

# Binding free energy analysis of galectin-3 natural ligands and synthetic inhibitors

Luke Newman<sup>1,2</sup> | Valerie Vaissier Welborn<sup>1,2</sup> 

<sup>1</sup>Department of Chemistry, Virginia Tech, Blacksburg, Virginia, USA

<sup>2</sup>Macromolecules Innovation Institute, Virginia Tech, Blacksburg, Virginia, USA

## Correspondence

Valerie Vaissier Welborn, Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA.

Email: [vwelborn@vt.edu](mailto:vwelborn@vt.edu)

## Funding information

National Science Foundation, Grant/Award Number: DMR-1933525

Review Editor: Lynn Kamertlin

## Abstract

Galectin-3–ligand complexes are characterized by halogen,  $\sigma$ -hole bonds, hydrogen bonds, cation- $\pi$  and CH- $\pi$  interactions. Here, we model these non-covalent interactions with the AMOEBA polarizable force field and conduct an absolute binding free energy analysis on leading galectin-3 inhibitors. Synthetic drug molecules GB0139, GB1107, and GB1211 were estimated to have binding free energies of  $-4.3$ ,  $-6.7$ , and  $-9.5$  kcal/mol respectively. This compares to  $-0.3$  and  $1.4$  kcal/mol for the natural ligands, N-acetyllactosamine type 1 and type 2, respectively. We calculated the electric fields projected along key bonds in each ligand to further rationalize these results. We find that while the hydroxyl groups of the natural ligands interact reasonably well with residues in galectin-3's binding pocket, structural dynamics weaken the binding pose and favor interactions with water, sometimes yielding to dissociation. In contrast, the more favorable binding energy of GB1211, leading inhibitor in clinical studies, is associated with strong and constant electric fields across the bonds investigated, suggesting a stiffer binding pose with a stabilizing  $\sigma$ -hole interaction.

## KEYWORDS

alchemical, AMOEBA, binding free energy, galectin, galectin-3, inhibitor, ligand, molecular dynamics

## 1 | INTRODUCTION

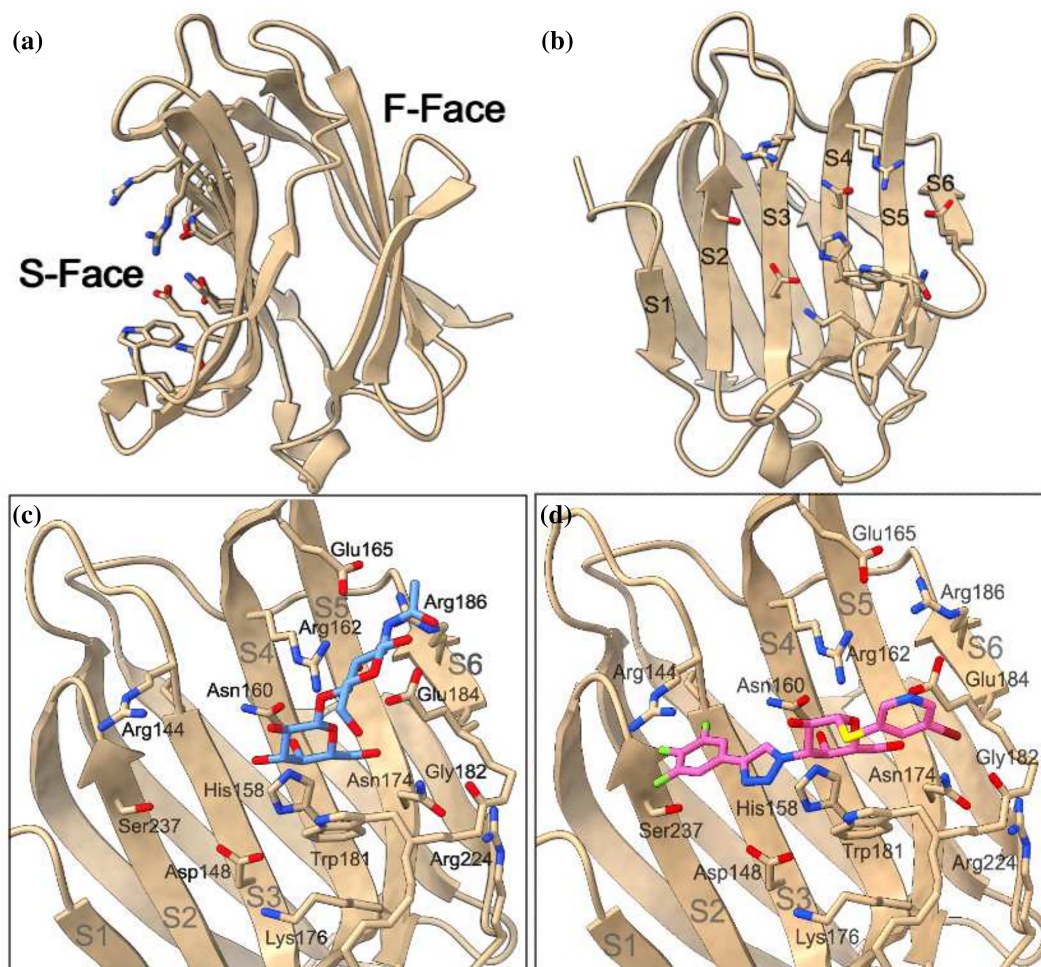
Galectins are a family of soluble proteins that bind  $\beta$ -galactose-containing carbohydrates and participate in cell–cell interactions, cell-matrix interactions, and transmembrane signaling (Demetriou et al., 2001; Markowska et al., 2011; Nabi et al., 2015). There are 15 known galectins in mammals and 12 known galectin genes in humans (Cummings et al., 2022; Verkerke et al., 2022). Each galectin contains one or two carbohydrate recognition domains (CRDs) that non-covalently bind  $\beta$ -galactose (Leffler et al., 2002). Among these, galectin-3 (Gal-3) has been extensively researched due to its overexpression in numerous disorders (Sciacchitano et al., 2018), including inflammation (Henderson & Sethi, 2009; Rabinovich &

Toscano, 2009; Thiemann & Baum, 2016), cardiovascular diseases (Boer et al., 2009; Lok et al., 2010; Sharma et al., 2004), cancer (Califice et al., 2004; Liu & Rabinovich, 2005; Ruvolo, 2016), and fibrosis (Dang et al., 2012; Henderson et al., 2006; MacKinnon et al., 2012; Nishi et al., 2007). Because of this strong association with various pathologies, Gal-3 has become a diagnostic biomarker (Dong et al., 2018) and, more recently, a target for therapeutic inhibition (Ahmed et al., 2023; Bouffette et al., 2023). Ideally, synthetic inhibitors would leverage non-covalent interactions in the CRD to out-compete natural ligands and suppress functionality to treat disease symptoms.

Gal-3 is the only chimera-type galectin with a long flexible N-terminus tail attached to the highly conserved and globular CRD. The CRD consists of approximately

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.



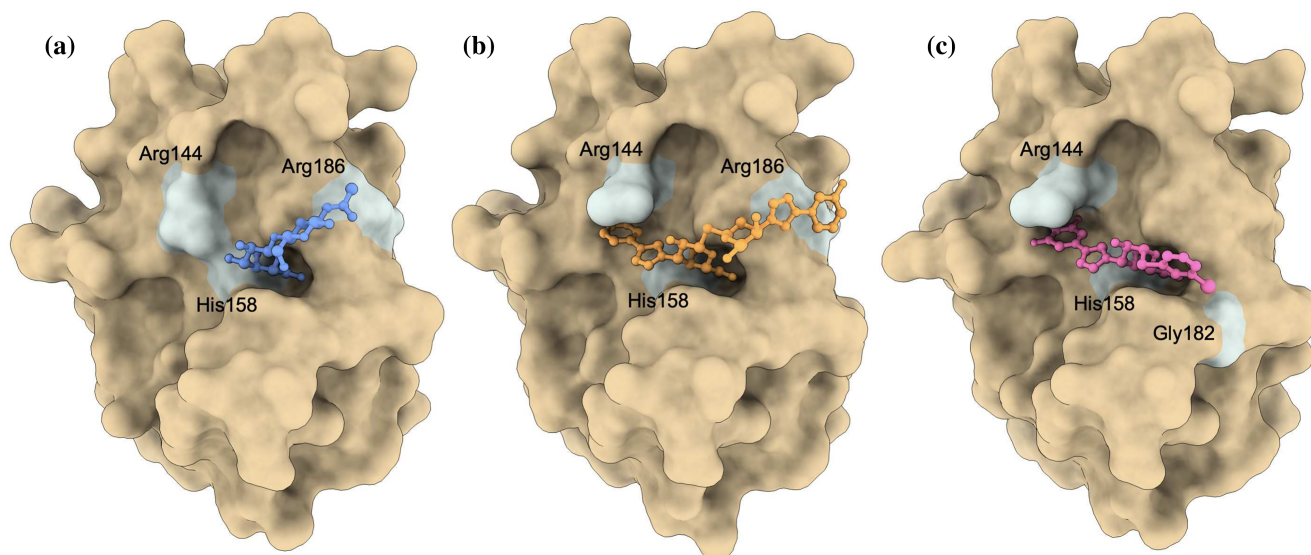
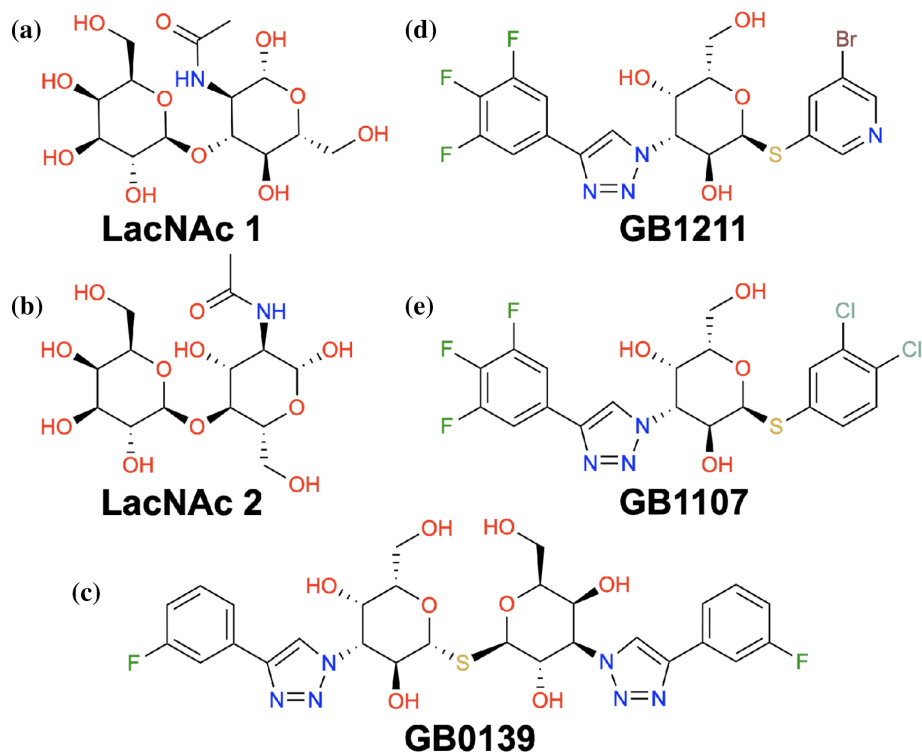
**FIGURE 1** (a) Visualization of the Gal-3 CRD with the anti-parallel beta sheets termed the S- and F-faces. Polar residues in the binding pocket on the S face are emphasized. (b) Direct view of the S-face, with the  $\beta$  strands labeled S1–S6. (c) Binding pose of Gal-3's natural ligand, LacNAc type 2 (PDB ID: 1KJL). (d) Binding pose of Gal-3's synthetic inhibitor, GB1211 (PDB ID: 7ZQX).  $\beta$  strands and nearby residues have been labeled. Hydrogens have been omitted for visual clarity.

135 residues that form two antiparallel beta sheets, termed the S-face and F-face (Figure 1a). Ligands bind to the slightly concave S-face, a groove spanning six  $\beta$ -strands (S1–S6 in Figure 1b) and containing a high density of polar residues (Seetharaman et al., 1998; Leffler et al., 2002). The  $\beta$ -galactose unit in Gal-3's natural ligands forms 2–5 hydrogen bonds and exhibits favorable van der Waals interactions with the S4 and S5  $\beta$ -strands. Diehl et al. also showed the importance of CH- $\pi$  bonding to the tryptophan residue directly below the  $\beta$ -galactose unit in the binding region, Trp181 (Diehl et al., 2024). Other units, such as glucosamine in N-acetyllactosamine (LacNAc, see Figure 2a,b), interact with residues of the S6  $\beta$ -strand or the solvent (Figure 1c).

Synthetic Gal-3 inhibitors have commonly retained the  $\beta$ -galactose unit while introducing side chains that have more reach to interact with residues on the S2 and S3  $\beta$ -strands (Figure 1d). For example, aromatic rings have been added to facilitate interactions with

Arg144, a key feature of these synthetic inhibitors (Sörme et al., 2005). One of the early successful Gal-3 inhibitors was a thio-disaccharide with substituted aromatic rings, named GB0139 (formerly TD139, see Figure 2c), that showed promise in reducing inflammation and fibrosis in mice models (Delaine et al., 2016; MacKinnon et al., 2012). GB0139 was designed to maintain the established hydrogen bonds of the  $\beta$ -galactoside unit with His158 and Asn174, while enhancing affinity with Arg144 and Arg186 through two outer aromatic rings. In recent years, a second generation of Gal-3 inhibitors was established with 1,3-substituted  $\alpha$ -D-monogalactopyranosides (Zetterberg et al., 2018, 2022). Affinity for these  $\alpha$  anomers is not shared by other galectins, enabling greater Gal-3 selectivity. Further, 1,3-substituted  $\alpha$ -D-monogalactopyranosides were reported to maintain the potent affinity of the previous synthetic ligands such as GB0139 while exhibiting a lower polar surface area, which allowed for oral administration (Zetterberg et al., 2022). Example molecules from

**FIGURE 2** Small-molecule Gal-3 ligands and inhibitors: (a) natural ligand LacNAc type 1, (b) synthetic inhibitor GB1211, (c) natural ligand LacNAc type 2, (d) synthetic inhibitor GB1107. Both natural ligands and synthetic inhibitors share the galactose unit.



**FIGURE 3** Representation of the Gal-3 binding pose of (a) disaccharides such as natural ligands LacNAc 1 and 2, (b) thiodigalactosides, such as the first generation synthetic inhibitor GB0139, (c) monogalactopyranosides, such as the second generation synthetic inhibitors GB1211 and GB1107. Generated from PDB crystal structures ID 1KJL, 5H9P, and 7ZQX respectively.

this second generation of Gal-3 inhibitors are GB1211 and GB1107 (Figure 2d,e). GB1211 and GB1107 are compounds 11d and 11b from Galecto Bio, with GB1211 currently in clinical studies (Aslanis et al., 2023). Interestingly, these inhibitors also introduced halogen atoms that are thought to engage in a highly directional  $\sigma$ -hole bond with Gly182 (Verteramo et al., 2024; Zetterberg et al., 2022). The binding poses of disaccharides, thiodigalactosides and monogalactopyranosides

in Gal-3's CRD, shown in Figure 3, highlights the difference in size and reach between natural and synthetic ligands.

Despite encouraging pharmacokinetics results for GB1211, a clear understanding of how these new compounds bind to Gal-3's CRD is missing. A detailed molecular picture of Gal-3's interactions with 1,3-substituted  $\alpha$ -d-monogalactopyranosides, that goes beyond the analysis of single structures and

crystallographic data, would enable further breakthroughs in the design of potent and selective inhibitors. In this work, we investigate an ensemble of binding configurations of two natural ligands LacNAc type 1 and 2 and the three aforementioned synthetic inhibitors GB0139, GB1211, and GB1107, using molecular dynamics (MD) simulations, binding free energy calculations, and electric field analysis. We use the polarizable AMOEBA force field (Shi et al., 2013) to better account for anisotropic non-covalent interactions such as CH- $\pi$ , cation- $\pi$ , and  $\sigma$ -hole interactions. We find that GB0139 interacts more favorably with Gal-3 than the natural ligands but still exhibits a relatively strong affinity for water. The second-generation inhibitors GB1211 and GB1107 bind Gal-3 even better than GB0139. Interestingly while both GB1211 and GB1107 interact with similar residues, GB1211 has a stiffer binding pose that enables lasting interactions with Gal-3's CRD and results in an optimized binding energy.

## 2 | RESULTS AND DISCUSSION

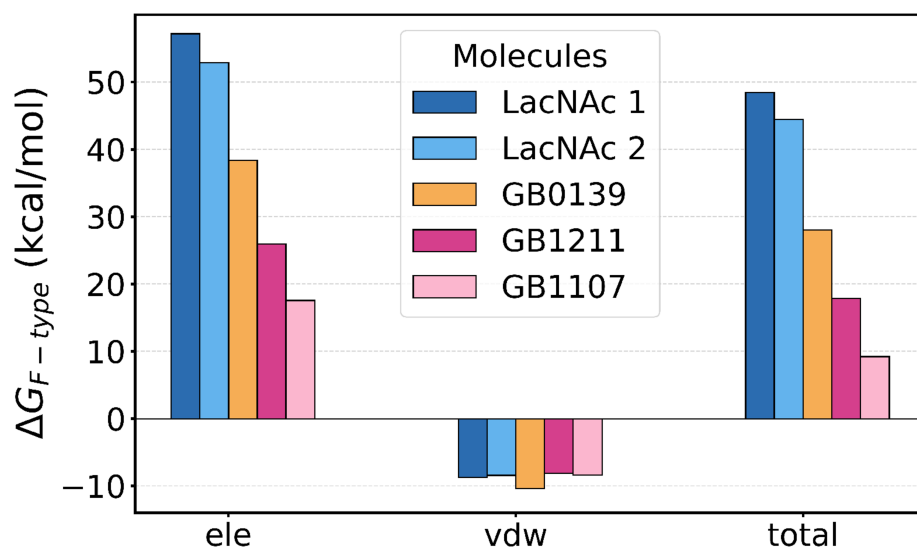
### 2.1 | Natural ligands and synthetic inhibitors in water

We present in Figure 4 the magnitude of the free energy difference ( $\Delta G_{F\text{-total}}$ ) associated with the removal of the ligands from water (i.e., the ligand desolvation free energy). The electrostatic ( $\Delta G_{F\text{-ele}}$ ) and van der Waals ( $\Delta G_{F\text{-vdw}}$ ) components of  $\Delta G_{F\text{-total}}$  are also shown. Note that, at first approximation,  $\Delta G_{F\text{-total}}$  can be considered as the opposite of the hydration free energy because the ligand intramolecular electrostatics are expected to be negligible. We observe that the natural ligands, LacNAc 1 and 2, have a much greater affinity for water, with a  $\Delta G_{F\text{-total}}$  more than twice that of the second-generation synthetic inhibitors GB1211 and

GB1107. Meanwhile, the first-generation thio-disaccharide inhibitor GB0139 proved to be a middle ground between the two groups. We also note that  $\Delta G_{F\text{-vdw}}$  is almost identical for all ligands such that the differences in  $\Delta G_{F\text{-total}}$  are primarily due to variations in  $\Delta G_{F\text{-ele}}$ . Interestingly, GB1211 shows almost twice the affinity for water compared to GB1107 ( $\Delta G_{F\text{-total}} = 17.9$  and 9.2 kcal/mol, respectively). To further rationalize these trends, we computed the electric fields generated by the water molecules onto key bonds in each ligand: the three common hydroxyl bonds (galactose O<sub>2</sub>-H, O<sub>4</sub>-H and O<sub>6</sub>-H), the unique carbon-halogen bonds (C-X), and the central carbon-fluorine bond (C-F).

The electric field projections are presented in Table 1. The electric field experienced by the galactose bonds O<sub>2</sub>-H, O<sub>4</sub>-H, and O<sub>6</sub>-H in LacNAc 1 and 2 ranges from 130 to 165 MV/cm. In this case, the electric field is oriented from O to H such that the electrons are favored to move in the opposite direction, from H to O. This indicates a favorable interaction with water where H is pulled in by the water molecules, mimicking the initiation of a deprotonation event.

The electric fields experienced by these hydroxyl bonds in GB0139 are comparable to those of the natural ligands. However, O<sub>2</sub>-H in GB1211 and GB1107 experiences electric fields stronger by about 10 MV/cm than O<sub>2</sub>-H in the other compounds. A similar observation can be made for O<sub>6</sub>-H in GB1211. Further, O<sub>4</sub>-H is subject to a much stronger (weaker) field in GB1211 (GB1107). This partially explains the higher water affinity of GB1211 compared to GB1107. In the case of the C-X bonds, negative electric field projections (i.e., an electric field oriented from X to C) indicate that electrons are favored to move towards the halogen atom. Therefore, the  $\sim -15$  MV/cm electric field projections along the C-F bond indicate that the electrons are favored to move outward, towards the F atom, increasing the electronegativity of these functional groups

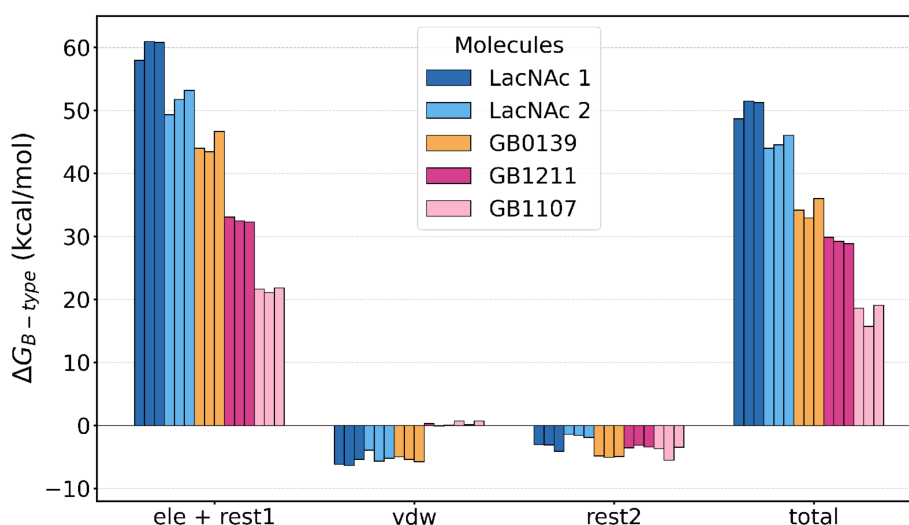


**FIGURE 4** Free energy differences along the path characterizing the free state (see Figure 7) of the ligands LacNAc 1, LacNAc 2, GB0139, GB1211 and GB1107. As described in Methods, the electrostatics (van der Waals) interactions were annihilated (decoupled) over 11 (18) states. The free energy differences between consecutive states were calculated with BAR, summed, and presented as  $\Delta G_{F\text{-ele}}$ ,  $\Delta G_{F\text{-vdw}}$  and  $\Delta G_{F\text{-total}} = \Delta G_{F\text{-ele}} + \Delta G_{F\text{-vdw}}$ . The data is also provided in Table S12.

**TABLE 1** Electric fields projected onto specific bonds, in mega volts per centimeter (MV/cm). Values reported are averages over 5 ns MD trajectories of each molecule in water. Positive projections mean that the electric field is oriented from the first atom (O or C) to the second (H, X, or F).

Bond	LacNAc 1	LacNAc 2	GB0139	GB1211	GB1107
O <sub>2</sub> –H	142.2	145.5	141.3	156.5	156.8
O <sub>4</sub> –H	149.4	131.7	140.2	185.5	89.9
O <sub>6</sub> –H	162.9	163.4	150.5	169.8	127.5
C–X	-	-	-	3.5	-4.5
C–F	-	-	-14.5	-15.9	-16.0

**FIGURE 5** Free energy differences along the path characterizing the bound state (see Figure 7) of the ligands LacNAc 1, LacNAc 2, GB0139, GB1211, and GB1107. As described in Section 4, the electrostatics (van der Waals) interactions were annihilated (decoupled) over 11 (20) states. The free energy differences between consecutive states were calculated with BAR, summed, and presented as  $\Delta G_{B\text{-ele+rest-1}}$ ,  $\Delta G_{B\text{-vdw}}$ ,  $\Delta G_{B\text{-rest-2}}$ , and  $\Delta G_{B\text{-total}} = \Delta G_{B\text{-ele+rest-1}} + \Delta G_{B\text{-vdw}} + \Delta G_{B\text{-rest-2}}$ . The data is also provided in Table S13.



almost identically in GB0139, GB1211, and GB1107. A similar effect is observed for the C–Cl bond in GB1107, although of a lesser magnitude (–4.5 MV/cm). However, the electric field projection is positive along the C–Br bond in GB1211, consistent with the formation of a  $\sigma$ -hole along this axis. Overall, the electric field data suggests that the substituted aromatic rings in the synthetic ligands disrupt electron distribution along O<sub>4</sub>–H, and to some extent O<sub>6</sub>–H, significantly increasing (decreasing) water affinity of that bond in GB1211 (GB1107). This is largely offset in LacNAc 1, LacNAc 2, and GB0139, by the additional 3–4 hydroxyl bonds of the additional galactose or glucosamine units. Note that the removal of hydroxyl bonds and overall decrease in polar surface area was a design feature in GB1107 and GB1211 to maximize tissue diffusion and delivery compared to GB0139 (Zetterberg et al., 2022).

## 2.2 | Natural ligands and synthetic inhibitors in Gal-3's CRD

We present in Figure 5 the magnitude of the free energy difference ( $\Delta G_{B\text{-total}}$ ) associated with the removal of the ligands from the protein environment. The electrostatic ( $\Delta G_{B\text{-ele+rest-1}}$ ), van der Waals ( $\Delta G_{B\text{-vdw}}$ ), and restraints ( $\Delta G_{B\text{-rest-2}}$ ) components of

$\Delta G_{B\text{-total}}$  are also shown. Similarly to the removal of the ligands from water,  $\Delta G_{B\text{-total}}$  is dominated by electrostatics. However,  $\Delta G_{B\text{-vdw}}$  is not constant for all ligands as was previously observed for  $\Delta G_{F\text{-vdw}}$ . Rather, LacNAc 1, LacNAc 2, and GB0139 show a similar, negative,  $\Delta G_{B\text{-vdw}}$  while the second-generation inhibitors, GB1211 and GB1107, exhibit a vanishingly small, positive,  $\Delta G_{B\text{-vdw}}$ . This increases the affinity of GB1211 and GB1107 for the protein, albeit to a smaller degree than  $\Delta G_{B\text{-ele+rest-1}}$ . The negative contributions from  $\Delta G_{B\text{-rest-2}}$  indicate that the release of the multiple distance constraints to a single COM constraint reduces the ligands' affinity for the protein environment. This suggests that the holo (ligand-bound) and apo (ligand-absent) forms of Gal-3 are slightly different, adapting to the substrate. This is especially true for the synthetic inhibitors where additional restraints were used and the magnitude of  $\Delta G_{B\text{-rest-2}}$  is greater. For LacNAc 2 however,  $\Delta G_{B\text{-rest-2}}$  is small (1.5–2 kcal/mol), indicating that Gal-3's CRD does not adapt to the presence of the ligand to the same extent as its synthetic counterparts. Despite this lesser CRD adaptation, LacNAc 1 and 2 have an overall greater  $\Delta G_{B\text{-total}}$ , followed by synthetic inhibitors GB1211 and GB1107.

We present in Table 2 the average electric field projections onto the galactose O<sub>2</sub>–H, O<sub>4</sub>–H, and O<sub>6</sub>–H bonds, as well as the projections onto the C–F and

**TABLE 2** Average electric fields projected onto specific bonds, in MV/cm, averaged over three replicates of 7 ns for each molecule bound to Gal-3.

	LacNAc 1	LacNAc 2	GB0139	GB1211	GB1107
O <sub>2</sub> –H prot.	7.7	4.1	–2.6	–0.8	–6.4
O <sub>2</sub> –H sol.	132.4	139.6	157.7	156.9	190.7
O <sub>4</sub> –H prot.	144.0	108.1	171.5	198.7	167.4
O <sub>4</sub> –H sol.	36.6	40.1	–3.5	–10.4	–17.5
O <sub>6</sub> –H prot.	129.2	126.0	165.3	126.6	118.4
O <sub>6</sub> –H sol.	7.1	21.2	9.3	–24.4	–25.1
C–X prot.	-	-	-	3.4	–10.2
C–X sol.	-	-	-	–5.9	–8.1
C–F prot.	-	-	–8.7	–21.9	–16.4
C–F sol.	-	-	–3.9	2.5	–3.8

Note: Positive projections indicate that the electric field is oriented from the first atom (O or C) to the second (H, X, or F). The contributions from each replicate are given in Figures S8–S12. The electric fields for each replicate are given in Tables S14 and S15.

unique C–X bonds. In each case, the total field is split into two parts: the contribution from the protein and the contribution from the solvent (water and ions). For example, we observe that the electric field projection onto the O<sub>2</sub>–H bond of all ligands is similar in magnitude to that of the free state and is primarily due to contributions from the solvent. However, even if the magnitude of the electric field along the O<sub>4</sub>–H bond in GB1211 is comparable to that in the free state, it is here dominated by the contribution from the protein. The projection is equally dominated by the protein contribution for the O<sub>4</sub>–H bond in the other compounds, but the total magnitude of the field is greater than it was in the free state. This would suggest that the protein environment is favored to stabilize these hydroxyl groups, especially in GB1107 where the difference between the field in the bound and free states reaches 60 MV/cm, compared to 10–20 MV/cm for the natural ligands and GB0139. In contrast, the electric field projections onto the O<sub>6</sub>–H bond in all ligands but GB0139 are significantly smaller in the bound state than the free state, suggesting that a pure water environment is preferred over the interaction with the CRD for that bond. The decomposition of the protein electric fields into contribution from individual residues (Figures S8–S12) confirms that O<sub>4</sub>–H and O<sub>6</sub>–H engage in hydrogen bonds with His158, Asn174, and Glu184, as reported in the literature (Zetterberg et al., 2018, 2022). Finally, the projections onto the C–F and C–X bonds in GB1211 and GB1107 are overall more negative, by about 5–10 MV/cm, in the bound state compared to the free state. In GB0139, the C–F bond experiences similar or slightly more positive fields in the bound state than in the free state. Overall, the electric field data in the bound state reveal that the protein replaces water sufficiently well to stabilize the hydroxyl groups, which explains why the natural ligands have a more positive  $\Delta G_{B-total}$  than the synthetic inhibitors. The projections across all bonds in GB1211 are slightly higher than in

GB1107 and GB0139, suggesting that the difference between the inhibitors is a combined effect rather than the result of a single interaction.

## 2.3 | Binding free energies

In this last section, we combine the total free energy difference in the free ( $\Delta G_{F-total}$ ) and bound ( $\Delta G_{B-total}$ ) states to compute the binding free energy of all five ligands ( $\Delta G^\circ$ ) from Equation (2). The results are presented in Table 3. On average,  $\Delta G^\circ = -0.3, 1.4, -4.3, -9.5,$  and  $-6.7$  kcal/mol for LacNAc 1, LacNAc 2, GB0139, GB1211, and GB1107, respectively. We observe that the synthetic inhibitors have a  $\Delta G^\circ$  between 5 and 9 kcal/mol stronger than the natural ligands. GB1211 was also observed to have a greater (>3 kcal/mol)  $\Delta G^\circ$  than GB1107 and GB0139, which aligns with the findings from Zetterberg et al (Zetterberg et al., 2022). Interestingly, the binding energy of GB0139 falls between that of the natural ligands and second-generation inhibitors, keeping with its shared qualities of both groups.

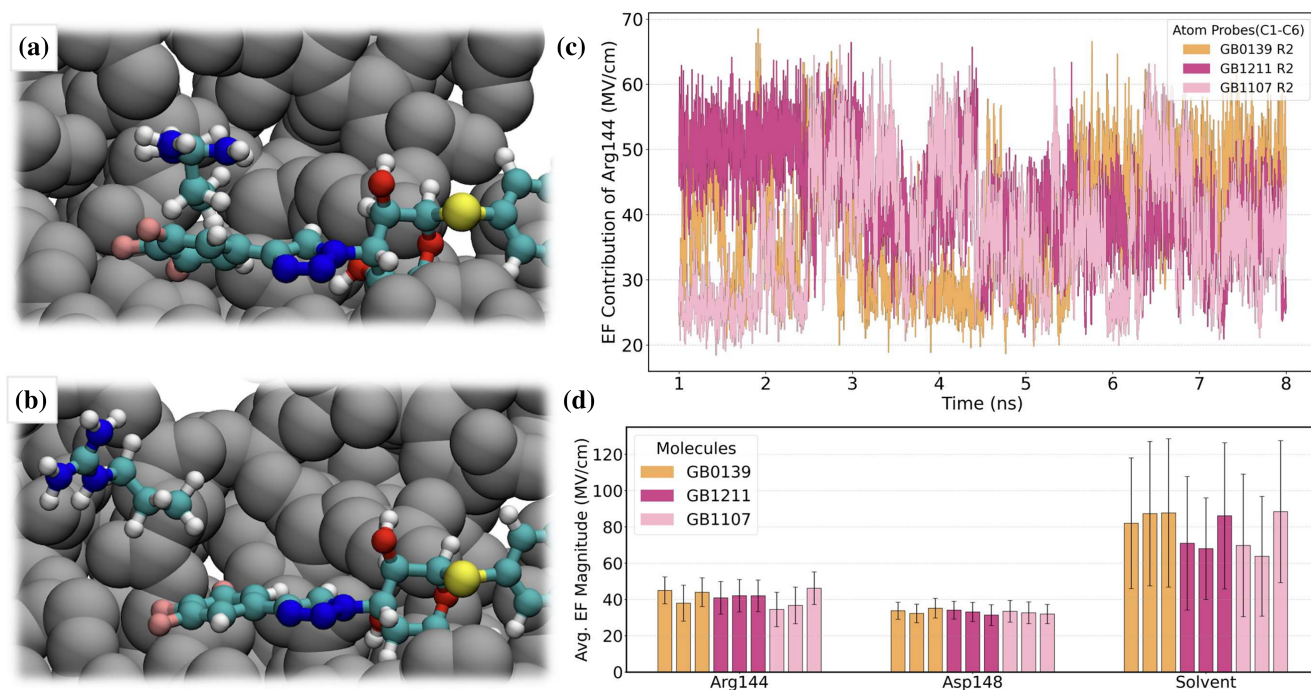
In addition to the  $\sigma$ -hole influence of the unique halogen atom in GB1211 and GB1107 described above, a key difference between GB1211 and GB1107 is the cation- $\pi$  interactions between the fluorinated ring and Arg144. Indeed, this interaction has been shown to be critical for synthetic Gal-3 inhibitors, and the addition of an aromatic ring substituent resulted in 5 times greater affinity than the base carbohydrate structure in experimental studies (Sörme et al., 2005). Sörme et al. report that the strength of the cation- $\pi$  interaction ranges from –0.4 to –2.4 kcal/mol, and that it does not depend entirely on the electrostatic interactions but also van der Waals and polarization effects. In Figure 6a,b, we show a snapshot representative of Arg144 engaging (disengaging) in (from) the cation- $\pi$  interaction with GB1107. To better quantify the strength of this

**TABLE 3** Binding free energy ( $\Delta G^\circ$ ) in kcal/mol for ligand replicates (R1–R3).

Replicates	LacNAc 1			LacNAc 2			GB0139			GB1211			GB1107		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
$\Delta G_{F\text{-total}}$ in kcal/mol	48.5 $\pm$ 0.3			44.5 $\pm$ 0.3			28.0 $\pm$ 0.4			17.9 $\pm$ 0.3			9.2 $\pm$ 0.3		
$\Delta G_{B\text{-total}}$ in kcal/mol	48.7	51.5	51.3	44.0	44.5	46.1	34.2	33.0	36.0	29.9	29.2	28.9	18.6	15.7	19.1
$\langle \Delta G_{B\text{-total}} \rangle$ in kcal/mol	50.5 $\pm$ 1.5			44.9 $\pm$ 1.1			34.4 $\pm$ 1.5			29.3 $\pm$ 0.5			17.8 $\pm$ 1.8		
$\Delta G_{\text{rest-3}}$ in kcal/mol	−1.7			−1.8			−2.0			−2.0			−1.9		
$\Delta G^\circ$ in kcal/mol	1.4	−1.3	−1.1	2.3	1.7	0.2	−4.1	−2.9	−5.9	−10.0	−9.4	−9.0	−7.5	−4.6	−8.0
$\langle \Delta G^\circ \rangle$ in kcal/mol	−0.3 $\pm$ 1.8			1.4 $\pm$ 1.4			−4.3 $\pm$ 1.9			−9.5 $\pm$ 0.8			−6.5 $\pm$ 2.1		
Data obtained from binding assays in kcal/mol	−5.50, −7.96 <sup>a</sup> (Hsieh et al., 2015)			−6.11, −8.84 <sup>a</sup> (Hsieh et al., 2015) −7.83 (Salomonsson et al., 2010) −5.04 (Seetharaman et al., 1998)			−10.71 (Delaine et al., 2016)			−10.36 (Zetterberg et al., 2022)			−10.13 (Zetterberg et al., 2022)		

Note: The components  $\Delta G_{F\text{-total}}$ ,  $\Delta G_{B\text{-total}}$  and  $\Delta G_{\text{rest-3}}$  are also shown (see Equation (2)).  $\langle \Delta G_{F\text{-total}} \rangle$  and  $\langle \Delta G^\circ \rangle$  are averages over replicas. Experimental binding free energies are converted from dissociation constants.

<sup>a</sup>Bilayer interferometry at 300 K, fluorescence anisotropy at 277 K.



**FIGURE 6** Cation- $\pi$  interaction between the fluorinated aromatic ring and Arg144 located in the Gal-3 CRD. (a) Arg144 is stacked on top of the ring in GB1107. (b) Arg144 has flipped outwards away from GB1107. (c) Time evolution of Arg144's contributed electric field using atomic probes C1–C6 (aromatic carbons) for replicate 2 of all synthetic inhibitors. (d) Arg144, Asp148, and solvent contributions to the magnitude of the electric field, averaged over time and the 6 aromatic carbons.

interaction across our various MD simulations, we computed the time evolution of the contribution from Arg144 to the electric field on each of the aromatic carbons in GB0139, GB1211, and GB1107. Note that Arg144 is the top contributor to the total field on these atoms, followed by Asp148. We show in Figure 6c that trajectories where Arg144 is engaging in the cation- $\pi$  interaction result in fields around 50 MV/cm, as for GB1211 R2. Meanwhile, the electric field drops to 20 MV/cm when Arg144 is not engaged in the

interaction, as it is the case for GB0139 R2 between 3 and 6 ns, GB1107 R2 during the first 2.5 ns, between 4.5 and 6.2 ns and after 7 ns. The data averaged over time and replicates is presented in Figure 6d, as well as the contribution from Asp148 and the solvent, for reference. We see that the contribution from Arg144 is very consistent across all replicates in GB1211, suggesting a sustainable binding pose. Meanwhile, Arg144 is 5–10 MV/cm lower in the second replicate of GB0139 and the first two replicates of GB1107, due to significant

**TABLE 4** Characteristic dihedral angles for LacNAc type 1 and 2 across replicates (see Figure S3 for a sketch of these angles). Angle 1 LacNAc 1: O<sub>5</sub>Gal–C<sub>1</sub>Gal–O<sub>3</sub>GlcNAc–C<sub>3</sub>GlcNAc | LacNAc 2: O<sub>5</sub>Gal–C<sub>1</sub>Gal–O<sub>4</sub>GlcNAc–C<sub>4</sub>GlcNAc. Angle 2 LacNAc 1: C<sub>1</sub>Gal–O<sub>3</sub>GlcNAc–C<sub>3</sub>GlcNAc–C<sub>4</sub>GlcNAc | LacNAc 2: C<sub>1</sub>Gal–O<sub>4</sub>GlcNAc–C<sub>4</sub>GlcNAc–C<sub>5</sub>GlcNAc. Data presented as average angle over the sampling MD (standard deviation). The evolution of each angle over time is given in Figures S4–S7.

	Ref (Hsieh et al., 2015; Seetharaman et al., 1998).	R1	R2	R3
Angle 1 (°)				
LacNAc 1	–60	–105.3 (21.7)	–69.0 (9.2)	–69.4 (8.2)
LacNAc 2	–66, –68	–97.1 (20.7)	–70.8 (9.9)	–70.3 (9.8)
Angle 2 (°)				
LacNAc 1	135	101.6 (18.2)	142.7 (7.8)	141.7 (7.6)
LacNAc 2	–108, –103	–137.0 (17.6)	–108.4 (7.5)	–108.5 (7.7)

amount of time where Arg144 faces away from the aromatic ring, breaking the cation- $\pi$  interaction. These results correlate with variations in  $\Delta G^\circ$  across replicates.

Finally, we find that  $\Delta G^\circ$  for the disaccharides (i.e., LacNAc 1, LacNAc 2, and GB0139) was overall more positive than the literature would suggest from converted dissociation constants. Further, our results show LacNAc 1 as the preferential binder, with LacNAc 2 having a positive binding free energy. The trend is contrary to findings from Hsieh et al., who found a binding preference of the Gal-3 CRD for LacNAc 2 over LacNAc 1 (Hsieh et al., 2015). Hsieh et al. associated the difference in binding between LacNAc 1 and 2 to a different binding pose of the N-acetylglucosamine unit, as indicated by two dihedral angles characterizing the glycosidic bond in the crystal structure (Hsieh et al., 2015). In Table 4, we show the average and standard deviations of these two angles in our MD simulations, compared with data acquired from crystal structure (Hsieh et al., 2015; Seetharaman et al., 1998).

The first dihedral angle provides the orientation of the  $\beta$ -galactose unit relative to the glycosidic bond. We observe an average of  $-70^\circ$  for LacNAc 1 and 2 in R2 and R3, which is in good agreement with the experimental values. However, R1 for both ligands showed a higher deviation and more negative angle due to larger molecule movements within the binding pocket. The second dihedral angle tracks the orientation of the N-acetylglucosamine unit relative to the galactose unit. Here again, R1 shows a higher deviation, but R2 and R3 show good agreement with the experimental data. Interestingly, this may explain the 2.5 kcal/mol difference in  $\Delta G^\circ$  observed between R1 and R2/R3 of LacNAc 1 (Table 3). The LacNAc 2 replicates also see a slight decrease in binding free energy compared to R1, however, these results are still more positive than expected from literature.

It is worth noting that the experimental binding energies are not a one-to-one comparison with our theoretical estimates, as the fluorescence polarization

methods require covalently attaching a fluorophore molecule to the ligand, which was not modeled in the simulations. In fact, Sörme et al. explicitly mentions that fluorescence polarization tends to overestimate the affinity for small carbohydrate-lectin systems, which would include Gal-3/LacNAc complexes (Sörme et al., 2004). The conversion from a dissociation constant to a binding affinity can also be a bad comparison when strong protein-ligand interactions occur outside of the canonical binding location. Indeed, our approach only sampled protein-ligand configurations close to the bound crystal structures and did not account for alternative binding sites along the groove in the S-face. Support for the existence of other, potentially more favorable, binding poses is that our first trial for LacNAc 2 resulted in the dissociation of the ligand after 39 ns of simulation time, directly after sampling the first replicate from 31 to 38 ns. Surprisingly,  $\Delta G^\circ$  from this replicate before dissociation was 0.63 kcal/mol. Alternatively, the underestimation of  $\Delta G^\circ$  for LacNAc 1 and 2 could point to a force field issue where AMOEBA overestimates the carbohydrates' affinity for water.

### 3 | CONCLUSION

Leading Gal-3 synthetic inhibitors were studied alongside natural carbohydrate ligands to rationalize binding affinities at the molecular scale. We used the AMOEBA polarizable force field along with alchemical transformations and electric field calculations to understand the ligands' affinity for water and Gal-3, thereby providing a deeper understanding of the resulting binding energies. We predicted the binding energy of inhibitor GB1211 to be  $-9.5$  kcal/mol on average, while GB1107 exhibited a lower affinity at  $-6.7$  kcal/mol, consistent with experimental data. Electric field calculations suggest that the stronger binding of GB1211 originates from stronger interactions between the protein and the hydroxyl bonds O<sub>4</sub>–H and O<sub>6</sub>–H, as well as a stronger cation- $\pi$  interaction with Arg144. However, our estimates for the disaccharides LacNAc 1, LacNAc 2, and GB0139 were

much more positive than experimental values, at  $-0.3$ ,  $1.4$ , and  $-4.3$  kcal/mol, respectively. This may suggest the existence of multiple binding poses for these natural ligands, not sampled in our MD that started from crystal structures. Alternatively, this could also come from a force field issue where AMOEBA overestimates the water affinity of carbohydrates. Further work will involve systematically scaling the van der Waals terms in AMOEBA to assess the later hypothesis and investigating the interaction profile of these ligands in other galectin members. Indeed, the relative difference between the electric field projections and binding energies of LacNAc 1, LacNAc 2, GB0139, GB1211, and GB1107 complexed in Gal-3 and other Gal forms will further assess the validity of our methods and provide a molecular picture of selectivity.

## 4 | METHODS

### 4.1 | System preparation

The crystal structures PDB ID 4XBN (Hsieh et al., 2015), 1KJL (Sörme et al., 2005), 5H9P (Hsieh et al., 2016), 7ZQX (Zetterberg et al., 2022), 6EOL (Zetterberg et al., 2018) were used to prepare the protein-ligand systems for LacNAc 1, LacNAc 2, GB0139, GB1211, and GB1107, respectively. The protein files were cleaned of crystallography artifacts and any missing atoms from terminal ends were added to result in 138 residues comprising the Gal-3 CRD but omitting the flexible N-terminus tail. Protons were added using the REDUCE program (Word et al., 1999). The histidines were consistently protonated on the  $\epsilon$  nitrogen (HIE) with the exception of His158, present at the binding site, which was protonated on the  $\delta$  nitrogen (HID) based on previous structural studies (Hsieh et al., 2015; Sörme et al., 2005), and NMR spectroscopy data (Manzoni et al., 2018). The TINKER8 software package (Rackers et al., 2018) was used to solvate the protein-ligand systems in a  $60^3 \text{ \AA}^3$  cubic water box, and the +4 total charge of the protein was balanced with 10  $\text{Na}^+$  ions and 14  $\text{Cl}^-$  ions. For the simulations of the ligand in water, the molecules were placed in a  $30^3 \text{ \AA}^3$  cubic box without ions.

### 4.2 | Molecular dynamics

POLYPE2 (Walker et al., 2022) was used to generate the AMOEBA parameters for the ligand molecules, while the AMOEBABIO18 parameters (Shi et al., 2013; Zhang et al., 2018) were used for the protein, water, and ions. GPU-enabled TINKER9 (Rackers et al., 2018; Wang, 2021) was used to minimize the energy of each system and run the MD simulations. The bound and free systems were run at 300 K in the NPT ensemble

with the Nosé-Hoover thermostat and barostat for an initial 30 ns equilibration period. The final atomic coordinates and velocities were used as inputs for the first round of thermodynamic state simulations, which we call replicate 1 (R1). The bound systems were further equilibrated to start a second replicate (R2) at 50 ns, then a third (R3) at 70 ns. Root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) plots are shown in Figures S1 and S2, respectively.

### 4.3 | Binding free energy calculations

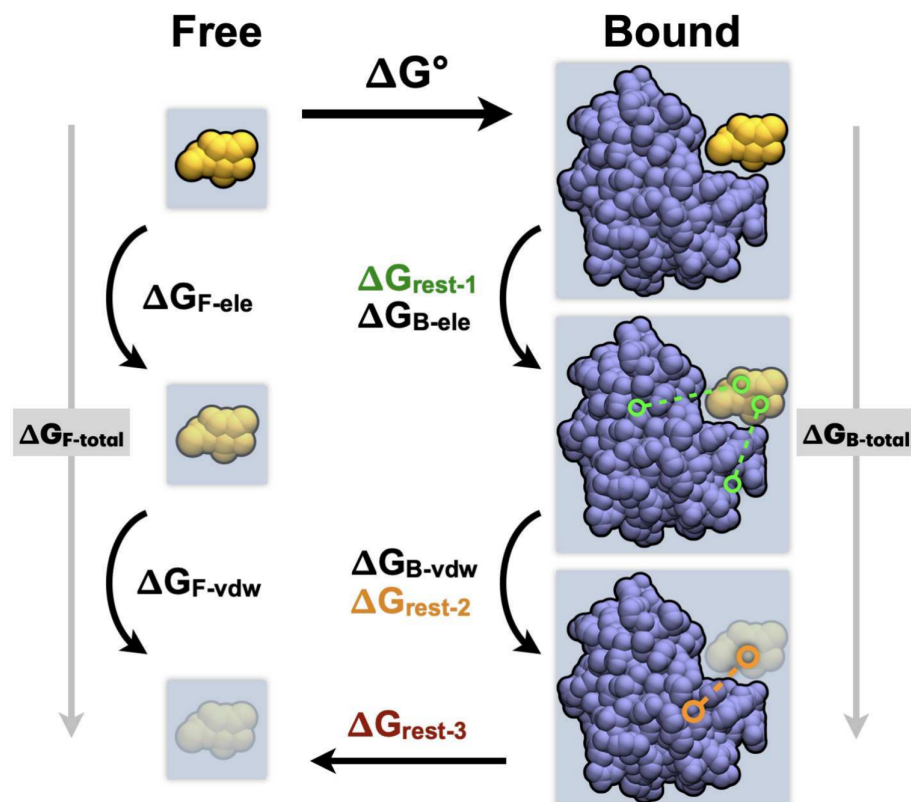
We calculate the binding free energy ( $\Delta G^\circ$ ) of Gal-3-ligand complexes according to the thermodynamic cycle presented in Figure 7. For each complex, the ligand is modeled in its free (i.e., in water) and bound (i.e., with Gal-3) states through multiple stages. First, the electrostatic interactions of the ligand are annihilated, meaning both intramolecular and intermolecular interactions are scaled off. Then, the van der Waals interactions between the ligand and its environment are decoupled, meaning the intermolecular van der Waals forces are scaled off while the intramolecular ones remain active. This scaling off of both electrostatics and van der Waals interactions results in a fully decoupled ligand that has effectively been moved to the gas phase. In this work, we repeat this process in both water and the protein environment. To prevent ligand dissociation during the stages where the electrostatics and van der Waals interactions are turned down, we use restraints to maintain its position and orientation in the binding pocket, as detailed below.

#### 4.3.1 | Restraints

In this work, we implement multiple distance restraints to sample the relevant configurational space through flat-bottomed harmonic potentials, defined as:

$$V_{\text{flat}}(d) = \begin{cases} k_{\text{flat}}(d - r_i)^2 & \text{if } d < r_i \\ 0 & \text{if } r_i \leq d \leq r_o \\ k_{\text{flat}}(d - r_o)^2 & \text{if } d > r_o \end{cases}$$

where  $k_{\text{flat}}$  is the force constant (kcal/mol),  $r_i$  and  $r_o$  the inner and outer radius in which the distance  $d$  is allowed to freely fluctuate. The radii were chosen to be around 3 standard deviations from the average value, with the intention to be non-invasive to the natural ligand interactions. The anchor residues were checked for low RMSF values and the anchor atoms were chosen to be heavy atoms (carbon, oxygen, or nitrogen). Specific details of atom indexes on which restraints were applied,  $r_i$ ,  $r_o$ , and  $k_{\text{flat}}$  are given in Tables S1–S3. In brief, the hydrogen bonds from the galactose unit for



**FIGURE 7** Absolute alchemical thermodynamic cycle showing the paths through the 'Free' (i.e., in water) and 'Bound' (i.e., in Gal-3) ligand states to compute the binding free energy  $\Delta G^\circ$ . In each path, the electrostatic and van der Waals interactions are annihilated and decoupled, respectively.

all ligands were mimicked by restraining the Gal-O<sub>4</sub> and Gal-O<sub>6</sub> atoms to nitrogen atoms in His158 and Asn174, respectively. The LacNAc molecules were additionally tethered by the anomeric O atom to the nitrogen in Arg162 to keep the carbohydrate against the protein without restraining the rotational freedom of the glucosamine unit along the glycosidic bond. Since the synthetic ligands are longer molecules, restraints were instead placed at both ends. The shared fluorinated aromatic ring was tethered to the backbone nitrogen of Gly238. For GB1107 and GB1211, one restraint mimicked the halogen bond between the chlorine or bromine atom and the carbonyl oxygen of Gly182. For GB0139, a restraint was placed to mimic the hydrogen bond from O<sub>2</sub> on the second galactose unit to Glu184. Due to the length of this inhibitor, a fifth restraint was added between the second aromatic ring and Arg186. The restraints were intended to mimic natural interactions and be non-invasive to the unbiased MD simulations.

To decouple the degrees of freedom between protein and ligand, we use a harmonic restraint, as implemented in TINKER9 (Rackers et al., 2018; Wang, 2021), between the center of mass of the ligand and protein binding region. The harmonic restraint potential is defined as:

$$V_{\text{harm}}(d) = k_{\text{harm}}(d - d_0)^2$$

where  $k_{\text{harm}}$  is the force constant ( $k_{\text{harm}} = 15$  kcal/mol here),  $d_0$  the equilibrium distance between selections, and  $d$  the current distance. The protein atom center of mass (COM) selection utilized heavy atoms within about 8–10 Å of the ligand COM, and the atom indices are listed in Table S4.

#### 4.3.2 | Thermodynamic window settings

The scaling off of the electrostatic and van der Waals interactions is achieved over multiple parallel MD simulations, also referred to as thermodynamic states, using a scaling parameter ( $\lambda$ ).  $\lambda$  is incrementally decreased from 1 (fully interacting ligand) to 0 (non-interacting ligand) over  $k$  steps.

For the free ligand in water, we used 29 total thermodynamic states (Table S5). States 1–11 were used to scale off the ligand's electrostatic interactions ( $\lambda_{\text{ele}} = 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.0$ ) and states 12–29 were used to scale off the van der Waals forces ( $\lambda_{\text{vdw}} = 0.975, 0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.0$ ). For the protein-bound ligand, we used 34 total thermodynamic states (Table S6). States 1–11 were used to scale off the electrostatic interactions ( $\lambda_{\text{ele}} = 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.0$ ) and states 12–31 were used to scale off the van der

Waals forces ( $\lambda_{\text{vdw}} = 0.975, 0.95, 0.9, 0.85, 0.8, 0.75, 0.725, 0.7, 0.65, 0.625, 0.6, 0.55, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.025, 0.0$ ). The flat-bottom harmonic protein–ligand restraints were activated linearly over the first five states ( $k_{\text{flat}} = 0, 10, 20, 30, 40$  kcal/mol) and were kept at  $k_{\text{flat}} = 40$  kcal/mol for the remainder of the steps. Once the ligand was fully decoupled ( $\lambda_{\text{ele}} = 0.0, \lambda_{\text{vdw}} = 0.0$ ) states 32–33 were used to scale down the flat-bottom restraint strength ( $k_{\text{flat}} = 20.5$  kcal/mol). In step 34, the multiple distance restraints were replaced with a single harmonic restraint ( $k_{\text{harm}} = 15$  kcal/mol) between the COM of the ligand and the COM of a selection of protein atoms in the binding site. See also Table S6 for tabulated values of  $\lambda_{\text{ele}}$ ,  $\lambda_{\text{vdw}}$ ,  $k_{\text{flat}}$ , and  $k_{\text{harm}}$  for each thermodynamic state.

Each  $\lambda$  thermodynamic state was run in parallel with the same initial coordinates and velocities. The free states were simulated for 6 ns total, 1 ns equilibration, and 5 ns sampling. The bound states were simulated for 8 ns total, 1 ns equilibration, and 7 ns sampling. The coordinates were printed every 1 ps for both systems, resulting in 5000 samples for each of the free states and 7000 samples for the bound states.

#### 4.4 | Computing free energy differences

The incremental change between successive  $\lambda$  allows for better phase-space overlap between neighboring thermodynamic states. This is necessary for the convergence of free energy estimators, such as the Bennett acceptance ratio (BAR) (Bennett, 1976; Ytreberg et al., 2006) used in this work. BAR estimates the free energy difference of each set of neighboring states using their respective Boltzmann factors. Here, we use the TINKER9 (Rackers et al., 2018; Wang, 2021) implementation of BAR where the statistical error is estimated through bootstrapping (error listed in Tables S7–S11 for each ligand). Considering the free energy difference,  $\Delta G_{\lambda_k \lambda_{k+1}}$ , between the two successive  $\lambda_k$  and  $\lambda_{k+1}$  states ( $k$  is the step number), it yields:

$$\begin{aligned} & \sum_{q_k=1}^{n_{\lambda_k}} \frac{1}{1 + \exp\left(\ln\left(\frac{n_{\lambda_k}}{n_{\lambda_{k+1}}}\right) + \beta \Delta U_{\lambda_k \lambda_{k+1}} - \beta \Delta G_{\lambda_k \lambda_{k+1}}\right)} \\ &= \sum_{q_{k+1}=1}^{n_{\lambda_{k+1}}} \frac{1}{1 + \exp\left(\ln\left(\frac{n_{\lambda_{k+1}}}{n_{\lambda_k}}\right) - \beta \Delta U_{\lambda_k \lambda_{k+1}} + \beta \Delta G_{\lambda_k \lambda_{k+1}}\right)}, \end{aligned} \quad (1)$$

where  $\beta = \frac{1}{k_B T}$ ,  $k_B$  the Boltzmann constant and  $T$  the temperature.  $n_{\lambda_k}$  and  $n_{\lambda_{k+1}}$  are the number of samples (i.e. MD simulation frames) from states  $\lambda_k$  and  $\lambda_{k+1}$ , respectively.  $\Delta U_{\lambda_k \lambda_{k+1}} = U_{\lambda_{k+1}}(q) - U_{\lambda_k}(q)$ , where  $U_i(q)$

is the potential energy of frame  $q$  in state  $i$ .  $\Delta G_{\lambda_k \lambda_{k+1}}$  must be numerically solved. In practice, BAR takes each MD frame from states  $\lambda_k$  and  $\lambda_{k+1}$ , calculates their energy with each of the states' potential energy functions, and iteratively solves for  $\Delta G_{\lambda_k \lambda_{k+1}}$  using Equation (1).

The energy difference between two stages of the thermodynamic cycle described in Figure 7 is computed by summing the  $\Delta G_{\lambda_k \lambda_{k+1}}$  over the number of steps  $k$  required to turn on/off the electrostatic or van der Waals interactions. The detailed  $\Delta G_{\lambda_k \lambda_{k+1}}$  for each segment of the thermodynamic cycle is given in Tables S7–S11 for each replicate.

Overall, we define the binding free energy ( $\Delta G^\circ$ ) as:

$$\begin{aligned} \Delta G^\circ &= \underbrace{\Delta G_{\text{F-ele}} + \Delta G_{\text{F-vdw}}}_{\Delta G_{\text{F-total}}} \\ &\quad - \left( \underbrace{\Delta G_{\text{B-ele}} + \Delta G_{\text{rest-1}} + \Delta G_{\text{B-vdw}} + \Delta G_{\text{rest-2}}}_{\Delta G_{\text{B-total}}} \right) \\ &\quad - \Delta G_{\text{rest-3}}, \end{aligned} \quad (2)$$

where  $\Delta G_{\text{F-ele}}$  is the free energy difference associated with the annihilation of the electrostatic interactions of the ligand in water (i.e., “free” state),  $\Delta G_{\text{F-vdw}}$  the free energy difference associated with the decoupling of the van der Waals forces of the ligand in water,  $\Delta G_{\text{B-ele}}$  the free energy difference associated with the annihilation of the electrostatic interactions of the ligand in the protein environment (i.e., bound state),  $\Delta G_{\text{B-vdw}}$  the free energy difference associated with the decoupling of the van der Waals forces of the ligand in the protein environment,  $\Delta G_{\text{rest-1}}$  is the free energy difference associated with turning on multiple distance restraints to maintain the ligand in the binding pocket and  $\Delta G_{\text{rest-2}}$  is the free energy difference associated with collapsing the multiple distance restraints to a single distance restraint. All of these free energy differences are computed through the BAR method. Finally,  $\Delta G_{\text{rest-3}}$  is the free energy associated with releasing the non-interacting ligand from the binding pocket to bulk solvent (Hermans & Wang, 1997). This term is a correction due to sampling and convergence issues of letting the ligand explore the entire simulation box and is defined as:

$$\Delta G_{\text{rest-3}} = -k_B T \ln\left(C^0 V_{\text{restraint}}\right), \quad (3)$$

where  $V_{\text{restraint}}$  is the volume integral of the restrained ligand and  $C^0$  the standard concentration (1 mol/L or 1 mol/1660 Å<sup>3</sup>). Here, we use the TINKER8 FREEFIX (Rackers et al., 2018) utility to compute  $\Delta G_{\text{rest-3}}$ , closing the alchemical cycle in Figure 7.

## 4.5 | Electric field calculations

We use ELECTRIC (Nash et al., 2020) to calculate the electric field for each ligand replicate from the AMOEBA MD sampling simulations (6 ns for the free ligands and 8 ns for the protein-bound molecules, all fully interacting).

The electric field projection along a bond  $ij$  between atoms  $i$  and  $j$  was calculated as:  $E_{\text{proj}}^{ij} = \frac{\vec{E}_i + \vec{E}_j}{2} \cdot \vec{u}^{ij}$  where  $E_{\text{proj}}^{ij}$  is the projected electric field,  $\vec{E}_i$  the electric field at atom  $i$ ,  $\vec{E}_j$  the electric field at atom  $j$ , and  $\vec{u}^{ij}$  is the unit bond axis between  $i$  and  $j$ . For the natural ligands (LacNAc 1 and 2), the projected electric field was analyzed along the O<sub>2</sub>–H<sub>2</sub>, O<sub>4</sub>–H<sub>4</sub>, and O<sub>6</sub>–H<sub>6</sub> bonds in the galactose unit. This selection was repeated for synthetic inhibitors, with additional projections along the C–X bond (X = bromine or chlorine) and the central C–F bond. For the fluorinated aromatic ring, the six carbons were used as electric field probes. Instead of projections, the magnitude of the electric field was computed.

### AUTHOR CONTRIBUTIONS

**Luke Newman:** Writing – original draft; investigation; methodology; validation; visualization; data curation.  
**Valerie Vaissier Welborn:** Funding acquisition; writing – original draft; writing – review and editing; formal analysis; supervision; conceptualization.

### ACKNOWLEDGMENTS

The authors thank GlycoMIP, a National Science Foundation Materials Innovation Platform funded through Cooperative Agreement DMR-1933525 for financial support. Protein and ligand images were generated using UCSF ChimeraX (Pettersen et al., 2021) and VMD (Humphrey et al., 1996).

### FUNDING INFORMATION

GlycoMIP, National Science Foundation Materials Innovation Platform, Funded through Cooperative Agreement DMR-1933525.

### CONFLICT OF INTEREST STATEMENT

There are no conflicts to declare.

### DATA AVAILABILITY STATEMENT

The data supporting this article have been included as part of Appendix S1. Input files, including Poltype 2 and Tinker 9 input with parameter files, are available at <https://github.com/WelbornGroup/>.

### ORCID

Valerie Vaissier Welborn  <https://orcid.org/0000-0003-0834-4441>

## REFERENCES

- Ahmed R, Anam K, Ahmed H. Development of galectin-3 targeting drugs for therapeutic applications in various diseases. *Int J Mol Sci.* 2023;24(9):8116. <https://doi.org/10.3390/ijms24098116>
- Aslanis V, Slack RJ, MacKinnon AC, McClinton C, Tantawi S, Gravelle L, et al. Safety and pharmacokinetics of GB1211, an oral galectin-3 inhibitor: a single- and multiple-dose first-in-human study in healthy participants. *Cancer Chemother Pharmacol.* 2023;91(3):267–80. <https://doi.org/10.1007/s00280-023-04513-y>
- Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys.* 1976;22(2):245–68.
- Boer dRA, Voors AA, Muntendam P, Gilst vWH, Veldhuisen vDJ. Galectin-3: a novel mediator of heart failure development and progression. *Eur J Heart Fail.* 2009;11(9):811–7.
- Bouffette S, Botez I, De Ceuninck F. Targeting galectin-3 in inflammatory and fibrotic diseases. *Trends Pharmacol Sci.* 2023;44(8):519–31. <https://doi.org/10.1016/j.tips.2023.06.001>
- Califice S, Castronovo V, Van Den Brùle F. Galectin-3 and cancer. *Int J Oncol.* 2004;25(4):983–1075.
- Cummings RD, Liu FT, Rabinovich GA, Stowell SR, Vasta GR. Galectins. *Essentials of glycobiology.* Volume 4: Chapter 36. New York: Cold Spring Harbor Laboratory Press; 2022.
- Dang Z, MacKinnon A, Marson LP, Sethi T. Tubular atrophy and interstitial fibrosis after renal transplantation is dependent on galectin-3. *Transplantation.* 2012;93(5):477–84.
- Delaine T, Collins P, MacKinnon A, Sharma G, Stegmayr J, Rajput VK, et al. Galectin-3-binding glycomimetics that strongly reduce bleomycin-induced lung fibrosis and modulate intracellular glycan recognition. *ChemBioChem.* 2016;17(18):1759–70.
- Demetriou M, Granovsky M, Quaggin S, Dennis JW. Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature.* 2001;409(6821):733–9.
- Diehl RC, Chorghade RS, Keys AM, Alam MM, Early SA, Dugan AE, et al. CH- $\pi$  interactions are required for human galectin-3 function. *JACS Au.* 2024;4(8):3028–37. <https://doi.org/10.1021/jacsau.4c00357>
- Dong R, Zhang M, Hu Q, Zheng S, Soh A, Zheng Y, et al. Galectin-3 as a novel biomarker for disease diagnosis and a target for therapy. *Int J Mol Med.* 2018;41(2):599–614. <https://doi.org/10.3892/ijmm.2017.3311>
- Henderson NC, Mackinnon AC, Farnworth SL, Poirier F, Russo FP, Iredale JP, et al. Galectin-3 regulates myofibroblast activation and hepatic fibrosis. *Proc Natl Acad Sci.* 2006;103(13):5060–5.
- Henderson NC, Sethi T. The regulation of inflammation by galectin-3. *Immunol Rev.* 2009;230(1):160–71.
- Hermans J, Wang L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *J Am Chem Soc.* 1997;119:2707–14.
- Hsieh T, Lin H, Tu Z, Huang B, Wu S, others. Structural basis underlying the binding preference of human galectins-1, -3 and -7 for Gal $\beta$ 1-3/4GlcNAc. *PLoS One.* 2015;10(5):e0125946.
- Hsieh TJ, Lin HY, Tu Z, Lin TC, Wu SC, Tseng YY, et al. Dual thiodigalactoside-binding modes of human galectins as the structural basis for the design of potent and selective inhibitors. *Sci Rep.* 2016;6(1):29457.
- Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14(1):33–8.
- Leffler H, Carlsson S, Hedlund M, Qian Y, Poirier F. Introduction to galectins. *Glycoconj J.* 2002;19:433–40.
- Liu FT, Rabinovich GA. Galectins as modulators of tumour progression. *Nat Rev Cancer.* 2005;5(1):29–41.
- Lok DJ, Van Der Meer P, Porte d IPWBA, Lipsic E, Van Wijngaarden J, Hillege HL, et al. Prognostic value of galectin-3,

- a novel marker of fibrosis, in patients with chronic heart failure: data from the DEAL-HF study. *Clin Res Cardiol.* 2010;99:323–8.
- MacKinnon AC, Gibbons MA, Farnworth SL, Leffler H, Nilsson UJ, Delaine T, et al. Regulation of transforming growth factor- $\beta$ -driven lung fibrosis by galectin-3. *Am J Respir Crit Care Med.* 2012;185(5):537–46.
- Manzoni F, Wallerstein J, Schrader TE, Ostermann A, Coates L, Akke M, et al. Elucidation of hydrogen bonding patterns in ligand-free, lactose- and glycerol-bound galectin-3C by neutron crystallography to guide drug design. *J Med Chem.* 2018;61(10):4412–20. <https://doi.org/10.1021/acs.jmedchem.8b00081>
- Markowska AI, Jefferies KC, Panjwani N. Galectin-3 protein modulates cell surface expression and activation of vascular endothelial growth factor receptor 2 in human endothelial cells. *J Biol Chem.* 2011;286(34):29913–21. <https://doi.org/10.1074/jbc.M111.226423>
- Nabi IR, Shankar J, Dennis JW. The galectin lattice at a glance. *J Cell Sci.* 2015;128(13):2213–9. <https://doi.org/10.1242/jcs.151159>
- Nash J, Barnes T, Welborn VV. ELECTRIC: electric fields leveraged from multipole expansion calculations in tinker rapid interface code. *J Open Source Softw.* 2020;5(54):2576.
- Nishi Y, Sano H, Kawashima T, Okada T, Kuroda T, Kikkawa K, et al. Role of galectin-3 in human pulmonary fibrosis. *Allergol Int.* 2007;56(1):57–65.
- Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 2021;30(1):70–82.
- Rabinovich GA, Toscano MA. Turning ‘sweet’ on immunity: galectin-glycan interactions in immune tolerance and inflammation. *Nat Rev Immunol.* 2009;9(5):338–52.
- Rackers JA, Wang Z, Lu C, Laury ML, Lagardère L, Schnieders MJ, et al. Tinker 8: software tools for molecular design. *J Chem Theory Comput.* 2018;14(10):5273–89. <https://doi.org/10.1021/acs.jctc.8b00529>
- Ruvolo PP. Galectin 3 as a guardian of the tumor microenvironment. *Biochim Biophys Acta Mol Cell Res.* 2016;1863(3):427–37. <https://doi.org/10.1016/j.bbamcr.2015.08.008>
- Salomonsson E, Carlsson MC, Osla V, Hendus-Altenburger R, Kahl-Knutson B, Öberg CT, et al. Mutational tuning of galectin-3 specificity and biological function. *J Biol Chem.* 2010;285(45):35079–91.
- Sciacchitano S, Lavra L, Morgante A, Olivieri A, Magi F, de Francesco G, et al. Galectin-3: one molecule for an alphabet of diseases, from a to z. *Int J Mol Sci.* 2018;19(2):379. <https://doi.org/10.3390/ijms19020379>
- Seetharaman J, Kanigsberg A, Slaaby R, Leffler H, Barondes SH, Rini JM. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-Å resolution. *J Biol Chem.* 1998;273(21):13047–52.
- Sharma UC, Pokharel S, Van Brakel TJ, van Berlo JH, Cleutjens JPM, Schroen B, et al. Galectin-3 marks activated macrophages in failure-prone hypertrophied hearts and contributes to cardiac dysfunction. *Circulation.* 2004;110(19):3121–8.
- Shi Y, Xia Z, Zhang J, Best R, Wu C, Ponder JW, et al. Polarizable atomic multipole-based AMOEBA force field for proteins. *J Chem Theory Comput.* 2013;9(9):4046–63.
- Sörme P, Arnoux P, Kahl-Knutsson B, Leffler H, Rini JM, Nilsson UJ. Structural and thermodynamic studies on cation-II interactions in lectin-ligand complexes: high-affinity galectin-3 inhibitors through fine-tuning of an arginine-arene interaction. *J Am Chem Soc.* 2005;127(6):1737–43. <https://doi.org/10.1021/ja043475p>
- Sörme P, Kahl-Knutsson B, Huflejt M, Nilsson UJ, Leffler H. Fluorescence polarization as an analytical tool to evaluate galectin-ligand interactions. *Anal Biochem.* 2004;334(1):36–47.
- Thiemann S, Baum LG. Galectins and immune responses—just how do they do those things they do? *Annu Rev Immunol.* 2016;34(1):243–64.
- Verkerke H, Dias-Baruffi M, Cummings RD, Arthur CM, Stowell SR. Galectins: an ancient family of carbohydrate binding proteins with modern functions. *Methods Mol Biol.* 2022;2442:1–40. [https://doi.org/10.1007/978-1-0716-2055-7\\_1](https://doi.org/10.1007/978-1-0716-2055-7_1)
- Verteramo ML, Ignjatovi MM, Kumar R, Wernersson S, Ekberg V, Wallerstein J, et al. Interplay of halogen bonding and solvation in protein–ligand binding. *iScience.* 2024;27(4):109636.
- Walker B, Liu C, Wait E, Ren P. Automation of AMOEBA polarizable force field for small molecules: Poltype 2. *J Comput Chem.* 2022;43(23):1530–42. <https://doi.org/10.1002/jcc.26954>
- Wang Z. Tinker9: next generation of Tinker with GPU support. Washington University in St. Louis; 2021. <https://github.com/TinkerTools/tinker9>
- Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol.* 1999;285(4):1735–47. <https://doi.org/10.1006/jmbi.1998.2401>
- Ytreberg FM, Swendsen RH, Zuckerman DM. Comparison of free energy methods for molecular systems. *J Chem Phys.* 2006;125(18):184114.
- Zetterberg FR, MacKinnon A, Brimert T, Gravelle L, Johnsson RE, Kahl-Knutson B, et al. Discovery and optimization of the first highly effective and orally available galectin-3 inhibitors for treatment of fibrotic disease. *J Med Chem.* 2022;65(19):12626–38. <https://doi.org/10.1021/acs.jmedchem.2c00660>
- Zetterberg FR, Peterson K, Johnsson RE, Brimert T, Håkansson M, Logan DT, et al. Monosaccharide derivatives with low-nanomolar lectin affinity and high selectivity based on combined fluorine–amide, phenyl-arginine, sulfur- $\pi$ , and halogen bond interactions. *ChemMedChem.* 2018;13(2):133–7. <https://doi.org/10.1002/cmdc.201700744>
- Zhang C, Lu C, Jing Z, Wu C, Piquemal JP, Ponder JW, et al. AMOEBA polarizable atomic multipole force field for nucleic acids. *J Chem Theory Comput.* 2018;14:2084–108.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Newman L, Vaissier Welborn V. Binding free energy analysis of galectin-3 natural ligands and synthetic inhibitors. *Protein Science.* 2025;34(6):e70143. <https://doi.org/10.1002/pro.70143>