Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

Pharmacophore
an International Journal
ISSN 2229-5402

PREDICTING AMORPHOUS SOLID DISPERSION PERFORMANCE USING MISCIBILITY, GLASS TRANSITION, AND DISSOLUTION DATA

Sanjay Kulkarni[1], Meenal Joshi[1]*, Rohan Patil[2], Aniket Deshmukh[2]

1. *Department of AI in Pharmaceutical Systems, Faculty of Pharmacy, Savitribai Phule Pune University, Pune, India.*
2. *Department of Computational Drug Discovery, Faculty of Pharmaceutical Technology, IIT Bombay, Mumbai, India.*

ARTICLE INFO

Received:
05 February 2025
Received in revised form:
04 May 2025
Accepted:
06 May 2025
Available online:
28 June 2025

Keywords: Amorphous solid dispersion, Predictive modelling, Drug–polymer miscibility, Glass transition temperature, Dissolution, Supersaturation

ABSTRACT

Amorphous solid dispersions can improve the oral delivery potential of poorly water-soluble drugs by stabilising the drug in a high-energy amorphous state. Their performance depends on drug–polymer miscibility, thermal mobility, and the ability to generate and maintain supersaturation during dissolution. Formulation screening remains strongly empirical, and individual measurements such as a single glass transition temperature or visual evidence of miscibility rarely provide a complete performance forecast. This limits rational formulation design because stability, dissolution, and precipitation are often interpreted separately. This manuscript describes a conceptual predictive model for estimating amorphous solid dispersion performance from miscibility, glass transition, and dissolution descriptors. The intended outputs are crystallization stability, supersaturation behaviour, and dissolution profile quality. A gradient-boosted regression framework is proposed using formulation-level inputs such as interaction parameters, measured or predicted glass transition temperature, drug loading, polymer characteristics, and early dissolution metrics. The model is intended as a decision-support tool rather than a replacement for experimental confirmation. Conceptually, the model could predict whether an amorphous solid dispersion would be expected to remain physically stable and whether it should maintain a useful supersaturation profile. It could also identify formulation variables most responsible for predicted failure or success. A predictive modelling workflow of this type could reduce the experimental burden of amorphous solid dispersion development by prioritising a smaller set of rational formulation candidates. The approach supports earlier, more integrated decision-making in amorphous formulation design.

To Cite This Article: Kulkarni S, Joshi M, Patil R, Deshmukh A. Predicting Amorphous Solid Dispersion Performance Using Miscibility, Glass Transition, and Dissolution Data. Pharmacophore. 2025;16(3):12-21. https://doi.org/10.51847/Oy5zb144XZ

Introduction

Poor aqueous solubility remains a central challenge in oral drug development, and amorphous solid dispersion technology is widely used to improve apparent solubility and dissolution-driven exposure. The formulation concept relies on converting a crystalline drug into an amorphous state while using a polymeric carrier to suppress recrystallization and maintain dissolution advantage, a design space that has been analysed through both thermodynamic and kinetic perspectives [1]. Because amorphous solubility advantage is not automatically translated into durable formulation performance, predictive models must consider stability and dissolution together rather than treating them as separate development questions [2]. A model-oriented ASD framework therefore begins with the recognition that bioavailability enhancement depends on both maintaining the amorphous phase and sustaining useful supersaturation after administration [3].

Current practice often depends on heuristic interpretation of miscibility, thermal analysis, and dissolution tests, but these measurements are not always combined into a unified quantitative prediction. Differential scanning calorimetry can provide glass transition information, yet the interpretation of Tg depends on composition, measurement method, and the extent to which the drug and polymer form a single amorphous phase [4]. Miscibility assessment is similarly complex because solubility parameters, Flory-Huggins interaction estimates, melting point depression, and phase diagrams may lead to different levels of confidence in drug–polymer compatibility [5]. Dissolution assessment adds another layer of complexity because non-sink

Corresponding Author: Sanjay Kulkarni. Department of AI in Pharmaceutical Systems, Faculty of Pharmacy, Savitribai Phule Pune University, Pune, India. E-mail: meenal.joshi@gmail.com

dissolution, supersaturation, and precipitation behaviour may reflect formulation performance more directly than conventional sink testing [6].

Recent ASD literature further supports the need for a predictive framework that links formulation composition, process route, polymer function, and solution-state performance in one model. Manufacturing-focused reviews show that hot-melt extrusion, spray drying, and related ASD preparation strategies can strongly influence downstream physical stability, meaning that preparation method should be treated as a formulation descriptor rather than as background information [7]. Polymer-centered reviews also emphasize that the carrier affects nucleation, crystal growth, phase separation, dissolution, and supersaturation maintenance, reinforcing the need to encode polymer properties beyond nominal polymer identity [8, 9]. At the same time, foundational miscibility work shows that drug–polymer solubility and miscibility are central to stabilization, while supersaturation theory indicates that dissolution advantage depends on maintaining a metastable concentration state that may influence oral absorption [10, 11]. Together, these studies justify a model architecture that does not predict ASD success from Tg, miscibility, or dissolution alone, but instead combines these domains into a unified performance estimate.

Machine learning has emerged as a useful conceptual tool for formulation science because it can integrate heterogeneous descriptors and learn non-linear relationships that are difficult to encode in simple rules. Earlier ASD-focused models have explored the prediction of physical stability and formulation behaviour using machine learning methods, demonstrating that formulation descriptors can be mapped to stability-related outcomes when curated datasets are available [12]. More recent work has extended this direction toward chemically stable hot-melt extrusion systems, high-throughput screening workflows, and deep learning approaches for imbalanced ASD stability classes [13–15]. However, a gap remains for a single predictive framework that jointly considers miscibility, Tg, and dissolution-derived performance indicators rather than focusing on one endpoint in isolation.

The central thesis of this manuscript is that a machine learning model trained on miscibility, glass transition, and dissolution descriptors could predict the long-term physical stability and dissolution performance of an amorphous solid dispersion. Such a model would not replace confirmatory studies, but it could provide a rational basis for selecting polymer type, drug loading, and processing route before resource-intensive formulation work begins [16]. By linking drug–polymer interaction descriptors to measured or predicted Tg and early dissolution behaviour, the model could rank candidate ASDs according to expected crystallization risk and supersaturation maintenance [17]. This creates a formulation-oriented decision framework in which empirical screening is guided by interpretable predictions rather than by isolated measurements.

Background

Amorphous Solid Dispersions and Their Instability Mechanisms

Amorphous solid dispersions are thermodynamically unstable systems because the amorphous drug possesses higher free energy than its crystalline counterpart, creating a driving force for molecular reorganisation and crystallization. Polymer carriers are used to slow this process by reducing molecular mobility, disrupting drug–drug interactions, and modifying nucleation and crystal growth pathways [18]. The same formulation must also perform during dissolution, where the drug may generate a spring-like supersaturation profile before polymer-mediated inhibition of precipitation provides the parachute effect [3]. Predictive modelling must therefore recognise that physical stability and dissolution advantage arise from related but non-identical mechanisms.

Drug–Polymer Miscibility and Glass Transition Temperature

Drug–polymer miscibility is commonly evaluated through thermodynamic and experimental indicators, including solubility parameter differences, Flory-Huggins interaction parameters, melting point depression, and temperature-composition phase behaviour [19]. These measurements are important because a miscible system is more likely to exhibit a single glass transition and reduced crystallization tendency under appropriate storage conditions [20]. Mixture Tg can be estimated using relationships such as the Gordon-Taylor or Couchman-Karasz equations, but experimental Tg remains sensitive to composition, preparation history, and analytical method [4]. In a predictive model, miscibility descriptors and Tg should therefore be treated as complementary features rather than interchangeable indicators of stability.

Dissolution Methodology for ASD Performance Assessment

Dissolution testing for ASDs should capture supersaturation generation, precipitation inhibition, and the duration of elevated apparent solubility rather than only the final dissolved fraction. Non-sink dissolution methods are especially relevant because they can reveal the maximum concentration, supersaturation ratio, and precipitation lag behaviour that better reflect the intended function of an ASD [6]. Studies of supersaturation maintenance using thermal and spectroscopic approaches have shown that dissolution behaviour depends on the interplay among drug amorphousness, polymer interactions, and precipitation kinetics [3]. These dissolution-derived variables provide natural model inputs because they translate formulation structure into performance-relevant kinetic signatures.

Prior Modeling Efforts for ASD Stability

Prior modelling efforts have included empirical stability rules, thermodynamic estimates of miscibility, molecular descriptor-based predictions, and machine learning workflows for ASD stability. Machine learning models have been used to predict the physical stability of solid dispersions, while other approaches have focused on chemical stability or high-throughput formation screening [12, 14, 15]. At the same time, thermodynamic critiques have cautioned that simple regular-solution assumptions and solubility-parameter methods may be insufficient when used alone for amorphous dispersion design [19, 21]. A more comprehensive predictive framework should therefore combine mechanistic descriptors with supervised learning rather than relying on a single theory or empirical threshold.

Supervised Learning for Formulation Design

Supervised learning is attractive for ASD formulation design because it can map formulation descriptors to observed stability or dissolution endpoints while accommodating non-linear interactions among drug loading, polymer type, miscibility, and Tg. Gradient boosting and related ensemble methods are particularly suitable conceptually because they can handle mixed numerical and categorical features, tolerate missingness more gracefully than many parametric models, and provide interpretable feature-importance information [12]. Similar machine learning strategies have been applied in pharmaceutical materials contexts such as spray-dried dispersion particle-size prediction and dissolution profile modelling, supporting their relevance to formulation-oriented prediction [22, 23]. For small formulation datasets, the model should be regularised, cross-validated, and interpreted cautiously to avoid converting dataset bias into design rules.

Model Development Overview

High-Level Prediction Pipeline

The proposed model begins with a formulation vector containing drug identity, polymer identity, drug loading, miscibility index, measured or predicted Tg, storage descriptors, and dissolution test variables. This vector is transformed into predicted performance outputs, including time-to-crystallization category, expected maximum supersaturation, precipitation tendency, and relative dissolution advantage, reflecting the multi-dimensional nature of ASD success [24]. The pipeline is conceptually aligned with previous attempts to use machine learning for ASD physical stability while extending the feature space to include dissolution and miscibility-derived descriptors [12]. The output should be interpreted as a decision-support estimate that prioritises formulations for confirmation rather than as a definitive claim of product performance.

Figure 1 illustrates the proposed end-to-end predictive workflow linking drug–polymer miscibility, glass transition behavior, dissolution kinetics, model interpretation, and formulation decision support for amorphous solid dispersion development.

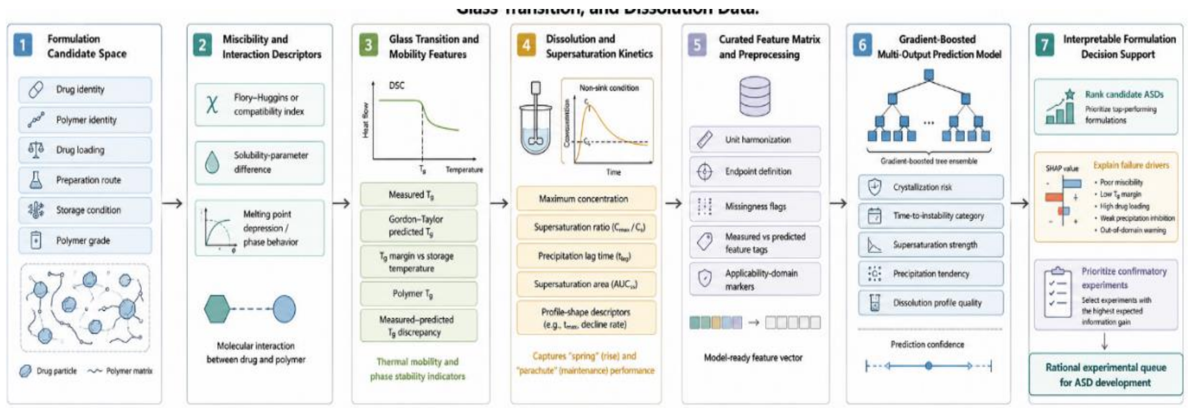

Figure 1. Integrated Predictive Modeling Workflow for Amorphous Solid Dispersion Performance Using Miscibility, Glass Transition, and Dissolution Data

Core Input Features

Core input features should include a drug–polymer interaction parameter or surrogate, the experimental Tg of the dispersion, the Gordon-Taylor predicted Tg, drug loading, polymer class, polymer Tg, hygroscopicity-related descriptors, and molecular-weight category. When direct Flory-Huggins parameters are unavailable, solubility parameter differences, melting point depression behaviour, or phase diagram descriptors can provide surrogate estimates of compatibility [5, 20]. Dissolution features should include maximum concentration, supersaturation ratio, precipitation lag time, and profile shape descriptors because these variables connect formulation structure to biopharmaceutical performance [6]. These inputs create a physically meaningful feature set in which miscibility, mobility, and dissolution kinetics are represented together.

Design Principles

The model should be designed for small, heterogeneous formulation datasets because ASD data are often compiled from literature reports, internal screens, and methodologically diverse experiments. This requires careful standardisation of units, storage conditions, dissolution media, polymer naming, and endpoint definitions before supervised learning can be meaningful [16]. Uncertainty estimation should be included so that predictions for sparse chemical regions are flagged as lower confidence rather than presented as equally reliable. Interpretability is also essential because formulators need to understand whether a prediction is driven mainly by drug loading, poor miscibility, low Tg, or an unfavourable dissolution profile [13].

Data Sources and Feature Engineering

Curating an ASD Performance Dataset

A curated ASD performance dataset would combine published reports and internal laboratory records containing drug–polymer combinations, preparation method, drug loading, stability observations, and dissolution outcomes. Published studies provide valuable examples of physical stability modelling, excipient selection, miscibility assessment, and dissolution behaviour, but their methods and reporting formats are heterogeneous [12, 16, 17]. Internal records may add formulation diversity and negative examples, although they must be harmonised with public data before model training. The curation process should prioritise traceable metadata such as storage temperature, relative humidity, polymer grade, preparation route, and dissolution medium because these variables can strongly influence apparent performance [6].

Encoding Miscibility and Thermal Features

Miscibility and thermal features should be encoded in a way that preserves their physicochemical meaning while allowing the model to learn non-linear effects. Directly measured Flory-Huggins parameters, melting point depression results, and phase-diagram classifications can be used when available, while solubility-parameter differences or computational compatibility scores can serve as secondary descriptors [21, 25]. Experimental Tg should be included alongside Gordon-Taylor or Couchman-Karasz predictions because disagreement between measured and predicted values may signal incomplete miscibility, phase separation, or measurement sensitivity [4]. Treating the Gordon-Taylor k-value as a feature could allow the model to learn formulation-specific deviations from ideal mixture behaviour.

Table 1 organizes the proposed amorphous solid dispersion feature space according to the physicochemical role, predictive meaning, and formulation decision relevance of each descriptor domain.

Table 1. Integrated ASD Feature Architecture Linking Formulation Inputs to Predictive Meaning

Feature domain	Representative variables	Physicochemical meaning	Predictive contribution	Formulation decision value
Drug–polymer composition	Drug identity; polymer identity; drug loading; polymer grade; preparation route	Defines the formulation system and the amount of stabilizing carrier available to suppress drug crystallization	Allows the model to learn composition-dependent stability and dissolution behavior across drug–polymer combinations	Helps identify whether performance risk is mainly driven by excessive drug loading, unsuitable polymer selection, or route-specific formulation limitations
Miscibility and interaction descriptors	Flory–Huggins interaction parameter; solubility-parameter difference; melting point depression; phase-diagram class; compatibility score	Represents thermodynamic compatibility and likelihood of forming a single amorphous phase	Supports prediction of phase separation, crystallization risk, and drug–polymer stabilization potential	Guides polymer screening by distinguishing promising compatibility signals from weak or uncertain miscibility evidence
Glass transition and molecular mobility	Experimental Tg; Gordon–Taylor predicted Tg; Couchman–Karasz predicted Tg; Tg margin relative to storage temperature; measured–predicted Tg discrepancy	Captures molecular mobility, plasticization risk, and deviation from ideal mixing behavior	Helps estimate whether the dispersion is likely to remain kinetically trapped under storage conditions	Supports selection of polymer fraction, storage conditions, and formulations requiring additional stability safeguards
Polymer functional attributes	Polymer Tg; hygroscopicity; molecular-weight class; hydrogen-bonding capacity; precipitation-inhibition tendency	Represents stabilizing and dissolution-modifying behavior of the carrier beyond nominal polymer name	Improves transferability across related excipients and reduces dependence on simple categorical polymer labels	Enables rational comparison of polymer grades and chemically related excipients before experimental screening
Dissolution and supersaturation descriptors	Maximum concentration; supersaturation ratio; supersaturation area; precipitation lag time; concentration–time profile shape	Describes whether the ASD generates and maintains useful apparent solubility during non-sink dissolution	Adds performance prediction beyond physical stability by modeling spring-and-parachute behavior	Helps avoid selecting formulations that are physically stable but biopharmaceutically weak
Storage and environmental descriptors	Storage temperature; relative humidity; accelerated stability condition; moisture-risk flag; storage duration	Captures environmental stressors that may reduce Tg, increase molecular mobility, or alter polymer interactions	Improves prediction of humidity-driven or temperature-driven instability	Supports interpretation of whether failure risk is formulation-intrinsic or storage-condition dependent

Data provenance and uncertainty markers	Literature vs internal source; measured vs predicted feature flag; missingness indicator; method-normalization flag; applicability-domain score	Preserves information about data reliability, measurement origin, and extrapolation risk	Prevents the model from treating all descriptors as equally certain or equally comparable	Helps formulators decide when a prediction is actionable and when confirmatory data are needed first

Defining Performance Labels

Performance labels should represent both physical stability and dissolution function because an ASD that remains amorphous but fails to generate useful supersaturation may not be an enabling formulation. Stability may be expressed conceptually as a binary class at a defined condition, an ordinal risk category, or a continuous time-to-recrystallization endpoint, while dissolution labels may capture maximum concentration, supersaturation area, and precipitation kinetics [3, 24]. Literature on rapid physical stability assessment and amorphous solubility advantage supports the need to define labels in ways that distinguish thermodynamic potential from kinetic durability [2, 26]. A multi-output labelling strategy would allow the model to predict whether a formulation is expected to be stable, dissolve effectively, and resist precipitation.

Predictive Model Architecture

Model Choice – Gradient-Boosted Trees

Gradient-boosted tree models such as XGBoost or LightGBM are conceptually suitable for ASD prediction because they can model non-linear interactions among drug loading, polymer properties, miscibility descriptors, Tg, and dissolution metrics. Ensemble models have already been used in ASD physical stability prediction, and related machine learning applications in formulation development suggest that tree-based methods can be practical for mixed descriptor sets [12, 23]. Their ability to handle missing or partially observed features is useful for literature-derived formulation datasets, where dissolution, Tg, and miscibility information are not always reported together. A multi-output implementation could estimate stability class and dissolution-related endpoints within a single formulation-ranking workflow.

Feature Vector Assembly and Pre-processing

Feature vector assembly should begin with continuous variables such as drug loading, Tg, miscibility descriptors, dissolution concentrations, and precipitation-time descriptors, followed by categorical variables such as polymer type, preparation method, and storage condition. Polymer identity can be encoded using physicochemical properties rather than only nominal categories, because polymer Tg, hygroscopicity, and molecular-weight descriptors may explain transferability across related excipients [5]. Missing dissolution or miscibility values should not be filled indiscriminately; instead, imputation should be constrained within chemically or methodologically comparable groups to avoid introducing artificial correlations. Pre-processing should also preserve uncertainty flags so the model can distinguish measured values from predicted or surrogate descriptors [27].

Output: Stability and Performance Metrics

The proposed model output should combine a stability probability or risk category with dissolution performance descriptors such as expected supersaturation strength and precipitation tendency. Prediction intervals derived through bootstrap, quantile regression, or related uncertainty methods would be useful because ASD datasets are often limited and chemically uneven [13]. This output structure would support ranking of candidate formulations while also indicating which predictions require cautious interpretation or confirmatory testing. By pairing physical stability prediction with dissolution behaviour, the model would better reflect the practical definition of ASD success than stability-only rules or dissolution-only screens [17].

Handling Limited Data and Physicochemical Constraints

Data Augmentation with Physicochemical Priors

Because ASD datasets are often sparse, physicochemical priors should be used as model inputs rather than treated as substitutes for experimental validation. Gordon-Taylor Tg estimates, Flory-Huggins interaction parameters, solubility-parameter differences, and phase-diagram classifications can enrich the feature space when direct measurements are unavailable or incomplete [5, 19]. Melting point depression and temperature-composition mapping can also provide structured information about miscibility boundaries that helps the model distinguish likely single-phase dispersions from systems at risk of phase separation [20, 28]. These priors should be encoded transparently so that the model can learn from theory-informed descriptors while preserving the distinction between measured and inferred information.

Transfer Learning Across Polymer Types

Transfer learning across polymer types would allow the model to share information among chemically related excipients while avoiding the assumption that all polymers behave identically. Polymer embeddings or hierarchical descriptors could represent shared attributes such as hydrogen-bonding capacity, glass transition temperature, hygroscopicity, and molecular-weight class, enabling predictions for less-studied grades to borrow strength from better-characterised ones [29]. This strategy is consistent with formulation studies showing that polymer compatibility and stabilisation effects depend on both chemical structure and drug-specific interactions rather than on broad polymer labels alone [30, 31]. In practice, transferability should be treated as conditional and uncertainty-weighted, especially when extrapolating to new drug–polymer chemistries.

Enforcing Physicochemical Monotonicity

Physicochemical constraints can improve model robustness by preventing predictions that contradict established formulation principles. For example, increasing drug loading would generally be expected to increase crystallization risk or reduce stabilising polymer availability, so the model could impose monotonic restrictions where supported by domain knowledge [18, 32]. Similar constraints could be applied to Tg-related mobility descriptors, where lower mixture Tg relative to storage temperature should not be interpreted as beneficial for physical stability [4]. These constraints should remain flexible enough to accommodate polymer-specific exceptions, because dissolution performance and precipitation inhibition may improve or deteriorate through mechanisms not captured by drug loading alone [33].

Model Interpretability and Formulation Insights

SHAP Analysis of Formulation Drivers

SHAP analysis could be used to explain whether a predicted ASD failure is driven primarily by poor miscibility, excessive drug loading, low Tg, or weak precipitation inhibition. Prior stability modelling work supports the value of interpretable machine learning in distinguishing formulation factors associated with physical instability, while deep learning approaches highlight the need for additional explanation when class imbalance or complex feature interactions are present [12, 13]. In this framework, high Flory-Huggins interaction values or unfavourable compatibility descriptors would be expected to increase crystallization-risk contributions, whereas favourable Tg and dissolution-profile features could moderate that risk [3, 19]. The purpose of interpretation is not only to justify a prediction but also to translate model logic into experimentally actionable formulation hypotheses. **Table 2** summarizes how SHAP-identified formulation drivers can be translated into practical ASD reformulation hypotheses and experimental prioritization decisions.

Table 2. Translating SHAP-Identified Formulation Drivers into Experimental ASD Design Actions

SHAP-Identified Driver	Interpretation of Model Logic	Practical Formulation Action
Poor drug–polymer miscibility	Predicted instability is mainly driven by weak molecular compatibility or high crystallization tendency.	Screen alternative polymers or use complementary miscibility evidence such as COSMO-based or NMR-supported assessment.
Excessive drug loading	The model suggests that the formulation exceeds the stabilizing capacity of the polymer matrix.	Reduce drug loading or test a graded drug–polymer ratio series to identify the stability threshold.
Low Tg or weak Tg margin	Physical instability may be linked to insufficient glass-transition protection under storage conditions.	Select a higher-Tg polymer, increase polymer fraction, or reconsider storage and packaging requirements.
Weak precipitation inhibition	Dissolution benefit may not translate into sustained supersaturation.	Prioritize polymers or excipient combinations with stronger precipitation-inhibition capacity.
Favorable dissolution but high stability risk	The model detects a trade-off between release performance and crystallization risk.	Move the candidate to a lower-priority experimental queue unless stability can be improved.

Translating Model Logic into Formulation Recommendations

The model could be used in silico to vary drug loading, polymer identity, or Tg-related descriptors and evaluate how predicted stability and dissolution behaviour respond. Such counterfactual analysis could identify the minimum polymer fraction expected to maintain stability, rank polymers by predicted stabilising effect, or flag cases where dissolution improvement may be offset by precipitation risk [16, 17]. Compatibility-oriented tools such as COSMO-based screening and nuclear magnetic resonance approaches could provide complementary evidence for these recommendations when miscibility remains uncertain [25, 30]. The recommended formulation set should therefore be viewed as a prioritised experimental queue rather than as a final prescription.

Integration Into Formulation Development Workflow

Pre-Screening before Hot-Melt Extrusion or Spray Drying

During pre-formulation, the predictive model could rank candidate drug–polymer combinations before hot-melt extrusion or spray drying is attempted. This would be especially useful because preparation route, residual crystallinity, particle attributes, and process conditions can influence ASD performance even when the nominal formulation composition is similar [23, 28]. Machine learning models developed for chemically stable hot-melt extrusion systems and high-throughput ASD formation screening show how formulation prediction can be positioned before resource-intensive manufacturing trials [14, 15]. In such a workflow, the model would narrow the formulation space and direct confirmatory experiments toward candidates with a stronger predicted balance of miscibility, Tg, and dissolution performance.

Linking Prediction to Quality-by-Design

A predictive ASD model could support Quality-by-Design thinking by converting formulation variables into an interpretable design space. Input ranges for drug loading, polymer grade, miscibility index, Tg margin, and dissolution-profile descriptors could be linked to predicted regions of acceptable stability and supersaturation performance [6, 33]. Because physical stability and dissolution advantage are both required for a viable ASD, the design space should represent trade-offs rather than a single pass-fail threshold [2]. The model would therefore provide a structured rationale for risk assessment, formulation selection, and later experimental verification.

Evaluation Strategy

Prediction Accuracy

Prediction accuracy should be evaluated separately for stability classification and dissolution-related regression outputs, but without relying solely on aggregate metrics. Stability classification could be assessed conceptually through discrimination, sensitivity, and specificity, while dissolution outputs could be evaluated by agreement between predicted and observed supersaturation behaviour on held-out drug–polymer combinations [12, 22]. External testing should emphasise chemical and formulation dissimilarity from the training set because random splits may overstate generalisability when similar systems appear in both training and test partitions [16]. The most informative evaluation would determine whether the model correctly prioritises formulations for follow-up rather than whether it merely reproduces patterns in a familiar dataset.

Benchmark against Physicochemical Rules

The proposed model should be benchmarked against conventional formulation rules, such as simple drug-loading cut-offs, Tg-margin criteria, or miscibility classifications based on solubility parameters alone. Such comparisons are important because critiques of regular-solution and solubility-parameter approaches show that simple compatibility estimates may be useful but incomplete for ASD design [19, 21]. The model should also be compared with thermodynamic and semi-empirical approaches that predict solubility enhancement or crystallization tendency from basic principles [24]. A useful predictive framework would be expected to add value by integrating miscibility, Tg, and dissolution features rather than by replacing expert rules with an opaque statistical score. **Figure 2** illustrates the proposed evaluation strategy, showing how model performance should be assessed across prediction accuracy, chemically dissimilar external testing, formulation-prioritisation utility, and benchmarking against conventional physicochemical rules.

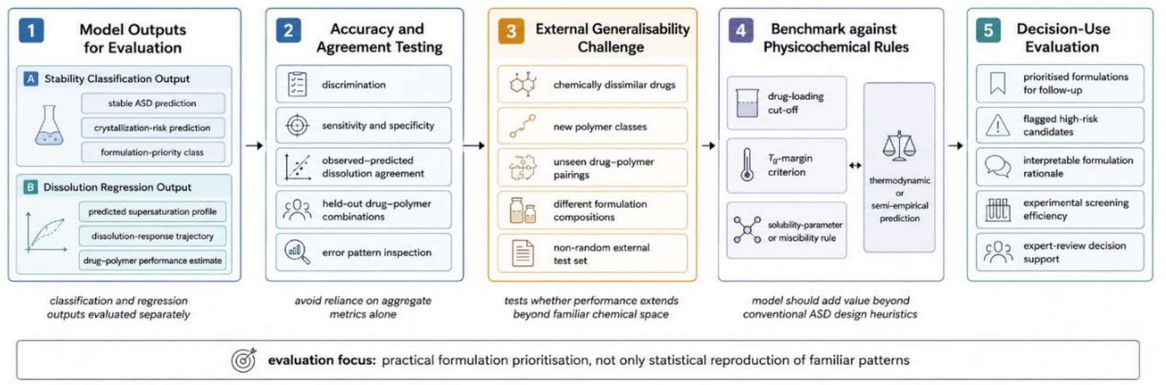

Figure 2. Evaluation Architecture for an ASD Prediction Model Integrating Accuracy Testing, External Generalisability, and Physicochemical Benchmarking

Prospective Experimental Validation

Prospective validation should involve selecting a small set of candidate ASDs predicted to span low, intermediate, and high performance risk, then testing whether observed stability and dissolution behaviour align with the model's ranking. The experimental design should include both formulations expected to succeed and formulations expected to fail, because negative examples are essential for assessing whether the model can discriminate risk rather than only confirm favourable cases [26]. Validation should also consider dissolution mechanism, precipitation behaviour, and release-profile interpretation, since a physically stable ASD may still underperform if it fails to sustain supersaturation [3, 33]. The outcome of such validation would be used to refine the model and update uncertainty estimates rather than to claim universal applicability.

Table 3 translates predicted stability, supersaturation, uncertainty, and interpretability outputs into formulation-development actions and confirmatory testing priorities.

Table 3. Prediction-to-Action Framework for Interpreting ASD Model Outputs

Model output pattern	Likely formulation interpretation	Key explanatory drivers to inspect	Recommended formulation action	Confirmatory experiment priority
Low crystallization risk + strong supersaturation	Candidate formulation is predicted to balance	Favorable miscibility; adequate Tg margin;	Advance to focused confirmatory stability and dissolution testing	High priority for experimental

maintenance + high confidence	amorphous stability with dissolution advantage	moderate drug loading; long precipitation lag		confirmation as a lead ASD candidate
Low crystallization risk + weak supersaturation generation	Formulation may remain amorphous but may not deliver meaningful dissolution benefit	High Tg margin but low maximum concentration; weak supersaturation ratio; limited spring effect	Consider polymer or surfactant strategy changes to improve dissolution performance	Medium priority; confirm only if stability is strategically valuable
High crystallization risk + strong initial supersaturation	Formulation may dissolve rapidly but fail because of phase separation, precipitation, or recrystallization risk	High drug loading; poor miscibility score; low Tg margin; short precipitation lag	Reduce drug loading, change polymer, or test precipitation-inhibition strategies	High priority as a failure-mode learning case rather than a direct lead
High crystallization risk + weak dissolution performance	Candidate is predicted to be poor across both stability and dissolution dimensions	Poor compatibility; unfavorable Tg; weak dissolution profile; high uncertainty in core features	Deprioritize unless there is a compelling therapeutic or manufacturing reason	Low priority; use as a negative control if needed for model validation
Intermediate stability risk + strong dissolution profile + moderate uncertainty	Candidate may be promising but requires targeted stability clarification	Conflicting miscibility indicators; measured–predicted Tg discrepancy; limited storage data	Add focused miscibility and accelerated stability testing before scale-up	Medium-to-high priority depending on chemical novelty
Stable prediction but high humidity sensitivity	Dry-state model may underestimate real storage risk under moisture exposure	Hygroscopic polymer; low effective Tg under humidity; missing relative-humidity data	Include moisture sorption, humidity-stressed stability, and packaging-risk assessment	High priority if intended for humid storage or accelerated testing
Favorable prediction but out-of-domain warning	Prediction may be unreliable because the formulation lies outside the training chemical or polymer space	Novel drug chemistry; rare polymer; unusual process route; sparse comparable examples	Treat prediction as hypothesis-generating only and expand experimental evidence	High priority for applicability-domain validation
Strong model disagreement with simple physicochemical rules	Machine learning signal may reveal interaction effects, or the model may be learning dataset bias	SHAP drivers; Tg margin; miscibility conflict; dissolution-profile features; missingness pattern	Review raw data and benchmark against rule-based expectations before acting	Medium priority; useful for model audit and mechanistic interpretation

Limitations

Generality across Wide Chemical Space

The model's generality is limited by the chemical diversity, polymer coverage, preparation routes, and reporting consistency of the data used to train it. Predictions for structurally unusual drugs, novel polymers, salts, co-amorphous systems, or poorly represented processing methods should be interpreted with caution because learned relationships may not extrapolate reliably beyond the training domain [27, 29]. Even sophisticated compatibility tools and hybrid modelling approaches remain sensitive to the assumptions used to represent drug–polymer interactions [25]. For this reason, the model should include applicability-domain assessment so that unfamiliar chemistries are flagged for experimental confirmation.

Environmentally-Driven Stability

Environmental effects represent a major limitation when moisture sorption, humidity exposure, and water-induced plasticisation are not included as explicit features. Water uptake can reduce effective Tg, increase molecular mobility, and alter drug–polymer interactions, meaning that predictions based mainly on dry-state miscibility and thermal properties may underestimate humidity-driven crystallization risk [1, 18]. Polymer hygroscopicity and storage relative humidity should therefore be encoded whenever available, especially for formulations intended for accelerated or long-term stability assessment. Without these features, the model should be considered primarily a predictor of composition-driven performance under controlled conditions rather than a complete stability model.

Conclusion

A predictive model linking miscibility, glass transition, and dissolution data could provide an integrated framework for amorphous solid dispersion design. By combining interaction descriptors, thermal mobility indicators, drug loading, polymer properties, and dissolution-profile features, the model would address the multi-factor nature of ASD performance. Its role would be to estimate likely crystallization risk and dissolution advantage before extensive formulation screening is performed. The main strength of this approach is its ability to combine information that is usually interpreted separately. Miscibility measurements, Tg analysis, and dissolution testing each provide partial insight, but a model-oriented framework can relate them to a shared formulation-performance objective. When paired with interpretability tools, the model could help formulators understand which variables are most responsible for predicted success or failure.

Important challenges remain before such a workflow can become a routine formulation tool. Broader datasets are needed to cover more drugs, polymers, preparation methods, storage environments, and dissolution protocols. Prospective validation is also essential because a model trained on historical or literature-derived records may not automatically generalise to new chemical space or new laboratory conditions.

Progress in predictive ASD design would benefit from open, standardised data sharing and collaborative benchmarking. Shared reporting formats for miscibility, Tg, stability, and dissolution outcomes would make formulation datasets more reusable and comparable. With transparent validation and careful domain-of-applicability assessment, predictive modelling could mature into a practical decision-support approach for amorphous formulation development.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Lin X, Hu Y, Liu L, Su L, Li N, Yu J, et al. Physical stability of amorphous solid dispersions: a physicochemical perspective with thermodynamic, kinetic and environmental aspects. J Pharm Res. 2018;35(6):125.
2. Ueda K, Moseson DE, Taylor LS. Amorphous solubility advantage: theoretical considerations, experimental methods, and contemporary relevance. J Pharm Sci. 2025;114(1):18-39.
3. Baghel S, Cathcart H, O'Reilly NJ. Understanding the generation and maintenance of supersaturation during the dissolution of amorphous solid dispersions using modulated DSC and 1H NMR. Int J Pharm. 2018;536(1):414-25.
4. Szabo E, Haraszti A, Zahonyi P, Vadas D, Csontos I, Nagy ZK, et al. Evaluation of different thermoanalytical methods for the analysis of the stability of naproxen-loaded amorphous solid dispersions. Pharmaceutics. 2022;14(11):2508.
5. Thakore SD, Akhtar J, Jain R, Paudel A, Bansal AK. Analytical and computational methods for the determination of drug-polymer solubility and miscibility. Mol Pharm. 2021;18(8):2835-66.
6. Hermans A, Milsmann J, Li H, Jede C, Moir A, Hens B, et al. Challenges and strategies for solubility measurements and dissolution method development for amorphous solid dispersion formulations. AAPS J. 2023;25(1):11.
7. Bhujbal SV, Mitra B, Jain U, Gong Y, Agrawal A, Karki S, et al. Pharmaceutical amorphous solid dispersion: a review of manufacturing strategies. Acta Pharm Sin B. 2021;11:2505-36.
8. Shi Q, Chen H, Wang Y, Wang R, Xu J, Zhang C. Amorphous solid dispersions: role of the polymer and its importance in physical stability and in vitro performance. Pharmaceutics. 2022;14(8):1747.
9. Zhang J, Guo M, Luo M, Cai T. Advances in the development of amorphous solid dispersions: the role of polymeric carriers. Asian J Pharm Sci. 2023;18(4):100834.
10. Qian F, Huang J, Hussain MA. Drug-polymer solubility and miscibility: stability consideration and practical challenges in amorphous solid dispersion development. J Pharm Sci. 2010;99(7):2941-7.
11. Taylor LS, Zhang GGZ. Physical chemistry of supersaturated solutions and implications for oral absorption. Adv Drug Deliv Rev. 2016;101:122-42.
12. Han R, Xiong H, Ye Z, Yang Y, Huang T, Jing Q, et al. Predicting physical stability of solid dispersions by machine learning techniques. J Control Release. 2019;311:16-25.
13. Lee H, Kim J, Kim S, Yoo J, Choi GJ, Jeong YS. Deep learning-based prediction of physical stability considering class imbalance for amorphous solid dispersions. J Chem. 2022;2022(1):4148443.
14. Jiang J, Lu A, Ma X, Ouyang D, Williams RO 3rd. The applications of machine learning to predict the forming of chemically stable amorphous solid dispersions prepared by hot-melt extrusion. Int J Pharm X. 2023;5:100164.
15. Lu T, Wu Y, Xiong P, Zhong H, Ding Y, Li H, et al. Combining high-throughput screening and machine learning to predict the formation of both binary and ternary amorphous solid dispersion formulations for early drug discovery and development. Pharm Res. 2025;42(4):697-709.
16. Zeneli E, Bohets H, Mebenga FN, Holm R, Tistaert C, Kuentz M. Guiding excipient selection for physically stable amorphous solid dispersions: a combined in-vitro in-silico approach. Eur J Pharm Sci. 2025;212:107152.
17. Zeneli E, Bohets H, Mebenga FN, Tistaert C, Holm R, Kuentz M. Guiding excipient selection for amorphous solid dispersions by combining an in vitro-in silico approach-II: supersaturation and drug release. Int J Pharm. 2025:125976.
18. Lehmkemper K, Kyeremateng SO, Bartels M, Degenhardt M, Sadowski G. Physical stability of API/polymer-blend amorphous solid dispersions. Eur J Pharm Biopharm. 2018;124:147-57.
19. Anderson BD. Predicting solubility/miscibility in amorphous dispersions: it is time to move beyond regular solution theories. J Pharm Sci. 2018;107(1):24-33.

20. Ousset A, Chavez PF, Meeus J, Robin F, Schubert MA, Somville P, et al. Prediction of phase behavior of spray-dried amorphous solid dispersions: assessment of thermodynamic models, standard screening methods and a novel atomization screening device with regard to prediction accuracy. Pharmaceutics. 2018;10(1):29.
21. Turpin ER, Taresco V, Al-Hachami WA, Booth J, Treacher K, Tomasi S, et al. In silico screening for solid dispersions: the trouble with solubility parameters and χFH. Mol Pharm. 2018;15(10):4654-67.
22. Au-Yeung L, Tseng CY, Tam YK, Tsai PA. Machine learning based quantitative structure-dissolution profile relationship. J Chem Inf Model. 2025;65(12):6273-86.
23. Schmitt JM, Baumann JM, Morgen MM. Predicting spray dried dispersion particle size via machine learning regression methods. Pharm Res. 2022;39(12):3223-39.
24. Skrdla PJ, Floyd PD, Dell'Orco PC. Predicting the solubility enhancement of amorphous drugs and related phenomena using basic thermodynamic principles and semi-empirical kinetic models. Int J Pharm. 2019;567:118465.
25. Antolovic I, Vrabec J, Klajmon M. COSMOPharm: drug-polymer compatibility of pharmaceutical amorphous solid dispersions from COSMO-SAC. Mol Pharm. 2024;21(9):4395-415.
26. Mistry P, Amponsah-Efah KK, Suryanarayanan R. Rapid assessment of the physical stability of amorphous solid dispersions. Cryst Growth Des. 2017;17(5):2478-85.
27. de Castro M, Cordeiro AS, Li M, Lübbert C, McColl C, Khurana J, et al. Advancing amorphous solid dispersions through empirical and hybrid modeling of drug-polymer solubility and miscibility: a case study using ibuprofen. Int J Pharm X. 2025:100373.
28. Moseson DE, Taylor LS. The application of temperature-composition phase diagrams for hot melt extrusion processing of amorphous solid dispersions to prevent residual crystallinity. Int J Pharm. 2018;553(1-2):454-66.
29. Iemtsev A, Hassouna F, Klajmon M, Mathers A, Fulem M. Compatibility of selected active pharmaceutical ingredients with poly(D,L-lactide-co-glycolide): computational and experimental study. Eur J Pharm Biopharm. 2022;179:232-45.
30. Coutinho AL, Hom K, Polli JE. Prediction of successful amorphous solid dispersion pairs through liquid state nuclear magnetic resonance. Mol Pharm. 2024;21(12):6153-65.
31. Pugliese A, Tobyn M, Hawarden LE, Abraham A, Blanc F. New development in understanding drug-polymer interactions in pharmaceutical amorphous solid dispersions from solid-state nuclear magnetic resonance. Mol Pharm. 2022;19(11):3685-99.
32. Song S, Wang C, Zhang B, Sun CC, Lodge TP, Siegel RA. A rheological approach for predicting physical stability of amorphous solid dispersions. J Pharm Sci. 2023;112(1):204-12.
33. Walter S, Mileo PG, Afzal MA, Kyeremateng SO, Degenhardt M, Browning AR, et al. Predicting the release mechanism of amorphous solid dispersions: a combination of thermodynamic modeling and in silico molecular simulation. Pharmaceutics. 2024;16(10):1292.