Available online **www.ijpras.com**

International Journal of Pharmaceutical Research & Allied Sciences, 2025, 14(3):83-92
https://doi.org/10.51847/zjqeltofID

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

From Sequence Design to a Manufacturable Biologic: A Translation Framework for Expression, Folding, Modification, Purification, and Stability

Nora Schneider[1]*, Tobias Frank[1], Andreas Müller[2], Christoph Meier[3]

[1] Department of Biologic Translation and Sequence Design, Faculty of Pharmacy, University of Freiburg, Freiburg, Germany.
[2] Department of Expression and Folding Optimization, Faculty of Pharmacy, Heidelberg University, Heidelberg, Germany.
[3] Department of Purification and Stability, Faculty of Pharmacy, University of Hohenheim, Stuttgart, Germany.

***Email:** nora.schneider@natfak.uni-freiburg.de*

ABSTRACT

Sequence design can optimize biologics for target engagement, specificity, format, and predicted structure, but sequence-level promise does not establish efficient expression, productive folding, consistent modification, robust purification, formulation, or stability. This non-empirical article proposes a sequence-to-product translation framework for recombinant therapeutic proteins, particularly antibodies and antibody-derived formats. Manufacturability is treated as a conditional interaction among sequence, host, cellular processing, bioprocess, product heterogeneity, purification, formulation, and intended use. The framework introduces translation debt for liabilities carried into later development; non-compensatory gates that prevent strength in one domain from masking failure in another; and conditional manufacturability readiness as a multidimensional judgement. Validation escalates from computational applicability and host-relevant expression to orthogonal molecular-quality, purification, formulation, scale-transfer, and prospective stage-gate evidence. The framework organizes reasoning and progression; it is not a validated predictive model, universal standard, clinical recommendation, or regulatory pathway.

Key words: *Pharmaceutical biotechnology, Biomanufacturing, Cell factory, Bioprocess, Biologic developability, Synthetic biology*

Received: 05 MAY 2025; **Revised:** 31 August 2025; **Accepted:** 01 September 2025

INTRODUCTION

Biologic sequence design has advanced faster than evaluation of whether a design can become a reproducible pharmaceutical product. Therapeutic proteins occupy overlapping distributions of solubility, self-interaction, hydrophobicity, thermal behavior, and related properties, so no single favorable attribute explains developability [1]. Sequence-derived assessments can flag charge, hydrophobicity, or aggregation liabilities when material is scarce, but their value depends on the sequences, structures, formats, and reference distributions used. They are risk screens, not proof that a design will express, fold, purify, formulate, and remain stable [2].

Developability evidence is also fragmented. Design teams emphasize function, cell-line teams titre and growth, downstream teams recovery and impurity clearance, and formulation teams concentration-dependent behavior. Integrated blueprints accordingly treat candidate quality as a combination of molecular properties and stage-specific evidence [3], but do not always explain how liabilities arise, propagate, become controllable, or cross decision boundaries.

The proposed Biologic-Translation Framework connects sequence with expression, host compatibility, folding, modification, purification, formulation, storage, and delivery. It defines constructs and directional relationships, derives stage-gated principles, and specifies evidence for testing them. The contribution is conceptual: it provides neither a dataset nor a validated prediction system, assay standard, regulatory rule, or manufacturing protocol.

Why sequence-level promise does not guarantee manufacturability

Sequence-level promise indicates potential function without obvious computational or structural liabilities. Manufacturability additionally requires efficient production, the intended molecular state, controllable heterogeneity, feasible purification, and acceptable formulation and delivery behavior. This distinction is especially important for fragments, multispecifics, and engineered formats, which can carry stability risks not inferred from full-length antibodies [4].

Developability is a context-sensitive landscape rather than a universal sequence boundary. Charge, hydrophobicity, surface composition, structure, and format create overlapping regions of advantage and liability [5]. A candidate may therefore shift into a higher-risk region after reformatting, concentration, host transfer, or process exposure.

Properties are also coupled. Reducing self-association may alter nonspecific interactions, stability, affinity, or expression. Machine-learning-guided substitutions can improve selected interaction liabilities without establishing performance across the manufacturing pathway [6]. The relevant distinction is between evidence for one property and evidence for the cross-domain consequences of molecular change.

Translation debt denotes unresolved liabilities carried forward, such as host-engineering needs, folding-limited secretion, heterogeneous modification, or concentration-dependent instability. It does not make every early liability disqualifying; it asks whether the liability is understood, detectable, controllable, and unlikely to reappear later. Sequence promise is therefore not equivalent to product manufacturability.

Proposed sequence-to-product translation framework

Sequence optimization is one layer of a larger translation system. Automated experimentation can improve solubility and conformational stability [7], while machine-learning approaches can co-optimize affinity and specificity within tested mutational domains [8]. Neither result establishes host compatibility, modification control, impurity clearance, or stability at the required concentration. Sequence design is therefore an initiating state tested across biological, process, and product contexts.

Figure 1 presents the original conceptual synthesis for translation pathway from designed sequence to manufacturable biologic, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

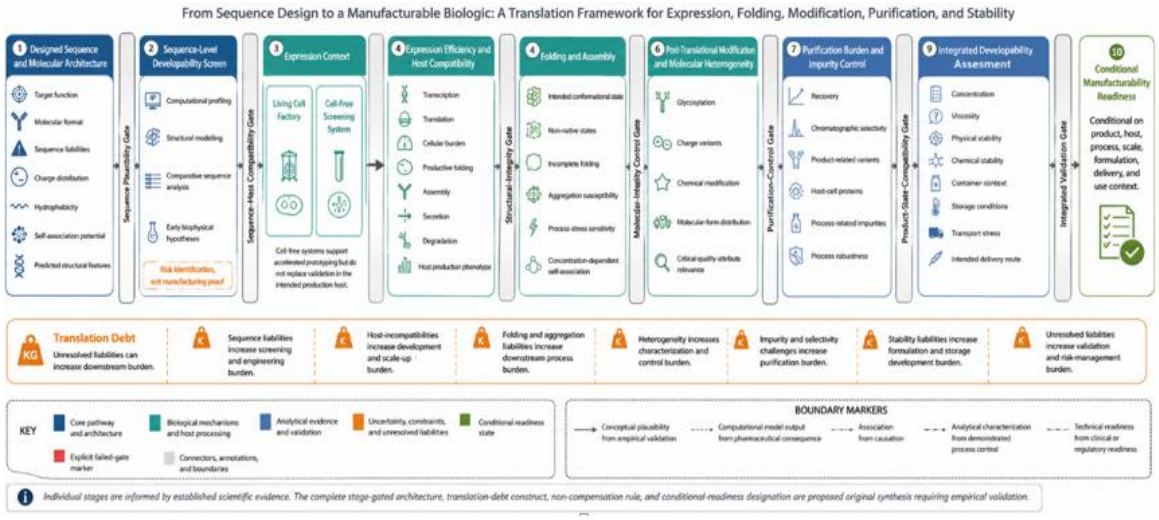

Figure 1. Translation Pathway from Designed Sequence to Manufacturable Biologic

The final layer is conditional manufacturability readiness, defined as a traceable judgement that a candidate has accumulated sufficient and internally coherent evidence to progress under stated product, host, process, scale, formulation, and delivery assumptions. The framework may be represented computationally as a mechanistically

constrained world model in which molecular and process states are connected by explicit dependencies, uncertainty, and validation status. Such a representation would organize evidence rather than simulate a complete cell factory or predict manufacturing success autonomously. Accordingly, no single layer—clinical-stage biophysical profiling, sequence optimization, or cell-free screening—can establish manufacturability because each interrogates a different translation context [1, 7, 9].

Expression efficiency and host compatibility

Titre alone incompletely represents expression. A cell factory must transcribe and translate the construct, support folding, assembly, modification, stress management, trafficking, and secretion. CHO expression can be limited at several linked points [10]; expression efficiency is therefore the ability of a specified hostâ€“constructâ€“process combination to generate the intended product at a usable rate and quality.

Host compatibility concerns whether metabolic, proteostatic, secretory, and quality-control systems accommodate the molecule without an unstable phenotype. Multi-omics work supports productive capacity as a systems-level cellular state [11]. Cell or molecular engineering may improve production but can also change resource allocation, stress, heterogeneity, or process sensitivity.

Mechanisms hidden by bulk measurements matter. Intracellular accumulation of a difficult-to-express antibody has been linked to incomplete folding and endoplasmic-reticulum-associated degradation [12]. Low extracellular yield therefore requires evidence distinguishing transcription, translation, folding, assembly, retention, and degradation.

The sequenceâ€“host compatibility gate integrates transcript, translation, intracellular and extracellular product, assembly, secretion, stress, degradation, cell performance, and recovered-product quality. It passes conditionally when dominant constraints are characterized and controllable, not merely when protein is detectable or one host gives a higher titre.

Folding, aggregation, and conformational stability

Productive folding is acquisition and maintenance of functional conformational and assembly states, not merely full-length expression or absence of precipitation. Directed selection can produce aggregation-resistant variants [13], but resistance remains assay-, format-, environment-, and stress-dependent.

Aggregation also reflects process history. Pumping, mixing, filtration, interfaces, temperature, and concentration gradients can populate non-native states; extensional flow can reveal aggregation absent under quiescent conditions [14]. One flow model does not reproduce every operation, but demonstrates that low-stress measurements may miss vulnerabilities.

Conformational stability differs from colloidal stability. A folded molecule may still exhibit self-attraction, viscosity, clustering, or aggregation at product concentration. Models combining descriptors and experiments can estimate high-concentration aggregation and viscosity risks [15], but remain bounded by their formats, formulations, ranges, and datasets.

The structural-integrity gate separates intrinsic folding, host-dependent assembly, process perturbation, and product-state self-association. It requires orthogonal, context-relevant evidence because aggregation risk depends on sequence, folding, processing stress, and formulation concentration [13-15]. Passing means major liabilities are controllable under specified conditions, not universally absent.

Post-translational modification and molecular heterogeneity

A biologic is usually a distribution of modified or degraded forms rather than one uniform entity. Post-translational modification is part of product formation, and glycosylation can be influenced during biosynthesis or selected downstream separations [16]. Host and culture conditions matter: media additives can alter CHO-derived IgG glycosylation [17], so the same coding sequence may yield different populations across cell lines and processes.

Feed additives can also change N-glycosylation through distributed cellular and product effects [18]. A desired glycoform shift must therefore be interpreted alongside other forms, productivity, stability, and separation needs.

The molecular-identity control gate asks what generates a variant, whether it affects function or stability, whether it is measured reproducibly, and whether its distribution is controllable. Not every detectable difference is a critical quality attribute. Product-specific evidence must connect heterogeneity to function, consistency, stability, or another justified consequence.

Purification burden and impurity control

Purification burden is the effort needed to recover the product while separating variants and process impurities. It reflects both feed-stream composition and molecular properties. Amino-acid substitution can alter monoclonal-antibody cation-exchange behavior, directly connecting design with separation performance [19]. Sequence changes may affect charge, exposure, conformation, retention, selectivity, and recovery, so purification consequences should be considered during design.

This connection complicates the assumption that downstream processing can compensate freely for upstream liabilities. A sequence change may alter charge distribution, surface exposure, conformational populations, or interactions with the stationary phase. These effects can influence retention, selectivity, peak shape, recovery, and the ability to separate closely related molecular variants. The significance of any substitution remains dependent on the molecule, buffer, resin, operating mode, and impurity profile, but the broader implication is that purification consequences should be considered during design rather than after sequence selection is complete.

Process-related impurities create a different but connected burden. Host-cell proteins may vary in abundance, physicochemical similarity to the product, interaction with the product, and susceptibility to removal. Evidence that buffer-system selection can improve host-cell-protein removal shows that impurity clearance depends on process chemistry rather than solely on the product sequence or upstream titre [20]. Purification performance must therefore be demonstrated through process-relevant experiments rather than inferred from a general platform precedent.

The proposed purification-control gate asks whether acceptable recovery, selectivity, impurity removal, and robustness can be achieved without excessive process complexity or unacceptable effects on product quality. Purification may reduce some forms of molecular heterogeneity and remove process-related impurities, but it cannot be assumed to repair incomplete folding, restore lost activity, or eliminate every upstream liability. A heavily compensatory purification process may represent accumulated translation debt even when final analytical results appear acceptable. The gate therefore considers not only whether purification is technically possible, but whether it remains controllable, scalable, and compatible with the intended product state.

Formulation, storage, and delivery stability

Formulation translates a purified molecular population into a product presentation that must remain usable over its intended storage and delivery period. This stage introduces determinants that may have been absent from earlier development, including high concentration, limited administration volume, container interactions, transport stress, repeated temperature exposure, and delivery-device constraints. For subcutaneous biologics, dose requirements can create a trade-off between high-concentration low-volume products and lower-concentration higher-volume delivery strategies [21]. The appropriate balance depends on the molecule, dose, route, administration system, and intended use.

High concentration can amplify molecular interactions that are negligible in dilute solution. Increased viscosity may complicate mixing, filtration, filling, pumping, syringeability, or injection, while stronger self-association may affect physical stability and recovery from interfaces. Formulation-development approaches consequently evaluate viscosity, aggregation, chemical degradation, solubility, excipient response, and manufacturability as connected properties rather than isolated measurements [22]. Early testing can identify risk, but it must approximate the intended product concentration and relevant processing conditions to support a meaningful decision.

Storage stability is similarly conditional. Accelerated studies may reveal degradation pathways or support formulation comparison, but they do not automatically establish long-term behavior under the intended container, closure, temperature, transport, and handling conditions. Delivery stability adds further constraints because a product may experience shear, interfacial contact, warming, dilution, device residence, or administration times not reproduced by routine analytical assays. The framework therefore defines product stability as preservation of justified physical, chemical, structural, and functional attributes within a specified formulation and use context.

The proposed product-state compatibility gate is passed when the molecule's formulation, storage, handling, and delivery risks are sufficiently characterized and controllable for the intended presentation. Stability in one buffer or at one concentration is not generalized to another product context. Nor does successful short-term analytical performance establish clinical suitability or regulatory acceptability. **Table 1** organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for integrated construct, components, evidence requirements, failure modes, and decision boundaries for from sequence design to a manufacturable biologic.

Table 1. Integrated construct, components, evidence requirements, failure modes, and decision boundaries for From Sequence Design to a Manufacturable Biologic.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Designed sequence and molecular architecture	Functional design may conceal liabilities associated with format, charge, hydrophobicity, or self-interaction	Amino-acid sequence, molecular format, structural model, intended function, comparative sequence landscape	**Proposed original synthesis:** sequence decisions establish an initial liability profile that may propagate through later translation domains	Earlier identification of design-associated risk and explicit change-impact hypotheses	Sequence and structural profiling, orthogonal biophysical testing, format-relevant comparison; sequence-based profiling provides risk evidence but not manufacturing proof [2]	Descriptor sensitivity, model-domain limitations, unrecognized format effects, coupled-property trade-offs	A favorable sequence profile does not establish host, process, formulation, or product compatibility
Expression and host compatibility	A biologically active sequence may express poorly, fold incompletely, or burden the production host	Construct design, host genotype and phenotype, expression system, culture conditions, cellular processing capacity	**Proposed sequence–host compatibility gate:** expression is an interaction among the construct, host, and bioprocess rather than an intrinsic sequence property	Mechanistic selection of host, construct, and process interventions	Transcript, translation, intracellular and extracellular product, assembly, secretion, stress, degradation, and recovered-product evidence; linked cellular bottlenecks are documented in CHO systems [10]	Host-specific behavior, unstable production phenotype, incomplete folding, intracellular retention, degradation	Detectable expression or high titre alone does not establish production of the intended molecular state
Folding, aggregation, and conformational stability	Early structural measurements may miss process-induced or concentration-dependent instability	Sequence, molecular format, host processing, process stress, concentration, formulation composition	**Proposed structural-integrity gate:** intrinsic, process-induced, and product-state risks are evaluated separately and then integrated	Prevention of progression based on one non-representative stability assay	Orthogonal folding, assembly, aggregation, self-association, flow-stress, and concentration-relevant studies; process stress can expose hidden aggregation pathways [14]	Assay dependence, stress-model mismatch, non-native-state populations, high-concentration effects	No single thermal, solubility, or aggregation result is sufficient to establish structural robustness
Post-translational modification and molecular heterogeneity	The produced biologic consists of a distribution of modified and variant forms	Host, culture medium, feeds, enzymes, residence time, process conditions, product susceptibility	**Proposed molecular-identity control gate:** variant origin, functional relevance, measurability, and controllability	Product-specific control of molecular-form distributions	Orthogonal characterization, functional linkage, process-intervention studies, and lot consistency; media conditions can alter	Unrecognized variants, intervention trade-offs, analytical blind spots, product-specific effects	Detectable heterogeneity is not automatically a critical quality attribute or an unacceptable defect

			determine significance		glycosylation [17]		
Purification burden and impurity control	Sequence and upstream conditions can produce difficult separations or persistent impurities	Product surface properties, charge, feed-stream composition, resin, buffer, operating conditions, impurity profile	**Proposed purification-control gate:** purification burden represents both intrinsic separation difficulty and accumulated upstream translation debt	Feasible recovery, selectivity, impurity clearance, and process robustness	Product recovery, chromatographic selectivity, product-related variant separation, host-cell-protein clearance, and process robustness; sequence substitution can alter chromatographic behavior [19]	Low recovery, poor selectivity, product–impurity association, excessive process complexity, scale sensitivity	Purification cannot be assumed to repair folding, functional, or stability failures generated upstream
Formulation, storage, and delivery stability	Purified material may fail at the required concentration, storage condition, container, or delivery route	Dose, concentration, excipients, pH, ionic environment, container, temperature, transport, device, route	**Proposed product-state compatibility gate:** product readiness depends on the intended formulation and delivery context	Alignment of molecular properties with product presentation and use	Concentration-relevant viscosity, aggregation, chemical stability, container, storage, transport, and delivery studies; concentration–volume trade-offs shape subcutaneous product design [21]	High viscosity, self-association, chemical degradation, interface effects, device incompatibility, context transfer failure	Stability in one dilute or accelerated condition does not establish suitability for the intended product presentation
Integrated developability assessment	Fragmented evidence may obscure cross-stage inconsistency or unresolved liabilities	Results from all preceding domains, uncertainty, intended use, process assumptions, validation status	**Proposed original synthesis:** a non-compensatory readiness vector records evidence state, uncertainty, and translation debt without producing a universal scalar score	Traceable candidate comparison and explicit progression rationale	Orthogonal cross-domain evidence, documented assumptions, conflict resolution, and prospective testing	False reassurance from averaging, hidden missing data, arbitrary weighting, platform-specific interpretation	Integrated assessment organizes evidence; it does not itself validate manufacturability or prescribe universal thresholds
Validation and stage-gated progression	Early evidence may not predict later scale, product, or use-context behavior	Assay relevance, external systems, process scale, independent products, prospective decisions	**Proposed evidence-escalation mechanism:** confidence increases through computational, experimental, process, scale, and product-	Testable framework decisions and revisable progression rules	External, prospective, cross-host, cross-format, scale-transfer, and decision-impact evidence	Retrospective bias, domain shift, surrogate-outcome failure, untested decision consequences	Passing one gate supports further evaluation and does not establish final technical, clinical, or regulatory readiness

context validation

Integrated developability assessment

Integrated developability assessment is required because the domains in the translation pathway generate evidence of different types, scales, and meanings. Sequence descriptors provide risk hypotheses; expression studies characterize host–construct compatibility; biophysical assays interrogate selected molecular states; purification studies test separation behavior; and formulation studies examine product-context performance. These outputs should not be converted automatically into a common numerical scale. Hierarchical analysis of multidimensional developability data can support candidate comparison and identify patterns that would be obscured by evaluating each measurement independently [23].

Integration also requires orthogonality. Computational and experimental methods often detect different liabilities, and apparent agreement may arise because methods share descriptors, assumptions, or source data. High-throughput in vitro and in silico workflows can accelerate risk identification, but their greatest value lies in combining distinct evidence streams rather than treating one as a replacement for the other [24]. Discordant results should therefore trigger investigation of assay relevance, molecular state, data quality, and context rather than simple averaging.

The framework proposes a conditional readiness vector containing the evidence state of each translation domain, the confidence attached to that evidence, unresolved uncertainty, and accumulated translation debt. This representation differs from a composite score because it preserves non-compensatory failures and prevents strong performance in one domain from hiding a critical weakness elsewhere. Integrated developability assessment therefore requires orthogonal, hierarchically organized evidence rather than a single scalar score [3, 23, 24]. The precise vector structure, terminology, and decision rules are proposed and require prospective testing.

Decision use depends on the purpose of the assessment. A candidate may be sufficiently characterized for exploratory expression while remaining unsuitable for stable cell-line development; it may progress to purification studies while carrying an explicit formulation risk; or it may require redesign because a structural liability lacks a credible control strategy. Integrated assessment should therefore document the intended next decision, the assumptions supporting progression, the evidence still missing, and the conditions that would reverse the decision. It does not classify a molecule as universally manufacturable.

Validation and stage-gated decisions

Validation must address both the individual evidence-generating methods and the framework's stage-gate decisions. An assay may be analytically reproducible yet poorly related to the later outcome it is intended to inform. Conversely, an observed correlation may be useful within one workflow but fail when the molecular format, platform, formulation, or outcome changes. Studies relating in vitro developability measurements to pharmacokinetic outcomes illustrate the potential value of translational correlates while also showing why relationships must be established for the relevant product and decision context [25].

Early experimental systems provide another validation layer. Cell-free protein synthesis can enable rapid production and screening of designed variants, including sequences generated through machine-learning workflows [26]. Such systems can test whether a design is expressible and functional under defined conditions, but they do not reproduce all features of an industrial host, secretory pathway, product-modification system, stable production process, or commercial scale. Cell-free performance should therefore support progression to host-relevant testing rather than substitute for it.

The proposed validation architecture contains escalating layers: computational applicability assessment; orthogonal laboratory confirmation; intended-host expression and molecular-quality evaluation; process-relevant stress, purification, and formulation testing; scale and lot transfer; and prospective evaluation of whether gate decisions predict later development consequences. Each layer should have a prespecified question, comparator, acceptance rationale, and rule for revising the framework. Relationships between early measurements and later outcomes may inform this design, but they remain product- and outcome-specific until independently confirmed [25].

A stage gate is passed when the evidence is sufficient for the next defined decision, not when uncertainty has been eliminated. A conditional pass may identify monitoring requirements or an explicit mitigation plan, whereas a failed gate may lead to redesign, host modification, process reassessment, or candidate discontinuation. The framework does not impose universal pass–fail thresholds because evidence relevance depends on the product,

modality, host, process, formulation, and intended use. Passing an early gate supports progression to another evidence layer; it does not establish final process robustness, clinical suitability, or regulatory acceptability.

Limitations and priorities

The framework is constrained first by the composition of the available evidence. Much of the detailed developability literature concerns monoclonal antibodies and antibody-derived formats produced in CHO cells. The proposed constructs may be relevant to other recombinant proteins, but their mechanisms, dominant liabilities, production hosts, post-translational modifications, purification strategies, and product presentations may differ. Separate evaluation would be required before extending the framework to vaccines, enzymes, microbial proteins, cell therapies, gene therapies, nucleic-acid products, or other biologic modalities.

Computational developability assessment introduces additional limitations. Model performance depends on data provenance, assay consistency, descriptor construction, structural representation, endpoint definition, and coverage of the intended molecular domain. Critical analyses of biopharmaceutical informatics emphasize that data quality, applicability domain, interpretability, and validation practice remain central obstacles to reliable decision use [27]. A computationally favorable candidate may therefore reflect interpolation within a familiar dataset rather than transferable evidence of process or product behavior.

Multimodal approaches may improve selected predictions by combining sequence, structure, experimental, or formulation information. Prediction of antibody viscosity using multimodal feature learning illustrates how multiple representations can be integrated for a defined property [28]. However, improved performance for one endpoint does not establish a comprehensive world model, causal understanding, or generalization to new formats, concentrations, hosts, and processes. The proposed world-model concept must remain constrained by explicit mechanisms, missing variables, domain boundaries, and uncertainty.

Research priorities should therefore focus on prospective rather than exclusively retrospective evaluation. The framework should be tested across independent molecular formats, cell factories, process scales, purification platforms, formulations, and delivery contexts. Studies should examine whether translation debt predicts later resource burden, whether non-compensatory gates improve candidate decisions, whether conditional readiness classifications remain stable as evidence accumulates, and whether rejected assumptions are revised transparently. Computational models should be evaluated outside their original development domains, with data and validation limitations documented explicitly [27]. Until such evidence exists, the framework remains an organizing and hypothesis-generating contribution.

CONCLUSION

The proposed Biologic-Translation Framework reframes manufacturability as a conditional sequence–host–process–product relationship rather than an intrinsic consequence of a promising molecular design. It connects sequence-level assessment with expression efficiency, productive folding, molecular heterogeneity, purification burden, formulation, storage, delivery, and integrated decision use. Translation debt identifies liabilities that propagate into later development stages, non-compensatory gates prevent unrelated strengths from obscuring critical failures, and conditional manufacturability readiness preserves the context and uncertainty of progression decisions. These constructs offer a coherent basis for organizing pharmaceutical-biotechnology evidence and deriving testable validation questions, but they do not establish a validated predictive system, universal assay hierarchy, manufacturing standard, clinical recommendation, regulatory judgement, or deployment-ready workflow. Their value depends on prospective examination across products, hosts, processes, scales, formulations, and intended uses.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Jain T, Sun T, Durand S, Hall A, Houston NR, Nett JH, et al. Biophysical properties of the clinical-stage antibody landscape. Proc Natl Acad Sci U S A. 2017;114(5):944-9. doi:10.1073/pnas.1616408114
2. Raybould MIJ, Marks C, Krawczyk K, Taddese B, Nowak J, Lewis AP, et al. Five computational developability guidelines for therapeutic antibody profiling. Proc Natl Acad Sci U S A. 2019;116(10):4025-30. doi:10.1073/pnas.1810576116
3. Mieczkowski C, Zhang X, Lee D, Nguyen K, Lv W, Wang Y, et al. Blueprint for antibody biologics developability. MAbs. 2023;15(1):2185924. doi:10.1080/19420862.2023.2185924
4. Condado-Morales I, Dingfelder F, Waibel I, Turnbull OM, Patel B, Cao Z, et al. A comparative study of the developability of full-length antibodies, fragments, and bispecific formats reveals higher stability risks for engineered constructs. MAbs. 2024;16(1):2403156. doi:10.1080/19420862.2024.2403156
5. Bashour H, Smorodina E, Pariset M, Zhong J, Akbar R, Chernigovskaya M, et al. Biophysical cartography of the native and human-engineered antibody landscapes quantifies the plasticity of antibody developability. Commun Biol. 2024;7(1):922. doi:10.1038/s42003-024-06561-3
6. Makowski EK, Wang T, Zupancic JM, Huang J, Wu L, Schardt JS, et al. Optimization of therapeutic antibodies for reduced self-association and non-specific binding via interpretable machine learning. Nat Biomed Eng. 2024;8(1):45-56. doi:10.1038/s41551-023-01074-6
7. Rosace A, Bennett A, Oeller M, Mortensen MM, Sakhnini L, Lorenzen N, et al. Automated optimisation of solubility and conformational stability of antibodies and proteins. Nat Commun. 2023;14(1):1937. doi:10.1038/s41467-023-37668-6
8. Makowski EK, Kinnunen PC, Huang J, Wu L, Smith MD, Wang T, et al. Co-optimization of therapeutic antibody affinity and specificity using machine learning models that generalize to novel mutational space. Nat Commun. 2022;13(1):3788. doi:10.1038/s41467-022-31457-3
9. Hunt AC, Rasor BJ, Seki K, Ekas HM, Warfel KF, Karim AS, et al. Cell-free gene expression: Methods and applications. Chem Rev. 2025;125(1):91-149. doi:10.1021/acs.chemrev.4c00116
10. Li ZM, Fan ZL, Wang XY, Wang TY. Factors affecting the expression of recombinant protein and improvement strategies in Chinese hamster ovary cells. Front Bioeng Biotechnol. 2022;10:880155. doi:10.3389/fbioe.2022.880155
11. Raab N, Zeh N, Kretz R, Weiß L, Stadermann A, Lindner B, et al. Nature as blueprint: Global phenotype engineering of CHO production cells based on a multi-omics comparison with plasma cells. Metab Eng. 2024;83:110-22. doi:10.1016/j.ymben.2024.03.007
12. Mathias S, Wippermann A, Raab N, Zeh N, Handrick R, Gorr I, et al. Unraveling what makes a monoclonal antibody difficult-to-express: From intracellular accumulation to incomplete folding and degradation via ERAD. Biotechnol Bioeng. 2020;117(1):5-16. doi:10.1002/bit.27196
13. Ebo JS, Saunders JC, Devine PWA, Gordon AM, Warwick AS, Schiffrin B, et al. An in vivo platform to select and evolve aggregation-resistant proteins. Nat Commun. 2020;11(1):1816. doi:10.1038/s41467-020-15667-1
14. Dobson J, Kumar A, Willis LF, Tuma R, Higazi DR, Turner R, et al. Inducing protein aggregation by extensional flow. Proc Natl Acad Sci U S A. 2017;114(18):4673-8. doi:10.1073/pnas.1702724114
15. Lai PK, Gallegos A, Mody N, Sathish HA, Trout BL. Machine learning prediction of antibody aggregation and viscosity for high concentration formulation development of protein therapeutics. MAbs. 2022;14(1):2026208. doi:10.1080/19420862.2022.2026208
16. Edwards E, Livanos M, Krueger A, Dell A, Haslam SM, Smales CM, et al. Strategies to control therapeutic antibody glycosylation during bioprocessing: Synthesis and separation. Biotechnol Bioeng. 2022;119(6):1343-58. doi:10.1002/bit.28066
17. Ehret J, Zimmermann M, Eichhorn T, Zimmer A. Impact of cell culture media additives on IgG glycosylation produced in Chinese hamster ovary cells. Biotechnol Bioeng. 2019;116(4):816-30. doi:10.1002/bit.26904
18. Kranjc J, Kramer L, Mikelj M, Anderluh M, Pišlar A, Brinc M. Modulating antibody N-glycosylation through feed additives using a multi-tiered approach. Front Bioeng Biotechnol. 2024;12:1448925. doi:10.3389/fbioe.2024.1448925
19. Saleh D, Hess R, Ahlers-Hesse M, Beckert N, Schönberger M, Rischawy F, et al. Modeling the impact of amino acid substitution in a monoclonal antibody on cation exchange chromatography. Biotechnol Bioeng. 2021;118(8):2923-33. doi:10.1002/bit.27798

20. Lakatos D, Idler M, Stibitzky S, Amann J, Schuschkewitz J, Krayl D, et al. Buffer system improves the removal of host cell protein impurities in monoclonal antibody purification. Biotechnol Bioeng. 2024;121(12):3869-80. doi:10.1002/bit.28844
21. Desai M, Kundu A, Hageman M, Lou H, Boisvert D. Monoclonal antibody and protein therapeutic formulations for subcutaneous delivery: High-concentration, low-volume vs. low-concentration, high-volume. MAbs. 2023;15(1):2285277. doi:10.1080/19420862.2023.2285277
22. Zarzar J, Khan T, Bhagawati M, Weiche B, Sydow-Andersen J, Alavattam S. High concentration formulation developability approaches and considerations. MAbs. 2023;15(1):2211185. doi:10.1080/19420862.2023.2211185
23. Metcalf KJ, Wo G, Zaragoza JP, Raoufi F, Baker J, Chen D, et al. Biologics developability data analysis using hierarchical clustering accelerates candidate lead selection, optimization, and preformulation screening. MAbs. 2025;17(1):2502127. doi:10.1080/19420862.2025.2502127
24. Jain T, Boland T, Vásquez M. Identifying developability risks for clinical progression of antibodies using high-throughput in vitro and in silico approaches. MAbs. 2023;15(1):2200540. doi:10.1080/19420862.2023.2200540
25. Jain T, Prinz B, Marker A, Michel A, Reichel K, Czepczor V, et al. Assessment and incorporation of in vitro correlates to pharmacokinetic outcomes in antibody developability workflows. MAbs. 2024;16(1):2384104. doi:10.1080/19420862.2024.2384104
26. Thornton EL, Boyle JT, Laohakunakorn N, Regan L. Cell-free protein synthesis as a method to rapidly screen machine learning-generated protease variants. ACS Synth Biol. 2025;14(5):1710-8. doi:10.1021/acssynbio.5c00062
27. Khetan R, Curtis R, Deane CM, Hadsund JT, Kar U, Krawczyk K, et al. Current advances in biopharmaceutical informatics: Guidelines, impact and challenges in the computational developability assessment of antibody therapeutics. MAbs. 2022;14(1):2020082. doi:10.1080/19420862.2021.2020082
28. Anapindi KDB, Liu K, Wang W, Yu Y, He Y, Hsieh EJ, et al. Leveraging multi-modal feature learning for predictions of antibody viscosity. MAbs. 2025;17(1):2490788. doi:10.1080/19420862.2025.2490788