Skepthical review: *Dynamic Multiscale Graph Analysis Reveals Structural Signatures of Peptide Aggregate Stability and Splitting*

Summary

The manuscript proposes a dynamic multiscale graph-theoretic pipeline to analyze peptide self-assembly from MD, demonstrated on a 500 ns trajectory of 30 KYFIL pentapeptides (Sec. 2.1–2.2). At each frame, a coarse-grained (CG) peptide–peptide contact graph is constructed (nodes=peptides; edges=heavy-atom contacts), and for each CG aggregate a fine-grained (FG) residue contact graph is built (nodes=residues within the aggregate; edges=heavy-atom contacts) (Sec. 2.3). Over an "equilibrium" window (primarily 100–500 ns; Sec. 2.2.2, 3.1), the authors compute time series and distributions of graph metrics including density, clustering, centralities, and Laplacian spectral quantities (Fiedler value) (Sec. 2.4, 3.3). They introduce composite order parameters combining largest-aggregate size with FG density to obtain more temporally stable descriptors (Sec. 2.6, 3.4), and they track aggregates across frames via overlap/Jaccard-based matching to define events (formation/dissolution/splitting/merging) and relate pre-event graph properties to splitting propensity and aggregate longevity (Sec. 2.7, 3.5). The multiscale framing is promising and timely, but several core methodological choices (contact definition, FG edge types, event tracking) and statistical practices are currently underspecified or internally inconsistent, limiting interpretability, reproducibility, and confidence in the reported "splitting signatures" and order-parameter claims.

Strengths

- Clear multiscale conceptualization linking peptide-level aggregates (CG connected components) to within-aggregate residue interaction structure (FG graphs) (Sec. 2.3).
- Broad, well-motivated set of graph diagnostics including spectral connectivity (Fiedler value), which can capture bottlenecks/fragmentation beyond density alone (Sec. 2.4, 3.3).
- Frame-resolved analysis and event-based framing (formation/dissolution/splitting/merging) are a natural fit to MD trajectories and help connect structure to dynamics (Sec. 2.7, 3.5).
- Composite order-parameter idea is potentially useful: combining an extensive descriptor (aggregate size) with an intensive internal descriptor (FG density) is a sensible route to reduce apparent noise (Sec. 2.6, 3.4).
- Overall manuscript organization is logical (simulation → equilibration → graph construction → metric characterization → OP design → event analysis), and figures generally align with the narrative (Sec. 1–4).

Major issues

1. **Contact cutoff definition is internally inconsistent between Methods and Results, yet it controls all CG/FG graph topologies and therefore every metric and event outcome. Sec. 2.2.2 states the heavy-atom cutoff (4.0 Å) was chosen from an RDF first minimum, whereas Sec. 3.1 states there was no distinct RDF minimum and 4.0 Å was chosen by convention. This ambiguity undermines the validity and reproducibility of the entire pipeline (Sec. 2.2.2, 2.3, 3.1; Abstract).**

   Recommendation: Unify the narrative in one place (preferably Sec. 2.2.2): explicitly state whether an RDF minimum exists and yields **4.0 Å**, or whether RDF was inconclusive and **4.0 Å** was selected heuristically. Include the RDF plot in the main text or SI with the chosen cutoff marked. Add a brief sensitivity analysis (e.g., **3.5/4.0/4.5 Å**, optionally **5.0 Å**) showing robustness of key conclusions: LCC size distribution (Sec. 3.3.1), CG/FG densities and λ_2 distributions (Sec. 3.3), composite OP CoV ranking (Sec. 3.4), and pre-split vs stable differences (Sec. 3.5.2).

2. **FG residue graphs likely conflate intra-peptide and inter-peptide contacts, which can dominate clustering and local connectivity for short peptides and confound interpretation of "aggregate cohesion." For a pentapeptide, intra-chain geometry can create short-range contacts/triangles that inflate FG clustering without reflecting inter-peptide packing; this may also contribute to the reported combination of high FG clustering and near-zero FG Fiedler values (Sec. 2.3.2, 3.3.2).**

   Recommendation: In Sec. 2.3.2 and Sec. 3.3.2, explicitly separate FG edges into (i) intra-peptide and (ii) inter-peptide residue contacts. Recompute (or at least report alongside current results) FG density/clustering/λ_2 on an inter-peptide-only residue graph within each CG aggregate. If you want a single multiscale descriptor, consider reporting both components or using a multiplex/weighted representation (e.g., weight edges by contact occupancy over a short window). Re-evaluate the composite OPs and pre-splitting signatures using inter-peptide FG density to ensure conclusions track aggregate packing rather than peptide-internal geometry (Sec. 3.4–3.5.2).

3. **Interpretation and computation of FG Fiedler values are unclear given frequent near-zero values. $\lambda_2 = 0$ indicates a disconnected graph; small positive λ_2 can arise from sparse graphs or bottlenecks and is sensitive to numerical tolerance. The manuscript discusses "fragmented internal networks" but does not report how often FG graphs are actually disconnected, what precision is used, or how λ_2 is defined for disconnected graphs and tiny graphs (Sec. 2.4.3, 3.3.2, 3.5.1).**

Recommendation: In Sec. 2.4.3 (and/or Sec. 2.5), provide mathematically total definitions: (i) how λ_2 is handled for disconnected graphs (e.g., set to 0 by definition; or compute per component and summarize), and (ii) how cases with $n < 2$ (or $n < 3$) nodes are handled. In Sec. 3.3.2, report: the fraction of FG graphs with > 1 connected component; the distribution of number of components; and λ_2 values with sufficient precision (consider log-scale plots). Also reconcile how a CG aggregate can be connected while the corresponding inter-peptide residue graph is disconnected—if this occurs, it should be explained (e.g., CG edge existence via any heavy-atom contact does not imply residue-graph global connectivity if many edges are intra-peptide; or it may indicate an implementation/indexing issue that should be ruled out).

4. **Aggregate tracking and event detection (formation/dissolution/splitting/merging) are under-specified, yet central to Sec. 3.5 conclusions. The manuscript mentions "maximum overlap" and a Jaccard criterion but does not fully specify thresholds, tie-breaking, how peptide exchange is distinguished from true split/merge, whether temporal smoothing is used to avoid cutoff flickering, and how event counts relate to one another (Sec. 2.7.1, 2.7.2, 3.5.1–3.5.2).**

   Recommendation: Expand Sec. 2.7.1 with a precise algorithmic specification: exact Jaccard threshold(s), tie-breaking rules, one-to-one vs one-to-many matching logic, and explicit definitions for split vs merge vs dissolve/form given overlaps. State whether edges/aggregates are temporally filtered (e.g., persistence for $\geq k$ frames) and justify the chosen minimum aggregate duration (currently 5 frames). Add a short worked example (figure or appendix) illustrating event labeling across 3–5 frames. Provide a robustness check: vary the Jaccard threshold and any persistence parameter and report how splitting-event counts and the pre-split signatures (Sec. 3.5.2) change.

5. **Statistical support for "pre-split vs stable" differences and longevity correlations is incomplete and may be biased by size dependence and temporal correlation. Sec. 2.7.2 mentions tests, and Sec. 3.5.2 uses "significantly different," but p-values/CI/effect sizes, sample sizes, multiple-comparison handling, and independence assumptions are not reported. In addition, density and λ_2 depend on node count, so stable vs splitting groups should be size-controlled (Sec. 2.7.2, 3.5.1–3.5.2).**

   Recommendation: In Sec. 3.5.1–3.5.2, for each comparison (CG density, CG λ_2, FG density, FG λ_2), report: group sizes, summary statistics, test choice and assumptions, test statistic, p-value, and an effect size (e.g., Cliff's delta or Cohen's d) with confidence intervals. Apply and report multiple-comparison correction if multiple metrics are tested. Control for aggregate size by matching (e.g., within size bins or matched sampling) or by regression including size as a covariate. Address temporal correlation by block bootstrapping, subsampling, or a mixed-effects/cluster-robust approach (at

minimum, show that conclusions persist under frame subsampling). Justify the "5-frame pre-split" window and repeat for alternative windows (e.g., 1/10/20 frames) to test robustness.

6. **Equilibration and analysis-window selection (100 ns cutoff) are justified mainly qualitatively; however, all distributions, OP assessments, and event statistics depend on stationarity over 100–500 ns. Some captions/text also suggest inconsistent windows (e.g., 100–250 ns), which further clouds what data underpin reported means/CoVs (Sec. 2.2.2, 3.1; multiple figure captions in Sec. 3).**

   Recommendation: In Sec. 2.2.2 and Sec. 3.1, add quantitative stationarity diagnostics: block-averaged RMSD/Rg and at least one key graph observable (e.g., LCC size, CG density) across successive windows (0–100, 100–200, 200–300, 300–400, 400–500 ns), and/or distribution comparisons early vs late in the chosen window. If drift exists, shift the window or report sensitivity of key results to alternative start times (e.g., 150 or 200 ns). Ensure every figure/caption explicitly states its time window and that summary statistics are computed consistently from the declared equilibrium range.

7. **Core methodological details needed for reproducibility and to assess feasibility are missing or ambiguous: frame stride/time between analyzed frames (needed to interpret "5 frames" and the reported 66,!772 frames), treatment of periodic boundary conditions, definition of "heavy atoms," and computational details (eigensolver choices/tolerances, runtime) (Sec. 2.1–2.5, 2.8).**

   Recommendation: Augment Sec. 2.1–2.5 and Sec. 2.8 to state: MD timestep, trajectory saving frequency, analyzed-frame stride, and the implied physical time represented by 5 frames; treatment of PBC (minimum-image convention and whether molecules are made whole); precise atom selections (confirm heavy atoms exclude H; exclude water/ions); and numerical details for spectral computations (library, solver, tolerances). Provide a compact pseudocode/flowchart from coordinates → contacts → graphs → metrics → tracking/events. Report approximate computational cost (wall time, hardware) and, if possible, provide code/data access or enough detail for reimplementation.

8. **Scope and claims: the study analyzes one peptide sequence (KYFIL), one system size (30 peptides), and one trajectory/protocol. Some discussion/conclusions generalize to "peptide sequences," "dynamic systems," and broader biomaterial implications without clearly delimiting what is demonstrated vs hypothesized (Sec. 2.1, 3.6, 4).**

   Recommendation: In Sec. 3.6 or Sec. 4, add an explicit limitations paragraph: single sequence, finite-size constraints on aggregate distributions, single-replica sampling uncertainty, and possible force-field/concentration dependence. Temper general claims to

hypotheses (e.g., "we expect these signatures may generalize…"). If feasible, add at least one additional replica or a short sensitivity study (e.g., different concentration/system size) to support generality; otherwise clearly frame as future work.

Minor issues

1. Composite order parameters are evaluated primarily via coefficient of variation, which can decrease due to cancellation between anticorrelated inputs rather than improved physical "stability." The manuscript does not quantify whether size and FG density compensate, nor whether the composite OP improves event discrimination (Sec. 2.6, 3.4).

   Recommendation: In Sec. 3.4, report correlation between LCC size and (inter-peptide) FG density; show OP behavior around exemplar split/merge events; and provide a simple predictive/discriminative check (e.g., ROC/AUC for predicting "split within next τ frames," or logistic regression) to demonstrate added value beyond LCC size alone. Clarify physical interpretation of the OP and its dynamic range (min/max).

2. Positioning relative to biomolecular network literature is thin; references skew toward astronomy/cosmology and generic graph-learning, making novelty within MD/residue-interaction-network communities harder to assess (Sec. 1, 4; References).

   Recommendation: Add a short related-work subsection (e.g., Sec. 1.1) summarizing residue interaction networks, MD dynamic networks, network-based allostery/community analysis, and prior network approaches to aggregation/self-assembly. Replace or motivate cross-domain citations and clearly articulate what is novel here (multiscale + dynamic tracking + event-linked signatures).

3. Several methodological design choices are not well justified: 5-frame minimum aggregate duration, choice of metrics (centralities are introduced but not used substantively), and clustering-coefficient variant is not always specified (Sec. 2.4–2.7, 3.3).

   Recommendation: Provide brief rationales for thresholds and metric selection in Sec. 2.5–2.7. Specify clustering definition (average local vs transitivity) consistently in Methods/captions. Either interpret centralities in Results (Sec. 3.3) or clearly label them exploratory and de-emphasize.

4. Figure clarity and consistency: multiple figures/captions lack units, panel labels, or statistical annotations; several plots suffer from overplotting; and some captions/text references appear mismatched (e.g., Figs. 4, 14, 18, 19; broad set noted across Sec. 3) which reduces actionability of the results.

   Recommendation: Systematically audit figures and captions: add units, panel labels, consistent metric naming (e.g., CG density vs CG_Density), and annotate key statistics (mean/SD/CoV; r and p where correlations are claimed; n for sample sizes). Re-

duce overplotting via transparency/hexbin/density plots and export at higher resolution/vector format. Fix any caption–figure mismatches and merge redundant figures where appropriate.

5. Event counts and relationships are hard to interpret (aggregate instances vs formation/dissolution vs split/merge counts), and "**5** frames" is not translated into time units (Sec. 3.5.1).

   Recommendation: Add a small table in Sec. 3.5.1 summarizing counts with definitions, and explicitly state the physical time corresponding to **5** frames given the analysis stride. Clarify whether splits are counted in addition to dissolutions/formations or as separate categories.

6. Uncertainty from MD protocol (single trajectory; thermostat/barostat and force-field dependence) is not connected to uncertainty in graph metrics and signatures (Sec. 2.1, 4).

   Recommendation: Add a short note (Sec. 2.1 or Sec. 4) acknowledging sampling/protocol uncertainty, and—if additional runs are not feasible—recommend future validation via independent replicas and/or alternative force fields.

Very minor issues

1. Abstract keywords are unrelated to the topic (astronomy/template artifacts such as "Astronomical object identification," etc.) (Abstract).

   Recommendation: Replace keywords with domain-relevant terms (e.g., peptide self-assembly, molecular dynamics, multiscale networks, residue contact networks, Laplacian spectra, aggregate stability).

2. Typos/formatting inconsistencies: line-break artifacts, inconsistent section heading formatting (e.g., stray symbols before Sec. 2.8), inconsistent $\mathbf{\AA}/\boldsymbol{\lambda_2}$ notation, and at least one truncated statistic (FG clustering CoV appears cut off) (Sec. 1, 2.8, 3.3.2).

   Recommendation: Proofread and standardize formatting throughout; ensure all reported statistics are complete (including full CoV values/precision) and symbols/units are consistently rendered.

3. Some captions contain vague interpretive phrases and inconsistent time-window statements (e.g., mixing **100–500 ns** and **100–250 ns**) (Sec. 3 figure captions).

   Recommendation: Edit captions to (i) state exactly what is plotted (variables, window, $\boldsymbol{n}$), (ii) provide one concise takeaway, and (iii) ensure time windows match the analysis described in Sec. 3.1.

4. Undefined-case handling for spectral metrics in tiny graphs (e.g., $\boldsymbol{\lambda_2}$ for $\boldsymbol{n < 2}$; handling multiple zero eigenvalues) is not explicitly stated (Sec. 2.4.3).

Recommendation: Add a brief explicit rule covering all edge cases so definitions are unambiguous and reproducible.

Key statements and references

- ⚠ **Graph theory has recently been applied to represent and quantify complex interaction networks beyond simple spatial proximity in scientific domains, motivating its use here to characterize peptide aggregates as multiscale contact networks rather than relying solely on distance cutoffs or aggregate size (Pavlou et al., 2023; Sun et al., 2024).**
- *Reference(s):* Pavlou et al., 2023, Sun et al., 2024
- *Justification:* Pavlou et al., 2023 and Sun et al., 2024 both apply graph theory to represent and quantify complex relationships that are not based on spatial proximity (similarity graphs of galaxy spectra; citation-based knowledge graphs), supporting the general claim that graph methods model complex interaction networks. However, neither paper discusses peptide aggregates, contact networks, or critiques of distance cutoffs/aggregate size in that context. Thus the peptide-specific motivation is not directly supported by these sources.

- ✖ **To ensure that the molecular dynamics analysis focused on equilibrium behavior, the authors followed the common practice of identifying an equilibrium window via global structural observables such as RMSD and radius of gyration, using prior work on equilibrium identification in dynamical systems as methodological precedent (Wang et al., 2015; Voit et al., 2024).**

- *Reference(s):* Wang et al., 2015, Voit et al., 2024
- *Justification:* Neither paper discusses molecular dynamics, RMSD, or radius of gyration, nor the practice of defining an equilibrium window with such observables. Wang et al., 2015 analyzes equilibrium points in gravitational potential fields of irregular celestial bodies, and Voit et al., 2024 develops a regulator model for equilibrium states of galactic atmospheres. These do not provide methodological precedent for MD equilibrium identification as claimed.

- ✖ **The definition of the system's radius of gyration as a descriptor of overall compactness and aggregate formation, and its use to identify a plateau signaling equilibration, follows established practice in the analysis of molecular and astrophysical systems where R_g is used as a compactness metric (Wadhwa et al., 2024, 2024a; Claret, 2023; Li et al., 2021).**

- *Reference(s):* Wadhwa et al., 2024, 2024a, Claret, 2023, Li et al., 2021
- *Justification:* The cited works treat the stellar (dimensionless) radius of gyration (k or β) as an internal-structure/compactness parameter for stars and its role in orbital stability (Wadhwa et al., 2024, 2024a; Li et al., 2021) and explicitly define β via $I \approx$

$(\beta R)^2 M$ (Claret, 2023). However, none of these papers discuss molecular systems, aggregate formation, or using R_g time-series plateaus to signal equilibration. Thus the cross-disciplinary claim and the plateau/equilibration use are unsupported by the provided references.

- **✖ The heavy-atom contact cutoff of 4.0 Å used to define edges in the peptide and amino-acid graphs was chosen by computing the radial distribution function $g(r)$ of inter-peptide heavy-atom pairs and then adopting a standard nonbonded-contact distance consistent with prior RDF-based contact definitions in simulation studies (Pan et al., 2011; Dong et al., 2025).**
- *Reference(s):* Pan et al., 2011, Dong et al., 2025
- *Justification:* Neither paper discusses peptides, amino-acid graphs, or a **4.0 Å** heavy-atom contact cutoff. Pan et al., 2011 analyzes the radial distribution function for inertial particles in turbulent flows; Dong et al., 2025 computes proton–proton RDFs in dense hydrogen. While both mention RDFs, neither provides an RDF-based criterion for biomolecular heavy-atom contacts or endorses a **4.0 Å** cutoff. Thus the claimed methodology and value are unsupported by these references.
- **✖ Spectral graph properties, in particular the Fiedler value λ_2 of the graph Laplacian, were employed as quantitative measures of connectivity and robustness to fragmentation in both coarse-grained and fine-grained peptide networks, following established spectral graph theory and partitioning methodologies (Pavlou et al., 2023; Strey et al., 2024; Eldén, 2023).**
- *Reference(s):* Pavlou et al., 2023, Strey et al., 2024, Eldén, 2023
- *Justification:* Eldén (2023) discusses spectral partitioning and relates the second eigenvalue (Fiedler value) to connectivity via Cheeger-type inequalities, but does not apply it to peptide networks. Pavlou et al. (2023) use spectral graph methods for galaxy spectra, not peptides, and do not employ λ_2 as a robustness metric. No attached paper analyzes coarse- or fine-grained peptide networks, so the claimed application is unsupported.

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: light

The paper's mathematics consists primarily of graph construction rules (contact-based adjacency), standard network statistics (density, clustering, centralities), Laplacian spectral quantities (Fiedler value), and simple composite order parameters (products and normalized sums). There

are very few explicit equations and no multi-step derivations; the main audit points are definition consistency, edge-case completeness, and internal consistency of stated formulas/time windows.

Checked items

1. ✔ **Graph Laplacian definition** (Sec. 2.4.1, p.4 ('$L = D - A$'))
    - **Claim:** Defines the (combinatorial) graph Laplacian as $L = D - A$ for the largest component.
    - **Checks:** definition consistency, notation consistency
    - **Verdict:** PASS; confidence: high; impact: critical
    - **Assumptions/inputs:** A is the adjacency matrix of a simple undirected graph (symmetric, no self-loops)., D is the diagonal matrix of node degrees computed from A.
    - **Notes:** The Laplacian definition is standard and consistent with subsequent use of symmetric eigendecomposition.
2. ✔ **Fiedler value identification** (Sec. 2.4.1, p.4 and Sec. 2.4.2, p.4 (λ_2 as second-smallest Laplacian eigenvalue))
    - **Claim:** Uses λ_2 (second smallest eigenvalue of L) as algebraic connectivity / robustness to fragmentation.
    - **Checks:** definition consistency, sanity/limiting case
    - **Verdict:** PASS; confidence: high; impact: critical
    - **Assumptions/inputs:** Eigenvalues are ordered nondecreasing., For connected graphs, $\lambda_2 > 0$; for disconnected graphs, $\lambda_2 = 0$.
    - **Notes:** Interpretation aligns with the mathematical property that $\lambda_2 = 0$ iff the graph is disconnected (with multiplicity equal to number of components).
3. ✔ **Connected components = aggregates** (Sec. 2.3.2, p.3 (Aggregate Identification))
    - **Claim:** Each connected component of the CG peptide graph represents a peptide aggregate.
    - **Checks:** definition consistency, logic consistency
    - **Verdict:** PASS; confidence: high; impact: critical
    - **Assumptions/inputs:** Edges represent peptide-peptide contacts at a fixed cutoff., Connectivity is an appropriate equivalence relation for aggregate membership.
    - **Notes:** The mapping from connectivity to aggregate membership is mathematically coherent for an undirected contact graph.
4. ✔ **FG graph domain restriction to CG aggregates** (Sec. 2.3.2, p.3 (FG Graph Construction per Aggregate))

- **Claim:** FG graphs are built separately within each CG connected component, with residues as nodes and residue-residue contact edges.
- **Checks:** logic consistency, set/membership consistency
- **Verdict:** PASS; confidence: high; impact: moderate
- **Assumptions/inputs:** Only residues belonging to peptides in that CG component are included., Edges represent residue contacts at the same cutoff.
- **Notes:** The nested construction is internally consistent (FG graphs are induced by residue sets per aggregate).

5. ✔ **Node-count consistency (5 residues per peptide)** (Sec. 3.2, p.6 (example: 24 peptides → 120 FG nodes; 6 peptides → 30 FG nodes))

    - **Claim:** FG node counts equal number_of_peptides_in_aggregate $\times 5$ residues.
    - **Checks:** dimensional/units (counting), definition consistency
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** KYFIL is a pentapeptide (5 residues)., Each residue corresponds to one FG node.
    - **Notes:** The counts are consistent with the stated peptide length and FG node definition.

6. ✔ **Graph density relationship to edge count** (Sec. 3.3.1, p.6–7 ('density, directly proportional to the number of edges for a fixed number of nodes'))

    - **Claim:** For fixed node count N (e.g., CG $N = 30$), density is directly proportional to number of edges.
    - **Checks:** definition consistency, algebraic relationship
    - **Verdict:** PASS; confidence: high; impact: moderate
    - **Assumptions/inputs:** Simple undirected graph density is $m/\binom{N}{2}$ (equivalently $2m/(N(N-1))$).
    - **Notes:** Given a fixed N, density is an affine scaling of m with constant factor $2/(N(N-1))$.

7. ✔ **Coefficient of variation definition and edge case acknowledgment** (Sec. 2.5, p.4 (CoV = std/mean) and Sec. 3.3.2, p.8 (CoV inflated when mean ~ 0))

    - **Claim:** Defines CoV as standard deviation divided by mean and notes issues when mean is near zero.
    - **Checks:** definition consistency, sanity/limiting case
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** Mean is nonzero for stable CoV; otherwise CoV can be arbitrarily large.

- **Notes:** The definition is correct and the caveat about near-zero means is mathematically appropriate.

8. ✔ **Interpretation of near-zero FG Fiedler values** (Sec. 3.3.2, p.8 (discussion around Fig. 10))
   - **Claim:** Near-zero FG λ_2 indicates disconnected or weakly connected FG graphs, implying fragmented internal contact networks.
   - **Checks:** sanity/limiting case, logic consistency
   - **Verdict:** PASS; confidence: medium; impact: moderate
   - **Assumptions/inputs:** λ_2 close to 0 corresponds to easy partitioning / low connectivity; $\lambda_2 = 0$ corresponds to disconnection.
   - **Notes:** The qualitative interpretation matches the mathematical role of λ_2. (A finer point: if the graph is disconnected, multiple zero eigenvalues occur; the text does not discuss multiplicity.)

9. ✔ **Composite OP as product of LCC size and FG density** (Sec. 2.6, p.4 and Sec. 3.4, p.9–10 (OP_Size_x_FG_Density))
   - **Claim:** Defines a composite order parameter as $OP =$ (size of largest CG aggregate) $\times$ (FG density of that aggregate).
   - **Checks:** dimensional/units consistency, definition consistency, range/sanity checks
   - **Verdict:** PASS; confidence: high; impact: moderate
   - **Assumptions/inputs:** Size is a nonnegative integer count., Density is dimensionless in $[0, 1]$ for a simple graph with $N > 1$.
   - **Notes:** The product is well-defined and dimensionless up to an overall count scaling. The paper's use (as a stability metric via CoV) is mathematically consistent.

10. ✔ **Composite OP using FG Fiedler value** (Sec. 3.4, p.9 (OP_Size $\times$ FG_Fiedler))
    - **Claim:** Defines OP = \text{LCC_Size} \times \text{FG_Fiedler} and observes it is near zero because FG_Fiedler is near zero.
    - **Checks:** algebraic relationship, sanity/limiting case
    - **Verdict:** PASS; confidence: high; impact: minor
    - **Assumptions/inputs:** FG_Fiedler can be zero when FG graphs are disconnected or nearly disconnected.
    - **Notes:** If one multiplicand is near zero, the product is near zero; the stated conclusion follows algebraically.

11. ⚠ **Global OP (normalized sum over aggregates)** (Sec. 2.6, p.4 (described verbally, no equation))

- **Claim:** A 'global OP' is defined as a normalized sum over aggregates of (aggregate size × FG property), divided by total peptides.
- **Checks:** missing-definition check, notation completeness
- **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
- **Assumptions/inputs:** A precise summation index set (aggregates per frame) exists., The FG property is defined for each aggregate (including edge cases).
- **Notes:** No explicit formula is given (indices, per-frame vs time-averaged definition, handling of aggregates with undefined spectral metrics). This prevents a full internal-consistency verification.

12. ✖ **Heavy-atom contact cutoff determination narrative** (Sec. 2.2.2, p.3 (RDF first minimum used to set cutoff) vs Sec. 3.1, p.5 (RDF inconclusive; cutoff adopted by common practice))

    - **Claim:** The paper claims both that the **4.0 Å** cutoff was determined from the RDF minimum and that the RDF had no distinct minimum so **4.0 Å** was chosen heuristically.
    - **Checks:** definition consistency, cross-section consistency
    - **Verdict:** FAIL; confidence: high; impact: critical
    - **Assumptions/inputs:** A single cutoff definition is used consistently across all constructions.
    - **Notes:** These statements cannot both be true as written. Since the cutoff defines the adjacency matrices, this inconsistency undermines the stated derivation/justification of the graphs (though the value used may still be consistent in code).

13. ⚠ **Equilibrium analysis window consistency** (Sec. 2.2.1, p.2 and Sec. 3.1, p.5 (**100–500 ns**) vs figure captions/text e.g., Fig. 8 caption p.8 ('250 ns simulation') and Fig. 13 caption p.10 ('**100–250 ns**'))

    - **Claim:** All reported statistics/time series are computed over the stated equilibrium window.
    - **Checks:** cross-reference consistency, definition consistency
    - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
    - **Assumptions/inputs:** A single time interval underlies computed means/std/CoVs unless explicitly stated.
    - **Notes:** The text and captions reference different time spans. Without clarification, it is unclear whether summary statistics are computed on **100–500 ns** or a shorter subset.

14. ⚠ **Spectral metric definition for tiny components** (Sec. 2.4.1–2.4.2, p.4 (computing λ_2 for largest component / each FG graph))

- **Claim:** λ_2 is computed for the largest component / each aggregate FG graph in all frames.
- **Checks:** edge-case completeness
- **Verdict:** UNCERTAIN; confidence: medium; impact: minor
- **Assumptions/inputs:** Graphs always have at least **2** nodes (so a second eigenvalue exists)., Or the implementation defines a convention for small graphs.
- **Notes:** The manuscript does not state how λ_2 is handled when a component/aggregate has $n < 2$ (undefined) or $n = 2$ (λ_2 exists but interpretation differs). This is a definitional completeness issue.

Limitations

- The PDF contains very few explicit equations and no step-by-step derivations; most 'math' is definitional. Where formulas (e.g., density) are described verbally rather than written, the audit can only check conceptual consistency, not exact implemented expressions.
- No explicit notation table is provided; some quantities (e.g., the global OP, Jaccard tracking threshold) are described without precise mathematical definitions, limiting verifiability.
- This audit does not validate any reported numeric values, plots, statistical test results, or computational outputs; it only checks internal analytic/definitional consistency.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

19 numeric consistency checks were executed and all passed. Checks covered time-window/framecount implied timestep, direction/magnitude of a reported RMSD SD reduction, partition sums, residue-to-node multiplications for FG graphs, graph density recomputation for CG and FG LCC graphs (with appropriate tolerance for ratio-of-means approximations), CoV recomputation for LCC size, proportionality checks of equal CoVs under fixed scaling, sanity checks for aggregate event counts, inequalities comparing pre-split vs stable group means, and repeated-constant consistency for the **4.0 Å** contact cutoff.

Checked items

1. ✔ **C1_frames_equilibrium_window** (Results §3.1 (page 5): 'trajectory segment from **100 ns**–**500 ns** ... encompassing **66,!772** frames')
    - **Claim:** Equilibrium phase from **100 ns** to **500 ns** encompasses **66,!772** frames.
    - **Checks:** time-window-to-framecount consistency (requires implied timestep)
    - **Verdict:** PASS

- **Notes:** Implied timestep computed for review: $dt \approx 0.0059906247$ ns/frame using $(n_{\text{frames}} - 1)$ and $dt \approx 0.0059905350$ ns/frame using n_{frames}; no explicit timestep was stated to match against.

2. ✔ **C2_rmsd_sd_reduction_ratio** (Results §3.1 (page 5): 'standard deviation ... decreased ... from 9.496 Å ... to 5.080 Å')
    - **Claim:** RMSD standard deviation decreased from 9.496 Å (first 100 ns) to 5.080 Å (subsequent 400 ns).
    - **Checks:** ratio/percent change recomputation
    - **Verdict:** PASS
    - **Notes:** Decrease confirmed: difference $= 4.416$ Å; percent decrease $\approx 46.5038\%$ (no reported percent to match).
3. ✔ **C3_first_equilibrium_frame_partition** (Results §3.2 (page 6): 'CG graph had 30 nodes ... 35 edges ... two connected components: 24 peptides and 6 peptides')
    - **Claim:** At 100 ns, two CG connected components have sizes 24 and 6 out of 30 peptides.
    - **Checks:** parts-sum-to-total
    - **Verdict:** PASS
    - **Notes:** $24 + 6 = 30$ exactly.
4. ✔ **C4_fg_nodes_from_peptides_and_residues_24mer** (Results §3.2 (page 6): '24-peptide aggregate ... FG graph with 120 nodes (amino acids)' and Results §3.1 (page 5): 'peptide entities ... blocks of 5 residues')
    - **Claim:** A 24-peptide aggregate yields 120 amino-acid nodes given 5 residues per peptide.
    - **Checks:** unit-consistent multiplication
    - **Verdict:** PASS
    - **Notes:** $24 \times 5 = 120$ exactly.
5. ✔ **C5_fg_nodes_from_peptides_and_residues_6mer** (Results §3.2 (page 6): '6-peptide aggregate ... FG graph with 30 nodes (amino acids)' and Results §3.1 (page 5): 'blocks of 5 residues')
    - **Claim:** A 6-peptide aggregate yields 30 amino-acid nodes given 5 residues per peptide.
    - **Checks:** unit-consistent multiplication
    - **Verdict:** PASS
    - **Notes:** $6 \times 5 = 30$ exactly.
6. ✔ **C6_cg_density_from_edges_and_nodes** (Results §3.3.1 (page 6): 'mean of 35.392' edges and 'density ... (mean 0.081)' for CG with 30 nodes)

- **Claim:** CG density mean **0.081** is consistent with mean edges **35.392** for $N = 30$.
- **Checks:** graph density recomputation
- **Verdict:** PASS
- **Notes:** Computed density $= 2E/(N(N-1)) \approx 0.0813609$, consistent with reported **0.081** (rounding).

7. ✔ **C7_cg_cov_edges_from_mean_and_implied_sd** (Results §3.3.1 (page 6): 'mean of **35.392**' and 'CoV **10.40%**' for CG edges)
    - **Claim:** CoV of CG edges is **10.40%** given mean **35.392** (implies a specific SD).
    - **Checks:** CoV-to-SD consistency
    - **Verdict:** PASS
    - **Notes:** Implied SD $=$ **CoV** $\times$ **mean** $\approx 0.104 \times 35.392 = 3.680768$ (SD not reported for direct comparison).
8. ✔ **C8_cg_density_cov_matches_edges_cov** (Results §3.3.1 (page 6): 'edges ... CoV **10.40%**' and 'density ... CoV **10.40%**')
    - **Claim:** CG density CoV equals CG edges CoV (both **10.40%**) for fixed N.
    - **Checks:** repeated-value consistency / proportionality
    - **Verdict:** PASS
    - **Notes:** Both CoVs stated as **10.40%**; exact match.
9. ✔ **C9_lcc_cov_from_mean_and_sd** (Results §3.3.1 (page 7): 'LCC averaged **24.196** ... standard deviation **5.529**, resulting in a CoV of **22.85%**')
    - **Claim:** LCC size CoV **22.85%** is consistent with mean **24.196** and SD **5.529** .
    - **Checks:** CoV recomputation
    - **Verdict:** PASS
    - **Notes:** Computed CoV $= 5.529/24.196 \approx 0.2285088$ vs reported **0.2285**.
10. ✔ **C10_fg_lcc_nodes_mean_from_cg_lcc_mean** (Results §3.3.2 (page 7): 'FG nodes ... mean **120.979**' and earlier 'LCC averaged **24.196** peptides' plus '**5** residues' per peptide (page **5**))
    - **Claim:** FG node mean **120.979** amino acids is consistent with mean LCC size **24.196** peptides and **5** residues per peptide.
    - **Checks:** cross-metric multiplication consistency
    - **Verdict:** PASS
    - **Notes:** Predicted mean $= 24.196 \times 5 = 120.98$, very close to reported **120.979** (rounding-level difference).

11. ✔ **C11_fg_lcc_cov_nodes_matches_cg_lcc_cov** (Results §3.3.2 (page 7): 'FG nodes ... CoV 22.85%' and Results §3.3.1 (page 7): LCC size CoV 22.85%)
    - **Claim:** FG node count CoV equals LCC size CoV (both 22.85%) because FG nodes are proportional to LCC size.
    - **Checks:** proportionality CoV equality
    - **Verdict:** PASS
    - **Notes:** Both CoVs stated as 22.85%; exact match.
12. ✔ **C12_fg_lcc_density_from_nodes_and_edges** (Results §3.3.2 (page 7): 'FG edges averaged 174.896' and 'FG density averaged 0.026' with 'mean 120.979 amino acids')
    - **Claim:** FG density mean 0.026 is consistent with mean edges 174.896 and mean nodes 120.979 for the FG LCC graph.
    - **Checks:** graph density recomputation (approximate, using means)
    - **Verdict:** PASS
    - **Notes:** Using ratio-of-means estimate: $2E/(N(N-1)) \approx 0.0240988$ vs reported mean density 0.026; treated as an approximate sanity check.
13. ✔ **C13_op_size_x_fg_density_mean_consistency** (Results §3.4 (page 9): 'OP_Size_x_FG_Density showed a mean of 0.585' and earlier means 'LCC averaged 24.196' and 'FG density averaged 0.026')
    - **Claim:** Mean(OP = \text{LCC_Size} \times \text{FG_Density}) of 0.585 is consistent with mean LCC size 24.196 and mean FG density 0.026.
    - **Checks:** approximate product-of-means sanity check
    - **Verdict:** PASS
    - **Notes:** Product of means $\approx 24.196 \times 0.026 = 0.629096$ vs reported mean(OP) 0.585; within broad tolerance given $\mathbb{E}[XY]$ need not equal $\mathbb{E}[X]\mathbb{E}[Y]$.
14. ✔ **C14_op_cov_to_sd_implied** (Results §3.4 (page 9): 'OP_Size_x_FG_Density ... mean 0.585 with ... CoV of 7.07%')
    - **Claim:** OP_Size_x_FG_Density CoV 7.07% implies a specific SD given mean 0.585.
    - **Checks:** CoV-to-SD implied
    - **Verdict:** PASS
    - **Notes:** Implied SD $= 0.0707 \times 0.585 \approx 0.0413595$ (SD not reported for direct comparison).
15. ✔ **C15_aggregate_event_counts_consistency** (Results §3.5 (page 10): '2832 aggregate instances ... formation (2830) and dissolution (2831) events ... 48 potential splitting events')

- **Claim:** Counts: **2832** tracked aggregate instances; **2830** formation events; **2831** dissolution events; **48** splitting events.
- **Checks:** count sanity relationships
- **Verdict:** PASS
- **Notes:** Heuristic sanity check: $|\textbf{formations} - \textbf{dissolutions}| = 1$ (within tolerance); weak upper-bound checks also passed.

16. ✔ **C16_split_vs_stable_ratio_cg_density** (Results §3.5.2 (page **12**): 'pre-split ... CG_Density (**0.1259**) ... stable ... (**0.2833**)')
    - **Claim:** Pre-split CG_Density **0.1259** is lower than stable CG_Density **0.2833**.
    - **Checks:** inequality + ratio difference
    - **Verdict:** PASS
    - **Notes:** $0.1259 < 0.2833$; ratio ≈ 0.4444; pre-split is $\approx 55.5595\%$ lower than stable (relative to stable).
17. ✔ **C17_split_vs_stable_ratio_cg_fiedler** (Results §3.5.2 (page **12**): 'pre-split ... CG_Fiedler_Value (**0.0758**) ... stable ... (**0.2524**)')
    - **Claim:** Pre-split CG_Fiedler_Value **0.0758** is lower than stable CG_Fiedler_Value **0.2524**.
    - **Checks:** inequality + ratio difference
    - **Verdict:** PASS
    - **Notes:** $0.0758 < 0.2524$; ratio ≈ 0.3003; pre-split is $\approx 69.9683\%$ lower than stable (relative to stable).
18. ✔ **C18_split_vs_stable_ratio_fg_density** (Results §3.5.2 (page **12**): 'pre-split ... FG_Density (**0.0307**) ... stable ... (**0.0715**)')
    - **Claim:** Pre-split FG_Density **0.0307** is lower than stable FG_Density **0.0715**.
    - **Checks:** inequality + ratio difference
    - **Verdict:** PASS
    - **Notes:** $0.0307 < 0.0715$; ratio ≈ 0.4294; pre-split is $\approx 57.0629\%$ lower than stable (relative to stable).
19. ✔ **C19_contact_cutoff_consistency** (Methods §2.2.2 and §2.3.1-2.3.2 (pages **3-4**) and Results §3.1-3.2 (pages **5-6**): cutoff '**4.0 Å**' repeated)
    - **Claim:** Heavy atom contact cutoff is consistently **4.0 Å** across methods and results.
    - **Checks:** repeated-constant consistency
    - **Verdict:** PASS
    - **Notes:** All extracted cutoff values matched **4.0 Å**.

Limitations

- Only parsed text from the provided PDF was used; numeric values embedded solely in plots/figures were not extracted or pixel-read.
- Several checks involving means of nonlinear functions (e.g., mean density vs density from mean edges/nodes; mean of product vs product of means) are only approximate sanity checks because the underlying frame-by-frame data are not available.
- Any validation requiring the MD trajectory, timestep, raw time series, or per-aggregate measurements cannot be performed from the PDF alone.
- One reported numeric item is truncated in the parsed text (FG clustering coefficient CoV), preventing verification of that CoV.
- Recomputing reported Pearson correlations is not possible without the underlying per-aggregate data.
- Statistical significance claims for splitting vs stable comparisons cannot be validated without p-values/test statistics, sample sizes, and/or raw distributions.