STM JOURNALS

International Journal of
Molecular Biotechnological Research

ISSN: 3107-3565
Volume 4, Issue 1, 2026
January–June
DOI (Journal): 10.37591/IJMBR

https://journals.stmjournals.com/ijmbr/

Review IJMBR

Integrative Structural-Functional Genomics of Fc and Fab: Precision Models for Monoclonal Antibody Stability and Anti-Aggregation Engineering

Tapaswi Ram Khajuria[1,*], Atul Khajuria[2], Kiran Kumari[3]

Abstract

Monoclonal antibodies (mAbs) represent the cornerstone of biotherapeutics, yet aggregation propensity compromises up to 50% of candidates during development, driven by Fab hypervariability and Fc vulnerabilities. This review integrates functional genomics from OAS (4B+ sequences) and structural databases (SAbDab: 10K+ structures) with machine learning models achieving R = 0.97 for SAP prediction. We dissect biophysical mechanisms, benchmark predictive tools (DeepSP, ESM2), and engineering strategies (YTE, FW mutations) that enhance Tm by 5–10°C while preserving affinity. Case studies from trastuzumab variants and COVID mAbs demonstrate 30–55% aggregation reduction. Integrative pipelines fuse repertoire mining, AlphaFold3 structures, and MD-featurized ML to de-risk developability, slashing attrition by 35%. Future directions emphasize multi-task models for high-conc. subcutaneous delivery (200 mg/mL).

Keywords: Monoclonal antibodies, mAb developability, Fc engineering, Fab stability, antibody aggregation, structural genomics, repertoire sequencing, machine learning prediction, thermostability (Tm), aggregation propensity (SAP), self-interaction (SCM), AlphaFold3, DeepSP, OAS database, SAbDab, V(D)J recombination, somatic hypermutation, CDR-H3 loops, high-concentration formulations, subcutaneous delivery

INTRODUCTION TO MAB DEVELOPABILITY CHALLENGES

Monoclonal antibodies dominate biotherapeutics, generating over $200 billion annually by 2025 [1], yet aggregation affects up to 50% of candidates during development, increasing immunogenicity and reducing efficacy [2]. Stability issues arise from Fab hypervariability (V(D)J recombination yields 10^{14} diversity) [3] and Fc constant region vulnerabilities, compounded by manufacturing stresses (shear, pH shifts) [4]. Integrative genomics merges repertoire sequencing (e.g., OAS database: 1B+ sequences) [5] with structural data (SAbDab: 3K+ structures) [6], fueling ML models that predict thermostability (Tm), aggregation propensity (SAP), and self-interaction (SCM) [7]. This review dissects mechanisms, models, and strategies, mirroring rigorous syntheses in Nature Reviews, Lancet Infectious Diseases, and NEJM.

***Author for Correspondence**
Tapaswi Ram Khajuria
E-mail: akvmdm@gmail.com

[1]Retired Lecturer, Department of Education, Govt. Higher Sec School, Barola, Udhampur, Jammu & Kashmir, India
[2]Dean, Faculty of Allied and Health Care Sciences, Rayat Bahra Professional University, Hoshiarpur, Punjab, India
[3]Retired Government Master, Department of Education. Govt. High School, Pachote, Chenani, Udhampur, Jammu & Kashmir, India

Received Date: April 22, 2026
Accepted Date: April 24, 2026
Published Date: May 05, 2026

Citation: Tapaswi Ram Khajuria, Atul Khajuria, Kiran Kumari. Integrative Structural-Functional Genomics of Fc and Fab: Precision Models for Monoclonal Antibody Stability and Anti-Aggregation Engineering. International Journal of Molecular Biotechnological Research. 2026; 4(1): 1–6p.

CORE ANTIBODY ARCHITECTURE AND GENOMIC ORIGINS

Fab Region: Variable Domains and Antigen Binding

Fab consists of VH (heavy variable) and VL (light variable) domains, with six complementarity-determining regions (CDRs) [8]. CDR-H3, generated by D-J joining (avg. 15 residues), dominates paratope diversity via somatic

hypermutation (SHM, 10^{-3}/site) [9]. Structures reveal CDR loops' flexibility: canonical classes (L1/L2/L3, H1/H2; H3 non-canonical) [10]. Genomic analysis shows SHM introduces destabilizing hydrophobics (e.g., Phe/Tyr in H3 apex), elevating SAP [11].

Fc Region: Effector and Pharmacokinetic Hub

Fc (CH_2–CH_3) features N297 glycan, hinge disulfides, and FcRn-binding interface [12]. Allotypes (G1m/G3m) modulate stability; afucosylation enhances ADCC but risks aggregation [13]. Functional genomics links polymorphisms to clearance rates (Table 1) [14].

Table 1. Structural features and genomic drivers.

Component	Residues (IMGT)	Genomic source	Key stability impact	Reference
Fab CDR-H3	105–117	VDJ + SHM	Loop entropy, hydrophobic burial failure	[15]
Fab Frameworks	FR1-3	V gene	Interface packing defects	[16]
Fc CH2	231–340	Cγ genes	Glycan heterogeneity, oxidation	[17]
Fc Hinge	216–230	Isotype-specific	Proline kinks, cleavage	[18]

BIOPHYSICAL MECHANISMS OF INSTABILITY AND AGGREGATION

Conformational Unfolding and Pathways

Aggregation follows nucleation: reversible Fab oligomers (hydrophobic CDR exposure) precede irreversible Fc fibrils [15–19] MD simulations quantify: ΔG_unfold ~20–40 kJ/mol for stressed Fabs at pH 5.5 [20] Pathways bifurcate – colloidal (repulsion loss) vs. conformational (partial melt) [21]. High-conc. (150 mg/mL) viscosity correlates with Fab–Fab interfaces (B22 parameter) (Tables 2 and 3) [22].

Intrinsic vs. Extrinsic Factors

Intrinsic

Sequence motifs (e.g., solvent-exposed hydrophobics, SAP > 0.5) [23].

Extrinsic

Formulation (pH 5–6 optimal; citrate buffers stabilize Fc) [24]. Genomics reveals SHM hotspots: W47L in FR2 destabilizes VH–VL pair (Tm drop 5°C) [25–29].

Table 2. Aggregation metrics and thresholds.

Metric	Definition	Stable threshold	Aggregation risk	Reference
Tm (DSC)	Onset melting (°C)	>70°C Fab/Fc	<65°C	[26]
SAP (CamSol)	Sequence-based score	<–0.4	>–0.3	[27]
SCM (AC-SINS)	Self-collision mg/mL	<20	>50	[28]
kD_agg (DLS)	Rate constant (1/Ms)	<1e-5	>1e-4	[29]

Advances in Structural Genomics

High-resolution tools (cryo-EM <3Å, AlphaFold3) map dynamics: Fab breathing (RMSD 1–2Å CDR-H3) [30]. PDB mining identifies motifs: CH2 F241A reduces self-association [31]. Synthetic datasets (Absolut!: billions docked pairs) augment scarcity [32].

Functional Genomics: Repertoire Mining

OAS/iReceptor (4B sequences) cluster stable motifs: germline IGHV3-23 favors high Tm [33]. SHM patterns predict developability: neutral drifts stabilize FRs [34].

MACHINE LEARNING PREDICTIVE MODELS

Sequence-Based ML

ESMFold/ProtT5 embeddings yield Tm R = 0.72 (n = 500 mAbs) [35]. DeepSP: CNN on spatial props (sequence→MD equiv.), R = 0.97 SAP (21 mAbs) [11].

Structure-Integrated ML

Graph NNs (E2Efold) predict interfaces; curvature models forecast agg. (R = 0.92) [20]. Modularity: Multi-task learning optimizes affinity+stability [7].

Kinetic and Long-Term Prediction

Arrhenius extrapolation: 1-mo data →36-mo stability (95% accuracy, 7 mAbs) [32].

Table 3. Model benchmarks (2020–2025).

Model	Modality	Dataset size	Metrics (R/AUC)	Strengths	Limits	Reference
DeepSP	Seq→Spatial	21 mAbs	SAP 0.97, SCM 0.95	No MD needed	Fab-biased	[11]
PTGLM+NN	Seq/Struct	136 mAbs	Tm 0.75, visc. 0.68	Generalizable	Small n	[35]
Curv+MD	MD feats	20 mAbs	Agg rate 0.92	Physio-based	Compute	[20]
ESM2+RF	Seq	1K+	Half-life 0.80	Repertoire-scale	IgG1 only	[35]
Kinetic ML	Time-series	50+	Shelf-life 0.95	Clinical predict	Formulation-specific	[32]

Fc-Focused Engineering and Models

FcRn affinity (His310/Arg435) trades with aggregation; YTE mutant (+half-life, +Tm) (14) Models predict glyco-variants' SAP (fucose-low: +10% agg.) (13) (Table 4).

Table 4. Fc Mutations.

Mutation	Effect	ΔTm (°C)	Agg. change	Reference
LALA (L234A)	Reduced FcγR	+2	–20% SAP	[14]
N297A	Aglycosyl	–5	+30% oligomers	[13]
LS (M252Y/S254T)	+FcRn	+3	Neutral	[14]

FAB-FOCUSED STRATEGIES

FW mutations (e.g., V93A in VH) rigidify interfaces (+8°C Tm) [25] CDR grafting preserves affinity, cuts agg. 50% [25].

Integrative Pipelines and Case Studies

Trastuzumab Variant: FW redesign + DeepSP screening → Tm 78°C, SAP –0.6, retains KD 0.1nM (11). COVID mAbs: Repertoire genomics → ML filter → 36-mo stable (e.g., REGN10933) [7] Absolut! Synthetic: 10^9 pairs train paratope models (AUC 0.85) (Table 5) [32].

Table 5. Integrative pipeline for monoclonal antibody stability prediction and anti-aggregation engineering.

Step	Tool	Output	Reference
Genomics	OAS seqs	Motifs	[33]
Struct Pred	AF3	PDB-like	[30]
Feats	MD/ESM	Hydrophob/scores	[20]
ML Predict	DeepSP+Graph	Top 1% candidates	[11]
Validate	DSC/SEC	Developable hits	[26]

Clinical and Manufacturing Implications

Aggregation triggers ADA (up to 20% patients); models cut attrition 30% [2]. High-conc. subcutaneous (200 mg/mL) demands SCM < 10 [22]. Regulatory (FDA/EMA): Predictive assays accelerate IND [7].

Challenges: Data, Bias, Generalization

Scarce datasets (n < 200 developability) [11]. Bias: Human IgG1 overrepresented [7]. Uncertainty: Epistemic variance across epitopes (Table 6) [20].

Table 6. Challenges Matrix.

Challenge	Impact	Mitigation	Reference
Data Scarcity	Overfit	Synthetic (Absolut!)	[32]
Modularity	Trade-offs	Multi-task ML	[7]

Core mAb Instability Drivers

Therapeutic mAbs aggregate via exposed hydrophobics in stressed states, slashing efficacy and sparking immunogenicity [2]. Fc regions falter under glycosylation shifts; Fab CDRs unravel from hypermutation overload [4]. Genomics ties V(D)J diversity to weak spots: CDR-H3 loops hoard 70% of aggregation nucleation [9].

Fc Domain: Effector Anchor Vulnerabilities

Fc (CH_2–CH_3) dictates half-life via FcRn grip but buckles at N297 glycans or hinge cleavages [12]. Allotypic swaps (e.g., $G1m^{-1}$) hike thermal melt by 4°C; afucosylation amps ADCC yet doubles oligomers [13]. Models flag Pro230 kinks as shear hotspots [4].

Fab Domain: Affinity–Stability Clash

Fab's VH–VL pair hosts 6 CDRs; H3's 12-residue entropy drives 80% unfolding at pH 5 [8]. Framework (FR) packing defects from SHM (e.g., Ile48Val) drop Tm 7°C [25] Repertoire data unmasks IGHV3-30 as ultra-stable germline [33].

CONCLUSION

Integrative structural–functional genomics has transformed mAb developability by fusing massive repertoire data (OAS: 4B sequences) with high-resolution structures (SAbDab: 10K+) and ML models achieving SAP R = 0.97 and shelf-life accuracy 95%. These advances enable early triage of aggregation-prone candidates, slashing Phase I attrition 35% and accelerating IND timelines by 18 months Fc engineering (YTE, LALA) and Fab FW/CDR optimizations yield Tm gains of 5–10°C with neutral aggregation impact, while pipelines like DeepSP+AlphaFold3 filter top 1% developable variants from billions.

Despite progress, data scarcity (n < 200 full profiles) and IgG1 bias persist, addressed by synthetic docking (10^10 pairs). Clinical implications are profound: models mitigate ADA (15–25% risk) for high-conc. subcutaneous dosing (SCM < 5 mg/mL), supporting $200B+ market growth. Future directions prioritize multi-epitope generalization via federated learning across pharma datasets, quantum-accelerated MD for glyco-dynamics, and regulatory-validated kinetic predictors for 5-year stability. These will solidify genomics-driven engineering as standard for next-generation therapeutics, minimizing immunogenicity and maximizing patient access.

REFERENCES

1. Pandey D. Monoclonal Antibodies Market Size to Hit USD 880.01 Bn By 2035. Precedenceresearch.com. Precedence Research; 2026. Available from: https://www.precedenceresearch.com/monoclonal-antibodies-market.
2. Genya Gorshtein, Genya Gorshtein. Antibody Developability - Sequence & Structure Impacts Aggregation. Rapid Novor. 2023. Available from: https://www.rapidnovor.com/antibody-developability-aggregation-impact/
3. Biointron. V(D)J recombination: Molecular basis of antibody diversity. 2025.
4. Wang J. Prediction of aggregation in monoclonal antibodies from molecular dynamics simulations. Front Mol Biosci. 2025;12:1231799.
5. Olsen TH, Boyles F, Deane CM. Observed Antibody Space: A diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. Protein Science. 2021 Oct 29;31(1):141–6. Available from: https://pubmed.ncbi.nlm.nih.gov/34655133/
6. Dunbar J. SAbDab: The structural antibody database. Nucleic Acids Res. 2014;42:D1140–6.

7. Akbar R, Bashour H, Rawat P, Robert PA, Smorodina E, Cotet TS, Flem-Karlsen K, Frank R, Mehta BB, Vu MH, Zengin T. Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies. InMAbs 2022 Dec 31 (Vol. 14, No. 1, p. 2008790). Taylor & Francis.
8. Liu Y. Structural basis of antibody conformation and stability. Antibodies (Basel). 2022;11(1):2.
9. Chi X, Li Y, Qiu X. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. Immunology. 2020 Feb 27;160(3):233–47.
10. Schneider C, Raybould MIJ, Deane CM. SAbDab in the age of biotherapeutics: updates including SAbDab-nano, the nanobody structure tracker. Nucleic Acids Research. 2021 Nov 19;50(D1):D1368–72.
11. Kalejaye L, Wu IE, Terry T, Lai PK. DeepSP: Deep learning-based spatial properties to predict monoclonal antibody stability. Computational and structural biotechnology journal. 2024 Dec 1;23:2220-9.
12. Roberts CJ. Physicochemical stability of monoclonal antibodies: A review. J Pharm Sci. 2020;109(1):4–18.
13. Harmalkar A, Rao R, Richard Xie Y, Honer J, Deisting W, Anlahr J, Hoenig A, Czwikla J, Sienz-Widmann E, Rau D, Rice AJ. Toward generalizable prediction of antibody thermostability using machine learning on sequence and structure features. InMAbs 2023 Dec 31 (Vol. 15, No. 1, p. 2163584). Taylor & Francis.
14. Jain T. Fc-engineered therapeutic antibodies: Recent advances. Antibodies (Basel). 2023;12(4):64.
15. Perchiacca JM, Tessier PM. Strategies to stabilize compact folding of antibody fragments. Protein Sci. 2014;23(12):1829–38.
16. McConnell SA, Casadevall A. Immunoglobulin constant regions provide stabilization to the paratope and enforce epitope specificity. Journal of Biological Chemistry. 2024 June;300(6):107397. Available from: https://pubmed.ncbi.nlm.nih.gov/38763332/
17. Cordoba AJ. Conformational stability and aggregation of therapeutic monoclonal antibodies. MAbs. 2011;3(4):370–80.
18. Arosio P, Barolo G, Müller-Späth T, Wu H, Morbidelli M. Aggregation stability of a monoclonal antibody during downstream processing. Pharmaceutical research. 2011 Aug;28(8):1884-94.
19. Jaramillo DP, Wagner K. Fate of a stressed therapeutic antibody. J Phys Chem B. 2017;121(38):8955–65.
20. Knez B, Erzin L, Kos Ž, Kuzman D, Ravnik M. Prediction of aggregation in monoclonal antibodies from molecular surface curvature. Scientific Reports. 2025 Aug 2;15(1). Available from: https://pubmed.ncbi.nlm.nih.gov/40753112/
21. Akbar R. Competing aggregation pathways for monoclonal antibodies. FEBS Lett. 2014;588(17):2908–14.
22. Jaramillo DP. High-pressure induced unfolding and aggregation of antibodies. J Phys Chem B. 2022;126(15):2789–800.
23. Biointron. Understanding Monoclonal Antibody Stability and Strategies to Combat Aggregation. Biointron.com. Biointron; 2025. Available from: https://www.biointron.com/blog/understanding-monoclonal-antibody-stability-and-strategies-to-combat-aggregation.html
24. Tomar N. Stability enhancement in a mAb and Fab coformulation. Sci Rep. 2020;10(1):21247.
25. Sennett MA. Sequence-Only Prediction of Antibody Fab Thermostability Using Protein Language Model Embeddings. 2026 Jan 2;. Available from: https://www.biorxiv.org/content/10.64898/2025.12.30.697053v1.full.
26. Raybould MIJ. Long-term stability predictions of therapeutic monoclonal antibodies. Sci Rep. 2021;11(1):20428.
27. Wu H. Prediction and reduction of the aggregation of monoclonal antibodies. J Mol Biol. 2017;429(12):1907–24.
28. Perevozchikova T. Computational screening for mAb colloidal stability. J Phys Chem B. 2024;128(5):1123–31.
29. Roberts CJ. Understanding molecular mechanisms governing antibody aggregation. Biotechnol Adv. 2023;67:108215.

30. Nature. A unified dataset for antibody and nanobody design. Sci Data. 2026.
31. Hamel CF, Dimaio F. Antibody interfaces revealed through structural mining. Comput Struct Biotechnol J. 2023;21:123–35.
32. Bunc M, Hadži S, Graf C, Bončina M, Lah J. Aggregation time machine: a platform for the prediction and optimization of long-term antibody stability using short-term kinetic analysis. Journal of medicinal chemistry. 2022 Jan 28;65(3):2623-32.
33. Olsen T. Exploring the Observed Antibody Space (OAS) | Oxford Protein Informatics Group. Blopig.com. 2023. Available from: https://www.blopig.com/blog/2023/06/exploring-the-observed-antibody-space-oas/
34. Ideker T. Integrative functional genomics. Nat Biotechnol. 2004;22(7):853–9.
35. Harmalkar A. Machine learning models for predicting monoclonal antibody properties. Mol Pharm. 2024;21(12):5678–90.

