Specialty Journal of Pharmacognosy, Phytochemistry, and Biotechnology
ISSN: 3062-441X
2024, Volume 4, Issue 2, Page No: 40-49

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

How Much Structure Has the Spectrum Actually Proven? A Confidence-Graded Workflow for Combining HRMS, NMR, Chiroptical Data, and Biosynthetic Plausibility

Mikko Lahtinen[1]*, Elina Salo[1], Juhani Virtanen[2]

[1] Department of Structure Elucidation and Spectral Confidence, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland.
[2] Department of Chiroptical Data and Biosynthetic Plausibility, Faculty of Pharmacy, University of Eastern Finland, Kuopio, Finland.

***E-mail** ✉ mikko.lahtinen@helsinki.fi

Received: 05 May 2024; **Revised:** 03 August 2024; **Accepted:** 04 August 2024

ABSTRACT

Natural-product structures are commonly communicated as finished molecular identities, yet the analytical evidence underlying a drawing is heterogeneous in what it can establish. High-resolution mass spectrometry constrains elemental composition and fragmentation-compatible substructures; NMR defines connectivity and, with suitable restraints or calculations, stereochemical relationships; chiroptical methods can support absolute configuration; and biosynthetic reasoning can test chemical coherence. None of these evidential roles is interchangeable. This article develops an original, proposed analytical confidence scheme that treats structure elucidation as a sequence of claim-specific inferences rather than a binary identified/not-identified decision. The scheme separates observed data, candidate generation, candidate discrimination, stereochemical assignment, plausibility assessment, conflict detection, and reporting. It also distinguishes incomplete evidence from contradictory evidence: missing support should lower resolution, whereas material incompatibility should reopen candidate space or trigger escalation to more discriminating measurements. The central proposal is that confidence should attach to the structural proposition actually tested—composition, constitution, relative configuration, absolute configuration, or full molecular assignment—rather than to the molecule as a whole. Orthogonality is therefore judged by whether a new modality contributes an independently constraining observable, not merely another score derived from the same assumptions. The framework is intended to improve transparency in natural-product structure assignment and revision, but it is not a validated probability model and does not supply universal numerical thresholds. Its applicability remains limited by sample availability, spectral quality, conformational complexity, candidate-space completeness, computational domain, and the feasibility of orthogonal validation.

Keywords: Natural products, Structure elucidation, HRMS, NMR, Chiroptical spectroscopy, Structural revision

How to Cite This Article: Lahtinen M, Salo E, Virtanen J. How Much Structure Has the Spectrum Actually Proven? A Confidence-Graded Workflow for Combining HRMS, NMR, Chiroptical Data, and Biosynthetic Plausibility. Spec J Pharmacogn Phytochem Biotechnol. 2024;4(2):40-9. https://doi.org/10.51847/NeG6SB9yyV

Introduction

A natural-product structure is often presented as a single resolved identity, but the route from experimental signal to molecular drawing contains multiple inferential steps. A decade-scale analysis of marine natural-product misassignments showed that errors can arise from incorrect constitutional interpretation, stereochemical assignment, spectral reasoning, or overreliance on apparently persuasive but incomplete evidence [1]. The important implication is not that natural-product structures are generally unreliable. Rather, the confidence appropriate to a structural claim should reflect which parts of the structure were directly constrained, which were computationally or chemically inferred, and whether plausible alternatives were actively excluded.

The granularity of the claim matters because “the structure” contains several propositions that may have unequal evidential support. An accurate elemental formula does not determine how atoms are connected; convincing connectivity does not necessarily specify every relative stereochemical relationship; and a relative configuration

does not establish absolute configuration. Treating these propositions as one confidence object can therefore hide where uncertainty resides. The proposed scheme uses the highest structurally meaningful resolution actually supported by the observations, while permitting a molecule to be well established at one level and provisional at another.

Structural revision provides a second reason to avoid binary language. Reviews of revised natural products show that accepted assignments can be overturned by reinterpreted spectra, synthesis, improved stereochemical analysis, or newly available orthogonal evidence [2]. Revision should therefore be treated as a normal corrective mechanism in structural science, not as an anomalous failure that can be prevented simply by accumulating more data of the same type. A structurally dense molecule may remain underdetermined even when its spectra are technically high quality if those spectra do not discriminate the critical alternatives.

Contemporary structure elucidation has expanded substantially through advanced NMR, computation, anisotropic measurements, crystallography, and integrated workflows, yet difficult cases still persist because material can be scarce, observables can be ambiguous, and structurally similar candidates can fit substantial portions of the same dataset [3]. The recurring problem is thus epistemic as well as technical: what exactly has each analytical observation established, and how far may the interpretation legitimately extend beyond it?

This distinction also affects how disagreement is handled. If two methods interrogate different structural properties, concordance can strengthen an assignment; if they depend on the same candidate set, conformer ensemble, or calculated observable, apparent agreement may be partly redundant. The article therefore uses "orthogonal" in an evidential rather than merely technological sense. The question is not whether two instruments or algorithms differ, but whether the second result adds an independently constraining reason to prefer, reject, or leave unresolved a structural proposition.

This article proposes a confidence-graded workflow for answering that question. The framework does not assign universal percentages to structures and does not treat one technology as intrinsically superior. Instead, it separates five claim levels—composition, constitution, relative configuration, absolute configuration, and integrated full-structure assignment—and asks whether each is supported, unresolved, or contradicted by the available evidence. The proposal further distinguishes direct analytical constraints from computational discrimination and biosynthetic plausibility. Confidence is increased only when an additional modality constrains a previously uncertain proposition without merely reproducing the assumptions of an earlier step. Conversely, unresolved contradiction triggers reassessment rather than averaging. This structure-specific treatment of evidence is intended to make uncertainty visible, guide escalation when necessary, and make subsequent structural revision easier to interpret.

Structure elucidation as a layered evidence problem

Automated annotation illustrates why structural confidence cannot be reduced to a single score. Machine-learning systems can extract useful structure-related information from MS and NMR data, but annotation remains an inferential task whose output depends on the information content of the spectra, the representation used, and the domain of the model [4]. A candidate name generated by an algorithm therefore belongs to a different evidential layer from the measured peaks that produced it.

Chemical reasoning adds another layer. Mechanistic and biosynthetic questions can expose implausible assignments or make one candidate more coherent than another, but plausibility should complement rather than replace analytical interrogation [5]. A biosynthetically elegant proposal is still vulnerable if a decisive spectral relationship is incompatible with it. In the proposed scheme, such reasoning functions as corroboration or challenge, not as a substitute for constitution- or stereochemistry-sensitive observations.

The same distinction applies to increasingly rapid spectral annotation. Convolutional neural networks have demonstrated that structurally diverse natural products can be annotated from learned spectral patterns with substantial efficiency [6]. That achievement changes search and prioritization, but it does not change the logical status of the evidence: model output is an interpretation of measured data, and its reliability can weaken when the target lies outside the represented chemical space.

This layered view also prevents evidence laundering across computational steps. Formula generation, database retrieval, spectral prediction, and probabilistic ranking may occur sequentially, yet several stages can inherit the same initial candidate-space assumptions. Multiple favorable outputs are therefore not automatically multiple independent confirmations. In the proposed framework, provenance is retained: an analyst should be able to

identify which measured observable supplied a constraint, which transformation interpreted it, and which structural proposition received the resulting support.

Accordingly, structure elucidation is treated here as layered inference. The first layer contains observations: accurate mass, isotope pattern, fragment ions, chemical shifts, correlations, coupling information, residual dipolar couplings, optical or vibrational responses, and other measured observables. The second contains candidate constraints and comparisons derived from those measurements. The third contains structural propositions about formula, connectivity, relative configuration, and absolute configuration. A final interpretive layer integrates chemical coherence, biosynthetic plausibility, and conflict across modalities. The proposed confidence grade attaches to each structural proposition at this third layer and is then carried into the integrated assignment. This prevents strong evidence for one proposition—for example, elemental composition—from being silently transferred to another, such as absolute configuration.

What different analytical modalities can establish

HRMS and tandem MS are powerful for narrowing composition and candidate space, but their principal evidential strength lies in mass accuracy and fragmentation-compatible structural information rather than complete stereochemical proof. SIRIUS demonstrates how fragmentation trees can convert tandem spectra into molecular-formula and candidate-structure information [7]. CANOPUS extends fragmentation evidence to systematic compound-class prediction, illustrating that class-level inference can be strong while exact constitutional identity remains unresolved [8]. MSNovelist further shows that tandem spectra can support de novo structure generation, but generated structures are hypotheses requiring downstream discrimination [9]. Retention order can add orthogonal chromatographic information and improve candidate ranking when it is measured under relevant LC conditions [10].

NMR provides more direct constraints on constitution and three-dimensional relationships, although its resolving power depends on the observables obtained. CASE-3D shows that isotropic and anisotropic NMR parameters can be combined to strengthen three-dimensional assignment [11]. Quantum-chemical NMR calculations provide a complementary means of comparing candidate structures against experimental shifts and couplings [12]. DP4+ is useful for discriminating candidates, but its probabilities remain conditional on the candidate set and computational protocol rather than representing unconditional probability that a structure is true [13]. Floating-chirality distance geometry can reduce otherwise large stereochemical search spaces when adequate NMR restraints are present [14]. DP4+App improves efficiency and standardization of those comparisons without creating a new independent evidence source [15]. Simulated HSQC matching likewise provides a discriminating spectral fingerprint, but a match score remains dependent on the quality of the candidate and simulation [16].

Chiroptical evidence addresses stereochemical propositions that mass spectrometry cannot resolve directly. Natural-product case studies demonstrate both the value and the difficulty of chiroptical assignment when conformational or spectral behavior complicates interpretation [17]. Empirical induced-circular-dichroism rules can also have sharply bounded substrate applicability [18]. Vibrational circular dichroism can support absolute configuration through comparison of experimental and calculated spectra [19]. Across such computationally assisted methods, conformational sampling is a critical boundary because inadequately represented conformer populations can distort the calculated observable used for comparison [20].

Biosynthetic plausibility occupies a different category. It can test whether a proposed scaffold, oxidation state, substitution pattern, or stereochemical outcome is chemically coherent with a credible pathway, but it normally operates downstream of direct analytical constraints [5]. The framework therefore treats biosynthesis as contextual corroboration or as a reason to reopen scrutiny when an assignment is chemically anomalous, not as an evidential shortcut around unresolved spectral ambiguity.

These modality-specific capabilities and limits are summarized in **Table 1**.

Table 1. Evidence layers contributed by major modalities in natural-product structure assignment

Evidence source	Structural proposition best constrained	What it can contribute	What it does not establish by itself
HRMS/MS	Composition; substructure/class	Formula, isotope compatibility, fragmentation-supported candidates	Unique constitution or stereochemistry
NMR	Constitution; relative configuration	Connectivity, proximity, coupling, anisotropic 3D constraints	Automatic absolute configuration

Computed NMR	Candidate discrimination	Relative fit among explicit structural hypotheses	Truth outside the candidate set
Chiroptical data	Absolute-configuration hypotheses	Stereo-sensitive experimental–computational agreement	Conformer-independent certainty
Biosynthetic plausibility	Integrated coherence	Chemical and pathway consistency	Override of contradictory analytical evidence

Note: This classification is proposed as an interpretive aid. The evidential value of each modality remains conditional on data quality, method applicability, and the specific structural proposition being tested.

Conflicting and incomplete structural evidence

The most important distinction in confidence grading is between evidence that is absent and evidence that is incompatible. A missing stereochemical restraint means that a proposition remains unresolved; it does not count against a candidate in the same way as an observed relationship that the candidate cannot reproduce. The reassignment of (+)-diplopyrone illustrates how convergent computational NMR analyses can expose an accepted stereochemical proposal as inconsistent with multiple candidate-discrimination metrics [21]. Such a result should reopen the assignment rather than merely lower a global confidence score.

Conflicts can also be localized. Diagnostic ^13C chemical-shift behavior was sufficient to support revision across several classes of prenylated aromatic natural products, showing that a recurring local inconsistency may reveal a larger structural error [22]. Database-supported analysis introduces another complication: if nearly identical spectral information becomes associated with divergent structures, the apparent existence of multiple literature precedents may reflect propagated assignment error rather than independent confirmation. Computational reassessment of triterpenoid records using NAPROC-13 and DFT demonstrates how such conflicts can be detected and adjudicated [23].

When disagreement survives routine reassessment, escalation to a more discriminating measurement can be necessary. The revision of wheldone combined DFT-GIAO chemical-shift calculations, 1,1-HD-ADEQUATE NMR, and X-ray crystallography to resolve a contested structure [24]. The general lesson is conditional: escalation is valuable when the disputed proposition is material to the final assignment and the additional modality supplies genuinely new constraints. It is not a requirement that every natural product undergo every available technique.

A contradiction must also be qualified before it is allowed to dominate the assignment. Apparent incompatibility may arise from peak misassignment, sample impurity, incorrect conformer populations, inappropriate computational settings, or a method applied outside its suitable chemical domain. The workflow therefore requires an explicit attempt to explain the discrepancy. If a credible methodological cause is demonstrated, the conflict can be downgraded; if not, the contradiction remains visible and prevents the disputed proposition from receiving the highest confidence state.

Within the proposed scheme, three states are therefore retained separately. Supported means the available evidence is compatible with a structural proposition and meaningfully discriminates it from relevant alternatives. Incomplete means the proposition has not been sufficiently constrained but has not encountered material counterevidence. Contradicted means at least one credible observation or orthogonal analysis is materially incompatible with the proposition and cannot be dismissed without a specific methodological reason. These states are applied claim by claim, not molecule by molecule.

Figure 1 integrates this layered treatment of evidence with the confidence, adjudication, revision, reporting, and uncertainty logic of the proposed framework.

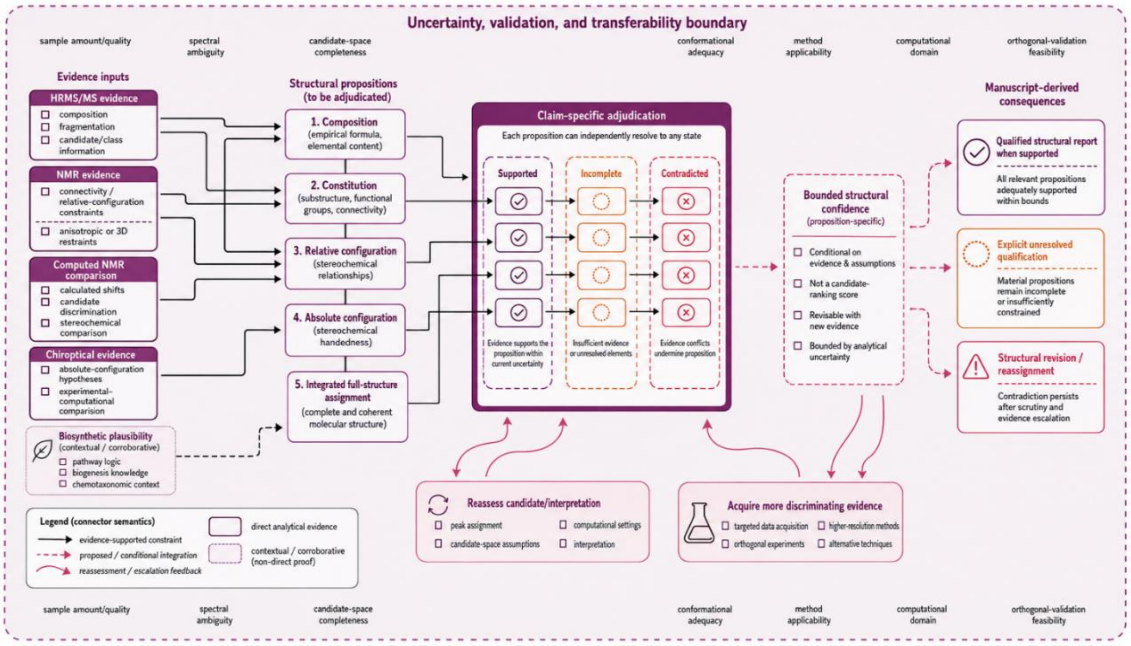

Figure 1. Proposed evidence architecture for converting multimodal spectra into bounded structural confidence

Confidence levels in natural-product assignment

Confidence should be assigned to the structural proposition actually tested rather than inherited from the apparent completeness of the final drawing. For absolute configuration, agreement between anisotropic NMR and chiroptical spectroscopy can provide mutually informative stereochemical constraints, as demonstrated for 12*S*-deoxynortryptoquivaline [25]. Likewise, an integrated ECD and NMR/DP4+ pipeline resolved difficult diphenazine assignments by combining observables that interrogate different aspects of the candidate structures [26]. These examples support multimodal convergence, but they do not validate a universal numerical scale.

The scheme proposed here therefore uses four qualitative confidence grades. Grade A, constrained, applies when the claimed structural element is directly and discriminatingly supported by suitable evidence and no material contradiction remains. Grade B, corroborated, applies when the proposition is supported by a primary constraint and strengthened by an additional, meaningfully distinct line of evidence. Grade C, provisional, applies when a preferred assignment fits the available data but important alternatives are insufficiently excluded. Grade D, unresolved or contested, applies when the available evidence cannot discriminate relevant alternatives or when credible contradiction persists. These grades describe evidential status, not probabilities of truth.

A grade is also bounded by the weakest indispensable proposition. If biological interpretation depends on absolute configuration, a well-established constitution cannot compensate for unresolved enantiomeric assignment. Conversely, uncertainty at a stereocenter that is irrelevant to a narrower constitutional claim should not downgrade that narrower claim. This asymmetry is deliberate: the grade follows the scope of the statement being made. It prevents both confidence inflation and unnecessary skepticism.

Two safeguards are essential. First, computational probabilities remain conditional on the proposed candidate space and calculation protocol [13]. A high ranking cannot upgrade an omitted alternative. Second, conformational inadequacy can affect calculated spectral observables across several methods and thereby create apparently convergent but correlated error [20]. Orthogonality must consequently be judged by the independence of the structural constraint, not by the number of software packages or output scores.

The highest grade for an integrated structure should not erase weaker grades for individual components. A constitution may be constrained while one remote stereocenter remains provisional; conversely, a stereochemical comparison is irrelevant if constitution is unsettled. Material contradiction also acts as a ceiling. Revision studies show that disagreement capable of overturning an assignment should prompt reconsideration [21], while genuinely more discriminating evidence can resolve such conflict when obtainable [24]. The proposed scheme therefore preserves a proposition-level confidence record alongside any final molecular drawing.

A workflow for combining orthogonal evidence

The proposed workflow begins by defining the structural question before selecting the adjudicating method. Accurate formula, constitution, relative configuration, and absolute configuration are treated as separate decision targets. Candidate generation then proceeds from direct observations, with prior chemical knowledge allowed to restrict search only when its role remains explicit. Conditional molecular-generation approaches demonstrate that prior knowledge can usefully constrain NMR-based candidate generation [27], but a prior can also exclude the correct structure if treated as unquestionable.

The next stage is candidate discrimination. Automated interpretation of routine one-dimensional NMR can narrow large structural search spaces [28], while integrated machine-learning and quantum-mechanical scoring can accelerate comparison among stereochemical alternatives [29]. These tools are most defensible when they make candidate testing more efficient rather than converting a ranking into confirmation. For every decisive score, the workflow asks what measurement generated the discriminating information, whether the true structure could lie outside the candidate set, and whether another favorable analysis inherits the same computational assumptions.

Evidence is then combined proposition by proposition. A new result raises confidence when it constrains the same disputed structural feature through a meaningfully different observable. Agreement derived from the same calculated conformer ensemble is recorded as dependent corroboration, whereas direct experimental constraints sensitive to a different structural property carry greater orthogonal value. Contradiction initiates a return path: recheck assignments, purity, candidate enumeration, conformer treatment, and method applicability; if the conflict survives, obtain a more discriminating measurement where feasible or retain the proposition as contested.

Operationally, orthogonality is tested with three questions: does the new result depend on a different measured observable; does it discriminate the alternatives responsible for the present uncertainty; and would the result remain informative if the earlier computational ranking were removed? A “no” answer does not make the result useless, but it places it in dependent corroboration rather than independent constraint. This distinction is particularly important when the same conformer ensemble feeds calculated NMR and chiroptical observables.

Biosynthetic reasoning enters after this analytical adjudication rather than functioning as an early veto. Revision of parvistemoamide altered the associated biosynthetic proposal, demonstrating that structural assumptions can propagate into pathway interpretation [30]. Biosynthesis can therefore strengthen coherence or identify a reason for additional scrutiny, but it should not rescue a candidate contradicted by credible structure-sensitive evidence.

Table 2 converts these principles into comparative decision states, and **Figure 2** shows how the workflow preserves feedback rather than forcing every case toward a definitive endpoint.

Table 2. Decision states for integrating orthogonal evidence in the proposed structure-elucidation workflow

Decision state	Evidential pattern	Appropriate action	Reporting consequence
Convergent	Distinct constraints support the same proposition	Retain candidate; assess confidence grade	State supported level explicitly
Dependent agreement	Favorable results share inputs or assumptions	Count as corroboration, not independence	Disclose dependence
Incomplete	Relevant alternatives remain insufficiently tested	Acquire discriminating evidence if feasible	Preserve provisional status
Contradicted	Credible evidence conflicts with the proposition	Reassess candidate, data, and method	Do not average away conflict
Unresolved after escalation	Conflict or ambiguity persists	Stop forced convergence	Report competing assignments or bounded uncertainty

Note: Decision states are proposed qualitative categories; they are not numerical probability thresholds.

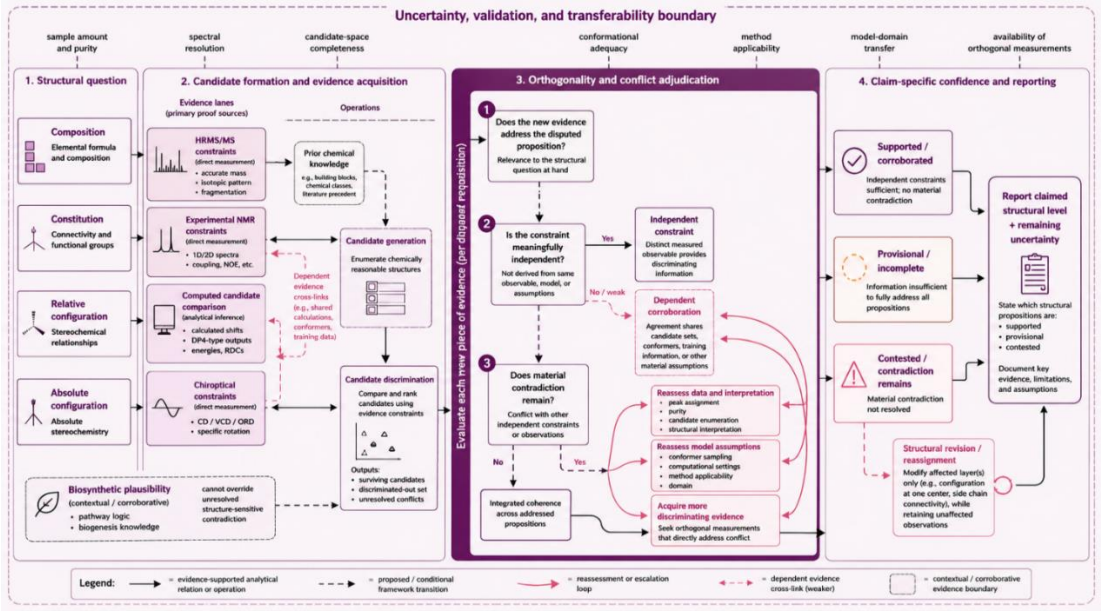

Figure 2. Proposed recursive workflow for combining orthogonal evidence without converting agreement into automatic proof

Implications for structural revision and reporting

A confidence-graded assignment changes the meaning of structural revision. Revision need not imply that every earlier observation was wrong; often, a subset of observations remains valid while the inference connecting them to a particular constitution or stereochemical arrangement changes. Separating proposition-level support therefore permits a revised structure to inherit genuinely stable evidence without inheriting unsupported certainty.

Biosynthetic findings can be valuable in this process when their evidential role is explicit. Elucidation of nannosterols A and B combined structural characterization with biosynthetic investigation, illustrating how biological pathway evidence can make a chemical interpretation more coherent [31]. Such coherence should be reported as corroboration, however, because pathway compatibility does not independently determine every structural or stereochemical element. The direction of inference matters: a well-supported structure can inform biosynthetic interpretation more securely than an attractive biosynthetic narrative can establish an underconstrained structure.

Recent revision of tinotufolins using empirical rules and DFT calculations likewise demonstrates that published assignments remain open to reassessment when systematic spectral inconsistencies emerge [32]. The proposed reporting standard follows from this revisability. Authors should state which structural level is being claimed, identify the observations that discriminate it, and retain visible qualification where alternatives are not adequately excluded. Computational ranking should be described as candidate discrimination, not “confirmation,” unless an independent validating observation justifies the stronger term.

Reporting should also distinguish evidence provenance. Experimental NMR, calculated shifts derived from the same NMR candidate set, and a DP4-type probability are related rather than three independent confirmations. Chiroptical calculations may share conformational assumptions with other computed observables. Biosynthetic plausibility is a contextual test. A compact confidence statement can therefore communicate more than a generic phrase such as “the structure was elucidated by spectroscopic methods”: for each material proposition, it can state whether support is constrained, corroborated, provisional, or contested and what limitation prevents stronger assignment.

For reporting, the scheme favors short claim-linked statements over a single global adjective. An assignment might be described as having constrained constitution, corroborated relative configuration, and provisional absolute configuration, followed by the unresolved reason. Such wording preserves the useful parts of an elucidation while preventing readers and databases from interpreting the entire structure as equally established. It also makes later revision more local: the corrected proposition can be identified without discarding evidence that remains valid.

This approach also improves future correction. When spectra, raw assignments, candidate alternatives, computational settings, and confidence qualifications are preserved, a later investigator can identify exactly which

inference needs revision. Structural transparency then becomes a form of error containment: uncertainty is attached to the proposition that carries it rather than propagated silently into databases, synthesis targets, biosynthetic schemes, or downstream biological interpretation.

Unresolved analytical challenges

The proposed scheme cannot remove information that the experiment never captured. Scarce material, mixtures, overlapping resonances, weak long-range correlations, and inaccessible derivatives may leave constitution or stereochemistry intrinsically underconstrained. Flexible natural products add another problem because calculated NMR or chiroptical observables depend on the sampled conformational ensemble [20]. Agreement generated from an incomplete ensemble may appear stronger than the underlying model warrants.

Automation creates a related transfer problem. Conditional structure generation and one-dimensional NMR machine learning can substantially organize candidate search [27]. Their outputs nevertheless depend on represented chemistry, training distributions, molecular complexity, and the candidate spaces accessible to the model [28]. External evaluation on chemically unfamiliar natural products is therefore needed before model ranking can be treated as generally transferable evidence. The same caution applies when several computational methods share training data, conformers, predicted quantities, or databases.

There is also no validated mapping from the qualitative grades proposed here to a numerical probability of correctness. Establishing one would require securely resolved and deliberately challenging reference structures, blinded application of the workflow, comparison across analysts, and explicit testing of cases with missing, conflicting, or misleading modalities. Prospective studies should examine whether proposition-specific grading improves calibration, identifies revisions earlier, or merely formalizes judgments that experienced spectroscopists already make.

Validation should also include deliberate “hard negatives”: candidate sets containing near-isomeric constitutions, remote stereochemical alternatives, and examples in which attractive biosynthetic reasoning points toward the wrong candidate. Otherwise, a confidence scheme may appear successful simply because easy cases dominate. The most informative benchmark would test whether analysts recognize when the evidence is insufficient, not only whether they choose the correct candidate when it is already represented.

Finally, analytical escalation has practical limits. Additional NMR experiments may require unavailable material; crystallography may be impossible; chiroptical signals may be weak or conformationally complex; and synthesis may be disproportionate to the structural question. The framework therefore permits a scientifically legitimate unresolved endpoint. Its purpose is not to guarantee a full assignment but to make the boundary between established structure and preferred interpretation explicit.

Conclusion

A molecular drawing compresses several different evidential claims into one visual object. The central proposal of this article is to unpack that compression. Composition, constitution, relative configuration, absolute configuration, and the integrated assignment should receive confidence only to the extent that each has been specifically constrained, and evidence for one level should not be transferred automatically to another.

The proposed workflow consequently separates observation from candidate generation, candidate generation from discrimination, and discrimination from validation. It also distinguishes incomplete evidence from active contradiction. Missing information leaves a proposition provisional; credible incompatibility requires reassessment, methodological explanation, additional discriminating evidence where feasible, or an explicitly contested outcome. Orthogonality is defined by the independence of the structural constraint rather than by the number of instruments, calculations, algorithms, or favorable scores. Biosynthetic plausibility contributes coherence but does not override structure-sensitive counterevidence.

This framework is intended as a reporting and reasoning discipline, not as a validated probability model. Its qualitative grades have not been prospectively calibrated, and their usefulness will depend on chemical class, sample quality, conformational behavior, candidate-space completeness, computational applicability, and access to orthogonal measurements. Those limitations are not reasons to collapse confidence back into a binary assignment. They are reasons to report structural knowledge at the resolution the evidence actually supports.

The practical endpoint is therefore deliberately modest: a natural-product structure should be presented not merely as the best available drawing, but as a set of explicit structural propositions whose support, dependence,

uncertainty, and remaining contradictions can be inspected and revised. Such reporting makes analytical confidence visible without pretending that uncertainty has been eliminated.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Shen SM, Appendino G, Guo YW. Pitfalls in the structural elucidation of small molecules. A critical analysis of a decade of structural misassignments of marine natural products. Nat Prod Rep. 2022;39(9):1803-32. doi:10.1039/d2np00023g
2. Chhetri BK, Lavoie S, Sweeney-Jones AM, Kubanek J. Recent trends in the structural revision of natural products. Nat Prod Rep. 2018;35(6):514-31. doi:10.1039/c8np00011e
3. DiBello M, Healy AR, Nikolayevskiy H, Xu Z, Herzon SB. Structure elucidation of secondary metabolites: Current frontiers and lingering pitfalls. Acc Chem Res. 2023;56(12):1656-68. doi:10.1021/acs.accounts.3c00183
4. Hu G, Qiu M. Machine learning-assisted structure annotation of natural products based on MS and NMR data. Nat Prod Rep. 2023;40(11):1735-53. doi:10.1039/d3np00025g
5. Brown PD, Lawrence AL. The importance of asking “how and why?” in natural product structure elucidation. Nat Prod Rep. 2017;34(10):1193-202. doi:10.1039/c7np00025a
6. Reher R, Kim HW, Zhang C, Mao HH, Wang M, Nothias LF, et al. A convolutional neural network-based approach for the rapid annotation of molecularly diverse natural products. J Am Chem Soc. 2020;142(9):4114-20. doi:10.1021/jacs.9b13786
7. Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, et al. SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods. 2019;16(4):299-302. doi:10.1038/s41592-019-0344-8
8. Dührkop K, Nothias LF, Fleischauer M, Reher R, Ludwig M, Hoffmann MA, et al. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nat Biotechnol. 2021;39(4):462-71. doi:10.1038/s41587-020-0740-8
9. Stravs MA, Dührkop K, Böcker S, Zamboni N. MSNovelist: De novo structure generation from mass spectra. Nat Methods. 2022;19(7):865-70. doi:10.1038/s41592-022-01486-3
10. Bach E, Schymanski EL, Rousu J. Joint structural annotation of small molecules using liquid chromatography retention order and tandem mass spectrometry data. Nat Mach Intell. 2022;4(12):1224-37. doi:10.1038/s42256-022-00577-2
11. Navarro-Vázquez A, Gil RR, Blinov K. Computer-assisted 3D structure elucidation (CASE-3D) of natural products combining isotropic and anisotropic NMR parameters. J Nat Prod. 2018;81(1):203-10. doi:10.1021/acs.jnatprod.7b00926
12. Marcarino MO, Zanardi MM, Cicetti S, Sarotti AM. NMR calculations with quantum methods: Development of new tools for structural elucidation and beyond. Acc Chem Res. 2020;53(9):1922-32. doi:10.1021/acs.accounts.0c00365
13. Marcarino MO, Cicetti S, Zanardi MM, Sarotti AM. A critical review on the use of DP4+ in the structural elucidation of natural products: The good, the bad and the ugly. A practical guide. Nat Prod Rep. 2022;39(1):58-76. doi:10.1039/d1np00030f
14. Immel S, Köck M, Reggelin M. NMR-based configurational assignments of natural products: How floating chirality distance geometry calculations simplify gambling with 2^N−1 diastereomers. J Nat Prod. 2022;85(7):1837-49. doi:10.1021/acs.jnatprod.2c00427

15. Franco BA, Luciano ER, Sarotti AM, Zanardi MM. DP4+App: Finding the best balance between computational cost and predictive capacity in the structure elucidation process by DP4+. Factors analysis and automation. J Nat Prod. 2023;86(10):2360-7. doi:10.1021/acs.jnatprod.3c00566
16. Priessner M, Lewis RJ, Johansson MJ, Goodman JM, Janet JP, Tomberg A. HSQC spectra simulation and matching for molecular identification. J Chem Inf Model. 2024;64(8):3180-91. doi:10.1021/acs.jcim.3c01735
17. Mazzeo G, Cimmino A, Masi M, Longhi G, Maddau L, Memo M, et al. Importance and difficulties in the use of chiroptical methods to assign the absolute configuration of natural products: The case of phytotoxic pyrones and furanones produced by Diplodia corticola. J Nat Prod. 2017;80(9):2406-15. doi:10.1021/acs.jnatprod.7b00119
18. Górecki M, Frelek J. A critical appraisal of dimolybdenum tetraacetate application in stereochemical studies of vic-diols by circular dichroism. J Nat Prod. 2020;83(4):955-64. doi:10.1021/acs.jnatprod.9b00800
19. Batista ANL, Valverde AL, Nafie LA, Batista JM Jr. Stereochemistry of natural products from vibrational circular dichroism. Chem Commun (Camb). 2024;60(76):10439-50. doi:10.1039/d4cc02481h
20. Zhou Y, Limbu I, Garson MJ, Krenske EH. Conformational sampling in computational studies of natural products: Why is it important? J Nat Prod. 2024;87(10):2543-49. doi:10.1021/acs.jnatprod.4c00852
21. Sarotti AM. In silico reassignment of (+)-diplopyrone by NMR calculations: Use of a DP4/J-DP4/DP4+/DIP tandem to revise both relative and absolute configuration. J Org Chem. 2020;85(17):11566-70. doi:10.1021/acs.joc.0c01563
22. Ren FC, Wang LX, Lv YF, Hu JM, Zhou J. Structure revision of four classes of prenylated aromatic natural products based on a rule for diagnostic 13C NMR chemical shifts. J Org Chem. 2021;86(16):10982-90. doi:10.1021/acs.joc.0c02409
23. Guerrero De León E, Sánchez-Martínez H, Morán-Pinzón JA, Del Olmo Fernández E, López-Pérez JL. Computational structural revision of elaeophorbate and other triterpenoids with the help of NAPROC-13. A new strategy for structural revision of natural products. J Nat Prod. 2023;86(4):897-908. doi:10.1021/acs.jnatprod.2c01135
24. Rangel-Grimaldo M, Earp CE, Raja HA, Wood JS, Mardiana L, Ho KL, et al. Wheldone revisited: Structure revision via DFT-GIAO chemical shift calculations, 1,1-HD-ADEQUATE NMR spectroscopy, and X-ray crystallography studies. J Nat Prod. 2024;87(8):2095-100. doi:10.1021/acs.jnatprod.4c00649
25. Doro-Goldsmith E, Song Q, Li XL, Li XM, Hu XY, Li HL, et al. Absolute configuration of 12S-deoxynortryptoquivaline from ascidian-derived fungus Aspergillus clavatus determined by anisotropic NMR and chiroptical spectroscopy. J Nat Prod. 2024;87(2):381-7. doi:10.1021/acs.jnatprod.3c01157
26. Zhuang Y, Yang F, Menon A, Song JM, Espinoza RV, Schultz PJ, et al. An ECD and NMR/DP4+ computational pipeline for structure revision and elucidation of diphenazine-based natural products. J Nat Prod. 2023;86(7):1801-14. doi:10.1021/acs.jnatprod.3c00306
27. Yao L, Yang M, Song J, Yang Z, Sun H, Shi H, et al. Conditional molecular generation net enables automated structure elucidation based on 13C NMR spectra and prior knowledge. Anal Chem. 2023;95(12):5393-401. doi:10.1021/acs.analchem.2c05817
28. Hu F, Chen MS, Rotskoff GM, Kanan MW, Markland TE. Accurate and efficient structure elucidation from routine one-dimensional NMR spectra using multitask machine learning. ACS Cent Sci. 2024;10(11):2162-70. doi:10.1021/acscentsci.4c01132
29. Tsai YH, Amichetti M, Zanardi MM, Grimson R, Hernandez Daranas A, Sarotti AM. ML-J-DP4: An integrated quantum mechanics-machine learning approach for ultrafast NMR structural elucidation. Org Lett. 2022;24(41):7487-91. doi:10.1021/acs.orglett.2c01251
30. Olivier WJ, Smith JA, Bissember AC. Structural revision of parvistemoamide: Informing biosynthetic proposals of Stemona alkaloids. Org Lett. 2022;24(31):5772-6. doi:10.1021/acs.orglett.2c02254
31. Akone SH, Hug JJ, Kaur A, Garcia R, Müller R. Structure elucidation and biosynthesis of nannosterols A and B, myxobacterial sterols from Nannocystis sp. MNa10993. J Nat Prod. 2023;86(4):915-23. doi:10.1021/acs.jnatprod.2c01143
32. Nishidono Y, Tanaka K. Structural revision of tinotufolins from Tinospora crispa leaves guided by empirical rules and DFT calculations. J Nat Prod. 2024;87(4):774-82. doi:10.1021/acs.jnatprod.3c00902