

The Role of Computational Chemistry and Cheminformatics in Modern Herbicide Discovery: A Critical Review

Akansha Agrwal^a, Aman Sahu^b and Neakanshika Chadha^{c*}

^a *Department of Applied Science, Krishna Institute of Engineering & Technology
(KIET), Ghaziabad, Delhi-NCR, Uttar Pradesh, India-201206*

^b *Department of Chemistry, Lajpat Rai college, Sahibabad, Uttar Pradesh, India-201005*

^c*Department of Chemistry, Jaypee Institute of Information Technology, Noida-Sec 62.*

Email: neakanshika.chadha@mail.jiit.ac.in

Abstract

Herbicides are prevalent chemical agents that can eliminate unwanted vegetation including weeds and specific grasses, or inhibit their growth to ensure the increase in agricultural productivity. The majority of pharmaceuticals drugs and herbicides are typically identified through a timeconsuming and expensive process of trial and error, which involves the evaluation of the potency of numerous compounds against a target in vitro. The tools of computational chemistry and techniques of cheminformatics acts as powerful tools that fasten the identification, design, and optimization of novel herbicidal molecules with increased cost-effectiveness. This comprehensive review addresses the current trends and challenges within herbicidal research along with illustrative examples including application in classical, cheminformatics and computational methodologies for herbicide innovation.

A broad literature survey conducted on the published literature on summarizing the various computational techniques and cheminformatics tools used in herbicide research including molecular docking (MolDock), molecular dynamics (MD) simulations, ADMET (Absorption Distribution Metabolism Excretion Toxicity) prediction, and machine learning (ML)-based QSAR (Quantitative structure analysis relationship) to evaluate their efficiency and restrictions in herbicide designing. This review forms a bridge between various computational tools and techniques with the herbicide designing in laboratory and cover the active challenges originated.

Keywords: Computational chemistry, Cheminformatics, Herbicide discovery, Molecular modelling, QSAR, Virtual screening

1. Introduction

In order to cater the demand of ever-growing population, the agricultural system faces intense pressure to guarantee the provision of food and resources without deteriorating the environment and the sustainability of the ecosystem[1]. The development and production of effective and environmentally friendly pesticides have been a challenge to feed the increasing population of our planet. The development of pesticides with a lower level of toxicity that encounter or control the pest insects, fungi and weeds is crucial. Amongst, weeds is the greatest threat to crop productivity which resulted in high yield losses across the main food production regions of the world [2]. Modern weed management has historically been based on the use of herbicides, but in recent years, herbicides have been challenged by the development of their resistant-to-herbicides species. Hence, there is an urgent need to discover new strategies to evaluate the effects of pesticides on the environment, and to certain rules and regulations to have safer agrochemicals [3], [4]. Conventional methods of herbicides discovery uses high-end empirical screening with vigorous trial and error processes, which are indeed time consuming, expensive and in most cases, inefficient. In addition to this, there is a huge pool of the chemical database of potential bioactive molecules, where random experimental screening would be unfeasible. In this regard, computational chemistry and cheminformatics have been born as ground-breaking techniques that can be used to design herbicidal agents rationally and in data-oriented way. Such techniques are also used to predict biological activity, target interaction, physicochemical properties, environmental viability, and toxicity profile prior to synthesis and intensive testing [5], [6].

Computational chemistry gives understanding about the interaction at the molecular level of herbicides and their target using quantum mechanical methods, molecular docking, and molecular dynamics simulations. . These methods are used to clarify the binding processes, estimation of free energies, residual studies of the molecules responsible for the inhibition, and to obtain a optimized structure. . At the same time, cheminformatics makes use of large chemical and biological sets with the help of quantitative structure-activity relationship (QSAR) modelling, machine learning, virtual screening, and similarity analysis to select putative candidates among large chemical libraries[7], [8]. Combination of the computational approaches with experimental validation has played a significant role to speed up the discovery pipelines of herbicides. The new wave of designing strategies is becoming the interplay of target-based design, phenotypic screening , and AI-driven predictions that will help to identify new modes of action and resolve resistance[9], [10]. In addition, computational tools assist in creating environmental friendly herbicides through offering early determinacy on ecotoxicity, biodegradable, and off-target impact. Nonetheless, significant advancements have been achieved, such as the lack of high-quality data sets, lack of interpretability of the models, emergence of resistance, and the necessity of an inter-disciplinary cooperation between computational scientists, chemists, biologists, and agronomists. Thus, this review critically focusses on methodology, current status, challenges , and further outlooks. However, most studies indicate the potential offered by computational chemistry and cheminformatics in the discovery of herbicides, but gaps exist in the literature[11]. Most of the reviews consist mostly of descriptions of available tools, but do not critically examine their predictive power, reproducibility, or real-world applicability. Standardized datasets are rarely used to benchmark on, and it is left to determine which modeling strategies can be the most reliable. Such emerging fields like AI-based de novo design, systems-level modeling of plants and evolution

of resistance remain underexplored, and environmental risk, regulatory action and sustainability measures are seldom part of computation. Further, small datasets of proprietary nature restrict scalability and model validation to teams. Future studies need to be comparative, data-driven, and interdisciplinary, in order to relate more fully molecular modelling and field performance and sustainable development of herbicides [12], [13]. To cover these loopholes, this review will comprehensively evaluate the current advances of computational chemistry and cheminformatics in discovering herbicides in modern times beyond the traditional descriptive studies. To assess predictive reliability, repeatability and practical utility of computational methodologies, notably molecular modelling, QSAR, virtual screening, and multiscale simulations, we critically examine them for future accountability. Recent innovations regarding the adoption of artificial intelligence, data-driven design and systems-level target identification, resistance-aware modelling and lead optimization is well discussed. Furthermore, the concern to environmental safety, regulatory views, and sustainability measures in computational processes to streamline discovery process in order to match agrochemical needs in the real world is also explored. This cumulative effort will be focused on proposing a prospective framework that enables the rational, efficient and sustainable production of the next-generation herbicides.

2. Overview of Computational Chemistry in Herbicide Discovery

The field of computational chemistry was included as an essential part of herbicide discovery that facilitates the molecularistic knowledge of herbicide-target interactions, predictive computer modelling of physicochemical properties, and guiding the rational design of lead compounds. Contrary to the traditional methods of empiricism, the computational methods enable researchers to access large chemical space effectively with minimize experimental activity. Combination of quantum chemical, molecular docking, virtual screening, and MD simulations has greatly helped in accelerating identification and optimization of new herbicidal agents[14]. Figure 1. describes a combined in silico-driven model for the rational discovery of new herbicide candidates. This process starts with data resource gathering of chemical databases, genomics/proteomics data, and resistance data which will inform target identification and selection of ligands. A number of computational methods, including molecular docking, molecular dynamics simulations and QSAR modelling, are then used to gain an understanding of binding interactions, predict activity and help lead optimization and design based on potency and selectivity and safety. The method also uses resistance analysis, such as mutant modelling and study of the resistance to guarantee effectiveness against the emerging weed targets[15]. The top applications are then taken through virtual screening and in silico and macro validation of biology. This agile pipeline eventually provides the production of sustainable, environmentally safe, agrochemicals that are more effective and less detrimental to the environment.

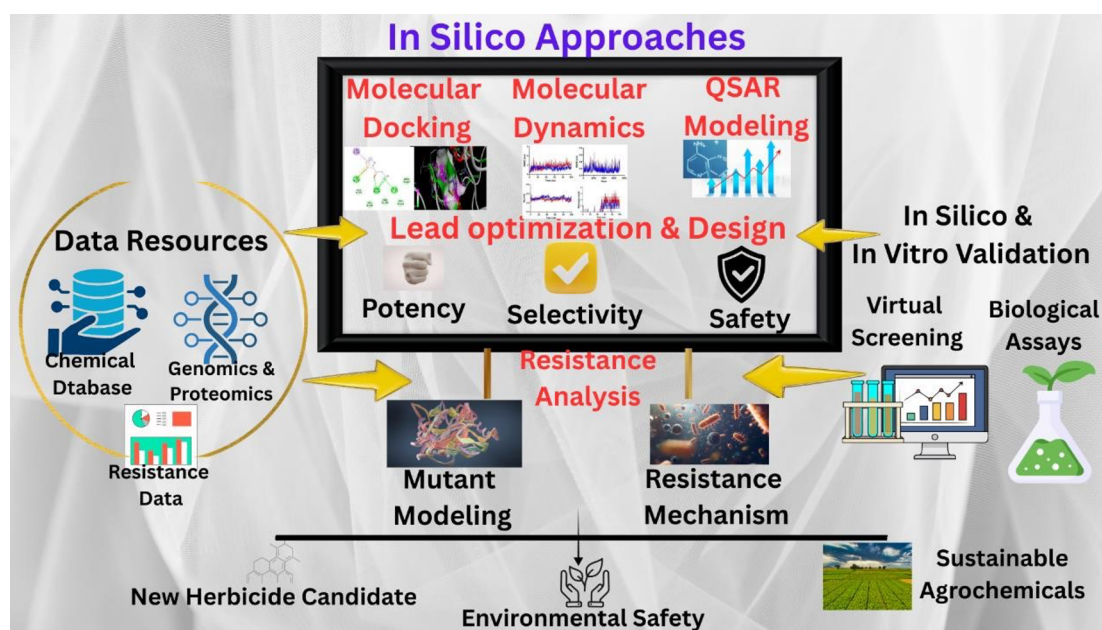


Figure 1. Workflow of Computational Chemistry in Herbicide Discovery

2.1. Quantum Chemical Calculations

Calculations at the quantum chemical level are essential in the current discovery of herbicides as they offer atomistic and electronic levels of understanding of the structure of the molecule, its reactivity and the mechanism of interaction with other molecules. The technique is founded on quantum mechanics and are employed to explain the behaviour of electrons in molecules that allows correct prediction of physicochemical and biological properties before the synthesis. Among the existing methods, the most popular is the Density Functional Theory (DFT) because of its balanced quality at computational scale[16]. In cases where extra precision is required, e.g. for predicting reaction mechanisms, ab initio methods, including Hartree-Fock (HF) and postHartree-Fock (e.g. MP2) can be applied. The use of quantum chemical calculations has been widely used during the process of discovering herbicides for refining the geometry of the molecules and for the identification of the most stable conformation. This is necessary since structural parameters such as bond lengths, bond angles and dihedral angles have a direct influence on the biological activity and binding affinity. Frontier molecular orbital analysis (HOMO-LUMO gap) provides information about the reactivity of the molecules on the molecular scale by considering donating and accepting nature of the electron of the compounds. A small HOMO-LUMO gap normally indicates increased chemical reactivity and the spatial distribution of the orbitals that helps in predicting potential site of interaction with target enzymes[17]. Molecular electrostatic potential mapping matches the distributions of electron density and electron deficiency, thus aiding in prediction of hydrogen bonding, electrostatic interaction and preferred binding interactions. It is also possible to study reaction pathways, activation energies and transition states using these computational methods, but it is important to note that they are critical in understanding degradation, plant metabolism and environmental persistence. Moreover, one can also calculate thermodynamic and electronic characteristics, including dipole moment,

polarizability, ionization potential, electron affinity, and global reactivity indices, and, in this way, provide useful correlations with solubility, permeability, and biological activity[18].

2.2. Molecular Docking (MolDock) and Virtual Screening (VS)

MolDock and VS are the core structure-based computational techniques utilized till date in herbicidal discovery in order to enable fast detection and optimization of bioactive compounds in reference to a given molecular name [19]. The binding configuration of a small ligand in the active site of a target protein like acetolactate synthase (ALS), 5-enolpyruvylshikimate3-phosphate synthase (EPSPS) or photosystem II, which are all target proteins of herbicides, are predicted using these strategies [20]. It measures the favoured binding orientation, pattern of interaction, and binding affinity of potential compounds of interest. It specifies the interactions in regards to hydrogen bonds, hydrophobic contacts, electrostatic interactions, and π - π stacking that are essential to the inhibitory activity. Docking results allow the rationalization of structural activity and to evaluate chemical structure changes to improve potency and selectivity. In addition, the comparative docking of wild -type enzymes with mutant enzymes also provides information on herbicide resistance modes.

Virtual screening (VS) is an expansion of docking which allows prioritization of potential candidates in large chemical libraries before synthesis and empirical validation [21]. The ligandbased screening paradigm, as well as the structure-based screening paradigm, is used to explore vast chemical space in an efficient manner. This goes a long way to reduce cost, time, and reliance on tedious laboratory screening programs. Although these advantages exist, docking and virtual screening are limited by the fidelity of scoring functions, the effects of protein flexibility and dependency on high-quality structural information. In this regard, such techniques are often combined with molecular dynamics simulations and experimental validation to enhance credibility. Overall, MolDock and VS have already become irreplaceable tools to hasten the process of rational herbicide design, discover new structures and overcome resistance barriers in the perspective of sustainable agrochemical development.

Table 1. highlights the MolDock and VC tools utilized in herbicide discovery including their type, features, advantages, and limitations. AutoDock and AutoDock Vina are open-source docking software used in academic studies since it is costless and reasonably accurate, and commercial applications like Glide, GOLD, and MOE offer more advanced scores, more accurate results, and workflows at the cost of a license. Dock can also still be used in mass screening but technical expertise is required. In virtual screening, many so-called user-friendly high-throughput screening engines like PyRx, or databases such as ZINC, make available large chemical libraries to select the candidates[22]. The workflow platforms like KNIME can be used to accomplish automation and hyphenate docking with the cheminformatics and machine learning models. In general, these tools are all in favour of efficient screening and ranking the potential herbicide before getting to the experimental validation.

Table 1. Molecular Docking and Virtual Screening Tools

S.N.	Tool	Type	Key Features	Advantages	Limitations	Reference
------	------	------	--------------	------------	-------------	-----------

1	AutoDock	Docking	Lamarckian genetic algorithm-based docking	Open-source, widely used, flexible ligand docking	Moderate speed, limited protein flexibility	[23]
2	AutoDock Vina	Docking	Improved scoring function and speed	Faster than AutoDock, good accuracy	Less customizable than AutoDock	[24]
3	Glide (Schrödinger)	Docking	HTVS, SP, XP precision modes	High accuracy, advanced scoring	Commercial, expensive	[25]
4	GOLD	Docking	Genetic algorithmbased docking	Good handling of ligand flexibility	Commercial license required	[26]
5	MOE Dock	Docking	Integrated docking and QSAR platform	User-friendly, multifunctional	Commercial software	[27]
6	DOCK	Docking	Grid-based docking	Suitable for large-scale screening	Requires expertise for setup	[28]
7	PyRx	Virtual Screening	GUI interface integrating AutoDock	Easy highthroughput screening	Limited advanced customization	[29]
8	ZINC Database	Virtual Screening Resource	Large library of purchasable compounds	Access to millions of compounds	Requires docking software integration	[30]

2.3. Molecular Dynamics Simulations (MDS)

The molecular dynamics simulations (MDS) provide a dynamic, time resolved view of herbicide target interaction that is over and above the snapshot perspectives revealed by molecular docking. MDS can be used to evaluate the stability of protein–ligand complexes and monitor their conformational dynamics and flexibility through continuous time simulation of atomic and molecular motion under physiological conditions. It can be of special interest in the discovery of herbicides, both to analyse the stability of the docked complexes w.r.t fine-tuning and discloses the

important hydrogen-bonding and hydrophobic interactions to determine solvents and temperatures effects on binding behaviour. Such computational experiments demonstrate that there are transient interactions and rearrangements of conformations that could have significant effects on inhibitory behaviour unlikely to be studied through rigid docking models.

MDS-based free energies, including MM-PBSA and MM-GBSA, provide more realistic prediction of binding affinity that adds to a dependable prioritization of potential candidates [31]. In addition, MDS allow one to study the mechanisms of resistance through comparison of interaction dynamics of wild-type and mutant target enzymes. In spite of their computer intensity, MD simulations together with quantum-chemical calculations and docking research increases prediction ability of structure-based herbicide design studies. Together, they offer a realistic setup of biological settings and have a critical role in sterilizing and optimizing possible herbicide candidates.

3. Role of Cheminformatics in Herbicide Design

Cheminformatics is a key field to the modern discovery of herbicides because it allows the analysis of data with its databases, the designation of the biological activity, and the rational investigation of chemical space [16]. As the use of chemical and biological databases rapidly expands, cheminformatics systems can help manage, interpret and model large data volumes in an efficient manner. These techniques are used to supplement computational chemistry techniques, dominant in statistical modelling, pattern recognition and predictive analytics, making them faster to identify promising herbicidal candidates at lowering costs of experiments.

3.1. Quantitative Structure–Activity Relationship (QSAR) Modelling

One of the most widely used cheminformatics approaches in the field of herbicide design is called quantitative structure-activity relationship (QSAR) modelling. The strategy defines mathematical correlations on the relationships between molecular descriptors (including hydrophobicity and electric values as well as steric factors and topological indexes) and the outcome biological activity[32]. With the creation of correlations between certain structural patterns and herbicidal activity, QSAR models allow the prediction of activity profiles of hitherto untapped compounds before their experimental preparation. Recent QSAR prototype combine advanced statistical analyses and machine-learning methods with the focus of improving predictive faithfulness. The models play key roles in optimization of leads, clarification of key functional groups that form the basis of bioactivity, and priorities of the candidate molecules which possess the desired physicochemical properties. However, effectiveness of QSAR predictions depends on a number of parameters such as integrity of the underlying dataset, judicious choice of descriptors and strict adherence to validation procedures. Table 2 gives an overview of a collection of computing tools required in the model development of QSAR and cheminformatics operations for predictive toxicology and molecular design. Tools like QSAR Toolbox (OECD)[33] and MOE offer one-stop shop environments to compute descriptors, predict toxicity and docking and QSAR toolbox gives regulatory acceptability, on the other hand, MOE offers a convenient multitask interface, both potentially limited in the flexibility of advanced machine learning or licensing[34]. Molecular descriptors and fingerprints descriptor-based methods, such as Dragon and PaDEL-Descriptor are capable of computing large numbers of molecular descriptors and fingerprints which are useful during model development, with PaDEL faces shortcomings at direct model development.

KNIME, WEKA, scikit-learn, and R (caret, random Forest) workflow and machine learning systems further extend the analysis of QSAR using customizable pipelines, statistical models as well as a variety of ML algorithms, with high flexibility and ability to achieve more analysis, but more technical skills and programming expertise are required[35]. Together, the tools provide a complete set of answers since they make it possible to generate descriptors, build models, validate and interpret these models in the context of contemporary data-driven agrochemical and pharmaceutical research.

Table 2. Quantitative Structure–Activity Relationship (QSAR) Modelling Tools

S.N.	Tool	Type	Key Features	Advantages	Limitations	Reference
1	QSAR Toolbox (OECD)	QSAR Platform	Toxicity prediction, readacross, regulatory focus	Free, regulatory acceptance	Limited advanced ML options	[36]
2	MOE	Integrated Modeling	Descriptor calculation, QSAR, docking integration	User-friendly, multifunctional	Commercial license required	[37]
3	Dragon	Descriptor Generator	Calculates >5000 molecular descriptors	Extensive descriptor library	Commercial, no built-in ML	[38]
4	PaDEL-Descriptor	Descriptor Tool	Calculates 1D, 2D, 3D descriptors & fingerprints	Free, widely used	Limited modeling capabilities	[39]
5	KNIME	Workflow Platform	QSAR modeling with ML integration	Highly customizable, integrates ML	Requires setup expertise	[40]
6	WEKA	Machine Learning Tool	Classification & regression algorithms	Free, multiple ML algorithms	Limited chemicals specific features	[41]
7	scikit-learn (Python)	ML Library	Regression, classification, validation tools	Flexible, powerful	Requires programming knowledge	[42]
8	R (caret, randomForest)	Statistical Platform	Advanced statistical modeling	Open-source, strong statistical control	Requires coding expertise	[43]

3.2. Database Mining and Chemical Space Exploration

The vast chemical space serves both as an opportunity as well as a challenge in the discovery of new herbicides. Database mining will enable the researchers to derive valuable information of

publicly and proprietary available chemical libraries. Similarity searching, substructure searching and clustering analysis are some of the techniques that are used to identify structurally related compounds and novel scaffolds. Computational tools of space exploration of chemicals allow visualization and mapping of a wide range of molecular structures, which allows identifying unexplored areas with possible herbicidal activity[44]. The combination of data mining with virtual screening techniques also helps to increase the possibility of finding new modes of action and addressing resistance concerns.

3.3. Machine Learning and Artificial Intelligence

The application of machine learning (ML) and artificial intelligence (AI) has dramatically changed the design of herbicides because it can model predictively using large and complex datasets[45], [46]. Random forests, support machines, neural networks, and deep learning models are algorithms used to forecast biological activity, toxicity, environmental fate, and physicochemical properties. AI techniques are used to assist the high-throughput screening, pattern recognition, and multiparameter optimization and makes it suitable, potent, selective and efficient to a large extent. Besides, the ML models also reveal unseen correlations in data that might not be necessarily observed with the standard statistical procedure. In spite of determinants, AI-based solutions need high-quality, high-curation data and validation measures to prevent overfitting and biasness. Combined with computational chemistry and experimental validation, cheminformatics and AI can present an effective framework in discovering herbicides in a sustainable and efficient way.

4. Integrating Computational Chemistry and Cheminformatics

Combining the cheminformatics and the computational chemistry will be an effective and synergistic strategy in herbicide discovery today. In this context, on one hand computational chemistry can give detailed mechanistic and molecular level understanding of herbicide-target interactions [47], on the other hand, cheminformatics can give access to mass analysis of data, predictive approaches and exploration of chemical space nucleus [48]. When combined with the other complementary methods, these enable the traditional trial-and-error method of herbicide development to be turned into a rational, data-driven, and efficient procedure.

Computational chemistry methods like quantum chemical methods, molecular docking, molecular dynamics methods provide high-resolution information in terms of molecular structure, electronic properties, interactions in the binding process and reaction mechanisms. These methods are however computationally extensive and are normally used on a few candidate molecules. By comparison, the cheminformatics methods such as QSAR modeling, database mining and machine learning enable a high-speed screening and prioritization of thousands to millions of compounds based on their predicted activities, toxicity, and physicochemical properties. A unified procedure will most often start with database mining and virtual screening to find possible herbicidal scaffolds. Prediction of biological activity and environmental behaviour is then followed using QSAR and machine learning models. Potential candidates are then subjected to molecular dynamics simulations and molecular docking as a measure of the binding stability and the mode of interaction at the target site. Quantum chemical calculations can also be used to obtain a finer

tuning of the molecular properties, optimize functional groups and evaluate reactivity or degradation reactions. At last, the chosen compounds are sent to the real experimental validation.

This multi-layered approach will improve prediction in accuracy, minimise the workload in experiments, and succeed in discovering herbicides with new mode of action and enhanced environmental safety. Notably, resistance management is also aided by the application of integration since comparative modelling of mutant target enzymes and wild ones is possible. Despite of huge success, there has been difficulties in harmonizing datasets, model interpretability, computational cost and reproducibility. It can be anticipated that in the future, automation will be implemented in pipelines, cloud computing, AI- driven multi-objective optimization, as well as closer integration between in silico predictions and high-throughput experimental platforms. In general, sustained and effective discovery of herbicides can be thoroughly addressed through the integration of both computational chemistry and cheminformatics, which will combine the fresh perspectives on this problem at a molecular level with massively predictive analytics to unlock the opportunities offered by the global agricultural context.

5. Major Developments in the field

Historical development of computational chemistry/ cheminformatics in the discovery of herbicides represents the gradual emergence of classical methods of correlation-based discovery to the modern and more sophisticated design-based technologies evident in artificial intelligence. Earlier in 1977, QSAR models that utilizes Hansch analysis, linear free-energy relationships, and the physicochemical descriptors including hydrophobicity (logP) were employed to predict the connection between molecular properties and herbicidal activity for the optimization of biological activity by Corwin Hansch and his co-workers. In 2001, a pivotal moment came with the revolution of structural biology that made it possible to visualise major targets of herbicides in detail including the targets ALS/AHAS, EPSPS as well as HPPD so that structure-based molecular docking and mechanistic insights into inhibition and resistance can be obtained. The study revealed that glyphosate establishes multiple H-bonds and electrostatic interactions with crucial amino acid residues that leads to the stabilization via shikimate pathway. In 2002, Growing interest in herbicide resistance and the need to find new active modes take advantage of integrated computational pipelines comprising popular QSAR and docking technique, especially in relation to HPPD and ALS inhibitors. Pang et al. reported the crystal structure of yeast acetohydroxyacid synthase (AHAS), a major enzyme involved in the biosynthesis of branched-chain amino acids and its catalytic role in the condensation reaction of pyruvate-derived substrates.

The growth of digital chemical data sets and better algorithms leads to the introduction of machine learning methods of chemical space analysis, activity prediction and compound prioritization. In 2010, Verma et al. succeeded by the creation of the three dimensional QSAR methods such as CoMFA, CoMSIA and pharmacophore modelling, which gave in-depth understanding with regard to the steric and electrostatic influences in activity. This study highlights the drawbacks of 3DQSAR such as alignment and sensitivity to conformational preferences, but also discussed that this computational method can be used in combination with molecular docking and other in silico approaches in medicinal chemistry. Dayan and his team discussed the current state of herbicide research and future opportunities in weed control. They review the progress in molecular biology, genomics, and high-throughput screening that facilitate the discovery of new targets, and the role

of natural products in the discovery of new herbicides. Tukur et al. performed molecular docking studies on a set of 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors with potential herbicidal properties. The team developed statistically robust QSAR models based on molecular descriptors of optimized molecules that shows a good fit between predicted and experimental inhibitory activities. In 2021, Oršolić et al. discussed the efficacy of ML for systematic understanding of herbicide chemical space, facilitate virtual screening and guide the rational design of novel agrochemicals with enhanced efficacy and safety. In 2022, Agrwal et al. reported the synthesis, molecular docking, and structure–activity relationship (SAR) analysis of a series of substituted dihydropyridine (DHP) derivatives as potential herbicides. In recent times, integrations of computational-experimental workflows have been used to integrate docking with ANN-QSAR, CoMFA, molecular dynamics simulations, ADMET prediction, and synthesis and bioassay validation to increase the predictive reliability. The present time is characterized by AI-based cheminformatics systems, deep learning, graph neural networks, and generative design systems that facilitate the stream of multi-target optimization and large scale virtual screening, which is a shift toward intelligent, data-driven, and highly automated herbicide discovery pipelines .

Table 3. Major Developments in the field

S.No	Major Development	Computational Approach	Target / Focus	Representative Reference
1	Quantitative structure–activity relationships (QSAR) introduced	Hansch analysis; linear free-energy relationships; hydrophobicity constants (logP)	General agrochemical optimization	[49]
2	3D-QSAR and steric/electrostatic field modeling	CoMFA; CoMSIA; pharmacophore mapping	Herbicide analog modeling	[50]
3	Structural biology revolution in herbicide targets	X-ray crystallography; structure-based modeling	ALS, EPSPS, HPPD	[51]
4	First high-resolution AHAS structural model	Crystal structure + docking	ALS/AHAS	[52]
5	Herbicide-binding sites revealed in plant AHAS	Structural elucidation; binding site mapping	ALS/AHAS	[53]
6	Modern discovery challenges & new MOAs	Review of computational strategies	Multiple targets	[54]
7	QSAR + Docking combined pipelines	Ligand-based + structure-based integration	HPPD inhibitors	[55]

8	Machine learning for chemical space analysis	ML classification; clustering	Herbicide chemical space	[56]
9	Integrated computational-experimental design	Docking; ANN-QSAR; CoMFA; synthesis & bioassay	AHAS inhibitors	[57]
10	Docking + Molecular Dynamics + ADMET	MD simulations; binding stability; toxicity screening	β -amino carbonyl herbicides	[58]
11	AI and cheminformatics platforms	Machine learning; data-driven discovery	Crop protection pipelines	[45]
12	Deep learning & generative design	Graph neural networks; generative models; AI-driven	Multi-target herbicide design	[59]
		virtual libraries		

6. Challenges and Future Directions

Although computation techniques has greatly improved, there are still a number of challenges in this field in new herbicide launch. This may includes poor access to quality datasets, overfitting of models, and non-interpretability of AI models that decreases predictive reliability. Additionally, premature arrangements and the computational expense of more sophisticated simulations in addition restrain accuracy. This is also accompanied by the demand of constant model inversion and safety assessment for evolving herbicide resistance and strict environmental regulations. Future directions are the formation of combined computational pipelines between AI, docking, molecular dynamics and quantum computations. Explainable AI, high-performance computing and multi-objective optimization will enhance efficiency and transparency. Increased interaction with the omics data and green chemistry is bound to be the force that will lead to safer, sustainable, and resistant-resilient herbicides.

7. Conclusion

Computational chemistry and cheminformatics have radically changed the contemporary herbicide discovery towards achieving a paradigm shift between the traditional trial and error experimentation to rational and data-oriented design. Such methods as quantum chemical calculations, molecular docking, molecular dynamics simulations, QSAR modeling, and machine learning allow screening potential herbicides quickly, optimizing, and scoring candidates just prior to synthesis. Combination of those strategies is able to increase the accuracy of predictions, decrease prices and development time and help to create more environmentally friendly and resistant-stressing herbicides. Despite the obstacles like quality of data and cost of computation and the interpretation of the model, there are still better developments in artificial intelligence and high-performance computing that would make this area stronger and stronger. Altogether, the combination of computational chemistry and cheminformatics represents a desalinating and

- [13] M. Quareshy, J. Prusinska, J. Li, and R. Napier, "A cheminformatics review of auxins as herbicides," *J. Exp. Bot.*, vol. 69, no. 2, pp. 265–275, Jan. 2018, doi: 10.1093/jxb/erx258.
- [14] R. Sooranna *et al.*, "Computational Workflow for Chemical Compound Analysis: From Structure Generation to Molecular Docking," *Sci. Pharm. 2026, Vol. 94, Page 9*, vol. 94, no. 1, p. 9, Jan. 2026, doi: 10.3390/scipharm94010009.
- [15] Q. Yu and S. B. Powles, "Resistance to AHAS inhibitor herbicides: Current understanding," *Pest Manag. Sci.*, vol. 70, no. 9, pp. 1340–1350, Sep. 2014, doi: 10.1002/ps.3710.
- [16] S. Lehoczki-Krsjak, M. Varga, Á. Szabó-Hevér, and Á. Mesterházy, "Translocation and degradation of tebuconazole and prothioconazole in wheat following fungicide treatment at flowering," *Pest Manag. Sci.*, vol. 69, no. 11, pp. 1216–1224, Nov. 2013, doi: 10.1002/ps.3486. [17] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules*. Oxford University Press; Clarendon Press, 1989. Accessed: Feb. 17, 2026. [Online]. Available: https://books.google.com/books/about/Density_Functional_Theory_of_Atoms_and_M.html?id=TYRMCAAQBAJ
- [18] P. Geerlings, F. De Proft, and W. Langenaeker, "Conceptual Density Functional Theory," *Chem. Rev.*, vol. 103, no. 5, pp. 1793–1873, 2003, doi: 10.1021/cr990029p.
- [19] H. Wang *et al.*, "Structure-Based Design, Virtual Screening, and Discovery of Novel Patulin Derivatives as Biogenic Photosystem II Inhibiting Herbicides," *Plants 2024, Vol. 13, Page 1710*, vol. 13, no. 12, p. 1710, Jun. 2024, doi: 10.3390/plants13121710.
- [20] S. O. Duke and F. E. Dayan, "Discovery of new herbicide modes of action with natural phytotoxins," in *ACS Symposium Series*, vol. 1204, American Chemical Society, 2015, pp. 79–92. doi: 10.1021/bk-2015-1204.ch007.
- [21] X. Y. Leng *et al.*, "Discovery of novel HPPD inhibitors: Virtual screening, molecular design, structure modification and biological evaluation," *Pestic. Biochem. Physiol.*, vol. 192, p. 105390, May 2023, doi: 10.1016/j.pestbp.2023.105390.
- [22] T. Sterling and J. J. Irwin, "ZINC 15 – Ligand Discovery for Everyone," *J. Chem. Inf. Model.*, vol. 55, no. 11, pp. 2324–2337, Nov. 2015, doi: 10.1021/acs.jcim.5b00559.
- [23] B. Ni, H. Wang, H. K. S. Khalaf, V. Blay, and D. R. Houston, "AutoDock-SS: AutoDock for Multiconformational Ligand-Based Virtual Screening," *J. Chem. Inf. Model.*, vol. 64, no. 9, pp. 3779–3789, May 2024, doi: 10.1021/acs.jcim.4c00136.
- [24] O. Trott and A. Olson, "Autodock vina: improving the speed and accuracy of docking," *J. Comput. Chem.*, vol. 31, no. 2, pp. 455–461, 2010, doi: 10.1002/jcc.21334.AutoDock.
- [25] R. A. Friesner *et al.*, "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy," *J. Med. Chem.*, vol. 47, no. 7, pp. 1739–1749, Mar. 2004, doi: 10.1021/JM0306430/SUPPL_FILE/JM0306430_S.PDF.
- [26] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, "Improved proteinligand docking using GOLD," *Proteins Struct. Funct. Genet.*, vol. 52, no. 4, pp. 609–623, Sep. 2003, doi: 10.1002/prot.10465.
- [27] Molecular Operating Environment (MOE), "Molecular Operating Environment (MOE) | MOEsaic | PSILO." Accessed: Nov. 23, 2025. [Online]. Available: <https://www.chemcomp.com/en/Products.htm>
- [28] W. J. Allen *et al.*, "DOCK 6: Impact of New Features and Current Docking Performance," *J. Comput. Chem.*, vol. 36, no. 15, p. 1132, Jun. 2015, doi: 10.1002/JCC.23905.
- [29] S. Dallakyan and A. J. Olson, "Small-Molecule Library Screening by Docking with PyRx," *Methods Mol. Biol.*, vol. 1263, pp. 243–250, 2015, doi: 10.1007/978-1-4939-2269-7_19.
- [30] B. I. Tingle *et al.*, "ZINC-22—A Free Multi-Billion-Scale Database of Tangible Compounds for Ligand Discovery," *J. Chem. Inf. Model.*, vol. 63, no. 4, pp. 1166–1176, Feb. 2023, doi: 10.1021/acs.jcim.2c01253.

- [31] S. Genheden and U. Ryde, "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities," *Expert Opin. Drug Discov.*, vol. 10, no. 5, pp. 449–461, May 2015, doi: 10.1517/17460441.2015.1032936.
- [32] Y. Dundar *et al.*, "Synthesis and Biological Evaluation of the Salicylamide and Salicylic Acid Derivatives as Anti-Estrogen Agents," *Med. Chem. (Los. Angeles)*., vol. 8, no. 3, pp. 481–490, Nov. 2012, doi: 10.2174/1573406411208030481.
- [33] D. Yordanova *et al.*, "Automated and standardized workflows in the OECD QSAR Toolbox," *Comput. Toxicol.*, vol. 10, pp. 89–104, May 2019, doi: 10.1016/j.comtox.2019.01.006.
- [34] E. Mombelli and P. Pandard, "Evaluation of the OECD QSAR toolbox automatic workflow for the prediction of the acute toxicity of organic chemicals to fathead minnow," *Regul. Toxicol. Pharmacol.*, vol. 122, p. 104893, Jun. 2021, doi: 10.1016/j.yrtph.2021.104893.
- [35] D. Fourches, E. Muratov, and A. Tropsha, "Trust, But Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research," *J. Chem. Inf. Model.*, vol. 50, no. 7, pp. 1189–1204, Jul. 2010, doi: 10.1021/ci100176x.
- [36] T. W. Schultz, R. Diderich, C. D. Kuseva, and O. G. Mekenyan, "The OECD QSAR Toolbox Starts Its Second Decade," in *Methods in Molecular Biology*, vol. 1800, Humana Press, New York, NY, 2018, pp. 55–77. doi: 10.1007/978-1-4939-7899-1_2.
- [37] F. Y. Alomari, A. A. Sharfalddin, M. H. Abdellattif, D. Domyati, A. S. Basaleh, and M. A. Hussien, "QSAR Modeling, Molecular Docking and Cytotoxic Evaluation for Novel Oxidovanadium(IV) Complexes as Colon Anticancer Agents," *Mol. 2022, Vol. 27, Page 649*, vol. 27, no. 3, p. 649, Jan. 2022, doi: 10.3390/molecules27030649.
- [38] A. Mauri, V. Consonni, M. Pavan, and R. Todeschini, "DRAGON software: An easy approach to molecular descriptor calculations," *Match*, vol. 56, no. 2, pp. 237–248, 2006.
- [39] C. W. Yap, "PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints," *J. Comput. Chem.*, vol. 32, no. 7, pp. 1466–1474, May 2011, doi: 10.1002/JCC.21707;JOURNAL:JOURNAL:1096987X;WGROU:STRING:PUBLICATION.
- [40] N. Tripathi, N. Bhardwaj, S. Kumar, and S. K. Jain, "A machine learning-based KNIME workflow to predict VEGFR-2 inhibitors," *Chem. Biol. Drug Des.*, vol. 102, no. 1, pp. 38–50, Jul. 2023, doi: 10.1111/cbdd.14250.
- [41] P. Kotak and H. Modi, "Enhancing the data mining tool WEKA," *Proc. 2020 Int. Conf. Comput. Commun. Secur. ICCCS 2020*, Oct. 2020, doi: 10.1109/ICCCS49678.2020.9276870.
- [42] F. Pedregosa FABIANPEDREGOSA *et al.*, "Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011, Accessed: Aug. 01, 2025. [Online]. Available: <http://scikit-learn.sourceforge.net>.
- [43] K. Siddiqui *et al.*, "Open Source Machine Learning Classifier Algorithm Using Random Forest in 'Caret' for R and Conventional Statistical Modeling: Preemptive Ivig on Day-1 in Pediatric PostHematopoietic Cell Transplantation Cytomegalovirus Reactivation," *Transplant. Cell. Ther.*, vol. 31, no. 2, p. S527, Feb. 2025, doi: 10.1016/j.jtct.2025.01.838.
- [44] D. Oršolić, V. Pehar, T. Šmuc, and V. Stepanić, "Comprehensive machine learning based study of the chemical space of herbicides," *Sci. Reports 2021 111*, vol. 11, no. 1, pp. 11479–, Jun. 2021, doi: 10.1038/s41598-021-90690-w.
- [45] Y. Djoumbou-Feunang *et al.*, "Cheminformatics and artificial intelligence for accelerating agrochemical discovery," *Front. Chem.*, vol. 11, p. 1292027, Nov. 2023, doi: 10.3389/fchem.2023.1292027.
- [46] S. Li, P. Zhang, R. Zhou, Y. Bu, and H. Zhang, "A machine learning-based QSAR model for predicting toxicity of commercial pesticides to *Eisenia fetida*," *Ecotoxicol. Environ. Saf.*, vol. 311, no. 1, p. 119804, Feb. 2026, doi: 10.1016/j.ecoenv.2026.119804.

- [47] G. Wu, Y. Chen, Y. Yao, W. Huang, and K. Wu, "Integrating network toxicology and molecular docking to uncover mechanisms of novel herbicide-induced neurodegeneration," *Pestic. Biochem. Physiol.*, vol. 216, p. 106821, Jan. 2026, doi: 10.1016/j.pestbp.2025.106821.
- [48] R. Parrondo-Pizarro, L. Menestrina, R. Garcia-Serna, A. Fernández-Torras, and J. Mestres, "Enhancing molecular property prediction through data integration and consistency assessment," *J. Cheminformatics* 2025 171, vol. 17, no. 1, pp. 163-, Oct. 2025, doi: 10.1186/s13321-025-01103-3.
- [49] C. Hansch, S. D. Rockwell, P. Y. Jow, A. Leo, and E. E. Steller, "Substituent constants for correlation analysis.," *J. Med. Chem.*, vol. 20, no. 2, pp. 304–306, Feb. 1977, doi: 10.1021/jm00212a024.
- [50] J. Verma, V. M. Khedkar, and E. C. Coutinho, "3D-QSAR in Drug Design - A Review," *Curr. Top. Med. Chem.*, vol. 10, no. 1, pp. 95–115, Jan. 2010, doi: 10.2174/156802610790232260.
- [51] E. Schönbrunn *et al.*, "Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 4, pp. 1376–1380, Feb. 2001, doi: 10.1073/pnas.98.4.1376.
- [52] S. S. Pang, R. G. Duggleby, and L. W. Guddat, "Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors," *J. Mol. Biol.*, vol. 317, no. 2, pp. 249–262, Mar. 2002, doi: 10.1006/jmbi.2001.5419.
- [53] J. A. McCourt, S. S. Pang, J. King-Scott, L. W. Guddat, and R. G. Duggleby, "Herbicide-binding sites revealed in the structure of plant acetohydroxyacid synthase," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 3, pp. 569–573, Jan. 2006, doi: 10.1073/pnas.0508701103.
- [54] F. E. Dayan, "Current Status and Future Prospects in Herbicide Discovery," *Plants* 2019, Vol. 8, Page 341, vol. 8, no. 9, p. 341, Sep. 2019, doi: 10.3390/plants8090341.
- [55] S. Tukur, G. A. Shallangwa, and A. Ibrahim, "Theoretical QSAR modelling and molecular docking studies of some 4-hydroxyphenylpyruvate dioxygenase (HPPD) enzyme inhibitors potentially used as herbicides," *Heliyon*, vol. 5, no. 11, p. e02859, Nov. 2019, doi: 10.1016/j.heliyon.2019.e02859.
- [56] Y. Wang, Y. Xiong, E. A. Lozano Garcia, Y. Wang, and C. J. Butch, "Drug Chemical Space as a Guide for New Herbicide Development: A Cheminformatic Analysis," *J. Agric. Food Chem.*, vol. 70, no. 31, pp. 9625-9636, Aug. 2022, doi: 10.1021/acs.jafc.2c01425.
- [57] A. Agrwal, A. Verma, N. Chantola, S. Verma, and V. Kasana, "Synthesis, molecular docking and extensive structure activity relationship of substituted DHP derivatives: a new class of herbicides," <https://doi.org/10.1080/03601234.2022.2062188>, vol. 57, no. 5, pp. 379–420, 2022, doi: 10.1080/03601234.2022.2062188.
- [58] S. Bhandari, A. Agrwal, V. Kasana, S. Tandon, Y. Boulaamane, and A. Maurady, "β-amino carbonyl derivatives: Synthesis, Molecular Docking, ADMET, Molecular Dynamic and Herbicidal studies.," *ChemistrySelect*, vol. 7, no. 48, p. e202201572, Dec. 2022, doi: <https://doi.org/10.1002/slct.202201572>.
- [59] S. O. Duke *et al.*, "New Approaches to Herbicide and Bioherbicide Discovery," *Weed Sci.*, vol. 72, no. 5, pp. 444–464, Sep. 2024, doi: 10.1017/wsc.2024.54.