SMART: Spectral Multimodal Adaptation via Routed Test-Time Tuning for Zero-Shot Molecular Elucidation

Anonymous authors
Independent working draft

Abstract

SMART adapts pretrained molecular generators to experimental spectra without target structural labels. Under heterogeneous acquisition shifts, supplying all spectra can reduce accuracy below that of a single informative view. We address this negative transfer by separating the supervision-source view from the Full inference view. A source-trained structural-utility router selects a frozen teacher; encoder InfoNCE preserves molecular correspondence, while shared-prefix decoder Jensen–Shannon alignment transfers conditional structural preferences. Persistent, formula-compatible predictions then support cross-view refinement, with a dynamic controller selecting checkpoints according to prediction drift and feasibility. End-to-end hybrid adaptation reaches 55.66% Top-1 on 3,669 SDBS molecules, 4.96 percentage points above the strongest zero-shot single-spectrum reference. Objective ablations and Chemotion evaluation examine complementary learning signals and stronger acquisition shift. These results show how selective supervision can recover the value of complete spectral observations. Code is available at https://anonymous.4open.science/r/SMART-57808/.

1 Introduction

Molecular elucidation integrates complementary constraints: NMR describes atomic environments, infrared spectra identify vibrational signatures, and mass spectra supply composition and fragmentation evidence. Simulated multispectral datasets and pretrained generators make joint inference scalable (Alberts et al., 2024; 2025; Su et al., 2026). Yet a generator trained on simulation can recover fewer experimental structures from all spectra than from a single informative modality. Figure 1 quantifies this negative transfer and its recovery. The difficulty lies in how evidence is interpreted. Acquisition shifts affect modalities unevenly, altering learned fusion rules and the allocation of trust (Yang et al., 2024; Chen et al., 2025; Mismetti et al., 2025). Carbon environments can remain informative while unfamiliar mass intensities distort the joint prediction. Additional observations still contain chemical constraints, but exploiting them requires adapting the inference rule.

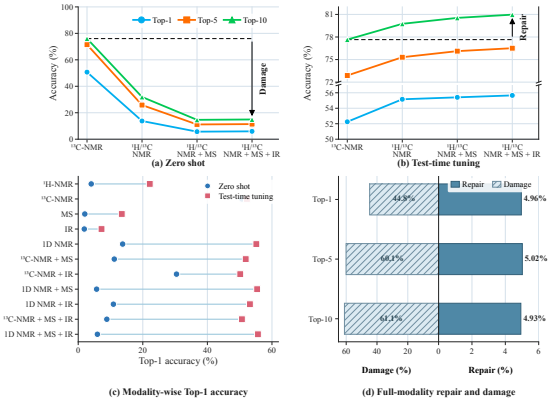

Figure 1: SDBS exact-match retrieval before and after adaptation. Repair/damage bars give percentage-point differences from zero-shot carbon NMR; figure-version details appear in Appendix G.

Unlabeled test-time learning offers a way to revise this rule (Sun et al., 2020; Wang et al., 2021). However, self-training a shifted Full predictor may reinforce its initial errors. SMART instead separates the *supervision-source view* from the *inference view*: a subset selected for structural utility supplies a frozen reference, while the Full student retains complementary observations that can improve it. This division connects reliable initial interpretation with broader evidence integration.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

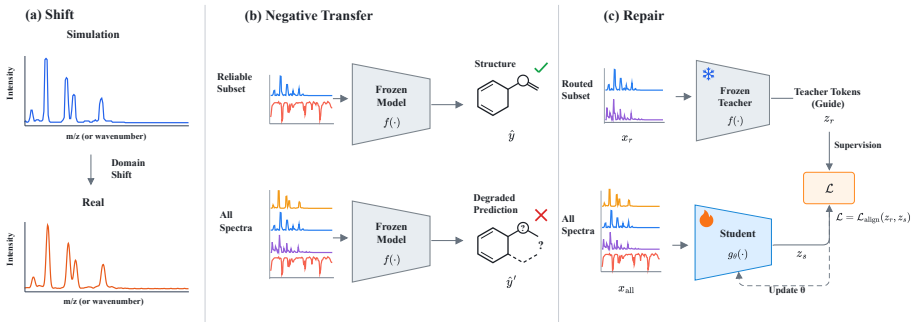

Figure 2: Heterogeneous acquisition shifts can turn complementary spectra into conflicting model evidence. SMART uses a routed teacher to guide adaptation with all available spectra.

A source-trained router selects the teacher view. Encoder InfoNCE then preserves molecule-specific correspondence across views, and decoder Jensen–Shannon (JS) alignment transfers conditional preferences among sequence continuations. Persistent, formula-compatible structures subsequently support refinement across spectral subsets. A dynamic checkpoint controller follows prediction drift during Stage 1; refinement begins after its prescribed budget ends.

Our contributions are:

- **Supervision decoupling for spectral adaptation:** a decision-theoretic account of negative transfer and a source-trained policy for selecting target supervision without target structural labels.
- **Hybrid molecular alignment:** encoder InfoNCE preserves cross-view distinctions and shared-prefix decoder JS connects them to structural generation.
- **Temporal selection and cross-view refinement:** a drift- and feasibility-driven controller identifies persistent, composition-compatible targets for learning from complementary subsets.

2 Related work

2.1 Structure learning from complementary spectra

Multispectral simulation supports generative structure elucidation (Alberts et al., 2024; 2025), and SpectraLLM extends language-based prediction across spectral inputs (Su et al., 2026). SMART adapts the pretrained mapping when acquisition conditions change. Mass-spectral adaptation (Mismetti et al., 2025) addresses a related challenge; our setting additionally requires choosing supervision under heterogeneous NMR, IR, and MS shifts.

2.2 Test-time learning and multimodal reliability

Test-time training derives supervision from deployment inputs (Sun et al., 2020); Tent, EATA, and SAR study entropy-based updates, efficiency, forgetting, and stability (Wang et al., 2021; Niu et al., 2022; 2023). Multimodal methods address reliability bias and selective adaptation (Yang et al., 2024; Chen et al., 2025). SMART applies view selection to autoregressive molecular generation, linking a stable teacher policy to representation and sequence learning. Adaptation is transductive over an unlabeled cohort: zero-shot describes initial deployment, followed by test-time parameter updates.

3 Problem setting and motivation

3.1 From spectral evidence to structural decisions

A molecule is observed through its formula F and spectra $x = (F, X_H, X_C, X_M, X_I)$, with $\mathcal{M} = \{H, C, M, I\}$. A view a retains a nonempty spectral subset: $x^a = (F, \{X_m : m \in a\})$. Formula accompanies every view; four available spectra give fifteen actions, including Full. The

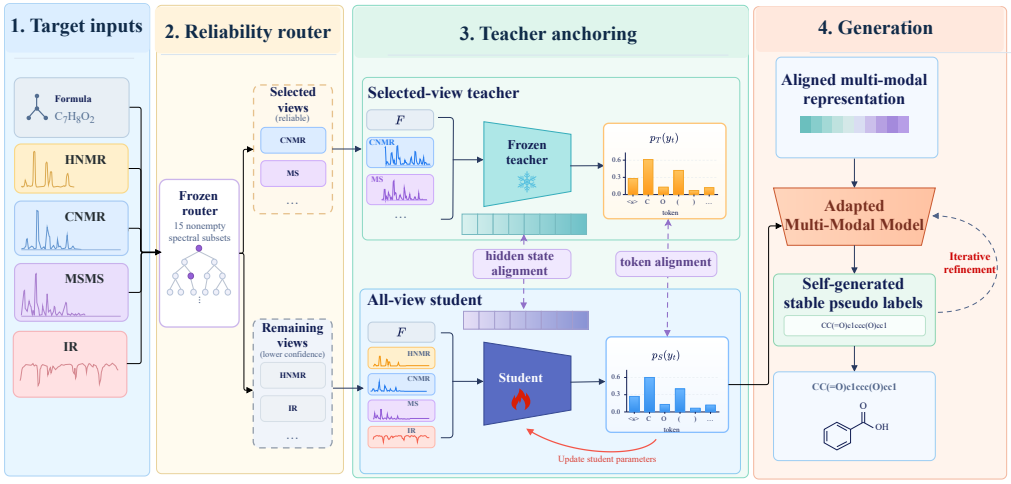

Figure 3: SMART workflow: source-trained routing, frozen-teacher alignment, and selective refinement. The student retains all available spectra and the observed formula; molecular sketches illustrate the stages.

source generator writes a SMILES sequence conditioned on this evidence and the preceding symbols:

$$p_{\theta_0}(y \mid x^a) = \prod_{t=1}^{T} p_{\theta_0}(y_t \mid y_{<t}, x^a). \tag{1}$$

Here θ_0 denotes the pretrained parameters. Canonicalization maps sequences to molecular graphs, and Top-k measures reference-graph recovery under a fixed decoding budget. Adaptation uses only target observations $\mathcal{U} = \{x_i\}$, without their structural labels. Its goal is to improve Full-view decisions while retaining useful partial-view behavior.

3.2 Available information and learned prediction risk

An ideal decision maker can ignore additional observations. Thus every subset rule remains available with Full input. For molecular graph Y, a common structural loss ℓ, and target distribution P_t, the optimal expected risk satisfies (Blackwell, 1953)

$$\begin{aligned} \mathcal{R}_t^*(a) &= \inf_f \mathbb{E}_{P_t}[\ell(Y, f(X^a))], \\ \mathcal{R}_t^*(\text{Full}) &\le \mathcal{R}_t^*(a), \qquad a \subseteq \mathcal{M}. \end{aligned} \tag{2}$$

The infimum ranges over measurable rules with the same output space. A Full rule can reproduce any subset rule by projection, establishing the inequality. Chemically, extra spectra offer opportunities to resolve ambiguities without removing existing constraints. Appendix C gives the proof and a complementary information-theoretic interpretation.

A fixed source generator instead applies a particular rule learned under P_s. Writing its decoded decision from view a as $g_{\theta_0}^a$, negative transfer occurs when

$$\Delta_t(a) = \mathcal{R}_t(g_{\theta_0}^{\text{Full}}) - \mathcal{R}_t(g_{\theta_0}^a) > 0, \qquad \mathcal{R}_t(g) = \mathbb{E}_{P_t}[\ell(Y, g(X))]. \tag{3}$$

Heterogeneous acquisition shifts can change which patterns the source model treats as decisive. A positive gap therefore reflects unreliable use of richer evidence, consistent with the Bayes-risk inequality. For example, carbon environments may remain recognizable while unfamiliar mass intensities change the joint prediction. Adaptation must correct this learned interpretation to recover the value of the additional spectra.

3.3 Choosing supervision under heterogeneous shift

A useful teacher view balances structural information with compatibility between its measurements and the source generator. Carbon NMR may supply a reliable initial constraint for one molecule; fragmentation may be essential for another. Reliability consequently depends on the view, molecule, and

generator together. SMART learns this structural utility on labeled source examples. The selected teacher guides target updates, while the Full student continues to observe complementary evidence that can improve the initial interpretation. This separation makes a useful partial decision a starting point for stronger joint inference.

4 SMART: LEARNING TO INTEGRATE SHIFTED SPECTRAL EVIDENCE

SMART separates the view supplying supervision from the observations used for inference (Figure 3). A source-trained router selects a frozen teacher; hybrid alignment transfers its molecular reference to a Full student; temporal consensus then supports refinement across spectral subsets.

4.1 STRUCTURAL-UTILITY ROUTING

4.1.1 LEARNING USEFUL VIEWS THROUGH SOFT STRUCTURAL SUPERVISION

Reliability is defined by structural usefulness. A clean spectrum can leave isomers unresolved, while a noisy measurement may contain a decisive fragment. We compare views of the same labeled source molecule after identity-preserving perturbations, separating acquisition effects from molecular difficulty. Weak, medium, and strong modality-specific profiles are calibrated from same-molecule residuals (Appendix D), connecting acquisition variation to the reliability imbalance studied in multimodal adaptation (Yang et al., 2024; Chen et al., 2025).

For each perturbed input $\tilde{x}_i$ and legal view a, the frozen generator produces five candidates. Exact recovery, candidate coverage, sequence likelihood, and chemical validity define its utility. Likelihood supplies continuous feedback when beam candidates miss the reference; the invalid-output penalty accounts for unusable structures. Appendix E specifies the reductions. The criterion measures evidence that this generator can convert into a structural decision. A view is rewarded for recovering the reference rather than merely producing a confident sequence. This task-facing comparison links the source perturbations to the eventual chemical objective, while allowing different molecules to rely on different combinations of measured constraints.

$$R_i(a) = \mathrm{Top1}_i(a) + 0.25\,\mathrm{Top5}_i(a) - 0.05\,\mathrm{NLL}_i(a) - 0.10\,\mathrm{Invalid}_i(a). \tag{4}$$

Evidence also depends on its companions: carbon environments and fragments may resolve each other's ambiguities, whereas unfamiliar background can disrupt a useful view. Individual utilities, pair interactions, and subset-size preferences express this dependence:

$$s_\rho(a, z) = \sum_m a_m u_m(z_m, z_g) + \sum_{m<n} a_m a_n v_{mn}(z_{mn}, z_g) + b_{|a|}(z_g). \tag{5}$$

Observed-input features z summarize modality-to-formula relations, pairwise encoder agreement, and availability. Unary terms describe individual usefulness, pair terms capture reinforcement or interference, and the size term models the preference for retaining observations. Unavailable actions are masked before normalization.

Several views can offer comparable guidance. Soft reward targets preserve these alternatives instead of declaring only one action correct:

$$\begin{aligned} q_i(a) &= \mathrm{softmax}_a(R_i(a)/\tau_o), \qquad p_\rho(a \mid z_i) = \mathrm{softmax}_a(s_\rho(a, z_i)/\tau_r), \\ \mathcal{L}_R &= \frac{1}{N}\sum_i D_{\mathrm{KL}}(q_i \| p_\rho). \end{aligned} \tag{6}$$

We use $\tau_o = .20$ and $\tau_r = 1$. Source validation utility selects the router, which remains frozen throughout target adaptation. Structural labels thus train the supervision policy on the source domain; target updates use only observed spectra and formulas.

4.2 Stage 1: hybrid alignment of molecular identity and generation

4.2.1 A stable teacher and discriminative encoder alignment

For an unlabeled target molecule, the router chooses a teacher view and the frozen source generator returns a beam-selected sequence:

$$a_i^* = \arg\max_{a \in \mathcal{A}_i} p_\rho(a \mid z_i), \qquad y_i^R = \mathrm{BeamTop1}_{10}\left(p_{\theta_0}(\cdot \mid x_i^{a_i^*})\right). \tag{7}$$

The student starts from the same weights and reads x_i^{Full}. The selected view may contain any legal subset, including Full. Teacher parameters $\theta_T = \theta_0$, representations, and logits are constants during optimization over student parameters ϕ: $\nabla_\phi k_i = \nabla_\phi g_{i,t}^T = 0$. This stable reference links target updates to a structural interpretation supported by the source generator.

Transferring this reference requires both recognition and generation (Figure 4). Consider a carbon-routed teacher paired with Full input containing a background-rich mass spectrum. Their peak patterns differ, but molecular identity is shared. Encoder alignment preserves the teacher's molecular distinctions while the student processes additional observations. Competing teacher representations make the correspondence discriminative: matching the current molecule must remain preferable to matching other records. The decoder then connects that representation to conditional structural decisions (van den Oord et al., 2018; Hinton et al., 2015; Agarwal et al., 2024).

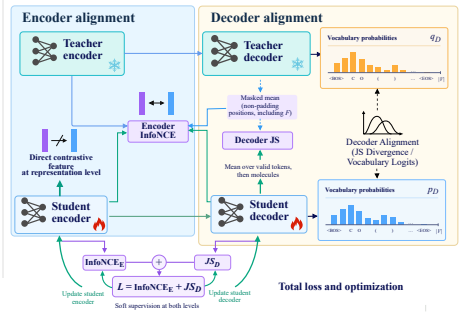

Figure 4: Hybrid alignment couples encoder InfoNCE with decoder JS under a frozen teacher. Their sum updates the full-modality student.

Pooled states u_i^S, u_i^T yield normalized queries $h_i = u_i^S/\|u_i^S\|_2$ and frozen keys $k_i = u_i^T/\|u_i^T\|_2$. For minibatch $\mathcal{C}$, eligible keys J_i include the positive i:

$$\ell_i^{\mathrm{NCE}} = -\log \frac{\exp(h_i^\top k_i/\tau_C)}{\sum_{j \in J_i} \exp(h_i^\top k_j/\tau_C)}, \qquad \mathcal{L}_E^{\mathrm{NCE}} = \mathbb{E}_{\mathcal{C}} \frac{1}{|\mathcal{C}|} \sum_{i \in \mathcal{C}} \ell_i^{\mathrm{NCE}}. \tag{8}$$

Minimization increases the positive pair's relative identification margin (van den Oord et al., 2018; Chen et al., 2020). Normalization focuses comparison on direction; temperature controls competition among nearby alternatives. Extra acquisition detail should therefore preserve molecule-specific evidence rather than displace the representation toward an unrelated record.

Known duplicate observations are excluded from each other's negatives using available metadata, while chemically distinct isomers remain competitors. This distinction matters when a shared elemental composition permits different connectivity. Within-formula retrieval can diagnose whether the representation retains information beyond composition. The intended invariance concerns molecular discrimination across views, rather than indiscriminate removal of spectral differences. For a carbon-guided reference, retaining atomic-environment distinctions can prevent unrelated background peaks from dominating the joint representation. Complementary spectra remain available to the student and can supply information absent from that reference. This makes the teacher a point of correspondence rather than a restriction on the information the student may encode. Structural retrieval evaluates whether that correspondence supports the correct molecule.

4.2.2 Decoder JS and joint structural learning

A partial structure can admit several plausible atom choices, branches, or ring closures. Soft distillation preserves their relative support. Both decoders read the same raw teacher prefix $y_{i,<t}^R$ and predict the full vocabulary:

$$q_{i,t} = \mathrm{softmax}(g_{i,t}^T/\tau_D), \qquad p_{i,t} = \mathrm{softmax}(g_{i,t}^S/\tau_D). \tag{9}$$

The common prefix fixes the structural context, so differences reflect the evidence provided by each view. The teacher sequence supplies context and valid positions; its probabilities supply the target. This retains alternatives beyond the winning token and keeps uncertainty in the supervision signal.

The common mixture gives a symmetric measure of conditional disagreement:

$$\mathrm{JS}(q,p) = \tfrac{1}{2} D_{\mathrm{KL}}(q\|m) + \tfrac{1}{2} D_{\mathrm{KL}}(p\|m), \qquad m = \tfrac{1}{2}(q+p). \tag{10}$$

With natural logarithms, JS is nonnegative and bounded by $\log 2$ (Lin, 1991). Teacher-favored continuations and the remaining vocabulary jointly contribute graded preferences. These probabilities describe tokens in a structural representation; they are not direct probabilities over complete molecular graphs. For example, a graded preference between branch continuations can be retained until later spectral evidence supports a more specific connectivity decision.

To give molecules equal weight despite different SMILES lengths, we average over valid positions before averaging across the cohort:

$$\mathcal{L}_D^{\mathrm{JS}} = \frac{1}{|\mathcal{U}|} \sum_i \frac{1}{T_i} \sum_{t=1}^{T_i} \mathrm{JS}(q_{i,t}, p_{i,t}). \tag{11}$$

Longer strings therefore do not dominate solely through token count. The decoder learns a sequence of conditional preferences, connecting the selected evidence to atom and connectivity decisions.

Encoder and decoder alignment address complementary consequences of acquisition shift. The first promotes correspondence between views of the same molecule; the second guides its expression as a structure. Stage 1 uses unit coefficients:

$$\mathcal{L}_1 = \mathcal{L}_E^{\mathrm{NCE}} + \mathcal{L}_D^{\mathrm{JS}}. \tag{12}$$

Both losses update the student encoder, and JS also updates the decoder. We use $\tau_C = .10$ and $\tau_D = 1$; gradients appear in Appendix C. No additional hard-sequence CE is applied in this stage. Preserving carbon-environment distinctions is useful when those distinctions influence subsequent structural choices. Conversely, soft sequence guidance benefits from a Full representation organized around molecular evidence. Joint learning maintains this connection while the student continues to observe spectra that can improve the teacher's initial interpretation.

4.3 Stage 2: dynamic temporal consensus and cross-view refinement

Stage 1 runs for E_1 epochs. A parallel controller observes its trajectory and chooses checkpoints without changing the alignment objective. Consensus is formed only after this budget ends, and Stage 2 starts from the final Stage-1 weights. If no structure is accepted, that final model is retained.

4.3.1 Sampling evolving structural hypotheses

A fixed probe set $\mathcal{B}$, selected by deterministic identity hashing, follows the same molecules during Stage 1. Full greedy decoding at epoch e gives canonical predictions $c_i^{(e)}$ and indicators $v_i^{(e)}$ for RDKit validity and agreement with the observed formula:

$$d_e = \frac{1}{|\mathcal{B}|} \sum_i \mathbf{1}[c_i^{(e)} \neq c_i^{(e-1)}], \qquad g_e = \frac{1}{|\mathcal{B}|} \sum_i v_i^{(e)}. \tag{13}$$

Drift d_e detects changing interpretations, including connectivity switches; feasibility g_e distinguishes chemically admissible predictions from persistent invalid outputs. Rapid change shortens the checkpoint interval, whereas sustained feasible agreement allows it to grow:

$$n_e = \begin{cases} n_{\min}, & g_e < \gamma, \\ \max(n_{\min}, \lfloor n_{e-1}/2 \rfloor), & d_e \geq \delta_{\text{high}}, \\ \min(n_{\max}, n_{e-1}+1), & r_e \geq h, \\ n_{e-1}, & \text{otherwise.} \end{cases} \tag{14}$$

Here r_e counts consecutive feasible probes with drift at most δ_{low}, and h is patience. A checkpoint is selected when elapsed epochs reach the updated n_e; epochs one and E_1 are always retained. Thus $\mathcal{E} = \{s_1 < \cdots < s_K\}$ changes with the observed trajectory. Exact replay and thresholds are specified in Appendix E.

4.3.2 Persistent targets supported by complementary views

A reusable interpretation should survive multiple parameter states and match measured composition. Full predictions enter the accepted set only if they remain the same valid canonical structure and satisfy the formula throughout:

$$\mathcal{P} = \{i : \text{canon}(\hat{y}_i^{(e)}) = s_i,\ v_i^{(e)} = 1 \text{ for every } e \in \mathcal{E}\}. \tag{15}$$

Canonical comparison groups different strings describing the same structure. Formula agreement enforces the observed elemental inventory, while temporal agreement selects interpretations that persist through optimization. The raw sequence from the ordered middle checkpoint becomes $\tilde{y}_i$. Selection combines chemical constraints with accumulated evidence of structural persistence; it does not guarantee that every accepted isomer is correct.

Stage 2 holds these targets fixed while changing the visible spectra. The molecular hypothesis remains shared across observations. Carbon-containing views constrain atomic environments, infrared-containing views emphasize functional groups, and mass-containing views supply fragmentation evidence. Learning the same structure from these combinations encourages compatible interpretations through a shared generator. Starting from the final Stage-1 student, we sample nonempty proper spectral subsets. With four available modalities, retained size is uniform from one to three and the subset is uniform within each size; formula always remains visible:

$$\begin{aligned} \mathcal{L}_2 &= \frac{1}{|\mathcal{P}|} \sum_{i \in \mathcal{P}} \mathbb{E}_{A \sim \pi} \left[-\frac{1}{T_i} \sum_t \log p_\phi(\tilde{y}_{i,t} \mid \tilde{y}_{i,<t}, x_i^A) \right], \\ \pi(A) &= \frac{1}{3\binom{4}{|A|}}, \qquad 1 \le |A| \le 3. \end{aligned} \tag{16}$$

Removing a spectral pathway makes the remaining observations support more of the prediction. Shared parameters and Stage-1 initialization connect these partial-view updates to Full inference. Final evaluation uses all observations over the complete cohort, testing whether selective supervision improves their joint use at deployment. This separates the chemical constraint used to accept a target from the varied evidence used to learn it.

5 Experiments

5.1 Evaluation protocol

We adapt the Multitasking generator (Alberts et al., 2025) on 3,669 SDBS molecules and 1,995 strictly selected Chemotion records (Tremouilhac et al., 2021; Chemotion, 2026). Target spectra and formulas are available during transductive adaptation; structures are used only for offline scoring. Beam-ten decoding reports Top-1/5/10 on each complete cohort. H/C/MS/IR denote proton NMR, carbon NMR, mass spectra, and infrared spectra; every view retains formula. The paired 500-record structural study has a distinct checkpoint lineage. Appendix A preserves supplementary experiments and Appendix G documents result provenance.

5.2 Recovering complementarity and improving the teacher

Table 1 evaluates the end-to-end InfoNCE–JS implementation. Unadapted Full obtains 5.94% Top-1 versus 50.70% for carbon NMR. SMART raises Full to 55.66%, recovering 49.72 points and exceeding the strongest initial subset by 4.96 points. Full Top-5/10 increase from 11.42/14.91% to 76.51/80.97%. Adapted carbon reaches 52.25%, leaving Full a 3.41-point advantage; H+C and H+C+MS reach 55.16% and 55.41%. These outcomes connect reliable initial constraints to gains from complementary evidence. Several adapted views support the recovery, and the best Top-1 endpoint in this table uses the complete observation.

The routed teacher, Stage 1, and Stage 2 achieve 48.87/53.09/55.66% Top-1 and 70.40/73.78/76.51% Top-5. Thus both learning a Full mapping and refining accepted targets improve on teacher selection. Under stronger Chemotion shift, Full Top-1 rises from .100% to 4.511%, and Top-5/10 from .201/.351% to 7.018/8.271%. Absolute accuracy remains low; C+MS instead leads Top-10 at 9.323%, distinguishing first-candidate recovery from shortlist coverage.

Table 1: End-to-end SDBS retrieval on 3,669 molecules: unadapted baseline and Stage 2 of the InfoNCE–JS hybrid pipeline. Values are percentages; differences are percentage points.

	multitasking			SMART			Δ (improvement)		
Modality	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑
H	4.03±0.24	8.72±0.38	11.31±0.45	22.16±0.61	45.13±0.88	55.38±1.02	+18.13	+36.41	+44.07
C	**50.70**±0.72	**71.49**±0.86	**76.04**±0.94	52.25±0.78	72.88±0.97	77.65±1.10	+1.55	+1.39	+1.61
MS	2.04±0.18	4.99±0.27	6.76±0.34	13.49±0.46	33.96±0.71	44.04±0.86	+11.45	+28.97	+37.28
IR	1.88±0.16	4.96±0.25	7.28±0.31	7.22±0.35	18.56±0.56	26.60±0.72	+5.34	+13.60	+19.32
H + C	13.76±0.48	25.81±0.67	31.92±0.79	55.16±0.74	75.31±0.95	79.75±1.09	+41.40	+49.50	+47.83
C + MS	11.17±0.44	21.18±0.61	26.11±0.73	51.87±0.68	72.85±0.89	77.54±1.03	+40.70	+51.67	+51.43
C + IR	30.44±0.56	50.23±0.74	56.72±0.88	50.18±0.72	70.92±0.94	75.74±1.08	+19.74	+20.69	+19.02
H + C + MS	5.70±0.35	11.15±0.52	14.64±0.63	55.41±0.81	76.12±1.03	80.54±1.17	+49.71	+64.97	+65.90
H + C + IR	10.90±0.39	20.90±0.58	27.56±0.69	53.18±0.76	73.81±0.98	78.82±1.12	+42.28	+52.91	+51.26
C + MS + IR	8.86±0.46	17.99±0.64	22.21±0.76	50.70±0.73	72.09±0.95	76.94±1.09	+41.84	+54.10	+54.73
H + C + MS + IR	5.94±0.42	11.42±0.59	14.91±0.71	**55.66**±0.84	**76.51**±1.06	**80.97**±1.20	**+49.72**	**+65.09**	**+66.06**

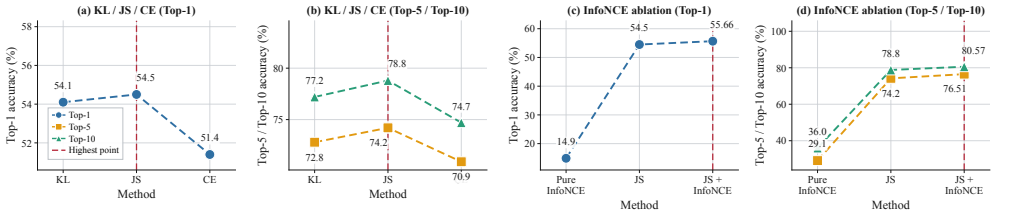

Figure 5: SDBS alignment-objective ablation. Hybrid InfoNCE–JS achieves the highest retrieval in this comparison.

5.3 Alignment and redistribution of spectral evidence

The objective comparison (Figure 5) gives 14.9% Top-1 for encoder InfoNCE alone, 54.5% for decoder JS, 54.1% for KL, and 51.4% for hard CE. Hybrid InfoNCE–JS reaches 55.66/76.51/80.57% Top-1/5/10, improving on JS by 1.16/2.31/1.77 points. Its Top-10 value belongs to the separate objective comparison; the main end-to-end table reports 80.97%. The gap between encoder InfoNCE and decoder JS highlights the decoder's role: identifying a related molecular representation still leaves the model to express atom choices and connectivity as a valid sequence, so a task-facing conditional signal is central to that translation. The ordering supports a division of labor: molecular correspondence is useful when linked to conditional structural decisions.

Hybrid alignment changes how observations interact inside the encoder rather than adding new measurements. The average attention summary identifies which observations receive more attention after adaptation, and the layerwise summary locates where that preference emerges; read together, they tie the redistribution of evidence to encoder depth while keeping structural prediction as the downstream test. Layerwise differences also indicate where interventions should act, turning the attention analysis into a guide for verifiable mechanism studies. Within-formula retrieval and branch-level error analyses can likewise test which chemical ambiguities account for the alignment gain, while matched repeated runs quantify its variability. Carbon-directed shifts vary with the query modality and the layer in which they arise, so the redistribution is not a single global reallocation. Contrastive alignment retains distinctions between molecules across views, and soft prediction alignment connects those distinctions to structural continuations; the paired losses keep both requirements active throughout adaptation.

A routed teacher provides an initial interpretation that the Full student improves: 48.87% Top-1 for the teacher, 53.09% after Stage 1, and 55.66% after Stage 2, with Top-5 following 70.40, 73.78, and 76.51%. The two stages add 4.22 and 2.57 Top-1 points respectively, so learning a Full mapping and refining accepted targets each contribute beyond selecting a teacher view.

Figure 6 relates these decisions to encoder evidence allocation. Final-layer mass and layerwise changes show a distributed reorganization, consistent with renewed use of composition and carbon-environment constraints. Attention is a descriptive diagnostic; controlled interventions are needed to establish which changes cause retrieval improvements. The independent depth profile (Figure 7) shows final-layer increases of 9.61 and 10.88 points toward formula and carbon, and decreases of 7.30 and 12.78 toward proton and MS. In the paired 500-record cohort, Full validity rises from 51.4% to 100%, functional-group similarity from .1351 to .3109, and MCES distance falls from 16.294 to 9.070. Carbon-only MCES increases from 8.272 to 9.484, showing that Full gains do not imply uniform improvement for every subset.

The router decomposition gives a related view of conditional usefulness. Carbon has a positive mean unary contribution (+.3673), while proton, MS, and IR have negative means; several pair terms remain positive despite unfavorable unary terms. A measurement that is difficult to interpret alone can therefore support another constraint. This interaction pattern explains why routing compares combinations rather than assigning a single global reliability ranking to the spectra.

Router diagnostics further separate usable syntax from correct structure: target validity, formula agreement, and exact recovery are 98.26/82.77/48.87%. The selected-view distribution is dominated by carbon, motivating a matched fixed-carbon control. Historical fixed checkpoint intervals retain 1,236–2,180 targets and yield 54.70–55.74% Top-1. These records motivate the dynamic controller; they do not establish its advantage over matched fixed schedules. Appendix A provides the complete interval, router, structural, and coefficient studies.

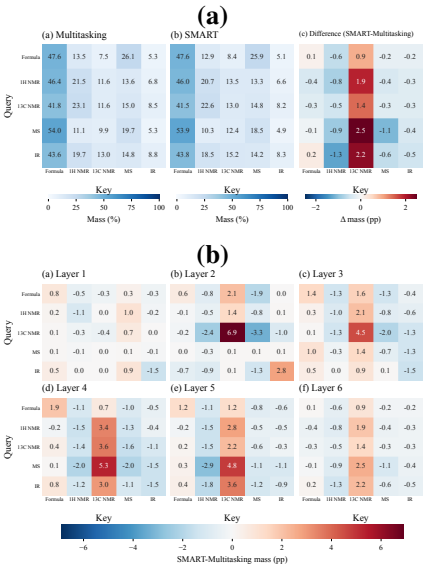

Figure 6: SDBS encoder attention: (a) final-layer mass (%); (b) layerwise changes (percentage points).

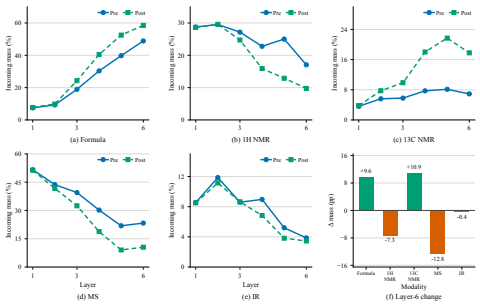

Figure 7: Encoder-depth attention profiles for the baseline and a historical Stage-2 checkpoint. Configuration and panel scales are detailed in Appendix G.

6 Conclusion

SMART addresses spectral negative transfer by separating the supervision-source view from full-modality inference. Structural-utility routing supplies a reference; encoder InfoNCE and decoder JS transfer molecular distinctions and conditional predictions; dynamic temporal consensus supports selective cross-view refinement. Experiments on real spectra show that this separation can recover the value of complementary measurements. The central principle is to adapt how the generator uses its evidence as acquisition conditions change.

AI USE STATEMENT

Generative AI tools (ChatGPT/Codex) assisted with manuscript drafting and revision, Chinese–English translation, literature discovery, interpretation of project records, mathematical exposition and derivation checks, pseudocode, and LaTeX typesetting. Bibliographic metadata were checked against primary publication records, and numerical results were retained from the experimental materials. The authors are responsible for the final scientific claims, citations, and presentation.

REPRODUCIBILITY STATEMENT

Appendix C states the assumptions and derivations; Appendix D specifies source perturbations; Appendix E describes input handling, loss reductions, and checkpoint selection. Appendix H provides algorithms, configuration tables, and complexity analysis. The accompanying materials include method code, numerical checks, original figure and table sources, and a result-version index. Target structures are reserved for offline evaluation.

ETHICS STATEMENT

The study concerns structure inference from public spectral resources. Dataset attribution and the separation between unlabeled adaptation and structural evaluation are maintained throughout the analysis. Generated structures are hypotheses for scientific interpretation; chemically valid predictions can still have incorrect connectivity or stereochemistry. Their use in experimental decisions requires the same evidential scrutiny as other computational structure proposals.

REFERENCES

Rishabh Agarwal, Nino Vieillard, Yongchao Zhou, Piotr Stanczyk, Sabela Ramos Garea, Matthieu Geist, and Olivier Bachem. On-Policy Distillation of Language Models: Learning from Self-Generated Mistakes. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/5be69a584901a26c521c2b51e40a4c20-Abstract-Conference.html.

Marvin Alberts, Oliver Schilter, Federico Zipoli, Nina Hartrampf, and Teodoro Laino. Unraveling Molecular Structure: A Multimodal Spectroscopic Dataset for Chemistry. In *Advances in Neural Information Processing Systems*, volume 37, pp. 125780–125808, 2024. doi: 10.52202/079017-3996. URL https://doi.org/10.52202/079017-3996.

Marvin Alberts, Nina Hartrampf, and Teodoro Laino. Automated Structure Elucidation at Human-Level Accuracy via a Multimodal Multitask Language Model. ChemRxiv preprint, 2025. URL https://doi.org/10.26434/chemrxiv-2025-q80r9. Preprint, posted 30 May 2025.

David Blackwell. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953. doi: 10.1214/aoms/1177729032. URL https://doi.org/10.1214/aoms/1177729032.

Chemotion. Chemotion. Chemotion website, 2026. URL https://chemotion.net/. Online resource; accessed 26 September 2026.

Mingcai Chen, Baoming Zhang, Zongbo Han, Wenyu Jiang, Yanmeng Wang, Shuai Feng, Yuntao Du, and Bingkun Bao. Test-Time Selective Adaptation for Uni-Modal Distribution Shift in Multi-Modal Data. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 9711–9727. PMLR, 2025. URL https://proceedings.mlr.press/v267/chen25ch.html.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 1597–1607. PMLR, 2020. URL https://proceedings.mlr.press/v119/chen20j.html.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, second edition, 2006. doi: 10.1002/047174882x. URL https://doi.org/10.1002/047174882x.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the Knowledge in a Neural Network. arXiv preprint arXiv:1503.02531, 2015. URL https://arxiv.org/abs/1503.02531.

J. Lin. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115. URL https://doi.org/10.1109/18.61115.

Laura Mismetti, Marvin Alberts, Andreas Krause, and Mara Graziani. Test-Time Tuned Language Models Enable End-to-end De Novo Molecular Structure Generation from MS/MS Spectra, 2025. URL https://arxiv.org/abs/2510.23746. arXiv preprint 2510.23746; revised January 2026.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofo Chen, Shijian Zheng, Peilin Zhao, and Mingkui Tan. Efficient Test-Time Model Adaptation without Forgetting. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162, pp. 16888–16905. PMLR, 2022. URL https://proceedings.mlr.press/v162/niu22a.html.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Zhiquan Wen, Yaofo Chen, Peilin Zhao, and Mingkui Tan. Towards Stable Test-Time Adaptation in Dynamic Wild World. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2302.12400.

Yunyue Su, Jiahui Chen, Zao Jiang, Zhenyi Zhong, Liang Wang, Qiang Liu, and Zhaoxiang Zhang. SpectraLLM: Uncovering the Ability of LLMs for Molecule Structure Elucidation from Multi-Spectra. In *Proceedings of the 14th International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/2aad9aaf0e14591f7b7d4770752eec89-Abstract-Conference.html.

Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei Efros, and Moritz Hardt. Test-Time Training with Self-Supervision for Generalization under Distribution Shifts. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 9229–9248. PMLR, 2020. URL https://proceedings.mlr.press/v119/sun20b.html.

Pierre Tremouilhac, Pei-Chi Huang, Chia-Lin Lin, Yu-Chieh Huang, An Nguyen, Nicole Jung, Felix Bach, and Stefan Bräse. Chemotion Repository, a Curated Repository for Reaction Information and Analytical Data. *Chemistry–Methods*, 1(1):8–11, 2021. doi: 10.1002/cmtd.202000034. URL https://doi.org/10.1002/cmtd.202000034.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation Learning with Contrastive Predictive Coding. arXiv preprint arXiv:1807.03748, 2018. URL https://arxiv.org/abs/1807.03748.

Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully Test-time Adaptation by Entropy Minimization. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=uXl3bZLkr3c.

Mouxing Yang, Yunfan Li, Changqing Zhang, Peng Hu, and Xi Peng. Test-time Adaptation against Multi-modal Reliability Bias. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/f25d75fc760aec0a6174f9f5d9da59b8-Abstract-Conference.html.

A Extended empirical results

A.1 Stage-wise retrieval and partial-view behavior

The complete stage comparison and Stage-1 partial-view results appear in Tables 2 and 3. Reading them together shows how the Full student changes relative to its routed teacher and how the update transfers to other visible subsets. The teacher is fixed before adaptation, whereas the two student stages alter the shared generator. The endpoint differences are reported with the evaluation views held explicit.

Table 2: SDBS retrieval for the routed teacher and the two student stages.

	Zero Shot			Test-Time Tuning					
				Stage 1			Stage 2		
Modality	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑
H	4.03±0.24	8.72±0.38	11.31±0.45	13.06±0.52	27.23±0.75	33.63±0.88	22.16±0.61	45.13±0.88	55.38±1.02
C	**50.70**±0.72	**71.49**±0.86	**76.04**±0.94	48.30±0.83	68.66±1.02	74.13±1.14	52.25±0.78	72.88±0.97	77.65±1.10
MS	2.04±0.18	4.99±0.27	6.76±0.34	9.35±0.41	24.72±0.68	31.97±0.82	13.49±0.46	33.96±0.71	44.04±0.86
IR	1.88±0.16	4.96±0.25	7.28±0.31	2.64±0.29	7.03±0.47	9.92±0.61	7.22±0.35	18.56±0.56	26.60±0.72
H + C	13.76±0.48	25.81±0.67	31.92±0.79	49.66±0.76	68.74±0.98	74.52±1.12	55.16±0.74	75.31±0.95	79.75±1.09
C + MS	11.17±0.44	21.18±0.61	26.11±0.73	46.20±0.69	66.78±0.91	72.47±1.05	51.87±0.68	72.85±0.89	77.54±1.03
C + IR	30.44±0.56	50.23±0.74	56.72±0.88	39.38±0.63	58.63±0.84	65.25±0.98	50.18±0.72	70.92±0.94	75.74±1.08
H + C + MS	5.70±0.35	11.15±0.52	14.64±0.63	52.74±0.82	**73.92**±1.04	**78.09**±1.18	55.41±0.81	76.12±1.03	80.54±1.17
H + C + IR	10.90±0.39	20.90±0.58	27.56±0.69	44.84±0.77	64.05±0.99	69.96±1.13	53.18±0.76	73.81±0.98	78.82±1.12
C + MS + IR	8.86±0.46	17.99±0.64	22.21±0.76	43.64±0.71	63.56±0.93	70.24±1.07	50.70±0.73	72.09±0.95	76.94±1.09
H + C + MS + IR	5.94±0.42	11.42±0.59	14.91±0.71	**53.09**±0.79	73.78±1.01	78.00±1.15	**55.66**±0.84	**76.51**±1.06	**80.97**±1.20

Table 3: SDBS retrieval before adaptation and after Stage 1. Values are percentages; differences are percentage points.

	multitasking			SMART			Δ (improvement)		
Modality	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑
H	4.03±0.24	8.72±0.38	11.31±0.45	13.06±0.52	27.23±0.75	33.63±0.88	+9.03	+18.51	+22.32
C	**50.70**±0.72	**71.49**±0.86	**76.04**±0.94	48.30±0.83	68.66±1.02	74.13±1.14	-2.40	-2.83	-1.91
MS	2.04±0.18	4.99±0.27	6.76±0.34	9.35±0.41	24.72±0.68	31.97±0.82	+7.31	+19.73	+25.21
IR	1.88±0.16	4.96±0.25	7.28±0.31	2.64±0.29	7.03±0.47	9.92±0.61	+0.76	+2.07	+2.64
H + C	13.76±0.48	25.81±0.67	31.92±0.79	49.66±0.76	68.74±0.98	74.52±1.12	+35.90	+42.93	+42.60
C + MS	11.17±0.44	21.18±0.61	26.11±0.73	46.20±0.69	66.78±0.91	72.47±1.05	+35.03	+45.60	+46.36
C + IR	30.44±0.56	50.23±0.74	56.72±0.88	39.38±0.63	58.63±0.84	65.25±0.98	+8.94	+8.40	+8.53
H + C + MS	5.70±0.35	11.15±0.52	14.64±0.63	52.74±0.82	**73.92**±1.04	**78.09**±1.18	+47.04	**+62.77**	**+63.45**
H + C + IR	10.90±0.39	20.90±0.58	27.56±0.69	44.84±0.77	64.05±0.99	69.96±1.13	+33.94	+43.15	+42.40
C + MS + IR	8.86±0.46	17.99±0.64	22.21±0.76	43.64±0.71	63.56±0.93	70.24±1.07	+34.78	+45.57	+48.03
H + C + MS + IR	5.94±0.42	11.42±0.59	14.91±0.71	**53.09**±0.79	73.78±1.01	78.00±1.15	**+47.15**	+62.36	+63.09

A.2 Router probabilities and selected views

The supplied routing distribution is strongly concentrated on carbon NMR: the displayed argmax shares are approximately 97.5% carbon, 2.1% IR, .3% proton, and .1% H+C+IR. This concentration is consistent with the strong zero-shot carbon baseline in the corresponding study. A direct fixed-carbon control is therefore essential for determining how much the router contributes beyond a dominant global view. The supplementary mechanistic router panel provides a complementary decomposition of action scores and structural success.

A.3 Candidate lists and stage-specific recoveries

The Top-5 examples show how adaptation changes both structure and rank. They are selected cases that illustrate different outcomes within the candidate list. Aggregate Top-k results provide the population-level complement: a candidate can be chemically closer yet still fail exact matching, or recover the correct structure below rank one. The stage-specific case records include molecules recovered only after refinement and molecules recovered by both student stages after the teacher failed. The supplied artwork shows the target, teacher interpretation, and observed spectra; the stage

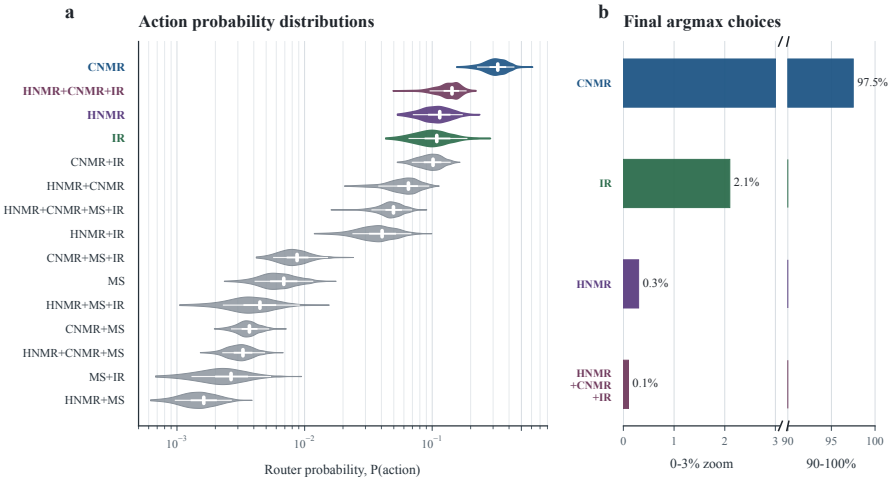

Figure 8: SDBS router probabilities and selected actions. Carbon NMR accounts for 97.5% of displayed argmax choices.

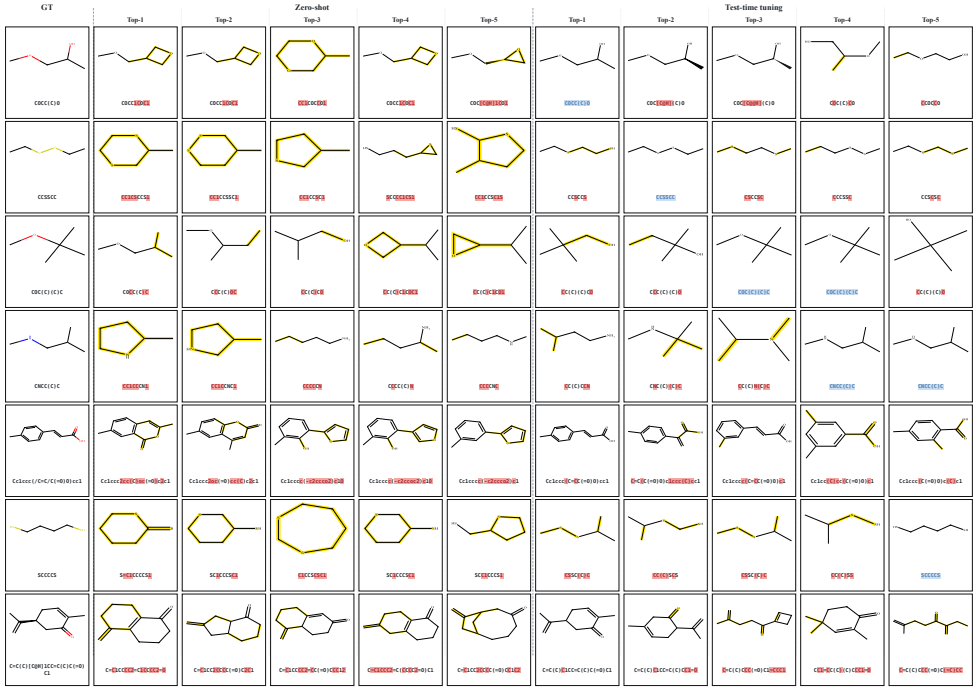

Figure 9: Selected SDBS Top-5 candidates with their original rank, connectivity, and stereochemical annotations.

outcome labels come from the accompanying prediction records. These examples are useful for designing broader analyses of functional groups and connectivity changes across the entire cohort.

A.4 Training trajectories and computational context

The trajectory panels summarize the supplied SDBS and Chemotion runs. A falling training loss describes progress on each stage's own objective. Since the objectives and accepted populations differ, their absolute loss levels should be interpreted within a stage. The lower panels use training-only compute estimates and timing proxies. End-to-end deployment cost additionally includes routing features, teacher decoding, probing, consensus decoding, and final evaluation. Reporting these components separately would make the tradeoff between structural gain and adaptation cost reproducible.

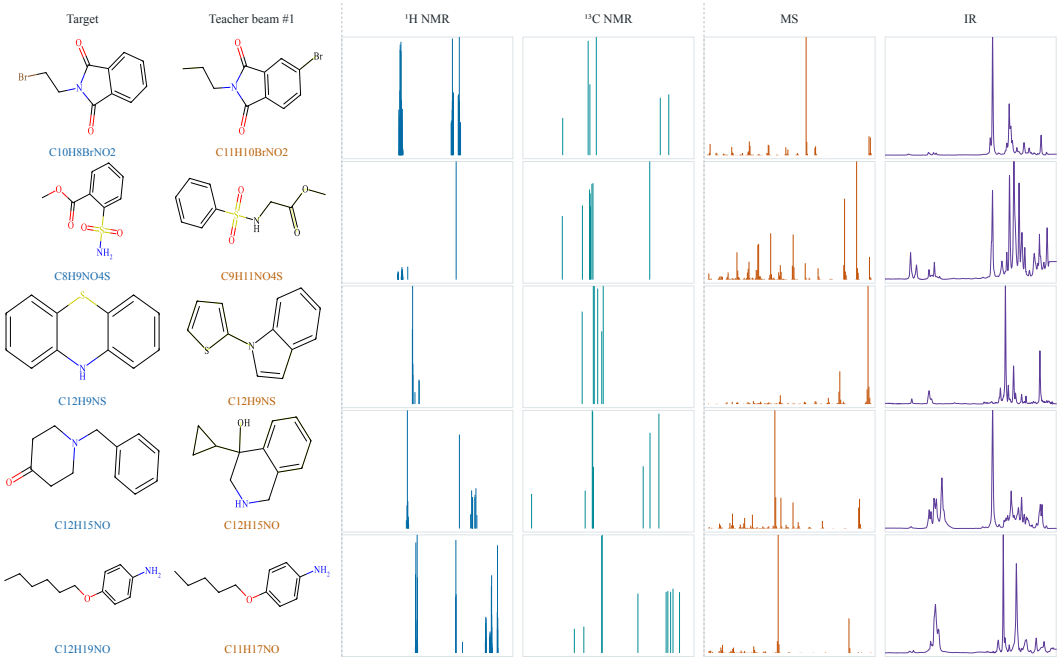

Figure 10: SDBS cases recovered by Stage 2 after teacher and Stage-1 errors. Structures and spectra follow the associated case records.

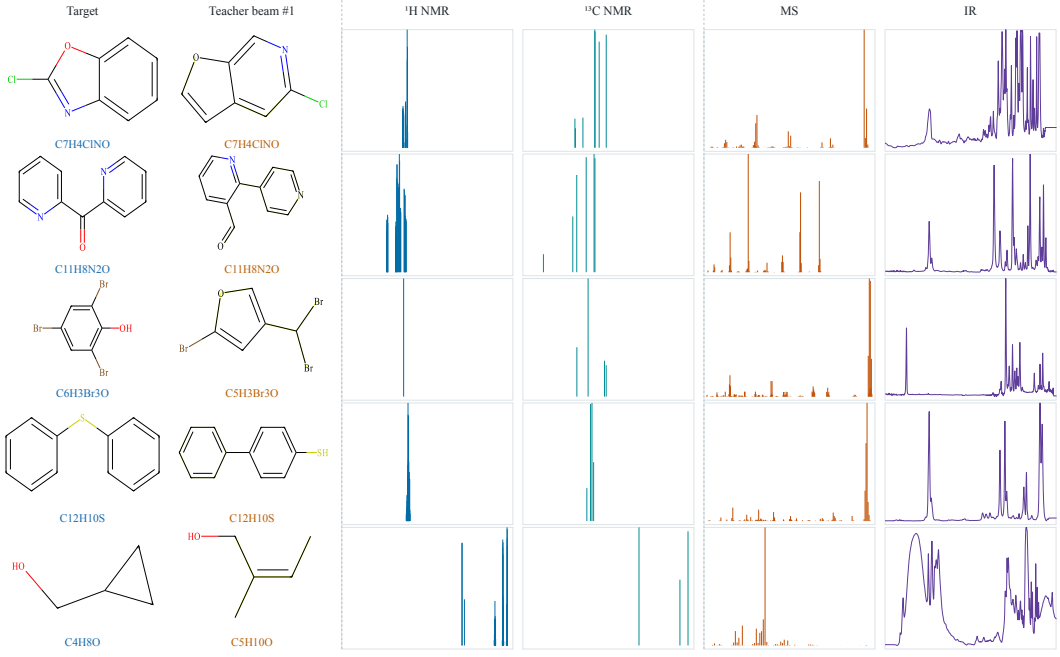

Figure 11: SDBS teacher errors recovered by both stages, including composition and connectivity ambiguities.

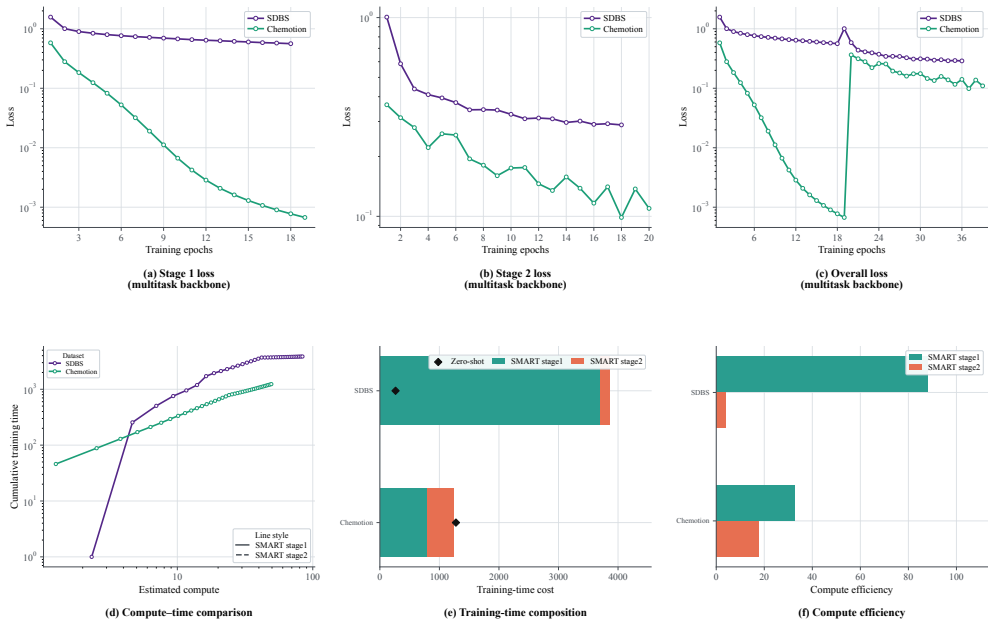

Figure 12: Training losses and computational summaries. Cost panels report training-only FLOP estimates and timing proxies.

A.5 Complete paired structure-quality metrics

Table 4: Complete unadapted structure-quality metrics on the paired 500-record SDBS cohort.

Input	Valid. ↑	ECFP4 ↑	Cosine ↑	MCES ↓	FG ↑	MACCS ↑	Fraggle ↑
H	70.4%	0.0358	0.0704	12.240	0.2143	0.0995	0.1068
C	63.2%	0.0321	0.0624	8.272	0.2064	0.0821	0.1080
MS	68.4%	0.0360	0.0690	21.974	0.1983	0.0976	0.1281
IR	81.2%	0.0297	0.0579	9.010	0.2074	0.0808	0.1186
H+C	64.6%	0.0353	0.0709	16.016	0.1967	0.0872	0.1032
H+C+IR	62.2%	0.0273	0.0557	10.150	0.1742	0.0660	0.0855
H+C+MS	55.6%	0.0308	0.0622	13.930	0.1663	0.0765	0.0981
H+C+IR+MS	51.4%	0.0251	0.0519	16.294	0.1351	0.0610	0.0678

Table 5: Complete SMART structure-quality metrics on the same paired 500-record SDBS cohort.

Input	Valid. ↑	ECFP4 ↑	Cosine ↑	MCES ↓	FG ↑	MACCS ↑	Fraggle ↑
H	98.6%	0.0573	0.1063	9.536	0.3547	0.1391	0.1120
C	98.2%	0.0541	0.1020	9.484	0.3390	0.1285	0.1101
MS	98.4%	0.0532	0.1000	9.820	0.3496	0.1299	0.1191
IR	100.0%	0.0522	0.0991	9.638	0.2865	0.1296	0.0602
H+C	98.4%	0.0560	0.1057	9.478	0.3537	0.1359	0.1146
H+C+IR	100.0%	0.0620	0.1159	9.356	0.3099	0.1224	0.0431
H+C+MS	98.8%	0.0538	0.1021	9.576	0.3459	0.1251	0.1157
H+C+IR+MS	100.0%	0.0665	0.1236	9.070	0.3109	0.1312	0.0408

Tables 4 and 5 retain all supplied metrics for the same paired 500-record cohort. The metrics cover validity, graph and fingerprint similarity, edit distance, and functional-group recovery, following the

structural evaluation dimensions used by Su et al. (2026). Invalid candidates receive the recorded zero-similarity treatment. MCES is interpreted in the opposite direction to similarity metrics: lower distance is better. Sample identities, canonicalization, and invalid-output handling must remain common when comparing the two tables.

B Supplementary empirical analyses

B.1 Evaluation setting and the evidence considered

We study transfer from simulated spectra to experimental measurements using the Multitasking generator of Alberts et al. (2025). The primary SDBS cohort contains 3,669 molecules ; a second cohort contains 1,995 Chemotion records under the project's strict input-selection protocol (Tremouilhac et al., 2021; Chemotion, 2026). Adaptation is transductive: the evaluation cohort supplies unlabeled spectra and observed formulas, while molecular structures are reserved for offline scoring. Unless otherwise specified, exact-match retrieval uses beam ten and reports Top-1/5/10 over the complete cohort. This protocol evaluates learning from deployment observations with structural supervision confined to the source domain. Tables use H/C/MS/IR for proton NMR, carbon NMR, mass spectra, and infrared spectra; the observed formula accompanies every view.

A paired 500-record SDBS analysis adds validity, graph similarity, and functional-group measures. These paired observations are analyzed within their own cohort, whose checkpoint lineage differs from the 3,669-molecule retrieval benchmark. Accuracy differences below are percentage points. Together, exact retrieval and structural diagnostics distinguish recovery of the correct molecule from improvements in the chemical plausibility of nearby candidates.

B.2 Can adaptation recover the value of complete observations?

The most demanding comparison is against the strongest original subset: restoring a damaged Full model is useful, but recovering complementarity requires going beyond that subset. In Table 1, unadapted Full reaches 5.94% Top-1, whereas carbon NMR reaches 50.70%. SMART raises Full to 55.66%, recovering 49.72 points over its original prediction and exceeding the carbon reference by 4.96 points. The result places the adapted joint interpretation above the strongest starting view, which is the central deployment objective.

Candidate coverage shows the same recovery. Full Top-5/10 rise from 11.42/14.91% to 76.51/80.97% in the supplied Stage-2 table. Adapted H+C and H+C+MS also reach 55.16% and 55.41% Top-1. Carbon-only adaptation reaches 52.25%, leaving a 3.41-point advantage for adapted Full. These comparisons connect the gain to a chemical opportunity: carbon environments provide a strong starting constraint, while other spectra can help resolve the alternatives left by that constraint. Several adapted views support the recovery, and the best Top-1 endpoint in this table uses the complete observation.

B.3 Does the model redistribute attention toward useful constraints?

The attention analyses connect recovered retrieval to changes in the encoder's allocation of evidence. Figure 6 combines the average and layerwise summaries; Figure 7 provides an independent depth profile.

Adaptation changes how observations interact inside the encoder. Panel (a) summarizes the final allocation of attention, while panel (b) shows where that allocation changes across depth. Carbon-directed changes vary with the query modality and layer, consistent with a distributed reorganization of evidence. Formula constrains composition and carbon NMR constrains local atomic environments, giving these shifts a chemical reference. Controlled spectral interventions can then test which redistributed interactions affect structural decisions.

The two panels answer complementary questions. The average summary identifies which observations receive more attention, and the layerwise summary locates where this preference emerges. Reading them together ties the change in evidence allocation to the encoder's depth while keeping structural prediction as the downstream test.

A separate checkpoint shows how aggregate attention changes with encoder depth. In the final layer, attention to formula and carbon increases by 9.61 and 10.88 percentage points, while proton and MS attention decreases by 7.30 and 12.78 points. This pattern is consistent with renewed emphasis on composition and carbon-environment constraints. Its contribution to prediction can be tested by perturbing background peaks and referencing while tracking the resulting molecular structures.

B.4 How does the student progress beyond its teacher?

A routed teacher provides an initial interpretation that the Full student can improve. Figure 13 reports Top-1 of 48.87% for the teacher, 53.09% after Stage 1, and 55.66% after Stage 2 in the supplied summary. Top-5 follows the progression 70.40%, 73.78%, and 76.51%. The two stages therefore add 4.22 and 2.57 Top-1 points, respectively. Learning a Full mapping and refining accepted targets each contribute beyond selecting a teacher view.

Chemically, a subset may identify a plausible family of structures while leaving connectivity or functional-group placement ambiguous. A student with all observations has the opportunity to revise that initial hypothesis. The stage-wise gains are consistent with this opportunity, and the structural cases below illustrate where it is realized. Per-molecule teacher–student transitions would further separate corrected interpretations from newly introduced errors. The aggregate takeaway is that the teacher supplies a useful starting judgment rather than the final performance ceiling.

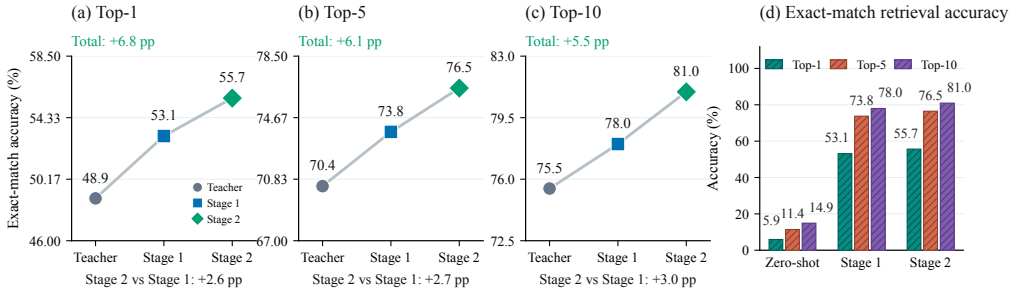

Figure 13: Stage-wise SDBS retrieval: the teacher uses its routed view and the student uses all spectra. Stage-2 Top-1/5/10 are 55.66/76.51/80.97%.

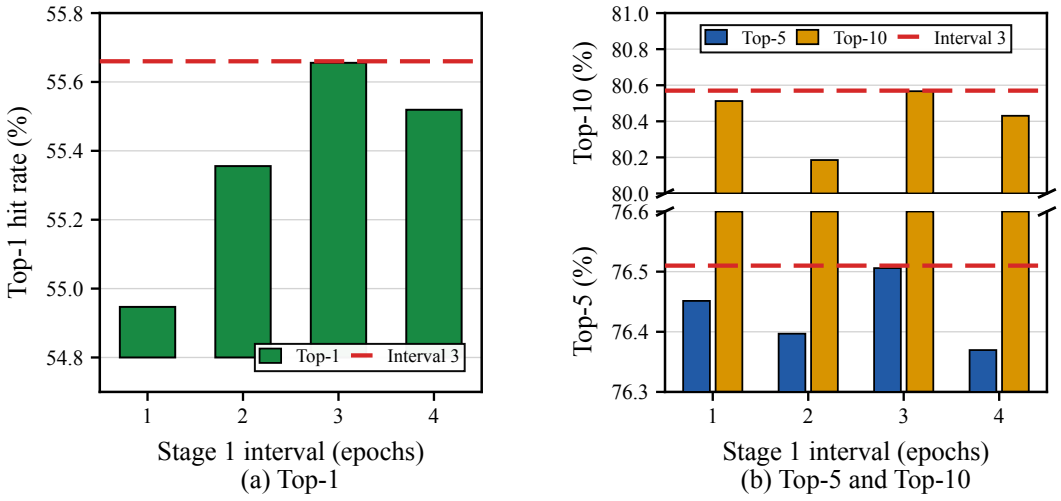

Figure 14: Display-adjusted checkpoint-interval comparison with a broken axis in (b). Counts are anchored to interval three, with interval four five hits lower; original results are listed in Appendix G.

B.5 When should structural persistence be assessed?

Checkpoint spacing determines how many model states an accepted interpretation must survive. Original intervals one through four retain 1,236, 1,716, 2,035, and 2,180 targets. Their Top-1 spans 54.70–55.74%; intervals three and four reach 55.41% and 55.74%. Wider spacing changes both se-

lection strictness and coverage. These original measurements motivate a trajectory-aware schedule; the supplied graphic uses the display adjustment identified below.

Figure 14 retains the supplied display-adjusted graphic; the numerical comparisons above use the original sweep. These records motivate observing structural change directly. A switch between formula-compatible isomers can be informative even when both predictions are valid, whereas repeating the same feasible structure may add little new evidence. The dynamic controller allocates observations around this distinction. Its matched comparison should hold adaptation budget and checkpoint count fixed, so that the value of observation timing can be separated from selection strictness.

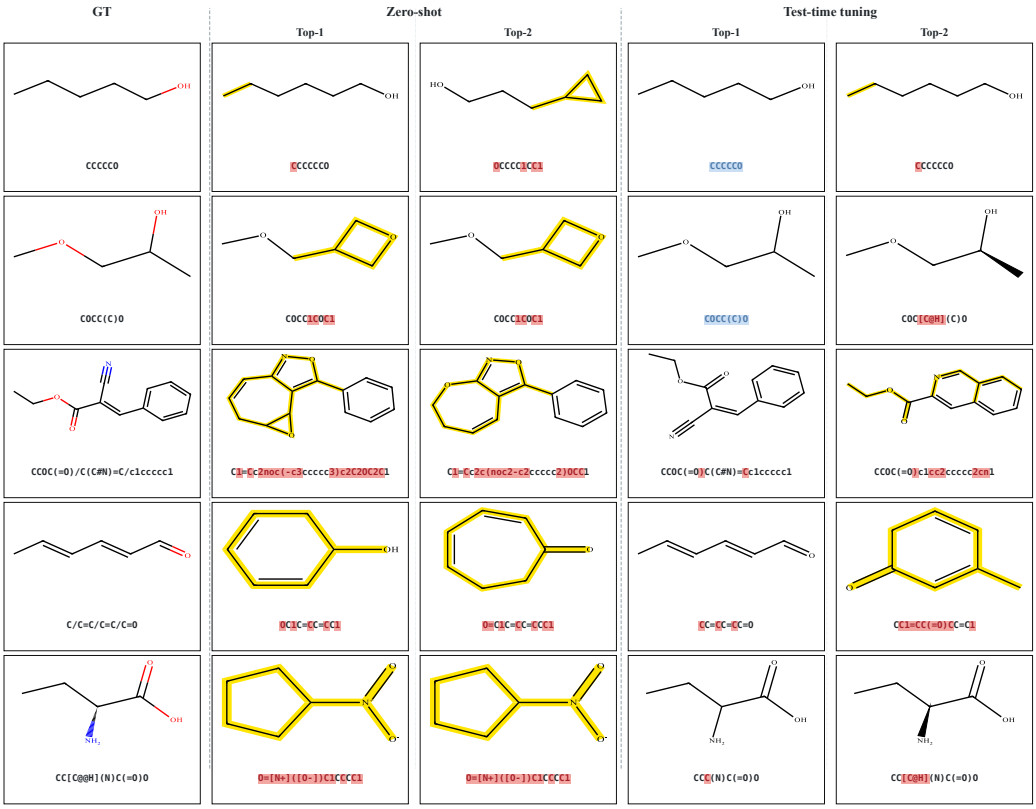

Figure 15: Representative SDBS Top-2 predictions illustrating changes in connectivity and stereochemistry.

B.6 Which aspects of molecular structure improve?

Candidate changes make retrieval chemically interpretable. These examples include functional-group placement, carbon connectivity, and oxygen-containing substructures. Some candidates recover connectivity while leaving stereochemical distinctions unresolved. They illustrate which ambiguities complementary spectra can help address and why molecular evaluation should distinguish composition, connectivity, and stereochemistry. Further candidate lists and stage-specific recovery cases appear in Appendix A.

The paired 500-record analysis in Table 6 examines this progression across several measures. Full validity rises from 51.4% to 100%, mean structural cosine similarity from .0519 to .1236, and functional-group similarity from .1351 to .3109; MCES distance falls from 16.294 to 9.070. The joint movement is more informative than validity alone: the model produces usable structures that also move closer to the reference in graph and functional-group space.

Partial-view effects remain heterogeneous. Carbon-only MCES increases from 8.272 to 9.484, even as the Full metrics improve. A shared generator can therefore improve its combined interpretation while changing a subset's structural errors in a different direction. Reporting the full set of paired

metrics preserves this distinction. The practical target is accurate molecular identification, for which validity, topology, and functional-group recovery provide complementary diagnostics.

Table 6: Paired structural metrics on 500 SDBS records. MCES is lower-is-better; other metrics are higher-is-better.

	Zero shot				Test-time tuning				Δ (improvement)			
Modality	Valid. ↑	Cosine ↑	MCES ↓	FG ↑	Valid. ↑	Cosine ↑	MCES ↓	FG ↑	Valid. (pp)	Cosine	MCES	FG
H	70.4%	0.0704	12.240	0.2143	98.6%	0.1063	9.536	0.3547	+28.2	+0.0359	-2.704	+0.1404
C	63.2%	0.0624	8.272	0.2064	98.2%	0.1020	9.484	0.3390	+35.0	+0.0396	+1.212	+0.1326
MS	68.4%	0.0690	21.974	0.1983	98.4%	0.1000	9.820	0.3496	+30.0	+0.0310	-12.154	+0.1513
IR	81.2%	0.0579	9.010	0.2074	100.0%	0.0991	9.638	0.2865	+18.8	+0.0412	+0.628	+0.0791
H+C	64.6%	0.0709	16.016	0.1967	98.4%	0.1057	9.478	0.3537	+33.8	+0.0348	-6.538	+0.1570
H+C+IR	62.2%	0.0557	10.150	0.1742	100.0%	0.1159	9.356	0.3099	+37.8	+0.0602	-0.794	+0.1357
H+C+MS	55.6%	0.0622	13.930	0.1663	98.8%	0.1021	9.576	0.3459	+43.2	+0.0399	-4.354	+0.1796
H+C+IR+MS	51.4%	0.0519	16.294	0.1351	100.0%	0.1236	9.070	0.3109	+48.6	+0.0717	-7.224	+0.1758

B.7 Why are representation and prediction alignment complementary?

Figure 5 tests whether useful representation geometry is sufficient for structural generation. InfoNCE alone reaches 14.9% Top-1, while decoder JS reaches 54.5%, KL 54.1%, and hard CE 51.4%. The difference highlights the decoder's role: identifying a related molecular representation still leaves the model to express atom choices and connectivity as a valid sequence. A task-facing conditional signal is central to that translation.

Combining InfoNCE with JS reaches 55.66/76.51/80.57% at Top-1/5/10, improving over JS by 1.16/2.31/1.77 points. This ordering supports the proposed division of labor. Contrastive alignment retains distinctions between molecules across views; soft prediction alignment connects those distinctions to structural continuations. Within-formula retrieval and branch-level error analyses can test which chemical ambiguities account for the gain, while matched repeated runs quantify its variability. In the supplied comparison, their combination produces the strongest structural retrieval.

B.8 How does acquisition shift affect transfer to another cohort?

Input-space differences in Figure 16 vary across modalities. Peak multiplicity changes the discrete evidence seen in NMR; intensity and background alter the prominence of mass-spectral entries; and infrared shifts change continuous signal shape and scale. These changes act on different chemical observables and justify separate perturbation mechanisms during source routing. The useful abstraction is heterogeneous acquisition variation, with its effect on structural utility evaluated after preprocessing.

Chemotion provides a severe out-of-distribution stress test: unadapted full-modality Top-1 is .100%, rising to 4.511% after adaptation, while Top-5/10 rise from .201/.351% to 7.018/8.271% (Table 7). The Top-1 endpoint is approximately 45.1 times the baseline, with an absolute gain of 4.411 percentage points. Under a fixed architecture and without target structural labels, the update recovers discrimination from a predictor whose initial exact-match decisions are nearly always wrong. The result demonstrates that usable structure-related information remains accessible through adaptation even in this low-accuracy regime.

The risk perspective clarifies the scientific value of this recovery. On the same target population, the Bayes component is common to both predictors; their difference reflects a reduction in the learned rule's structural error. Routing and refinement therefore address a recoverable part of the deployment mismatch within the existing generator. Reporting the relative gain alongside absolute retrieval makes this a test of adaptation under severe shift. Further improvement in candidate quality and exact identification remains the central objective for this cohort.

The retrieval endpoint also matters. Adapted Full has the highest Top-1 in the table, whereas C+MS reaches 9.323% Top-10 and exceeds Full at that cutoff. For a chemist inspecting a shortlist, complementary carbon and fragmentation evidence can therefore remain useful even when another view ranks the leading candidate better. The result supports recovery of the leading Full decision while showing that candidate diversity and rank quality deserve separate optimization.

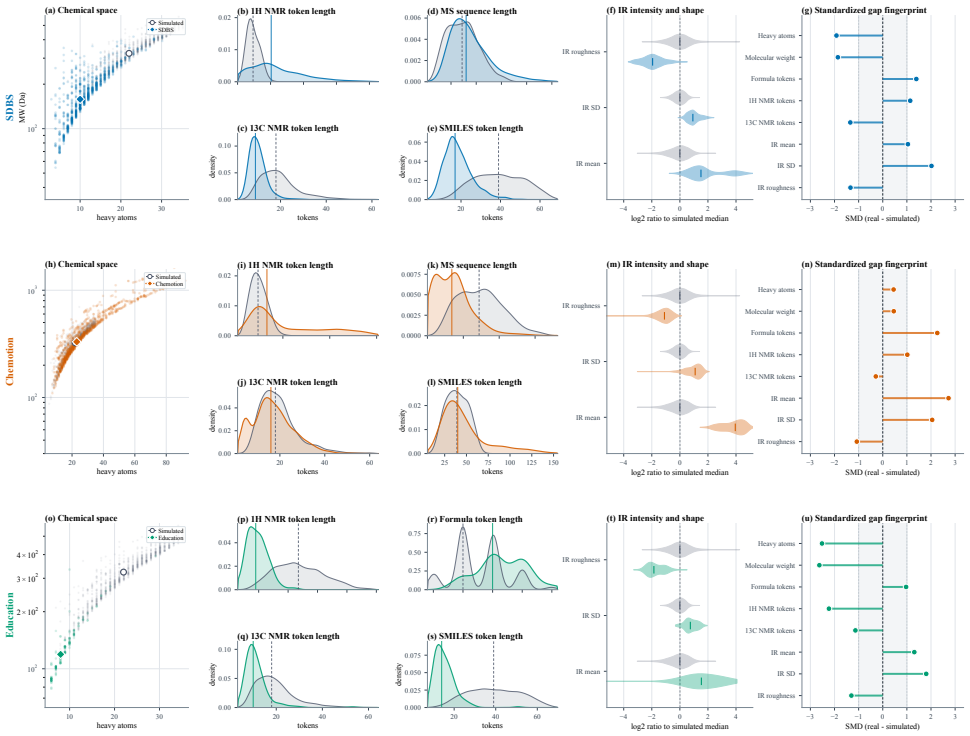

Figure 16: Input-space differences across simulated and experimental spectral sources. Education uses observed formula length in the position assigned to MS length for other cohorts.

Table 7: Retrieval before and after adaptation on 1,995 Chemotion records. Values are percentages; differences are percentage points.

	Multitasking			SMART			Δ (improvement)		
Modality	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1 ↑	Top-5 ↑	Top-10 ↑	Top-1	Top-5	Top-10
H	0.150±0.02	0.201±0.03	0.301±0.04	2.356±0.08	4.411±0.12	5.714±0.15	2.206	4.210	5.413
C	3.609±0.06	6.366±0.09	8.020±0.12	3.759±0.10	7.368±0.14	8.972±0.18	0.150	1.002	0.952
MS	0.251±0.02	0.451±0.03	0.602±0.05	1.554±0.07	3.409±0.10	4.361±0.13	1.303	2.958	3.759
IR	0.050±0.02	0.251±0.03	0.301±0.04	1.855±0.06	3.860±0.09	5.564±0.12	1.805	3.609	5.263
H + C	0.201±0.03	0.401±0.05	0.652±0.07	4.060±0.09	**7.419**±0.13	8.672±0.16	3.859	**7.018**	8.020
H + MS	0.050±0.04	0.100±0.06	0.150±0.08	2.206±0.08	4.311±0.11	5.414±0.14	2.156	4.211	5.264
H + IR	0.050±0.01	0.150±0.02	0.150±0.03	2.406±0.08	4.662±0.11	5.564±0.15	2.356	4.512	5.414
C + MS	1.454±0.05	3.008±0.08	3.759±0.11	4.110±0.10	**7.419**±0.14	**9.323**±0.18	2.656	4.411	5.564
C + IR	1.404±0.05	3.158±0.09	3.910±0.12	4.160±0.11	7.068±0.15	8.521±0.19	2.756	3.910	4.611
MS + IR	0.000±0.02	0.100±0.03	0.201±0.04	1.704±0.07	3.459±0.10	4.261±0.13	1.704	3.359	4.060
H + C + MS	0.150±0.05	0.351±0.08	0.501±0.11	4.361±0.10	7.368±0.15	8.571±0.19	4.211	**7.017**	8.070
H + C + IR	0.050±0.04	0.351±0.07	0.551±0.09	4.160±0.11	7.218±0.15	8.622±0.20	4.110	6.867	8.071
H + MS + IR	0.050±0.01	0.150±0.03	0.201±0.05	2.256±0.09	4.411±0.13	5.664±0.17	2.206	4.261	5.463
C + MS + IR	0.602±0.06	1.253±0.09	1.855±0.13	3.910±0.11	7.419±0.16	9.273±0.21	3.308	6.166	7.418
H + C + MS + IR	0.100±0.05	0.201±0.08	0.351±0.11	**4.511**±0.12	7.018±0.17	8.271±0.22	**4.411**	6.817	7.920

B.9 What does the router capture about chemical evidence?

The score decomposition in Figure 17 connects individual evidence with its context. Carbon has a positive mean unary contribution (+.3673), while proton, MS, and IR have negative means. Several pair terms are positive despite unfavorable unary terms. This is a useful form of conditional reliability: a measurement that is difficult to interpret alone can become useful alongside another constraint. The decomposition gives the router a way to express that dependence within its low-order action model.

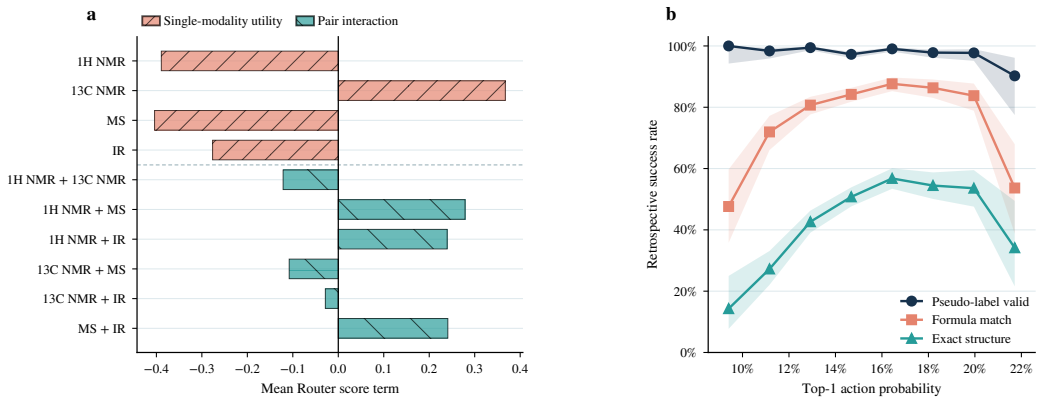

Figure 17: Router utilities on 8,795 source episodes and outcome rates on 3,669 target records. Shading shows Wilson 95% intervals across action-probability bins.

Target bins relate maximum action probability to validity, formula agreement, and exact recovery. Their overall rates are 98.26%, 82.77%, and 48.87%. The separation has a chemical interpretation: satisfying elemental inventory and basic molecular validity still leaves many possible isomers. The small highest-probability tail also shows declining structural success. Action concentration can guide which evidence to consult, while a structural confidence estimate must account for the ambiguity remaining after that choice.

B.10 How strongly should each learning signal act?

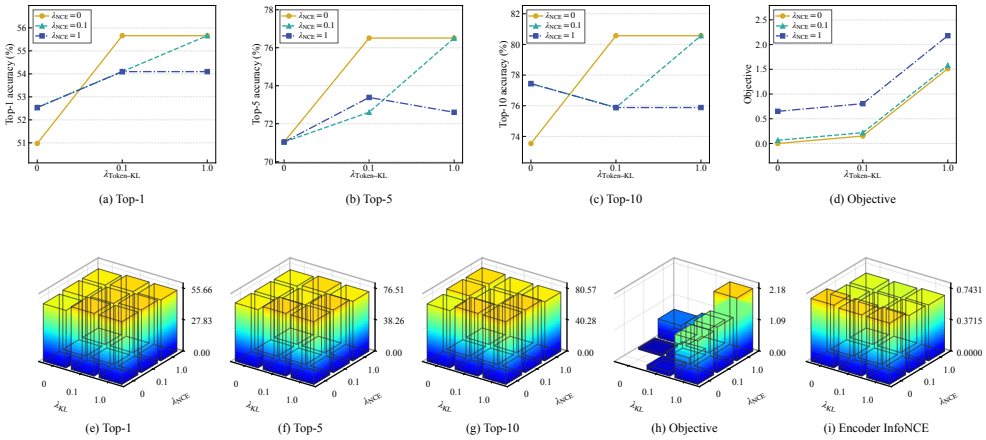

Figure 18: Display-shifted token-KL/InfoNCE pilot with 128 records and one epoch. Accuracy offsets and unshifted results appear in Appendix G; losses retain their original scale.

The sensitivity pilot varies decoder token-KL and encoder InfoNCE on 128 carbon-routed records for one epoch. Its figure uses shifted presentation surfaces, so we interpret the unshifted measurements here. With both coefficients zero, Top-1 is 8.59375%; decoder KL alone reaches 13.28125%, InfoNCE alone at .1 reaches 10.15625%, and the (.1,.1) combination reaches 11.71875%. Adding a second signal changes the optimization trajectory, and its relative strength matters even when both objectives have useful individual motivations.

The chemical question is how much of the teacher's representation should be preserved while the student learns to express a structure from Full evidence. Excessive emphasis on a coarse reference could hinder distinctions supplied by other spectra; insufficient emphasis could leave that reference ineffective. The KL pilot motivates measuring gradient norms and compatibility in the hybrid JS study, whose full-cohort point estimates appear separately in Figure 5. A useful balance should be judged by recovered structures and preserved distinctions, alongside the two loss values.

C Theoretical analysis

The analysis follows the same question as the method: how can a useful interpretation from one view improve decisions based on a broader observation? We first distinguish information from a learned fusion rule, then examine what the two alignment gradients transfer, and finally connect selection and subset learning to structural risk. Each result identifies a relationship that can be measured in the corresponding experiment.

C.1 Additional observations and Bayes decision risk

A Full observation retains every decision available from a subset. This nesting gives a general reason to seek better evidence integration under shift.

Proposition 1 (Risk under nested observations)**.** *Let $X^a = \pi_a(X^{\text{Full}})$ for a measurable projection π_a, and let the two decision problems share an output space and a nonnegative loss with finite expected risk for at least one rule. If all measurable decision rules are permitted, then $\mathcal{R}_t^*(\text{Full}) \le \mathcal{R}_t^*(a)$.*

Proof. For every subset rule f_a, the Full rule $f_a \circ \pi_a$ makes the same decision on every record. Their expected losses are equal under P_t. Taking the infimum over Full rules gives a value no larger than the risk of any such composed rule; taking the infimum over f_a establishes the result. □

This comparison concerns optimal expected loss at the target distribution. A fixed source decoder occupies only one point in each decision class and can react unfavorably to extra inputs. For discrete molecular identity with finite entropy, Eq. (17) follows by expanding both mutual informations as conditional-entropy differences; the remainder is a nonnegative conditional mutual information. The two results isolate the bottleneck addressed by SMART: a learned rule can lose access to useful evidence even when that evidence remains present.

The conditional mutual-information chain rule gives the corresponding information statement (Cover & Thomas, 2006):

$$I_t(Y; X_{\mathcal{M}} \mid F) = I_t(Y; X_a \mid F) + I_t(Y; X_{\mathcal{M}\setminus a} \mid F, X_a). \tag{17}$$

C.2 Contrastive alignment as relative identification

Molecular identity must remain distinguishable when the student sees additional acquisition detail. The contrastive objective expresses this requirement through the positive pair's advantage over competing records. For one student query, it becomes

$$\ell_i^{\text{NCE}} = \log\left[1 + \sum_{j \in J_i \setminus \{i\}} \exp\left(\frac{h_i^\top k_j - h_i^\top k_i}{\tau_C}\right)\right]. \tag{18}$$

This form makes the comparison explicit: each negative contributes according to its similarity relative to the positive. Minimization encourages the student to identify the matching teacher representation among the available alternatives. The geometry depends on relative margins, so exact feature reconstruction is unnecessary. A common orthogonal transformation of queries and keys preserves the objective.

Proposition 2 (Finite-batch range)**.** *For unit vectors, $M_i = |J_i| \ge 1$, and $\tau_C > 0$,*

$$\log[1 + (M_i - 1)e^{-2/\tau_C}] \le \ell_i^{\text{NCE}} \le \log[1 + (M_i - 1)e^{2/\tau_C}]. \tag{19}$$

Proof. Each dot product lies in $[-1, 1]$, so each difference in Eq. (18) lies in $[-2, 2]$. Bound each exponential and apply the increasing logarithm. For $M_i = 1$, both bounds equal zero. □

The range depends on temperature and candidate count. Batch composition therefore determines how much competition the encoder faces. In chemical datasets, a batch containing closely related isomers asks a different identification question from one dominated by unrelated formulas, even at the same nominal batch size.

C.3 Encoder gradients and teacher diversity

All derivatives below are taken with respect to the student parameters or student activations, conditional on fixed teacher parameters $\theta_T = \theta_0$. Teacher keys and logits are constants in this optimization problem. A near-matching alternative should influence identification more than a distant molecule. The softmax weights w_{ij} implement this preference, and the fixed-teacher gradient shows its direction explicitly:

$$\frac{\partial \ell_i^{\mathrm{NCE}}}{\partial h_i} = \frac{1}{\tau_C}\left(\sum_j w_{ij} k_j - k_i\right), \qquad \frac{\partial \ell_i^{\mathrm{NCE}}}{\partial u_i^S} = \frac{I - h_i h_i^\top}{\|u_i^S\|_2}\frac{\partial \ell_i^{\mathrm{NCE}}}{\partial h_i}. \tag{20}$$

The weighted key average emphasizes difficult alternatives. Normalization projects the update onto the tangent space of the unit sphere. Since the norm of a convex combination of unit vectors is at most one,

$$\|\partial_{h_i} \ell_i^{\mathrm{NCE}}\|_2 \le 2/\tau_C, \qquad \|\partial_{u_i^S} \ell_i^{\mathrm{NCE}}\|_2 \le 2/(\tau_C \|u_i^S\|_2). \tag{21}$$

The encoder Jacobian determines how these feature gradients propagate to model parameters. Pooled norms and gradient norms are therefore useful diagnostics alongside the loss itself. Together with within-formula margins, they connect the strength of adaptation to the molecular distinctions represented by the teacher.

Teacher diversity supplies the distinctions on which contrastive learning depends. If all keys are identical, every query has loss $\log M_i$ and zero feature gradient. If r keys coincide with the positive, the loss has a floor of at least $\log r$. Known duplicate records should consequently be removed from each other's negative sets. Undocumented repeated structures can likewise create chemically equivalent competitors. The implemented objective is naturally interpreted as paired-record identification; its discrimination should be assessed with the available identity metadata. A mutual-information interpretation applies under the joint-positive and independent-marginal-negative sampling assumptions used in contrastive predictive coding (van den Oord et al., 2018).

C.4 Decoder JS: range and gradient

Decoder alignment transfers the distribution of support over possible continuations of a partial structure. Its discrepancy has a common scale regardless of vocabulary size, which helps compare positions with different levels of ambiguity.

Proposition 3 (Bounded decoder discrepancy)**.** *With natural logarithms and normalized distributions,* $0 \le \mathrm{JS}(q, p) \le \log 2$*, and equality at zero holds exactly when* $p = q$*.*

Proof. Nonnegativity follows from KL. The mixture $m = (p + q)/2$ satisfies $m_j \ge p_j/2$ and $m_j \ge q_j/2$. Each KL term is therefore at most $\log 2$. Zero requires equality of both distributions with the mixture. □

The same bounds hold after averaging across valid positions and molecules. They characterize the decoder term; encoder InfoNCE has the batch-dependent range above.

The correction at a token depends on its relative support in the two distributions. For a fixed teacher and positive student probabilities, differentiating JS gives

$$a_j = \frac{\partial\, \mathrm{JS}(q,p)}{\partial p_j} = \tfrac{1}{2}\log\frac{2p_j}{q_j + p_j}, \qquad \frac{\partial\, \mathrm{JS}(q,p)}{\partial g_j} = \frac{p_j}{\tau_D}\left(a_j - \sum_k p_k a_k\right). \tag{22}$$

Differentiating both KL terms through their shared mixture yields the first expression; applying the softmax Jacobian yields the second. The logit gradient centers each token's relative support a_j against its probability-weighted mean. The resulting update redistributes probability among competing continuations while respecting normalization, and its temperature controls the scale of that redistribution. Extremely small student probabilities can attenuate the logit gradient through the factor p_j, making token-level recovery worth measuring explicitly.

C.5 Local agreement and sequence generation

A conditional distribution matters to generation through the decisions it changes. Total variation bounds the change in the probability of any vocabulary event. With $\mathrm{TV}(p, q) = \frac{1}{2}\|p - q\|_1$, Pinsker's inequality applied to p, m and q, m gives

$$\mathrm{TV}(p, q) \le \sqrt{2\,\mathrm{JS}(p, q)}. \tag{23}$$

If the teacher's leading token exceeds its runner-up by margin γ, then $\mathrm{JS}(p, q) < \gamma^2/8$ is sufficient to preserve that leading token. Each probability changes by at most TV, so the margin can decrease by at most twice TV. Small-margin decisions require closer agreement, which makes them useful locations for examining ambiguous atom, branch, or ring-closure choices. The inequality links a local alignment value to the stability of the corresponding token decision.

Free-running generation also depends on the prefixes that are visited. Agreement at each prefix of a teacher greedy trajectory yields the same greedy sequence by induction under common termination rules. SMART uses a beam-selected prefix sequence, and its decoder objective averages over those particular contexts. We therefore evaluate complete structures directly, alongside conditional-distribution agreement.

The same molecule can be written with different SMILES strings, so structural agreement should group equivalent serializations. For full sequence distributions P, Q and the deterministic graph map c, including an invalid outcome, the log-sum inequality gives

$$\mathrm{JS}(c_{\#}P, c_{\#}Q) \le \mathrm{JS}(P, Q). \tag{24}$$

The graph map pools probability assigned to equivalent strings, so agreement at the sequence-distribution level carries through to structure. The implemented loss samples selected teacher contexts; full-sequence comparisons additionally account for the probability of reaching those contexts. This motivates evaluating both conditional agreement and complete generated graphs.

C.6 Interaction between the two encoder gradients

Representation recognition and structural generation share encoder parameters. Their local interaction can be measured through the angle between $g_C = \nabla_{\phi_E}\mathcal{L}_E^{\mathrm{NCE}}$ and $g_D = \nabla_{\phi_E}\mathcal{L}_D^{\mathrm{JS}}$:

$$\|g_C + g_D\|^2 = \|g_C\|^2 + \|g_D\|^2 + 2\langle g_C, g_D\rangle. \tag{25}$$

The inner product measures local compatibility of the two updates. It motivates recording gradient cosine and norm ratio in matched experiments. The method combines two complementary learning tasks; whether their gradients cooperate at a particular point is an empirical property of the model and batch. This distinction connects the sensitivity pilot to a precise mechanism study.

C.7 What mean-KL routing learns

The router summarizes many source episodes through a compressed input representation. Episodes that share the same features contribute together to the learned action preference. For an unrestricted conditional router, minimizing population mean KL yields

$$p^*(a \mid z) = \mathbb{E}[q_i(a) \mid Z = z]. \tag{26}$$

Condition on z, collect all terms independent of p, and write the remaining cross-entropy with the conditional mean target. This establishes the minimizer. When reward is determined by z, a realizable optimum preserves its action ordering. With compressed features, averaging soft probabilities and averaging rewards can favor different actions.

For example, let action B have reward zero and action A have reward one with probability .15 and $-.1$ otherwise. A has positive expected reward .065. At oracle temperature .20, however, its mean soft probability is

$$.15(1 + e^{-5})^{-1} + .85(1 + e^{.5})^{-1} \approx .4699056. \tag{27}$$

The example shows why validation utility complements distribution-fitting loss: the deployed decision is an action, and its value depends on the structural outcomes it produces.

At fixed z, the four unary, six pair, and four size terms have a 12-dimensional score space. To see this, require a coefficient vector to produce zero on every action. Singleton equations give $u_m = -b_1$, pair equations give $v_{mn} = 2b_1 - b_2$, triples require $b_3 = 3b_2 - 3b_1$, and Full requires $b_4 = 6b_2 - 8b_1$. Two free coefficients remain among fourteen columns, giving rank twelve. Removing the common softmax shift leaves eleven effective dimensions, compared with fourteen for an unrestricted distribution over fifteen actions. The sharing structure trades unrestricted interactions for parameter reuse across related views.

C.8 A conditional view-utility transfer bound

Source perturbations are useful when they preserve which views support good structural judgments in the target domain. Denote source and target conditional utilities on common legal actions by $\bar{R}_s(z, a)$ and $\bar{R}_t(z, a)$. If their maximum discrepancy is $\delta(z)$ and the selected action $\hat{a}$ has source regret at most $\epsilon(z)$, adding and subtracting source utilities gives

$$\max_a \bar{R}_t(z, a) - \bar{R}_t(z, \hat{a}) \le 2\delta(z) + \epsilon(z). \tag{28}$$

The two δ terms account for the target-optimal action and the selected action; the remaining difference is bounded by source regret. The assumptions identify the quantity that augmentation should preserve: decision-relevant action utility. Matching spectral marginals is useful when it helps preserve this relationship. Measuring both terms would distinguish an imperfect selector from a shift in what the generator can infer from each view.

C.9 Temporal selection, coverage, and cost

Temporal agreement changes the population that supplies supervision. Its quality therefore depends on both the errors retained and the number of molecules accepted. For acceptance event A, wrong-structure event W, and realized schedule S,

$$\mathbb{P}(W \mid A, \mathcal{E} = S) = \frac{\mathbb{P}(W \cap A \mid \mathcal{E} = S)}{\mathbb{P}(A \mid \mathcal{E} = S)}. \tag{29}$$

Thus accepted-label precision must be interpreted together with coverage. Checkpoint errors that are independent after conditioning on the selection context permit a product bound on the numerator. Along a single training trajectory, shared structural preferences can instead preserve systematic errors: if a fraction α always predicts the same wrong valid isomer and the rest always predicts the truth, all records pass consensus and the accepted error remains α for every checkpoint count. The relevant chemical challenge is repeated selection of the same wrong isomer. Accepted-set precision and isomer-level error persistence measure that challenge directly.

The interval limits also specify how much of the trajectory is observed. With first and final epochs retained and each realized gap at most $n_{\max}$,

$$1 + \left\lceil \frac{E_1 - 1}{n_{\max}} \right\rceil \le K \le E_1. \tag{30}$$

The lower bound follows by summing $K - 1$ gaps to $E_1 - 1$; the upper bound allows at most one checkpoint per epoch. For N target records and B probe records, caching selected-epoch probe outputs gives $E_1 B + K(N - B)$ greedy record decodes. Sequence length and hardware still determine runtime.

C.10 Cross-view learning and Full prediction

Partial spectra can support Full prediction when the shared parameters learn compatible structural decisions from both. This relation is captured locally by gradient agreement. For an L-smooth Full loss L_F and a sampled-view gradient g_A, the descent lemma gives

$$L_F(\phi - \eta g_A) - L_F(\phi) \le -\eta \langle \nabla L_F, g_A \rangle + \tfrac{1}{2} L \eta^2 \|g_A\|^2. \tag{31}$$

Positive gradient agreement is one mechanism through which partial-view training can benefit Full. The relation motivates direct Full evaluation during controlled studies, together with comparisons among Full-only, include-Full, and exclude-Full refinement. Those experiments measure whether

different spectral pathways reinforce the same structural decisions through the shared architecture and its Stage-1 initialization.

D Acquisition mechanisms and source episodes

D.1 Calibration populations and the purpose of perturbation

The documented calibration uses 794,403 source rows, 2,669 real train/validation rows, and 1,059 same-molecule pairs per modality. Same-molecule pairing helps separate acquisition and representation changes from changes in molecular identity. The router then learns from how these disturbances alter the usefulness of each input view. Chemotion and nmrXiv complement this analysis through audits of representation, lengths, and input boundaries. Their observations are kept separate from the SDBS paired residuals.

The three severity values .20, .55, and .90 index weak, medium, and strong profiles. They coordinate training exposure across different disturbance mechanisms. Parameters come from project residual summaries and the implemented profiles; they retain modality-specific units. The connection to the downstream task is evaluated through the action rankings produced by the frozen generator.

D.2 Proton NMR: changing the granularity of observed peaks

The paired proton summaries report recall .7465, precision .6180, missing fraction .2535, extra/source ratio .6435, displacement summary .0388, and split factor 4.36. This combination motivates both missing lines and an expansion of model-visible entries. A multiplet represented as one group in one source may appear as several local lines in another representation. The profile therefore combines retention, local splitting, jitter, and additional entries while preserving the source molecular target.

The reported target peak-count means are 14, 20, and 34, with log-scale standard deviations .30, .42, and .50. If a lognormal sampler targets arithmetic mean m, its location must be $\mu = \log m - \sigma^2/2$, because $\mathbb{E}N = \exp(\mu + \sigma^2/2)$. Numerical values are retokenized through the generator's legal input range. Counts and shifts should be summarized after tokenization as well as before it, since quantization and truncation determine the actual evidence reaching the model.

D.3 Carbon NMR: shared referencing and local shifts

Carbon residual summaries report recall .8977, precision .9192, missing fraction .1023, extra-line fraction .0820, and median common shift $-.062$ ppm. A common reference offset moves the spectrum while preserving peak spacings; local variation changes the relationships between individual environments. The profile expresses these two effects separately:

$$\tilde{\delta}_j = \delta_j + \Delta + \epsilon_j, \quad \Delta \sim \mathcal{N}(0, \sigma_r^2), \quad \epsilon_j \sim \mathcal{N}(0, \sigma_l^2). \tag{32}$$

Independent local components give cross-peak residual covariance σ_r^2 and marginal variance $\sigma_r^2+\sigma_l^2$. A common shift thus preserves relative spacings more strongly than independent noise of the same marginal variance.

Across the three tiers, reference scales are .06/.18/.20, local jitter scales .10/.18/.35, and extra-line ratios .02/.08/.14. Retention ranges are .95–1, .83–1, and .75–.90. These parameters create varying reliability rather than treating carbon as universally clean. Whether carbon remains the preferred view is determined by structural reward.

D.4 Mass spectra: useful peaks within a background tail

Mass-spectral summaries combine recall .9301 with precision .1996, an extra/source ratio of 4.90, and a real median peak count of 82. This motivates retaining molecule-related evidence while varying the number and intensity of background peaks. The tiers use keep probabilities .95/.88/.75, target count means 56/82/122, count standard deviations 9/18/31, intensity exponents 2.4/3.2/4.0, and mass-position jitter .20/.50/.90 in the implementation's coordinates.

Intensity affects whether a fragment stands out from the tail of weak peaks. To vary this prominence for relative intensity I in the normalized range, we use

$$I' = 100 \left[\frac{\max(I, .1)}{100} \right]^{\gamma} \tag{33}$$

suppresses weak peaks for $\gamma > 1$. Its logarithmic slope is γ above the clipping floor. Background peak count and background influence should therefore be measured separately: a large number of weak entries may have a different effect from a few prominent peaks. Acquisition-specific ionization differences motivate further transfer tests beyond this empirical profile.

D.5 Infrared spectra: smooth changes in shape and scale

The IR summaries report standard-deviation ratio 2.075, roughness ratio .450, and median best-orientation correlation .163. Peak location, width, and baseline affect how a vibrational pattern reaches the model in different ways. We represent these effects with a functional perturbation:

$$\tilde{I}(\nu) = g(\nu)[K_\sigma * I](\nu + \Delta(\nu)) + b_0 + b_1\nu + b_2\nu^2 + r(\nu) + \eta(\nu). \tag{34}$$

Broadening, coordinate warp, regional gain, baseline, broad residual bands, and local variation affect different aspects of the signal. Their combination models core acquisition disturbances while retaining model-visible patch coverage. This lets a source molecule keep its target identity as the presentation of its vibrational evidence changes.

The tier standard-deviation ratios are 1.60/2.08/2.78, roughness ratios .58/.45/.33, gain scales .18/.35/.60, and warp parameter pairs $(.020, .008)$, $(.040, .018)$, $(.070, .030)$. Warp coordinates follow the implementation. The inherited backbone's 1,800-value interface and a separately audited 400-point Chemotion representation require an explicit preprocessing map when used together.

D.6 Evaluating the usefulness of empirical episodes

A useful audit follows the transformation from measurement disturbance to tokenized observation and finally to action utility. Peak-matching tolerances, intensity thresholds, axis orientation, missing-value rules, and molecular pairing determine the residual summaries. Length caps and vocabulary bins determine which changes remain visible to the generator. The reward comparison then shows whether those changes alter structural judgment.

This chain suggests controlled studies that vary one mechanism at a time: common versus local NMR shifts, background count versus intensity, and IR baseline versus warp. Such studies can explain which domain differences matter to view selection and which are absorbed by preprocessing. They also provide a direct route from chemical intuition to measurable router behavior.

E Implementation and reproducibility

E.1 Backbone and information available during adaptation

The molecular backbone uses a shared encoder–decoder with width 512, six encoder layers, six decoder layers, and eight attention heads. Its infrared input contains 1,800 values grouped into 24 patches of 75. The recorded vocabularies contain 132 formula, 1,113 proton-NMR, 2,102 carbon-NMR, 3,143 MS, and 44 SMILES tokens. The architecture and training settings are summarized in Tables 8–9.

Source molecular structures supervise the router's action rewards. During target adaptation, the model accesses observed spectra, formula, stable record identities, and its own predictions. Formula is retained in every routed and dropout view and also supplies the composition check for consensus. True target SMILES belongs to the offline evaluator. The adaptation loader contains observed inputs and teacher outputs; the evaluator reads structural labels separately.

The router record uses 43,978 episodes: 35,183 for training and 8,795 for validation. Weak, medium, and strong counts are 2,048, 30,819, and 11,111. Training uses 100 epochs, batch 256, AdamW,

learning rate 3×10^{-4}, and seed 3247. Source validation utility selects the checkpoint. The curriculum reweights perturbation tiers; payload counts describe the stored examples, while sampling weights determine their effective exposure.

E.2 Feature summaries and legal actions

The `source_reference_free_v1` schema includes each modality's pooled cosine to the formula representation and $\log(1+\text{token count})$; final-layer and across-layer mean cosines for modality pairs; available-modality fraction and log total token count; and an availability mask. Features use the frozen input encoder and observed inputs. The schema supplies the unary, pair, and global parts of the structured scorer without a stored source-reference bank.

For m available spectra, $2^m - 1$ nonempty actions are legal. Unavailable actions receive masked logits before normalization. Formula is outside the four-bit action vector. A record with no spectra is sent to a separately labeled formula-only source prediction and receives no SMART update. This keeps the defined action space valid while retaining the record in evaluation accounting.

E.3 Reward components and reductions

The reward combines exact retrieval with sequence compatibility and candidate validity. The accompanying implementation specifies these two continuous components as

$$\begin{aligned} \mathrm{NLL}_i^{\mathrm{ref}}(a) &= -\frac{1}{T_i}\sum_t \log p_{\theta_0}(y_{i,t} \mid y_{i,<t}, \tilde{x}_i^a), \\ \mathrm{Invalid}_i^{\mathrm{ref}}(a) &= \frac{1}{5}\sum_{j=1}^{5} \mathbf{1}[\text{slot } j \text{ absent or invalid}]. \end{aligned} \tag{35}$$

The fixed source target uses the same tokenization for every action. The likelihood average includes EOS and excludes BOS and padding. Likelihood uses untempered probabilities. Five ranked raw beam slots are retained, including duplicate structures. Top-1 compares the first canonical candidate with the source truth, and Top-5 records a match in any valid slot. Each reward payload records its reduction settings alongside the action scores.

E.4 Stage-1 representations and eligible negatives

A routed view and a Full view contain different numbers of tokens but describe the same molecule. Pooling places their encoder states in a common feature space. With states H_i^S, H_i^T and valid-position indicators $m_{i,l}^S, m_{i,l}^T$, the reference uses

$$u_i^b = \frac{\sum_l m_{i,l}^b H_{i,l}^b}{\sum_l m_{i,l}^b}, \qquad b \in \{S, T\}. \tag{36}$$

Pooling permits different view lengths while keeping a common feature width. Normalize with a numerical norm floor, and detach the teacher representation. Padding is excluded; all valid encoder positions, including formula positions, contribute in this reference. Modality-balanced pooling is a separate ablation because it changes the relative influence of long and short observations.

Each student's positive is its own routed teacher. Other records provide in-batch negatives, except known duplicate identities established from metadata available before adaptation. Formula equality alone is insufficient to remove a negative because isomers share formulas. The denominator always contains the positive. A singleton eligible set has zero contrastive loss, while decoder JS still supplies supervision. Log eligible-set size to make such changes visible across batches.

E.5 Soft decoder supervision and numerical evaluation

The saved teacher sequence fixes raw prefixes and valid target positions. Teacher and student use the same tokenizer and vocabulary order. Full-vocabulary teacher logits are retained before beam pruning so that alternatives remain available for soft supervision. BOS supplies context, EOS is a

valid target when present, and padding is excluded. We average token losses within each molecule and then across molecules.

Soft alignment depends on preserving small probabilities as well as the leading token. Log probabilities are therefore computed in float32 even when model inference uses mixed precision. For $\ell_q = \log q$ and $\ell_p = \log p$, the stable mixture computation is

$$\log m = \operatorname{logaddexp}(\ell_q, \ell_p) - \log 2, \quad \mathrm{JS}(q,p) = \tfrac{1}{2}\sum_v q_v(\ell_{q,v} - \log m_v) + \tfrac{1}{2}\sum_v p_v(\ell_{p,v} - \log m_v). \tag{37}$$

The teacher runs in evaluation mode without gradients. The student updates all permitted encoder–decoder parameters using the sum in Eq. (12). The reference uses unit loss coefficients and $\tau_C = .10$, $\tau_D = 1$; temperatures act solely inside the stated probability computations. These temperatures and coefficients define the stated hybrid-loss configuration.

The SDBS Stage-1 configuration uses 12 epochs, batch 16, learning rate 10^{-5}, AdamW, bf16, gradient clipping .8, four workers, and seed 3247. The student is optimized with the hybrid objective in Eq. (12). A cache of teacher features, logits, prefixes, actions, and masks can avoid repeated teacher inference. Cache validity depends on model, tokenizer, record identity, and decoding configuration.

E.6 Dynamic interval update and exact replay

The controller starts at $n_1 = 3$, uses $n_{\min} = 1$, $n_{\max} = 4$, $\delta_{\text{low}} = .02$, $\delta_{\text{high}} = .10$, $\gamma = .80$, and patience $h = 2$. At most 256 fixed probe records are selected by deterministic identity hashing. Table 10 collects the controller settings and their roles.

After each epoch, Full greedy probe predictions determine drift and formula-compatible validity. The first epoch is selected unconditionally. For each later epoch, increase the stability counter only when $d_e \le \delta_{\text{low}}$ and $g_e \ge \gamma$; otherwise reset it. Apply Eq. (14) in the displayed order. Reset the counter after low feasibility, high drift, or an interval increase. Select the current epoch if its distance from the last selected epoch reaches the updated interval, or if it is the final epoch. Consequently, the state n_e and the realized inter-checkpoint gap are recorded separately.

An invalid prediction receives a fixed sentinel for drift calculation and fails the feasibility indicator. This combination distinguishes repeated invalid outputs from useful structural stability. All epoch weights may be saved, while full-cohort decoding is required only for selected epochs. Store probe identities, canonical signatures, feasibility flags, interval before and after update, stability counter, selection reason, and the resulting epoch set. These records permit exact controller replay independently of training.

E.7 Consensus, target choice, and Stage-2 views

Join selected-checkpoint predictions by record identity. Accept a record only when every selected prediction is RDKit-valid, matches its observed formula, and has the same declared canonical structure. Specify stereochemistry and fragment handling in the evaluator. With ordered selected epochs $s_1, \ldots, s_K$, use the raw target from $s_{\lfloor K/2 \rfloor + 1}$. Canonicalization compares structures; the retained raw token sequence supplies training supervision.

Initialize Stage 2 from the final Stage-1 weights with an explicitly recorded optimizer reset. If no targets pass, return the Stage-1 model and record zero coverage. For four available modalities, uniform retained size from one to three gives probability $1/12$ for each singleton or triple, and $1/18$ for each pair. Each modality has marginal retention probability $1/2$. The reference shares one sampled view per batch and derives randomness from the epoch seed.

For $m \ge 2$ available modalities, the natural extension samples nonempty proper subsets by first choosing retained size uniformly from $1, \ldots, m-1$. A single available spectrum has no such subset and is skipped in strict exclude-Full refinement. All records, including rejected or skipped ones, remain in final Full-view evaluation.

E.8 End-to-end procedure

1. Build same-molecule source perturbation episodes; evaluate all legal views with the frozen generator and fit the structured router to reward-derived soft targets.
2. Freeze the selected router and source teacher. Save one target action and beam-10 teacher prefix sequence per record.
3. Train the Full student with pooled-feature InfoNCE plus shared-prefix decoder JS. Probe each epoch and update the checkpoint interval.
4. Decode Full inputs at selected checkpoints and form the exact, formula-compatible consensus set. Keep each accepted raw target and its source epoch.
5. Refine the final Stage-1 student on fixed accepted targets under non-Full views. Evaluate the final model with the declared beam, canonicalization, and complete cohort denominator.

The end-to-end implementation applies the hybrid objective directly to the shared molecular encoder–decoder. Fixed routes and teacher outputs feed Stage 1; selected checkpoint predictions define the targets used by Stage 2. The accompanying method code specifies the loss reductions and scheduling logic, and numerical checks verify their gradients and state transitions. Algorithms 1–2 summarize the full data flow.

F Component evaluation and diagnostic protocols

F.1 Disentangling routing, alignment, and temporal selection

A matched study should first compare unadapted Full, fixed-carbon inference, router-only inference, fixed-carbon-teacher adaptation, and routed-teacher adaptation. This progression isolates view selection from the benefit of training. Router utility should be reported both overall and for records on which its selected action differs from carbon. Action frequencies, source validation utility, and offline target regret help explain the outcome.

Next, hold teacher actions and prefix sequences fixed while comparing hard CE, decoder KL, decoder JS, encoder InfoNCE, and InfoNCE plus JS. Match optimizer, epochs, batch order, and decoding. Report per-molecule prediction transitions: teacher-wrong to student-correct, teacher-correct to student-wrong, both correct, and both wrong. Their difference recovers the net Top-1 gain and shows whether the student improves through useful disagreement.

For temporal selection, compare every epoch, fixed intervals, a uniform schedule with matched checkpoint count, and the dynamic controller. Coverage and accepted-label precision describe the selected targets; all-cohort accuracy describes deployment. Since retaining epoch one can reject later improvements, a predeclared warm-up-start alternative is also informative. Schedule settings should be fixed without selecting them on target structural labels.

F.2 Representation and probability diagnostics

Measure positive cosine, nearest-negative cosine, contrastive margin, effective negative count, and pooled-feature norm. Stratify these quantities by molecular size and formula multiplicity. A formula-dominated representation could separate different compositions yet struggle among isomers, which makes within-formula analysis especially useful.

For the decoder, inspect JS across sequence positions and structural branches, together with teacher entropy. Evaluate calibration on free-running candidate correctness using a declared confidence definition. Average shared-prefix agreement and candidate-level calibration answer different questions and should be reported separately. Record shared-encoder gradient cosine and norm ratio to connect the coefficient study to optimization behavior.

F.3 Connecting spectral interventions to structural decisions

A direct chemical mechanism study should perturb acquisition detail while preserving the measured molecule. For each record, change MS background intensity, apply a common NMR reference offset, or vary a smooth IR baseline within a predeclared range. Measure the corresponding route changes,

matched-representation margin, shared-prefix JS, and free-running structural prediction. A useful adaptation should preserve task-relevant distinctions across these nuisance changes, and the same record should serve as its own control.

A second intervention should alter informative evidence, such as masking a characteristic IR band or a selected carbon environment. Compare this response with the acquisition perturbations at matched input scale. The contrast distinguishes selective robustness from a model that simply ignores a modality. Functional-group and connectivity error slices then reveal whether the affected evidence supports the expected type of structural decision.

These interventions should use only observed inputs to construct perturbations; reference structures enter the offline analysis of errors. Predeclared perturbation ranges, molecule-group bootstrap intervals, and chemically interpretable slices make the comparison reproducible. The endpoint is a measured relationship between a spectral change and a structural judgment.

F.4 Statistical comparisons and total cost

Use paired molecular outcomes for confidence intervals on accuracy differences. Resample molecular identity groups when multiple observations share a molecule. Repeat adaptation across declared seeds and report both between-run variation and paired sampling uncertainty. The source tables provide mean-plus-dispersion summaries; paired sampling intervals and between-seed variation are distinct statistical quantities.

The supplied cost panels estimate training-only FLOPs and timing proxies. The cost model in Appendix H includes teacher computation, student updates, probe decoding, selected-state decoding, and storage. Source router construction and target adaptation have separate costs; every probe call contributes even when no full-cohort checkpoint is selected. Wall time and peak memory additionally depend on hardware and sequence lengths. This decomposition identifies the computational trade-off between observation frequency, accepted-label quality, and retrieval performance.

G Experimental versions and figure provenance

The asset manifest records original identifiers, locations, and hashes. All PDF artwork and table values are preserved; the attention overview and layerwise changes form panels (a)/(b).

G.1 SDBS endpoints across recorded versions

The main SDBS table evaluates the integrated InfoNCE–JS pipeline over all 3,669 molecules and reports Full 55.66/76.51/80.97%; hybrid-objective and motivation graphics use 55.66/76.51/80.57%. The stage-comparison plot rounds 80.97% to 81.0%. Table-based Top-10 gains are 2.97 points over Stage 1 and 5.50 over the teacher.

Motivation-figure repair/damage bars show marginal accuracy differences; per-record transitions measure corrected and newly damaged predictions. Depth attention uses a pseudo-CE, interval-3 checkpoint with Full 55.41/76.12/80.54%. Average/layerwise attention uses a different aggregation transcribed from figure summaries; changes are interpreted within each figure.

G.2 Display-adjusted interval and pilot figures

The interval graphic anchors interval three to 2,042/2,807/2,956 hits among 3,669 records and places interval four five hits lower, changing the original ranking. Text comparisons use the original sweep; the caption identifies the transformation.

The KL/InfoNCE pilot uses 128 records, one epoch, seed 3247, batch 16, learning rate 10^{-5}, beam ten, and carbon routes. Display offsets of 42.37875/55.41625/55.57000 points shift raw Top-1/5/10 maxima 13.28125/21.09375/25.0% to 55.66/76.51/80.57%; losses remain raw. This coefficient study is separate from the full-cohort hybrid comparison.

G.3 Router, structure metrics, and Chemotion

Router panels use 8,795 source-validation episodes and 3,669 target records; target bin sizes are 63/246/720/987/842/505/265/41. Formula enters selection; target structures enter offline scoring. Bin counts contextualize uncertainty in probability–success relationships.

Structure-quality tables pair 500 records with sampling seed 20260923. These metrics use a distinct evaluation cohort and are compared within that cohort.

Chemotion uses the strict v8 cohort of 1,995 records: Stage 1 epoch 19, Stage 2 up to epoch 20, and epoch 18 selected by minimum pseudo-CE. Mean-plus-dispersion cells retain the values reported in the source tables.

The encoder-depth panels use separate vertical scales, so changes are compared within each panel.

H Algorithm, configuration, and computational cost

The summaries below connect the mathematical objectives to the execution order and the configuration of the molecular generator. Algorithms 1 and 2 distinguish source-supervised router construction from target adaptation using observed spectra and formulas. Tables 8–10 collect the architecture, training, and scheduling settings.

Algorithm 1: Structural-utility router training

Input: labeled source molecules $\mathcal{S}$, frozen generator θ_0, perturbation profiles, legal action vocabulary.
Split source molecules into training and validation groups by molecular identity.
For each sampled molecule and severity profile, generate $\tilde{x}_i = \mathcal{T}_\omega(x_i)$ while retaining y_i.
Extract observed-input features z_i; enumerate every nonempty subset of available spectra, retaining formula.
For each action, decode five candidates with θ_0 and compute structural reward from Eq. (4).
Store $q_i = \mathrm{softmax}(R_i/\tau_o)$ over legal actions; optimize the router using mean KL.
Select the router by source-validation utility and freeze ρ. **Output:** the fixed teacher-view policy.

Algorithm 2: Hybrid alignment, temporal selection, and cross-view refinement

Input: unlabeled target inputs $\mathcal{U}$, frozen ρ, θ_0, budgets E_1, E_2, controller settings.
Choose a_i^* once per record. Generate a beam-10 teacher sequence and obtain teacher keys and shared-prefix vocabulary probabilities.
Initialize $\phi \leftarrow \theta_0$; select the fixed hash-based probe $\mathcal{B}$; set $\mathcal{E} = \varnothing$, $n = 3$, $r = 0$.
For each epoch $e = 1, \ldots, E_1$, update the Full student on minibatches using $\mathcal{L}_E^{\mathrm{NCE}} + \mathcal{L}_D^{\mathrm{JS}}$.
Greedily decode the fixed probe. For $e > 1$, compute (d_e, g_e), update the stable-run counter, then apply Eq. (14).
Select e if $e = 1$, $e = E_1$, or $e - s_{\text{last}} \geq n_e$; append it to $\mathcal{E}$ and preserve its weights. Reset the counter after low feasibility, high drift, or an interval increase.
After E_1, decode selected states on all records. Accept only identical valid canonical structures with matching formulas; retain the raw middle-checkpoint sequence.
If the accepted set is empty, return the final Stage-1 model. Otherwise initialize Stage 2 from it and reset optimizer state.
For E_2 refinement epochs, sample proper nonempty spectral subsets and optimize the fixed accepted targets with Eq. (16).
Output: adapted generator, selected epochs, consensus coverage, and complete-cohort predictions under the fixed evaluation protocol.

H.1 Time and storage requirements

Let N_s be the number of source episodes, N the target cohort size, B the training minibatch size, B_p the probe size, P the generator parameter count, n the encoder input length, and T the output

Table 8: Molecular generator and routing architecture.

Component	Setting	Role
Shared hidden width	$d = 512$	Common teacher/student coordinates
Encoder / decoder depth	$6/6$	Shared multimodal sequence model
Attention heads	8	Per-layer attention
IR interface	$1800 = 24 \times 75$	24 patches, 75 values each
Formula / H / C vocabularies	$132/1113/2102$	Modality-specific input tokens
MS / SMILES vocabularies	$3143/44$	Input and generated sequences
Router action space	15	All nonempty subsets of four spectra
Action-score components	$4 + 6 + 4$	Unary, pair, and size scores
Adapted parameters	Student encoder/decoder	Teacher and router remain fixed

Table 9: Training and evaluation settings. E_2 denotes the cohort-specific refinement budget.

Setting	Value	Scope
Source episodes	43,978	35,183 train; 8,795 validation
Router optimizer	AdamW	100 epochs; minibatch 256
Router learning rate	3×10^{-4}	Seed 3247
Oracle / router temperature	$.20/1$	Soft action targets and prediction
Reward / teacher beam	$5/10$	Source reward / target teacher
SDBS Stage-1 budget	$E_1 = 12$	3,669 unlabeled molecules
Student optimizer	AdamW, 10^{-5}	Minibatch 16; seed 3247
Precision / clipping	bf16 / .8	Four loading workers
Alignment coefficients	$1/1$	Encoder InfoNCE / decoder JS
Alignment temperatures	$.10/1$	τ_C/τ_D
Refinement view size	$1, 2, 3$ uniformly	Uniform subset at each size; formula retained
Chemotion budgets	$E_1 = 19, E_2 \leq 20$	1,995 records; epoch 18 selected by pseudo-CE
Evaluation beam / cutoffs	$10/\{1, 5, 10\}$	Full-cohort structural exact match

Table 10: Dynamic checkpoint controller settings and acceptance rules.

Quantity	Value	Meaning
Fixed probe size	$\lvert\mathcal{B}\rvert \leq 256$	Deterministic record-identity hash
Initial / bounded interval	$3/[1, 4]$	$n_1/[n_{\min}, n_{\max}]$
Low / high drift	$.02/.10$	$\delta_{\text{low}}/\delta_{\text{high}}$
Feasibility threshold	.80	γ: valid and formula-compatible
Stable-probe patience	$h = 2$	Consecutive low-drift feasible probes
Mandatory states	$1, E_1$	First and final Stage-1 epochs
Consensus agreement	All selected states	Canonical structure, validity, formula
Target source	$s_{\lfloor K/2 \rfloor + 1}$	Raw sequence; canonicalization only for comparison

length. For encoder and decoder depths L_e, L_d, conventional dense-attention forward computation has order

$$C_f(n, T) = O\big[L_e(nd^2 + n^2 d) + L_d(Td^2 + T^2 d + Tnd) + Td|V|\big]. \tag{38}$$

This expression assumes feed-forward width proportional to d. A cached beam decoder has cost $G_b(n, T)$ of the order of encoder computation plus b decoder paths; search and vocabulary selection add implementation-dependent overhead. Backpropagation changes the constant factor, so we write C_{tr} for one student training example and C_T for its frozen-teacher forward cost.

Table 11 separates source preparation, target optimization, and structural selection. The encoder InfoNCE similarity matrix costs $O(B^2 d)$ per minibatch; decoder JS adds $O(BT|V|)$. Both are evaluated on model outputs already computed for alignment. The dynamic controller trades complete-cohort decodes against fixed-probe observations, rather than eliminating the probe cost.

Table 11: Asymptotic costs by stage; $N_P = |\mathcal{P}|$ and $K = |\mathcal{E}|$.

Stage	Time	Additional storage
Source reward construction	$O(N_s\, 15\, G_5)$	$O(N_s(\dim z + 15))$ scores/features
Teacher preparation	$O(N(G_{10} + C_T))$	$O(N(d+T\|V\|))$ cached targets
Stage-1 optimization	$O(E_1 N C_{\text{tr}})$	$O(P)$ weights/optimizer plus activations
Hybrid-loss arithmetic	$O(E_1 N(Bd + T\|V\|))$	$O(B^2 + BT\|V\|)$ temporary arrays
Probe and selected decoding	$O((E_1 B_p + K(N - B_p))G_1)$	$O(NT)$ streamed consensus state
Stage-2 refinement	$O(E_2 N_P C_{\text{tr}})$	$O(N_P T)$ accepted targets
Checkpoint files	$O(KP)$ writes	$O(KP)$ disk; $O(E_1 P)$ if all epochs kept

The decoding count reuses probe predictions at selected epochs. If teacher outputs are recomputed during training, add $O(E_1 N C_T)$ instead of storing their full cache. Streaming consensus keeps the current canonical signature, validity flag, and chosen raw target per record; retaining every checkpoint prediction uses $O(KNT)$ disk space. Actual runtime also depends on view lengths, beam size, mixed precision, and hardware. These symbolic costs describe the algorithm independently of the training-only timing proxies plotted in the empirical analysis.

Chemical parsing and canonicalization add $O(KNC_{\text{chem}})$ to consensus evaluation, where C_{chem} is the per-structure processing cost.