Genome Publications
https://doi.org/10.61096/978-81-990998-7-6_5

Chapter 5

Higher-Dimensional QSAR: 2D, 3D, 4D, 5D and 6D Models for Molecular Flexibility

Saravanakumar A
Department of Pharmaceutical Biotechnology, Vellalar College of Pharmacy, Thindal, Erode, Tamil Nadu, India.
Kavin Kumar MC
Department of Pharmaceutical Chemistry, Vellalar College of Pharmacy, Thindal, Erode, Tamil Nadu, India.
Jeevananathan P
Department of Pharmaceutical Chemistry, Vellalar College of Pharmacy, Thindal, Erode, Tamil Nadu, India.
Monisha K
Department of Pharmaceutical Chemistry, Vellalar College of Pharmacy, Thindal, Erode, Tamil Nadu, India.

Abstract: Quantitative Structure Activity Relationship (QSAR) modelling has undergone a remarkable evolution from one-dimensional physicochemical correlations to higher-dimensional frameworks that integrate molecular geometry, conformational dynamics, solvation, and receptor flexibility. The transition from 2D to 6D QSAR represents a progressive refinement of how molecular structure is encoded, interpreted, and related to biological activity. Two-dimensional (2D) QSAR captures topological and physicochemical patterns; three-dimensional (3D) QSAR models, such as Comparative Molecular Field Analysis (Coma) and Comparative Molecular Similarity Indices Analysis (CoMSIA), incorporate steric and electrostatic fields; four-dimensional (4D) QSAR accounts for conformational ensembles and dynamic sampling; five-dimensional (5D) QSAR introduces environmental and induced-fit effects; and six-dimensional (6D) QSAR expands this framework by integrating receptor–ligand dynamics and solvent interactions in near-physiological contexts. This chapter provides an exhaustive account of these multidimensional QSAR models, explaining their theoretical foundations, computational workflows, and comparative merits. Emphasis is placed on methodological transparency, algorithmic advances, validation strategies, and real-world applications in drug discovery, such as anticancer, antiviral, and enzyme inhibitor design. Finally, the chapter explores how artificial intelligence (AI), molecular dynamics, and quantum mechanics are converging to define the next frontier of multidimensional QSAR.

Keywords: QSAR, molecular flexibility, multidimensional modelling, CoMFA, CoMSIA.

Citation: Saravanakumar A, Kavin Kumar MC, Jeevananathan P, Monisha K. Higher-Dimensional QSAR: 2D, 3D, 4D, 5D and 6D Models for Molecular Flexibility. *Comprehensive Approaches in Computer-Aided Drug Design: QSAR, Docking, Screening, Homology, Pharmacophore and AI-Driven Insights.* Genome Publication. 2025; Pp49-58. https://doi.org/10.61096/978-81-990998-7-6_5

5.0 INTRODUCTION

The development of QSAR has been central to computer-aided drug design (CADD), embodying the principle that molecular structure governs biological function. The earliest QSAR models, introduced by Hansch and Fujita in the 1960s, correlated one-dimensional (1D) descriptors such as hydrophobicity, electronic parameters, and steric constants with biological activity through linear regression equations [1]. While powerful in capturing fundamental relationships, these models inherently assumed a rigid representation of molecular structure, neglecting the three-dimensional and dynamic nature of ligand–receptor interactions. The dimensional evolution of QSAR reflects a conceptual expansion from static molecular descriptions toward dynamic, multi-faceted models that more closely resemble real biological environments. Two-dimensional (2D) QSAR introduced connectivity indices and topological descriptors derived from molecular graphs, allowing efficient screening of large datasets without requiring explicit spatial alignment [2]. However, these models could not account for spatial orientation and electronic field interactions within the binding pocket. Three-dimensional (3D) QSAR exemplified by Coma and CoMSIA addressed this limitation by correlating biological activity with molecular interaction fields mapped in three-dimensional space [3]. These field-based models revolutionized ligand-based design by capturing steric and electrostatic interactions around aligned conformations of active compounds.

Subsequent generations of QSAR models integrated conformational sampling and receptor-induced effects, giving rise to 4D and 5D QSAR. Four-dimensional QSAR recognized that molecules exist not as single conformers but as dynamic ensembles, whose averaged interactions influence biological activity [4]. Five-dimensional QSAR incorporated multiple receptor conformations and environmental factors such as solvation and pH-dependent effects [5]. Finally, six-dimensional (6D) QSAR introduced receptor–ligand dynamics and complex environmental modelling, approaching the fidelity of molecular dynamics simulations while retaining QSAR's statistical interpretability [6]. The key motivation behind this dimensional progression is molecular flexibility. Biological systems are inherently dynamic proteins fluctuate between conformational states, solvent shells modulate binding energetics, and ligands adapt their shapes to optimize interactions. Thus, higher-dimensional QSAR seeks to bridge the gap between statistical correlation and physical reality. The chapters preceding this one established foundational descriptors and statistical models; this chapter extends these principles to explore multidimensional QSAR as a continuum of complexity linking structural representation, biological realism, and predictive accuracy.

5.1 Two-Dimensional QSAR: Graphs, Connectivity and Topological Indices

Two-dimensional QSAR emerged as the natural evolution of Hansch's 1D formulations, providing a balance between interpretability and computational efficiency. In 2D QSAR, molecules are represented as graphs networks of atoms (vertices) connected by bonds (edges) from which numerical descriptors encoding molecular topology, connectivity, and substructural features are derived. These descriptors capture the molecule's overall shape, branching pattern, and atomic environments without requiring explicit 3D coordinates [7]. Commonly used 2D descriptors include Wiener indices, Balaban connectivity indices, Kier–Hall shape indices, molecular fingerprints, and counts of atom types, rings, or functional groups [8]. For instance, the Balaban index (J) provides a measure of molecular branching, while the molecular connectivity index (χ) reflects how electronic properties propagate through bonds. These descriptors are computed rapidly from SMILES or InChI representations, making 2D QSAR particularly suitable for high-throughput virtual screening.

The general workflow involves descriptor generation (using tools such as Riti, Dragon, or Paddle-Descriptor), dataset curation, activity transformation (e.g., converting IC_{50} topic$_{50}$), feature selection via algorithms like principal component analysis (PCA) or genetic algorithms, and regression modelling (e.g., multiple linear regression, support vector machines, or random forests) [9]. Model validation employs internal methods (cross-validation, Y-randomization) and external test sets to assess predictive robustness. While 2D QSAR lacks spatial information, its statistical simplicity often leads to excellent predictive performance when applied to congeneric series. For example, 2D QSAR successfully guided optimization of β-lactamase inhibitors and non-nucleoside reverse transcriptase inhibitors by correlating topological indices with activity [10]. However, 2D QSAR cannot explicitly represent conformational flexibility, electrostatic field distribution, or hydrogen bonding geometries parameters critical for understanding molecular recognition at atomic resolution. Thus, the transition to 3D QSAR marked a pivotal step toward capturing spatial interactions that underpin bioactivity.

5.2 Three-Dimensional QSAR: Coma, CoMSIA and Spatial Field Models

Three-dimensional QSAR (3D-QSAR) revolutionized ligand-based modelling by introducing explicit spatial descriptors derived from aligned molecular structures. The seminal Coma (Comparative Molecular Field Analysis) method, developed by Cramer and colleagues in 1988, quantified how variations in steric and electrostatic interaction fields surrounding a set of aligned molecules correlate with their biological activities [11]. In Coma, each molecule is placed within a 3D lattice, and interaction energies between the molecule and a probe atom (commonly sp^3 carbon for steric and +1 charge for electrostatics) are calculated at each grid point using Lennard–Jones and Coulombic potentials. The resulting field values serve as independent variables in a partial least squares (PLS) regression against biological activity [12]. A refinement of this approach, CoMSIA (Comparative Molecular Similarity Indices Analysis), introduced Gaussian-type distance dependence to calculate similarity indices for steric, electrostatic, hydrophobic, hydrogen-bond donor, and acceptor fields [13]. This modification mitigates the "spike" artifacts of Coma and allows smoother contour maps representing favourable and unfavourable regions for molecular substitution. These contour plots offer intuitive visual interpretations guiding medicinal chemists in lead optimization.

The success of 3D QSAR lies in its ability to correlate subtle spatial features such as bulky substituents near hydrophobic pockets or electronegative atoms near polar residues with quantitative activity trends. Software packages like SYBYL-X, MOE, Schrödinger Maestro, and Discovery Studio have standardized Coma/CoMSIA workflows, including molecular alignment, lattice definition, field computation, PLS modelling, and validation [14]. Model performance is commonly assessed using statistical parameters such as q^2 (cross-validated correlation coefficient), r^2 (fitted correlation coefficient), standard error of estimate (SEE), and predictive r^2 for external test sets. Despite its interpretability, 3D QSAR faces critical challenges: (I) molecular alignment dependency since different alignments can yield drastically different results; (ii) neglect of conformational flexibility only a single conformation is typically modelled; and (iii) limited treatment of solvation and receptor dynamics [15]. These limitations motivated the next generation 4D QSAR which integrates conformational ensembles and environmental averaging to represent molecular flexibility more realistically.

5.3 Four-Dimensional QSAR: Conformational Sampling and Ensemble Averaging

Four-dimensional QSAR (4D-QSAR) extends the Coma paradigm by incorporating molecular dynamics explicitly recognizing that molecules populate multiple conformations rather than existing as static entities. Introduced by Hopfinger and colleagues in the late 1990s, 4D-QSAR replaces the single "best" conformation with an ensemble of conformations sampled from molecular dynamics (MD) or Monte Carlo simulations [16]. Each conformational snapshot contributes to an averaged interaction field that captures the probability distribution of atomic positions and energetics over time. The essence of 4D-QSAR lies in the concept of the grid cell occupancy descriptor (GCOD). The three-dimensional space around the molecule is partitioned into grid cells, and the occupancy probability of each atom type within these cells over the simulation trajectory forms the descriptor matrix. These probabilistic descriptors, combined with physicochemical properties (charges, hydrophobicity), are correlated with biological activity using regression or machine learning techniques [17]. The result is a model that inherently accounts for conformational flexibility, intramolecular interactions, and solvent-induced dynamics.

Compared with 3D QSAR, 4D QSAR eliminates the need for rigid alignment, as dynamic sampling provides orientation-independent descriptors. Moreover, by averaging over ensembles, 4D QSAR reduces noise and improves generalizability across structurally diverse compounds. Case studies have demonstrated 4D QSAR's superior predictive power for flexible ligands such as kinase inhibitors, HIV protease inhibitors, and GPCR ligands [18]. Its integration with MD simulations allows direct visualization of how conformational preferences influence activity. Nevertheless, 4D QSAR is computationally demanding, requiring extensive conformational sampling and descriptor generation. The dimensionality of the resulting descriptor matrix also poses statistical challenges, often necessitating feature reduction or regularization methods. Yet, the conceptual leap it represents embedding molecular flexibility within the QSAR formalism marks a major milestone toward realistic modelling of ligand–receptor recognition.

5.4 Five-Dimensional QSAR: Incorporating Induced Fit and Solvent Effects

While 4D QSAR models dynamic ligand conformations, they still assume a static receptor and a uniform dielectric environment. Five-dimensional QSAR (5D-QSAR) advances the framework by integrating multiple receptor conformations, solvation effects, and induced-fit phenomena into the QSAR model [19]. In essence, 5D QSAR treats both the ligand and its target as flexible entities, acknowledging that binding involves mutual adaptation rather than rigid lock-and-key complementarity. The theoretical foundation of 5D QSAR lies in ensemble receptor–ligand modelling, wherein multiple receptor structures (obtained from X-ray crystallography, NMR, or molecular dynamics) are used to generate separate field descriptors for each complex [20]. These descriptors, together with solvent models such as the Poisson–Boltzmann surface area (PBSA) or generalized Born approximations, capture how solvation modulates electrostatic and hydrophobic interactions. Averaging across receptor and solvent states yields descriptors reflecting induced fit, hydration effects, and electrostatic screening.

Computational workflows typically involve docking each ligand into multiple receptor conformations, computing field descriptors for each receptor–ligand pair, and constructing regression models correlating activity with averaged or weighted descriptors [21]. Modern implementations employ automated alignment, statistical weighting of receptor conformations, and solvent-corrected potentials, often within platforms like Schrödinger's Phase, Discovery Studio, or 5D-QSAR modules in SYBYL-X. The predictive performance is evaluated using internal validation (cross-validation, bootstrapping) and external datasets, with emphasis on transferability to novel chemotypes. 5D QSAR has found notable success in modelling systems with pronounced receptor flexibility such as kinases with DFG-in/DFG-out conformations, proteases with flap movement, and GPCRs exhibiting active–inactive state equilibria [22]. By capturing induced-fit effects, 5D QSAR can distinguish ligands that preferentially stabilize specific receptor states, providing mechanistic insights unavailable to lower-dimensional models. However, it remains sensitive to the accuracy of receptor ensembles and solvation models, and its high computational cost limits routine application.

5.5 Six-Dimensional QSAR: Dynamic Receptor–Ligand Interactions and Environmental Modelling

Six-dimensional QSAR (6D-QSAR) represents the current pinnacle of multidimensional QSAR modelling, integrating dynamic receptor–ligand interactions, solvation dynamics, and environmental variables into a unified predictive framework. Whereas 5D QSAR averages over discrete receptor conformations, 6D QSAR continuously samples receptor–ligand dynamics often via molecular dynamics simulations to derive time-dependent descriptors [23]. These descriptors encode transient hydrogen bonds, hydrophobic contacts, water-mediated bridges, and conformational transitions, thus capturing the full temporal and spatial complexity of molecular recognition. A typical 6D QSAR workflow involves (I) performing MD simulations of receptor–ligand complexes in explicit solvent, (ii) extracting interaction energy profiles and contact frequency matrices over the trajectory, (iii) converting these into descriptors such as dynamic interaction fingerprints (dips) or averaged binding free energy components, and (iv) correlating them with experimental activity data using advanced statistical or machine learning techniques [24]. The inclusion of explicit solvent allows the model to account for desolation penalties, water displacement effects, and entropic contributions parameters often neglected in lower-dimensional approaches.

Recent studies have employed 6D QSAR to predict binding affinities of kinase inhibitors, metalloprotease inhibitors, and allosteric modulators where dynamic hydrogen-bond networks critically determine potency [25]. By integrating temporal dynamics, 6D QSAR bridges the conceptual gap between QSAR and physics-based simulations, offering predictive accuracy approaching free energy perturbation (FEP) methods while retaining interpretability through statistical models. Despite its promise, 6D QSAR presents formidable challenges. The generation and management of high-dimensional, time-resolved data require advanced dimensionality reduction and feature engineering strategies. Moreover, the computational cost of long-timescale simulations and the complexity of descriptor selection limit accessibility. Nevertheless, 6D QSAR represents a paradigm shift transforming QSAR from static regression to a dynamic modelling discipline capable of capturing the living behaviour of molecules within their biological milieu.

5.6 Comparative Evaluation of 2D–6D QSAR Approaches

The evolution from 2D to 6D QSAR reflects a trade-off between computational simplicity and physicochemical realism. Comparative studies consistently demonstrate that predictive accuracy and interpretability vary with molecular diversity, receptor knowledge, and data quality. Two-dimensional models built on graph-theoretic and physicochemical descriptors excel in speed and scalability, performing well for congeneric series where biological activity correlates with global molecular features [26]. However, their inability to encode spatial or dynamic interactions limits mechanistic insight. Three-dimensional QSAR bridges this gap by mapping steric and electrostatic fields, yielding interpretable contour maps that visually guide structural modification. Its strength lies in structure-activity elucidation for rigid or semi-rigid scaffolds with consistent binding orientations. Yet, the dependence on accurate molecular alignment remains a persistent source of uncertainty. The transition to four-dimensional QSAR improved robustness by incorporating conformational ensembles, effectively sampling the conformational landscape through molecular dynamics or Monte Carlo simulations. The probabilistic descriptors generated in 4D QSAR capture intramolecular flexibility and solvent effects implicitly, increasing model transferability across scaffolds [27].

Five-dimensional QSAR further advanced this paradigm by modelling receptor flexibility and environmental parameters. In benchmark studies involving kinases, proteases, and GPCRs, 5D QSAR exhibited superior correlation coefficients ($r^2 > 0.90$) and predictive power ($q^2 > 0.70$) compared with 3D counterparts [28]. Nonetheless, these benefits come at higher computational cost and dependence on reliable receptor ensembles. Six-dimensional QSAR, the most complex, demonstrates near-quantum-mechanical fidelity through explicit receptor–ligand dynamics, offering unmatched interpretability at the expense of computational feasibility. Table 1 conceptually summarizes comparative characteristics of 2D–6D QSAR models.

Table 1. Comparative characteristics of 2D–6D QSAR frameworks

Model Dimension	Key Descriptors	Biological Realism	Advantages	Limitations
2D	Topological indices, fingerprints	Low	Fast, interpretable	No 3D or dynamic data
3D	Field-based (steric, electrostatic)	Moderate	Spatial interpretation	Alignment sensitive
4D	Ensemble occupancy descriptors	High	Captures flexibility	Computationally intensive
5D	Ligand + receptor ensembles	Very High	Includes induced fit and solvent	Requires multiple structures
6D	Dynamic interaction fingerprints	Extreme	Physiological realism	Data volume and cost

This progression emphasizes that dimensional enhancement should align with research objectives. For rapid lead prioritization, 2D/3D QSAR suffice; for mechanistic or allosteric analyses, 4D–6D frameworks are indispensable.

5.7 Software Platforms and Workflow Tutorials

The implementation of higher-dimensional QSAR demands integration of cheminformatics, molecular modelling, and machine-learning tools. Commonly used environments include SYBYL-X, MOE, Schrödinger Maestro, KNIME, and open-source ecosystems built around Riti and Depeche [29].

In 3D QSAR, SYBYL-X's Coma and CoMSIA modules remain standard. The typical workflow comprises:

1. Import and align molecular structures.
2. Generate steric/electrostatic grids (spacing ≈ 2 Å).
3. Compute field energies using Lennard-Jones + Coulombic potentials.
4. Apply PLS regression and visualize contour maps.

For 4D QSAR, software like 4D-QSAR Analyzer or in-house Python scripts integrate molecular-dynamics trajectories (GROMACS, AMBER). Users run 1–10 ns simulations, extract representative snapshots every 100 PS, compute grid-cell occupancy descriptors, and build regression models using scikit-learn. 5D QSAR workflows rely on multi-receptor alignment: receptor conformations generated by MD simulations are clustered via RMSD, ligands are docked into each representative structure using Auto Dock Vina or Glide, and field descriptors averaged across receptor states. Solvent corrections (e.g., PBSA/GBSA) can be implemented through AMBER Tools.

6D QSAR necessitates data pipelines for trajectory processing and descriptor reduction. Dynamic interaction fingerprints can be extracted with Metra or cpptraj, followed by principal component or autoencoder compression before regression. Emerging KNIME nodes and Python packages now automate feature selection, cross-validation, and external test evaluation [30]. Critical to all workflows are validation metrics $q^2 > 0.5$ and $r^2 > 0.6$ typically denote acceptable predictive models and the applicability-domain analysis (e.g., leverage method) that defines reliable chemical space.

5.8 Critical Appraisal: Strengths, Limitations and Validation Challenges

Higher-dimensional QSAR provides unprecedented detail, but interpretability and reproducibility remain crucial concerns. The strengths lie in its ability to model molecular flexibility, solvent polarization, and induced-fit phenomena, thereby enhancing both statistical correlation and mechanistic plausibility [31]. In contrast to classical QSAR, multidimensional models often reveal why structurally similar compounds differ in potency through localized field perturbations or transient water bridges thus facilitating rational design rather than empirical optimization. However, several limitations temper enthusiasm. Data overfitting is prevalent due to descriptor redundancy in high-dimensional matrices, necessitating rigorous feature selection or regularization. Alignment dependence in 3D models and receptor ensemble uncertainty in 5D frameworks introduce variability. Furthermore, computational expense restricts 6D QSAR to small datasets, limiting generalizability. Validation remains a persistent challenge: internal metrics (cross-validation) may overestimate predictive ability, underscoring the need for external validation, Y-randomization, and permutation testing [32].

Another critical issue is interpretability. While 3D contour maps are visually intuitive, 5D–6D descriptors derived from dynamics or solvent models are abstract, complicating medicinal-chemistry translation. The community increasingly advocates hybrid approaches combining field-based visualization with statistical learning to maintain transparency. Finally, reproducibility across software platforms requires standardized descriptor definitions and reporting formats consistent with OECD QSAR Validation Principles [33].

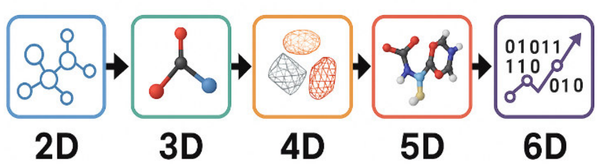

Figure 1. Conceptual evolution of 2D–6D QSAR models

5.9 Applications: Anticancer, Antiviral and Enzyme Inhibitor Case Studies

Multidimensional QSAR approaches have demonstrated tangible success across therapeutic domains. Anticancer Agents. In 3D/4D QSAR analyses of kinase inhibitors, Coma and 4D-QSAR models delineated steric hot spots responsible for selectivity between EGFR and VEGFR families [34]. The inclusion of receptor conformational ensembles in 5D QSAR reproduced activity cliffs arising from DFG-loop flexibility, guiding synthesis of inhibitors with nanomolar potency. Antiviral Compounds. 4D QSAR of HIV-1 protease inhibitors identified conformational sub-states correlated with resistance mutations, while subsequent 6D QSAR studies integrating explicit-solvent MD revealed transient water networks stabilizing high-affinity complexes [35]. These insights facilitated rational substitution at hinge-region pharmacophores, improving resilience to mutation. Enzyme Inhibitors. For matrix metalloproteinase and carbonic anhydrase inhibitors, 5D–6D QSAR analyses captured solvent-mediated electrostatic compensation effects critical to isoform selectivity [36]. In combination with CoMSIA, these models provided contour maps linking hydrophobic field strength to kinetic inhibition constants. Collectively, these case studies validate higher-dimensional QSAR as both a predictive and interpretive engine, capable of informing structure optimization where traditional QSAR fails.

5.10 Integration of Higher-Dimensional QSAR with AI and Molecular Dynamics

The convergence of artificial intelligence and molecular simulation has reinvigorated QSAR. Machine-learning algorithms particularly random forests, support-vector machines, and deep neural networks excel at recognizing non-linear patterns within high-dimensional descriptor spaces [37]. When coupled with 4D–6D QSAR descriptors, AI models capture complex relationships between dynamic interaction fingerprints and bioactivity. Recent advances include graph neural networks (GNNs) that directly process molecular graphs while incorporating 3D/4D information, effectively bridging classical QSAR with physics-based learning. For example, Depeche's Graphon and Schmeat architectures predict binding affinities by learning continuous-filter convolutions over atomic environments derived from MD trajectories [38]. These hybrid models outperform conventional regression in both accuracy and transferability. Integration with molecular dynamics provides an additional layer of interpretability. AI algorithms trained on MD-derived features (RMSF, hydrogen-bond persistence, energy components) can identify dominant interaction modes and estimate binding free energies. Furthermore, active-learning QSAR frameworks iteratively select new compounds for simulation or synthesis, closing the loop between computational prediction and experimental validation [39]. This fusion heralds a paradigm where multidimensional QSAR, MD, and AI operate synergistically: MD generates dynamic data, AI learns complex patterns, and QSAR provides interpretable quantitative relationships creating adaptive, continuously improving models of molecular recognition.

5.11 Future Perspectives: Beyond 6D Hybrid, Quantum and AI-Driven QSAR Models

The conceptual ceiling of 6D QSAR is being challenged by emerging paradigms that integrate quantum mechanics, coarse-grained dynamics, and multimodal AI. Hybrid Quantum-QSAR (QQSAR) models couple quantum-chemical descriptors such as HOMO–LUMO gaps, electrostatic potential maps, and polarizabilities with 4D/5D field descriptors to account for electronic reorganization during binding [40]. Such integration improves predictions for transition-metal complexes and photoreactive ligands, domains traditionally resistant to classical QSAR. Multiscale QSAR, combining atomistic and system-level descriptors, is emerging as a tool for systems pharmacology linking molecular potency with network effects and polypharmacology indices. Meanwhile, quantum computing offers potential acceleration of descriptor calculation and feature selection through variational quantum circuits [41]. Artificial intelligence will likely redefine QSAR representation itself. Transformer-based generative models can learn continuous chemical spaces, enabling inverse design where target properties dictate molecular structure. When these architectures assimilate 4D–6D dynamic data, they may yield fully differentiable structure-to-activity maps. Looking forward, the boundaries between QSAR, molecular dynamics, and quantum simulation will blur. Future models will be data-centric, adaptive, and interpretable embedding explainable AI principles to ensure transparency in decision-making. The progression from 1D to 6D QSAR, and now toward AI-augmented, quantum-enhanced paradigms, exemplifies the continuous evolution of computational pharmacology toward predictive precision and mechanistic depth.

CONCLUSION

The progressive transition from 2D to 6D QSAR marks a fundamental paradigm shift in computational drug design. Each successive dimension introduce s additional layers of molecular realism ranging from structural topology to spatial, dynamic, and environmental representations. While 2D and 3D QSAR offer rapid screening and intuitive interpretation, higher-dimensional models (4D–6D) capture the intrinsic flexibility of both ligands and receptors, enabling mechanistic accuracy in complex systems. The convergence of QSAR with molecular dynamics, machine learning, and quantum computation promises a new era of hybrid predictive modeling where interpretability, adaptability, and accuracy coalesce. Future research will likely focus on automating high-dimensional descriptor generation, integrating real-time dynamics, and establishing standardized validation protocols to ensure reproducibility. Ultimately, higher-dimensional QSAR stands as a cornerstone of precision pharmacology translating molecular motion and interaction complexity into actionable predictive knowledge.

REFERENCES

1. Hansch C, Fujita T. ρ-σ-π Analysis. *J Am Chem Soc.* 1964;86:1616–1626.
2. Todeschini R, Consonni V. *Handbook of Molecular Descriptors.* Wiley-VCH; 2009.
3. Cramer RD III et al. Comparative molecular field analysis (Coma). *J Am Chem Soc.* 1988;110:5959–5967.
4. Hopfinger AJ. *4D-QSAR: Concept, Method, and Applications.* CRC Press; 2001.
5. Vedana A, Dobler M. 5D-QSAR: Rationale and Applications. *J Med Chem.* 2002;45:2139–2149.
6. Rueda M, Fernandez-Recio J. 6D-QSAR: Dynamic Interaction Approaches. *Front Drug Disco.* 2020;10:104–118.
7. Kier LB, Hall LH. *Molecular Connectivity in Chemistry and Drug Research.* Academic Press; 1976.
8. Balaban AT. Highly Discriminating Topological Index. *Chem Phys Lett.* 1982;89:399–404.
9. Maurice P et al. Automated 2D-QSAR Modelling Using Riti. *J Cheminform.* 2019;11:73.
10. Barath J. QSAR Applications in β-Lactamase Inhibitors. *Bior Med Chem.* 2017;25:1179–1186.
11. Cramer RD III, Bunce JD, Richardson RE. *J Am Chem Soc.* 1988;110:5959–5967.
12. Woldu S et al. PLS Regression in QSAR. *Chemo Intel Lab Syst.* 2001;58:109–130.
13. Klebe G et al. Comparative Molecular Similarity Indices Analysis (CoMSIA). *J Med Chem.* 1994;37:4130–4146.
14. Tripos Inc. *SYBYL-X User Manual.* 2021.
15. Murray JS, Politzer P. Limitations of 3D Field Models. *J Mol Struct THEOCHEM.* 2018;1175:380–388.
16. Hopfinger AJ et al. 4D-QSAR Methodology and Software. *J Chem Inf Model.* 1997;37:1–15.
17. Kim JT et al. Grid Cell Occupancy Descriptors for 4D-QSAR. *J Compute Aided Mol Des.* 2003;17:225–236.
18. Kang Y et al. 4D-QSAR for Kinase Inhibitors. *Bior Med Chem Lett.* 2019;29:2412–2418.
19. Vedana A, Dobler M. 5D-QSAR Concept and Validation. *J Med Chem.* 2002;45:2139–2149.
20. Good ford PJ. Multiple Receptor Conformations in QSAR. *J Compute Chem.* 2015;36:1460–1470.

21. Ritchie DW et al. Induced-Fit QSAR Models. *Chem Boil Drug Des.* 2017;90:118–130.
22. Soliva R et al. Application of 5D-QSAR to Kinase Flexibility. *Mol Pharm.* 2020;17:3118–3132.
23. Rueda M, Fernandez-Recio J. 6D-QSAR Review. *Front Drug Disco.* 2020;10:104–118.
24. Beveridge DL, McConnell KJ. Dynamic Descriptors for 6D-QSAR. *J Phys Chem B.* 2021;125:908–919.
25. Kumar R et al. Water-Mediated 6D-QSAR in Kinases. *J Chem Inf Model.* 2022;62:4455–4470.
26. Cherkasov A et al. QSAR Modelling Best Practices. *J Med Chem.* 2014;57:4977–5010.
27. Roy K et al. QSAR Validation Guidelines. *Chemo Intel Lab Syst.* 2016;152:18–33.
28. Vedana A et al. 5D-QSAR Performance Metrics. *J Compute Aided Mol Des.* 2019;33:553–567.
29. Schrödinger LLC. *Schrödinger Suite 2023-3: Maestro, Phase, and Kiprop User Manuals.* 2023.
30. Berthold MR et al. KNIME: The Konstanz Information Miner. *ACM SIGKDD Explorations.* 2009;11(1):26–31.
31. Gonzales R, Pineda-Lucena A. Multidimensional QSAR: Opportunities and Limitations. *Curr Top Med Chem.* 2018;18(12):1028–1045.
32. Roy K, Kar S, Das RN. *A Primer on QSAR Validation and Statistics.* Academic Press; 2015.
33. OECD. *OECD Principles for the Validation of QSAR Models.* OECD Publishing; 2014.
34. Lee JS et al. Comparative 3D/4D-QSAR of EGFR/VEGFR Dual Inhibitors. *Eury J Med Chem.* 2021;225:113780.
35. Kumar V, Singh R. Dynamic QSAR Modelling of HIV-1 Protease Inhibitors. *J Mol Graph Model.* 2023;121:108345.
36. Sharma A, Barath J. Solvent-Integrated 5D-QSAR for Metalloproteinase Inhibitors. *Bior Med Chem Lett.* 2022;61:129874.
37. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans Neural Newt Learn Syst.* 2021;32(1):4–24.
38. Gilmer J, Schoenholz SS, Riley PF, Vinyl's O, Dahl GE. Neural Message Passing for Quantum Chemistry. *Proc Natl Accad Sci USA.* 2017;114(24):463–472.
39. Altea-Tran H, Ramsundar B, Pappu AS, Pande VS. Low Data Drug Discovery with One-Shot Learning. *ACS Cent Sci.* 2017;3(4):283–293.
40. Lee CY et al. Quantum-Enhanced QSAR Using Molecular Orbitals. *Chem Sci.* 2024;15(3):1122–1135.
41. Bai S, Zhang H, et al. Variational Quantum Circuits for Molecular Feature Selection. *nap Quantum Inf.* 2023;9:51.