

□ **What is Pharmacoinformatics?**

- Discipline where technology intersects with any aspects of drug delivery, from the basic sciences to the clinical use of medications in individuals and populations.
- Pharmacy Informatics, a subject of pharmacoinformatics, typically refers to the interface of technology with the practice of pharmacy.

- Includes pharmacy technologies in the preparation, delivery and management of medication use within health care delivery systems.
- Includes all health care systems related to medications and clinical support. (Barcode and RF packaging, smart pump delivery of medications, pharmacy departmental application, IV solution, compounding system, robotics used in the preparation and dispensing of medications.)

- Includes analysis and research related to data, generated by the use of these systems.
- Includes the development of new technologies that improve the Quality and safety of therapeutic care related to medication use.
- The work in pharmacoinformatics can be broadly divide into two categories-
 - a) Scientific aspects
 - b) Service aspects.

- The scientific aspects deals with the drug discovery and development activities whereas the service oriented aspects are more patient centric.

- Pharmacoinformatics subject feeds on many emerging information technologies like
 - ❖ Neuroinformatics
 - ❖ Biosystem informatics
 - ❖ Metabolome informatics
 - ❖ Chemical reaction informatics
 - ❖ Toxicoinformatics
 - ❖ Genome informatics
 - ❖ Healthcare informatics etc.

- **Neuroinformatics :-**

- Work is focused on the integration of neuroscientific information from the level of the genome to the level of human behavior.
- A major goal is to produce digital capabilities for a web-based information management system in the form of databases and associated data management tools.

- Some of the databases developed in neuroinformatics are,
 - Surface Management System (SuMS)
 - The Brain Architecture Management System (BAMS) etc.

- **Biosystem informatics :-**

- An integrated approach of biosystem as a whole unit and information technology to study and understand the function of biological system.
- The biological system can be at any level- cellular, cell, organ, tissue, or organism.

- **Genome informatics:-**
- It is a field encompasses the various methods and algorithms for analyzing and extracting biologically relevant information from the rapidly growing biological and essential sequence databases.
- Genome informatics came into existence with the initiation of Human Genome Project (HGP)

- Toxicoinformatics:-

- It involves the use of information technology and computational science for the prediction of toxicity of chemical molecules in the living systems.
- Two basic approach used
 - a) Based on modeling Structure activity relationship (SAR).
 - b) Rule based methods.

- Chemical reaction informatics:-
 - Enable a chemist to explore synthetic pathways, quickly design, and record completely new experiments from scratch or by beginning with reactions found in the reaction databases.

- Metabolome informatics:-
 - In the field drug discovery metabolome informatics can contribute to target identification, mechanism of action, and pathway of drug toxicity.
 - The efforts in this field can be divide in two categories,
 - a) Drug metabolism informatics
 - b) Metabolism pathway informatics

- **Healthcare informatics:-**
 - The major components of healthcare informatics,
 - a) Electronic Health Record (EHR) Systems
 - b) Hospital Information Systems (HIS)
 - c) Decision Support Systems (DSS) etc.

- Pharmacoinformatics has been involved in the following areas of pharmaceutical practice :

- Research and development
- Patient profiling
- Decision support
- Medication Information Systems (MIS)
- Telepharmacy

- Research and development:-
 - Pharmaceutical firms spend a lot of time and money, researching and developing new drugs and in the process , a lot of data in addition to the data that is being produced by bioinformatics.

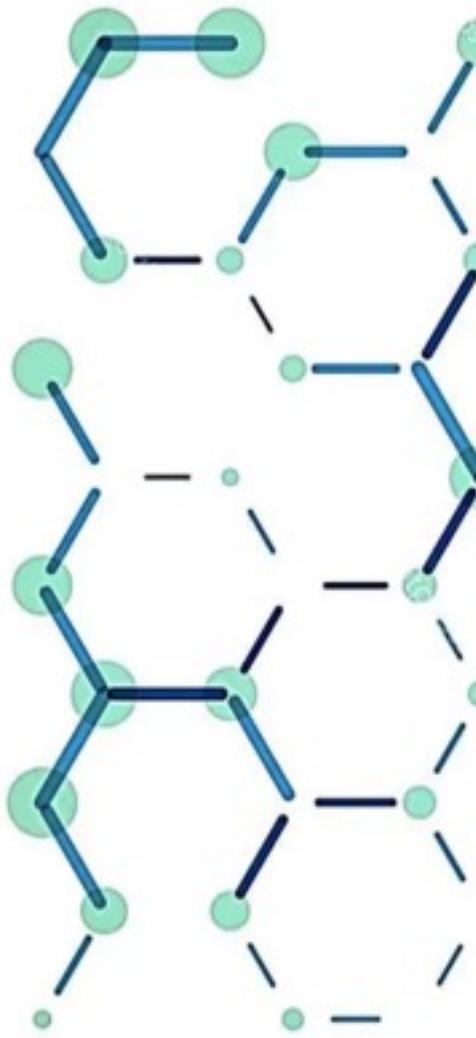
- Patient profiling:-
 - It is now easier to keep, track of medication that patients are taking presently, as well as medications they might have taken in the past, exact details such as past prescription and allergies, can be extracted from computer systems.

- **Decision Support:-**

- Contributed to the development of decision support systems in areas such as ,
 - a) Choosing potential medications for a condition
 - b) Offering alternatives based on past medical history and non pharmacological conditions such as cost.
 - c) Listing of indications and contraindications.

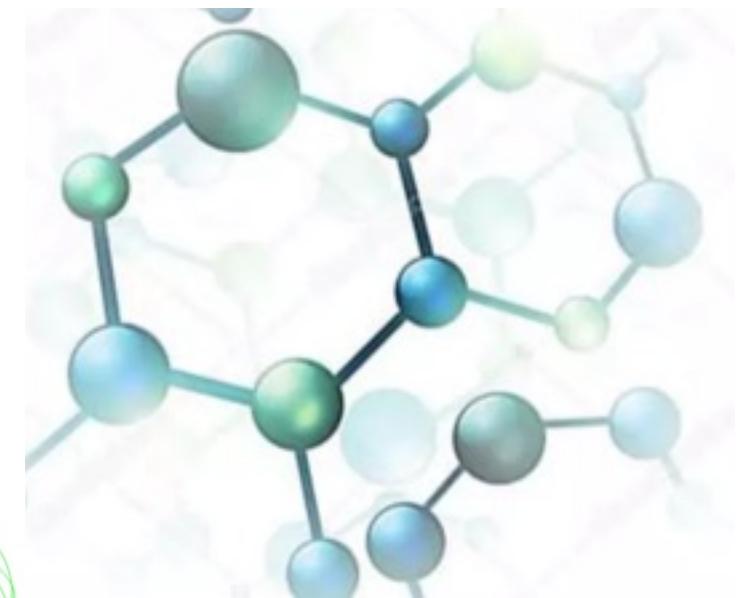
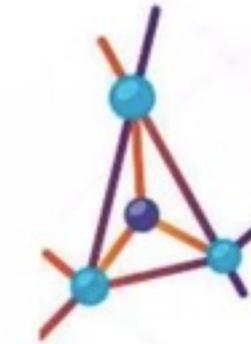
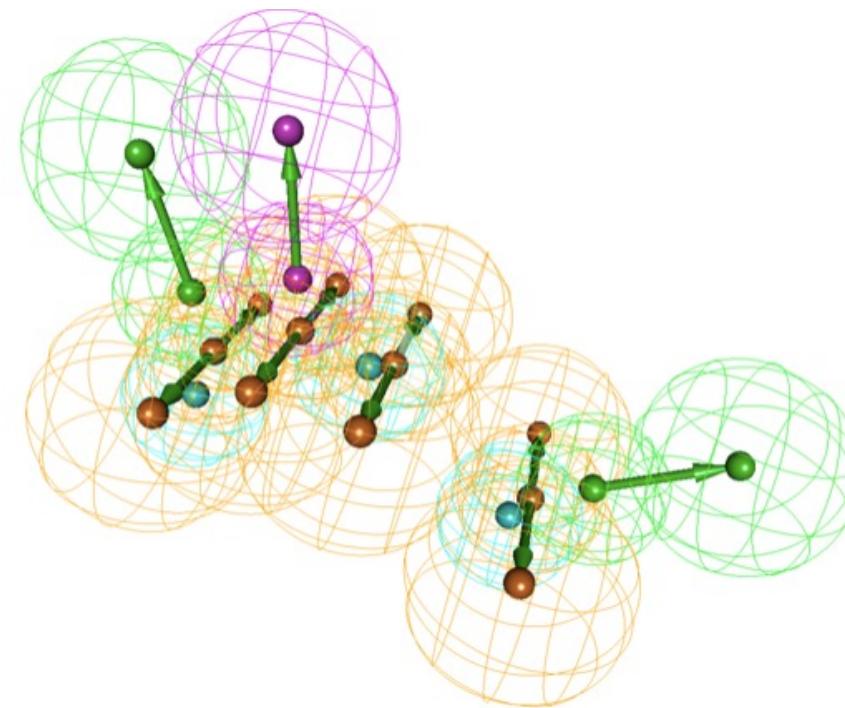
- Medication Information Systems (MIS) :-
 - Offers health information about medications.
 - Include the composition of drugs, their uses, risks and side effects and food and drug interactions.

- **Telepharmacy:-**
- This is the provision of pharmaceutical care through the use of computer systems, to the patients at remote locations by a pharmacist.



Pharmacophore Mapping

"Understanding its importance in Drug Discovery"

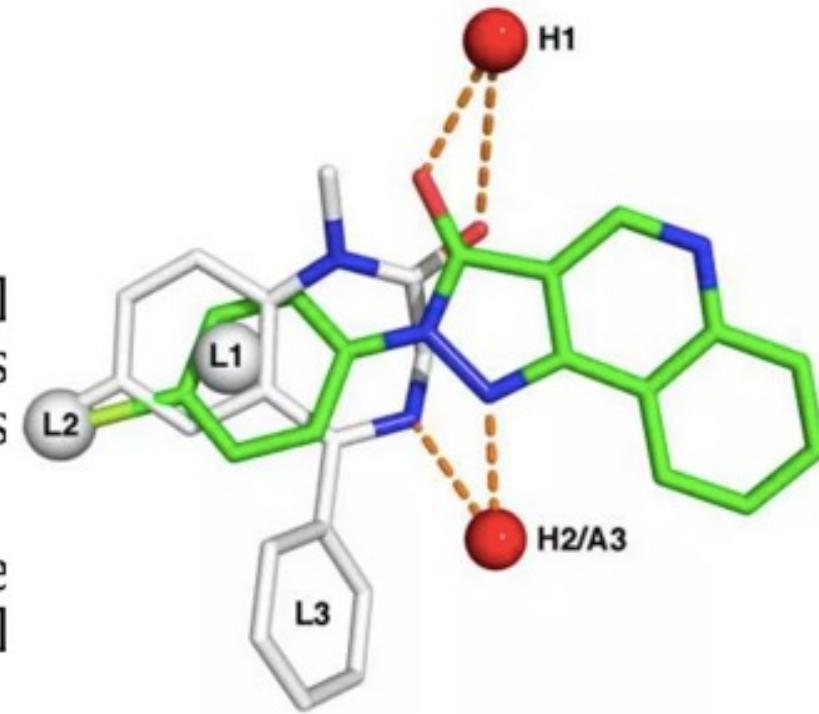


Introduction

- Definition of Pharmacophore

A pharmacophore is a three-dimensional arrangement of atoms in a molecule that is responsible for its biological activity, specifically its interaction with a target protein or receptor.

- A part of a molecular structure that is responsible for a particular biological or pharmacological interaction that it undergoes.
- First introduced in 1990 by Paul Herilich.
- A Pharmacophore is a representation of generalized molecular features including:
 - 3D (hydrophobic groups, charged/ionizable groups, hydrogen bond donors/acceptors)
 - 2D (substructures)
 - 1D (Physical or biological)



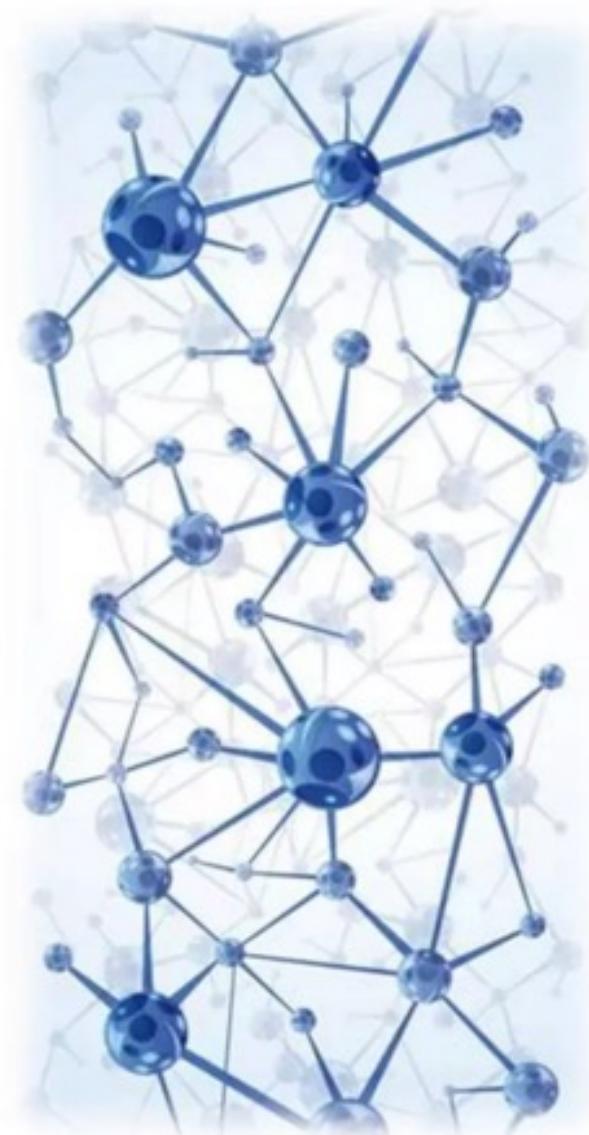
An example of a pharmacophore model of the benzodiazepine binding site on the GABA receptor. The red spheres labeled H1 and H2/A3 are, respectively, hydrogen bond donating and accepting sites in the receptor, while L1, L2, and L3 denote lipophilic binding sites.

Pharmacophore Mapping Vs Pharmacophore Modeling

Pharmacophore Mapping	Pharmacophore Modeling
Identifies and characterizes the key molecular features required for a drug molecule to bind to a target receptor and exert a therapeutic effect.	Involves the construction of a 3D model of the pharmacophore using computational methods.
Analyzes the structures of known ligands that bind to the target receptor to identify the common features that are essential for binding.	Predicts the binding affinity of new ligands to the target receptor based on the pharmacophore model.
Generates a pharmacophore map that can be used to guide the design of new drug molecules optimized for binding to the target receptor	Can be used to design new ligands with improved binding properties.
Typically based on experimental data	Can be based on a combination of experimental and computational data

Pharmacophore Mapping

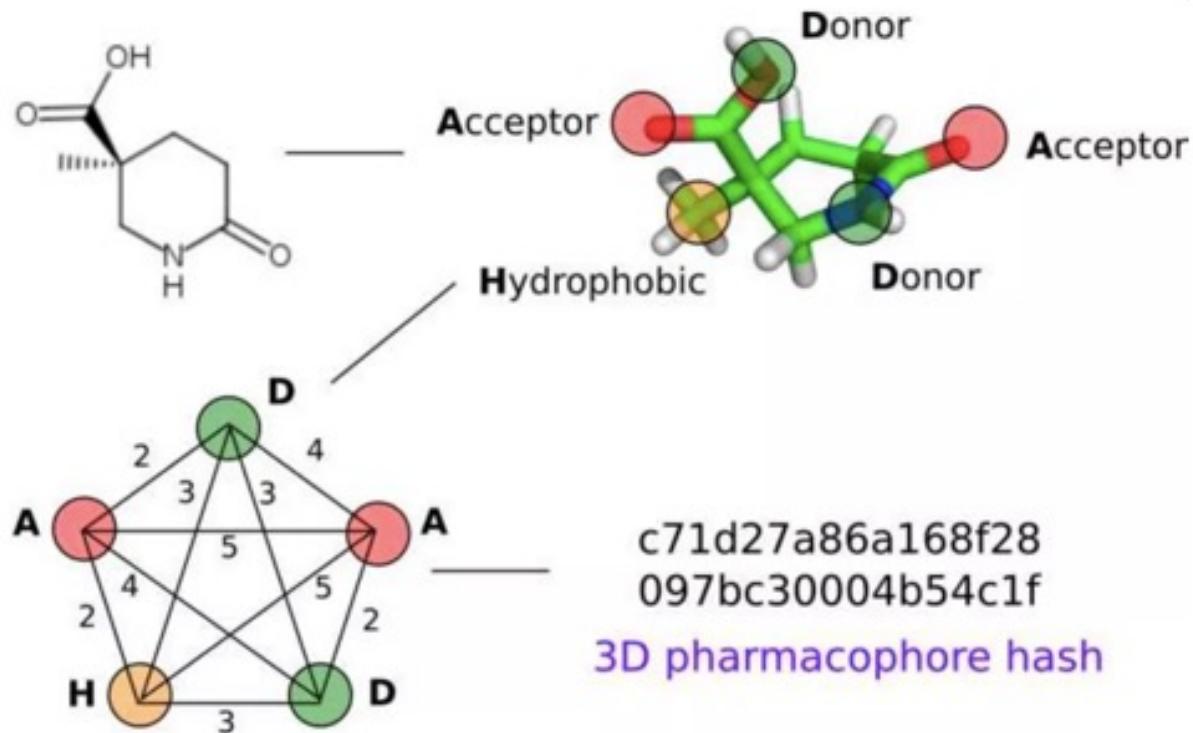
- It is a computational technique used in drug discovery.
- Its main purpose is to identify and optimize molecules that bind to a specific target.
- This technique is based on the identification of key molecular features, or pharmacophores, that are necessary for a molecule to interact with the target.
- Molecular interactions between ligands and receptors are represented by a set of pharmacophoric features, such as **hydrogen bond donors** and **acceptors**, **aromatic rings**, and **hydrophobic regions**.
- The pharmacophoric features are used to create a three-dimensional model of the binding site, which can be used to identify molecules that are likely to bind to the target.



Types of Pharmacophore Mapping

- **Ligand-based pharmacophore mapping:**

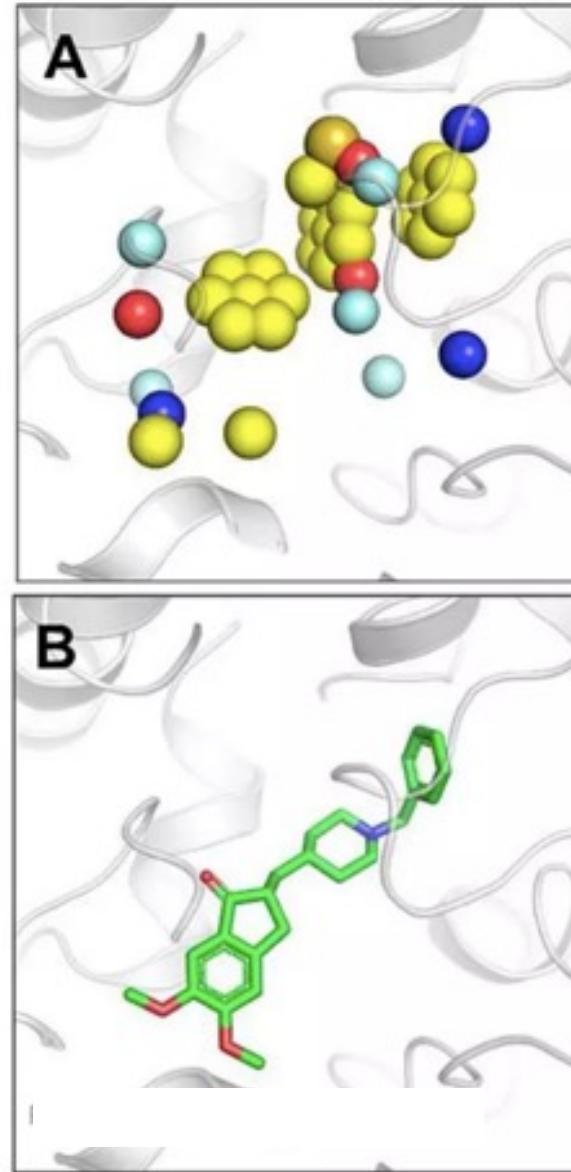
This approach uses the structural and physicochemical properties of a set of known ligands that bind to a target to identify common pharmacophoric features. The resulting pharmacophore model can be used to screen compound libraries for molecules with similar properties and potentially high binding affinity.



Types of Pharmacophore Mapping

- **Receptor-based pharmacophore mapping:**

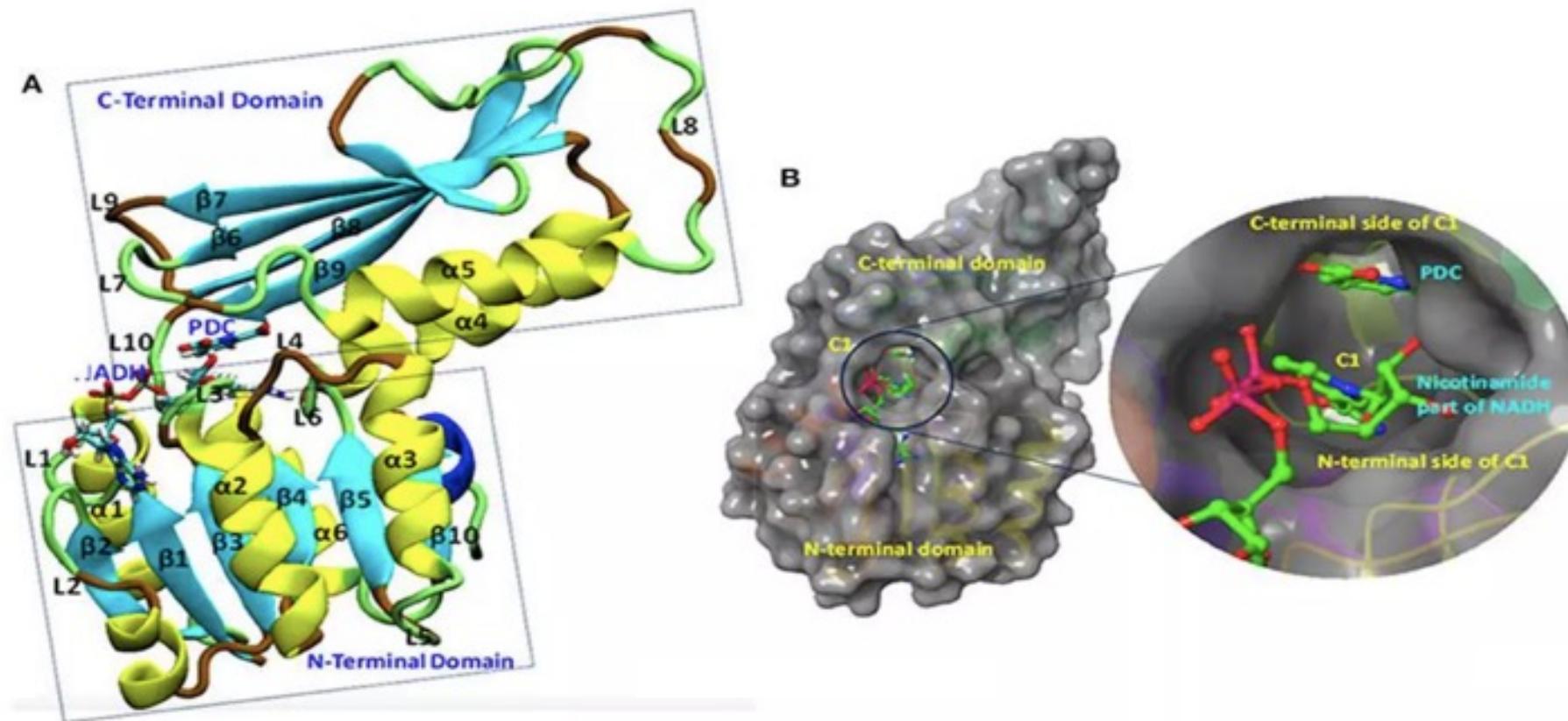
This approach involves the generation of a pharmacophore model based on the three-dimensional structure of a target receptor, which can be obtained from X-ray crystallography or other experimental methods. The model can be used to design or screen compounds that interact with the receptor.



- Blue sphere represents HBD pharmacophore.
- Red sphere represents HBA pharmacophore
- Single yellow sphere represents hydrophobic pharmacophore.
- Single yellow sphere that surrounded by six planar yellow spheres represents aromatic pharmacophore.
- Sulfur sphere represents positive, orange sphere represents negative, and palecyan sphere represents the root of h-bond.

Types of Pharmacophore Mapping

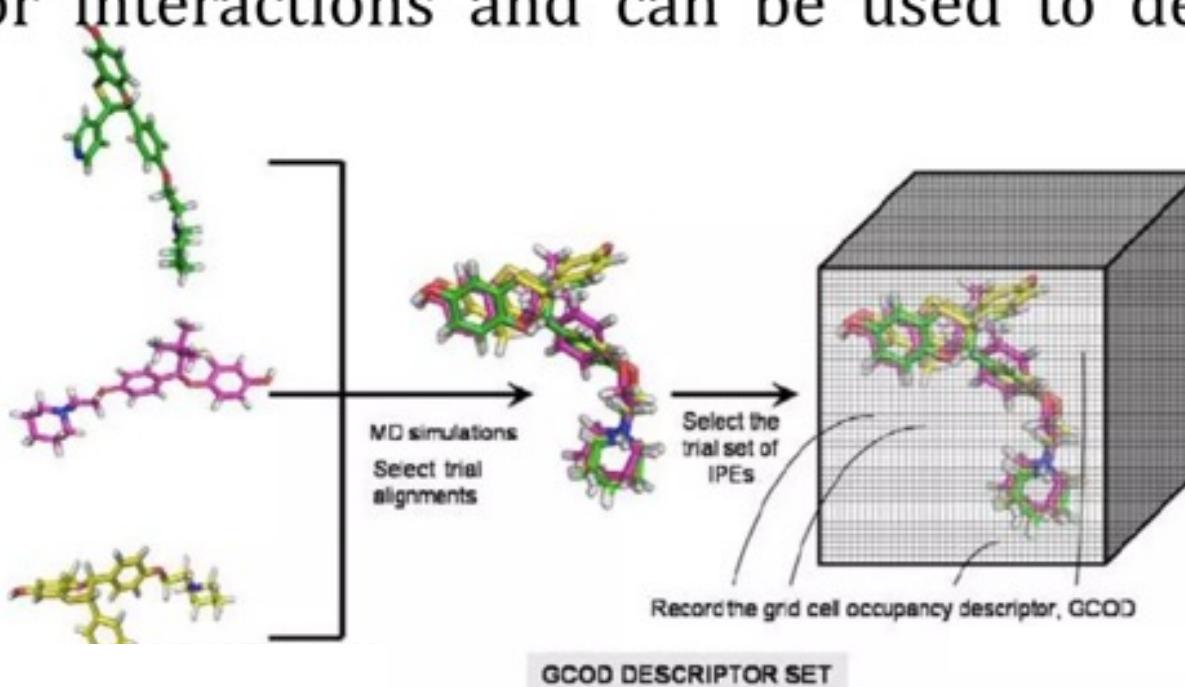
- **Hybrid pharmacophore mapping:** This approach combines ligand- and receptor-based pharmacophore models to take advantage of both types of information. It can be particularly useful when the three-dimensional structure of the target is not available or when the ligand binding site is poorly defined.



Types of Pharmacophore Mapping

- **4D-QSAR pharmacophore mapping:**

This approach involves the use of quantum mechanics and molecular dynamics simulations to generate a dynamic, time-dependent pharmacophore model that accounts for the flexibility and motion of both the ligand and the receptor. It can provide insights into the molecular mechanisms of ligand-receptor interactions and can be used to design compounds with specific properties.

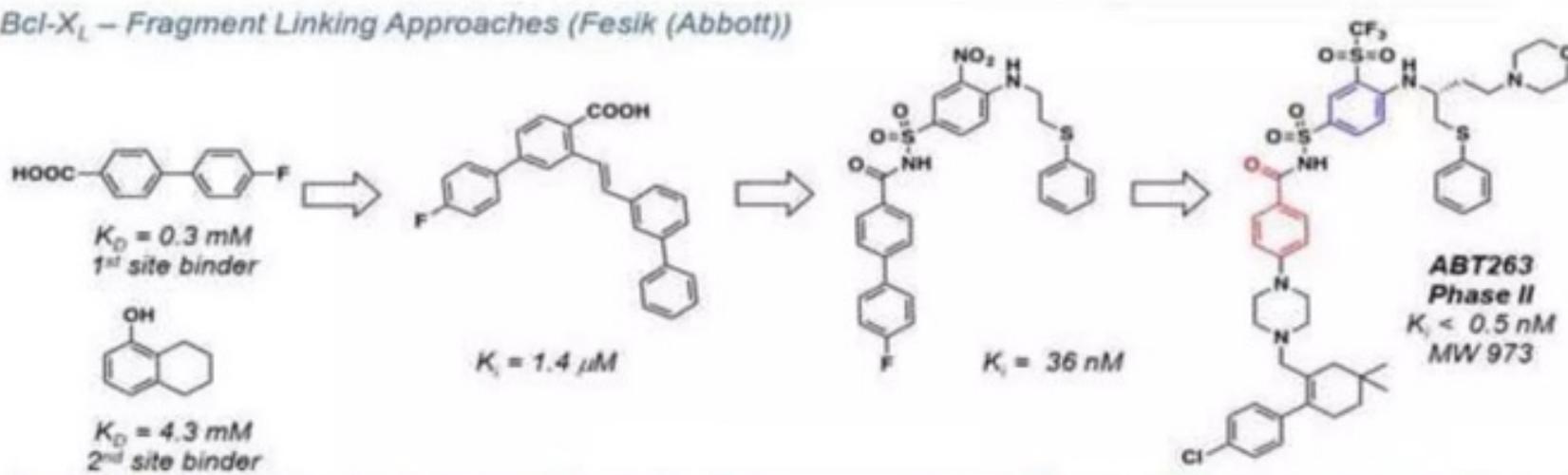


Types of Pharmacophore Mapping

- **Fragment-based pharmacophore mapping:**

This approach involves the identification of key functional groups or fragments that are important for ligand binding and optimization. The fragments can be assembled into larger compounds with the desired pharmacophoric features and properties. This approach can be particularly useful for designing drug candidates with improved potency, selectivity, and pharmacokinetic properties.

Bcl-X_L – Fragment Linking Approaches (Fesik (Abbott))



Methods and Tools

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

- The process of deriving pharmacophore is known as pharmacophore mapping.

It consist of three steps

- (1) Identifying common binding element that are responsible for the biological activity;
- (2) Generating potential conformations that active compound may adopt; and
- (3) Determining the 3D relationship between pharmacophore element in each conformation generated

CONFORMATIONAL SEARCH

- Conformation generally means **structural arrangement**
- Conformations are different three-dimensional structures of molecules that arise from :
 - Rotation about single bonds (torsion angles)
 - Different rings conformations
- The biological activity of molecules is strongly dependent on their conformation
- Done by exploring the energy surface of a molecule and determining the conformation with minimum energy
- Conformational analysis is needed to identify the ideal conformation of a molecule
- If the torsion angles are incremented in steps of 30° , this means that a molecule with 5 rotatable bonds will have $12^5 \approx 250K$ conformations

Applications

- **Lead identification:** Pharmacophore mapping can be used to identify lead compounds that possess the desired pharmacological properties.
- **Lead optimization:** Pharmacophore mapping can help in the optimization of lead compounds by identifying key structural features that are responsible for the desired activity.
- **Drug design:** Pharmacophore mapping can be used to design new drugs by identifying the key structural features required for the desired activity.
- **Virtual screening:** Pharmacophore mapping can be used to screen large databases of compounds to identify potential hits that match the desired pharmacophore.
- **Scaffold hopping:** Pharmacophore mapping can be used to identify new scaffolds that can be used to develop compounds with the desired activity.

Applications

- **Binding site identification:** Pharmacophore mapping can help in the identification of the binding site of a protein or receptor, which can aid in the design of ligands with the desired activity.
- **Structure-activity relationship (SAR) analysis:** Pharmacophore mapping can be used to analyze SAR data to identify key structural features that are responsible for the desired activity.
- **Fragment-based drug design:** Pharmacophore mapping can be used to identify fragments that can be used to build larger molecules with the desired activity.
- **Protein-ligand interaction analysis:** Pharmacophore mapping can help in the analysis of protein-ligand interactions to identify key structural features that are responsible for binding.
- **Toxicity prediction:** Pharmacophore mapping can be used to predict the toxicity of compounds by identifying structural features that are associated with toxicity.

Limitations and challenges

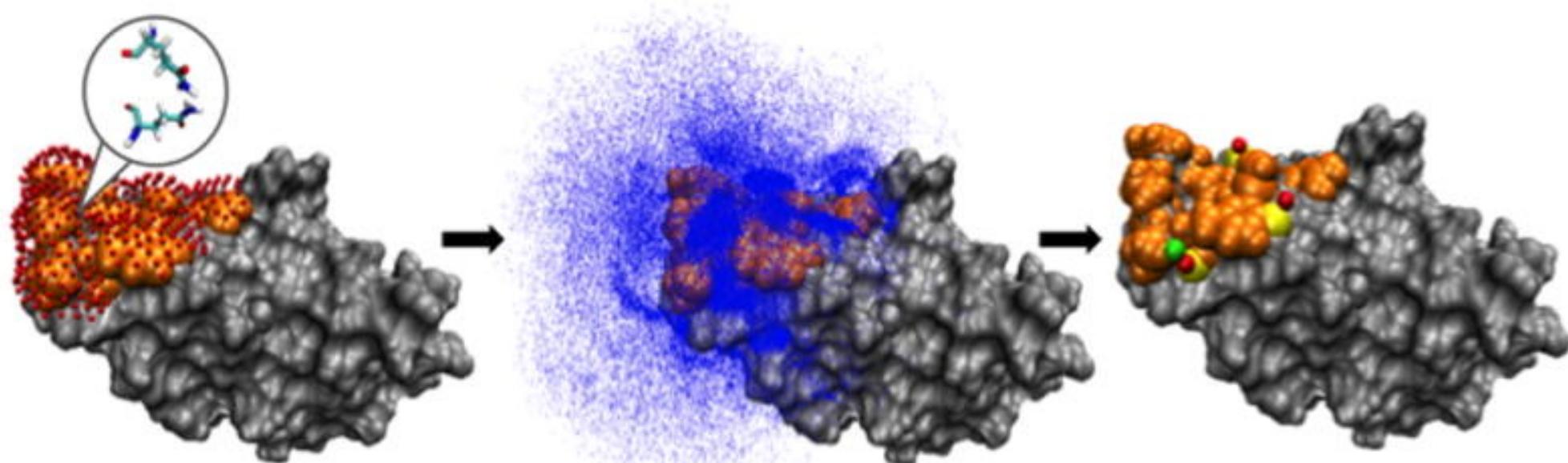
- **Limited accuracy:** The accuracy of pharmacophore mapping is limited by the quality and quantity of available experimental data. Inaccurate or incomplete data can lead to the generation of inaccurate pharmacophore models.
- **Overfitting:** Pharmacophore models can sometimes be overfitted to a specific set of ligands or receptor structures, resulting in poor performance when applied to new data.
- **Complexity of targets:** Some targets, such as protein-protein interactions or membrane proteins, are more difficult to model using pharmacophore mapping due to their complex structures and dynamic nature.

Limitations and challenges

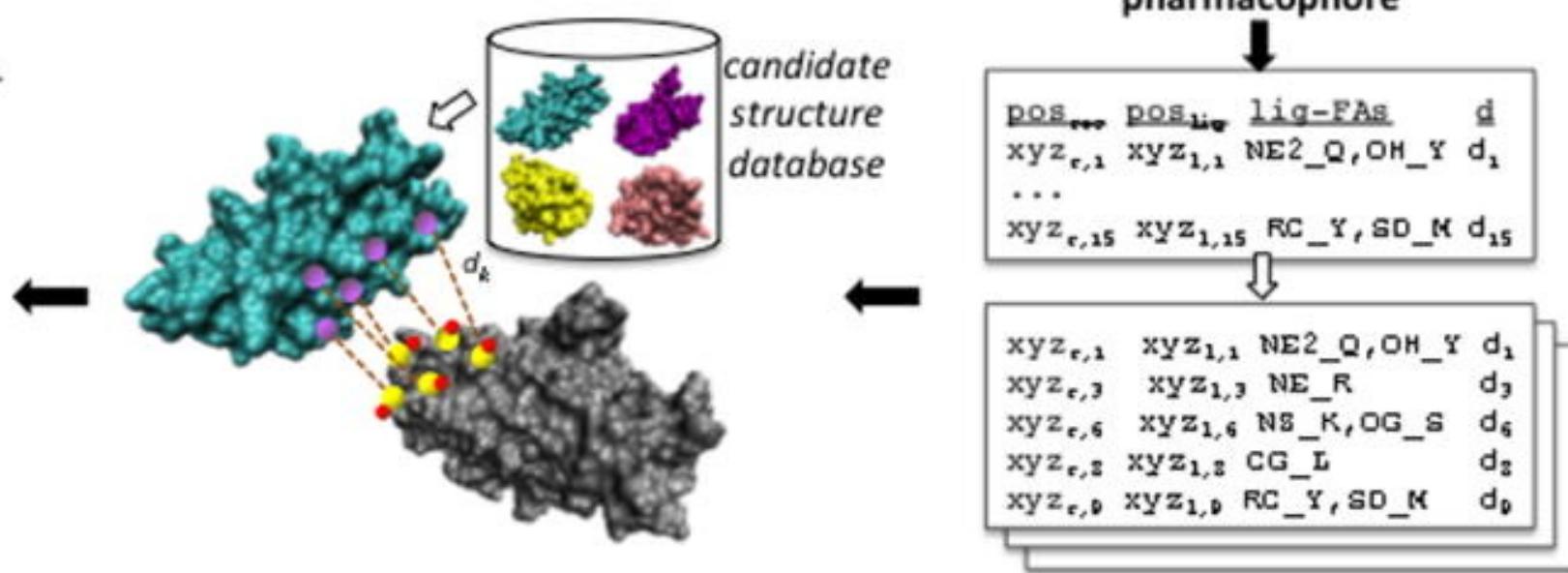
- **Computational requirements:** Pharmacophore mapping requires significant computational resources and can be time-consuming, especially when large numbers of molecules or targets are being studied.
- **Structural variability:** Different conformations or binding modes of a ligand can lead to variations in the pharmacophore model, making it difficult to accurately represent the binding site.
- **Chemical diversity:** Pharmacophore mapping may not accurately represent the chemical diversity of a ligand dataset, leading to biased or incomplete pharmacophore models.
- **Validation:** The validation of pharmacophore models can be challenging, and the performance of the model may be affected by the choice of validation methods and datasets.

Future Directions

- Integration with **machine learning** to improve accuracy and efficiency.
- Use of **big data** to identify new drug targets and lead compounds.
- Application to **protein-protein interactions** for the discovery of new therapeutic targets.
- Inclusion of **multiple binding modes** for better representation of ligand-target interactions.
- Integration with other modeling techniques to provide a more comprehensive understanding of ligand-target interactions and improve drug design accuracy.



<u>CognatePDB</u>	<u>%-rank</u>
1EAJ.A	1
1F5W.A	2
1P6A.B	3
1KAC.B	4
2WBW.B	6
2J12.B	8
2W9L.A	9
2J1K.A	16
...	



```
contactThresholdX_HYD = 3; #2  
distCutoff_HYD      = 10;  #5
```

```
cut2_HYD = $distCutoff_HYD * $distCutoff_HYD
```

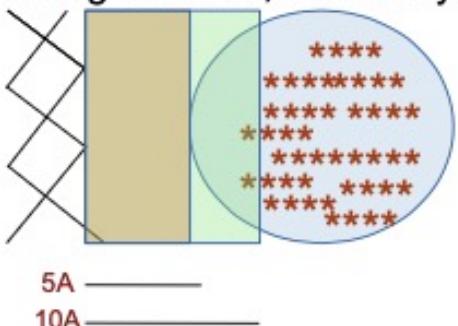
It means we are consider all probes within 10 angstrom.

```
contactThreshold = $contactThresholdX * $nHydAtom (number of hyd.atoms)
```

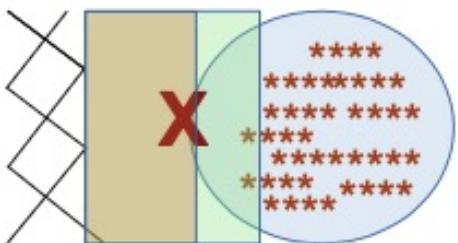
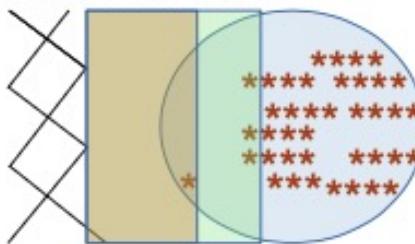
```
ContactThreshold = 3 * 1 =3
```

Assign FA to closest Mesh, MaxAssign_count = 5 A ## dist.sq 25

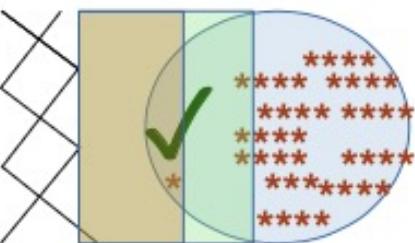
Within this 10 angstrom bin, how many times a probe FA exists.



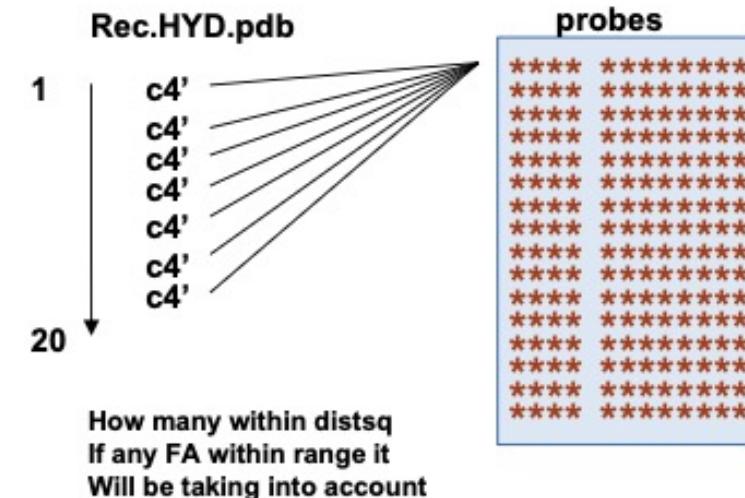
2 times FA within distance cutoff



Countc ++



Script1 Application



So Now If I increase the contactThresholdX
It means I am reducing the chances to identify the FA contacts.

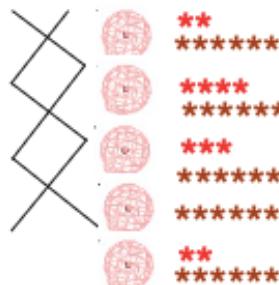
How

Suppose I increased it to 7, 8 etc, its means, I need at least 7, or 8 times C4' FA present within distance cutoff range

This is another reason due to which **countc was mostly 0**.

- 3, 4 contactThresholdX is fine
- Need to increase distance cutoff
- Need to increase MaxAssign_count

identify total countn, countc against each mesh point



→ 4/7, 6/7 i=2, fa = C4'_e, c = 0.57, n = 0.86

	Countn				Countc			
	C4'_e	C4'_t	C4'_l	C4'_j	C4'_e	C4'_t	C4'_l	C4'_j
	0.86	0.50	1.77	1.67	0.57	0.00	0.00	0.45

Script2 Calculate AE ratios

1. Each FA countn files called here and calculate AvgFA

i = 2	****	C4'_e, C4'_t, C4'_l, C4'_j	countn_avgFAs = 1.2	Calculate AvgFAs for against each mesh point
		0.86 0.50 1.77 1.67		

2. Cycle through each interaction type (HYD) to compute A/E ratio

Each FA countc files called here to figure out which mesh is expected to have non-zero countc

****	C4'_e, C4'_t, C4'_l, C4'_j	→	C4'_e, C4'_t, C4'_l, C4'_j	
	0.57 0.00 0.00 0.45		1 0.00 0.00 1	Countc values > 0 is converted to 1

MeshXYZ

ref_qX
ref_qY
ref_qZ

C4'_e, C4'_t, C4'_l, C4'_j bOccupied

i=0, 1
i=1, 0
i=2, 1
i=3, 1

1 means interact, 0 not interact

Int.atomXYZ

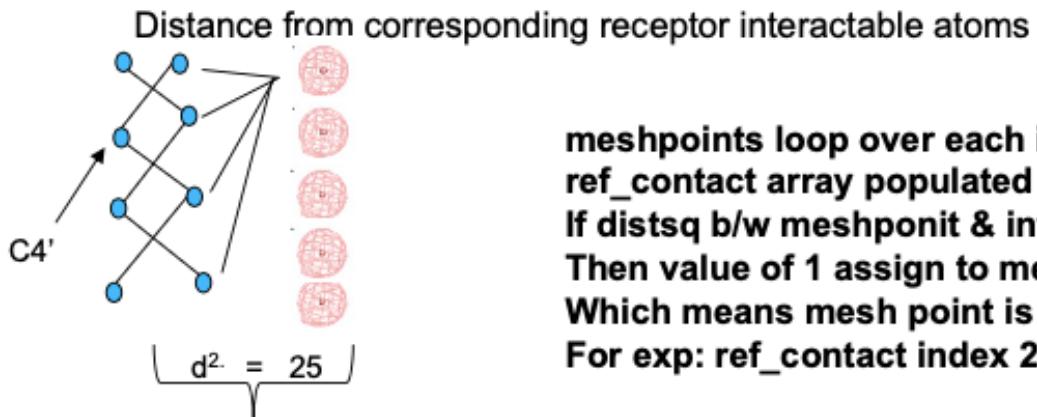
0.57 0.00 0.00 0.00	
0.00 0.00 0.00 0.00	
0.00 0.00 0.60 0.00	
0.00 0.00 0.00 0.10	

Total contact = 277

ref_recX
ref_recY
ref_recZ

distCutoff_HYD = 5

Now here we have AvgFAs (countn), and bOccupied (countc) against each mesh point



meshpoints loop over each interactable atom
ref_contact array populated (bContactReceptor)
If distsq b/w meshponit & interactable atoms < 25
Then value of 1 assign to meshpoint
Which means mesh point is assignable to any interface atom.
For exp: ref_contact index 240 = 1

bContactReceptor

1,
1,
0,
1,
0

of mesh point near interactable atoms = 413

natoms = 20

Mesh_Index	bContactReceptor	bOccupied	AvgFA	totalN
0,	1,	0,	1.25,	1.25,
1,	1,	0,	2.78,	4.03,
2,	1,	0,	0.78,	4.82,
3,	1,	0,	0.78,	5.60,
4	0	0	1.53	6.82

\$totalN += countn_AvgFA[\$i]
\$sum++

totalN= 1.25 + 0 = 1.25
totalN= 1.25 + 2.78 = 4.03
totalN= 4.03 + 0.78 = 4.82



Calculate geometric factor against each mesh_Index

1.25 (AvgFA countn) / 427.2967 (totalN) = 0.002925

2.78 (AvgFA countn) / 427.2967 (totalN) = 0.006519

427.2967 # of meshpoints merged = 492

Calculate countc_expected

Add-up all countc for each FA, (C4'_e, C4'_t, C4'_l, C4'_j)

$$\begin{aligned}\text{Countc_expected for each meshpoint} &= \text{TotalC} * \text{geometric Factor} \\ &= 138.15 * 0.002925 = .3045\end{aligned}$$

Each meshpoint expected value depends on geometric factor

Actual countc is the value comes from countc files i.e countc.c5_C4'_e.dat

Condition if value of actual countc = 0 then AE ratio is 0

else

countc_expected=0.254126604052102 total=138.152 geometricFactors=0.00183947104676083

Countc [90] = 0.429 , countc_expected [90] = 0.254126604052102, aeRatio=1.68813494203088

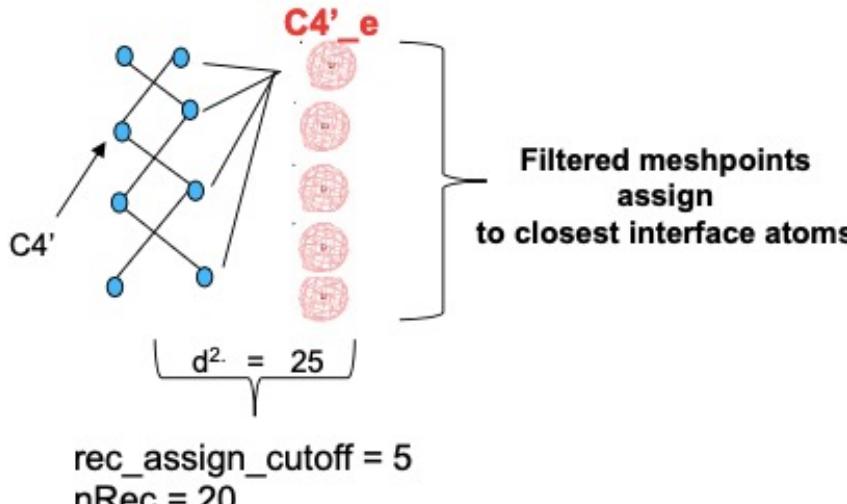
```
aeRatio_cutoff = 1.0;  
countc_cutoff = 1.0;  
clusterPCT   = 20; (20 %)
```

1. Call the countc & aeRatio files for each FA
Pick only mesh point indices which meet the cutoff criteria

	C4'_e,	C4'_t,	C4'_l,	C4'_j
Filtered index HYD	80,	12,	1,	19,
	91,	17,	15,	45,
	99,	39,	200,	67,
	300,	350,	318,	212,
	413	427	639	389

Load receptor interface file (rec.HYD.pdb)

- Generate atomTags for interface atoms such as 1-DA-C4', 2-DC-C4' etc
- Load Ref_xyz coordinates
- recFAhyd.list (load receptor sites = HYD)



Ref_assign2Rec_mesh

```
{  
    '3.DC.C4\' => [  
        80  
    ],  
    '4.DG.C4\' => [  
        150,  
        152  
    ],  
    '19.DG.C4\' => [  
        153  
    ]  
};
```

Ref_assign2Rec_faMask

```
{  
    '3.DC.C4\' => [  
        C4'_l  
    ],  
    '4.DG.C4\' => [  
        C4'_l,  
        C4'_e  
    ],  
    '19.DG.C4\' => [  
        C4'_l  
    ]  
};
```

Ref_assign2Rec_aeRatio

```
{  
    '3.DC.C4\' => [  
        5.01  
    ],  
    '4.DG.C4\' => [  
        2.78,  
        3.53  
    ],  
    '19.DG.C4\' => [  
        4.67  
    ]  
};
```

```

'10.DG.C4\' => [
    → 551,
    553,
    554,
    551,
    553,
    → 555,
    → 516,
    551,
    552,
    553,
    551,
    553
],
];
This is ref_assigned2Rec_faMask = $VAR1
'10.DG.C4\' => [
    → 'C4\'_l',
    'C4\'_l',
    'C4\'_l',
    'C4\'_j',
    'C4\'_j',
    → 'C4\'_j',
    → 'C4\'_e',
    'C4\'_e',
    'C4\'_e',
    'C4\'_e',
    'C4\'_t',
    'C4\'_t'
],
];
;
This is ref_assigned2Rec_aeRatio = $VAR1 = {
'10.DG.C4\' => [
    → ' 5.011',
    ' 2.735',
    ' 3.536',
    ' 1.943',
    ' 3.108',
    → ' 10.333',
    → ' 5.686',
    ' 1.440',
    ' 4.683',
    ' 1.915',
    ' 2.892',
    ' 2.296'
],
];

```

nMember=12

topPCTunique = C4'_j, C4'_e, C4'_l

Representative: Closest mesh to geometric center of assigned mesh

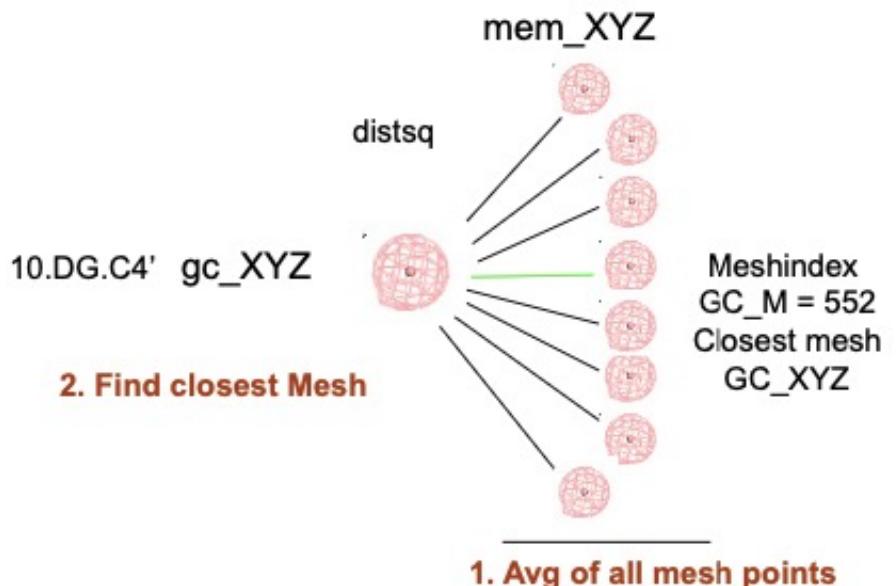
mem_X, mem_Y, mem_Z (mesh-ref.xyz_)
 gc_X, gc_Y, gc_Z (avg of no. of nMember XYZ)

Find the closest mesh to gc
 dX, dY, dZ (dX = mem_X[\$n] – gc_X)
 my \$distsq = \$dx*\$dx + \$dy*\$dy + \$dz*\$dz;

```

This is dx = $VAR1 = '0.5049166666666666';
This is dy = $VAR1 = '-0.7500833333333333';
This is dz = $VAR1 = '0.59675';
This is distsq = $VAR1 = '1.17367640972222'

```



<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8439153/>

[https://www.sciencedirect.com/science/article/pii/S09692126
16303392](https://www.sciencedirect.com/science/article/pii/S0969212616303392)