Chemistry

E-ISSN No : 2581-6195 | Volume : 9 | Issue : 7 | July 2026

# STRUCTURE–ACTIVITY RELATIONSHIP AND GREEN CHEMISTRY ASSESSMENT OF BIOACTIVE HETEROCYCLIC COMPOUNDS: AN INTEGRATED EXPERIMENTAL AND COMPUTATIONAL APPROACH

**Dalwadi Malav[1], Dr. Mikita Shah[1]**

## ABSTRACT

Bioactive heterocyclic compounds provide adaptable molecular frameworks for exploring antioxidant and antimicrobial properties, yet the prioritization of promising analogues requires more than ranking biological response alone. This article presents an integrated workflow that links molecular modification, experimental activity profiling, computational interpretation, and green chemistry assessment. Structure–activity relationship (SAR) analysis is framed around changes in electronic character, steric demand, hydrogen-bonding capacity, lipophilicity, polarity, and scaffold geometry. Computational tools—including molecular descriptor calculation, density functional theory (DFT), and molecular docking where a justified biological target is available—are positioned as complementary methods for interpreting experimental trends rather than substitutes for biological validation. Sustainability is evaluated using isolated yield, process mass intensity (PMI), E-factor, solvent burden, catalyst recovery, reaction time, and energy use under explicitly defined process boundaries. Worked-example tables illustrate how compounds can be compared through a multi-parameter decision framework without implying that the example values are empirical. The proposed approach supports transparent selection of candidates that balance activity, physicochemical suitability, synthetic practicality, and environmental performance. Its principal recommendation is to use experimental evidence, computational hypotheses, and sustainability metrics together, while retaining clear distinctions between measured data, predicted properties, and illustrative scenarios.

**KEYWORDS:** Heterocyclic Compounds, Structure–Activity Relationship, QSAR, Molecular Docking, DFT, Green Chemistry, Process Mass Intensity, Antioxidant Activity, Antimicrobial Activity, Lead Prioritization

[1] Department of Chemistry, School of Pure & Applied Sciences, Sabarmati University, Ahmedabad, Gujarat, India

**How to Cite:**
Dalwadi Malav & Dr. Mikita Shah(2026), Structure–Activity Relationship and Green Chemistry Assessment of Bioactive Heterocyclic Compounds: an Integrated Experimental and Computational Approach, International Educational Journal of Science & Engineering (IEJSE),Vol: 9, Issue: 7, 01-07

Research Paper

## 1. INTRODUCTION AND BACKGROUND OF THE STUDY

Heterocyclic compounds are central to medicinal and bioorganic chemistry because ring heteroatoms and modular substitution patterns permit systematic adjustment of molecular recognition, polarity, hydrogen bonding, conformation, and electronic distribution. Nitrogen-, oxygen-, and sulfur-containing scaffolds can therefore serve as useful starting points for analogue design. Nevertheless, the presence of a heterocycle or a favorable computational score does not establish biological efficacy. A defensible assessment requires reproducible synthesis, structural confirmation, appropriate controls, and biological testing using methods suited to the endpoint.

Classical quantitative structure–activity relationship approaches relate changes in molecular structure to changes in measured response. Substituent constants and physicochemical descriptors can help formulate hypotheses about electronic, hydrophobic, and steric effects, but correlations are sensitive to the quality and size of the dataset, the chemical domain represented, and the consistency of the

Copyright© 2026, IEJSE. This open-access article is published under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License which permits Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material) under the Attribution-NonCommercial terms.

biological assay [1,2]. A valid SAR interpretation therefore considers matched molecular pairs where possible, separates observed trends from proposed mechanisms, and avoids overgeneralization from small compound libraries.

Computational chemistry extends this analysis by describing molecular geometry, charge distribution, frontier orbital properties, hydrogen-bonding features, conformational preferences, and possible ligand–target interactions. Molecular docking can help generate structural hypotheses when a relevant target and reliable structure are available; however, docking scores are approximate and pose prediction, protein flexibility, protonation, solvent effects, and scoring-function limitations can affect rankings [3]. Computational predictions should consequently be evaluated alongside measured activity and, when appropriate, orthogonal experimental evidence.

Green chemistry assessment provides a second axis of evaluation. A route with high isolated yield may still require excessive solvent, difficult purification, high energy input, or stoichiometric reagents that generate substantial waste. Metrics such as atom economy, reaction mass efficiency, PMI, and E-factor capture different aspects of material efficiency, while solvent selection guides and catalyst-reuse studies help characterize process choices [4–7]. No single metric represents the entire environmental profile; comparisons must define the system boundary and include reaction, work-up, and purification consistently.

The purpose of this article is to describe a combined framework in which biological response, SAR, computational interpretation, and process sustainability inform candidate prioritization. Because no compound-specific raw data were provided, the tables include clearly labelled illustrative values and qualitative examples rather than fabricated experimental claims.

## 2. AIMS AND OBJECTIVES

- To establish a systematic approach for relating structural modifications in heterocyclic compounds to antioxidant and antimicrobial responses.
- To define a complementary computational workflow using physicochemical descriptors, DFT-derived properties, and molecular docking where target biology supports its use.
- To identify how green chemistry metrics can be calculated and interpreted alongside biological and molecular properties.
- To propose a transparent multi-criteria framework for prioritizing compounds without allowing a single assay, docking score, or yield value to dominate the decision.
- To demonstrate data reporting using pre-filled illustrative tables while clearly distinguishing example values from verified experimental observations.

## 3. METHODOLOGY AND INTEGRATED ASSESSMENT FRAMEWORK

### 3.1 Molecular library design and experimental characterization

A useful analogue set begins with a defined parent scaffold and a small number of purposeful structural changes. Substituent selection may include electron-donating groups, electron-withdrawing groups, halogens, heteroaromatic replacements, hydrogen-bond donors or acceptors, and changes in linker length or ring fusion. Whenever feasible, matched pairs should differ in one principal feature so that observed activity changes are more interpretable. Compound identity should be confirmed using suitable combinations of FTIR, 1H and 13C NMR, mass spectrometry, melting point, elemental analysis, and chromatographic purity assessment. A proposed structure should not be considered confirmed solely on the basis of one diagnostic peak.

### 3.2 Biological activity and quality assurance

Antioxidant profiling may include DPPH radical scavenging, ABTS radical-cation decolorization, or FRAP reducing-capacity measurement. These assays report related but non-identical chemical responses, so assay conditions, concentration ranges, reaction time, solvent composition, blanks, reference standards, replicate counts, and data normalization should be stated [8–10]. Concentration–response curves are preferable to comparing a single concentration. IC50 values should only be reported when the measured response spans the relevant response range and the fitting method is documented.

Antimicrobial activity should be evaluated using an appropriate standardized procedure, organism identity, strain designation, inoculum preparation, medium, incubation conditions, and positive and negative controls. MIC values obtained through dilution methods are generally more informative than an unstandardized inhibition-zone comparison, although the choice depends on the study design and organism. Biological results should be reported with replicate information and uncertainty, and any apparent inactivity should be interpreted in light of solubility, stability, and assay compatibility.

### 3.3 Computational interpretation

Descriptor analysis can include molecular weight, calculated lipophilicity, topological polar surface area, hydrogen-bond donor/acceptor counts, rotatable bonds, and formal charge. DFT calculations may be used to examine optimized geometry, electrostatic potential, charge distribution, and frontier orbital characteristics, provided the method, basis set, solvent treatment, and convergence criteria are reported. Docking is appropriate only when a biologically relevant macromolecular target is justified; the protein preparation, ligand protonation, binding-site definition, search settings, validation procedure, and pose-inspection criteria should be documented. Predicted interactions are mechanistic hypotheses, not direct proof of binding or activity.

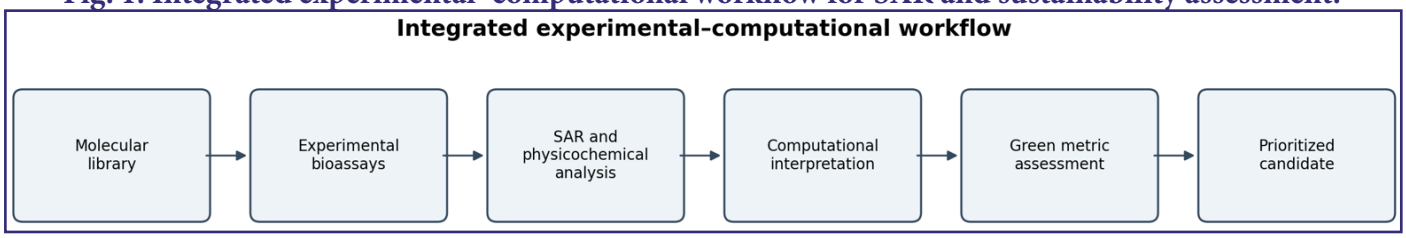
### 3.4 Green chemistry metrics and process boundaries

A fair sustainability comparison should state whether solvent used in reaction, extraction, washing, chromatography, and recrystallization is included. PMI is calculated as the total mass of included process inputs divided by the mass of isolated product. E-factor is the mass of waste divided by the mass of product; its value depends on whether water and other streams are counted. Atom economy is a stoichiometric property of the reaction equation and does not reflect excess reagents or purification burden. Reaction mass efficiency includes practical yield and stoichiometric efficiency. Energy consumption, catalyst recovery, solvent hazard, and repeatability should be reported separately when they are not represented adequately by a mass metric [4–7].

### 3.5 Integrated interpretation and candidate prioritization

The framework uses sequential evidence checks rather than a single opaque score: confirm chemical identity and purity; establish reproducible biological response; examine matched-pair SAR; inspect computational explanations; calculate process metrics within a defined boundary; then rank candidates using explicit criteria. A compound with strong activity but poor solubility, uncertain identity, or an impractical synthesis should not automatically outrank a slightly less active compound with more reliable evidence and a substantially improved process profile.

**Fig. 1: Integrated experimental–computational workflow for SAR and sustainability assessment.**




## 4. RESULTS AND DISCUSSION: WORKED-EXAMPLE ANALYSIS FRAMEWORK

The following tables are illustrative worked examples intended to show how a future dataset could be organized. The compound labels, descriptor values, activity values, and process values are not reported measurements. They must be replaced by verified observations and computational outputs before the manuscript is represented as an experimental research article.

**Table 1: Illustrative heterocyclic analogue design and intended SAR question**

| Code | Scaffold / modification | Primary structural question | Key descriptors or evidence to examine |
|---|---|---|---|
| H1 | Thiazole analogue with para-hydroxyphenyl substituent | Does an additional hydrogen-bond donor and phenolic group alter radical-scavenging response? | H-bond donor count; polarity; phenolic O–H assignment; concentration–response |
| H2 | Pyrazole analogue with para-methoxyphenyl substituent | Does an electron-donating substituent change electronic distribution and assay response? | Substituent constants; calculated charge distribution; methoxy signal confirmation |
| H3 | Benzimidazole analogue with para-halogen substituent | Does halogen substitution alter lipophilicity and antimicrobial response? | Calculated logP; polarizability; solubility; matched-pair MIC comparison |
| H4 | Unsubstituted parent scaffold | What is the baseline response against which analogue changes are compared? | Parent identity; purity; matched assay conditions; baseline descriptors |

**Table 2: Illustrative descriptor profile for demonstrating comparative interpretation**

| Code | MW (g/mol) | cLogP | tPSA (Å²) | HBD | HBA | Interpretive note |
|---|---|---|---|---|---|---|
| H1 | 280.3 | 2.1 | 82.4 | 2 | 5 | Example only: greater polarity and donor capacity may influence solubility and hydrogen bonding |
| H2 | 292.3 | 2.7 | 61.8 | 1 | 5 | Example only: methoxy substitution changes acceptor count and electronic character |
| H3 | 310.2 | 3.4 | 48.6 | 1 | 3 | Example only: greater lipophilicity may affect permeability but can reduce aqueous solubility |
| H4 | 250.3 | 1.8 | 55.2 | 1 | 3 | Example only: parent analogue provides a reference for matched-pair comparison |

**Table 3: Illustrative SAR matrix showing how observed activity would be interpreted**

| Comparison | Illustrative response pattern | Potential interpretation | Required check before conclusion |
|---|---|---|---|
| H1 vs H4 | H1 shows stronger DPPH response at matched concentrations | Phenolic functionality may contribute to hydrogen transfer or electron transfer | Confirm purity, solubility, sample blank, concentration dependence and replicate variation |
| H2 vs H4 | H2 shows modestly improved ABTS response | Electronic effects may influence assay-specific reactivity | Do not infer a universal antioxidant effect from one assay; compare orthogonal assays |
| H3 vs H4 | H3 shows lower MIC in a hypothetical matched assay | Halogen substitution may alter target interaction or cell access | Confirm strain, standardized MIC method, solubility and independent repeats |
| All analogues | Activity order differs between antioxidant assays | Assays probe different chemical mechanisms and conditions | Report each endpoint separately and avoid collapsing into a single unqualified activity label |

**Table 4: Illustrative green chemistry assessment for two alternative routes**

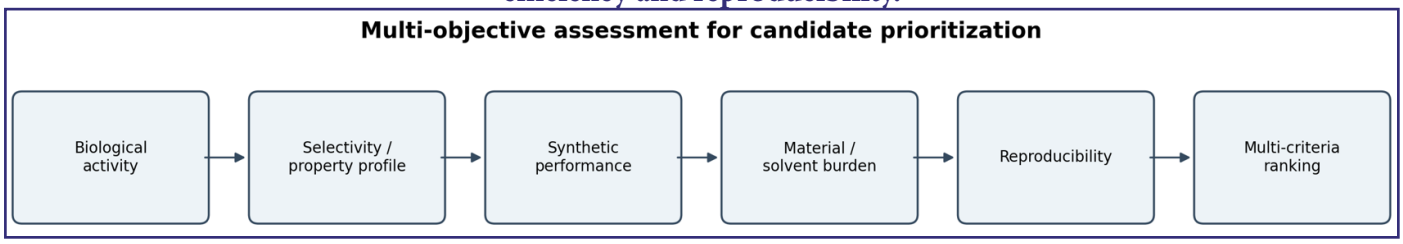
| Metric | Conventional route (example) | Candidate greener route (example) | Interpretation and boundary |
|---|---|---|---|
| Isolated yield | 72% | 86% | Example values only; correct for dry product and state purity basis |
| Reaction time | 6 h | 2 h | Example values only; report heating and cooling conventions consistently |
| Total process solvent | 30 mL per batch | 15 mL per batch | Example values only; include work-up and purification solvent in both routes |
| PMI | 18 kg input/kg product | 10 kg input/kg product | Example values only; define included inputs and product mass basis |

| | | | |
|---|---|---|---|
| E-factor | 17 kg waste/kg product | 9 kg waste/kg product | Example values only; explicitly state treatment of water and recovered solvent |
| Catalyst recovery | Not applicable / homogeneous | 85% recovered by mass | Example only; recovered mass alone does not establish retained catalytic activity |
| Energy use | Not measured | Not measured | No energy-saving claim can be made until power and equipment duration are measured or transparently estimated |

**Table 5: Illustrative multi-criteria decision matrix for candidate prioritization**

| Criterion | Suggested evidence | Illustrative weighting | Decision rule |
|---|---|---|---|
| Identity and purity | Spectral confirmation and purity result | Gatekeeper | Do not rank a compound as a lead until identity and adequate purity are established |
| Biological activity | Replicated concentration–response or standardized MIC | 30% | Prefer reproducible activity supported by controls and uncertainty estimates |
| SAR coherence | Matched-pair comparisons and plausible structure–property trend | 15% | Treat small or inconsistent differences cautiously |
| Computational support | Descriptors, validated QSAR or carefully interpreted docking/DFT | 15% | Use as corroboration; predictions must remain within applicability limits |
| Synthetic efficiency | Yield, reaction time, PMI, E-factor, purification burden | 25% | Reward lower material burden and reproducible, practical isolation |
| Safety and robustness | Solvent hazard, catalyst recovery, repeatability, scalability | 15% | Penalize hazardous inputs and fragile procedures; document trade-offs |

**Fig. 2: Multi-objective assessment for prioritizing compounds across activity, properties, process efficiency and reproducibility.**




### 4.1 Interpreting structure–activity patterns

SAR is strongest when comparisons are made between closely related analogues tested under the same conditions. If a para-substituent is the only structural change, a response difference can more plausibly be associated with that modification than when several structural features change simultaneously. Even in matched pairs, the explanation may involve several linked effects: electron distribution, lipophilicity, hydrogen bonding, conformation, aggregation, solubility, or assay interference. Therefore, a trend should first be described as an observation and only then discussed as a mechanistic possibility.

A substituent that improves one antioxidant assay may not improve another because DPPH, ABTS, and FRAP differ in reaction chemistry and solvent conditions. Likewise, a compound with a lower predicted docking energy is not necessarily more antimicrobial. A docking result is most useful when the target is biologically relevant, the protocol is validated against known ligands or structural evidence, and predicted interactions can be connected to testable hypotheses. The strongest SAR account combines experimental endpoints with carefully bounded interpretation.

### 4.2 Interpreting computational outputs

Descriptor analysis is useful for detecting obvious differences in molecular size, polarity, lipophilicity, and hydrogen-bonding capacity. It does not by itself establish drug-likeness or biological mechanism. DFT-derived quantities can help compare electronic features, but their interpretation depends on the computational method and optimized structures. QSAR modelling requires a sufficient, chemically

relevant dataset, consistent endpoint definitions, internal validation, external validation where feasible, and an explicit applicability domain. OECD guidance emphasizes validation principles for regulatory and predictive use of QSAR models [2]. For a small exploratory library, descriptive plots and matched-pair analysis may be more defensible than fitting a complex predictive model.

**4.3 Interpreting sustainability alongside activity**

Sustainability should be evaluated for the complete process rather than the reaction flask alone. A high-yield reaction can have a high PMI if large solvent volumes or extensive chromatography are required. A low E-factor can be misleading if water or solvent recovery is excluded from the boundary. Similarly, catalyst recovery by mass should be complemented by activity retention across reuse cycles. Solvent selection should consider health, safety, and environmental profiles rather than assuming that any bio-based or aqueous solvent is automatically benign [5,6]. Reporting multiple complementary indicators makes trade-offs visible and reduces the risk of overstating greenness.

**4.4 Practical lead-prioritization logic**

Candidate prioritization should begin with non-negotiable quality gates: confirmed identity, adequate purity, reproducible assay response, and absence of obvious assay interference. Compounds that pass these gates can be compared using biological potency, selectivity where measured, physicochemical suitability, SAR coherence, computational support, and process sustainability. Weightings in Table 5 are an example of a transparent decision aid, not a universal scoring system. They should be defined before ranking, justified by the study objectives, and tested for sensitivity to alternative weights. A promising candidate remains a preliminary lead until independent replication, selectivity testing, toxicity evaluation, and appropriate follow-up studies support further development.

## 5. LIMITATIONS AND RECOMMENDATIONS

The central limitation of the present manuscript is that it is a framework and worked-example article rather than a report of a completed compound-specific experimental campaign. The numerical values are illustrative and must not be interpreted as evidence that the proposed analogues were synthesized or tested. In a completed study, all compounds should be described using exact chemical identities and analytical data; all biological values should include methods, controls, replicates and uncertainty; and all computational outputs should identify software, versions, parameters, structures and validation procedures.

Future work should prioritize a manageable analogue library, matched-pair design, standardized biological assays, orthogonal antioxidant methods, and appropriate antimicrobial reference controls. Computational models should be matched to the size and quality of the dataset, and docking should be restricted to justified targets. Green metrics should use consistent boundaries that include purification and work-up. Where feasible, solvent recovery, catalyst recycling, energy measurement, and repeatability should be documented. These practices would enable the proposed framework to produce evidence that is reproducible, transparent, and suitable for independent assessment.

## 6. CONCLUSION

An integrated SAR and green chemistry framework provides a practical way to assess bioactive heterocyclic compounds across biological response, molecular properties, computational interpretation, and synthetic sustainability. Its value lies in combining distinct evidence streams without treating any one of them as definitive. Experimental activity establishes the observed response; SAR organizes structural trends; computation proposes explanations and prioritizes testable hypotheses; and green chemistry metrics characterize the resource and waste profile of the route. The resulting decisions should be based on verified data, explicit process boundaries, reproducible assays, and transparent multi-criteria rules. The illustrative tables in this manuscript demonstrate reporting structure only; measured and validated compound-specific data are required before scientific claims can be made about particular analogues or routes.

## REFERENCES

1. Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. Aromatic Substituent Constants for Structure–Activity Correlations. Journal of Medicinal Chemistry 1973, 16(11), 1207–1216. https://doi.org/10.1021/jm00269a003

2. OECD. Guidance Document on the Validation of (Quantitative) Structure–Activity Relationship [(Q) SAR] Models. OECD Series on Testing and Assessment No. 69; OECD Publishing: Paris, 2014. https://doi.org/10.1787/9789264085442-en
3. Paggi, J. M.; Pandit, A.; Dror, R. O. The Art and Science of Molecular Docking. Annual Review of Biochemistry 2024, 93, 389–410. https://doi.org/10.1146/annurev-biochem-030222-120000
4. Sheldon, R. A. Metrics of Green Chemistry and Sustainability: Past, Present, and Future. ACS Sustainable Chemistry & Engineering 2018, 6(1), 32–48. https://doi.org/10.1021/acssuschemeng.7b03505
5. Prat, D.; Hayler, J.; Wells, A. A Survey of Solvent Selection Guides. Green Chemistry 2014, 16, 4546–4551. https://doi.org/10.1039/C4GC01149J
6. Byrne, F. P.; Jin, S.; Paggiola, G.; Petchey, T. H. M.; Clark, J. H.; Farmer, T. J.; Hunt, A. J.; McElroy, C. R.; Sherwood, J. Tools and Techniques for Solvent Selection: Green Solvent Selection Guides. Sustainable Chemical Processes 2016, 4, 7. https://doi.org/10.1186/s40508-016-0051-z
7. Jimenez-Gonzalez, C.; Ponder, C. S.; Broxterman, Q. B.; Manley, J. B. Using the Right Green Yardstick: Why Process Mass Intensity Is Used in the Pharmaceutical Industry to Drive More Sustainable Processes. Organic Process Research & Development 2011, 15(4), 912–917. https://doi.org/10.1021/op200097d
8. Brand-Williams, W.; Cuvelier, M. E.; Berset, C. Use of a Free Radical Method to Evaluate Antioxidant Activity. LWT – Food Science and Technology 1995, 28(1), 25–30. https://doi.org/10.1016/S0023-6438(95)80008-5
9. Re, R.; Pellegrini, N.; Proteggente, A.; Pannala, A.; Yang, M.; Rice-Evans, C. Antioxidant Activity Applying an Improved ABTS Radical Cation Decolorization Assay. Free Radical Biology and Medicine 1999, 26(9–10), 1231–1237. https://doi.org/10.1016/S0891-5849(98)00315-3
10. Benzie, I. F. F.; Strain, J. J. The Ferric Reducing Ability of Plasma (FRAP) as a Measure of "Antioxidant Power": The FRAP Assay. Analytical Biochemistry 1996, 239(1), 70–76. https://doi.org/10.1006/abio.1996.0292