Computers in Biology and Medicine 204 (2026) 111550

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Overcoming structural complexity in Galectin-3BP through an integrative computational antibody design workflow

Andrielly H.S. Costa [a,b], Eduardo M. Gaieta [a,b], Aline O. Albuquerque [b], Julia S. Souza [b], Diego S. Almeida [a,b], Jean V. Sampaio [a,b], Patrick England [c], Geraldo R. Sartori [d], João H.M. Silva [b,e,*]

[a] Postgraduate Program in Computational and Systems Biology, Oswaldo Cruz Foundation (Fiocruz), 21040-900, Rio de Janeiro, Brazil
[b] Structural and Functional Biology in Biopharmaceuticals Group – Fiocruz Ceará, 61760-000, Eusébio, Brazil
[c] Molecular Biophysics platform. Centre for Technological Resources and Research, Institut Pasteur, Paris, France
[d] São Carlos Institute of Physics, University of São Paulo, São Carlos, São Paulo, Brazil
[e] Pasteur-Fiocruz Center on Immunology and Immunotherapy, Oswaldo Cruz Foundation (Fiocruz), 61760-000, Eusébio, Brazil

ARTICLE INFO

Keywords:
Galectin-3 binding protein
Antibody engineering
Glycoprotein
Epitope screening
Gaussian accelerated molecular dynamics

ABSTRACT

Galectin-3 binding protein (Gal-3BP) is a clinically relevant oncology target, with overexpression associated with poor prognosis across multiple tumor types. However, its therapeutic exploration has been hindered by extensive glycosylation, conformational heterogeneity, and context-dependent oligomerization, which restrict epitope accessibility. Antibody-based strategies remain promising for targeting such complex proteins, yet their development is costly and experimentally demanding. To address these challenges, we established an integrative *in-silico* workflow tailored to the specific structural and biophysical features of Gal-3BP combining validated methodologies of structural prediction, molecular dynamics (MD) simulations, and antibody engineering. By mapping Gal-3BP across oligomeric states and characterizing its N-glycan conformational diversity, we identified two glycan-free epitopes within the BACK domain, termed E1 and E2. Scaffold selection using 3D Zernike descriptors–based similarity search identified BDBV-43 as a compatible candidate for E1. For E2, which lacked similarity-based matches, naïve repertoire mining retrieved the unmatured antibody E2-Ab1, broadening the set of viable templates. Engineering approaches included point mutations in BDBV-43 and full CDR swapping in E2-Ab1. Iterative refinement yielded variants with improved interaction profiles and robust stability during heated MD simulations. Furthermore, Gaussian accelerated MD (GaMD) revealed reorganized conformational landscapes together with modest shifts in the underlying free-energy profiles for the engineered antibodies relative to their native scaffolds, in line with the interpretative limits of GaMD reweighting. Collectively, this study positions Gal-3BP as a tractable therapeutic target and presents optimized antibody candidates capable of engaging epitopes minimally affected by glycan shielding, illustrating the potential of integrative computational pipelines for antibody design against structurally complex proteins.

1. Introduction

Cancer remains one of the leading causes of mortality worldwide, reflecting its biological complexity and the multitude of mechanisms that sustain tumor progression [1]. Over the past decades, immunotherapy has emerged as a transformative strategy, harnessing the immune system to recognize and eliminate malignant cells. Checkpoint inhibitors targeting pathways like PD-1/PD-L1 or the CTLA-4 axis have exemplified the clinical potential of this approach, paving the way for the exploration of novel immunomodulatory targets [2]. Among these, Galectin-3 binding protein (Gal-3BP, UniProt ID: Q08380) has attracted increasing attention due to its involvement in cell adhesion, extracellular matrix remodeling, and immune regulation [3]. Beyond its widespread overexpression in diverse tumors, Gal-3BP enrichment in cancer-derived extracellular vesicles facilitates antibody accessibility and intratumoral uptake, positioning it as a compelling candidate for therapeutic intervention in oncology [4]. Previous preclinical studies have demonstrated the feasibility of targeting Gal-3BP using

* Corresponding author. Structural and Functional Biology in Biopharmaceuticals Group – Fiocruz Ceará, 61760-000, Eusébio, Brazil.
E-mail address: joao.martins@fiocruz.br (J.H.M. Silva).

https://doi.org/10.1016/j.compbiomed.2026.111550
Received 12 December 2025; Received in revised form 27 January 2026; Accepted 8 February 2026
Available online 13 February 2026

antibody–drug conjugates (ADCs) to deliver cytotoxic payloads to the tumor microenvironment, with efficacy demonstrated across diverse models, including neuroblastoma (SK-N-AS), melanoma (A375m), hepatocellular carcinoma (HepG2), and even in human fibroblasts [5–7].

Gal-3BP is composed of three well-defined domains: the scavenger receptor cysteine-rich (SRCR) domain, the Broad-Complex, Tramtrack, Bric-à-brac/Poxvirus and Zinc finger (BTB/POZ) domain, and the C-terminal Kelch (BACK) domain. These are followed by a terminal extension (Phe361–Asp585) lacking detectable homology to any known protein [5]. Despite this domain characterization, its complete structure has yet to be experimentally determined, representing a significant bottleneck to target-oriented antibody design. Moreover, Gal-3BP extensive glycosylation and context-dependent heterogeneity pose significant obstacles to the design of selective therapeutics targeting epitopes that are both accessible and functionally relevant [8,9]. The SP-2 antibody, for example, was reported to target Gal-3BP galectin-binding domain, responsible for glycan-mediated interactions with Galectin-3 [6,10]. However, the substantial glycan variability across tissues and disease states can undermine the predictability of antibody-antigen engagement and, ultimately, limit antibody efficacy [6,10]. Thus, monoclonal antibodies (mAb) development against Gal-3BP requires innovative strategies capable of overcoming glycan-driven variability, ensuring precise recognition of non-glycosylated epitopes.

Regardless of these challenges, antibodies remain a cornerstone of oncology therapeutics due to their specificity, long half-life, and ability to engage immune effector mechanisms [11]. This specificity is driven by the structural diversity of the complementarity-determining regions (CDRs), whose length and sequence variability arise from V(D)J recombination and somatic hypermutation. Among these regions, CDR-H3 exhibits the greatest diversity and structural flexibility, allowing the precise recognition of distinct epitopes [12,13]. Exploiting such diversity for rational antibody development has been accelerated by advances in computational approaches that allow efficient modeling of protein structures, prediction of antigen–antibody interactions, and identification of candidate epitopes.

Established methods, such as molecular docking and molecular dynamics simulations, combined with advances in AI-based structure prediction, have provided unprecedented insights into conformational landscapes that are often difficult to capture experimentally, streamlining candidate prioritization and reducing the cost and time required for experimental screening [14–16]. Among the applications empowered by these predictive tools is antibody repurposing, which gained prominence during the COVID-19 pandemic to rapidly identify existing antibodies with cross-reactivity to new antigens [17]. Such cross-reactivity frequently arises from sequence homology or structural mimicry between epitopes and can be assessed *in silico* through sequence alignment, structural superposition, paratope–epitope docking, and other predictive workflows [18]. More recently, the FP-Zernike framework, which represents protein surfaces as three-dimensional (3D) Zernike descriptors and quantifies similarity via Euclidean distance, has emerged as a promising approach for structure comparison and holds a potential for application in the identification of cross-reactive scaffolds [19].

Building on these advances and considering the central role of Gal-3BP in cancer biology alongside the practical limitations of current therapeutic development approaches for heavily glycosylated targets, we developed a customized workflow grounded in validated *in silico* techniques to overcome the obstacles inherent to Gal-3BP's structural complexity. Our strategy integrates surface-based 3D Zernike descriptors to identify complementary antibody scaffolds [20], with naïve repertoire mining and iterative engineering to refine epitope targeting. This approach prioritizes functionally relevant, glycan-independent epitopes and enables a systematic exploration of antibody design space through iterative computational refinement. As a result, it generates optimized candidates for subsequent *in vitro* validation while reducing the experimental burden traditionally associated with antibody development. Although specifically tailored to the unique challenges posed by Gal-3BP, the concepts and methodologies applied here may inform future design efforts targeting other comparably complex and heterogeneous antigens.

2. Methods

We developed an *in silico* workflow integrating validated computational tools to propose novel antibody candidates capable of engaging with Gal-3BP and prevent its oligomerization (Fig. 1 A). The workflow begins with structural modeling of Gal-3BP, followed by the identification of accessible, non-glycosylated epitopes. These regions served as entry points for antibody discovery while minimizing interference from the protein's extensive glycan shielding.

From this stage, the pipeline branched into two complementary scaffold selection strategies (Fig. 1 B). The first one relied on surface-similarity screening across the AbSet database of antigen-antibody pairs [27]. By identifying epitopes on Gal-3BP that resemble these templates, we retrieved antibody candidates that interact with geometrically similar epitopes. When no matches were detected within the 2.5 Euclidean distance cutoff, we applied a second strategy consisting of iterative docking rounds using naïve antibody scaffolds, selecting variants using computational binding-energy scores and quantitative interface metrics. Both routes converged on targeted mutagenesis or CDR-swapping steps to refine binding specificity and strengthen interactions with the selected epitope. Ultimately, all engineered candidates underwent validation through heated molecular dynamics simulations, allowing the assessment of complex stability and ranking the most promising binders under computational stress conditions.

2.1. Gal-3BP three-dimensional modeling

Gal-3BP amino acid sequence (Q08380) was retrieved from UniProt [21], excluding the signal peptide. Monomeric and multimeric forms of Gal-3BP were modeled with AlphaFold2 and AlphaFold3, respectively [22]. The best model for each protein form was selected using AlphaFold's own local and global confidence metrics (predicted Local Distance Difference Test, pLDDT; Predicted Aligned Error, PAE). Additional comparative evaluation was performed with VoroMQA [23] to assess model quality based on the inter-atomic contacts.

Epitope identification was guided by a four-step selection strategy. First, GlycoSHIELD [24] generated conformational ensembles of the poly-N-Acetyllactosamine (poly-LacNAc) glycan linked to residues 69, 125, 192, 362, 398, 551, 580 of Gal-3BP. The resulting glycosylated models were then evaluated using CPPTRAJ from AmberTools24 [25] to quantify changes in protein solvent accessible surface area (ΔSASA) following glycan addition. Second, structural confidence scores were used to exclude poorly defined and highly flexible segments. Third, experimentally supported predictions from UniProt, based on Protein Epitope Signature Tag (PrEST) fragments [26], were integrated to refine candidate regions. Beyond accessibility, patches were prioritized when located at or near dimerization interfaces, allowing antibody binding to sterically block or allosterically modulate oligomer formation.

2.2. Antibodies repurposing

Aiming to identify a starting scaffold for optimization, we applied the antibody repurposing approach, which relies on the identification of antibodies previously characterized against other targets that exhibit potential cross-reactivity due to high structural similarity between epitopes. To implement this strategy, a similarity search was performed in the AbSet dataset [27], using the protein surface (PS) mode of the FP-Zernike software [20] to compute 3D Zernike descriptors for epitope–paratope pairs (Almeida et al., to be published) . Two accessible epitopes of Gal-3BP were used as queries for the similarity assessment.

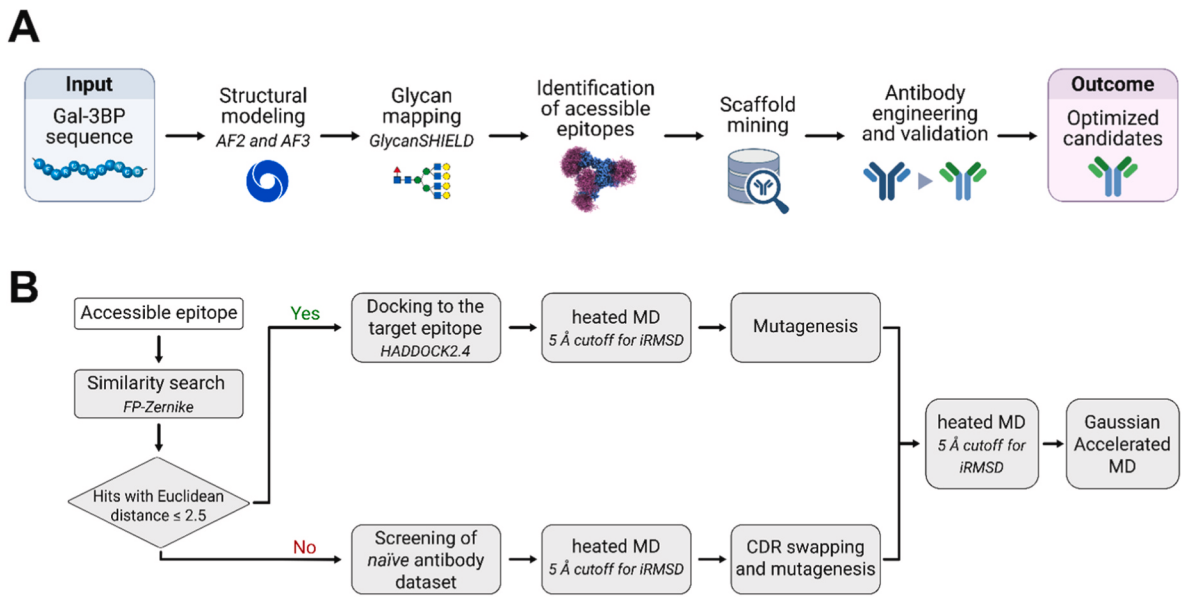

Fig. 1. Overview of the computational pipeline for antibody engineering against Gal-3BP. (A) Linear schematic of the workflow, beginning with structural modeling of Gal-3BP and identification of non-glycosylated, accessible epitopes. (B) Detailed approach for antibody scaffold selection converging on mutagenesis or CDR-swapping refinement steps and validation through heated molecular dynamics simulations.

Antibodies interacting with epitopes structurally similar to those of Gal-3BP, based on a Euclidean distance cutoff of 2.5, were curated according to their original targets. Selected antibodies were subsequently subjected to blind docking to assess whether the repurposed candidates could recognize the new epitope in the absence of predefined interaction constraints. For the blind docking, CDR residues from Martin's definition [28] were designated as active, while all solvent-exposed residues across the antigen surface were defined as passive, guiding the HADDOCK2.4 screening process [29]. Refinement of the binding pose was then achieved through driven docking using the same program and the top-ranked pose was selected through the REF15 scoring function [30].

2.3. Creation of the naïve antibody dataset

For epitopes that failed to generate cross-reactive candidates, we developed a systematic pipeline to identify naïve antibodies compatible with the target epitopes. Unique human germline pairs of the variable heavy (VH) and light (VL) chains were selected from the PyIgClassify database [31] on January 18, 2023. In total, 946 non-redundant pairs were retrieved from antibody structures deposited in the Protein Data Bank (PDB) [32]. To catalogue the sequences corresponding to each germline pair, we queried the Observed Antibody Space (OAS) database [33] on July 12, 2023. Filters were applied to select sequences meeting the following criteria: human species, non-vaccine, and obtained from naïve B cells. Sequences containing any non-canonical amino acid were excluded, remaining 799 germline pairs in our dataset. The antibody sequences were renumbered according to the Martin scheme using a local installation of ANARCI [34]. For positions with missing residues, the most probable amino acid was restored using the AbLang 1.0 language model [35].

2.4. Selection of VH/VL sequence combinations based on CDR-H3

To reduce computational costs, we selected a representative sequence per germline pair with the following procedure: (i) within the group of sequences associated with each germline pair, the most frequent CDR-H3 length was identified; (ii) among sequences of this length, the most probable CDR-H3 sequence was determined using the equation below:

$$P = \prod_{n=1}^{N} P_n^r$$

where P is the product operator, N is the length of the CDR-H3, r is the residue under consideration, n is the position of r within the CDR-H3, and P_n^r is the number of sequences with residue r at position n. The positional entropy was calculated based on the dataset curated by Sampaio et al. [36], providing the probability of residue occurrence at each position of a CDR-H3 of a given length. For each germline pair, the VH/VL sequence combination with the highest P value was selected as representative. The three-dimensional structures of these representative VH/VL pairs were then generated using ABodyBuilder2 [37].

2.5. Antigen–antibody docking protocol

HADDOCK2.4 was employed to generate antigen-antibody complexes. Antibody active residues were defined based on the CDRs, while epitope residues were designated as active. Although HADDOCK2.4 provides its own scoring function to rank the generated poses, all models were additionally rescored using REF15 scoring function to establish a complementary consensus. The underlying rationale was that if different scoring functions rank a given pose highly, the likelihood of it representing a near-native binding mode increases.

Docking poses were filtered in two stages:

First stage. To minimize the selection of decoys, poses were filtered according to the following conditions:

1. Poses with REF15 < 0 kcal mol^{-1};
2. Poses present both in the most populated cluster and in the cluster with the best HADDOCK score. If an antibody had more than one pose satisfying the above conditions, the pose with the lowest REF15 value was selected.

Second stage. To ensure diversity among the ranked candidate poses for molecular dynamics simulations, three pose sets were selected:

(i) the top 10 poses ranked by REF15;
(ii) the top 10 poses ranked by ΔHADDOCK score; and

(iii) the set of poses for which both the ΔHADDOCK score and REF15 values were ≤ -15 [30].

The ΔHADDOCK score reflects the energy difference between the first and second lowest energy clusters of the same antibody. The rationale for selecting poses based on ΔHADDOCK is that if the second cluster exhibits an energy comparable to the first, the most populated cluster is not necessarily the most representative binding mode. Conversely, if a significant energy difference is observed (about −15 HADDOCK arbitrary units), this suggests that the first cluster is not only the most energetically favorable, but also the most representative. A similar approach was successfully applied to distinguish true binding conformations from decoys in antibody-integrin poses produced by ClusPro [38].

2.6. Validation of the antibody/Gal-3BP interaction mode through hMD

The docking-generated complexes were subjected to pose validation through heated molecular dynamics (hMD) simulations using the AMBER24 package [39]. The FF14SB, TIP3P, and GLYCAM06 [40–42] force fields were applied for protein, water, and carbohydrates, respectively. Residues protonation states were assigned using PDB2PQR at pH 6.8 [43]. Complexes were centered in a 12 Å octahedral box, neutralized with Na+ or Cl– counterions, and adjusted to an ionic strength of 0.15 M.

To prepare topology and coordinate files for production dynamics, the systems were subjected to a relaxation protocol divided into 10 steps [44,45]. Initially, water molecules were minimized with 1,000 steps using the steepest descent method, followed by 9,000 steps of conjugate gradient minimization under weak restraints of 1 kcal/(mol·Å^2). The system was then heated to 310 K under constant volume (NVT) for 0.5 ns, applying positional restraints of 100 kcal/(mol·Å^2) to all atoms except hydrogens and water molecules. Water density equilibration was performed under constant pressure (1 bar) with 100 kcal/(mol·Å^2) restraints for 1 ns. Subsequently, restraints were reduced to 10 kcal/(mol·Å^2), and the system was equilibrated for an additional 1 ns. A second minimization was carried out for 10,000 steps using the conjugate gradient method, applying restraints only to backbone atoms with a force constant of 10 kcal/(mol·Å^2). System release was performed in three successive stages of 1 ns each under constant pressure and temperature (NPT), gradually decreasing backbone restraints from 10 kcal/(mol·Å^2) to 1 kcal/(mol·Å^2) and finally to 0.1 kcal/(mol·Å^2). The equilibration protocol was completed with 1 ns of unrestrained molecular dynamics (MD) under NPT conditions. Coulomb, Lennard-Jones, and short-range neighbor list cutoffs were set to 8 Å.

Production dynamics were conducted using the CUDA implementation of the Partial Mesh Ewald Molecular Dynamics (PMEMD) engine [46], combined with hydrogen mass repartitioning (HMR) [47], which enabled simulations with a 4 fs timestep. Temperature and pressure were controlled using the Langevin thermostat and Monte Carlo barostat, respectively. Following the heated MD protocol described by Radom et al. [48], simulations were conducted at 310 K for the initial 30 ns, followed by sequential temperature increases to 330 K and 360 K for 12.5 ns each, and finally to 390 K for the last 15 ns. Trajectory processing and analysis was performed with CPPTRAJ from the AmberTools24 package [49]. Interface RMSD (iRMSD) was computed using a 5 Å cutoff to discriminate native-like binding poses from decoys [38,48]. Interface residues were automatically defined as those located within 8 Å of the binding partner.

2.7. Antibody modification

Antibody modifications were performed using the structural optimization program Ab-SELDON [36], which seeks to mimic the biological process of affinity maturation to enhance antigen binding. The program relies on the analysis of data from human antibodies collected from the Observed Antibody Space (OAS) database to maximize interactions within the microenvironment. Ab-SELDON performs both conformational and sequence sampling through CDR grafting, in which each CDR from the original antibody is replaced by that of another naïve antibody to explore alternative conformations capable of maximizing interactions. This step is followed by point mutation guided by statistical data derived from memory antibody repertoires. Because OAS provides limited representation of λ-chain memory repertoires, κ-chain (IgKV) antibodies were prioritized during earlier selection steps.

Ab-SELDON offers modular execution of its design pipeline. Given that antibodies retrieved through the repurposing workflow already possess affinity-matured CDRs, only the mutagenesis module was applied to these candidates, introducing localized modifications while preserving the global architecture of the paratope. In contrast, the naïve antibodies, whose CDRs remain unmatured, were subjected to both CDR swapping and mutagenesis, aiming to refine loop conformations and enhance chemical complementarity to the target epitope. All Ab-SELDON parameters were kept at default values as recommended by the developers. As CDR swapping can introduce loops of variable length and potentially perturb the local secondary structure of the epitope, we subsequently redocked the optimized CDR-swapped candidates against Gal-3BP. The resulting complexes were evaluated using the REF15 scoring function and were further inspected for their global orientation relative to the antigen.

To complement this strategy, we also employed RosettaAntibodyDesign (RAbD) [50], which iteratively modifies CDR sequences and conformations to increase binding affinity. A total of 100 outer cycles were executed, generating five designs per run. Proline and cysteine substitutions were disallowed to prevent structural instability, and all other parameters followed default configurations. As with Ab-SELDON, repurposed antibodies underwent mutagenesis cycles (using the seq_design_cdrs flag), whereas naïve antibodies were also subjected to CDR grafting (via the graft_design_cdrs flag). Designs that did not incorporate any modifications were removed from further consideration.

Finally, the resulting antibody panels were evaluated using BioPhi [51] to assess humanness. CDR definition and antibody numbering were defined according to the IMGT scheme, and the OASis prevalence threshold was set to "relaxed" (≥10% subjects). OASis Identity and Percentile were evaluated, after which all retained designs were subjected to hMD simulations. Three independent simulation replicas were performed for each system, allowing a robust evaluation of the conformational stability of each modified antibody.

2.8. Evaluation of the potential mean force through Gaussian Accelerated MD (GaMD)

To enhance conformational sampling and characterize the free-energy landscape of the antibody–antigen complexes before and after antibody modification, Gaussian Accelerated Molecular Dynamics (GaMD) was employed [52]. Simulation systems, previously prepared for hMD validation, were first subjected to conventional MD equilibration for 400 ps, followed by an additional 400 ps of unbiased MD during which potential statistics were collected, including the maximum, minimum, average, and standard deviation. This was followed by 3.6 ns of GaMD equilibration, during which boost potentials were applied but potential statistics were not updated. Subsequently, three independent 200 ns production dual-boost GaMD simulations were performed for each system. The sigma0P and sigma0D parameters, which define the upper limits of the standard deviation for the dihedral and total potential boosts, respectively, were set to 6 kcal/mol. All force fields and simulation parameters were maintained from the hMD protocol. While CPPTRAJ was used for trajectory processing and analysis, antibody–antigen interaction profiles were quantified with PLIP 3.12. Frames were extracted every 400 steps from each trajectory, yielding 100 uniformly spaced snapshots per replicate. Residue numbering and chain identifiers were standardized to match the reference PDB.

Interaction frequencies were then averaged across frames. Conformational variability was assessed via principal component analysis (PCA) of antibody-antigen interface Cα atoms using pyPcazip [53]. Finally, 2D PyReweighting was applied to compute the systems' corresponding potentials of mean force (PMFs) [54] using the two first principal components as collective variables (CVs) for potential reweighting. For systems subjected to extensive CDR swapping and subsequent redocking, large variations in antibody orientation precluded projection into a common antibody-based PCA space. In these cases, PCA and subsequent energetic reweighting were restricted to antigen interface atoms to enable consistent comparisons between pre- and post-modification systems. In parallel, PCA was also performed across the three GaMD replicates of each system using antibody–antigen interface atoms to capture internal conformational variability and to extract centroids of the most populated clusters for conformational comparison.

3. Results

3.1. BACK domain identified as a key epitope region

The development of monoclonal antibodies (mAbs) targeting Gal-3BP is constrained by the intrinsic complexity of this protein, characterized by oligomerization and context-dependent glycosylation. In order to direct antibody design towards accessible epitopes, we first modeled Gal-3BP structure using state-of-the-art modeling tool, AlphaFold3, and incorporated N-linked glycosylation conformers (Fig. 2). Prior to glycan modeling, we evaluated the predicted self-assembly states of Gal-3BP (Supplementary Fig. 1) focusing on the mature protein, comprising residues Val19–Asp585 and excluding the signal peptide.

The predicted aligned error (PAE) map (Fig. 2 B) indicated that the dimeric model is globally well resolved, with elevated uncertainty primarily confined to the SRCR domain. This pattern is consistent with previous reports describing the intrinsic flexibility of this region [8,9], responsible for preventing steric hindrance during dimerization (Supplementary Fig. 1). pLDDT scores further supported the reliability of the model, with most residues scoring above 80 (Fig. 2C, Supplementary Fig. 2). Lower confidence was observed in the loop regions spanning Ser386–Arg416, prompting the exclusion of these segments from subsequent analyses related to epitope mapping. These results are also corroborated by the predominance of high-quality scores obtained with VoroMQA (Supplementary Fig. 3, blue).

Nevertheless, expanding the assembly beyond the dimer did not yield improvements in structural quality. The tetrameric, hexameric, and octameric models exhibited substantially lower local confidence scores (Supplementary Fig. 1). Most of these predicted assemblies also failed to recapitulate experimentally observed states detected by microscopy, with the dimer and the hexamer being the only exceptions [8]. Therefore, the dimeric state was prioritized for subsequent analyses, as it not only represents the predominant oligomeric form adopted in extracellular context, but also stabilizes disordered loops that remain solvent-exposed in the monomer.

Next, to capture the steric and conformational variability introduced by distinct glycan ensembles, we added poly-LacNAc to seven N-linked glycosylation sites of Gal-3BP (Fig. 2 A, purple diamond). The glycans' conformational ensemble reveals a broad and dynamic spatial footprint, generating a fluctuating shield that can hinder access to the protein surface, depending on glycan conformation and the surrounding microenvironment. Notably, most glycans are positioned in flexible regions, leaving substantial areas of the BACK domain accessible for interactions without glycan interference (Fig. 2 D). Thus, targeting structurally rigid, glycan-free regions may provide a viable strategy for designing monoclonal antibodies capable of disrupting multimer formation and modulating Gal-3BP function.

To advance from structural modeling to epitope selection, we focused on regions with reliable confidence scores and minimal glycan occlusion, excluding low-confidence or inaccessible segments such as the Ser386–Arg416 loop and the SRCR domain (Fig. 2 B). Complementary UniProt-based experimental evidence from Protein Epitope Signature Tag (PrEST) fragments (Supplementary Figure 4 A) identified the BACK domain as immunogenic, further supporting its relevance as a therapeutic target. In addition, candidate sites were prioritized when

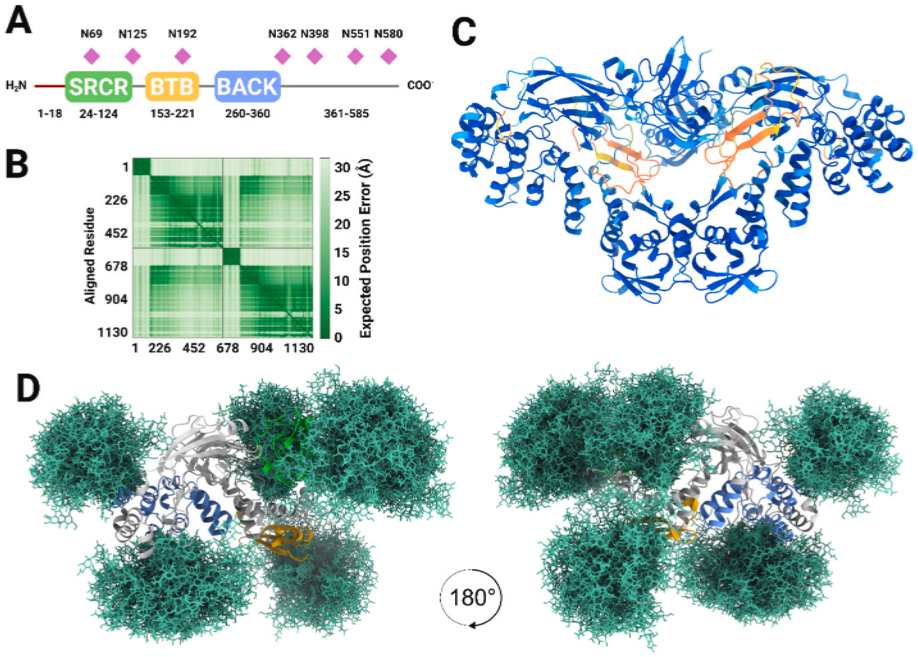

Fig. 2. Structural evaluation of Gal-3BP. (A) The linear domain organization of Gal-3BP, highlighting the glycosylation sites (pink diamonds). (B) Predicted aligned error (PAE) for the dimeric form of Gal-3BP. (C) Tridimensional arrangement of dimer. The protein is colored by the predicted local distance difference test (pLDDT) in dark blue (pLDDT > 90), blue (90 > pLDDT > 70), yellow (70 > pLDDT > 50), and orange (pLDDT < 50). (D) Structure of Gal-3BP (cartoon representation, gray) with seven distinct N-glycan conformers (sticks, green) generated with GlycoSHIELD. The SRCR, BTB/POZ, and the BACK are colored green, yellow, and blue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

located at or near dimerization interfaces (SRCR and BTB/POZ domain), as antibody binding at these locations could sterically hinder or allosterically modulate oligomer formation.

This four-stage integrative strategy converged on two principal regions of interest. The first one (E1) encompasses residues Thr220, Phe227, Asp254, Ser256, Phe257, Met259, Pro260, Tyr264, and Gln292–Val299. The second one (E2) is situated within the C-terminal BACK/Kelch domain, where residues Trp258, Leu265, Ala292, Leu294, Glu295, Leu297, Cys298, Leu302, and Ser330–Glu337 emerged as the most promising epitope candidates (Supplementary Figure 4 B). While glycan shielding markedly reduced overall protein surface accessibility, with a global ΔSASA decrease of approximately 1299 Å^2, the ΔSASA values for E1 and E2 were 2.9 Å^2 and 7.0 Å^2, respectively, indicating that these regions remain largely exposed and reinforcing their suitability as accessible epitope sites (Supplementary Fig. 5).

3.2. *Shape-based screening revealed BDBV-43 as a compatible scaffold for Gal-3BP E1*

Following epitope definition, a surface patch corresponding to each candidate epitope was constructed and queried the AbSet dataset of antibody–antigen complexes to identify antibodies that recognize geometrically similar surfaces, leveraging the central role of shape complementarity in paratope-epitope recognition and enabling the repurposing of existing antibody scaffolds for the new target [55]. To efficiently perform this comparison, we employed 3D Zernike descriptors, which represent local surface geometries as rotation-invariant feature vectors. This descriptor-based strategy bypasses the limitations of traditional alignment methods and supports scalable matching between our epitope models and large collections of structurally characterized epitope-paratope pairs [20].

Applying a Euclidean distance cutoff of 2.5 to retrieve structurally similar epitopes without compromising precision, the Zernike-based search identified five antibody structures whose epitopes share significant similarity with E1, corresponding to PDB IDs 6B3S, 6VGR, 7KEW, 7UVO, and 8EPA (Supplementary Table 1). To reduce the set of candidates, we removed structures containing unresolved regions or sequence gaps, non-human antibodies to mitigate immunogenicity risk, and those paired with λ light chains, given their limited representation in OAS memory repertoires, which would restrict downstream optimization.

The search converged on a single scaffold from the PDB ID 7KEW. Notably, this structure corresponds to BDBV-43, a broadly neutralizing antibody that targets the glycan cap of ebolavirus, a structurally complex and highly glycosylated protein [56]. To assess whether this scaffold could engage the selected Gal-3BP epitope, we first performed blind antibody-antigen docking. Visual inspection of the generated binding modes showed that 32 out of 250 poses were directed to the target site. The spontaneous engagement of BDBV-43 with this region, without any predefined restraints, indicates an inherent compatibility. To refine the binding orientation, we next carried out site-directed docking. Among the top 250 water-refined antibody-antigen models, the final pose was selected based on a REF15 score ranking (Supplementary Table 2).

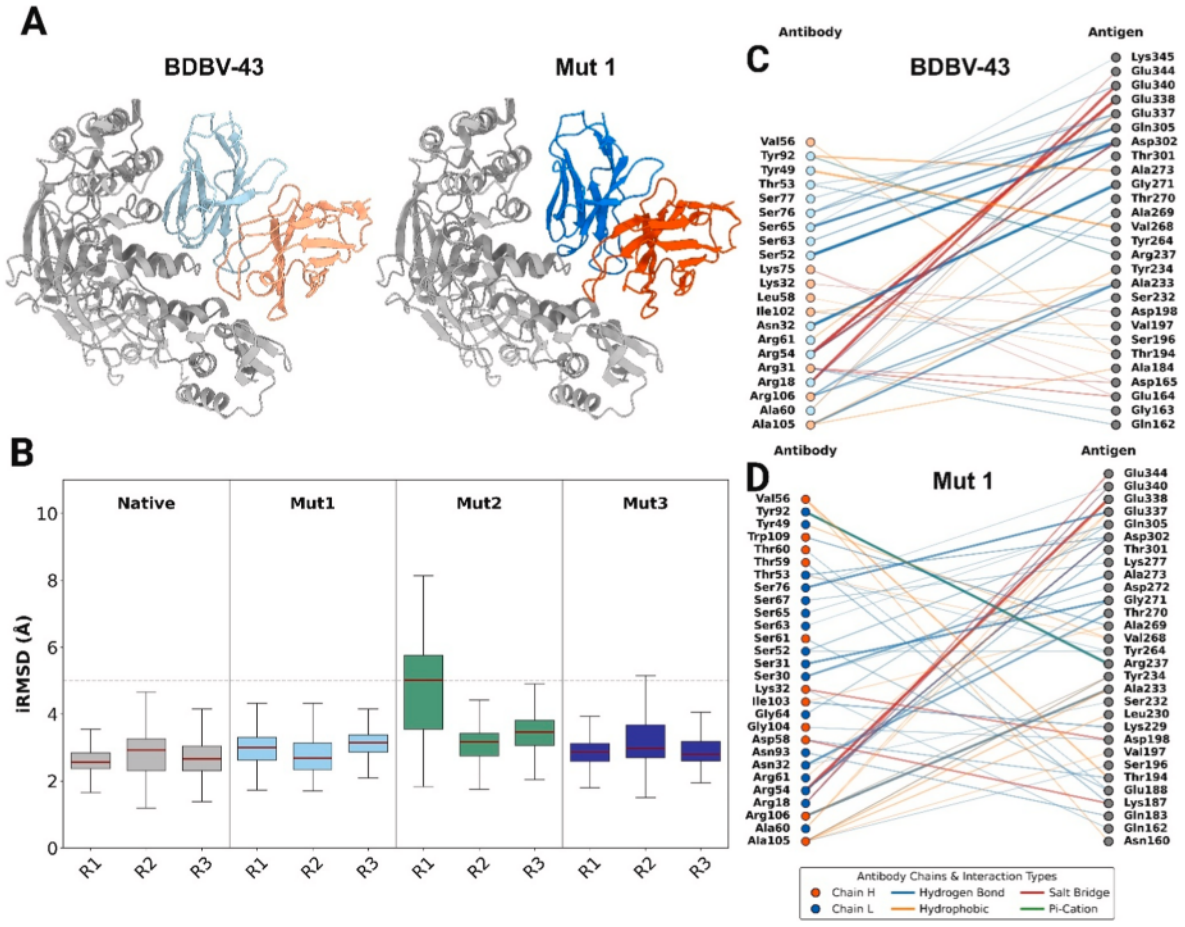

Fig. 3. BDBV-43 antibody engineering. (A) Three-dimensional representation of the parental antibody (light colors) and the Ab-SELDON-generated mutant, termed Mut1 (dark colors). (B) Boxplot representation of the RMSD values for the heated simulations of the complexes containing Gal-3BP and the BDBV-43 antibody (native) or the three variants generated with Ab-SELDON. Interaction network between Gal-3BP (gray) and (C) the native antibody or the (D) optimized candidate (heavy and light chains represented in orange and blue, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

While molecular docking can generate near-native binding conformations, its main limitation remains in the difficulty of distinguishing correct models from decoys. To evaluate the reliability of our docking poses, we employed heated molecular dynamics (hMD), a simulation strategy that subjects the system to sequential thermal perturbations [48]. Under this protocol, kinetically stable models remain intact, whereas incorrect models destabilize and exhibit increased interface RMSD. The unmodified BDBV-43 docked to Gal-3BP displayed remarkable stability during hMD, maintaining a consistent structural arrangement throughout the simulations, with RMSD values below the 5 Å threshold (Fig. 3 B). These results support that shape compatibility, identified via 3D Zernike descriptors and reinforced by molecular docking, translates into a stable interaction in dynamic conditions, confirming BDBV-43 as a suitable starting scaffold for subsequent antibody optimization.

In silico affinity maturation was performed using Ab-SELDON (Fig. 3) and RosettaAntibodyDesign (Supplementary Fig. 6), incorporating targeted mutational rounds to enhance local paratope–epitope complementarity while preserving the overall surface compatibility of the BDBV-43 parental scaffold (Fig. 3 A, light color). RAbD produced five designs whereas Ab-SELDON yielded one refined antibody per run, motivating the execution of three replicates. A subsequent filtering step removed models without any introduced sequence modifications relative to BDBV-43, resulting in the exclusion of four RAbD mutagenesis outputs. We also discarded antibodies that exceeded the 5 Å RMSD threshold in any of the hMD simulation replicates (Fig. 3 B). After applying these criteria, one candidate remained, Mut1 from Ab-SELDON. The selected BDBV-43 variant retained most of the original binding interface, preserving the initial paratope–epitope shape complementarity. Mut1 also exhibited OASis percentile and identity scores comparable to those of the native antibody, indicating that the introduced modifications did not substantially affect its humanness profile (Supplementary Table 3).

Interaction analysis was performed on the enhanced sampling trajectories obtained via Gaussian accelerated MD (GaMD). Despite the limited number of modifications introduced by Ab-SELDON (two-point mutations), the engineered antibody exhibited a strengthened interaction profile. This enhancement was built upon the already robust interface of native BDBV-43 with Gal-3BP, increasing the number of unique interactions from 161 in the native complex to 165 in the engineered variant (Fig. 3C and D). Polar interactions remained predominant, with hydrogen bonds accounting for approximately 67% of total contact.

Most key interaction pairs were preserved, including Arg54–Glu338 and Ser76–Glu337. Additionally, the L58D mutation in CDR-H2 established a persistent salt bridge with Lys187, observed in 50% of the sampled conformations. Despite the CDR-L3 Y91D mutation not introducing novel stable interactions, the light chain contributed to several persistent contacts. In CDR-L2, Arg54 formed both a hydrogen bond and a salt bridge with Asp302 (83% and 79% prevalence, respectively). In CDR-L3, Tyr92 engaged Arg105 through a hydrogen bond and a π-cation interaction (55% and 49% prevalence, respectively). CDR-H3 continued to play a central stabilizing role, with Ile103, Arg106, and Trp109 forming hydrogen bonds with Lys229, Ala233, and Ala269 at 38%, 75%, and 34% prevalence, respectively (Supplementary Fig. 7, Supplementary Fig. 8). Collectively, these interactions support the sustained kinetic stability observed throughout the simulations.

3.3. Naïve repertoire screening and optimization uncovered new binders for the E2 epitope

For E2, no suitable antibody scaffold met the 2.5 Euclidean distance cutoff in the Zernike-based similarity search. We therefore screened a curated, non-redundant database of naïve antibody structures to identify candidates with higher complementarity to the target region. Docking simulations were performed, and the top-ranked structures were selected based on the consensus between REF15 scoring and the ΔHADDOCK score (Supplementary Table 4). This screening process yielded three candidates: 3711081 (E2-Ab1), 765571 (E2-Ab2), and 1491303 (E2-Ab3). Among them, only antibody E2-Ab1 remained below the 5 Å RMSD cutoff during hMD, whereas the others failed to remain within the desired stability threshold (Fig. 4 B). This outcome is consistent with the behavior of antibodies produced by naïve B cells, which undergo affinity maturation and clonal expansion only after antigen encounter and therefore often display limited initial stability [57].

A combined strategy of CDR swapping followed by targeted mutagenesis was subsequently adopted aiming to enhance the affinity and stability of the selected scaffolds. Ab-SELDON and RAbD were used for antibody design, and REF15 scoring served as the primary ranking metric, enabling the selection of the highest-performing variant for each scaffold which was then evaluated through hMD (Fig. 4 B). Additional filtering involved assessing humanness with BioPhi (Supplementary Table 3). Notably, the substantial reduction in both OASis Identity and Percentile for the RAbD-generated designs indicated that the introduced modifications rendered these variants with a significantly less human-like profile both relative to the 9-mer peptides represented in OAS and to therapeutic antibodies included in the percentile calculation. Given the potential increased risk of immunogenicity, these sequences were not considered for further analysis. The Ab-SELDON CDR-swap outputs, which displayed better OASis Identity and Percentile scores, were prioritized and subsequently evaluated based on their structural stability and epitope-interaction profiles.

The hMD evaluation showed that only the E2-Ab1 variant consistently stayed below the 5 Å threshold across all replicates, whereas the remaining variants exhibited pronounced deviations, with RMSD values reaching or surpassing 10 Å (Fig. 4 B). Notably, although the naïve E2-Ab3 initially exceeded the 5 Å cutoff, successive optimization rounds reduced its RMSD average to within the stability threshold. Although this improvement was still accompanied by pronounced deviations across replicates, it demonstrated the effectiveness of the adopted antibody design strategy. Nonetheless, considering the high computational cost associated with enhanced sampling techniques required to recover the free-energy landscape, we restricted our selection to the most stable mutant of antibody E2-Ab1 for all subsequent analyses.

E2-Ab1, which already demonstrated notable stability in its naïve form, retained this favorable behavior after CDR swapping. The incorporation of 20 engineered modifications, primarily within CDR-H3 and CDR-L3, reoriented the antibody into a new binding configuration while preserving its overall alignment with the targeted epitope. This rearrangement led to the loss of specific native contacts, including the salt bridge between antibody Arg36 and antigen Glu382. However, this loss was offset by the emergence of new persistent hydrogen bonds in the remodeled interface. Consistent with this, key antigen residues remained engaged, including Asp473, Phe469, and Phe361 (Fig. 4 A–C, D).

GaMD trajectory analysis revealed that CDR-H3 Tyr101 established multiple hydrophobic contacts with Leu390 and Tyr394, maintained in 50% and 30% of the analyzed frames, respectively. Notably, the interaction between Tyr101 and Tyr394 occasionally adopted a geometry compatible with π–π stacking, as defined by distance and angular criteria, which was detected in ~13% of the frames, highlighting this pair as a newly identified hotspot critical for stabilizing the interface. Despite the polar character of the introduced S56R mutation, the residue at position 56 formed novel hydrophobic interactions with Thr385 and Pro469, observed in 49% and 22% of the frames (Supplementary Fig. 9, Supplementary Fig. 10).

Nevertheless, the most persistent contacts were located in CDR-L1, where Tyr29, Ser30, and Ser31 established stable hydrogen bonds and hydrophobic interactions with Gal-3BP Gln374, all exceeding 77% prevalence. Altogether, these results indicate that our optimization workflow, which seeks to emulate key aspects of affinity maturation, effectively refined interfacial interactions and improved the kinetic

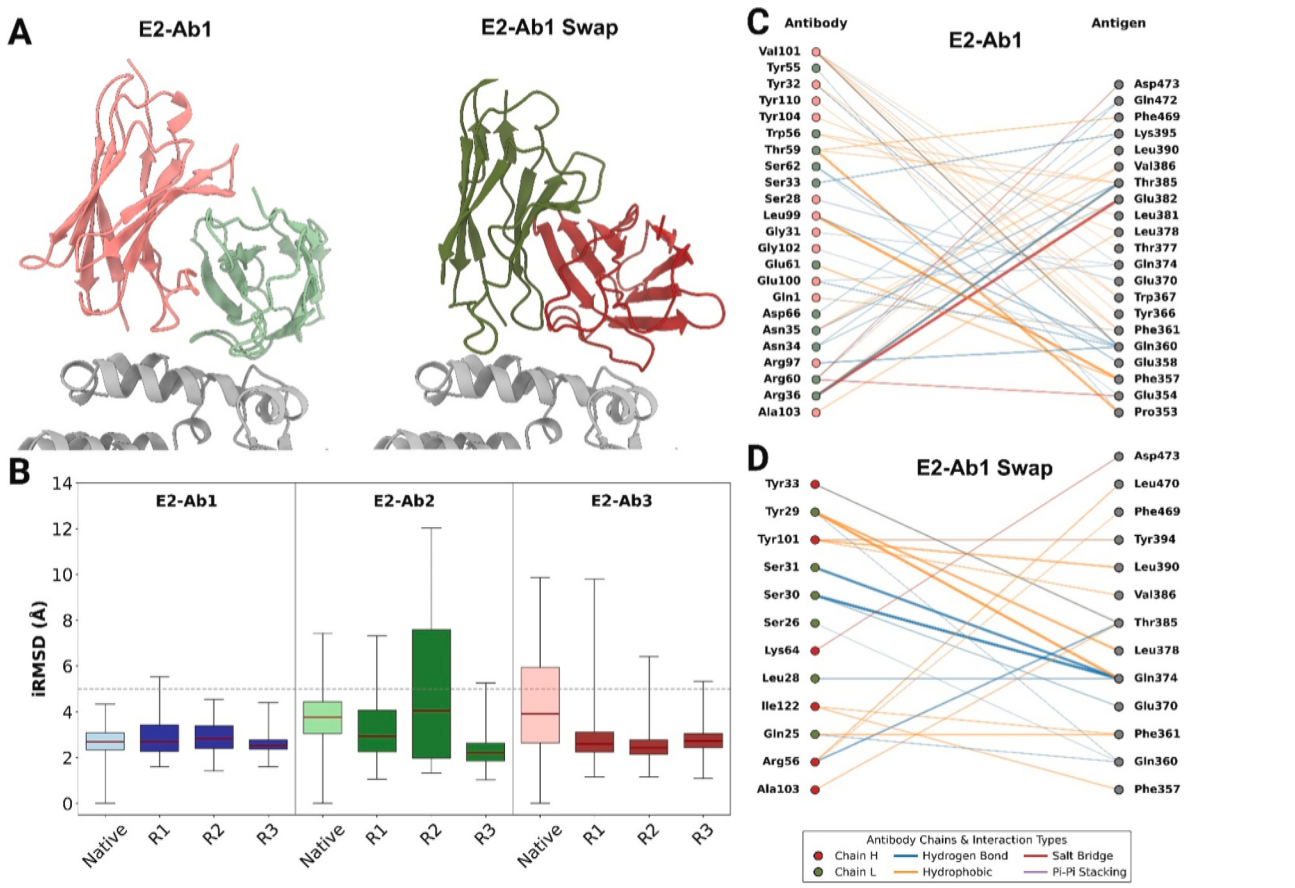

Fig. 4. Naïve antibody mining and optimization toward Epitope E2. (A) Three-dimensional representation of the parental E2-Ab1 antibody (light colors) and of the Ab-SELDON-generated variant (dark colors) (E2-Ab1 Swap). (B) Boxplot representation of RMSD values from hMD simulations of the native naïve scaffolds 3711081 (E2-Ab1), 765571 (E2-Ab2), and 1491303 (E2-Ab3), along with their respective mutants. Interaction network between Gal-3BP (gray) and (C) the native antibody or the (D) optimized candidate (heavy and light chains are represented in red and green, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

stability of geometrically compatible antibody scaffolds identified through naïve repertoire mining.

3.4. *Conformational and energetic landscape of the mutant antibodies*

To characterize the free energy landscape and identify the dominant conformational states sampled during the simulations, the potential of mean force (PMF) was calculated using collective variables derived from principal component analysis (PCA) of the Gaussian Accelerated MD trajectories. The antibody-antigen interface was compared between the native BDBV-43 (Fig. 5 A) and the variant Mut1 (Fig. 5 B). Projection onto the essential subspace revealed a clear separation between these trajectories, with minimal PC1 overlapping and a marked conformational restriction along PC2 in the complex with the mutated antibody (Fig. 5).

The PMF surface of BDBV-43 was characterized by a broad and heterogeneous distribution, with multiple shallow minima spread over an extended region of the conformational space, indicating substantial flexibility and the coexistence of several interfacial substates which allows local rearrangements while preserving the overall interfacial architecture across independent simulations (Fig. 5 A, Supplementary Fig. 11A and B). In contrast, the Mut1-Gal-3BP complex displayed a PMF profile dominated by a well-defined and localized free energy basin (Fig. 5 B). Notably, the global minimum was deeper than that observed for BDBV-43, indicating lower relative free energy and enhanced thermodynamic stabilization of the optimized complex. This stabilization was accompanied by a reduced conformational dispersion, reflecting the reshaping of the accessible landscape toward a more restricted (Supplementary Fig. 11 B, J) and energetically favorable ensemble. Interaction analyses directly support this interpretation, demonstrating not only an increase in the total number of unique interfacial interactions but also the emergence of specific contacts involving the mutated residue L58D. Additional vector analysis along PC1 and PC2 revealed an increased contribution from motions associated with the BTB/POZ domain, which correlated with glycan displacement events observed during the simulation (Supplementary Fig. 12). The separation of conformational states and the modified contribution of the BTB/POZ domain indicate that the mutation affects both structural flexibility and energetic preferences of the system.

For E2-Ab1 extensive CDR swapping and redocking led to pronounced variations in antibody orientation, preventing projection into a shared antibody-based PCA space. Consequently, the analysis was focused on antigen interface atoms before and after optimization. As expected, the modifications introduced in the optimized E2-Ab1 did not substantially alter the range of conformations explored by the antigen interface (Fig. 5C and D, Supplementary Fig. 11C and D). Instead, they reshaped the energetic landscape stabilizing specific interfacial states, with the optimized E2-Ab1 exhibiting lower relative free-energy values compared to the native complex. These results recapitulate the trends observed for Mut1 relative to BDBV-43 and suggest that the twenty engineered alterations exerted a discernible energetic influence on the system.

The deepening of the free-energy minimum observed upon optimization correlated with the interaction analysis, which revealed a

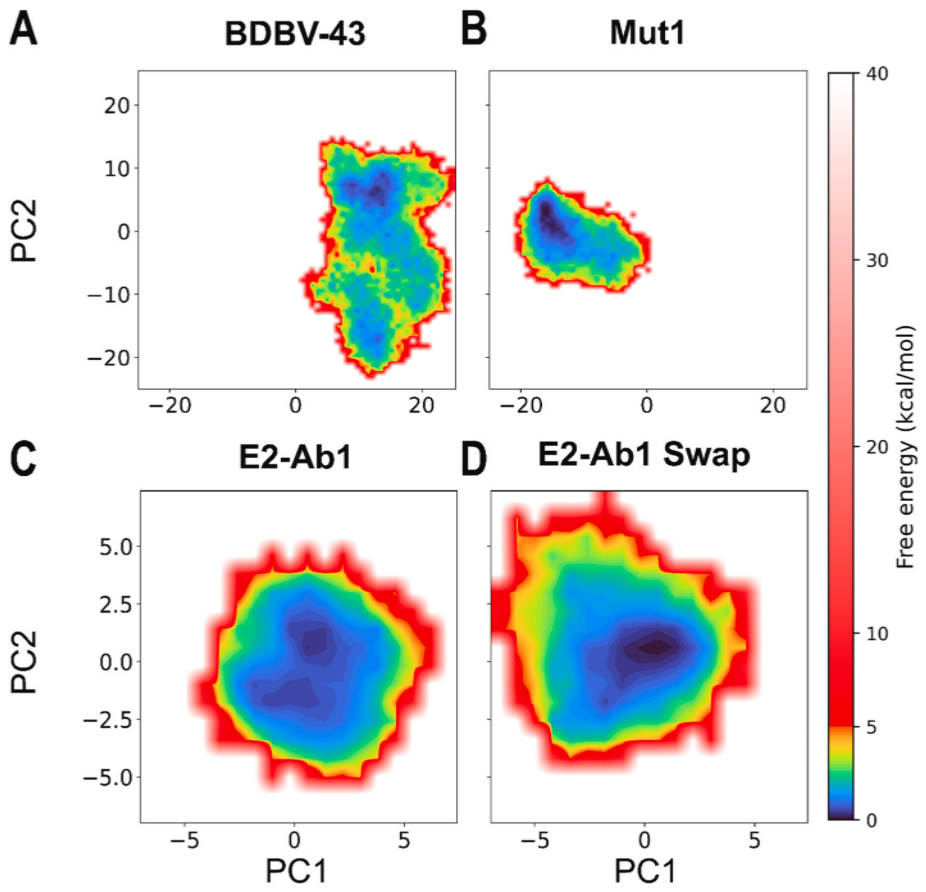

Fig. 5. Potential of Mean Force distribution derived from Principal Component Analysis of interface residues (A) Native BDBV-43 Gal-3BP complex. (B) Optimized BDBV variant Mut1. (C) Native E2-Ab1-Gal-3BP complex. (D) Optimized E2-Ab1 generated by CDR swapping. For E2-Ab1 systems, PCA was restricted to Gal-3BP interface atoms.

reorganization of the interfacial contact network characterized by the replacement of native salt bridges with persistent hydrogen bonding and hydrophobic hotspots. Consistent with this behavior, PCA of the antibody–antigen interface across the three GaMD replicates of the native E2–Ab1-Gal-3BP complex showed limited overlap in the sampled conformational space, possibly reflecting a more nonspecific recognition mode (Supplementary Fig. 11C, D, G, H). In contrast, the interface with the CDR-swapped antibody exhibited a highly consistent conformational profile across replicates, reflecting a more constrained and stable interaction mode. Considered alongside the mutation-driven reorganization identified for Mut1, these findings suggest that rationally introduced alterations can modulate the accessible conformational landscape while preserving the energetic features that govern interface stability.

4. Discussion

Gal-3BP structural characterization remains incomplete, creating significant challenges for the rational design of therapeutic antibodies. To date, no full-length structure has been deposited in the Protein Data Bank and only individual protein domains have been experimentally resolved: the SRCR (PDB ID 1BY2) and the BTB/POZ domains (PDB ID 6GFB) [58,59]. The lack of a complete structural model limits a direct evaluation of the protein dynamics and hampers the identification of antigenic surfaces, thereby requiring computational predictions to approximate its topological organization [11]. AlphaFold, a tool that has reshaped the landscape of protein modeling by accurately predicting even structurally complex proteins, provided a strong structural baseline for Gal-3BP. Recent benchmarking indicated that AlphaFold2 and AlphaFold3 perform similarly in global structure prediction, with AlphaFold3 providing modest LDDT improvements for multidomain proteins [60]. Accordingly, AlphaFold2 and AlphaFold3 were used to model the monomeric and multimeric forms of Gal-3BP, respectively, yielding global folds consistent with previous microscopy observations [8].

The AF-Q08380-F1-v6 model presented high-confidence regions (pLDDT >80) while still revealing segments where predictive accuracy is reduced. Intrinsically flexible or post-translationally modified regions, such as the Ser386–Arg416 loop, showed low local confidence scores and elevated predicted aligned error (PAE), underscoring uncertainty in the absence of explicit glycan modeling. This limitation is consistent with reports that AlphaFold2 does not capture conformational changes driven by post-translational modifications [61], and that regions with low pLDDT scores (<50) should be interpreted as intrinsically disordered or highly flexible rather than misfolded [62]. Moreover, such segments may undergo structural rearrangements upon glycosylation or partner binding, adopting more ordered conformations under specific extracellular contexts [9]. This observation emphasizes both the potential regulatory role of such loops and the importance of integrating glycan modeling to anticipate conformational plasticity.

Hyperglycosylation and glycan heterogeneity add substantial complexity to both the structure and function of Gal-3BP. The protein contains multiple N-linked glycosylation sites whose occupancy varies across cell types and pathophysiological contexts, thereby modulating interactions with lectins and extracellular matrix components [6,10,59]. In practice, this variability generates a dynamic glycan shield capable of masking protein surfaces, making probabilistic assessment of glycan coverage essential for rational antibody targeting. Similar observations in other heavily glycosylated systems, such as the SARS-CoV-2 spike

protein, demonstrate that glycans can restrict epitope accessibility and influence the effectiveness of neutralizing antibodies, underscoring the importance of ensemble-based structural analyses [63]. Consistently, computational strategies based on the insertion of N-linked glycans have been successfully employed for high-throughput epitope profiling [64].

In this context, our investigation revealed two accessible epitopes within the BACK domain that are minimally perturbed by glycan occupancy and therefore represent promising target sites for antibody engagement. Importantly, these regions coincide with oligomerization hotspots, suggesting that their recognition could interfere with Gal-3BP assembly and disrupt biologically relevant oligomeric states [3,9]. A similar strategy was applied to block the Stimulator of interferon genes (STING) protein, whose activity requires oligomer formation. Preventing STING assembly hampered downstream inflammatory signaling [65], exemplifying how disrupting oligomerization can effectively modulate protein function.

Transitioning from structural biology to antibody engineering, our design strategy builds upon established computational affinity maturation frameworks, ranging from early docking- and force-field–based mutagenesis approaches to more recent automated workflows that prioritize rapid variant generation and cost efficiency [66,67]. While conceptually similar, our workflow integrates distinct validated methodologies for epitope mapping, scaffold selection, antibody optimization, and candidate validation. Similarity search using 3D Zernike descriptors provided a robust approach for scaffold selection, yielding an antibody with putative cross-reactivity that displayed remarkable intrinsic stability. This result aligns with the notion that shape complementarity is a key determinant of antibody recognition [68]. Consequently, for antibody repurposing applications, only minimal engineering may be required to preserve essential paratope geometry. In such instances, targeted mutagenesis can be applied to fine-tune local contacts and enhance chemical complementarity. This hypothesis was further supported by the stability of Mut1, containing two mutations relative to BDBV-43, during heated molecular dynamics (hMD) simulations [48]. Such protocol, which employs a stepwise temperature-ramping scheme to distinguish native-like binding poses from decoys, has been successfully used in antibody–antigen systems, including bispecific T-cell engager antibodies against CD3/CD117 and antibody–integrin complexes, and is grounded on prior correlations between molecular dynamics–predicted thermal stability and experimental thermostability [69–71].

Principal Component Analysis (PCA) followed by energetic profiling revealed that optimization steered the antibody–antigen interface toward newly accessible low-energy conformational states. Projection of GaMD trajectories onto the dominant principal components showed increased consistency in the conformational space sampled by replicates containing optimized candidates, emphasizing the importance of replicate simulations for robust convergence and sampling assessment when using structurally sensitive techniques such as PCA [72].

Compared with the native antibody, the optimized BDBV-43–derived variant led to a redistribution of interface conformational populations, characterized by pronounced restriction along PC2 and concurrent sampling of previously inaccessible regions along the PC1 axis. This behavior, together with the improved interaction profile, is consistent with a shift toward interface-competent conformations. The ability of PCA to capture such reorganization of the conformational landscape has been previously demonstrated [73,74]. The corresponding free energy landscapes, constructed using the leading principal components as reaction coordinates, highlighted the stabilization of specific low-energy basins in the variants derived from BDBV-43 and E2-Ab1, consistent with effective interface optimization toward thermodynamically favored states.

It is important to recognize that the reliability of the PMF reconstruction relies strongly on the Gaussian approximation (second-order cumulant expansion), and that GaMD reweighting is susceptible to the influence of outliers and incomplete sampling [52,75]. As a result, subtle energetic shifts should be interpreted with caution. Even so, the preliminary energetic assessment indicated that minimal paratope modifications can still induce measurable energetic alterations, potentially translating into the modulation of the binding affinity [12,13]. This is particularly relevant for the repurposed antibody, where such subtle adjustments may help redirect its affinity toward the new antigen.

Despite the trends observed in the PCA analysis, it is essential to consider potential sources of bias beyond the introduced mutations. The starting GaMD structures differed between systems (originating from docking for the antibody scaffolds and from the optimization pipeline for the mutants) and may remain kinetically trapped throughout the simulation, limiting transitions to alternative states despite the application of boost potential. Moreover, as the systems were processed by successive mutation and minimization steps, part of the observed differences may arise from cumulative rearrangement rather than solely from the mutations themselves.

Nevertheless, it is also important to recognize that the native and mutant antibodies may intrinsically occupy distinct conformational spaces and never converge to each other's structural states during the interaction process. Such behavior would support the hypothesis that the mutation directly modulates conformational sampling and stabilizes an alternative energetic landscape. In addition, the glycan linked to Asn192 increased the flexibility in the BTB/POZ domain, located close to the antibody interface. This enhanced mobility likely contributes to the dynamic behavior detected in both PCA and PMF analyses and must be considered an intrinsic factor influencing collective motions and energy calculations, rather than an effect exclusively associated with the mutations.

For epitopes that lacked cross-reactive scaffolds, we propose that starting from an unmatured antibody and mimicking the natural maturation trajectory of B-cell receptors could yield variants with improved stability and affinity while preserving epitope specificity. Despite recent advances in artificial-intelligence–based methods, including protein language models and diffusion approaches applied to *de novo* antibody design, experimental validation of the antibodies engineered by these techniques remain relatively scarce. Moreover, diffusion-based methods such as RFAntibody scale in runtime as $O(N^2)$, where N corresponds to the number of residues in the system [76]. For large protein targets like Gal-3BP, containing more than 600 residues, the computational expense remains substantial even when the structure is truncated to include only the relevant interaction domains.

By contrast, established strategies such as CDR swapping and targeted mutagenesis possess a broader track record of success in affinity and specificity optimization, as observed in engineering efforts against the SARS-CoV-2 Spike protein [77] and in the optimization of antibodies targeting PAD4 [78]. These approaches, which combine conformational and sequence variation with iterative selection, were successful in establishing stabilizing interactions for the E2-Ab1 antibody with a solvent-accessible epitope of Gal-3BP. Going forward, hybrid workflows that integrate repertoire mining, CDR grafting, and AI-based methods can build on our current strategy to accelerate the translation of computational designs into experimentally validated therapeutic candidates.

Overall, the translational implications of our study are direct, as antibodies designed to bind glycan-free surfaces located at or near oligomerization interfaces are most likely to modulate Gal-3BP function with robustness and specificity. Two promising antibodies (modified BDBV-43 and E2-Ab1) were proposed, which may represent a novel avenue for therapeutic development targeting Gal-3BP. In summary, our results demonstrate that integrating validated *in silico* approaches with biologically grounded epitope selection criteria enables a rational and efficient pipeline for antibody discovery against structurally complex targets such as Gal-3BP, while remaining complementary to downstream experimental characterization.

5. Conclusion

In this study, we presented a computational framework for anti-Gal-3BP antibody engineering that integrates atomistic modeling, structural comparison, and targeted redesign to refine antigen recognition. By employing strategies based on 3D Zernike descriptors and surface complementarity, we identified a structurally similar scaffold suitable for engineering. The subsequent implementation of mutational refinement highlighted the potential of combining structural information with *in silico* assays to enhance antibody complementarity toward Gal-3BP. Importantly, this dual strategy allowed us to retain favorable structural features of the native scaffold while simultaneously introducing new stabilizing contacts that improved the binding interface.

Nonetheless, we also evaluated an end-to-end pipeline that starts from naïve repertoire sequences and emulates key aspects of B-cell affinity maturation — including the consequences of V(D)J recombination and somatic hypermutation — to reshape paratopes toward a predefined epitope. Applying this workflow to BTB domain of Gal-3BP yielded a viable design candidate for a site with surface distinct from all antibody-antigen complexes with known structure, demonstrating that repertoire-based maturation can complement repositioning strategies.

Given the relevance of Gal-3BP as a therapeutic target in cancer, the antibody designs described here hold potential to interfere with tumor progression, adhesion, and metastasis upon validation in preclinical and clinical settings. Beyond their immediate therapeutic implications, this work highlights the value of computational design as a powerful complement to experimental antibody discovery. Although developed for Gal-3BP, the integrative workflow presented here is broadly applicable and provides a generalizable framework to support early-stage antibody engineering against other structurally challenging targets.

CRediT authorship contribution statement

Andrielly H.S. Costa: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Eduardo M. Gaieta:** Methodology, Data curation. **Aline O. Albuquerque:** Writing – review & editing, Methodology, Investigation, Data curation. **Julia S. Souza:** Visualization, Investigation, Data curation. **Diego S. Almeida:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Jean V. Sampaio:** Methodology, Data curation. **Patrick England:** Writing – review & editing, Methodology, Funding acquisition. **Geraldo R. Sartori:** Writing – review & editing, Supervision, Conceptualization. **João H.M. Silva:** Writing – review & editing, Supervision, Resources, Funding acquisition, Formal analysis, Conceptualization.

Ethics statement

This study used computational methods and public data. No human participants, human tissue, or animal models were involved. Therefore, ethical approval from an institutional review board or ethics committee was not required.

Funding

This research was funded by the Coordination for the Improvement of Higher Education Personnel (CAPES), an institution affiliated with the Brazilian Ministry of Education. The authors acknowledge the support of the Tripartite Fiocruz/Pasteur/USP funding scheme, and National Council for Scientific and Technological Development – CNPq, under project Universal (grant 405134/2023-1). The funding bodies did not play any role in this study's design or development, the analysis and interpretation of data, or this manuscript's writing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: João Hermínio Martins da Silva reports was provided by National Council for Scientific and Technological Development. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

All the data acquisition and analysis were performed through the computational cluster facilities of the Oswaldo Cruz Foundation – Fiocruz Ceará (Brazil), to whom we express our sincere gratitude for their invaluable support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2026.111550.

Data availability

Public dataset: AbSet, https://github.com/SFBBGroup/AbSet. Softwares used: Amber 24, http://ambermd.org/AmberMD.php; AmberTools 24, http://ambermd.org/AmberTools.php; Chimera 1.19, https://www.cgl.ucsf.edu/chimera/; Ab-SELDON, https://github.com/SFBBGroup/Ab-SELDON. The data and additional resources relevant to this study are available from the corresponding author upon request.

References

[1] S.C. Warwar, L.M. Janczewski, G.M. Rodriguez, J.D. Wayne, D.J. Bentrem, Trends in immunotherapy (IO) use and survival among patients with high-incidence stage IV cancers across the United States, J. Surg. Oncol. 131 (7) (2025 Jun 1) 1455–1466 [Internet] [cited 2025 Aug 13]. Available from:/doi/pdf/10.1002/jso.28084.

[2] N. Wellhausen, J. Baek, S.I. Gill, C.H. June, Enhancing cellular immunotherapies in cancer by engineering selective therapeutic resistance, Nat. Rev. Cancer 249 (9) (2024) 614–628 [Internet]. 2024 Jul 24 [cited 2025 Apr 30];24, https://www.nature.com/articles/s41568-024-00723-5.

[3] T. Sasaki, C. Brakebusch, J. Engel, R. Timpl, Mac-2 binding protein is a cell-adhEMBO J.esive protein of the extracellular matrix which self-assembles into ring-like structures and binds beta1 integrins, collagens and fibronectin, EMBO J [Internet] 17 (6) (1998 Mar 3) 1606 [cited 2024 Sep 3]. Available from:/pmc/articles/PMC1170508/?report=abstract.

[4] I. Cela, V.C.A. Caponio, E. Capone, M. Pinti, M. Mascitti, L. Togni, et al., LGALS3BP is a potential target of antibody-drug conjugates in oral squamous cell carcinoma, Oral Dis. 30 (4) (2024 May 1) 2039–2050 [cited 2026 Jan 27]. Available from:/doi/pdf/10.1111/odi.14719.

[5] V. Loimaranta, J. Hepojoki, O. Laaksoaho, A.T. Pulliainen, Galectin-3-binding protein: a multitask glycoprotein with innate immunity functions in viral and bacterial infections, J. Leukoc. Biol. 104 (4) (2018 Oct 1) 777–786 [cited 2024 Sep 3], https://onlinelibrary.wiley.com/doi/full/10.1002/JLB.3VMR0118-036R.

[6] F. Giansanti, E. Capone, S. Ponziani, E. Piccolo, R. Gentile, A. Lamolinara, et al., Secreted Gal-3BP is a novel promising target for non-internalizing antibody–drug conjugates, J. Contr. Release 294 (2019 Jan 28) 176–184.

[7] N. Tinari, M. D'Egidio, S. Iacobelli, M. Bowen, G. Starling, C. Seachord, et al., Identification of the tumor antigen 90K domains recognized by monoclonal antibodies SP2 and L3 and preparation and characterization of novel Anti-90K monoclonal antibodies, Biochem. Biophys. Res. Commun. 232 (2) (1997 Mar 17) 367–372.

[8] S. Hellstern, T. Sasaki, C. Fauser, A. Lustig, R. Timpl, J. Engel, Functional studies on recombinant domains of Mac-2-binding protein, J. Biol. Chem. 277 (18) (2002) 15690–15696.

[9] S.A. Müller, T. Sasaki, P. Bork, B. Wolpensinger, T. Schulthess, R. Timpl, et al., Domain organization of Mac-2 binding protein and its oligomerization to linear and ring-like structures, J. Mol. Biol. 291 (4) (1999 Aug 27) 801–813.

[10] T.W. Lin, H.T. Chang, C.H. Chen, C.H. Chen, S.W. Lin, T.L. Hsu, et al., Galectin-3 binding protein and Galectin-1 interaction in breast cancer cell aggregation and metastasis, J. Am. Chem. Soc. 137 (30) (2015 Aug 5) 9685–9693 [cited 2025 Sep 22]. Available from:/doi/pdf/10.1021/jacs.5b04744?ref=article_openPDF.

[11] J. Cheng, T. Liang, X.Q. Xie, Z. Feng, L. Meng, A new era of antibody discovery: an in-depth review of AI-driven approaches, Drug Discov. Today 29 (6) (2024 Jun 1)

103984 [Internet] [cited 2025 Dec 8], https://www.sciencedirect.com/science/article/abs/pii/S1359644624001090?casa_token=nQUGszgpRVkAAAAA:LzeO0uZRm-kUiakIYuxk_glKZeblvFaJvDbsAH65Gc4z1CYmSW4YLFHgJvVAf-49RUGG8dCHTzs.

[12] A. Sarker, A.S. Rathore, M.F. Khalid, R.D. Gupta, Structure-guided affinity maturation of a single-chain variable fragment antibody against the Fu-bc epitope of the dengue virus envelope protein, J. Biol. Chem. 298 (4) (2022 Apr 1) 101772 [cited 2025 Dec 9], https://pmc.ncbi.nlm.nih.gov/articles/PMC8956951/.

[13] W. Ye, X. Liu, R. He, L. Gou, M. Lu, G. Yang, et al., Improving antibody affinity through in vitro mutagenesis in complementarity determining regions, J. Biomed. Res. 36 (3) (2022 May 1) 155 [cited 2025 Dec 8], https://pmc.ncbi.nlm.nih.gov/articles/PMC9179109/.

[14] A. Roy, S. Nair, N. Sen, N. Soni, M.S. Madhusudhan, In silico methods for design of biological therapeutics, Methods 131 (2017 Dec 1) 33–65.

[15] M.F. Erasmus, L. Spector, F. Ferrara, R. DiNiro, T.J. Pohl, K. Perea-Schmittle, et al., AIntibody: an experimentally validated in silico antibody discovery design challenge, Nat. Biotechnol. 42 (11) (2024 Nov 1) 1637–1642 [cited 2025 Sep 22], https://www.nature.com/articles/s41587-024-02469-9.

[16] J.R. Devasani, G. Guntuku, P. sarabu, M.K.K. Muthyala, M.S. Palla, V. M. Subrahmanyam, Integrative and emerging models in antibody research: a comprehensive review, Antib. Ther. 8 (4) (2025 Aug 28) 317–335, https://doi.org/10.1093/abt/tbaf018 [cited 2025 Sep 22].

[17] P. Rawat, D. Sharma, A. Srivastava, V. Janakiraman, M.M. Gromiha, Exploring antibody repurposing for COVID-19: beyond presumed roles of therapeutic antibodies, Sci. Rep. 11 (1) (2021) 10220, 111 [Internet]. 2021 May 13 [cited 2025 Dec 7], https://www.nature.com/articles/s41598-021-89621-6.

[18] C. Maguire, C. Wang, A. Ramasamy, C. Fonken, B. Morse, N. Lopez, et al., Molecular mimicry as a mechanism of viral immune evasion and autoimmunity, Nat. Commun. 151 (1) (2024) 9403 [Internet]. 2024 Oct 30 [cited 2025 Dec 7];15, https://www.nature.com/articles/s41467-024-53658-8.

[19] S. Cherian, P. Shil, A.C. Mishra, Antigen-antibody docking reveals the molecular basis for cross-reactivity of the 1918 and 2009 influenza A/H1N1 pandemic viruses, Bioinformation 6 (1) (2011 Mar 2) 35 [Internet] [cited 2025 Dec 7], https://pmc.ncbi.nlm.nih.gov/articles/PMC3064850/.

[20] J. Qi, C. Feng, Y. Shi, J. Yang, F. Zhang, G. Li, et al., FP-Zernike: an open-source structural database construction toolkit for fast structure retrieval, Genom. Proteom. Bioinform. 22 (1) (2024 May 9), https://doi.org/10.1093/gpbjnl/qzae007 [Internet] [cited 2025 Sep 22].

[21] A. Bateman, M.J. Martin, S. Orchard, M. Magrane, R. Agivetova, S. Ahmad, et al., UniProt: the universal protein knowledgebase in 2021, Nucl. Acids Res. 49 (2021 Jan 8) [cited 2021 Nov 30](D1):D480–9, https://academic.oup.com/nar/article/49/D1/D480/6006196.

[22] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, et al., Highly accurate protein structure prediction with AlphaFold, Nat 596 (7873) (2021) 5967873 [Internet]. 2021 Jul 15 [cited 2024 Sep 10]:583–9, https://www.nature.com/articles/s41586-021-03819-2.

[23] K. Olechnovič, C.D.S. Venclovas, VoroMQA web server for assessing three-dimensional structures of proteins and protein complexes, Nucl. Acids Res. 47 (W1) (2019 Jul 7) W437 [cited 2024 Sep 10] Available from:/pmc/articles/PMC6602437/.

[24] Y.X. Tsai, N.E. Chang, K. Reuter, H.T. Chang, T.J. Yang, S. von Bülow, et al., Rapid simulation of glycoprotein structures by grafting and steric exclusion of glycan conformer libraries, Cell [Internet] 187 (5) (2024 Feb 29) 1296–1311.e26 [cited 2025 Sep 22], https://www.cell.com/action/showFullText?pii=S0092867424001016.

[25] D.A. Case, H.M. Aktulga, K. Belfon, D.S. Cerutti, G.A. Cisneros, V.W.D. Cruzeiro, et al., AmberTools, J. Chem. Inf. Model. 63 (20) (2023 Oct 23) 6183–6191 [Internet] [cited 2025 Dec 8]. Available from:/doi/pdf/10.1021/acs.jcim.3c01153?ref=article_openPDF.

[26] M. Zeiler, W.L. Straube, E. Lundberg, M. Uhlen, M. Mann, A protein epitope signature tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines, Mol. Cell. Proteomics 11 (3) (2012 Mar 1) [Internet] [cited 2025 Sep 22] O111.009613, https://www.sciencedirect.com/science/article/pii/S1535947620305053?via%3Dihub.

[27] D.S. Almeida, M.V. Almeida, J.V. Sampaio, E.M. Gaieta, A.H.S. Costa, F.F. A. Rabelo, et al., AbSet: a standardized data set of antibody structures for machine learning applications, J. Chem. Inf. Model. 65 (10) (2025 May 26) 4767–4774 [Internet] [cited 2025 Sep 22]. Available from:/doi/pdf/10.1021/acs.jcim.5c00410?ref=article_openPDF.

[28] K.R. Abhinandan, A.C.R. Martin, Analysis and improvements to kabat and structurally correct numbering of antibody variable domains, Mol. Immunol. 45 (14) (2008 Aug 1) [cited 2026 Jan 27]:3832–9, https://www.sciencedirect.com/science/article/abs/pii/S0161589008002046.

[29] R.V. Honorato, M.E. Trellet, B. Jiménez-García, J.J. Schaarschmidt, M. Giulini, V. Reys, et al., The HADDOCK2.4 web server for integrative modeling of biomolecular complexes, Nat. Protoc. 19 (11) (2024 Nov 1) 3219–3241 [Internet] [cited 2025 Sep 22], https://www.nature.com/articles/s41596-024-01011-0.

[30] R.F. Alford, A. Leaver-Fay, J.R. Jeliazkov, M.J. O'Meara, F.P. DiMaio, H. Park, et al., The rosetta all-atom energy function for macromolecular modeling and design, J. Chem. Theor. Comput. 13 (6) (2017 Jun 13) 3031–3048 [cited 2025 Sep 22]. Available from:/doi/pdf/10.1021/acs.jctc.7b00125?ref=article_openPDF.

[31] J. Adolf-Bryfogle, Q. Xu, B. North, A. Lehmann, R.L. Dunbrack, PyIgClassify: a database of antibody CDR structural classifications, Nucl. Acids Res. 43 (D1) (2015 Jan 28), https://doi.org/10.1093/nar/gku1106 [cited 2025 Sep 22]:D432–8.

[32] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The protein data bank, Nucl. Acids Res. 28 (1) (2000 Jan 1) 235–242 [cited 2019 Nov 6], https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/28.1.235.

[33] T.H. Olsen, F. Boyles, C.M. Deane, Observed antibody space: a diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences, Protein Sci. 31 (1) (2022 Jan 1) 141–146 [cited 2025 Sep 22]. Available from:/doi/pdf/10.1002/pro.4205.

[34] J. Dunbar, C.M. Deane, ANARCI: antigen receptor numbering and receptor classification, Bioinformatics 32 (2) (2016 Jan 15) 298–300, https://doi.org/10.1093/bioinformatics/btv552 [Internet] [cited 2025 Sep 22].

[35] T.H. Olsen, I.H. Moal, C.M. Deane, AbLang: an antibody language model for completing antibody sequences, Bioinforma Adv. 2 (1) (2022 Jan 10), https://doi.org/10.1093/bioadv/vbac046 [Internet] [cited 2025 Sep 22].

[36] J.V. Sampaio, A.H.S. Costa, A.O. Albuquerque, J.S. Souza, D.S. Almeida, E. M. Gaieta, et al., Ab-SELDON: leveraging diversity data for an efficient automated computational pipeline for antibody design, J. Chem. Inf. Model. 66 (3) (2026 Jan 20) 1895–1905. Available from:/doi/pdf/10.1021/acs.jcim.5c01924?ref=article_openPDF.

[37] B. Abanades, W.K. Wong, F. Boyles, G. Georges, A. Bujotzek, C.M. Deane, ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins, Commun. Biol. 61 (1) (2023) 1–8, 2023 May 29 [Internet] [cited 2024 Sep 10];6, https://www.nature.com/articles/s42003-023-04927-7.

[38] B. Chaves, G.R. Sartori, D.C.A. Vasconcelos, W. Savino, E.R. Caffarena, V. Cotta-De-Almeida, et al., Guidelines to predict binding poses of antibody-integrin complexes, ACS Omega 5 (27) (2020 Jul 14) 16379–16385 [cited 2021 Dec 7], https://pubs.acs.org/doi/full/10.1021/acsomega.0c00226.

[39] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the amber biomolecular simulation package, Wiley Interdiscip. Rev. Comput. Mol. Sci. 3 (2) (2013 Mar 1) 198–210 [cited 2019 Nov 4]. Available from:.

[40] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theor. Comput. 11 (8) (2015 Aug 11) 3696–3713 [cited 2019 Nov 4], http://www.ncbi.nlm.nih.gov/pubmed/26574453.

[41] K.W. Lexa, G.B. Goh, H.A. Carlson, ParameterChoice matters: validating probe parametersfor use in mixed-solvent simulations, J. Chem. Inf. Model. 54 (8) (2014 Aug 25) 2190 [Internet] [cited 2021 Nov 13]. Available from:/pmc/articles/PMC4144759/.

[42] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. González-Outeiriño, C.R. Daniels, B.L. Foley, et al., GLYCAM06: a generalizable biomolecular force field. Carbohydrates, J. Comput. Chem. 29 (4) (2008 Mar 1) 622–655 [cited 2024 Sep 10], https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.20820.

[43] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, et al., Improvements to the APBS biomolecular solvation software suite, Protein Sci. 27 (1) (2018 Jan 1) 112–128 [Internet] [cited 2024 Sep 10]. Available from:/doi/pdf/10.1002/pro.3280.

[44] L. Fallon, K.A.A. Belfon, L. Raguette, Y. Wang, D. Stepanenko, A. Cuomo, et al., Free energy landscapes from SARS-CoV-2 spike glycoprotein simulations suggest that RBD opening can be modulated via interactions in an allosteric pocket, J. Am. Chem. Soc. 143 (30) (2021 Aug 4) 11349–11360 [Internet] [cited 2024 Sep 10], https://pubs.acs.org/doi/abs/10.1021/jacs.1c00556.

[45] Relaxation of explicit water systems. https://ambermd.org/tutorials/basic/tutorial13/index.php#Run_scr.

[46] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald, J. Chem. Theor. Comput. 9 (9) (2013 Sep 10) 3878–3888 [cited 2021 Nov 12], https://pubs.acs.org/doi/abs/10.1021/ct400314y.

[47] C.W. Hopkins, S. Le Grand, R.C. Walker, A.E. Roitberg, Long-time-step molecular dynamics through hydrogen mass repartitioning, J. Chem. Theor. Comput. 11 (4) (2015) 1864–1874 [cited 2021 Nov 12], https://pubs.acs.org/doi/abs/10.1021/ct5010406.

[48] F. Radom, A. Plückthun, E. Paci, Assessment of ab initio models of protein complexes by molecular dynamics, PLOS Comput. Biol. 14 (6) (2018 Jun 1) e1006182 [cited 2021 Dec 6], https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006182.

[49] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theor. Comput. 9 (7) (2013 Jul 9) 3084–3095 [cited 2022 Feb 3], https://pubs.acs.org/doi/abs/10.1021/ct400341p.

[50] J. Adolf-Bryfogle, O. Kalyuzhniy, M. Kubitz, B.D. Weitzner, X. Hu, Y. Adachi, et al., RosettaAntibodyDesign (RAbD): a general framework for computational antibody design, PLoS Comput. Biol. 14 (4) (2018 Apr 1) e1006112 [cited 2025 Sep 22], https://pmc.ncbi.nlm.nih.gov/articles/PMC5942852/.

[51] D. Prihoda, J. Maamary, A. Waight, V. Juan, L. Fayadat-Dilman, D. Svozil, et al., BioPhi: a platform for antibody design, humanization, and humanness evaluation based on natural antibody repertoires and deep learning, mAbs 14 (1) (2022 Dec 31) [Internet] [cited 2026 Jan 22], https://www.tandfonline.com/doi/pdf/10.1080/19420862.2021.2020203.

[52] J. Wang, P.R. Arantes, A. Bhattarai, R.V. Hsu, S. Pawnikar, Y.M. Huang, et al., Gaussian accelerated molecular dynamics: principles and applications, Wiley Interdiscip. Rev. Comput. Mol. Sci. 11 (5) (2021 Sep 1) e1521 [cited 2021 Nov 30], https://onlinelibrary.wiley.com/doi/full/10.1002/wcms.1521.

[53] A. Shkurti, R. Goni, P. Andrio, E. Breitmoser, I. Bethune, M. Orozco, et al., pyPcazip: a PCA-Based toolkit for compression and analysis of molecular simulation data, SoftwareX 5 (2016 Jan 1) 44–50 [Internet] [cited 2019 Nov 4], https://www.sciencedirect.com/science/article/pii/S2352711016300036.

[54] Y. Miao, W. Sinko, L. Pierce, D. Bucher, R.C. Walker, J.A. McCammon, Improved reweighting of accelerated molecular dynamics simulations for free energy calculation, J. Chem. Theor. Comput. 10 (7) (2014 Jul 8) 2677–2689 [cited 2025 Jun 16]. Available from:/doi/pdf/10.1021/ct500090q.
[55] D. Kuroda, J.J. Gray, Shape complementarity and hydrogen bond preferences in protein–protein interfaces: implications for antibody modeling and protein–protein docking, Bioinformatics 32 (16) (2016 Aug 15), https://doi.org/10.1093/bioinformatics/btw197 [Internet] [cited 2025 Dec 8]:2451–6.
[56] C.D. Murin, P. Gilchuk, P.A. Ilinykh, K. Huang, N. Kuzmina, X. Shen, et al., Convergence of a common solution for broad ebolavirus neutralization by glycan cap-directed human antibodies, Cell Rep. 35 (2) (2021 Apr 13) 108984 [Internet] [cited 2026 Jan 22], https://www.sciencedirect.com/science/article/pii/S2211124721002989?via%3Dihub.
[57] A.J. MacLean, L.P. Deimel, P. Zhou, M.A. ElTanbouly, J. Merkenschlager, V. Ramos, et al., Affinity maturation of antibody responses is mediated by differential plasma cell proliferation, Science 387 (6732) (2024 Jan 24) 413 [Internet] [cited 2026 Jan 27], https://pmc.ncbi.nlm.nih.gov/articles/PMC11938350/.
[58] Lodermeyer V, Ssebyatika G, Passos V, Ponnurangam A, Malassa A, Ewald E. Crossm the Antiviral Activity of the Cellular Glycoprotein LGALS3BP/90K Is Species Specific.
[59] E. Capone, S. Iacobelli, G. Sala, Role of galectin 3 binding protein in cancer progression: a potential novel therapeutic target, J. Transl. Med. 19 (1) (2021) 1–18 [Internet] [cited 2022 Dec 8], https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-021-03085-w.
[60] C. Peng, W. Ni, Q. Liu, G. Hu, W. Zheng, A comprehensive benchmarking of the AlphaFold3 for predicting biomacromolecules and their interactions, Briefings Bioinf. 26 (6) (2025 Nov 1) [Internet] [cited 2026 Jan 27], https://pubmed.ncbi.nlm.nih.gov/41313605/.
[61] Z. Yang, X. Zeng, Y. Zhao, R. Chen, AlphaFold2 and its applications in the fields of biology and medicine, Signal Transduct. Targeted Ther. 81 (1) (2023) 115 [Internet]. 2023 Mar 14 [cited 2026 Jan 27];8, https://www.nature.com/articles/s41392-023-01381-z.
[62] K.M. Ruff, R.V. Pappu, AlphaFold and implications for intrinsically disordered proteins, J. Mol. Biol. 433 (20) (2021 Oct 1) 167208 [cited 2026 Jan 27], https://www.sciencedirect.com/science/article/pii/S0022283621004411?via%3Dihub.
[63] D. Wang, Z. Zhang, J. Baudys, C. Haynes, S.H. Osman, B. Zhou, et al., Enhanced surface accessibility of SARS-CoV-2 omicron spike protein due to an altered glycosylation profile, ACS Infect. Dis. 10 (6) (2024 Jun 14) 2032–2046 [cited 2025 Sep 22]. Available from:/doi/pdf/10.1021/acsinfecdis.4c00015?ref=article_openPDF.
[64] P.J. Greisen, L. Yi, R. Zhou, J. Zhou, E. Johansson, T. Dong, et al., Computational design of N-linked glycans for high throughput epitope profiling, Protein Sci. 32 (10) (2023 Oct 1).
[65] F. Humphries, L. Shmuel-Galia, Z. Jiang, J.Y. Zhou, L. Barasa, S. Mondal, et al., Targeting STING oligomerization with small-molecule inhibitors, Proc. Natl. Acad. Sci. U. S. A. 120 (33) (2023 Aug 15) e2305420120 [Internet] [cited 2025 Dec 8]. Available from:/doi/pdf/10.1073/pnas.2305420120?download=true.
[66] M. Ahmed, Y. Goldgur, J. Hu, H.F. Guo, N.K.V. Cheung, In silico driven redesign of a clinically relevant antibody for the treatment of GD2 positive tumors, PLoS One 8 (5) (2013 May 16) e63359 [cited 2026 Jan 27], https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0063359.
[67] S.R. Krishnan, D. Sharma, Y. Nazeer, M. Bose, T. Rajkumar, G. Jayaraman, et al., rAbDesFlow: a novel workflow for computational recombinant antibody design for healthcare engineering, Antib. Ther. 7 (3) (2024 Jul 1) 256 [cited 2026 Jan 27], https://pmc.ncbi.nlm.nih.gov/articles/PMC11384895/.
[68] A. De Lauro, L. Di Rienzo, M. Miotto, P.P. Olimpieri, E. Milanetti, G. Ruocco, Shape complementarity optimization of antibody–antigen interfaces: the application to SARS-CoV-2 spike protein, Front. Mol. Biosci. 9 (2022 May 20) 874296 [cited 2025 Dec 8], www.frontiersin.org.
[69] Z.A. Rollins, T. Widatalla, A.C. Cheng, E. Metwally, AbMelt: learning antibody thermostability from molecular dynamics, Biophys. J. 123 (17) (2024 Sep 3) [cited 2025 Dec 8]:2921–33, https://www.sciencedirect.com/science/article/pii/S0006349524003850#sec4.
[70] G.J. Bekker, B. Ma, N. Kamiya, Thermal stability of single-domain antibodies estimated by molecular dynamics simulations, Protein Sci. 28 (2) (2019 Feb 1) 429–438 [cited 2025 Dec 8]. Available from:/doi/pdf/10.1002/pro.3546.
[71] Kadir FFN. Abdul, M.A. Che Nordin, A.N. Shuid, Omar MT. Che, Modified heated CGMD simulations for discovering stable docked conformations of BiTE antibody against CD3 and CD117/c-kit, Mol. Simul. 50 (14) (2024 Sep 21) 1019–1038 [cited 2026 Jan 27], https://www.tandfonline.com/doi/abs/10.1080/08927022.2024.2378833.
[72] D. Roccatano, Principal component analysis of molecular dynamic trajectories: concepts, tools, and applications, Wiley Interdiscip. Rev. Comput. Mol. Sci. 15 (6) (2025 Nov 1) e70060 [cited 2026 Jan 27]. Available from:/doi/pdf/10.1002/wcms.70060.
[73] J. Chen, J. Wang, W. Yang, L. Zhao, J. Su, Activity regulation and conformation response of janus kinase 3 mediated by phosphorylation: exploration from correlation network analysis and markov model, J. Chem. Inf. Model. 65 (8) (2025 Apr 28) 4189–4205 [Internet] [cited 2026 Jan 27], https://pubs.acs.org/doi/abs/10.1021/acs.jcim.5c00096.
[74] J. Chen, J. Wang, W. Yang, L. Zhao, G. Hu, Conformations of KRAS4B affected by its partner binding and G12C mutation: insights from GaMD trajectory-image transformation-based deep learning, J. Chem. Inf. Model. 64 (17) (2024 Sep 9) 6880–6898 [Internet] [cited 2026 Jan 27], https://pubs.acs.org/doi/abs/10.1021/acs.jcim.4c01174.
[75] Y. Miao, V.A. Feher, J.A. Mccammon, Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. https://pubs.acs.org/sharingguidelines, 2015.
[76] N.R. Bennett, J.L. Watson, R.J. Ragotte, A.J. Borst, D.L. See, C. Weidle, et al., Atomically accurate de novo design of antibodies with RFdiffusion, Nat (2025) 1–11, 2025 Nov 5 [Internet] [cited 2025 Dec 9], https://www.nature.com/articles/s41586-025-09721-5.
[77] X. Yang, H. Duan, X. Liu, X. Zhang, S. Pan, F. Zhang, et al., Broad sarbecovirus neutralizing antibodies obtained by computational design and synthetic library screening, J. Virol. 97 (7) (2023 Jul 27) [cited 2025 Sep 20]. Available from:/doi/pdf/10.1128/jvi.00610-23?download=true.
[78] X. Zhou, S. Kong, A. Maker, S.G. Remesh, K.K. Leung, K.A. Verba, et al., Antibody discovery identifies regulatory mechanisms of protein arginine deiminase 4, Nat. Chem. Biol. 20 (6) (2024 Jun 11) 742–750 [Internet] [cited 2025 Sep 20], https://www.nature.com/articles/s41589-023-01535-8.

Glossary

Gal-3BP: Galectin-3 binding protein
SRCR: Scavenger receptor cysteine-rich
BTB/POZ: Broad-Complex, Tramtrack, Bric-à-brac/Poxvirus and Zinc finger
BACK: C-terminal Kelch
ADCs: Antibody–drug conjugates
CDRs: Complementarity-determining regions
mAB: Monoclonal antibodies
PDB: Protein Data Bank
MD: Molecular dynamics simulations
hMD: Heated molecular dynamics simulations
GaMD: Gaussian Accelerated Molecular Dynamics simulations
PMEMD: Partial Mesh Ewald Molecular Dynamics
iRMSD: interface Root Mean Square Deviation
OAS: Observed Antibody Space
RAbD: RosettaAntibodyDesign