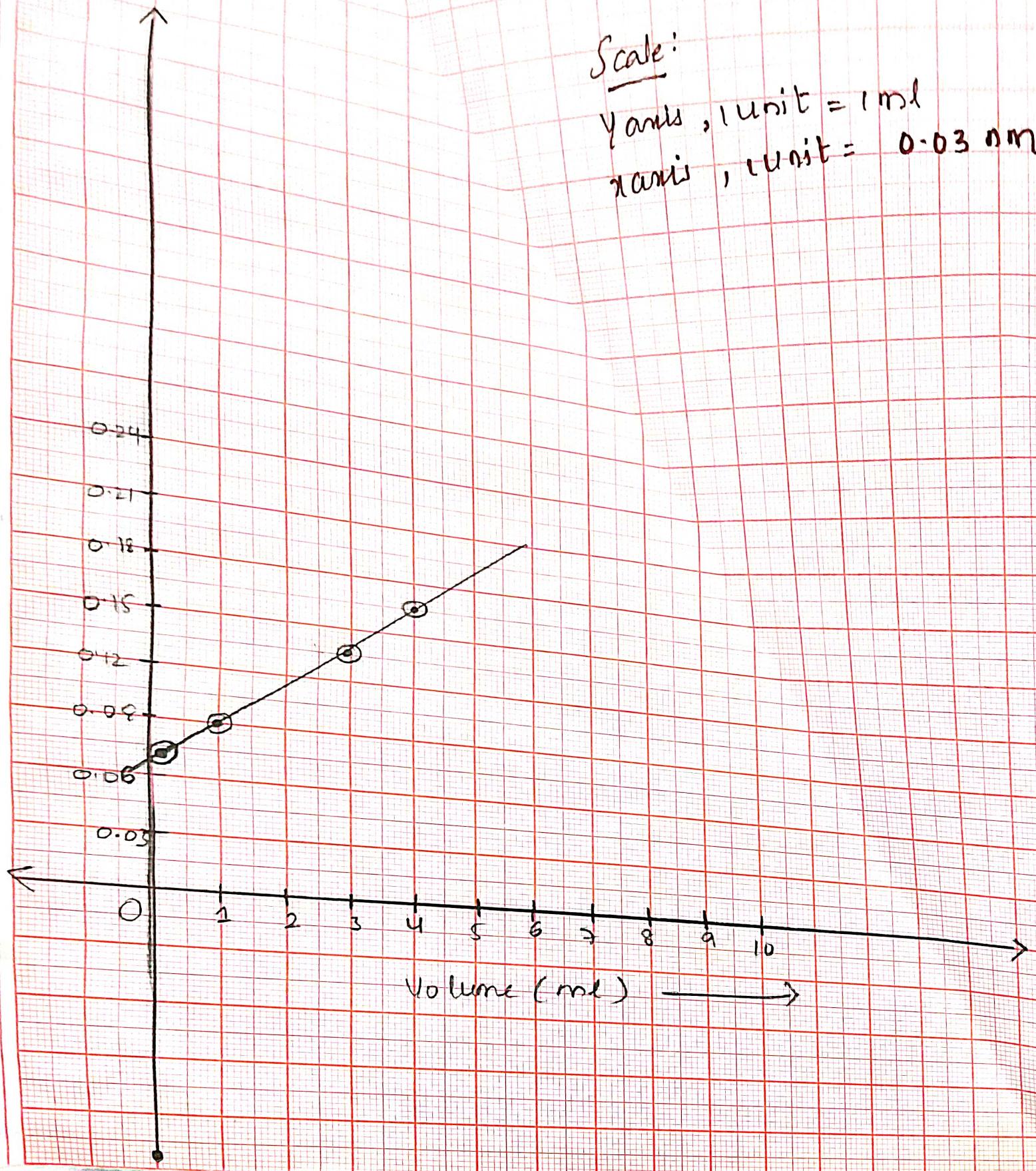
Result! Thus, the protein sample quantity is analysed.

26/12/23

Scale:

Y axis, 1 unit = 1 ml

X axis, 1 unit = 0.03 nm



31/5/28 03. swiss model

Page No. 9

Aim: To model the 3D structure of a protein that doesn't have any pre-existing structure in PDB

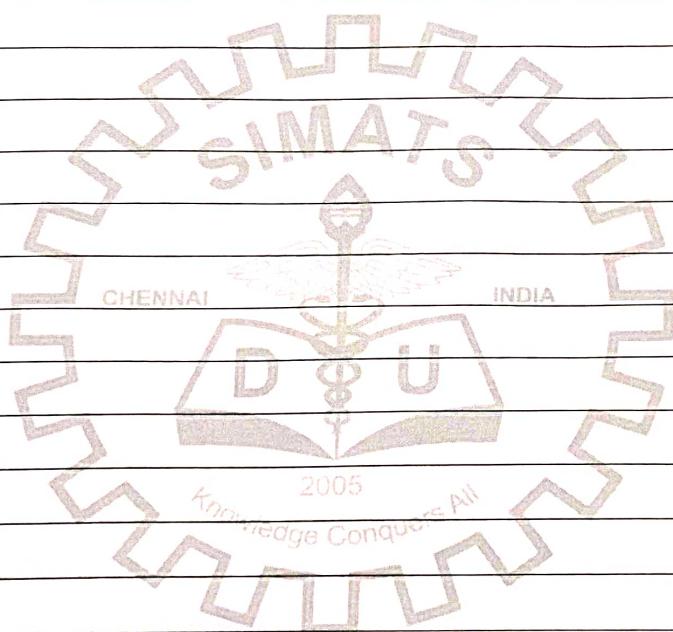
principle:

Swiss model is a automated protein structure homology-modelling server, accessible via the Enasy web server, or from the program Deep View. It is a structural bioinformatics web server dedicated to homology modelling of 3D protein structures.

Homology modelling aims to build three-dimensional protein structure models using experimentally determined structures of related family members as templates.

procedure:

- 1) Go to www.uniprot.org and search for the protein of interest.
- 2) From the results displayed, find the one that doesn't have any existing structure in PDB.
- 3) Copy in FASTA Sequence into a text file.
- 4) Go to <https://swissmodel.enasy.org/> and select "start modelling".
- 5) Copy and paste the FASTA sequence from the text file.
- 6) Click on "search for templates".
- 7) Choose an appropriate template (most probably the first one from the list) and click on "Build models".
- 8) The model is built and the file can be downloaded as a Zip file.



Result:

The 3D structure of a protein that doesn't have any pre-existing structure in PDB has been modelled successfully using Swiss model.

Q3/13
83/13

Aim: To predict the structure of a protein using homology method if approach in modeles.

Tool used: Modeller, chimera

principle: Modeller is a computer method used for homology modelling to produce models of protein tertiary structures and quaternary structures.

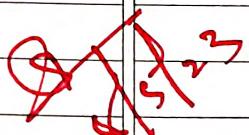
Template selection, sequence alignment, Model Building, loop refinement and validation.

Procedure:

- 1) Go to "UNIPROT" database.
- 2) Search the given "Uniprot Id" and download the Sequence of protein in FASTA format.
- 3) Copy the protein sequence and paste it in a text file named as "Target.ali".
- 4) Now, copy the protein sequence and paste in "BLAST 0" tool in NCBI for template selection.
- 5) In standard database option, choose "protein Data Bank" (Pdb) and click "BLAST".
- 6) List of Sequences producing significant alignments are shown with query coverage, percentage identity, accession number and other details.
- 7) Based on properties listed, choose the template:
 - i) Query coverage should be $\geq 80\%$.
 - ii) Percentage identity should be $\geq 30\%$.
 - iii) Sequence should be from similar family of organisms.
 - iv) Structure of template should be determined using X-ray crystallography.
 - v) Resolution of structure should be high $\leq 2\text{ \AA}$

- 8) Based on these, a template should be selected. Copy its accession number from BLAST results.
- 9) Paste the accession number in PDB search and download the pdb file.
- 10) Now, open the python script for sequence alignment and rewrite the code with your pdb file name in "align_20" file.
- 11) Run the module. The aligned sequence results are generated in the same folder.
- 12) Now, open the python script for building the model named as "Model_ligand".
- 13) Rewrite the code with your file names and alignment result files. Run the module.
- 14) The target protein models are generated in the same folder.
- 15) Choose the model with least mol pdf value which is modular objective function.
- 16) View the model using chimera.
- 17) Save the structure as Pdb file.
- 18) Validate the structure in Ramachandran plot of SAV 6.0 server by uploading the Pdb file.
- 19) When the errors is 0 and validation is $\geq 98\%$ The structure is valid.

Result: Hence the structure of R₂O₂ protein of Aedes aegypti is developed with 3 RBD's modelled and viewed by chimera.



10/15/25

05. Determination of PI of protein using expasy protoparam

Page No. 13

Aim: To determine the PI of protein using Expasy protoparam tool.

Tool used: Expasy - protoparam

Principle: protoparam is a tool which allows the compilation of various physical and chemical parameters for a given protein stored in swiss - prot (or) TrEMBL or for a user entered protein sequences.

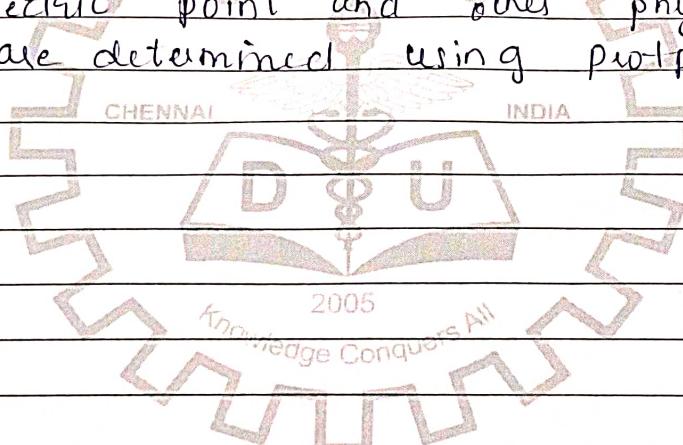
Accordingly, proteins are positively charged at a pH below their PI and negatively charged at a pH above their PI.

Procedure:

- 1) Search for Expasy protoparam in the browser.
- 2) Go to NCBI database and search for your protein of interest.
- 3) Retrieve the protein sequence.
- 4) Paste the sequence in the box given in protoparam tool.
- 5) Click on "Compute parameters" option.
- 6) Various parameters about the protein like number of amino acids, molecular weight, theoretical PI, amino acid composition, positive and negative residues count, atomic composition, aliphatic index, etc are given.
- 7) Hence, the physiochemical properties of the protein are computed.

Result:

- * Hence, the theoretical pI of the protein selected (melanin) is found to be 6.15. Below this pI , the protein is positively charged and above this, the protein negatively charged.
- * ~~the isoelectric point and other physiochemical properties are determined using protparam.~~



QX/10/12

Aim: To explore genescript using nucleotide sequence.

principle:

→ protein Isoelectric point :-

→ Calculates the theoretical pI for the protein sequence you enter. Use protein Isoelectric point when you want to know approximately where on a 2-D gel a particular protein will be found.

→ protein Molecular weight:

→ Accepts one or more protein sequences and calculates molecular weight. You can append copies of commonly used epitopes and fusion protein using the supplied dict.

→ protein Stats:

→ Returns the number of occurrences of each residue in the sequence you enter. Percentage totals are also given for each residue, and for certain groups of residues, allowing you to quickly compare the results obtained for different sequences.

→ Reverse Translate:

→ Accepts a protein sequence as input and uses a codon usage table to generate a DNA sequence representing the most likely non-degenerate coding sequence.

Procedure:

- 1) Go to NCBI and select 'protein' from the drop down menu.
- 2) Search for 2 proteins of interest.
- 3) Open their FASTA Sequences and copy them into a notepad.
- 4) Open genescript and select 'protein Isoelectric point'.
- 5) Copy the FASTA sequences from the notepad and paste it in the Genescript site.
- 6) Click on submit for the results to be displayed.
- 7) Repeat the steps 5 and 6 by selecting "protein Molecular weight" in step 4.
- 8) Repeat the steps 5 and 6 by selecting 'protein stats' in step 4.
- 9) Repeat the steps 5 and 6 by selecting 'reverse translate' in step 4.

Result:

PI, molecular weight of protein, protein stats and reverse translate of amino acids are determined using genescript.

~~Q17~~
16/5/13

1st 5/23

07. Exploring genescript using amino acid sequence (Page No. 10)

-Aim: To explore genescript using amino acid sequence.

Principle:

protein Isoelectric point → calculates the theoretical PI for the protein sequences you enter. The protein Isoelectric point when you want to know approximately where on a 2-D gel a particular protein will be found.

protein molecular weight → Accepts one (or) more protein sequences and calculates molecular weight. Use protein molecular weight when you wish to predict the location of a protein of interest on a gel in relation to set of protein standards.

protein stats → Returns the no. of occurrences of each residue in the sequence you enter. Percentage totals are also given for each residue, and for certain group of residues, allowing you to quickly compare the results obtained for different sequences.

Reverse Translate → Accepts a protein sequence as input and uses a codon usage table to generate a DNA sequence representing the most likely non-degenerate coding sequence. A consensus sequence derived from all the possible codons for each amino acid is also returned. Use reverse translate when designing PCR primers to anneal to an unsequenced coding sequence from a related species.

Tools Required: Genescript

procedure:

- 1) Go to NCBI and select 'Protein' from the drop-down menu.
- 2) Search for 2 proteins of interest
- 3) Open their FASTA sequences and copy them into a notepad.
- 4) Open genescipt and selected "protein Isoelectri C point".
- 5) Copy the FASTA sequences from the notepad and paste it in the Genescipt site.
- 6) Click on submit for result to be displayed
- 7) Repeat the steps 5 and 6 by selecting 'protein molecular weight' in step 4.
- 8) Repeat the steps 5 and 6 by selecting 'Protein stats' in step 4.
- 9) Repeat the steps 5 & 6 by selecting 'Reverse translate' in step 4.

Result: The DNA stat, DNA pattern find, group DNA and range extract DNA parameter of gene / protein of interest is determined using genescipt.

Ques
Ans

Aim: To predict the ligand binding site for the protein or protein subchains by using the software Swiss Dock.

principle: Swiss Dock and S3DB are developed by Aurelio Guasidicis, Vincent Zoete and Olivier Michelin. From the molecular modelling group of the Swiss Institute of Bioinformatics in Lausanne, Switzerland. SWISS DOCK, a webservice to predict the molecular interactions that may occur between a target protein and a small molecule. S3DB, a database of manually curated target and ligand structures, inspired by the ligand-protein database.

- 1) Many binding modes are generated either in a base or in the vicinity of all target cavities.
- 2) Simultaneously, their "CHARMM" energies are estimated on a grid.
- 3) The binding modes with the most favourable energies are evaluated with FACTS, and clustered.
- 4) The most favourable clusters can be visualized online and downloaded on your computer.

procedure:

- 1) Open PDB and search for protein of interest.
- 2) Downloaded the PDB file of a protein in PDB format. Save the file
- 3) Open pubchem software, downloaded the ligand molecule which fits with the protein of interest in the SDF format
- 4) Go to open BABEL and select SDF format in the PDB column & paste the ligand file which was downloaded.

- 5) Select the PDB format in the output column and click convert button
- 6) Results will be displayed in the output column in PDB format.
- 7) Copy the PDB format and save it in notepad as filename .pdb.
- 8) open swiss dock click on the submit docking button.
- 9) In the target selection upload the file of protein which was downloaded from pdb.
- 10) In the ligand Selection column upload the ligand file which is in the pdf format.
- 11) Then start docking by submitting the mail Id's.

~~Result: Docking is done and finally the structure of uploaded ligand and protein in PDB format~~

✓✓✓✓✓

Aim: To study about separating a mixture of chemical substance.

Principle: Thin-layer chromatography is a chromatography technique used to separate non-volatile mixtures. Thin-layer chromatography is performed on a sheet of glass, plastic or aluminium foil, which is coated with a thin layer of adsorbent material, usually silica gel, aluminium oxide or cellulose. This layer of adsorbent is known as the stationary phase. After the sample has been applied on the plate, a solvent or solvent mixture is drawn up the plate via capillary action. Because different analytes ascend the TLC plate at different rates, separation is achieved. The mobile phase has different properties from the stationary phase.

To quantify the results, the distance travelled by the substance being considered is divided by the total distance travelled by the mobile phase. This ratio is called the retardation factor. In general, a substance whose structure resembles the stationary phase will have low R_f , while one that has a similar structure to the mobile phase will have high retardation factor. Retardation factors are characteristic, but will change depending on the exact condition of the mobile and stationary phase. Thin layer chromatography can be used to monitor the progress of a reaction, identify compounds present in a given mixture, and determine the purity of a substance. Specific examples of these applications include:-

Analysing ~~cosmetics~~ and fatty acids, detection of pesticides or insecticides in food and water, analysing the dye composition of fibres of bennies,

assaying the radiochemical purity of radiopharmaceuticals, or identification of medicinal plants and their constituents.

plate preparation:

TLC plates are usually commercially available, with standard particle size ranges to improve reproducibility. They are prepared by mixing the adsorbent, such as silica gel, with a small amount of inert binder like calcium sulphate and water. This mixture is spread as a thick slurry on an unreactive carrier sheet, usually glass, thick aluminium foil or plastic. The resultant plate is dried and activated by heating in an oven for thirty mins at 110°C. The thickness of the absorbent layer is typically around 0.1 - 0.25 mm for analytical purposes and around 0.5 - 2.0 mm for preparative TLC.

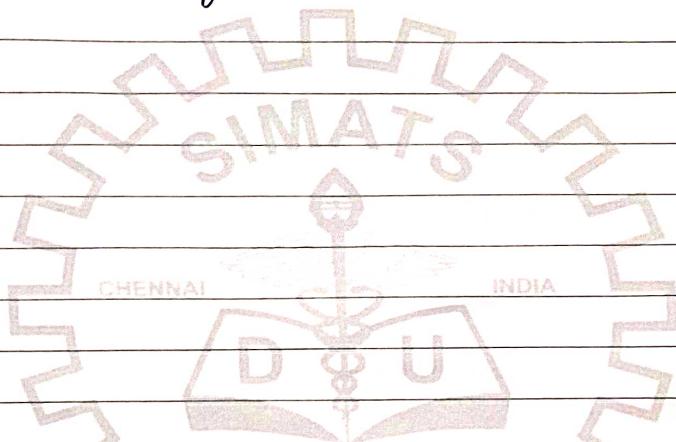
Technique: the process is similar to paper chromatography with the advantage of faster runs, better separations, and the choice of different stationary phases. Because of its simplicity and speed TLC is often used for monitoring chemical reactions and for the qualitative analysis of reaction products.

To run a thin layer chromatography plate, the following procedure is carried out:

~~Using a capillary, a small spot of solution containing the sample is applied to a plate, about 1.5 centimetre from the bottom edge. The solvent is allowed to completely evaporate off to prevent it from interfering with sample interactions with the mobile phase in the next step.~~

A small amount of an appropriate solvent is poured into a glass beaker or any other suitable transparent container to a depth of less than 1 centimetre. The container is closed with a cover glass or any other lid and is left for a few mins to let the solvent vapours ascend the filter paper and saturate the air in chamber.

Result: The analyte was identified using TLC.



✓ 22/23
22/23

Aim: To validate the protein structure built or retrieved using ramachandran plot in saves 6.0 servers.

Tool used: SAVES 6.0

Principle: The ϕ/ψ plot of amino acid residues in a peptide is called the ramachandran plot. It involves plotting ϕ -axis to predict the possible conformation of the peptide. The angle spectrum in each axis is from -180° to $+180^\circ$.

procedure:

- * Go to SAVES 6.0 server in the browser.
- * Click on 'choose file' option in the server page.
- * Choose the protein structure you need to validate which is saved in pdb format (.pdb).
- * Once the file is uploaded, click 'Run program'.
- * Click 'start' in ERRAT program to view the overall quality factor of the protein.
- * If the "ERRAT" overall quality factor is > 50 , the model is considered as high quality model.
- * Now click 'start' in PROCHECK program to view the evaluations of the structure which include errors, warnings and passes.

- * click on "results" and choose 'Ramachandran plot' to view the plot for the structure.
 - * The plot gives the details about favoured, allowed and disallowed regions for the residues for structure.
 - * The structure is valid if the no. of disallowed regions are zero.
 - * If the disallowed regions is not equal to zero, then the structure is invalid.
- Result:**
- * For the protein structure of melanin 1670 protein. The ERRAT score is 95.941 and the Ramachandran plot showed 'zero disallowed regions'.
 - * Hence, the structure is valid, with respect to the Ramachandran plot structure validation.

Ques 13