Module 1:

Steps:

★ Approaches for *In Silico* Drug Designing:

There are two approaches utilized for *in silico* drug designing:

1. Ligand-based Drug Designing:

- Ligand-based drug designing is also known as 'Indirect Drug Designing'. It relies on the knowledge of other molecules (Ligands) that bind to the biological target of interest. In this approach, we don't have to worry about the structure of the biological target (target protein's structure). We are only concerned about the small molecules that have been reported to get bound to the specific biological target.
- Then this known molecule can be utilized to develop a pharmacophore model. This pharmacophore model defines the minimum necessary structural characteristics that a molecule should possess in order to bind to the biological target.

  [In pharmacophore modeling, the features of the known molecule that is responsible for a biological effect is determined and used to search for similar molecules via what is called a similarity search. The molecules found are called the leads, they are checked for their ADME/T properties then they are optimized to derive a drug candidate.**]**
- This small molecule can also be utilized to develop a 2D or 3D Quantitative Structure-Activity Relationship (QSAR), which in turn is used to generate new leads that can bind to the target of interest.

  [QSAR employs statistics and analytical tools to investigate the relationship between the structure of the known ligand and their corresponding effects on interaction with the target. With this information it becomes possible to search for other molecules with similar characteristics.**]**

2. Structure-based Drug Designing:

- Structure-based drug designing is also known as 'Direct Drug Designing'. It relies on the knowledge of the 3D structure of the biological target obtained through experimental methods like x-rays crystallography, NMR, or electron microscopy.
- However, in case you don't have the experimentally determined structure of a target protein, it is possible to create a 3D structure of a target using different Bioinformatics tools and approaches such as Homology Modelling, *ab initio,* and protein threading methods.

  [These different approaches and their relative tools and how these approaches and tools are being utilized for structure prediction of proteins will be discussed in the later modules of the workshop.]

★ Steps followed in Ligand-based CADD:

- The first step in Ligand-based CADD is the formation of both the **2D/3D QSAR** and a **pharmacophore model** of the molecule (Ligand).
- Then **virtual screening** of the identified molecule against large chemical libraries in order to find out more leads of the molecule.
- Then you've to go through both *in vitro* and *in vivo* assays for the **analysis and optimization of the lead** compounds.
- Then synthesize your optimal compounds and report your observations.

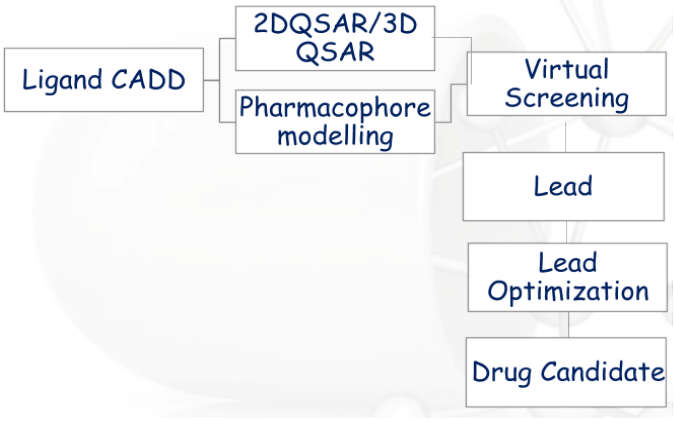

Figure 13: Shows flow diagram of steps followed in Ligand-based CADD.

★ **Example:**

- Let's take the example of **Remdesivir**, which is a promising antiviral drug as shown by various literature, used widely against a large array of RNA viruses including SARS/MERS-CoV, and recent studies have revealed that **Remdesivir** can be used as possible post-infection treatment against COVID-19.

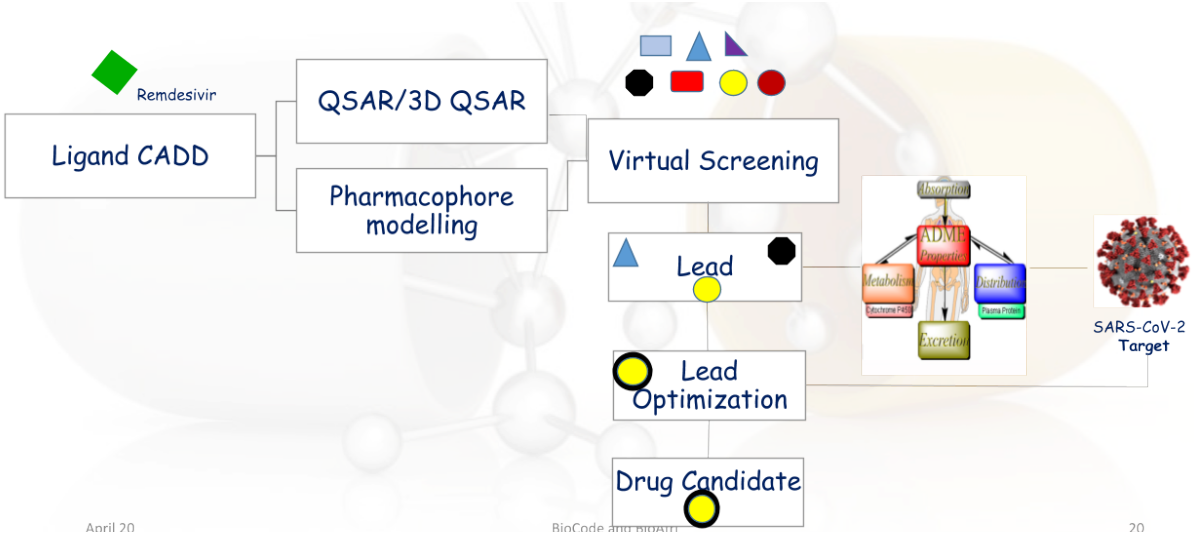

Figure 14: Shows the flow diagram of the steps followed in designing/selecting a suitable ligand against SARS-CoV-2.

- So, **Remdesivir** is our small known molecule, against which virtual screening has been done which includes similarity searching to generate lead compounds.
- Then the ADME/T properties of the lead compound are evaluated to filter out optimal leads.
  [ADME/T stands for Adsorption, Distribution, Metabolism, Excretion, and Toxicity of the lead compounds.]
  [The analysis of ADME/T properties using different tools will be demonstrated in the later modules of the workshop.]
- After selecting the optimal leads, the molecule is then docked against the target protein of SARS-COV-2 and the best hit is selected.
- Then the ligand-target interactions are studied using Molecular Dynamics approaches.
- Afterward, the leads are optimized to improve bioavailability, absorption, distribution, metabolism, excretion, and reduced toxicity of the drug to get the drug candidate.

★ **Steps followed in Structure-based Drug Designing:**

- Structure-based drug designing depends upon the knowledge of the biological target's structure.

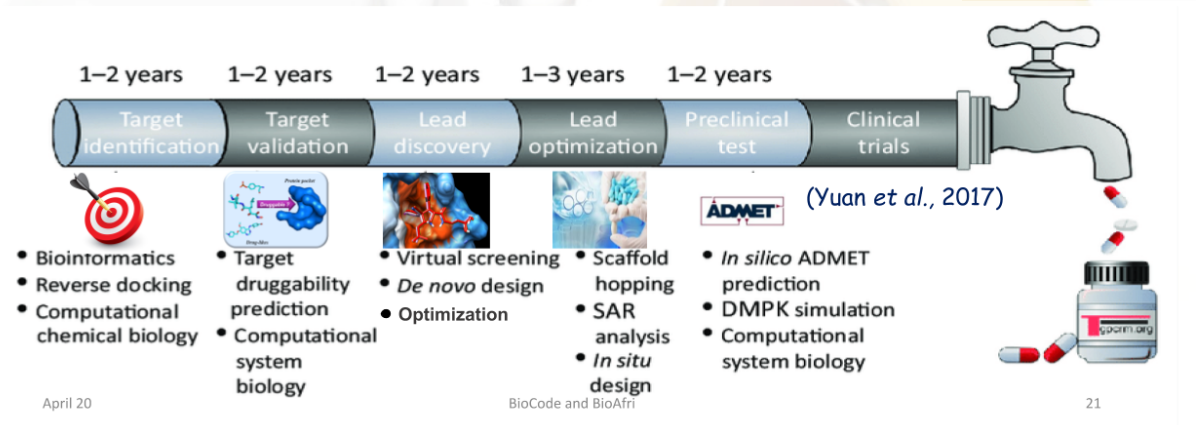

Figure 15: Shows general steps involved in Structure-based CADD.

- The first step is the **target identification** which can be identified from literature, via reverse docking techniques (i.e., by evaluation of the binding site of an already present co-crystallized ligand.)
- The 3D structure of the protein target can be obtained via different *in vitro* and *in silico* approaches which includes:

In Vivo Approaches	X-rays crystallography, NMR, Electron Microscopy
In Silico Approaches	Protein Data Bank (PDB) and other protein structure prediction tools.

[if there's no experimentally determined structure available of your target protein, you can use different protein structure prediction tools based on homology modeling or other methods.]

- Once the target protein molecule has been identified the next step is to **validate the target** by evaluating its druggability.
  [Druggability- the ability of the target protein's activity to be altered by the small molecule (i.e., drug).
- The druggability can be determined by precedence, prediction, or the structure of the target protein.
    - Precedence- the druggability of the target depends upon the condition if other members of the target protein family are druggable or not.]
    - Prediction of the target- the amino acid sequence of the target of interest helps in the evaluation of druggability.
    - Structure prediction involves the identification of the cavities or binding pockets present on the target protein by calculating the physicochemical and geometric properties of the binding pockets of the target.
- The next step is the **lead discovery**, which can be done via 'virtual screening'.
- Another way of lead discovery is '*de novo'* designing of the ligand which utilizes the information of the active site composition of the target and the orientation of various amino acids at the binding site of the target in order to design the ligands that are specific to that particular protein target.
- The next step is the **lead optimization** which refers to increasing the affinity, selectivity, efficacy/potency, and oral bioavailability of the lead molecule. This can be done via 'scaffold hopping' or 'SAR' (Structure-Activity Relationship) analysis.
- The next step is the **preclinical trial** which involves testing the lead either *in vitro* or *in vivo* in animal models to develop an 'Investigational New Drug (IND)'. Using different computational tools to check and evaluate the ADME/T properties can also serve this purpose.
- Once the lead elapses the preclinical testing, the IND is then tested in humans.

★ **Example:**

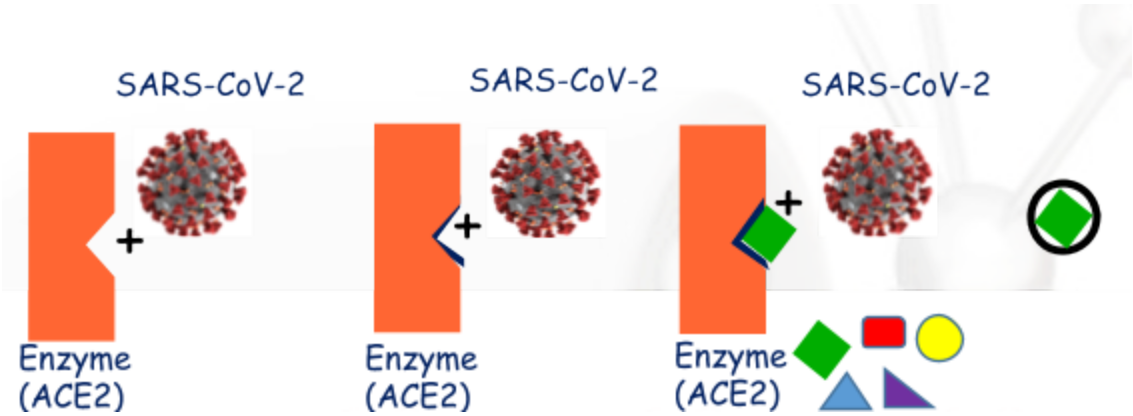

Figure 16: Target (ACE2) and drug (ligand) interaction to prevent the activity of SARS-CoV-2.

- The first step shows us our target. Our target here is the Angiotensin-converting enzyme 2 receptor (ACE2). The target was derived from the reports in the literature which have shown that the spike protein of SARS-COV-2 binds to this receptor to get into the cells of the lungs.
- Once the structure of our target is determined, the next step is to validate the target by checking for its druggability.
- After that, the next step is to screen for the compounds that will bind to the target with stronger affinity than the spike protein in the SAR-COV-2.
- So let's assume there are different molecules competing for the binding site of this receptor with the spike protein. So eventually one binds better than the spike protein of SAR-COV-2 than other molecules and is taken as the lead compound.
- The next step is to optimize the leads to improve its bioavailability and ADME/T properties.
- Once it is ascertained then the drug is checked for drug-likeness and its ADME/T properties.
- Then these are moved for the clinical trial.

★ Computational tools for Drug Designing:

- Following are some Databases and tools that are being used in the various step of Computer-Aided Drug Designing (CADD):

1. Databases:

- Databases are large repositories of organized information so that it can be easily accessed, managed, and updated.
- We can get the structure of the protein as well as ligands from different databases, some of which are described here:

Databases	Functions
PubChem	To find the structure of ligands
Zinc	To find the structure of ligands
PDB	For the retrieval of target protein structure
BindingMOAD	For the retrieval of target protein structure
PDBbind	Information on the experimentally measured binding affinity values for protein-ligand complexes is available here
SMP (Small Molecular Pathway)	For the retrieval of information about pathway elucidation and pathway discovery of small molecules in metabolomics, transcriptomics, proteomics and systems biology

2. Draw tools:

- Small molecules or drugs are composed of their chemical component which is responsible for the actions they elicit.
- In drug discovery, we get the structures of ligands by drawing their chemical structures. The preparation of proteins/ligands is also an important step in drug design.
- Following are some tools that are being utilized for drawing such chemical structures and protein visualization:

Drawing Tools	ChemDraw, MarvinSketch, ACD/ChemSketch, Pymol and UCSF Chimera
Visualization Tool & Analysis	Pymol and UCSF Chimera.

3. Homology modeling tools:

- If you don't have the experimentally determined structure of a target protein then it might be possible to create a homology model of the target protein based on the experimental structure of a related (Homologous) protein via homology modeling.

- There are various tools available for the prediction of a target protein structure using Homology Modelling techniques which include:

Servers	ROBETTA, I-TASSER, LOMETS, and SwissModel
Programs	MODELLER

4. Binding site prediction tools:

- One of the most important steps in target identification is binding site prediction.
- Following are some tools that are being utilized for this purpose:

Servers	Pocketome, 3DLigandSite, metaPocket, and FINDSITE
Programs	CAVER.

5. Molecular Docking tools:

- Molecular docking involves the interaction of two or more molecules to predict the stable complex. Depending upon binding properties of ligand and target, it predicts the three-dimensional structure of any ligand-target complex.
- Following are some tools that are being used for this purpose:

Servers	AutoDock Vina, iGemdock and GOLD
Programs	SwissDock, DockingServer and 1-ClickDocking

6. Molecular Dynamics (MD) Simulation Tools:

- If you want to derive, represent, and manipulate the structures and reactions of molecules, and those properties that are dependent on these three-dimensional structures, you need to analyze them via molecular dynamics. MD simulation studies also help in understanding the interaction between docked ligand and protein complex.

- Following are some tools available for MD simulation studies *in silico*:

Servers	SwissParam and SwissSideChain
Programs	GROMACS, CHARMM-GUI, and AMBER

7. **Ligand screening tools:**
    - Ligand virtual screening can be done using the following tools:

Servers	Swiss similarity, AnchorQuery and Ligandscout
Programs	Discovery Studio

8. **Target prediction tools:**
    - This can be done using servers like:
        - Swiss Target Prediction.
        - MolScore-Antibiotics.
9. **ADME/T analysis tools:**
    - Drug Likeness is a qualitative concept used in drug design for how "druglike" a substance is with respect to factors like bioavailability. This depends upon the pharmacodynamics and pharmacokinetics analysis of a drug.
    - Such an analysis can be done using the following tools:

Servers	VolSurf and Gastroplus
Programs	Swiss ADME, OSIRIS Property Explorer, Molinspiration and PACT-F

★ **Step followed in the process of Virtual Screening (VS):**

- The first step of VS involves **screening** the selected compounds against the libraries of compounds available online.
- These libraries/Databases include:
    - Zinc Database.

    - Maybridge.
    - Selleckchem.
    - Asinex.
    - ChemBridge.
    - DrugBank.
    - Natural Compounds.
- The second step is **similarity searching** which involves the matching Chemical, Biological, and pharmacological features of two compounds to generate pharmacophore models. This can be done by performing E-Pharmacophore screening,
- Following is the list of tools that are being used for the screening of pharmacophore:

Servers	SHAFTS and PharmaGist
Programs	Catalyst, LigandScout, Pharmer, SHAFTS, PharmaGist, Phase-Schrodinger

- The third step is **Docking** which is the most important step of VS. In this step, ligands are docked against the target protein in different conformations. These conformations are then given specific scores based on the scoring algorithm used or the tool used for docking.
- The ligand should exactly fit within the target's binding site just like the lock-and-key concept used in the case of enzymes.
- The fourth and most crucial step is **cherry-picking** in which we identify the lead compounds that are promising drug-like molecules. This can be done by considering different parameters like the H-bonding between ligand and target molecules, Binding affinity values, or docking scores.