International Journal of Medical and Pharmaceutical Sciences
Open access peer-reviewed Journal
Volume 2, Issue 9, Page 440-459.
https://doi.org/10.5281/zenodo.22917842

Review Article

Green CADD: Sustainable Approaches in Drug Discovery

Akash Kundekar*[1], Nilesh Chougule[1], Prasad Yadav[2]

[1]*Department Of Pharmacy Ashokrao Mane Institute of Pharmacy, Ambap*
[2]*Assistant Professor, Department of Pharmaceutical Chemistry, Ashokrao Mane Institute of Pharmacy, Ambap*

The implementation of sustainable techniques throughout the process of discovering and developing new drugs has speed up because there is an increasing demand for novel pharmaceutical methods that are also ecologically responsible. Green chemistry concepts, which prioritize safer reactions, renewable materials, and waste minimization, are gradually replacing traditional methods, which are characterized by high energy consumption, dangerous solvents, and excessive waste formation. Effective, low-resource therapeutic discovery has been made possible by parallel developments in computer-aided drug design (CADD), which include molecular docking, virtual screening, QSAR modeling, pharmacophore creation, and MD simulations. Predictive accuracy is further improved, experimental burden is decreased, and early ADMET and toxicity evaluation are supported by combining machine learning and artificial intelligence. When combined, these technologies create a contemporary framework where sustainability and therapeutic innovation meet, facilitating quicker, greener, and more affordable medication development.

Keywords: Computer-Aided Drug Design (CADD); Quantitative Structure–Activity Relationship (QSAR); In Silico Drug Discovery; ADME/Toxicity Prediction; Molecular Docking; Pharmacophore Modeling; Virtual Screening; Machine Learning; Artificial Intelligence; Drug Discovery.

INTRODUCTION

The drugs industry is at a turning point where environmental responsibility and innovation have to merge. Despite effective in producing life-saving drugs, traditional drug discovery and development paradigms have been characterized by resource-intensive procedures, high waste production, and notable environmental effects. These traditional methods have mostly depended on energy-intensive processes, petroleum-derived chemicals, hazardous solvents, and linear production models that give emphasis to therapeutic results at the expense of environmental effects. Pharmaceutical development has an impact on the environment in a number of ways. Even trace amounts of active pharmaceutical ingredients (APIs) that reach aquatic environments through human urine and factory effluents have been shown to be ecologically harmful. According to estimates, pharmaceutical production can produce between 25 and 100 kg of garbage for every kilogram of API generated, with manufacturing processes producing hazardous waste at ratios that are substantially greater than other chemical industries. Additionally, a significant portion of greenhouse gas emissions are caused by the carbon footprint of energy-intensive research, development, and production facilities. As a result, rather than treating environmental concerns as coincidental compliance difficulties, pharmaceutical companies have begun to incorporate sustainability ideas into their primary initiatives for research and development. The scientific literature shows that technologies that improve therapeutic innovation and diminish environmental consequences are being developed at a rapid pace. Synthetic routes have been revolutionized by green chemistry principles, hit discovery and optimization strategies have been transformed by computational methodologies, and manufacturing

improvements have significantly increased resource efficiency. These developments raise the prospect of a pharmacological paradigm in which therapeutic and environmental goals are complementary rather than antagonistic. This article's primary goal is to examine the contemporary sustainable practices used throughout the pharmaceutical value chain, from product lifecycle management and production to early discovery. [1,2,3]

The Role of Sustainability In Modern Pharmaceutical Production:

Figure no :1 Sustainable Pharmaceutical Development

A growingly important component of contemporary pharmaceutical manufacturing is sustainability, as the sector struggles with environmental issues. There have been calls for greener, more responsible manufacturing techniques in the pharmaceutical industry as a result of increased awareness of pollution, resource depletion, and climate change. According to (Rame, Purwanto, and C. sudarno,2024) sustainability in pharmaceutical production is both morally and financially required since it can save expenses, boost productivity, and improve a company's reputation. Sustainability in pharmaceutical production refers to enhancing resource utilization, eliminating waste, and minimizing environmental damage. Conventional pharmaceutical manufacturing techniques use a lot of energy and produce a lot of waste, frequently containing dangerous compounds that need to be thrown away correctly to prevent harming the environment. However, sustainability emphasizes the implementation of greener, better-performing methods and technologies that lessen the production process's total ecological impact. Using sustainable energy sources, conserving water, and implementing zero- waste plans are a few examples of these activities Furthermore, market and regulatory pressures are becoming more and more associated to sustainability in pharmaceutical manufacture. Investors and regulatory agencies have come to empathize with in recent years how important

environmental stewardship is to long-term corporate performance. Pharmaceutical businesses have been encouraged to reevaluate their manufacturing processes as a result of the implementation of stronger environmental restrictions, such as enforced waste management procedures and limits on greenhouse gas emissions. Additionally, there is a growing desire from consumers for eco-friendly items. Pharmaceutical businesses are under pressure from investors, patients, and healthcare professionals to adopt more sustainable practices; those that don't run the danger of losing market share and tarnishing their reputations. There are also major advantages for public health from sustainable pharmaceutical production. For instance, lessening the negative effects of medicine manufacture on the environment can help avoid health issues linked to pollution. Frequently found in wastewater and the environment, pharmaceuticals have the potential to negatively impact ecosystems and wildlife pharmaceutical businesses can help create cleaner air and water by lowering the number of these substances released into the environment through using more eco-friendly production methods. [4,5,6]

Integration of Green Chemistry in Drug Discovery:

The term "green chemistry" first appeared in the early 1990s, when Paul Anastassi and John Warner defined it. Because it may deal with chemical innovation and simultaneously accomplish economic and environmental goals, green chemistry is gaining traction in the scientific community. Consequently, the "design of chemical products and processes that reduce or eliminate the use and generation of hazardous substances" is the definition of "green chemistry." Though it is primarily dedicated to a more sustainable approach to the environment, green chemistry, which is also called sustainable chemistry, emphasizes using technology to lessen pollution and decrease the use of non-renewable resources. It aims to find affordable methods that produce better results and work more effectively. [7,8,9] The scientist Paul Anastas and his associates first presented the fundamental ideas of green chemistry in the 1990s. These guidelines offer a foundation to establish chemical processes that are safer, more efficient, and resource-efficient while also being more economically feasible. The following twelve principles serve as the cornerstone of this strategy, even though the precise uses of green chemistry may differ among businesses

- pollution Prevention: This idea promotes the creation of procedures that reduce or completely do away with the production of trash. Waste can be cut at the source by emphasizing the economical use of resources and energy, which lowers disposal expenses and protects the environment.

- Atom economy: “Atom economy” refers to the process design that makes sure every atom from the raw materials is used efficiently in creating the final product This lessens the requirement for surplus solvents and reagents, which are frequently thrown away as waste.

- fewer dangers Hazardous Chemical Synthesis: According to this theory, fewer toxic and dangerous reaction conditions, solvents, and reagents should be chosen. It encourages the use of safer chemicals and steers clear of materials that endanger the environment and public health.

- Energy-efficient design: Green chemistry promotes the use of energy-efficient procedures that reduce the demand for high pressures or temperatures. Reaction conditions with a lesser amount of energy usage and more efficient catalysts can accomplish this.

- Utilization of Renewable Feedstocks: Green chemistry relies heavily on the use of biodegradable, renewable materials as raw materials. This guarantees the sustainability of the chemical process and lessens reliance on non-renewable resources like petroleum chemical reaction.

- Cut Down on Derivatives: Avoid using needless derivatization reactions, like protecting groups. These reactions frequently produce waste and need for more reagents.

- Green chemistry encourages the use of catalysts to accelerate chemical processes more effectively, frequently in milder environments, hence

lowering the demand for energy- intensive procedures and stoichiometric reagents.

- Design for Degradation: After their useful lives, chemical-related products should be made to decompose into non-toxic forms. This lessens the chance of pollutants that linger in the environment contaminating the environment over time.

- Analysis in Real Time to Prevent Pollution: Process control, ensuring that pollutants are stopped at their source and emissions are kept to a minimum, requires real-time perceiving and analytical methods.

- Safer Chemistry for Accident Prevention: The idea promotes creating chemical processes and products that are safer by nature, lowering the chance of spills, fires, or toxic releases, and preventing accidents or hazardous situations.

Green chemistry principles, which emphasize waste reduction, energy efficiency, and the use of non-toxic and renewable materials, provide a revolutionary approach to drug development and production when applied to the pharmaceutical industry. These principles also align with the growing demand for more sustainable practices in the pharmaceutical sector. Additionally, they open the door for new developments in medication synthesis, including the possibility of more economical and environmentally friendly production techniques. [4,5,6]

Integration Of Green Chemistry And Cadd In Drugs Discovery:

In the pharmaceutical business, computer-aided drug design, or CADD, has become a key strategy for finding and creating novel therapeutic agents. The incorporation of green sustainability principles into CADD processes has grown in significance as environmental concerns have increased. This review of the literature attempts to investigate how CADD and green sustainability intersect with drug discovery, emphasizing methods, findings, and implications for further study. CADD includes a range of computational methods that support drug candidate design and optimization. Extensive laboratory effort is frequently required for traditional drug development techniques, which can be costly and harmful to the environment. The goal of integrating green chemistry concepts into CADD is to increase the overall sustainability of drug development processes while minimizing waste and energy usage. The contribution of CADD to the advancement of green sustainability has been emphasized in numerous research. For example, researchers can assess enormous libraries of compounds using virtual screening approaches without requiring considerable physical production, which reduces material waste. Furthermore, the identification of safer and more effective therapeutic candidates may result from the prediction of biological activity and toxicity made accessible by molecular docking and quantitative structure-activity relationship (QSAR) modeling Additionally, it has been demonstrated that using machine learning algorithms in CADD may improve medication design while upholding sustainable standards. Machine learning has the ability to cut down on the time and resources needed for drug development by using massive datasets to find patterns and forecast results more effectively than conventional techniques

CADD Key Techniques and Approaches:

The potent and multidisciplinary discipline of computer-aided drug design (CADD) plays a crucial role in modern drug discovery. It blends mathematical methods with biological expertise to find and enhance possible medication options. This incorporation Using various approaches enhances CADD's flexibility and effectiveness in the pharmaceutical sector. CADD's scope and adaptability stem from the abundance of of methods and strategies that support this field. The efficacy of this field is based on its several approaches, which include drug prediction and molecular modeling metabolism. Adhering Lipinski's rule is essential in CADD in order to optimum oral drug likelihood, where chemicals ideally reduce criteria violations such as suppliers and acceptors of hydrogen bonds, molecular weight, and lipophilicity. The method of drug development is accelerated and improved by the strategic integration of CADD concepts and following to drug-likeness criteria, illustrating the adaptability and significant significance of CADD in a variety of sectors including the pharmaceutical industry [10,11,12]

Structure-Based Approach In Drug Design:

If the target's spatial structure is established, straightforward methods or structure- based drug design (SBDD) may be implemented. The traits and characteristics of the macromolecule's spatial structure can be used to build compounds with capabilities that are complementary to the target area. Protein 3D structure can be confirmed via NMR, X-ray crystallography and in silico homology-based prediction methods. Once the three-dimensional structure is known, the binding/active site of the protein can be found. Typical techniques in SBDD include structure-based pharmacophore modeling, molecular docking, virtual screening (SBVS), and molecular dynamics (MD) simulations [13,14,15]

Virtual Screening:

In the process of developing drugs, virtual screening has been frequently employed for scaffold hopping, lead optimization, and lead identification. It offers a quick and affordable substitute for high-throughput screening in the search for novel drugs. The computational approaches for virtual screening fall into two main categories: (1) Ligand- based drug design methodologies, such as ligand similarity, and (2) structure-based drug design methodologies, such as ligand docking. Even though protein-ligand docking is intended to use the target protein's three-dimensional (3D) protein structure to predict the binding modes and affinities of ligands to the target, ligand similarity methods take advantage of the fact that ligands that mimic an active ligand are more likely to be active than random ligands [16,17,18,19].

Structure Based Virtual Screening:

The most common method used for in silico drug discovery is structure-based virtual screening (SBVS). SBVS aims to find the best way for two chemicals to interact and create a stable complex. It does this by using evaluation functions to measure the strength of non-covalent interactions between a ligand and its biological target. SBVS has the advantage of saving money and time by eliminating the need to screen millions of tiny compounds. The molecule may be computationally assessed prior to production since its physical existence is not necessary [13,14,15].

Target-based virtual screening (TBVS), another designation for structure-based virtual screening (SBVS), intends to forecast the most suitable ligand-target interaction to generate a complex. the top of the list shows the most promising compounds, which is based on the ligands' affinity for the target. To predict how target will interact with each chemical component using computer stimulations, SBVS strategies require knowledge of the target protein's three-dimensional structure. In this approach, scientists select chemicals from a database and sort them by how strongly they attach to the receptor site. Molecular docking stands out among SBVS techniques as a consequence of its low computational cost and successful outcomes. The analyzed molecules are arranged based on their affinity to earn the receptor site During SBVS, researchers can find ligands that are more likely to show some pharmacological activity with the molecular target.The probability that a binding site will describe the affinity between the target and ligand is confirmed using score functions. The crucial element of the docking process in this procedure is a trustworthy scoring function

Using SBVS has both benefits and drawbacks. Some of the benefits include:

1. The time and expenditure required to screen millions of tiny molecules has decreased.
2. The molecule can be computationally tested even before it is synthesized considering its physical existence is not mandatory.
3. SBVS are able to benefit from a number of tools.

The disadvantages can be highlighted as the following:

1. Certain tools are far more efficient in certain situations than others
2. Because it has challenged to define parameters for the intricacy of ligand receptor binding interactions, it is challenging to precisely anticipate the correct binding position and classification of substances.
3. Both false positives and false negatives may result from it. [20,21,22].

Molecular Docking:

Its in-silico technique forecasts where ligands or small molecules will be arranged within the target protein's active region. It has been frequently employed in virtual screening for lead component optimization and currently is primarily used to accurately estimate how well ligands bind to their receptors and how strong that connection is with their receptors. These methods consist of three objectives that are related to one another: bio affinity, virtual screening, and binding pose prediction. The search algorithm and scoring function for ligand creation and analysis form the foundation of these methods' tools [23,24,25].

There are two steps to the docking problem.

1. investigating the ligands' conformational space once they attach to target molecules.

2. This set is being scored, or sorted predicated on the estimated binding affinity. [26,27,28]

Docking Techniques:

Docking is a technique used in chemical simulations that predicts a molecule's preferred orientation by binding to an additional molecule to generate a stable complex. By using a score system, the preferred orientation can be utilized to forecast how strongly two molecules would associate or bind.

Table 1 Docking tools [23]

Docking software	Shape algorithm
Dock	Shape fitting
Auto dock	Lacker Kian algorithm genetic algorithm
Gold	Genetics algorithm
Glide	Monte Carlo sampling
Ligand fit	Monte Carlo sampling

Search Algorithms:

The binding site is carefully investigated for ligand orientations and conformations using search methods. A well-designed docking methodology will produce the most realistic ligand location at the binding site as well as the most feasible ligand conformations. While flexible docking adds conformational degrees of flexibility to the ligands' translations and rotations, rigid docking uses translational and rotational degrees of freedom to explore various ligand locations at the active binding site. The chemistry and geometry of the involved atoms are checked the genetic algorithm incremental construction and other methods are used by search algorithms to predict the correct conformation of ligands. According to Ruiz-Tagle et al. there are three categories of algorithms that take ligand flexibility into account: deterministic, stochastic, and systematic. To achieve superior outcomes, some software combines multiple of these strategies. Systematic search algorithms take advantage of how molecules can move and change. They usually build these molecules step by step at the place where they bind The number of assessments the algorithm must carry out rises as the degree of freedom (rotatable bonds) grows. The algorithm must make more assessments when the degree of freedom (rotary connections) is increased, which lengthens the time needed to execute the method. Termination criteria are added to the algorithm to stop it from attempting solutions that are in the space known to provide incorrect answers, thereby cutting down on execution time. Software that employs systematic search algorithms includes DOCK 6, FLEXX and Glide Stochastic search techniques explore multiple potential conformations by randomly altering the ligand's spatial conformation, often altering one system degree of freedom at a time The unpredictability of reaching a decent solution is the primary issue with stochastic algorithms. Therefore, multiple separate stochastic algorithm executions are typically carried out in order to minimize this issue. Stochastic research algorithms include the genetic algorithms used by GOLD and AutoDock4 and the Monte Carlo (MC) methods used by Glide and MOE. The initial state is in assigned to figuring out the movement that can be performed to create the following state during the execution of a deterministic search algorithm; this movement typically needs to be equal to or less than the initial state's energy.

Deterministic algorithms have the drawback of frequently becoming stuck in local minima due to their inability to overcome obstacles; however, there are workarounds for this issue, including raising the simulation temperature. Deterministic algorithms are exemplified by energy minimization techniques. Another deterministic search strategy is molecular dynamics (MD), which is employed by DOCK 6 . The runtime becomes a limiting factor for simulations given that structure databases can contain millions of ligands and targets, despite MD's high processing demands and promises of improved findings and full-system flexibility. [20,21,22]

Scoring Functions:

Molecular docking software uses scoring functions to mathematically estimate the strength of non- covalent interactions between a ligand and its molecular target. one of the most crucial elements of sbvs is a scoring function, that has mainly in charge of forecasting the binding affinity between a target and its ligand candidate. consequently, the primary factor impacting the success or failure of docking tools is the scoring functions. generally speaking, molecular docking uses scoring functions in three significant ways. they can first be utilized to ascertain the target-ligand conformation and the ligand binding location. allosteric sites can be found using this method. secondly, they can be employed to forecast the ligand-protein binding affinity. third, they can also be applied to lead optimization. according to the majority of authors categorize scoring functions into three types: force field (ff), empirical, and knowledge- based. liu and wang distinguish between two additional categories of scoring functions: hybrid approaches and machine-learning-based approaches.

Different Types of Scoring Functions:

A. Force-Field Based:

According to the principles of molecular mechanics and derived from experimental data, the force field scoring functions are based on the intermolecular interactions between the ligand and target atoms, including the van der Waals, electrostatic, and bond stretching/bending/torsional force interactions. [20,21,22] When combined with thermodynamic integration techniques or free energy perturbations, in general, force-based scoring functions are reliable for determining binding free energies. The contributions from numerous interaction energy categories, encompassing as hydrogen bonds, electrostatic interactions, and van der Waals interactions between ligand and protein atoms, are summed up by force field-based scoring functions. These scoring functions lack the capacity to compute entropic and solvation terms because their formulation was designed to describe the enthalpic gas phase for energetics and structures [26,27,28]. Gold score and Sybyl/D-Score are a few documented force-field scoring functions. [20,21,22]

B. Empirical:

It is believed that binding energies can be roughly expressed through the sum of separate uncorrelated variables is the basis of the design of empirical scoring functions. As a sum of multiple parameterized functions, these scoring functions are able to replicate experimental data, including binding energies and conformations. Although empirical functions are based on approximations that resemble force-field functions, their terms are frequently easy to evaluate, which makes them appealing. Regression analysis with experimentally measured binding energies and maybe X-ray structural data yields the coefficients of the various terms. These approaches' reliance on the molecular data sets for fitting and regression analysis is a drawback. This can frequently result in distinct weighting factors for the different terms. Therefore, it is difficult to recombine terms from differently fitted scoring functions into a new scoring function. Terms that account for non-bonded interactions can be implemented in empirical scoring functions in a variety of ways. Non-enthalpic contributions, such the so-called rotor term, which approximates entropy penalties upon binding from a weighted sum of the number of rotatable bonds in ligands, can be included in empirical scoring functions. There is which is distinct from the typical empirical scoring function that is derived from the statistical preference of atom pairs according to the Boltzmann and Helmholtz constants. The total free binding energy, that involves hydrogen bond, contact surface, and entropic contributions, influences LUDI and Flex X, F-Score adds another variable to account for aromatic interactions; and Chem score assesses contacts

between pairs of hydrophobic atoms; Fresno, the software used for peptide docking, explicitly accounts for ligand de solvation and a continuous electrostatic mode is employed to compute de solvation energy [26,27,28]. Several widely used empirical scoring functions are Sybyl-X/F-score Glide-Score and DOCK 6 empirical force field. [20,21,22]

C. Knowledge-Based:

merely focusing on binding energies, knowledge-based scoring systems aim to replicate experimental structures. Protein-ligand complexes are represented using comparatively basic atomic interaction-pair potentials in knowledge-based functions. Depending on their chemical surroundings, several atom-type interactions are established. Thus, knowledge-based scoring systems aim to implicitly capture binding effects that are challenging to model explicitly, just like empirical techniques do. The computational simplicity of many knowledge-based scoring algorithms is a key draw since it enables effective screening of enormous compound data sets. Their derivation focuses primarily on information that is implicitly embedded within a small number of protein ligand complex structures, which is a drawback. Solvent-accessibility modifications to pairwise potentials arealso included in Potential of Mean Force (PMF) and Drug Score. Another scoring system in this family that assesses protein-ligand interactions using pairwise atom potentials is called SMOG (Small Molecule Growth). [26,27,28] Para Docks is one program that employs a knowledge-based scoring function. [20,21,22]

D. Consensus Scoring:

The most recent development in this subject has been the introduction of consensus scoring methodologies, which use very divergent properties to increase performance in virtual screening according to the shortcomings of present scoring functions. Consensus scoring increases the likelihood of finding "true" ligands by combining data from several assessments to offset flaws in individual values. X-CSCORE, which includes the GOLD- like, DOCK-like, Chem Score, PMF, and Flex X scoring functions, is a prime example of a consensus scoring implementation [26,27,28]. Furthermore, have proposed machine-learning-based techniques as a fourth category of scoring functions. The ability of machine learning-based techniques to produce accurate predictions has drawn interest. Although many studies have improved SBVS algorithms using machine learning, we are unaware of any medications created by fusing machine learning with SBVS. Nonetheless, a new antibiotic that can stop the growth of E. coli bacteria was found by some researchers using machine learning approaches. These methods have been applied in quantitative structure-activity relationship (QSAR) analysis to forecast a range of small molecule compounds' physical-chemical (like the molecule's hydrophobicity and stereochemistry), biological (like its activity and selectivity), and pharmaceutical (like its absorption and metabolism) characteristics. Modern QSAR analyses may be utilized for create statistical models that determine protein-ligand binding scores across these kinds of scoring functions. SFCscoreRF, RF-Score-VS ID-Score, SVR-KB, SVR-EP, NN Score 2.0 and C Score are a few examples of this type of scoring functions. Because they incorporate two or more of the previously established scoring function types— force field (FF), empirical, knowledge-based, and machine learning-based—into a single scoring function, some hybridized scoring functions are difficult to categorize into any of the aforementioned groups. They are referred to as hybrid scoring functions as a result. A multiple linear regression fitting approach yields two or more scoring function components, which are then linearly combined to create the hybrid scoring function. For instance, the Galaxy Dock score function combines knowledge-based, empirical, and physics-based score terms, offering the benefits of each. Decoy posture discrimination tests showed an improvement in performance as a result. Some recently published examples of this type of scoring function include Galaxy Dock BP2, which combines force field, empirical, and knowledge-based scoring functions; I Score, which combines empirical and force-field scoring functions; SMoG2016, which combines knowledge-based and empirical scoring functions; and hybrid scoring, which combines force field and machine learning scoring functions. [20,21,22]

Protein Structure Prediction:

Essential molecules, proteins perform a role in many biological processes. Since a protein's three-dimensional structure plays a major role for determining its activity, protein structure prediction or modeling is essential. Moreover, the amino acid makeup of a protein determines its three-dimensional structure. Protein structure resolution experiments utilizing X-ray crystallography or NMR spectroscopy are costly, time- consuming, and intricate. As a result of developments in algorithms and computational tools, a model was constructed from amino acid sequences utilizing theoretical knowledge of protein structure, dynamics, and folding. Three types of protein structure prediction techniques can be distinguished (a) homology modeling; (b) threading; and (c) ab initio methods (de novo). Homology modeling, which uses a comparable known protein structure as a framework to predict an unknown structure, is the most efficient computer method for protein structure prediction [13,14,15].

Homology Modeling:

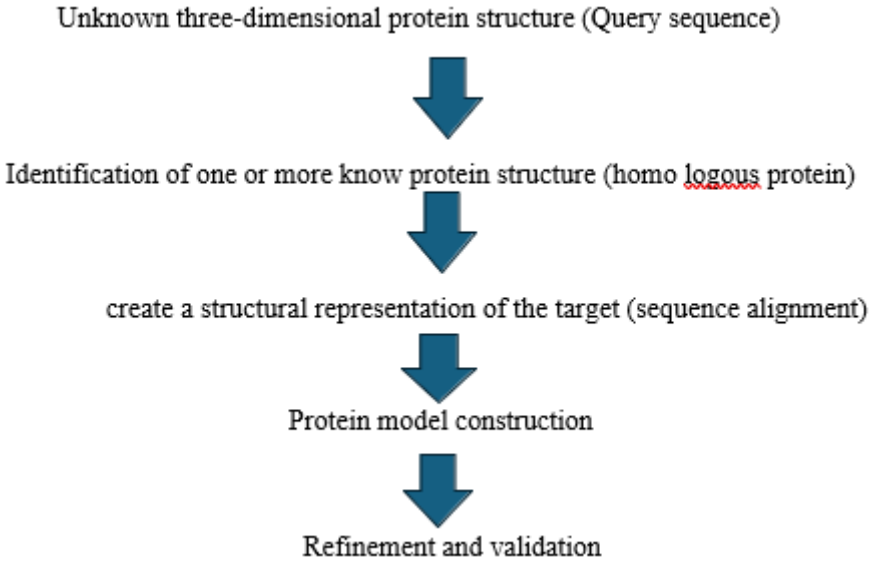

Figure no2: structure prediction by homology modeling [29,30]

It is necessary to precisely match the amino acid sequence of the protein with the appropriate amino acid residues of the folding motif in order to thread a sequence across a fold. This technique's primary objective is to identify the most likely fold from a given sequence or suitable sequences that could fold into a particular structure. The quantity of usable folds whose structures are precisely known down to the atomic level is what defines threading performance. The predictions were made using threading techniques, which employ methods for matching sequences with three-dimensional forms to ascertain the correct folding of a given sequence from a variety of options. Ab initio (de novo) protein structure prediction is a method for assessing the 3D structure of a protein when there isn't an experimentally determined structure of a similar or homologous protein. With this approach, the energy function directs the formation of the protein structure. First-principles physics and chemistry rules, along with the idea that a protein's natural structure always stays at the lowest energy level, provide the foundation of ab initio (from scratch) approaches. Nevertheless, ab initio modeling has little precision and typically only works with small proteins (120 residues).

Table No.2 The most well-known homology modeling methods used in drug development are summarized below [13,14,15].

Sr no	Name	Application
1	prime	Homology modeling evaluation and refining of the produced model using the energy function

2	LOMETS	Tertiary structure prediction with a local meta threading server
3	BHAGEERATHH	Methods of ab initio folding and homology are com bind
4	SWlSS-MODEL	Segment assembly / Local similarity
5	ROBETTA	Rosetta homology modeling and fragment assembly from scratch with Ginzu domain prediction
6	ESyPred3D	3D modeling template identification and alignment
7	Raptor X	Protein 3Dmodeling remote homology discovery and binding site prediction
8	Fold X	Protein design and energy calculations
9	HH-suite	Template detection alignment 3D modeling
10	l-TASSER	Reassembling fragment structure via threading

Binding Site and Cavity Prediction:

One of the key components of a protein target's three-dimensional structure is the binding the user must concentrate on the area of the protein structure where a ligand can bind and occupy its optimal position while conducting a protein-ligand interaction analysis. The investigation of protein-ligand complexes found in structural databases and published literature are two sources of binding site knowledge. The cavity prediction algorithms can therefore give the likely location of the binding site in a protein because the binding site residues are in the protein's cavity. Numerous cavity prediction methods, including CastP, Q-Site Finder, and COACH, have been developed using geometric, evolutionary, or energy-based methodologies. Evolutionarily related proteins are thought to share a binding site because it has been conserved throughout evolution, according to an evolutionary-based approach. While energy-based approaches use a probe to identify the best binding areas on proteins, geometry-based approaches rely on properties including shape, hydrophobic surface, compactness, and charged surface residues [31,32,33].

Pharmacophore Models:

Depending on the input data that is utilized in model development, there are two methods for creating pharmacophore models: "structure-based" and "ligand-based" pharmacophore modeling. The structure-based method finds compounds that may be utilized as drugs by using the structural information of the target proteins, such as enzymes or receptors. The ligand-based approach, on the other hand, relies solely on the physicochemical properties of known ligand molecules for drug development, which includes developing 3D pharmacophore models and modeling quantitative structure-activity relationships (QSAR) or quantitative structure-property relationships (QSPR). Several factors, including data availability, data quality, computational resources, and the planned use of the developed pharmacophore models, influence which approach is best to employ.

Structure-Based, Pharmacophor Modelling:

The 3D structure of a macromolecule target is a necessary precondition for producing a structure-based pharmacophore, which is why the structure-based method got its name. The creation or discovery of new medications can benefit greatly from the important atomic-level features provided by a protein's three-dimensional structure. As previously stated, a pharmacophore is an abstract representation that depicts the stereo-electronic features of a ligand that render a ligand bioactive toward a particular target. This kind of data can be taken from the protein target's three-dimensional structure in either its holo or apo form. The foundation of pharmacophoric modeling is the idea that biological activity on the same target results from shared chemical functions and a comparable spatial arrangement. The pharmacophoric model uses geometric concepts like spheres, planes, and vectors to depict the chemical properties of a molecule that can interact with its ligand. Hydrophobic regions (H), positively and negatively ionizable groups (PI/NI), aromatic groups (AR), hydrogen bond acceptors (HBAs), hydrogen bond donors (HBDs), and metal coordinating areas are the most significant pharmacophoric feature types. To depict the dimensions and form of the binding pocket, extra size limitations in the form of shapes or exclusion volumes (XVOL)-forbidden areas can be introduced. The models themselves can be beneficial

tools for identifying commonalities between molecules since they concentrate on chemical functions rather than actual atoms. Chemically different compounds can cause comparable biological outcomes because pharmacophore activity is independent of the scaffold. To select molecules of interest for the future vs. or in the field of chemo-informatics, these models can be employed as a query to search the huge libraries of chemicals available on the computational platform. [34,35]

De Novo Methodology for Drug Design:

De novo drug design creates novel chemical compounds based merely on a biological target (a receptor) or the target's known active binders (ligands with favorable binding or inhibitory activities). However, the quality of a homology model is dictated on template structure and sequence similarity. De novo drug design is a drug development technique that includes generating an entirely new chemical compound with a specific biological function. A variety of computational techniques and methodologies are used in the process of creating and improving new pharmacological candidates. One of the key components of this strategy is the application of methodologies like linking and development approaches. Linking techniques combine existing scaffolds or fragments to form new molecules. This technique selects scaffolds or fragments based on their demonstrated efficacy against the intended target. The way the fragments are joined maximizes the activity of the final molecule. On the other hand, growth processes start with a tiny precursor unit and work their way up to a molecule. This method is commonly used when there are no preexisting fragments that can serve as the basis for design. The development process is guided by the target's known characteristics as well as the physical and chemical characteristics necessary for optimal activity. Growth and coupling strategies are essential for de novo drug design success. They offer a methodical approach to the investigation of chemical space and the discovery of molecules with the best activity in relation to a certain goal. The possibility of creating novel medications by applying these techniques grows as long as computational technology advancements continue. Web servers and software used in de novo drug design follow the conventional process. The preparation of molecules (sampling), assessment of the produced compounds, and characterization of the receptor or ligand's active site are the main components of de novo drug design. pharmaceutical pharmacophore using modeling Ligand-based and structure-based approaches are the two main methods used in de novo drug design. Nuclear magnetic resonance (NMR), electron microscopy, or x-ray crystallography are frequently used to determine a receptor's three-dimensional structure. When developing new medications, homology modeling can be used to determine the best configuration while taking the receptor's structure into account [36].

Molecular Dynamics Simulation:

Analyzing the motion of atoms and molecules across a conformal space is the main goal of molecular dynamics (MD) simulation, a computer-based simulation technique. MD modeling can be used to verify the stability of protein-ligand complexes in a particular artificial environment. Therefore, to investigate the steady-state nature and conformal stability of the protein-ligand complexes, researchers performed a 50 ns MD simulation. By combining the force on the system and the speeds of the atoms over time.MD use Newton's second rulea of motion to predict new molecular system configurations. By reducing the likelihood that a molecular system get stuck in a local lowest energy region during a simulation, MD simulation approaches enable thorough conformational space sampling. As a result, when it comes to drug design employing in silico techniques, molecular dynamics simulation remains superior, faster, more logical, and more widely available. Furthermore, every atom's position and velocity inside the system are recorded at every instant in time, which is challenging to accomplish using any experimental method. It typically explains the atomic and molecular properties of the protein, interactions between drugs and targets, chemical solvation, and conformational changes that a receptor displays in various contexts. The primary drawback, though, is the simulated time, which for a big system is currently in the order of nanoseconds. [13,14,15]

Table no 3 The most frequently used software programs and force fields in molecular dynamics (MD) stimulations

Software	Processing Units	Force Field Employed	Characteristics	A
NAMD	CPU/GPU	CHARMM, AMBER	Free molecular dynamics software designed for high- performance simulations of large biomolecular systems	P, L, C, N, L, P
Abalone	GPU	AMBER, OPLS	Allows simulations of long time periods	P, C, N
CHARMM	CPU/GPU	CHARMM	Used for macromolecular simulations including energy minimization, MD, Monte Carlo simulations	P, L, C, N
AMBER	CPU/GPU	AMBER ff99SBildn, LIPID14	Package of programs for molecular dynamics simulations of proteins and nucleic acids	P, N, L-P
ACEMD	GPU	CHARM, AMBER, OPLS	The world's fastest MD engine for a single workstation	P, C, N
LAMMPS	GPU	CHARMM, AMBER, OPLS, GROMACS	Potential for soft/solid-state materials and coarse-grain simulations	P, L, C, N
TINKER	GPU	AMBER, CHARMM, OPLS, MMFF	General package for molecular mechanics and dynamics; features for biopolymers	P, N
DESMOND	GPU	AMBER, CHARMM, OPLS	Computes energies/forces for fixed-charge fields in biomolecular simulations	P, L, P
GROMACS	CPU/GPU	GROMOS 43A1-S3, MBER99SB-ILDN	Ease of use and exceptional performance	P, L, C, N, L, P
MOLDY	CPU	CHARMM	Enables simulations of millions of atoms using short- range potentials	P, N

P=protein, C= carbohydrates, L= lipids, A= most recently used tools, L-P= lipid protein complexes, N=nucleic acid, [37,38]

Molecular Mechanics/Qm Calculations:

A powerful computational method for analyzing chemical and biological processes at the atomic and molecular level is quantum mechanics/MM computations. The QM/MM method combines the speed of MD simulations with the precision of quantum mechanical computations. This technique is useful for analyzing chemical and biological systems in solutions or enzymes because it makes it possible to compute thermodynamic characteristics and evaluate reaction kinetics. Accurate assessments of reaction energy and reaction pathways are provided by QM/MM computations. Calculations in molecular mechanics and quantum mechanics often examine interactions between and within molecules in a system's static state without accounting for solvation. One of the main benefits of the QM/MM approach is its ability to simulate better systems than those that can only be simulated using quantum mechanical ab initio techniques. [36]

Ligand Based Approach In Drug Design

The targeted protein's 3D structure is unknown, but the ligands that bind to the desired location are known. This can be used to create molecules with all the necessary structural characteristics for binding to a target site or pharmacophore models. These methods are often quantitative structure activity relationships (QSARs) and pharmacophore-based approaches. Therefore, it is presumed that substances with comparable structures also have similar biological actions. engagement with the target protein An indirect method of drug desing known as ligand based drug design gives information on other compounds that bind to the intended biological target. A pharmacophore model that specifies the bare

minimum of required structural features may be derived using these. A chemical need to have certain properties in order to impede the target medicine. Knowing what binds to a biological target can help build a model of it. These models can then be utilized to create new molecular entities that interact with the target. QSAR, from which it is possible to infer a correlation between the computed properties of compounds and their experimentally determined biological activity. Using these QSAR relationships, we can predict how active a new analogue will be [23,24,25].

Ligand Based Virtual Screening:

Ligand-based virtual screening approaches identify and optimize leads by using the information found in known active ligands rather than the structure of a target protein. When the target protein's three-dimensional structure is unavailable, ligand-based techniques are the sole option. As an instance, this is usually the case with protein structures resolved in the apo form or targets of G-protein-coupled receptors (GPCRs). In actuality, one frequently learns that a group of ligands are active against the target of interest without knowing the protein structure of the target. Therefore, ligand-based virtual approaches—that is, identifying new ligands by comparing candidate ligands to existing active compounds—can be used. To choose a query for virtual screening or alignment in ligand-based design, the known active chemicals are gathered using a ligand-based approach. A strong similarity metric and a trustworthy scoring system are two crucial components of a ligand-based computational approach. Furthermore, a huge number of possible ligands should be screened quickly and accurately using the computational method. In addition to swiftly identifying a limited number of active compounds from a library that contains a large number of inactive compounds, a scoring approach for ligand-based screening should be able to distinguish active compounds from inactive ones during the ranking phase. [16,17,18]

Pharmacophore Based Modeling:

A pharmacophore is a molecular frame that explains the essential characteristics that give a molecule its biological activity. To improve knowledge of ligand-protein interactions, pharmacophore models are created. They can be used to find novel compounds that meet pharmacophore standards and are therefore anticipated to be active. If the target structure is unavailable, pharmacophore models can be constructed utilizing the structural details of the active ligands that bind to the target. This method is referred to as ligand-based pharmacophore modeling. Pharmacophore models can be constructed utilizing the target's structural characteristics when the target's structure is known. This method is referred to as structure- based pharmacophore modeling. The pattern of characteristics known as pharmacophore is what gives a substance its biological activity. This demonstrates that characteristics rather than chemical groups are the main focus of the pharmacophore idea. A pharmacophore pattern can be created from any atom or set of atoms in a molecule that exhibits characteristics linked to molecular recognition. patterns of molecular pharmacophores can be donors of hydrogen bonds (HBD), Hydrogen bond acceptors (HBA), hydrophobic characteristics, aromatic rings, positive and negative characteristics, and their combinations A variety of pharmacophore modeling tools are available. Software for creating pharmacophore models includes HipHop, Hypo Gen, Pharmer, PHASE, GASP, Pharma Gist, Pharm Mapper, MOE, Ligand Scout, and GALAHAD. With the use of such softwares, pharmacophore modeling has been employed at the various stages of the drug discovery process Virtual screening, ligand profiling, drug target fishing, docking and ADMET (absorption, distribution, metabolism, excretion, toxicity) prediction are among its popular application areas In the drug development process, two main pharmacophore modeling techniques are employed: two types of pharmacophore modeling: ligand-based and structure-based. A collection of active ligands is used to design novel ligands in the ligand-based pharmacophore modeling approach. This technique is employed if the target structure is not available. Similarly, when the target protein's structure is known, the structure-based pharmacophore method is used. The first active ligands in ligand-based pharmacophore modeling are found through database searches or the available literature. A training set and a test set are created from the data set. The training set ligands are then subjected to feature analysis. The alignment of the active ligands allows for the detection of shared characteristics. The creation of

pharmacophore models and their ranking are the following stages. Finally, pharmacophore model validation is carried out, and based on the outcomes, the optimal pharmacophore model is chosen. The initial stage in structure-based pharmacophore modeling is to choose and prepare the target protein structure. Predicting the binding location is the second stage. The binding site amino acids' complementary the chemical makeup and their configurations are then determined through meticulous analysis. Following this, the pharmacophore features are produced, which ought to be enhanced by the modified tools in the programs used. Lastly, the key pharmacophore characteristics that cause the activity are chosen (7). The software applications Ligand Scout (26), MOE (27), Pocket v2 (28) and Snooker (29) are frequently used for structure-based pharmacophore modeling. Pharmacophore modeling also uses a variety of servers and software programs. [39,40]

Table 4. overview of the most popular pharmacophore modeling software used in the process of designing drugs with computers. [13]

Sr No	Program	Application
1	Open3DQSAR	Exploration of pharmacophores using high-throughput chemometric analysis
2	Align-it	Pharmacophore alignment
3	Catalyst	Pharmacophore modeling
4	MOE	Pharmacophore modeling
5	Ligand Scout	Pharmacophore modeling
6	Phase	Pharmacophore modeling
7	Quasi	Pharmacophore modeling
8	Pharma gist	A website for the discovery of ligand-based pharmacophores
9	Pharmer	Pharmacophore search
10	FLAP	The fingerprints are characterized by pharmacophoric properties

Quantitative Structure-Activity Relationships (QSARs)Modeling:

Relationship between quantitative structure and activity (QSAR) The term "quantitative structure-property relationship" (QSPR) refers to the broad applicability of the QSAR concept, which has primarily been utilized in drug discovery and development to correlate molecular information with biological activities as well as other physicochemical properties. Typical molecular parameters that take into consideration topology, steric effects, hydrophobicity, and electrical characteristics can be ascertained theoretically by computational chemistry or experimentally through testing. [29,30] Utilizing a QSAR model constructed from an inventory of known ligands, the QSAR techniques are applied in CADD to estimate a compound's biological activity. The QSAR model establishes a connection between a group of chemicals' physicochemical characteristics and biological activity. A set of ligands (training set: 80%) that are effective against a specific target protein are used to create QSAR models. A test set (20% data) is used for internal validation of the QSAR model, and data from additional experimental sources may also be used for external validation. The biological activity of a new compound can be predicted using a variety of ID, 2D, 3D, etc. descriptors, including spatial, electronic, and structural properties, among many others. where as some descriptors, among them the number of H-bond accepters, can be calculated more easily using molecular connectivity, other descriptors may require more specialized measurements. Using quantum chemical equations, it is crucial to get descriptors that can be related to the chemical reactivity of pharmaceuticals, such as the lowest unoccupied molecular orbital energy and the highest occupied molecular orbital energy. To determine a connection between biological activity and structural characteristics, methods such as clustering, artificial neural networks, regression, and multivariate data analysis (principal component analysis, partial least square) are employed. Significant changes in biological activity are caused by slight variations in structural properties. In addition to reducing the amount of chemicals needed for in vivo research, QSAR enables a wide range of chemical products to

be prioritized as an in-silico technique based on their desired biological activity. Additionally, QSAR analysis indicates the structural and chemical modifications needed in a lead molecule to maximize its drug-suitable activity. Several tools are utilized to develop the QSAR analysis, including Gusar, Bio PPSy, Chem Des, and PHASE. [11] Nevertheless, QSAR modeling has many drawbacks. For instance, if there are few molecules in the training set, the data might not fully represent all of the characteristics, making it impossible to predict which compounds will be the most active [13,14,15].

Table no 5. Techniques and mathematical equations employed inQSAR modeling and drug desing[13]

Sr No	Equation	Techniques	Activity
1	nonlinear	Random forest	A better and more reliable estimate
2	linear	K-nearest neighbor	Simple
3	linear	Multiple linear regression	Simple
4	nonlinear	Artificial neural network	Works well with nonlinear data
5	linear	Partial least squares	Performs effectively on data including a big dataset
6	nonlinear	Decision tree	Extremely interpretable
7	nonlinear	Support vector machine	A most effective approach for classification and regression

Table no 6. QSAR TOOLS [41]

Tools	Description	Availability
Pharma QSAR	Pharma QSAR is a 3D(QSAR) software package that builds statistical models (COMFA, COMSIA and Hyphar) based on data obtained from experimental assays	COMMERCIAL
OECD QSAR Toolbox	The OECD QSAR toolbox is a free software and assessing chemical substance that uses computational methods as an alternative to animal testing	FREE
CORAL	QSPR/QSAR analysis for substance represented by simplified molecular input line entry system (SMILES)by the monte Carlo method	FREE
Auto QSAR	Auto QSAR automates the creation of high quality predictive QSAR models and makes their application trivially simple	COMMERCIAL
GUSAR	GUSAR software was developed to create QSAR/QSPR models on the basis of the appropriate training sets represented as SD file contained data about chemical structure and endpoint in quantitative terms	FREE

Similarity Search:

The fundamental idea behind fingerprint-based or similarity-based approaches is to choose new compounds for the target based on their physical and chemical resemblance to existing medications. Ligand similarity search techniques are straightforward yet efficient strategies predicated on the idea that compounds with similar structures typically have comparable binding characteristics Information regarding the actions of known binders of the target is not taken into consideration by these similarity metrics. Using similarity searches, a G-protein- coupled target GPR30 specific agonist that activates GPR30 was created. A 2D score plus a 3D structure similarity component made up the final similarity score that was applied [42].

In Silico Prediction of ADMET Properties and Drug Safety:

ADMET (absorption, distribution, metabolism, excretion, and toxicity) is linked to Lipinski's rule, which stipulates that an oral active medication typically has no more than one infraction of the following elements. A molecule can have no more than five hydrogen bond donors, or nitrogen-hydrogen and oxygen-hydrogen bonds combined.

- There are no more than ten hydrogen bond acceptors (all nitrogen or oxygen atoms) in a molecule.
- A molecule's molecular weight (MW) is less than 500 dal tons, or 800 grams.
- A molecule's octanol-water partition coefficient (Log P) cannot be more than 5.
- A molecule's polar surface area (PSA) cannot exceed 190 Å².
- The molar refractivity (MR) of a molecule typically falls between 40
- A molecule usually consists of around 20 to 70 atoms in total.
- There are no more than ten rotatable bonds in a molecule overall. [29,30]

Analysis of ADME Properties and Toxicity:

The website Swiss-ADME lets users draw their own ligand or drug molecule or add SMILES data from PubChem. It offers parameters like water solubility Log S, drug-likeness rules and lipophilicity (In pharmacology, the term "pharmacokinetic" describes the ADME/T between a pharmaceutical drug. Due to inadequate ADME features, about half of all drug candidates fail preclinical trial testing. Numerous promising therapeutic compounds are presently undergoing preclinical ADMET screening as a result of recent approaches and advancements in the drug detection procedure. Possible oral medication candidates' physicochemical properties are provided in the Swiss-ADME section based on five distinct rules established by Lipinski, Ghose, Veber, Egan, and Muegge. For compounds that were moderately soluble, the reference value of Log S was between 4 and -6, while for molecules that were extremely soluble, it was between 2 and -4. Based on the findings, every molecule is categorized as either extremely soluble or moderately soluble. According to all of these factors, 51320 is nearly a drug-like molecule [43].

Toxicity Analysis:

Before the medication candidate undergoes a clinical trial, toxicity assessment is a crucial step in better lead chemical selection. Any undesired or unfavorable chemicals or compounds that affect the human body or system are measured for toxicity. The conventional drug design method involves a number of animal tests to examine a substance's toxicity; these are costly, time-consuming, and involve ethical considerations. Comparing computer-aided toxicology tests to conventional These are efficient and cost-effective means of eliminating potentially hazardous chemicals and reducing the quantity of biological testing that is required. ecotoxicity, skin sensitivity, irritation, carcinogenicity, genotoxicity, and other to determine toxicity, endpoints are employed in either one or more dose trials. to ascertain how chemicals affect people, animals, plants, or the surroundings. inhibition of the potassium ion-related human ether-à-go-go-related gene (h ERG) For instance, channels can result in severe cardiac arrhythmia and are harmful to the heart. Therefore, early detection of putative h ERG inhibitors or non-inhibitors may be important. an important part in reducing cardiotoxicity [13,14].

Table 7. Overview of Widely Used ADME Analysis Tools in the Drug Desing Process [13]

Sr no	Program	Description
1	LIVERTOX	Hepatotoxicity prediction
2	ADME Tab	ADMET methodically by making use of the ADMET database
3	E Mol Tox	Molecular toxicity prediction
4	Pre ADMET	This web-based program determines the likelihood of carcinogenicity and toxic potency.
5	VNN	ADMET forecasts
6	Chem Tree	It is used to forecast ADME Tox characteristics
7	QikProp	used to predict traits connected to ADMET
8	Swiss ADME	Calculate physicochemical properties and forecast ADME
9	Meta base	is an inexpensive Excel-based radio analytical LIM for ADME/PK research.
10	TOPKAT	utilized in the prediction of toxicity
11	DSS Tox	It is a publicly accessible database of dispersed structure toxicity.

Machine Learning and Ai-Driven Approaches In CADD:

Machine Learning and Artificial Intelligence: The New Frontier

In Drug Discovery:

The technological renaissance that defines the 21st century has borne witness to the meteoric rise of Machine Learning (ML) and Artificial Intelligence (Al). These computational realms, known for their data-driven decision-making capabilities, have begun to significantly influence the sphere of Computer Aided Drug Design (CADD), reshaping the contours of drug discovery Machine Learning, a subset of Al, hinges on algorithms that can learn patterns from vast data sets without being explicitly programmed for specific tasks In drug discovery, Ml. has been instrumental in predicting molecular properties, understanding drug-receptor interactions, and forecasting biological responses based chemical structures. Techniques such as deep learning, which uses neural net-works modeled after the human brain, show immense potential for predicting complex drug-related outcomes with remarkable accuracy

Machine Learning Applications and Their Implications In CADD:

- Drug-Drug Interaction Prediction:

Understanding how a new drug may interact with other medications a patient may be taking is one of the challenges in drug discovery. Machine learning algorithms can analyze vast databases of known drug-drug interactions to forecast potentially hazardous combinations for novel compounds.

- **Repurposing Drugs:**

Finding novel therapeutic uses for already-approved medications is known as drug repurposing. Machine learning can find possible new targets for current drugs by evaluating large datasets, which reduces the time and expense involved in conventional drug discovery.

- **Drug Design Using Generative Adversarial Networks (GANs):**

A generator and a discriminator neural network are trained simultaneously in GANs, a type of Al. The discriminator assesses the molecular structures produced by the generator. As time passes, the generator gains proficiency in producing workable and possibly bioactive molecular structures that may be produced and examined in a laboratory.

- **Toxicology Prediction:**

Unexpected toxicity is one of the main causes of drug candidates' failure in clinical trials. By examining past data on drug-induced toxicities, machine learning models can assist in anticipating possible negative effects and weeding out potentially hazardous substances early in the discovery process. Furthermore, QSAR models correlate a molecule's structure with its probable toxicity by using descriptors such as molecular weight, lipophilicity, and electronic characteristics to predict toxicological effects. In computational toxicology, other characteristics including solubility, metabolic stability, and toxicophores, identification offer thorough insights that help prioritize chemicals for experimental testing and identify hazards early.More than only the acceptance of novel innovations is indicated by the incorporation of AI and ML into CADD. Using massive data and computing capacity to inform decisions at every stage of drug discovery, it signifies a paradigm shift from conventional hypothesis- driven research to data-driven discovery. Despite the fact that these technologies hold up the prospect of revolutionizing drug discovery, obstacles still exist. Concerns including data quality, AI model interpretability, and the requirement for experimental validation remain central to this integration. Essentially, a new era in drug discovery is ushered in by the combination of ML, AI, and CADD. a time of greater effectiveness, lower expenses, and quicker access to efficient treatments for individuals in need. [10,11,12]

Table no 8: Applications of Artificial Intelligence Tools In Drug Discovery [41]

Sr No	Name Of Tool	Description	Availability
1	Chemical VAE	Chemical design automation using a variational autoencoder (VAE))	Free
2	AphaFoll	A deep neural network is used to predict a protein's tertiary structure.	Free
3	Chemputer	Provides a step-by-step guide to creating a compound	Commercial
4	SMARTCyp	Site of metabolism prediction for CYP450 enzymes	Free
5	Hit Dexter	Use a combination of the AutoDock scoring approach and random forest (RF) to predict small molecule binding affinity with medicine.	Free
6	InnerOuterRNN	Neural networks with inner and outer recursion are used to forecast biological, chemical, and physical characteristics.	Free
7	DeltaVina	Use a combination of the AutoDock scoring approach and random forest (RF) to predict small molecule binding affinity with medicine.	Free
8	NNScore	Using a neural network-based scoring function, predict protein-ligand interaction affinity.	Free
9	Junction Tree VAE	To estimate small molecule binding affinity with medications, use the AutoDock score approach with random forest (RF).	Free
10	ORGANIC	ML algorithm for de novo design of organic molecules and polymers	Free
11	OML	Quantum machine learning Python toolki	Free
12	Open Drug Discovery Toolkit (ODDT's)	Using the random forest score (RF)- Score and the NNScore, a chemoinformatics pipeline has been developed	Free
13	PPB2	Using the nearest neighbor and machine learning algorithms, predict the target of the query molecule	Free
14	DIA-NN	Tool for proteomic data processing	Free
15	REINVENT	RNN (recurrent neural network) and RL (recurrent learning) are used to create a Commercial new molecule (reinforcement learning	Commercial
16	XenoSite	Predictor of Metabolism and reactivity of small molecules	Free

Table no 9: overview Deep learning tools used in drugs discovery [41]

Sr no	Name of tool	Description
1	Deep docking	A unique deep learning framework that is suitable for docking billions of molecule structure in a speedy and precise manner
2	Gluon	It is a deep learning library developed by AWS and Microsoft that enables developers to create train and deploy machine learning models in the cloud
3	Scikit learn	It is the most helpful library for machine learning in python, the sklearn package includes several useful methods for machine learning and statistical modeling such as classification regression clustering and dimensionality reduction
4	MX Net	It is a scalable deep learning framework that works with several deep learning models, including convolutional neural networks (CNNs) and long short-term memory networks (LSTMs).
5	Tensor flow	It particularly focuses on artificial neural networks as well as interference of deep neural networks.
6	Py torch	A machine learning framework that speeds up the transition from research prototyping to production deployment
7	Potential Net	Use a graph convolutional neural network, to predict binding affinity (CNN)

8	Deep Chem	For compound identification, an open-source Python library employs a deep learning method.
9	Deep Neural Net-QSAR	Using a hierarchical deep neural network, predict molecular activity (DNN)
10	Deep Tox	Using a deep learning algorithm, predict the toxicity of chemical substances
11	Tox_(R)CNN	A deep CVNN approach assessed the cytotoxicity of medicines
12	Convqsarfast	CNN approach to predict molecular characteristics
13	Neural graph fingerprint	Using CNN, predicts the properties of new compounds
14	PADME	Uses feed forward neural networks to predict drug target interactions

CONCLUSION:

The combination of sustainability principles, green chemistry, and advanced computational approaches is revolutionizing the modern pharmaceutical environment by allowing for cleaner, faster, and more efficient drug development. Sustainable manufacturing approaches reduce environmental impact while increasing production, whereas green chemistry reduces hazardous waste, optimizes resource utilization, and promotes safer synthesis routes. Simultaneously, Computer-Aided Drug Design (CADD) tools such as virtual screening, molecular docking, pharmacophore modeling, QSAR analysis, and molecular dynamics considerably speed up lead identification and optimization while requiring fewer experiments. The emergence of machine learning and artificial intelligence strengthens this ecosystem by increasing prediction accuracy, allowing for medication repurposing, anticipating toxicity, and building novel compounds using generative models. Collectively, these breakthroughs represent a paradigm shift toward eco-friendly, data-driven, and highly efficient drug development, opening the way for a future in which therapeutic innovation and environmental stewardship coexist.

REFERENCES

1. Ahmed SA, Pinjari NS. A Review of Sustainable Paradigms in Contemporary Drug Discovery and Development. Journal of Pharma Insights and Research. 2025 Apr 5;3(2):001-18.
2. Yingngam B, Aslam MS, Haghi AK, editors. Sustainable Pharmaceutical Product Development and Optimization Processes: From Eco-Design to Supply Chain Integrity. Springer Nature; 2025 Jan 30.
3. Verma[1] N, Kumar A, Hussain S, Fatima SK. Healing the world: Pharmaceutical sustainability initiatives.
4. Ogbuagu OO, Mbata AO, Balogun OD, Oladapo O, Ojo OO, Muonde MU. Sustainable pharmaceutical supply chains: Green chemistry approaches to drug production and distribution. IRE Journals. 2024 Oct;8(4):761-7.
5. Msingh R, Pramanik R, Hazra S, Uni J. Role of green chemistry in pharmaceutical industry: A review. J. Univ. Shanghai Sci. Technol. 2021; 23:291-9.
6. Koenig SG, Bee C, Borovika A, Briddell C, Colberg J, Humphrey GR, Kopach ME, Martinez I, Nambiar S, Plummer SV, Ribe SD. A green chemistry continuum for a robust and sustainable active pharmaceutical ingredient supply chain. ACS Sustainable Chemistry & Engineering. 2019 Sep 19;7(20):16937-51
7. Stefanache A, Marcinschi A, Marin GA, Mitran AM, Lungu II, Miftode AM, Crivoi F, Lacatusu D, Baican M, Cioanca O, Hancianu M. Green Chemistry Approaches in Pharmaceutical Synthesis: Sustainable Methods for Drug Development. AppliedChem. 2025 Jun 17;5(2):13.
8. Kumar R, Maurya A. Green Chemistry Techniques for Sustainable Pharmaceutical Synthesis. Journal of Drug Discovery and Health Sciences. 2024 Dec 30;1(04):187-200.
9. Zuin VG, Eilks I, Elschami M, Kümmerer K. Education in green chemistry and in sustainable chemistry: perspectives towards sustainability. Green Chemistry. 2021;23(4):1594-608.
10. Niazi SK, Mariam Z. Computer-aided drug design and drug discovery: a prospective analysis. Pharmaceuticals. 2023 Dec 22;17(1):22.

11. Nascimento IJ, de Aquino TM, da Silva-Júnior EF. The new era of drug discovery: The power of computer-aided drug design (CADD). Letters in Drug Design & Discovery. 2022 Nov 1;19(11):951-5.
12. Ece A. Computer-aided drug design. BMC chemistry. 2023 Mar 24;17(1):26.
13. Hasan MR, Alsaiari AA, Fakhurji BZ, Molla MH, Asseri AH, Sumon MA, Park MN, Ahammad F, Kim B. Application of mathematical modeling and computational tools in the modern drug design and development process. Molecules. 2022 Jun 29;27(13):4169.
14. Balasubramanian K. Mathematical and computational techniques for drug discovery: promises and developments. Current topics in medicinal chemistry. 2018 Dec 1;18(32):2774-99.
15. Makhouri FR, Ghasemi JB. Combating diseases with computational strategies used for drug design and discovery. Current Topics in Medicinal Chemistry. 2018 Dec 1;18(32):2743-73.
16. Hamza A, Wei NN, Zhan CG. Ligand-based virtual screening approach using a new scoring 6 function. Journal of chemical information and modeling. 2012 Apr 23;52(4):963-74.
17. Hamza A, Wei NN, Hao C, Xiu Z, Zhan CG. A novel and efficient ligand-based virtual screening approach using the HWZ scoring function and an enhanced shape-density model. Journal of Biomolecular Structure and Dynamics. 2013 Nov 1;31(11):1236-50.
18. Ye WL, Shen C, Xiong GL, Ding JJ, Lu AP, Hou TJ, Cao DS. Improving docking-based virtual screening ability by integrating multiple energy auxiliary terms from molecular docking scoring. Journal of Chemical Information and Modeling. 2020 Apr 30;60(9):4216-30.
19. Andricopulo AD, Guido RV, Oliva G. Virtual screening and its integration with modern drug design technologies. Current medicinal chemistry. 2008 Jan 1;15(1):37-46.
20. Maia EH, Assis LC, De Oliveira TA, Da Silva AM, Taranto AG. Structure-based virtual screening: from classical to artificial intelligence. Frontiers in chemistry. 2020 Apr 28; 8:343.
21. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH. Structure-based virtual screening for drug discovery: a problem-centric review. The AAPS journal. 2012 Mar;14(1):133-41.
22. Li Q, Shah S. Structure-based virtual screening. InProtein bioinformatics: from protein modifications and networks to proteomics 2017 Feb 2 (pp. 111-124). New York, NY: Springer New York.
23. Chandershekar A, Bhaskar A. A review on computer aided drug design (CAAD) and it's implications in drug discovery and development process. International Journal of Health Care and Biological Sciences. 2020 Jun 30:27-33.
24. Jamkar VB, Bogar AA, Butle SR, Choudhari RC, Ramteke SP. Revolutionizing Drug Discovery: A Comprehensive Review of Computer-Aided Drug Design (CADD) Tools in Modern Drug Development.
25. Purohit D, Makhija M, Pandey P, Kumar S, Kumar S, Dutt R, Kaushik D, Kumar P, Kumar S. Role of computer-aided drug design in the discovery and development of new medicinal agents a review. Int J Pharm Sci. 2018:1405-5.
26. Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN. Virtual screening in drug discovery-a computational perspective. Current protein and peptide science. 2007 Aug 1;8(4):329-51.
27. Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. Current medicinal chemistry. 2013 Aug 1;20(23):2839-60.

Cite: Akash Kundekar*, Nilesh Chougule, Prasad Yadav, Green CADD: Sustainable Approaches in Drug Discovery, Int. J. Med. Pharm. Sci., 2026, 2 (9), 440-459. https://doi.org/10.5281/zenodo.22917842