

Accelerating Polymorph Screening with AI & ML: A New Era in Drug Development

Anantha Rajmohan Muthusamy¹, Diwakar Chauhan^{1*}, Arvind Kumar Jain^{2*}, Meenakshi Sundaram Somasundaram³ and Amit Singh⁴

¹School of Basic Sciences, Galgotias University, Greater Noida, 203201, INDIA.

²School of Basic and Applied Sciences, IILM University, Knowledge Park II, Greater Noida, 201306, Uttar Pradesh, INDIA

³Department of Industrial Chemistry, Alagappa University, Karaikudi, Tamilnadu, 630003, INDIA

⁴Department of Chemistry, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, 110016, INDIA

*Corresponding Author: chauhan.diwakar@gmail.com, jain1dcy@gmail.com



www.jrasb.com || Vol. 4 No. 2 (2025): April Issue

Received: 03-04-2025

Revised: 13-04-2025

Accepted: 18-04-2025

ABSTRACT

Polymorph screening plays a crucial role in pharmaceutical development, influencing the solubility, stability, and bioavailability of active pharmaceutical ingredients (APIs). Traditional screening methods are time-consuming, labor-intensive, and often yield unpredictable results. Recent advancements in artificial intelligence (AI), machine learning (ML), and computational modeling such as molecular dynamics (MD) and density functional theory (DFT) have revolutionized this process, enabling faster and more accurate predictions of polymorphs, solvates, hydrates, cocrystals, and salts. The integration of AI-driven and computational models in polymorph screening is examined in this paper, with an emphasis on their potential applications in predicting the stability of amorphous solids, solubility of APIs, and the solvates, hydrates and cocrystals for the drug development. We also discussed how thermodynamic viability of solvate and hydrate formation and desolvation kinetics can be analyzed using computational techniques. AI and ML provides increase polymorph screening's success rate, which will impact the drug polymorphs selection for the manufacturing and regulatory compliance. The most recent advancements, challenges, and contemporary approaches in the use of AI/ML, MD and DFT in solid-state drug development are the main topics of this review.

Keywords- Intelligence, Computational polymorph screening, Experimental Polymorph screening, Machine learning, Molecular dynamics.

I. INTRODUCTION

Polymorphism in the pharmaceutical industry denotes the ability of an active pharmaceutical ingredients (APIs), to exist in many crystalline forms. These polymorphic forms can significantly affect APIs solubility, stability, manufacturability, and regulatory compliance [1]. The solubility and dissolution rates of drugs in polymorphic forms differ, which directly affects bioavailability. Specifically, the metastable polymorphs are generally more soluble than their stable counterparts. In the case of the antiviral drug Ritonavir, the metastable polymorph progressively transformed into a more stable polymorph during storage; hence, the drug was

reformulated because the less soluble polymorph affected bioavailability [2]. While a stable polymorph ensures a longer shelf life and different polymorphs may differ in their resistance to environmental elements including light, humidity, and temperature, degradation in an unstable form may compromise the drug's efficacy and safety. [3]. Thus, selecting the most stable polymorph is essential to maintaining the drug's efficacy.

Similarly, variations in the crystal packing, hardness, and compressibility of polymorphic APIs affect their mechanical properties, which in turn affect their pharmaceutical production processes, such as tablet formation, compaction, and powder flowability [4]. Optimizing polymorph selection ensures efficient

manufacturing and consistent drug product quality. In order to ensure batch-to-batch consistency and drug efficacy, regulatory agencies such as the European Medicines Agency (EMA) and the U.S. Food and Medicine Administration (FDA) require thorough characterization of polymorphs, including X-ray diffraction (XRD), differential scanning calorimetry (DSC), and infrared spectroscopy (IR) techniques [5]. To comply with regulations and gain market acceptance, it is crucial to maintain control over polymorph selection and manufacturing since post-approval changes in polymorphic form may result in regulatory barriers that require additional research to demonstrate stability and bioequivalence.

II. POLYMORPH SCREENING

Polymorph screening is a process to identify and characterize all the different crystalline forms of a compound, with the goal of selecting the most suitable polymorph for drug development. The discovery of polymorphs for ritonavir [6] and carbamazepine [7] highlighted the importance of careful screening procedures, which include identifying the thermodynamically most stable form, having the right ability to be compressed and flow well for easier manufacturing, and ensuring the selected polymorph doesn't change in ways that could compromise its therapeutic efficacy [8-17].

III. EXPERIMENTAL POLYMORPH SCREENING

Combinations of theoretical and experimental approaches are used in practice to obtain stable polymorphs for manufacturing (Figure 1). These techniques look at the many screening techniques used in polymorph screening, including as crystallization, computer modeling, solid-state characterization, and stability studies.

3.1 Crystallization Techniques

Crystallization is a widely used screening mechanism that involves altering the crystallization conditions to obtain different polymorphs [18]. The following screening methods employed in the polymorph screening.

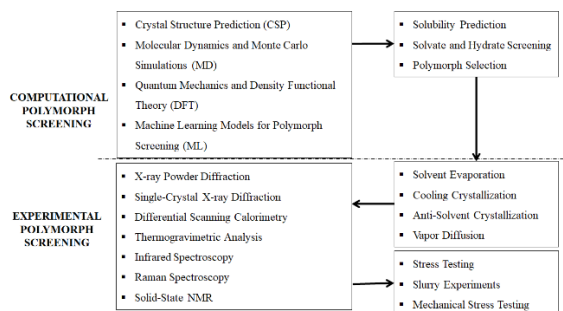


Figure 1: Summary of experiments and characterization in polymorph screening

A common way to do initial screening is the slow or rapid solvent evaporation, where the API is mixed with a solvent and then allowed to evaporate either slowly or quickly, leading to the formation of different crystal forms. In the cooling crystallization experiments, crystallization is accelerated by progressively lowering the solubility of a supersaturated solution. The resulting polymorph is significantly influenced by the chosen solvent and rate of cooling. Anti-Solvent Crystallization was carried out by reducing the solubility of the API by adding an anti-solvent, which caused different polymorphs to precipitate. Similar to this, vapor diffusion involves subjecting a solution to controlled humidity or solvent vapors, which over time can cause polymorphic changes. These crystallization techniques provides different polymorphs of a drug molecule, which can then be characterized and evaluated further.

3.2 Solid-State Characterization Techniques

The structural and physicochemical characteristics of the polymorphs are determined analytically using the following characterization techniques. X-ray powder diffraction (XRPD) and Single-Crystal X-ray Diffraction (SCXRD) are valuable tools for identifying and characterizing polymorphism due to its ability to differentiate between crystal structures based on their unique diffraction patterns. DSC can be used to determine the stability of polymorphs and the transition temperatures by examining the changes in heat flow during phase transitions. Thermogravimetric Analysis (TGA) is a technique that is useful for identifying solvates and hydrates that lose solvent molecules at specific temperatures. Similar to the PXRD, polymorphs can be easily identified using IR and Raman spectroscopy. Even when XRPD patterns look alike, small differences in the arrangement of molecules in different crystalline forms can be spotted using solid-state nuclear magnetic resonance (SSNMR).

3.3 Stability and Transformation Studies

To determine its resilience in different conditions, several stress tests are conducted. Researchers use slurry conversion experiments in different solvents to track the changes between polymorphs and select the most stable form. The long-term stability of the polymorph is determined by exposing it to different humidity and temperature conditions. Polymorphic transitions can be brought on by compression, grinding, and milling. Evaluating these conditions ensures the stability of the selected polymorph during the formulation and manufacturing. These studies aid in determining which polymorph is most suitable for manufacturing and obtaining regulatory approval.

IV. COMPUTATIONAL POLYMORPH SCREENING

Historically, polymorph screening has depended on extensive experimental techniques like crystallization and solid-state characterization. These techniques are resource- and time-intensive and often fall short of identifying all possible polymorphic forms. Recently, computational methods have become powerful tools for speeding up polymorph screening [19], enabling extremely precise polymorph, solubility, and stability predictions. Our study examines the ways in which several computational methods, like machine learning (ML) models, quantum mechanical techniques, molecular dynamics (MD) simulations, and crystal structure prediction (CSP)[20-34], improve the success rate of polymorph screening.

4.1 In Silico Techniques for Polymorph Screening:

Understanding the stability and transition paths of polymorphs under various environmental conditions, such as temperature, pressure, and humidity, can be accomplished with the help of Monte Carlo (MC) and MD simulations. A typical MD and MC simulation procedure is shown in Figure 3. MD simulations monitor atomic movements throughout time, allowing for the prediction of polymorph interconversion and solubility behavior [35]. Monte Carlo methods, on the other hand, employ statistical sampling to determine the most likely crystal arrangements. These simulations have been useful in forecasting phase transitions and hydration/dehydration processes in medicinal molecules.

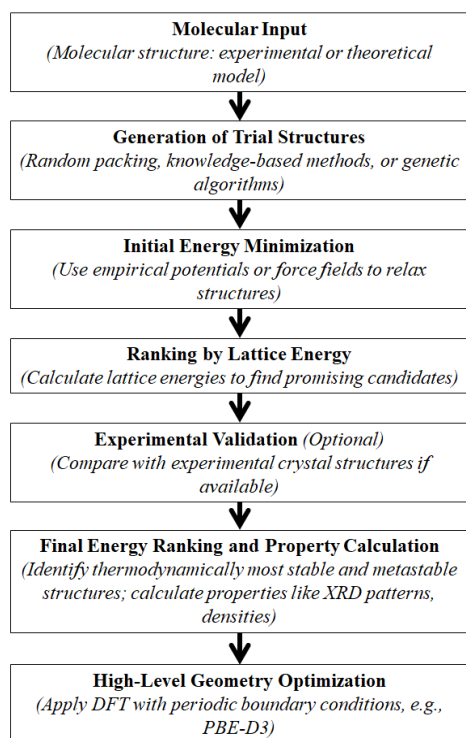


Figure 2: Workflow for CSP

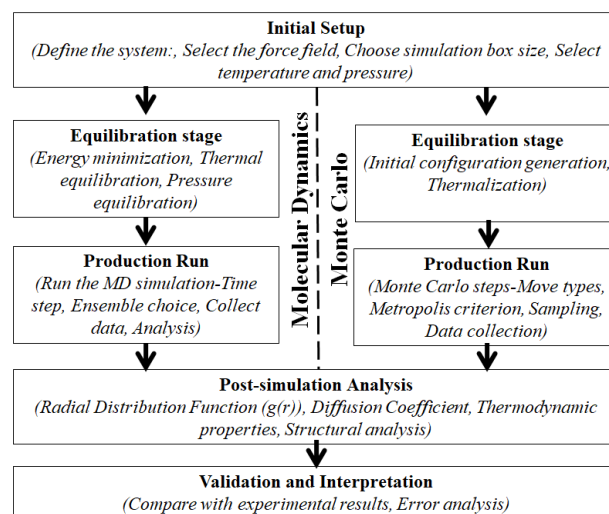
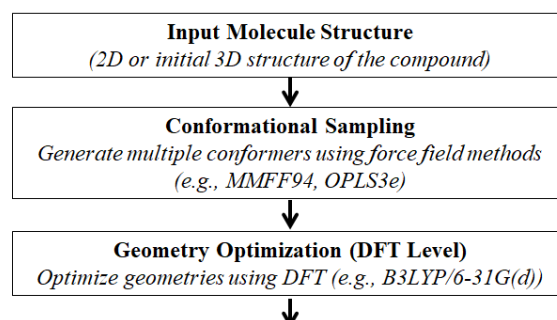


Figure 3: Workflow for Molecular Dynamics and Monte Carlo Simulations

Quantum Mechanics and DFT: Quantum mechanical approaches, particularly DFT, enable precise computations of intermolecular interactions, electron density distribution, and lattice energies in polymorphic systems [36]. Figure 4 illustrates a typical workflow for the DFT approach. Unlike the empirical force fields employed in MD simulations, DFT gives a first-principles approach to determining polymorph stability. It helps forecast how hydrogen bonding, π - π stacking, and van der Waals interactions affect crystal packing. Researchers can gain a more comprehensive understanding of polymorphic landscapes by combining DFT, CSP, and MD techniques.

Polymorph screening has been transformed by significant advances in AI and ML. Figure 5 illustrates a typical workflow for an ML model. ML models trained on experimental and computational datasets may predict polymorph formation using chemical descriptors, crystallization conditions, and solvent parameters [37]. Potential polymorphs were classified and ranked using neural networks, support vector machines (SVM), and random forest techniques (RF). AI-powered screening technologies not only speed up medication discovery but also minimize the need for costly and time-consuming experimental trials.



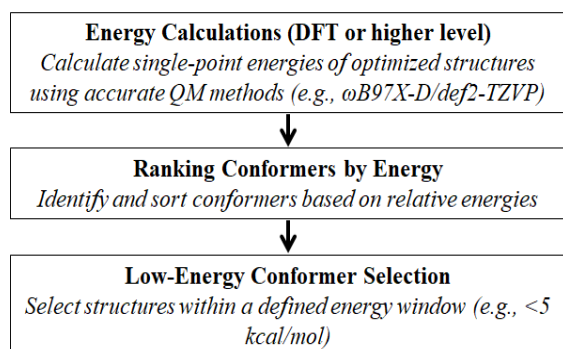


Figure 4: Workflow for Predicting Low-Energy Conformers using Quantum Mechanics and DFT Methods

One of the most difficult issues in polymorph screening is forecasting the solubility and stability of various crystal forms [38]. Computational approaches such as molecular docking, COSMO-RS (Conductor-like Screening Model for Real Solvents), and hybrid AI models can be used to estimate API solubility in different solvents and pH buffers. These predictions are critical in determining the most appropriate polymorphic form for formulation development. Furthermore, in silico techniques help evaluate the stability of amorphous solids and guide the design of amorphous solid dispersions (ASDs) to enhance medication bioavailability.

Computational approaches have been used successfully in salt and co-crystal screening, in addition to classical polymorph screening [39]. Salt production can improve API solubility and stability, but determining the correct counter ion is difficult. By analyzing ion-pair interactions, in-silico methods like molecular docking and quantum mechanical computations can anticipate advantageous salt combinations. Similarly, co-crystal screening uses molecular recognition algorithms to find candidate cofomers that boost therapeutic efficacy. Furthermore, hydrate and solvate prediction models evaluate the possibility of water or solvent integration into the crystal lattice, ensuring that the most stable form is selected for medication development.

Several pharmaceutical companies have successfully discovered polymorphs by incorporating computational techniques into their drug development pipelines. For instance, CSP has proven crucial in forecasting olanzapine polymorphs [40], while AI-powered solubility models have improved formulations for medications that are not very soluble. Because they understand the importance of computational data, regulatory bodies are promoting the use of in silico techniques to assist with patent applications and formulation choices.

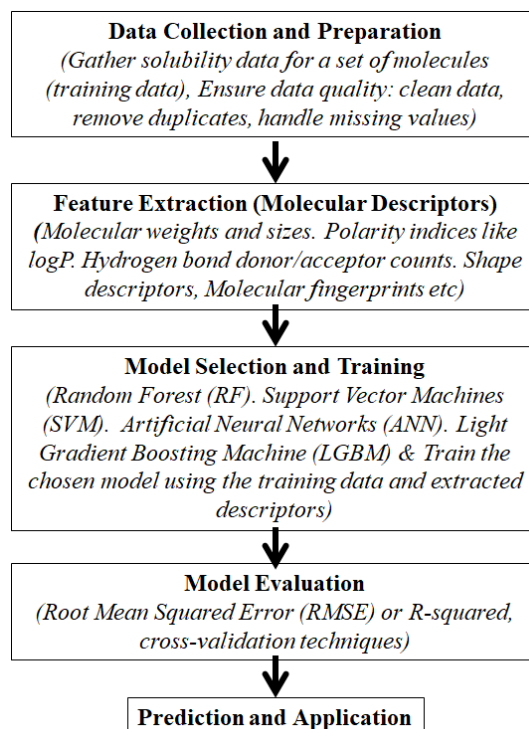


Figure 5: Workflow for Machine Learning Models for Polymorph Screening

4.2 In Silico Applications in Solubility Prediction:

When it comes to the development, formulation, and bioavailability of APIs, solubility is a crucial physicochemical characteristic. In drug discovery, poor solubility is a major obstacle that can result in clinical development failures. In silico methods, such as quantum mechanics, molecular modeling, AI and ML [41], have transformed solubility prediction, cutting down on experimental expenses and speeding up medication development. The developments in computer methods for predicting solubility are covered in this mini-review along with their consequences for pharmaceutical sciences.

Molecular descriptors and quantitative structure-property relationship (QSPR) models create correlations between molecular descriptors and solubility. Molecular weight, hydrogen bonding capacity, lipophilicity, and polar surface area are all examples of descriptors [42]. SVM [43], RF [44, 45], and deep learning models [46, 47] are three machine learning algorithms that increase QSPR accuracy by learning complicated patterns from big datasets.

Simulations for MD shed light on solute-solvent interactions, hydration free energy, and solubility trends. MD assists in predicting solubility differences across pH ranges and temperatures by mimicking molecular behavior in various solvents [48]. Advanced MD methods, such as enhanced sampling strategies, improve solubility estimates even further [49].

AI/ML models use massive chemical datasets to forecast solubility trends. Deep learning architectures,

such as convolutional neural networks (CNNs) [50] and graph neural networks (GNNs) [51], extract molecular information from chemical structures to improve prediction accuracy. Transfer learning and generative models aid in the discovery of novel drug-like compounds with high solubility.

DFT and quantum mechanics computations reveal information about electronic structure, charge distribution, and hydrogen bonding interactions that influence solubility [52]. Quantum mechanical approaches predict solubility in diverse solvents by estimating Gibbs free energy changes during solvation. The COSMO-RS predicts solubility in various solvents using quantum chemistry and statistical thermodynamics [53]. It computes activity coefficients and phase equilibria, which aids in solvent selection and formulation optimization.

Applications in Pharmaceutical Sciences: Formulating and choosing excipients for APIs is made easier by the ability to predict solubility in different solvents. *In silico* methods can be used to identify stable polymorphs and hydrates. The selection of cofomers and counterions to enhance solubility in salt and cocrystal design is aided by computational screening. Making predictions about pH-dependent solubility aids in the formation of prodrugs and salts.

4.3 *In Silico* Applications in Amorphous Solid Forms:

Amorphous solid forms (ASFs) are commonly employed in the pharmaceutical industry to improve the solubility and bioavailability of medicines that are poorly water-soluble [54]. ASFs, unlike crystalline forms, lack long-range molecular organization, making stability and dissolving qualities difficult to predict. *In silico* methods, such as molecular modeling [55], machine learning (ML) [56], and quantum chemistry, have considerably improved our understanding and prediction of ASFs. This mini-review focuses on computational techniques used for predicting amorphous forms, assessing stability, and designing formulations.

MD simulations can offer atomistic information about molecule packing, intermolecular interactions, and glass transition temperature (T_g) estimation. Advanced MD techniques, such as coarse-grained modeling and increased sampling methods, aid in predicting the physical stability of ASFs over time [57].

AI/ML algorithms use extensive datasets to forecast the possibility of a molecule becoming an amorphous solid. Deep learning and support vector machine (SVM) models help forecast glass-forming capability, recrystallization tendencies, and formulation stability. ML can also be used to optimize polymer selection for ASDs [58].

DFT calculations help us comprehend the molecular energetics, hydrogen bonding, and structural flexibility that drive amorphization. Quantum physics predicts the feasibility of ASF production by comparing the energy differences between crystalline and amorphous forms [59]. T_g is an important characteristic

for amorphous drug stability. T_g is estimated using *in silico* methods such as group contribution models and hybrid AI-based approaches, which consider molecular structure, hydrogen bonding strength, and molecular weight.

Application in Pharmaceutical Science: Amorphous Drug Development: Predicting API Amorphization for Increased Solubility. Formulation Development: AI-powered screening of polymers and excipients for ASDs. Stability Prediction: MD and ML models can predict physical stability and recrystallization problems. Bioavailability Enhancement: Optimizing ASFs for improved dissolution profiles.

4.4 *In Silico* Applications in Solvate and Hydrate Screening:

Solvates and hydrates are significant in pharmaceutical development because they influence the solubility, stability, and bioavailability of APIs. Traditional experimental methods for solvate and hydrate synthesis are laborious and time-consuming. *In silico* techniques such as molecular modeling, quantum mechanics, ML, and AI have emerged as useful tools for expediting and optimizing solvate and hydrate screening [60–70]. This mini-review examines the most recent computational developments in predicting solvate and hydrate formation, as well as the implications for pharmaceutical research and development. Solvates and hydrates are solid APIs with solvent or water molecules in their crystal structure. These solvates and hydrates can drastically change the physicochemical characteristics of APIs, requiring careful screening during pharmaceutical development. Computational methods have changed this process by predicting the formation of solvates and hydrates, lowering the need for experiments, and increasing efficiency. MD simulations, which offer insights into lattice stability and molecular interactions, are used to model how APIs behave in various solvent environments. MD makes it easier to comprehend solvation kinetics, desolvation mechanisms, and hydrate and solvate stability in various environments. The thermodynamic viability of solvate and hydrate formation is predicted using quantum mechanical techniques, such as DFT [72–74]. To identify the most stable solid form, these calculations offer helpful information on lattice energies, binding energies, and intermolecular forces. It is becoming more common to use AI and ML models to predict solvate and hydrate production based on extensive data. Based on chemical structures, crystallographic data, and thermodynamic considerations, these models provide precise predictions [75–81]. Two supervised learning methods that have shown potential in accurately identifying solvates and hydrates are neural networks and support vector machines. Forecasting the likelihood of solvate production based on solvent-API interactions is known as computational solvent screening. Solvents are ranked according to their compatibility with APIs utilizing techniques like molecular docking and COSMO-RS

[82]. Computational techniques can be used to estimate desolvation kinetics, which provides insight into how solvates or hydrates change into anhydrous forms. Comprehending these systems facilitates formulating techniques that yield stable pharmaceuticals [83]. Predictive models help choose the right solvent for crystallization, preventing undesirable solvates that could impair the effectiveness of medications. By computationally screening many solvents, researchers can optimize crystallization parameters for the desired solid form [84].

4.5 In Silico Applications in Polymorph Selection:

A key element of pharmaceutical research is polymorph selection because different crystal forms of an API might exhibit distinct physical, chemical, and biological characteristics. Using conventional experimental methods for polymorph screening takes a lot of time and money. In silico techniques like energy reduction, ML, and computer modeling have emerged as efficient alternatives to accelerate and optimize polymorph selection. This mini-review highlights the latest computational methods for using AI-driven rapid screening procedures, grading polymorph stability, simulating mechanical and dissolving properties, and forecasting polymorphic transitions [85–89].

The ability of a material to crystallize in several structural forms, each with its own stability, solubility, and bioavailability, is known as polymorphism. To guarantee consistent medication performance, the most thermodynamically and kinetically stable polymorph must be found. By forecasting stability, transformation conditions, and formulation behavior, computational techniques are essential for simplifying polymorph selection [90–94].

The relative stability of various polymorphic forms is determined using energy minimization approaches, such as DFT and molecular mechanics. Computational models can rank polymorphs and determine the best stable crystalline form by evaluating lattice energy, hydrogen bonding, and intermolecular interactions.

Molecular simulations and thermodynamic models are used to predict phase transitions between polymorphs under a range of temperature and pressure conditions. Solid-state stability control is enhanced by computational methods such as MD and metadynamics, which enable the identification of polymorphic transformation routes. For formulation development, in silico simulations provide important insights into mechanical properties such as compressibility, tabletability, and hardness. Dissolution modeling can be used to select the best medication candidates for increasing bioavailability and predict solubility behavior across a wide range of polymorphic forms. Machine learning techniques utilize large amounts of thermodynamic and crystallographic data to predict the polymorphism tendencies of APIs. Algorithms like neural networks and SVMs greatly lower the need for

experimental testing by identifying patterns in molecular structures and accurately predicting potential polymorphs.

The time and expense of experimental polymorph screening are decreased by in silico polymorph screening. Through the use of data-driven methods for finding stable polymorphs, the predicted accuracy is increased. This approach enables the documentation and assessment of risks for regulatory filings. By focusing on the most promising polymorph candidates, guided experimental design helps prioritize experimental screening efforts.

Challenges and Future Directions:

Despite substantial advances, computational polymorph screening still faces numerous hurdles. Current CSP approaches are limited in computing cost and force field precision, resulting in occasional differences between predicted and experimental structures. Machine learning models rely on big, high-quality datasets for training, which are not always available. Furthermore, polymorph stability is affected by kinetic parameters that are difficult to estimate just using thermodynamic calculations. Future research should prioritize enhancing AI algorithms, creating hybrid computational-experimental procedures, and incorporating real-time experimental feedback into prediction models.

V. CONCLUSION

Polymorphism is critical in pharmaceutical development, affecting medication solubility, stability, manufacturability, and regulatory acceptance. Understanding and regulating polymorphic forms ensures optimal medication performance, patient safety, and global regulatory compliance. Computational in silico technologies have revolutionized polymorph screening by making it possible to predict crystal shapes, solubility, and stability more quickly and economically. CSP, MD, DFT, and machine learning are all effective approaches for selecting the most stable polymorphs prior to experimental validation. As computational capabilities improve, combining AI-driven models with experimental methodologies will increase the efficiency and accuracy of polymorph screening. The pharmaceutical sector can speed up drug discovery, save prices, and enhance the formulation of life-saving pharmaceuticals by implementing these in silico technologies.

REFERENCES

- [1] Karpinski, P.H., (2006) Polymorphism of Active Pharmaceutical Ingredients. *Chem. Eng. Technol.*, 29, 233-237.
- [2] Censi R, Di Martino P. (2015) Polymorph Impact on the Bioavailability and Stability of

- Poorly Soluble Drugs. *Molecules*. 20 (10):18759-76.
- [3] Panda, Ramakant, S. Lankalapalli (2023). Bioavailability and polymorphic stability challenges affecting drug product's potential: a critical evaluation and pertinent solution. *Asian Journal of Pharmaceutical and Clinical Research*, 16, 11, 9-23
- [4] Shiyi Zhou, Tongtong Yang, Cheng Qian, Fei Wu, Yanlong Hong, Xiao Lin, (2024) An update on solid-state form, crystal modification and transition issues related to tablet manufacturing, *Journal of Drug Delivery Science and Technology*, 100, 106135.
- [5] Ana Simões, Francisco Veiga, and Carla Vitorino, (2024) Question-based review for pharmaceutical development: An enhanced quality approach, *European Journal of Pharmaceutics and Biopharmaceutics*, 195, 114174.
<https://www.pharmafocuseurope.com/articles/ritonavir-how-the-discovery-of-a-new-polymorph-changed-drug-development-forever>
- [6] Sanjay R. Chemburkar, John Bauer, Kris Deming, Harry Spiwek, Ketan Patel, John Morris, Rodger Henry, Stephen Spanton, Walter Dziki, William Porter, John Quick, Phil Bauer, John Donaubauer, B. A. Narayanan, Mauro Soldani, Dave Riley, Kathryn mcfarland, (2000) Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development. *Org. Proc. Res. Dev.* 4, 5, 413–417
- [7] Dedong Wu, Emma S.E. Eriksson, Sten O. Nilsson Lill, James F. McCabe, Christoph Bauer, Michelle L. Lamb, (2025) Discovery of the most stable form of an adenosine receptor antagonist through virtual polymorph screening and targeted crystallization, *Journal of Pharmaceutical Sciences*, 114 (2), 829-840.
- [8] Qin Shi, Haibiao Chen, Yanan Wang, Jia Xu, Ziyang Liu, Chen Zhang, (2022) Recent advances in drug polymorphs: Aspects of pharmaceutical properties and selective crystallization, *International Journal of Pharmaceutics*, 611, 121320.
- [9] Banga, S., Chawla, G., Varandani, D., et al. (2007). Modification of the crystal Habit of celecoxib for improved processability. *J. Pharm. Pharmacol.*, 59, 29–39.
- [10] Khomane, K.S. and Bansal, A.K. (2013) Weak hydrogen bonding interactions influence slip system activity and compaction behavior of pharmaceutical powders. *J. Pharm. Sci.*, 102, 4242–4245.
- [11] Joiris, E., Di Martino, P., Berneron, C., et al. (1998). Compression behavior of orthorhombic paracetamol. *Pharm. Res.* 15, 1122–1130.
- [12] Singhal, D. (2004). Drug polymorphism and dosage form design: a practical perspective. *Adv. Drug Delivery Rev.* 56, 335–347.
- [13] Picker-Freyer, K.M., Liao, X., Zhang, G., and Wiedmann, T.S. (2007) Evaluation of the compaction of sulfathiazole polymorphs. *J. Pharm. Sci.*, 96: 2111–2124.
- [14] Summers, M.P., Enever, R.P., and Carless, J.E. (1976). The influence of crystal form on the radial stress transmission characteristics of pharmaceutical materials. *J. Pharm. Pharmacol.* 28, 89–99.
- [15] Khomane, K.S. and Bansal, A.K. (2013) Effect of particle size on in-die and out-of-die compaction behavior of ranitidine hydrochloride polymorphs. *AAPS PharmSciTech*, 14, 1169–1177.
- [16] Yin, X.-Z., Wu, L., Li, Y., et al. (2016). Visualization and quantification of deformation behavior of clopidogrel bisulfate polymorphs during tableting. *Sci. Rep.*, 6: 21770.
- [17] Ann Newman. (2013). Specialized Solid Form Screening Techniques Org. Process Res. Dev. 17, 3, 457–471
- [18] Abramov, Yuriy. (2016) Computational Pharmaceutical Solid State Chemistry.
- [19] G. R. Desiraju. (2002) Cryptic crystallography. *Nature Materials*. 1 (2): 77–79.
- [20] S. M. Woodley, R. Catlow; Catlow. (2008) "Crystal structure prediction from first principles." *Nature Materials*. 7 (12): 937–946.
- [21] L. Pauling. (1929) The principles determining the structure of complex ionic crystals. *Journal of the American Chemical Society*. 51 (4): 1010–1026.
- [22] Martonak R., Laio A., Parrinello M. (2003) Predicting crystal structures: The Parrinello-Rahman method revisited. *Physical Review Letters*. 90 (3): 75502.
- [23] Martonak R., Donadio D., Oganov A. R., Parrinello M.; Donadio; Oganov; Parrinello. (2006) Crystal structure transformations in SiO₂ from classical and ab initio metadynamics. *Nature Materials*. 5 (8): 623–626.
- [24] Oganov, A. R.; Glass, C. W. (2006). Crystal structure prediction using ab initio evolutionary techniques: principles and applications. *Journal of Chemical Physics*. 124 (24): 244704.
- [25] Pickard, C. J.; Needs, R. J. (2006) High-Pressure Phases of Silane. *Physical Review Letters*. 97 (4): 045504.
- [26] Price, Sarah L. (2014) Predicting crystal structures of organic compounds. *Chemical Society Reviews*. 43 (7): 2098–2111.
- [27] Stone, Anthony. (2013) The Theory of Intermolecular Forces. Oxford University Press.
- [28] Nyman, Jonas; Day, Graeme M. (2015) Static and lattice vibrational energy differences

- between polymorphs. *Crystengcomm.* 17 (28): 5154–5165.
- [29] K. Sanderson. (2007) Model predicts structure of crystals. *Nature.* 450 (7171): 771.
- [30] Day, Graeme M.; Cooper, Timothy G.; Cruz-Cabeza, Aurora J.; Hejczyk, Katarzyna E.; Ammon, Herman L.; Boerrigter, Stephan X. M.; Tan, Jeffrey S.; Della Valle, Raffaele G.; Venuti, Elisabetta; Jose, Jovan; Gadre, Shridhar R.; Desiraju, Gautam R.; Thakur, Tejender S.; Van Eijck, Bouke P.; Facelli, Julio C.; Bazterra, Victor E.; Ferraro, Marta B.; Hofmann, Detlef W. M.; Neumann, Marcus A.; Leusen, Frank J. J.; Kendrick, John; Price, Sarah L.; Misquitta, Alston J.; Karamertzanis, Panagiotis G.; Welch, Gareth W. A.; Scheraga, Harold A.; Arnautova, Yelena A.; Schmidt, Martin U.; Van De Streek, Jacco; et al. (2009) Significant progress in predicting the crystal structures of small organic molecules—a report on the fourth blind test. *Acta Crystallographica B.* 65 (Pt 2): 107–125.
- [31] M. A. Neumann, F. J. J. Leusen, J. Kendrick; Leusen; Kendrick. (2008) A Major Advance in Crystal Structure Prediction. *Angewandte Chemie International Edition.* 47 (13): 2427–2430.
- [32] Reilly, Anthony M.; Cooper, Richard I.; Adjiman, Claire S.; Bhattacharya, Saswata; Boese, A. Daniel; Brandenburg, Jan Gerit; Bygrave, Peter J.; Bylsma, Rita; Campbell, Josh E.; Car, Roberto; Case, David H.; Chadha, Renu; Cole, Jason C.; Cosburn, Katherine; Cuppen, Herma M.; Curtis, Farren; Day, Graeme M.; Stasio, Robert A.; Dzyabchenko, Alexander; Van Eijck, Bouke P.; Elking, Dennis M.; Van Den Ende, Joost A.; Facelli, Julio C.; Ferraro, Marta B.; Fusti-Molnar, Laszlo; Gatsiou, Christina Anna; Gee, Thomas S.; De Gelder, Rene; Ghiringhelli, Luca M.; et al. (2016) Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallographica B.* 72 (4): 439–459.
- [33] Dybeck, Eric C.; Abraham, Nathan S.; Schieber, Natalie P.; Shirts, Michael R. (2017) Capturing Entropic Contributions to Temperature-Mediated Polymorphic Transformations Through Molecular Modeling. *Journal of Chemical Theory and Computation.* 17 (4): 1775–1787.
- [34] Josef M. Gallmetzer, Jakob Gamper, Stefanie Kröll, Stefanie Kröll, Thomas S. Hofer, (2025) Comparative Study of UMCM-9 Polymorphs: Structural, Dynamic, and Hydrogen Storage Properties via Atomistic Simulations, *J. Phys. Chem. C*, 129, 11, 5645–5655
- [35] Arturo Sauza-de la Vega a, Leonardo J. Duarte bc, Arnaldo F. Silva b, Jonathan M. Skelton d, Tomás Rocha-Rinza a and Paul L. A. Popelier , (2022) Towards an atomistic understanding of polymorphism in molecular solids, *Phys. Chem. Chem. Phys.*, 24, 11278–11294
- [36] Zuoxuan Zhu, Yuan Zhang, Zhixuan Wang, Weiwei Tang, Jingkang Wang, and Junbo Gong. (2024) Artificial Intelligence Assisted Pharmaceutical Crystallization, *Cryst. Growth Des.* 24, 10, 4245–4270
- [37] Firaha, D., Liu, Y.M., van de Streek, J., et al. (2023). Predicting crystal form stability under real-world conditions. *Nature*, 623, 324–328.
- [38] Rafael Barbas, Mercè Font-Bardia, Anant Paradkar, Christopher A. Hunter, and Rafael Prohens. (2018) Combined Virtual/Experimental Multicomponent Solid Forms Screening of Sildenafil: New Salts, Cocrystals, and Hybrid Salt–Cocrystals, *Cryst. Growth Des.* 18, 12, 7618–7627.
- [39] Sean Askin, Jeremy K. Cockcroft, Louise S. Price, Andrea D. Gonçalves, Min Zhao, Derek A. Tocher, Gareth R. Williams, Simon Gaisford, Duncan Q. M. Craig, (2019) Olanzapine Form IV: Discovery of a New Polymorphic Form Enabled by Computed Crystal Energy Landscapes. *Cryst. Growth Des.* 19, 5, 2751–2757.
- [40] Francesca Cenci, Samir Diab, Paola Ferrini, Catajina Harabaiju, Massimiliano Barolo, Fabrizio Bezzo, and Pierantonio Facco (2024), Predicting drug solubility in organic solvent mixtures: A machine-learning approach supported by high-throughput experimentation, *International Journal of Pharmaceutics*, 660, 124233.
- [41] Nastaran Meftahi, Michael L. Walker, and Brian J. Smith, (2021) Predicting aqueous solubility by QSPR modeling, *Journal of Molecular Graphics and Modelling*, 106, 107901.
- [42] Cristianini, N., Ricci, E. (2008). Support Vector Machines. In: Kao, MY. (eds.) *Encyclopedia of Algorithms*. Springer, Boston, MA.
- [43] Ho, Tin Kam. (1995) Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, QC, 14–16, 278–282.
- [44] Ho TK. (1998) The Random Subspace Method for Constructing Decision Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 20 (8): 832–844.
- [45] Shakhawath Hossain, Aleksei Kabedev, Albin Parrow, Christel A.S. Bergström, Per Larsson, (2019) Molecular simulation as a computational pharmaceutics tool to predict drug solubility, solubilization processes and partitioning, *European Journal of Pharmaceutics and Biopharmaceutics*, 137, 46–55

- [46] Lunna Li, Tim Totton, and Daan Frenkel. (2017). Computational methodology for solubility prediction: Application to the sparingly soluble solutes. *J. Chem. Phys.* 146, 214110.
- [47] Sumin Lee, Myeonghun Lee, Ki-Won Gyak, Sung Dug Kim, Mi-Jeong Kim, Kyoungmin Min, (2022) Novel Solubility Prediction Models: Molecular Fingerprints and Physicochemical Features vs Graph Convolutional Neural Networks, *ACS Omega*, 7, 14, 12268–12277
- [48] Ahmad, W.; Tayara, H.; Shim, H.; Chong, K.T. (2024) solpredictor: Predicting Solubility with Residual Gated Graph Neural Network. *Int. J. Mol. Sci.* 25, 715.
- [49] Renato Contreras, José Muñoz-Espinoza, and Bruno Sánchez (2024). Theoretical models of solubility for organic solvents. A conceptual density functional theory approach, *Journal of Molecular Liquids*, 403, 124736.
- [50] Isabella W. Cordova, Gabriel Teixeira, Paulo J. A. Ribeiro-Claro, Dinis O. Abranches, Simão P. Pinho, Olga Ferreira, João A. P. Coutinho, (2024) Using Molecular Conformers in COSMO-RS to Predict Drug Solubility in Mixed Solvents, *Ind. Eng. Chem. Res.* 63, 21, 9565–9575
- [51] Akosua B. Anane-Adjei, Esther Jacobs, Samuel C. Nash, Sean Askin, Ramesh Soundararajan, Mary Kyobula, Jonathan Booth, and Andrew Campbell, (2022) Amorphous solid dispersions: Utilization and challenges in preclinical drug development within AstraZeneca, *International Journal of Pharmaceutics*, 614, 121387.
- [52] George Zografi and Ann Newman, (2017) Interrelationships between Structure and the Properties of Amorphous Solids of Pharmaceutical Interest, *Journal of Pharmaceutical Sciences*, 106, 1, 5-27.
- [53] Kirk Swanson, Shubhendu Trivedi, Joshua Lequieu, Kyle Swanson, Risi Kondor. (2020) Deep learning for automated classification and characterization of amorphous materials. *Soft Matter*, 16, 435-446.
- [54] Bookwala, M., DeBoyace, K., Buckner, I.S., et al. (2020). Predicting Density of Amorphous Solid Materials Using Molecular Dynamics Simulation. *AAPS PharmSciTech*, 21, 96.
- [55] Tao, J.L., Wang, Y.J. (2024) An Artificial Intelligence Constitutive Model for Amorphous Solids Utilizing Graph Neural Networks. *JOM*, 76, 5777–5784
- [56] Ata Madanchi, Emna Azek, Karim Zongo, Laurent K. Béland, Normand Mousseau, Lena Simine, (2025) Is the Future of Materials Amorphous? Challenges and Opportunities in Simulations of Amorphous Materials, *ACS Phys. Chem Au*, 5, 1, 3–16.
- [57] Loschen C, Klamt A. (2016) Computational Screening of Drug Solvates. *Pharm Res* 33(11):2794-804.
- [58] Ipsha Shruti, Mubarak Almehairbi, Zeinab M. Saeed, Tamador Alkhalid, Wesam A. Ali, Rampal Vishwakarma, Sharmarke Mohamed, and Deepak Chopra (2022). Unraveling the Origin of Solvate Formation in the Anticancer Drug Trametinib: Insights from Crystal Structure Analysis and Computational Modeling. *Cryst. Growth Des.* 22, 10, 5861–5871
- [59] Asif A. Malik, Zeinab M. Saeed, Ishtiyah Ahmad, Tamador Alkhalid, Praveen B. Managutti, Sharmarke Mohamed, and Aijaz A. Dar. (2024) Polymorphism Steered Thermochromism in a Sulfonamide. *ACS Applied Optical Materials*, 2 (8) , 1709-1720.
- [60] Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. (2008) Towards Prediction of Stoichiometry in Crystalline Multicomponent Complexes. *Chem.—Eur. J.* 14, 8830–8836
- [61] Mohamed, S.; Li, L. (2018) From serendipity to supramolecular design: assessing the utility of computed crystal form landscapes in inferring the risks of crystal hydration in carboxylic acids. *CrystEngComm*, 20, 6026–6039
- [62] Braun, D. E.; Bhar dwaj, R. M.; Florence, A. J.; Tocher, D. A.; Price, S. L. (2013) Complex Polymorphic System of Gallic Acid-Five Monohydrates, Three Anhydrides, and over 20 Solvates. *Cryst. Growth Des.* 13, 19–23
- [63] Abramov YA, Loschen C, Klamt A. (2012) Rational coformer or solvent selection for pharmaceutical cocrystallization or desolvation. *J Pharm Sci.* 101(10):3687-97.
- [64] Cavallari C, Santos BP, Fini A. (2013) Olanzapine solvates. *J Pharm Sci.* 102(11):4046-56.
- [65] Tshepelevitsh S, Hernits K, Leito I. (2018) Prediction of partition and distribution coefficients in various solvent pairs with COSMO-RS. *J Comput Aided Mol Des.* 32(6):711-722.
- [66] Zhang X, Wang C, Zhou L, Yang W, Zhou L, Bao Y, Zhang M, Hou B, Xu Z, Yin Q. (2018) Effects of Hydrogen Bond Acceptor Ability of Solvents on Molecular Self-Assembly of Sulfadiazine Solvates. *J Pharm Sci.* 107(11):2823-2828.
- [67] Shaikh TR, Shelke N, Tothadi S. (2023) Multicomponent Solvate Crystals of 3,5-Dinitrobenzoic Acid and Acetamide and CSD Analysis of Solvates. *ACS Omega.* 8(27):24644-24653.

- [68] Ying Yang, Bingjie Hu, Markus A. Lill, (2014) Analysis of Factors Influencing Hydration Site Prediction Based on Molecular Dynamics Simulations, *J. Chem. Inf. Model.* 54, 10, 2987–2995
- [69] Healy, A. M., et al. (2002). Solid-state transformations of drugs: mechanochemical approaches to solvate formation. *J. Pharm. Sci.*, 91(3), 884–891.
- [70] Marenich, A. V., Cramer, C. J., Truhlar, D. G. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B*, 113(18), 6378–6396
- [71] Llinàs, A., Goodman, J. M. (2008). Polymorph control: past, present, and future. *Drug Discovery Today*, 13(5-6), 198–210.
- [72] Shihui Sun, Jian-Rong Wang, Lei Tang, Zhehuan Fan, Xutong Li, Zifu Huang, Weizhong Han, Qinggong Wang, Jie Yu, Zhiyin Xie, Wenzhi Yang, Zeen Yang, Mingyue Zheng, Dingyan Wang, Xuefeng Mei, Xiaomin Luo, (2025) Prediction of Solvate Based on Graph Attention Network, *Cryst. Growth Des.* 25, 2, 297–308
- [73] Wang, H.; Lian, D.; Zhang, Y.; Qin, L.; Lin, X. Gognn: (2020) Graph of graphs neural network for predicting structured entity interactions, *IJCAI-20*, 7, 1317–1323.
- [74] Wang, Y.; Min, Y.; Chen, X.; Wu, J. (2021) Multi-view Graph Contrastive Representation Learning for Drug-Drug Interaction Prediction. In *Proceedings of the Web Conference 2021*, ACM, 2921–2933.
- [75] Takieddin, K.; Khimyak, Y. Z.; Fábrián, L. (2016) Prediction of Hydrate and Solvate Formation Using Statistical Models. *Cryst. Growth Des.* 16 (1), 70–81
- [76] Bhardwaj, R. M.; Reutzel-Edens, S. M.; Johnston, B. F.; Florence, A. J. (2018) A random forest model for predicting crystal packing of olanzapine solvates. *CrystEngComm* 20 (28), 3947–3950
- [77] Johnston, A.; Johnston, B. F.; Kennedy, A. R.; Florence, A. J. (2008) Targeted crystallisation of novel carbamazepine solvates based on a retrospective Random Forest classification. *CrystEngComm*, 10 (1), 23–25,
- [78] Jia, L. H.; Guo, M. X.; Dai, J. Y.; Chen, W.; Zhou, L.; Yin, Q. X. (2023) Assessing Solvate Prediction Approaches: A Case of Spironolactone. *Cryst. Growth Des.* 23 (2), 832–841.
- [79] Loschen C, Klamt A. (2015) Solubility prediction, solvate and cocrystal screening as tools for rational crystal engineering. *J Pharm Pharmacol.* 67(6):803-11.
- [80] Tavagnacco L, Schnupf U, Mason PE, Sabounji ML, Cesàro A, Brady JW. (2011) Molecular dynamics simulation studies of caffeine aggregation in aqueous solution. *J Phys Chem B.* 115(37):10957-66.
- [81] Xie, Y.; Shi, G.; Sun, J.; Li, S.; Gao, W.; Hu, Y.; Zu, C.; Tang, W.; Gong, J. (2022) Computational Screening and Experimental Validation on Multicomponent Crystals of a New Class of Janus Kinase (JAK) Inhibitor Drug with Improved Solubility. *Crystals*, 12, 1722
- [82] Karfunkel, H. R.; Gdanitz, R. J. (1992) Ab Initio prediction of possible crystal structures for general organic molecules. *J. Comput. Chem.* 13 (10), 1171–1183
- [83] Spek, A. (2009) Structure validation in chemical crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* 65, 148–155,
- [84] Mohamed, S. (2018) Integrating Computed Crystal Energy Landscapes in Crystal Form Discovery and Characterisation, in *Understanding Intermolecular Interactions in the Solid State: Approaches and Technique*, 1–31
- [85] Wicker, J. G. P.; Crowley, L. M.; Robshaw, O.; Little, E. J.; Stokes, S. P.; Cooper, R. I.; Lawrence, S. E. (2017) Will they co-crystallize? . *CrystEngComm.* 19, 5336–5340
- [86] Cruz-Cabez5336–5340Taylor, E.; Sugden, I. J.; Bowskill, D. H.; Wright, S. E.; Abdullahi, H.; Tulegenov, D.; Sadiq, G.; Davey, R. J. (2020) Can solvated intermediates inform us about nucleation pathways? The case of β -pABA. *CrystEngComm*, 22, 7447– 7459,
- [87] Hoja, Johannes and Ko, Hsin-Yu and Neumann, Marcus A. and Car, Roberto and DiStasio, Robert A. and Tkatchenko, Alexandre, (2019) Reliable and practical computational description of molecular crystal polymorphs, *Sci. Adv.* 5, eaau3338.
- [88] Zhou, D., Bier, I., Santra, B. et al. (2025) A robust crystal structure prediction method to support small molecule drug development with large scale validation and blind study. *Nat Commun*, 16, 2210.
- [89] Gregory J. O. Beran, (2023) Frontiers of molecular crystal structure prediction for pharmaceuticals and functional organic materials, *Chem. Sci.*,14, 13290-13312
- [90] Vasileiadis, M., Pantelides, C. C. & Adjiman, C. S. (2015) Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule. *Chem. Eng. Sci.* 121, 60–76
- [91] Gelbrich T, Braun DE, Ellern A, Griesser UJ. (2013) Four polymorphs of methyl paraben: structural relationships and relative energy differences. . *Cryst. Growth Des.* 13,1206–17

-
- [92] Pantelides CC, Adjiman CS, Kazantsev AV. (2014) General computational algorithms for ab initio crystal structure prediction for organic molecules. In Prediction and Calculation of Crystal Structures, ed. S Atahan-Evrenk, A Aspuru-Guzik, 22–58. Berlin:: Springer
- [93] Podeszwa R, Rice BM, Szalewicz K. (2009) Crystal structure prediction for cyclotrimethylene trinitramine (RDX) from first principles. Phys. Chem. Chem. Phys. 11::5512–18
- [94] Martoňák R, Laio A, Parrinello M. (2003) Predicting crystal structures: the Parrinello–Rahman method revisited. Phys. Rev. Lett. 90::075503