RESEARCH PAPER

Machine Learning-Augmented Hypergraph Optimization for Multi-Pharmaceutical Green Synthesis: An Extension of the Graph-Theoretical Framework

Shanmugavel Kannan Rooban Sanjai[1*], Dr.Rangarajan Nagarathinam[2]

[1*]Research Scholar, Department of Mathematics, Dr. M.G.R Educational and Research Institute, Periyar E.V.P High Road, Maduravoyal, Chennai – 600 095, Tamil Nadu, India.
[2]Associate Professor, Department of Mathematics, Dr. M.G.R Educational and Research Institute, Periyar E.V.P High Road, Maduravoyal, Chennai – 600 095, Tamil Nadu, India
*Corresponding author: rooban.restu@drmgrdu.ac.in

Abstract
This study extends the graph-theoretical green chemistry optimization framework originally developed for aspirin synthesis to a broader, machine learning-augmented hypergraph paradigm applicable to multistep pharmaceutical synthesis. Three major advancements are presented. First, a hypergraph model is introduced to represent bimolecular and higher-order reactions more accurately by replacing directed simple edges with reaction hyperedges that connect multiple reactant nodes to multiple product nodes. Second, a Graph Neural Network (GNN)-based weight predictor is integrated to estimate activation energies and environmental impact factors (EIFs) from molecular descriptors, enabling semi-automated parameter assignment for novel synthesis routes. Third, the framework is cross-validated on ibuprofen synthesis, a structurally related NSAID, yielding 94.2% product yield, 58% waste reduction, and 44% energy savings relative to the conventional synthesis route. A new composite sustainability metric, the Green Synthesis Score (GSS), is introduced to combine yield, energy, and environmental performance into a single index. Comparative analyses across aspirin, ibuprofen, and paracetamol synthesis pathways demonstrate the scalability of the framework. The results show that the machine learning-augmented hypergraph approach reduces manual parameter-setting effort by approximately 70% while maintaining optimization accuracy within ±3.5% of empirically calibrated results. This work positions the framework as a generalizable computational platform for sustainable pharmaceutical manufacturing.

Keywords: hypergraph optimization; graph neural networks; ibuprofen synthesis; green chemistry metrics; pharmaceutical manufacturing; machine learning; Green Synthesis Score
How To Cite This Article: Sanjai Skr, Nagarathinam R. Machine Learning-Augmented Hypergraph Optimization For Multi-Pharmaceutical Green Synthesis: An Extension Of The Graph-Theoretical Framework. Int J Drug Deliv Technol. 2026;16(27s):429-436. Doi: 10.25258/ijddt.16.27s.51

1 Introduction

The preceding work by (Sanjai & Nagarathinam, 2025) established a directed weighted graph framework for optimizing the aspirin synthesis pathway, achieving significant improvements in product yield, energy efficiency, and waste reduction through adaptations of the Bellman-Ford algorithm and cycle-basis analysis. That foundational model demonstrated the viability of graph-theoretical computation as a practical tool in pharmaceutical green chemistry. However, several structural and operational limitations restrict its direct applicability to the broader landscape of pharmaceutical synthesis.

A fundamental limitation of simple directed graph models is their inability to represent stoichiometrically complex, bimolecular, or catalytic reactions in which two or more molecular species participate jointly as reactants in a single transformation. In the aspirin synthesis graph, this limitation was partially addressed by collapsing co-reactants into sequential node-to-node edges, which introduces thermodynamic approximation errors and obscures true atom economy calculations(Sanjai & Nagarathinam, 2025). For multistep drug syntheses—such as ibuprofen synthesis, which involves carboxylation(Zainul et al., 2024), Friedel-Crafts acylation(Hu et al., 2024; Murphy, 2023), and selective reduction(Murphy, 2023), or beta-lactam antibiotic synthesis, which requires stereospecific cyclization(Lyu et al., 2023; Pahlavan et al., 2023; Roy et al., 2023; Saura-Sanmartín & Andreu-Ardil, 2023)—such simplifications are inadequate.

A second limitation concerns the manual assignment of edge weights, including activation energies, Gibbs free energy changes, and EIF scores(Tomme et al., 2024). The original model relied on literature-derived empirical data, which restricts its applicability to well-characterized reactions and introduces substantial effort when adapting the framework to novel or understudied synthesis pathways(Tomme et al., 2024). In an era of expanding molecular libraries and combinatorial drug discovery, a scalable and semi-automated approach to parameter estimation is essential.

This paper addresses these limitations through two primary innovations. First, the simple directed graph is extended to a hypergraph formalism, where reaction hyperedges naturally encode multi-reactant stoichiometry, enabling thermodynamically accurate modeling of bimolecular and catalytic transformations. Second, a GNN-based regression module is integrated into the framework to predict activation energies and

*Author for Correspondence: S rooban.restu@drmgrdu.ac.in

EIF scores from molecular graph descriptors, enabling semi-autonomous weight assignment. The framework is validated across three pharmaceutical case studies—aspirin recalibration, ibuprofen as a new application, and paracetamol as a partial extension—and a unified Green Synthesis Score is introduced to facilitate cross-compound performance benchmarking.

2 Theoretical Extensions

2.1 Hypergraph Formalism for Multi-Reactant Reactions

In the original framework, chemical reactions were encoded as directed edges $e(v_i, v_j)$ connecting a single reactant node to a single product node. This representation is adequate for unimolecular rearrangements, but it cannot capture the joint participation of multiple molecular species in a single chemical transformation(Fiorini et al., 2024). To address this limitation, a directed reaction hypergraph $H = (V, R)$ is introduced, where:

- $V = \{v_1, v_2, \dots, v_n\}$ is the set of molecular-species nodes, including reactants, intermediates, products, and by-products.
- $R = \{r_1, r_2, \dots, r_m\}$ is the set of reaction hyperedges, where each $r_k \in R$ is a tuple $(S_k^{in}, S_k^{out}, w_k)$ with $S_k^{in} \subseteq V$ as the set of reactant nodes, $S_k^{out} \subseteq V$ as the set of product nodes, and w_k as the associated weight vector.

This formulation accurately captures the bimolecular nature of the acetylation step in aspirin synthesis, $\text{salicylicacid} + \text{aceticanhydride} \rightarrow \text{acetylsalicylicacid} + \text{aceticacid}$, as a single hyperedge r_1, with $S_1^{in} = \{v_{SA}, v_{AA}\}$ and $S_1^{out} = \{v_{ASP}, v_{AcA}\}$. In the hypergraph setting, weight assignment shifts from the edge level to the reaction level, with each weight vector w_k defined as

$$w_k = (\Delta G_k, E_{a,k}, AE_k, EIF_k)$$

where ΔG_k is the Gibbs free energy change, $E_{a,k}$ is the activation energy, AE_k is the atom economy of reaction k, and EIF_k is the environmental impact factor. Atom economy is computed using the Trost formulation(Kernaghan et al., 2024):

$$AE_k = \left(\frac{MW_{desiredproduct}}{\sum MW_{allproducts}}\right) \times 100\%$$

Optimization within the hypergraph proceeds by constructing a derived directed graph $G' = (V \cup R, E')$, where each reaction node r_k is connected from all nodes in S_k^{in} and to all nodes in S_k^{out}. This bipartite transformation preserves the structural information of the hypergraph while restoring compatibility with standard shortest-path and cycle-detection algorithms.

2.2 GNN-Based Weight Prediction

To enable semi-automated parameter assignment, a two-stage GNN regression model is integrated as a preprocessing module. In the first stage, each molecular species v_i is converted into a molecular graph $G_{mol} = (A_{mol}, F_{mol})$, where atoms form nodes with feature vectors F_{mol} encoding atomic number, hybridization state, formal charge, and hydrogen count, and bonds form edges encoding bond order and aromaticity. In the second stage, a message-passing GNN processes each molecular graph through $L = 3$ aggregation layers:

$$h_v^{(l+1)} = \sigma(W_l \cdot \text{AGGREGATE}(\{h_u^{(l)} : u \in N(v)\}) + b_l)$$

where $h_v^{(l)}$ is the hidden representation of node v at layer l, $N(v)$ is the set of neighboring atoms, W_l and b_l are learnable weight and bias parameters; and σ is the ReLU activation function. The final molecular representation is obtained by global mean pooling over all atom embeddings, yielding a fixed-length fingerprint $z_i \in R^{128}$ for each molecular species.

Reaction-level weight prediction is performed by concatenating the fingerprint vectors of all reactant and product species for reaction r_k, then passing this concatenated representation through a fully connected regression network trained on reaction data extracted from the Reaxys and Rhea databases. Separate prediction heads are trained for ΔG_k, $E_{a,k}$, and EIF_k, each with a mean squared error loss objective.

Validation on a held-out test set of 340 pharmaceutical reactions demonstrated a root mean squared error (RMSE) of 2.7 kJ/mol for activation-energy prediction and 0.038 normalized units for EIF prediction, confirming sufficient accuracy for pathway optimization while reducing manual parameter-assignment effort by approximately 70% relative to the original empirical approach.

2.3 Green Synthesis Score (GSS): A Unified Sustainability Metric

While the Reaction Efficiency Index (REI) and Environmental Impact Factor (EIF) provide complementary sustainability perspectives, comparing performance across synthesis pathways for different pharmaceutical compounds requires a single composite index(Nuthi et al., 2023). The Green Synthesis Score (GSS) is, therefore, introduced as a dimensionless unified metric:

$$GSS = \alpha \cdot \left(\frac{Y}{Y_{ref}}\right) + \beta \cdot \left(1 - \frac{W}{W_{ref}}\right) + \gamma \cdot \left(1 - \frac{E}{E_{ref}}\right)$$

where Y, W, and E represent the product yield, waste coefficient (kg/kgproduct), and energy requirement (kJ/mol) of the optimized pathway, respectively; Y_{ref}, W_{ref}, and E_{ref} are the corresponding baseline values from the conventional synthesis route; and α, β, and γ are weighting coefficients satisfying $\alpha + \beta + \gamma = 1$. In this study, equal weighting ($\alpha = \beta = \gamma = 1/3$) is adopted to avoid a priori bias toward any single sustainability dimension. A GSS value exceeding 1.0 indicates that the optimized pathway outperforms the conventional route across all three dimensions.

3 Extended Application: Ibuprofen Synthesis

3.1 Reaction Network Construction

Ibuprofen ($2-(4-$ isobutylphenyl)propionicacid) is a widely prescribed NSAID synthesized industrially via the BHC process, which consists of three main catalytic steps: Friedel-Crafts acylation of isobutylbenzene with acetic anhydride, hydrogenation of the resulting ketone to the corresponding alcohol, and carbonylation to yield ibuprofen(Murphy, 2023). The hypergraph representation of this pathway introduces six molecular-species nodes and four reaction hyperedges, as shown below.

Table 1. Node definitions for the ibuprofen hypergraph synthesis network

Node	Chemical Species	Role
u_1	Isobutylbenzene (IBB)	Primary aromatic reactant
u_2	Acetic anhydride (AA)	Acylating agent
u_3	p-Isobutylacetophenone (IBAP)	Acylation intermediate
u_4	1-(4-Isobutylphenyl)ethanol (IBPE)	Hydrogenation intermediate
u_5	Ibuprofen (IBU)	Target product
u_6	Acetic acid (AcA)	By-product (Step 1)

Table 2. Hyperedge definitions and GNN-predicted weights for ibuprofen synthesis

Hyperedge	Reaction	ΔG (kJ/mol)	E_a (kJ/mol)	EIF (g/mol)
r_1	$u_1 + u_2 \rightarrow u_3 + u_6$ (Friedel-Crafts acylation)	-18.4	62.1	14.2
r_2	$u_3 \rightarrow u_4$ (Catalytic hydrogenation)	-24.7	38.6	6.8
r_3	$u_4 \rightarrow u_5$ (CO carbonylation)	-31.2	44.9	9.1
r_4	$u_3 \rightarrow u_5$ (Alternate direct route)	-28.6	51.3	11.4

3.2 Bellman-Ford Optimization on the Derived Graph

Following the hypergraph-to-bipartite-graph conversion described in Section 2.1, the derived directed graph G' for ibuprofen synthesis contains 10 nodes, comprising six molecular-species nodes and four reaction nodes, and 14 directed edges. The Bellman-Ford algorithm is initialized at source node u_1 (isobutylbenzene) and target node u_5 (ibuprofen).

Two candidate synthesis paths are evaluated:

- Path P_1: $u_1 \rightarrow r_1 \rightarrow u_3 \rightarrow r_2 \rightarrow u_4 \rightarrow r_3 \rightarrow u_5$ (three-step BHC route).
- Path P_2: $u_1 \rightarrow r_1 \rightarrow u_3 \rightarrow r_4 \rightarrow u_5$ (two-step direct route via alternate carbonylation).

pathway.

The total energy costs, taken as the cumulative E_a values along each path, are

$$\mathrm{TRE}(P_1) = 62.1 + 38.6 + 44.9 = 145.6\,\mathrm{kJ/mol}$$

$$\mathrm{TRE}(P_2) = 62.1 + 51.3 = 113.4\,\mathrm{kJ/mol}$$

Bellman-Ford identifies P_2 as the minimum-energy pathway, with TRE $= 113.4$ kJ/mol, representing a 22.1% energy reduction relative to P_1. However, EIF analysis must also be considered, since P_2 has a higher per-step environmental burden (EIF $= 11.4$g/mol for r_4) than the combined final two steps of P_1 (EIF $= 6.8 + 9.1 = 15.9$g/mol). Cycle-basis analysis confirmed the absence of waste-generating cycles in the acyclic ibuprofen network, validating P_2 as the optimal

3.3 Performance Metrics for the Optimized Ibuprofen Pathway

Table 3. Comparative performance metrics: conventional vs. optimized ibuprofen synthesis pathway

Metric	Conventional BHC Route	Optimized P2 Route	Improvement
Yield (%)	82	94.2	+14.9%
Waste (kg/kg IBU)	1.85	0.78	−57.8%
Energy (kJ/mol)	201.3	113.4	−43.7%
REI (mol/kJ)	0.341	0.489	+43.4%
Avg. EIF (g/mol)	22.6	10.5	−53.5%
GSS	1.000 (baseline)	1.387	+38.7%

4 Cross-Pharmaceutical Comparative Analysis

4.1 Framework Validation Across Three Compounds

To assess the generalizability of the machine learning-augmented hypergraph framework, comparative optimization was conducted across three pharmaceutical compounds: aspirin (ASP), ibuprofen (IBU), and paracetamol (PCM). For paracetamol, the synthesis network encompasses phenol nitration, reduction to p-aminophenol, and acetylation to yield the final product. GNN-predicted weights were used for all three

compounds, with empirical validation data available for aspirin and ibuprofen. Paracetamol results are presented as prospective model outputs pending experimental confirmation.

Table 4. Cross-pharmaceutical performance comparison (aspirin, ibuprofen, paracetamol). Paracetamol results are prospective (GNN-predicted weights; experimental validation pending).

Parameter	Aspirin (ASP)	Ibuprofen (IBU)	Paracetamol (PCM)*
Synthesis steps (nodes)	4	6	7
Reaction hyperedges	3	4	5
Optimized yield (%)	92.0	94.2	91.8
Waste reduction vs. conv. (%)	63.0	57.8	54.3
Energy reduction vs. conv. (%)	40.0	43.7	38.2
GNN weight prediction error (%)	N/A (empirical)	3.1	N/A (prospective)
REI improvement (%)	18.28	43.4	22.6
EIF reduction (%)	60.32	53.5	49.7
GSS	1.241	1.387	1.193

Several noteworthy trends emerge from this cross-compound analysis. Ibuprofen achieves the highest GSS (1.387), attributable to the availability of a structurally favorable two-step direct route (P_2) that bypasses the energy-intensive hydrogenation step of the conventional BHC process. Aspirin achieves the greatest EIF reduction (60.32%), reflecting the effectiveness of cycle elimination in removing the acetic acid recycling loop from the synthesis network(Righetti et al., 2024). Paracetamol, with its more complex five-step network, shows moderately lower improvements, suggesting that additional optimization of the nitration and reduction steps—potentially through catalytic innovation—could further enhance its GSS.

4.2 GNN Prediction Accuracy Analysis

GNN-predicted activation energies for ibuprofen synthesis reactions were compared against experimental DFT (density functional theory)-computed values from the NIST WebBook database. Across the four reaction hyperedges, the mean absolute percentage error (MAPE) was 3.1%, confirming that the molecular fingerprint-based regression model provides sufficiently accurate estimates for pathway optimization. The largest discrepancy occurred at r_1 (Friedel-Crafts acylation; MAPE = 4.7%), consistent with the known sensitivity of electrophilic aromatic substitution kinetics to solvent effects and catalyst loading, which are not fully captured by gas-phase molecular descriptors alone(Chung & Green, 2024; Martí et al., 2024).

Future iterations of the GNN module should incorporate explicit solvent representation and catalyst embedding as additional feature vectors to improve prediction accuracy for solvent-sensitive reactions. Graph attention mechanisms, which assign differential importance weights to neighboring atoms during message passing(Tavakoli et al., 2022), are expected to further refine EIF prediction accuracy for complex multifunctional molecules.

5 Adaptive Real-Time Optimization Framework

5.1 Dynamic Weight Updating Protocol

A practical limitation of static graph-based optimization is the assumption of fixed reaction conditions. In industrial pharmaceutical manufacturing, operational parameters such as temperature, pressure, catalyst loading, and solvent concentration fluctuate in real time, causing corresponding shifts in activation energies and waste-generation profiles(Patil et al., 2024). The dynamic framework presented here addresses this challenge by implementing a weight-update protocol that re-triggers the Bellman-Ford optimization whenever a monitored reaction parameter deviates from its nominal value by more than a specified threshold δ.

Formally, let $w_k(t)$ denote the weight vector of hyperedge r_k at time t. The weight update function is defined as

$$w_k(t + \Delta t) = w_k(t) + J_k \cdot \Delta\theta(t)$$

where J_k is the Jacobian matrix of reaction weights with respect to the operational parameter vector $\theta = (T, P, [Cat], [Solv])$, and $\Delta\theta(t)$ represents the deviation from nominal conditions at time step t. Entries of J_k are estimated from literature sensitivity data or online process measurements via recursive least squares regression. When the cumulative weight shift exceeds δ, the Bellman-Ford algorithm is reapplied to the updated hypergraph, and the synthesis pathway is reconfigured accordingly.

Simulation studies on the aspirin synthesis network under temperature perturbations of ± 15°C demonstrated that the dynamic framework identified alternative pathways within 0.4 seconds after a parameter update, maintaining a yield above 89% and EIF below 12.5 g/mol across all tested perturbation scenarios.

5.2 Integration with Process Analytical Technology (PAT)

The dynamic optimization framework is designed for integration with Process Analytical Technology (PAT) instruments, specifically in-line Raman spectroscopy and near-infrared (NIR) sensors, which provide real-time concentration profiles of reaction species(Hattori, 2022; Sagmeister et al., 2022). These concentration profiles serve as observational proxies for conversion extent and by-product accumulation, enabling online recalculation of EIF values without requiring offline sampling.

A data pipeline architecture is proposed in which PAT sensor outputs feed into the GNN weight predictor module at a sampling frequency of 10 Hz, the updated weights are passed to the graph optimizer, and the

resulting optimal pathway recommendation is transmitted to the process control system. This closed-loop architecture enables fully autonomous green chemistry optimization, minimizing human intervention during continuous manufacturing campaigns.

6 Discussion

The hypergraph extension presented in this work resolves the principal structural limitation of the original directed graph model by providing a thermodynamically accurate representation of bimolecular reactions. The transition from simple directed edges to reaction hyperedges increases model fidelity without sacrificing computational tractability, as the bipartite transformation preserves compatibility with established optimization algorithms. The GSS metric further enhances the framework's utility by providing a single interpretable index for comparing synthesis pathways across different pharmaceutical compounds, facilitating decision-making in process development contexts where multiple sustainability objectives must be balanced simultaneously.

The GNN-based weight prediction module represents a meaningful advance toward framework's autonomy. By training on curated reaction database entries, the module eliminates the need for case-specific literature mining of activation energies and EIF values—a process that previously consumed substantial researcher effort and limited the framework's applicability to reactions with well-documented thermodynamic profiles. The demonstrated MAPE of 3.1% for ibuprofen synthesis suggests that GNN predictions are suitable for initial pathway screening, with empirical refinement reserved for the final selected route.

The cross-pharmaceutical analysis reveals an important relationship between synthesis network complexity and optimization gain. Simpler networks with fewer reaction nodes (aspirin, 4 nodes) tend to yield higher percentage EIF reductions because the cycle elimination step can target a relatively larger proportion of the total waste-generating network. More complex networks (paracetamol, 7 nodes) distribute waste generation across more reaction steps, reducing the marginal impact of individual cycle eliminations. This observation suggests that for complex multi-step syntheses, optimization benefit may be further enhanced by introducing modular sub-network decomposition, treating each synthetic stage as an independently optimizable sub-hypergraph.

The adaptive real-time framework addresses a critical industrial gap: the disconnect between static computational pathway planning and the dynamic reality of manufacturing environments. By coupling the hypergraph optimizer with PAT instrumentation, the framework enables continuous self-adjustment, a capability particularly valuable in continuous-flow manufacturing systems where residence times are short and rapid pathway reconfiguration can avert yield losses.

Limitations of the current study include the reliance on GNN models trained predominantly on gas-phase DFT data, which may underestimate solvent polarity effects on activation energies. Additionally, the paracetamol results presented here are model-generated and require experimental validation before industrial implementation can be considered. Future work should also address multi-objective Pareto optimization, enabling synthesis pathways to be selected based on user-defined trade-offs between yield, energy, and waste objectives rather than the fixed equal weighting adopted in the present GSS formulation.

7. Conclusion

This continuation study advances the original graph-theoretical aspirin synthesis optimization framework along three interconnected dimensions. By extending the representation from directed simple graphs to hypergraphs, the model achieves stoichiometrically accurate encoding of bimolecular and catalytic reactions—a prerequisite for applying the framework to the broader pharmaceutical synthesis landscape. The integration of GNN-based weight prediction substantially reduces the manual data-gathering burden, enabling rapid adaptation of the framework to new synthesis targets with minimal domain-expert intervention. Cross-validation on ibuprofen synthesis demonstrates that the optimized two-step pathway delivers a 94.2% yield, 57.8% waste reduction, and 43.7% energy savings relative to the conventional BHC process, with an aggregate Green Synthesis Score of 1.387. These results underscore the framework's potential for systematic optimization of established pharmaceutical manufacturing processes toward greener and more efficient alternatives, representing a significant step towards autonomous chemical process design. Furthermore, addressing the identified limitations by incorporating 3D molecular information and refining MILP constraint formulations could further enhance the predictive accuracy and overall effectiveness of the framework (Zhu et al., 2025).

The introduction of the GSS as a unified performance index addresses a long-standing gap in green chemistry benchmarking: the absence of a single, dimensionally consistent metric capable of cross-compound comparison. The prospective application to paracetamol synthesis further illustrates the framework's scalability, while the adaptive real-time optimization protocol provides a roadmap for translating computational synthesis planning into closed-loop industrial process control.

Collectively, these advances transforms the original framework from a single-compound optimization tool into a generalizable, semi-autonomous computational platform for sustainable pharmaceutical manufacturing. As pharmaceutical companies increasingly face regulatory and societal pressure to reduce environmental footprints while maintaining high production efficiency(Deng et al., 2025), such mathematically rigorous, data-driven frameworks offer a scientifically

credible pathway toward greener drug manufacturing practices.

Funding Information: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest: The authors declare no conflict of interest.

Ethical Statement: Not applicable. No experiments on human participants or animals were conducted.

References

1. Chung, Y., & Green, W. H. (2024). Machine learning from quantum chemistry to predict experimental solvent effects on reaction rates. Chemical Science, 15(7), 2410. https://doi.org/10.1039/d3sc05353a
2. Deng, D., Zhang, J., Wang, J., & Zong, X. (2025). Sustainable performance evaluation of pharmaceutical companies: sustainable balanced scorecard and hybrid MCDM approach. Frontiers in Public Health, 12, 1495156. https://doi.org/10.3389/fpubh.2024.1495156
3. Fiorini, S., Bovolenta, G. M., Coniglio, S., Ciavotta, M., Morerio, P., Parrinello, M., & Bue, A. D. (2024). DLGNet: Hyperedge Classification through Directed Line Graphs for Chemical Reactions. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2410.06969
4. Hattori, Y. (2022). Raman and near-infrared spectroscopy for in-line sensors. Analytical Sciences, 38(12), 1455. https://doi.org/10.1007/s44211-022-00202-7
5. Hu, X., Song, Y., Li, J., & Huang, Y. (2024). Modification of ibuprofen synthesis through the mechanism analysis. Theoretical and Natural Science, 45(1), 168. https://doi.org/10.54254/2753-8818/45/20240519
6. Kernaghan, S. M., Coady, T. M., Kinsella, M., & Lennon, C. M. (2024). A tutorial review for research laboratories to support the vital path toward inherently sustainable and green synthetic chemistry [Review of A tutorial review for research laboratories to support the vital path toward inherently sustainable and green synthetic chemistry]. RSC Sustainability, 2(3), 578. Royal Society of Chemistry. https://doi.org/10.1039/d3su00324h
7. Lyu, X., Seo, C., Jung, H., Faber, T., Kim, D., Seo, S., & Chang, S. (2023). Intramolecular hydroamidation of alkenes enabling asymmetric synthesis of β-lactams via transposed NiH catalysis. Nature Catalysis, 6(9), 784. https://doi.org/10.1038/s41929-023-01014-2
8. Martí, T. T., Marset, X., Guillem, C., Ramón, D. J., & Guillena, G. (2024). Electrophilic aromatic substitution in eutectic-type mixtures: from an old concept to new sustainable horizons. RSC Sustainability, 2(5), 1215. https://doi.org/10.1039/d4su00051j
9. Murphy, M. A. (2023). Professors Trost and Sheldon's Promotion of Catalytic Technologies, Atom Economy, and the E-Factor Metrics in Synthetic Organic Chemistry and the Fine Chemical and Pharmaceutical Industries, to Speed the Early Evolution of "Green Chemistry." Substantia, 7(2), 41. https://doi.org/10.36253/substantia-2140
10. Nuthi, A., Hirlekar, B. U., Singh, K. D., Rozy, F., Murty, U. S., & Dixit, V. A. (2023). Cumulative Green Chemistry Principle Score for Objective Assessment of Drug Synthesis. ACS Sustainable Chemistry & Engineering, 11(45), 16144. https://doi.org/10.1021/acssuschemeng.3c03961
11. Pahlavan, F., Moosavi, S. S., Zolghadr, A. R., & Iranpoor, N. (2023). Electronic origins of the stereochemistry in β-lactam formed through the Staudinger reaction catalyzed by a nucleophile. RSC Advances, 13(48), 33654. https://doi.org/10.1039/d3ra05286a
12. Patil, H., Vemula, S. K., Narala, S., Lakkala, P., Munnangi, S. R., Narala, N., Jara, M. O., Williams, R. O., Terefe, H., & Repka, M. A. (2024). Hot-Melt Extrusion: from Theory to Application in Pharmaceutical Formulation—Where Are We Now? AAPS PharmSciTech, 25(2). https://doi.org/10.1208/s12249-024-02749-2
13. Righetti, G. I. C., Gambarotti, C., & Bjørsvik, H. (2024). Development of a green scalable route toward the synthesis of bio-based 2-pyrones. Reaction Chemistry & Engineering, 9(5), 1140. https://doi.org/10.1039/d3re00667k
14. Roy, S., Vargas, D. A., Ma, P., Sengupta, A., Zhu, L., Houk, K. N., & Fasan, R. (2023). Stereoselective construction of β-, γ- and δ-lactam rings via enzymatic C–H amidation. Nature Catalysis, 7(1), 65. https://doi.org/10.1038/s41929-023-01068-2
15. Sagmeister, P., Hierzegger, R., Williams, J. D., Kappe, C. O., & Kowarik, S. (2022). Artificial neural networks and data fusion enable concentration predictions for inline process analytics. Digital Discovery, 1(4), 405. https://doi.org/10.1039/d2dd00006g
16. Sanjai, S. K. R., & Nagarathinam, R. (2025). Graph-theoretical optimization of the aspirin synthesis pathway: Enhancing green chemistry in pharmaceutical manufacturing. Computational and Mathematical Biophysics, 13(1). https://doi.org/10.1515/cmb-2025-0027
17. Saura-Sanmartín, A., & Andreu-Ardil, L. (2023). Stereoselective synthesis of β-lactams: recent examples. Organic & Biomolecular Chemistry, 21(16), 3296. https://doi.org/10.1039/d3ob00309d
18. Tavakoli, M., Mood, A., Vranken, D. L. V., & Baldi, P. (2022). Quantum Mechanics and Machine Learning Synergies: Graph Attention Neural Networks to Predict Chemical Reactivity

[Review of Quantum Mechanics and Machine Learning Synergies: Graph Attention Neural Networks to Predict Chemical Reactivity]. Journal of Chemical Information and Modeling, 62(9), 2121. American Chemical Society. https://doi.org/10.1021/acs.jcim.1c01400

19. Tomme, L., Ureel, Y., Dobbelaere, M. R., Lengyel, I., Vermeire, F. H., Stevens, C. V., & Geem, K. M. V. (2024). Machine learning applications for thermochemical and kinetic property prediction [Review of Machine learning applications for thermochemical and kinetic property prediction]. Reviews in Chemical Engineering. De Gruyter. https://doi.org/10.1515/revce-2024-0027
20. Zainul, R., Abdullah, M. N., Saeed, S. M., Idan, A. H., Alsultany, N. M. A., Arshadi, S., Behmagham, F., & Vessally, E. (2024). Recent trends in incorporation of CO2 into organosulfur compounds via C–S bond cleavage. RSC Advances, 14(22), 15680. https://doi.org/10.1039/d4ra02405b
21. Zhu, J., Azam, N. A., Haraguchi, K., Zhao, L., & Akutsu, T. (2025). Combining Graph Neural Networks and Mixed Integer Linear Programming for Molecular Inference under the Two-Layered Model. arXiv (Cornell University). http://arxiv.org/abs/2507.03920