

Computational Biology

Further analyses of APRIL/APRIL-receptor/glycosaminoglycan interactions by biochemical assays linked to computational studies

Mateusz Marcisz^{2,3}, Bertrand Huard⁴, Agnieszka G Lipska² and Sergey A Samsonov^{1,2}

²Faculty of Chemistry, University of Gdańsk, ul. Wita Stwosza 63, Gdańsk 80-308, Poland, ³Intercollegiate Faculty of Biotechnology of UG and MUG, ul. Abrahama 58, Gdańsk 80-307, Poland, and ⁴TIMC-IMAG, Université Grenoble-Alpes, CNRS UMR, La Tronche 5525, France

¹To whom correspondence should be addressed: Tel: +48-58-523-51-66; Fax: +48-58-523-50-12; e-mail: sergey.samsonov@ug.edu.pl

Received 4 January 2021; Revised 23 February 2021; Editorial Decision 23 February 2021; Accepted 23 February 2021

Abstract

A proliferation-inducing ligand (APRIL) is a member of the tumor necrosis factor superfamily. APRIL is quite unique in this superfamily for at least for two reasons: (i) it binds to glycosaminoglycans (GAGs) via its positively charged N-terminus; (ii) one of its signaling receptors, the transmembrane activator and CAML interactor (TACI), was also reported to bind GAGs. Here, as provided by biochemical evidences with the use of an APRIL deletion mutant linked to computational studies, APRIL–GAG interaction involved other regions than the APRIL N-terminus. Preferential interaction of APRIL with heparin followed by chondroitin sulfate E was confirmed by *in silico* analysis. Both computational and experimental approaches did not reveal the heparan sulfate binding to TACI. Together, computational results corroborated experiments contributing with atomistic details to the knowledge on this biologically relevant trimolecular system. Additionally, a rigorous high-throughput analysis of the free energy calculations data was performed to critically evaluate the applied computational methodologies.

Key words: APRIL, ELISA, flow cytometry, glycosaminoglycans, MM/GBSA

Introduction

Heparin (HP), chondroitin sulfate-C (CSc) and -E (CSe) belong to glycosaminoglycans (GAGs). GAGs are long, linear, anionic and periodic polysaccharides playing a crucial role via interactions with a variety of proteins in the extracellular matrix processes. They are made of repeating disaccharide unit consisting of an amino sugar and a uronic acid or galactose (Varki *et al.* 2015). Those saccharide units may manifest different sulfation patterns that influence the polysaccharide conformational and binding properties (Habuchi *et al.* 2004). In many cases, protein–GAG interactions are considered as nonspecific and mostly electrostatic-driven due to the high negative charge of the polysaccharides and positive charge of the protein-binding sites

(Imberty *et al.* 2007). Some of the proteins that interact with GAGs belong to the group of growth factors (Uciechowska-Kaczmarzyk *et al.* 2018; Bojarski *et al.* 2019) and chemokines (Derler *et al.* 2017; Nordsieck *et al.* 2018; Penk *et al.* 2019). In case of fibroblast growth factors (FGF), GAGs may form complexes with FGF1 (Digabriele *et al.* 1998) and FGF2 (Faham *et al.* 1996). GAGs can enhance the activity of the growth factors by either changing their conformation or by binding multiple FGFs and thus facilitating oligomerization of FGFR receptors, which plays a role in cell signaling (Mason 1994; Faham *et al.* 1996). Our recent computational study implementing a microsecond-scale molecular dynamics (MD) simulations (Bojarski *et al.* 2019) showed few novel insights on the

HP-FGF1 interactions and helped to deepen the knowledge on the topic of protein–GAG molecular systems. In particular, it was shown that the length of the simulation could be crucial for the calculation of the protein–GAG molecular system dynamic and energetic parameters, while conformational selection mechanism of binding as well as recognition specificity determinant were proposed from these long simulations for this complex. Another recent computational study on HP (Uciechowska-Kaczmarzyk et al. 2018) shows its effect on the dynamics of vascular endothelial growth factor (VEGF), a key player in the angiogenesis and regenerative processes, arthritis and cancer (Risau 1997). In this work, it was demonstrated that GAG binding could induce global conformational changes of a protein target, rendering its capability to bind its receptor on the cell membrane (Uciechowska-Kaczmarzyk et al. 2018). One more growth factor that is known to bind HP and heparan sulfate is transforming growth factor β 1 (TGF- β 1) (McCaffrey et al. 1992; Lyon et al. 1997). TGF- β 1's role is to regulate proliferation, adhesion, cell migration and differentiation (Moustakas et al. 2001). There is evidence that sulfated hyaluronan derivatives are able to bind this growth factor and so can enhance or inhibit its activity depending on GAG sulfation and the order of binding events in its tertiary complex with its receptors (Van Der Smissen et al. 2013; Koehler et al. 2017). Second important group of proteins that interacts with GAGs are chemokines (Derler et al. 2017; Crijns et al. 2020). They are a large group of predominantly proinflammatory cytokines, and they may influence the cell in a variety of different ways—while some of them can promote angiogenesis, tumor growth and metastasis, others can inhibit them (Luster 1998). GAGs are also known to bind these molecules mediating the activation of the leukocytes by affecting the binding of a respective chemokine to a G protein-coupled receptor (Larsen et al. 1989). Many computational and experimental studies successfully investigated GAGs' effects and interactions with chemokines/cytokines, such as CXCL-8 (Gandhi and Mancera 2011; Joseph et al. 2015; Nordsieck et al. 2018), IL-10 (Künze et al. 2014; Gehrcke and Pisabarro 2015; Künze et al. 2016), CXCL-12 (Panitz et al. 2016) and CXCL-14 (Penk et al. 2019). This proved theoretical approaches not only as useful complementation to experimental studies but also as an important and stand-alone work.

A proliferation-inducing ligand (APRIL) is a member of the tumor necrosis factor (TNF) superfamily (Hahne et al. 1998). It is produced first as a transmembrane protein before being processed by furin proteases to act as a soluble factor (López-Fraga et al. 2001). Notably, APRIL also binds to the GAGs of the heparan sulfate (HS) class (Hendriks et al. 2005; Ingold et al. 2005). Such binding is quite unique in the TNF superfamily since only one other member, ectodysplasin A, has been reported to date to bind GAG (Swee et al. 2009). APRIL binding to GAG allows its oligomerization to efficiently signal into target cells (Kimberley et al. 2009). Indeed, unlike other TNF ligands such as the TNF itself, the soluble APRIL trimer formed by noncovalent, mostly hydrophobic, interactions between beta-pleated sheets is not active to signal its two receptors transmembrane activator and CAML interactor (TACI) and B-cell maturation antigen (BCMA) (Bossen et al. 2008). The numerous negatively charged sulfate residues along the heparan chain of GAG create a platform with multiple binding sites for APRIL, hence mediating oligomerization. The GAG-binding region of APRIL has been located in its N-terminus that contains a stretch of positively charged lysine, three and four in the human and mouse molecules, respectively. Addition of a cross-linking anti-flag antibody renders trimeric flag-tagged APRIL signalization active, indicating that APRIL does not need a high order of oligomerization to signal. One of the APRIL signaling receptor, TACI but not BCMA, was further shown to interact with GAGs

(Bischof et al. 2006; Sakurai et al. 2007; Moreaux et al. 2009). Such ternary complex between a ligand, a coreceptor and a receptor may resemble the one described for FGF/FGF-R (Pomin 2016). APRIL main cellular targets are the antibody-producing plasmocytes (Baert et al. 2018). On the surface of these cells, TACI and/or BCMA, depending on their stage of differentiation, are present as the APRIL signaling receptors. They also express ubiquitously a unique GAG, CD138, also known as syndecan-1 (Wijdenes et al. 1996). CD138 has a mixed composition of HS and chondroitin (CS) chains (Kokenyesi and Bernfield 1994). However, only HS on CD138 appears to play a role on APRIL binding (Matthes et al. 2015). Recently, a new target cell for APRIL has been identified in the central nervous system with astrocytes (Baert et al. 2019). Notably, APRIL binds to CS GAG on astrocytes, and selectivity in APRIL binding according to CS types was observed. Here, we further investigate the binding of APRIL to HP and CS GAG in a computational and experimental study.

Materials and methods

Structures

GAG Structures. All the GAG structures—heparin (HP) tetramer/dp4 and hexamer/dp6 (dp stands for degree of polymerization), chondroitin-4,6-sulfate (CSe) dp4 and dp6, chondroitin-6-sulfate (CSc) δ dp4 and dp6—were constructed from the building blocks of the sulfated GAG monomeric units' libraries (Pichert et al. 2012) that are compatible with the AMBER16 package (Case et al. 2018). GLYCAM06 force field (Kirschner et al. 2008) and literature data (Huige and Altona 1995) were the sources of GAGs' charges.

Protein Structures. The structure from PDB ID 4ZCH (2.43 Å) (Schuepbach-Mallepell et al. 2015) of the single-chain human APRIL protein from the APRIL–BAFF–BAFF complex was used for the construction of truncated human variant trimer— H_{115} APRIL (136 amino acid residues starting from HIS 115). For this, chimera (Pettersen et al. 2004) and Modeller (Šali and Blundell 1993) were used to obtain the model of the human H_{115} APRIL trimer based on its murine homolog—PDB ID 1XU1 (1.90 Å) structure (Hymowitz et al. 2005). Later, using AMBER software package, minimization and 10 ns equilibration by MD simulation were performed to obtain the structure used for the further studies (see the details on the MD protocols Molecular dynamics).

In case of the full-length—natural form— A_{105} APRIL protein variant, additional 10 amino acid residues were added to the N-terminus (146 amino acid residues in total, starting from ALA 105) of each monomer. Coarse-grained modeling UNRES (from UNited RESidue) software (Liwo et al. 1997) was used to predict the structure of the trimer (see the protocol Calculation of full-length APRIL protein model in UNRES). Five different models corresponding to the energy minima were obtained. To choose the best one in terms of potential GAG binding, docking of HP dp4 to APRIL trimer and MD simulations were performed. Models with the overall lowest binding free energy were chosen for further analysis. To calculate energies, molecular mechanics generalized Born surface area (MM/GBSA) and linear interaction energy (LIE) analysis methods were used (see the protocols below).

Peptide fragment corresponding to the N-terminus of the full-length human variant and missing in the truncated one (ALA-VAL-LEU-THR-GLN-LYS-GLN-LYS-LYS-GLN) was minimized, and extensive MD simulation of 12 μ s was performed in order to analyze its structural properties.

Calculation of full-length APRIL protein model in UNRES

In order to calculate the conformations of the N-terminal fragments of each chain, we applied a coarse-grained multiplexed replica exchange molecular dynamics (MREMD) (Sugita and Okamoto 2000; Young and Pande 2003) approach implemented in UNRES (Adcock and McCammon 2006; Czaplewski et al. 2009). The protocol was similar to that used in our previous work (Vallet et al. 2018; Porthoff et al. 2019). Distance restraints were imposed on protein, except the first 10 amino acids of each chain. MREMD simulation consisted of 40 trajectories run at temperatures from 240 K to 350 K, with two trajectories for every temperature. Each trajectory consisted of 1.6×10^7 MD steps with 4.89 fs length. Only conformations from the second part of the simulation were taken into further analysis, with the use of the weighted histogram analysis method (WHAM) (Kumar et al. 1992). The next step was minimum variance cluster analysis (Ling and Späth 1981) of the conformational ensemble at $T = 300$ K, which enabled us to obtain five clusters, ranked according to summary probabilities of the ensembles and containing the most probable structures to the cluster with the least probable structures. For each cluster, one representative structure, closest to the cluster centroid, was selected as the representative conformation.

Electrostatic potential calculations

Amber16 PBSA (Poisson–Boltzmann surface area) program was used to calculate and visualize the solvent-mediated electrostatic potentials of different variants of the APRIL protein using a default 1 Å grid spacing step. Later on, it was visualized using VMD program (Humphrey et al. 1996) to assess the regions of potential GAG binding to APRIL protein. Previously, this method proved to be successful in GAG-binding regions' predictions (Samsonov and Pisabarro 2016).

Molecular docking

Autodock 3 (Morris et al. 1998) was used for docking simulations. This program with this particular version yielded the best results among other docking programs for protein–GAG systems (Samsonov and Pisabarro 2016; Uciechowska-Kaczmarzyk et al. 2019). Maximum gridbox size was used ($126 \text{ Å} \times 126 \text{ Å} \times 126 \text{ Å}$), which covered the entire APRIL region of predicted GAG-binding region with the default grid step of 0.375 Å. Lamarckian genetic algorithm was used for 1000 independent runs. The size of 300 for the initial population and 10^5 generations for termination conditions were chosen. 9995×10^5 energy evaluations were performed. Clustering was performed using DBSCAN algorithm (Ester et al. 1996) on top 50 docking results. The metric used for clustering accounted for the equivalence of the atoms of the same atomic type, which is more appropriate than the classical RMSD for periodic ligands (Samsonov et al. 2014). One to two clusters of each GAG's docking solutions were chosen for the further analysis. Each time clustering parameter was chosen individually to obtain one to three representative clusters.

Molecular dynamics

Every all-atom MD simulation of different APRIL variants complexes obtained from molecular docking was performed using AMBER16 software package. Truncated octahedron TIP3P periodic box of 8 Å water layer from the box's border to solute was used to solvate complexes. Charge was neutralized with Cl^- counterions.

Cysteines were connected to form appropriate disulfide bridges according to the structure from the PDB (PDB ID 1XU1) (Hymowitz et al. 2005; Schuepbach-Mallepell et al. 2015). Energy minimization was performed preceding the production MD runs: 500 steepest descent cycles and 10^3 conjugate gradient cycles with 100 kcal/mol/Å² harmonic force restraints, continued with 3×10^3 steepest descent cycles and 3×10^3 conjugate gradient cycles without any restraints. Following minimization steps, the system was heat up to 300 K for 10 ps with harmonic force restraints of 100 kcal/mol/Å². Then, the system was equilibrated at 300 K and 10^5 Pa in the isothermal isobaric ensemble for 500 ps. Afterward, the actual MD run was carried out in the same isothermal isobaric ensemble for either 10 ns or 100 ns (except for the 10 amino acid residues peptide, in which case, 12 μs MD run was performed). Particle mesh Ewald method for treating electrostatics and SHAKE algorithm for all the covalent bonds containing hydrogen atoms were implemented in the MD simulations.

Binding free energy calculations

To calculate free energy and per-residue energy decomposition, the obtained trajectories from MD simulations were analyzed using AMBER16 by two approaches—MM/GBSA model $\text{igb} = 2$ (Onufriev et al. 2002) and LIE analysis with dielectric constant of 80, performed by CPPTRAJ scripts. Particular frames taken for this analysis varied for different simulations to be representative in terms of the structural convergence.

Dynamic molecular docking

Compatibly with the MD simulations described Molecular dynamics, AMBER16 was used for dynamic molecular docking (DMD). This method allows for full flexibility of both the receptor and ligand as well as for taking into account the explicit solvent (Samsonov et al. 2014). First, a ligand was placed at 20–30 Å from the surface of the protein (50–60 Å from the center of the protein), which was significantly more than the cutoff value = 8 Å that was used in the MD to avoid any influence on the dynamics of the ligand at the beginning of the docking run. Then, the truncated octahedron TIP3P periodic box of 4 Å water layer from the box's border to solute was used to solvate the complex, and the charge was neutralized using Cl^- counterions. Disulfide bonds between cysteines were created accordingly with the structure from the PDB. Minimization and equilibration runs were performed as described previously in the Molecular dynamics section. After equilibration, the distances from ligand (O4S atom of residue 47Y) to protein (O atom of residue GLY403) were calculated and were assigned to the initial distance in the targeted MD run. Afterward, first 4 ns MD run was performed with the biased potential of 200 kcal/mol/Å² applied to the above-mentioned atoms of the receptor and the ligand as described in our previous work (Samsonov et al. 2014), employing Jarzynski procedure (Jarzynski 1997; Park and Schulten 2004). This step was repeated 100 times to obtain 100 different docking poses. Next, structures from final frames from all runs were taken as the starting structures for the unbiased MD simulation; each protein–GAG complex was solvated with 8 Å layer of TIP3P water in a form of truncated octahedron; Na^+ or Cl^- counterions were added and disulfide bonds were created. Minimization and equilibration steps were performed once more under the same conditions as in the first step of the DMD simulation. Finally, 10 ns MD production runs were carried out in the same isothermal isobaric ensemble for each

of the 100 docking poses. The obtained structures were considered as docking poses and were scored used the MM/GBSA protocol for all 10 ns of the MD production run.

Cells and reagents

L363 and 293 T HEK cells were obtained from the American tissue culture collection. All cells were propagated in RPMI 1640 medium (Gibco BRL, US) containing 10% fetal calf serum (Eurobio, France). Biotinylated anti-flag (clone M2) was from Sigma (US). The anti-TACI (1A1, rat immunoglobulin G2a [IgG2a]) and soluble human Thy-1 (aa 20-130), BCMA (aa 2-54) TACI (aa 2-118) fused to human Fc were obtained from EnzoLife sciences (Switzerland). Expression constructs for Fc-tagged and flag-tagged human A_{105} APRIL (aa 105-246) and H_{115} APRIL (aa 115-246) have been described previously (Baert et al. 2018). SDS-PAGE analysis was performed to confirm dimerization of APRIL trimer proteins (Supplementary Figure S1). Fc-tagged and flag-tagged proteins were produced transiently in 293 T HEK cells following polyethylenimine-based transfection in serum-free opti-mem medium (Gibco BRL). Fc-tagged and flag-tagged proteins were purified with Protein-A Sepharose (GE Healthcare, US) and anti-flag-Sepharose (Sigma, US) and acidic elution with 0,1 M glycine, pH: 2.5, respectively. Positive fractions were pooled and dialyzed against PBS before use. The plasmid encoding for human TACI has been described by (Schwaller et al. 2007). Heparin (Liquemin, 5000 IU/mL, Drossapharm, Switzerland) was used at 1/500. Heparan sulfate proteoglycan (HSPG) from mouse basement membrane sarcoma were obtained from Sigma. HSPG were biotinylated with biotin hydrazide and carbodiimide as previously described (Ahmed and Huard 2021).

Enzyme-linked immunosorbent assay

Proteins coating was performed overnight at 10 μ g/mL in 50 μ L of PBS at 4°C. Plates were blocked in PBS with 1% BSA for 1 h at RT. Incubation with ligands were performed for 1 h at RT. Washing buffer was PBS with 1% BSA, 0.05% Tween 20. Streptavidin conjugated to horseradish peroxidase (R&D Systems, US) and 3,3',5,5'-tetramethylbenzidine (Sigma) were used to reveal binding of biotinylated HSPG. Reaction was stopped with H_2SO_4 2 N. Optical density was read at 450 nm on a VICTOR multilabel plate reader (Perkin Elmer, UK).

Flow cytometry

$0,5 \times 10^6$ cells were stained at 4°C in PBS with 1% BSA for 30 min. Secondary reagents included goat antirat (anti-TACI, 10 μ g/mL) and goat antihuman immunoglobulin G (IgG) (Fc-APRIL, TACI-Fc) conjugated to Alexa 488. Streptavidin conjugated to Alexa 488 (BD Biosciences) was used to detect HSPG and anti-flag binding. Washes were performed with PBS. Fluorescence was acquired on a BD Accuri C6 flow cytometer.

Results and discussion

In this work, APRIL–GAG interactions were comprehensively analyzed. For all the experiments, trimers of APRIL were used (with the addition of the hexameric variants that were used for biological experiments) to reproduce the conditions in the cell. It was shown in enzyme-linked immunosorbent assay (ELISA) and flow cytometry experiments that HSPG binds to both wild-type and truncated variants of APRIL protein. Potential region of GAG binding on the

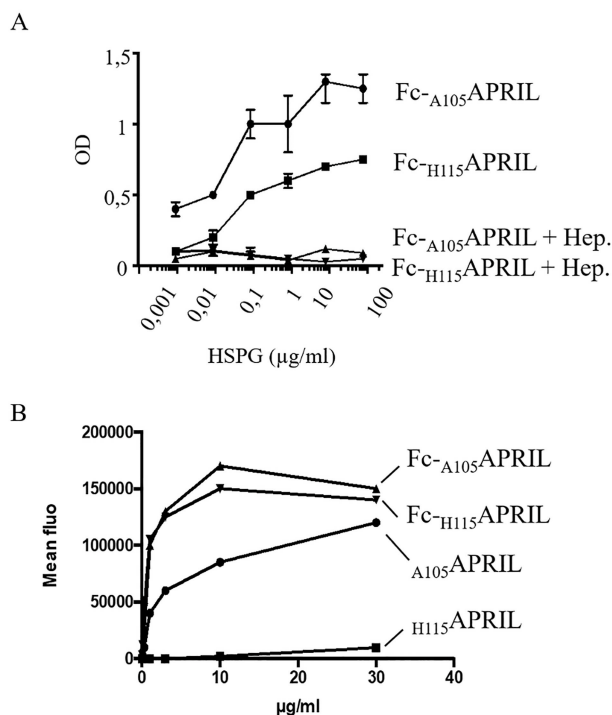


Fig. 1. Binding of oligomerized H_{115} APRIL to HSPG. (A) The binding of increasing concentrations of biotinylated HSPG to the indicated forms of coated APRIL was tested by ELISA. In some conditions, heparin was used as a competitor. Results are shown with the mean and min/max values of duplicates. Three independent experiments were performed. (B) Binding of trimeric APRIL and hexameric Fc-APRIL on HSPG expressed at the surface of HEK cells was assessed by flow cytometry. The results are representative of at least five for (A) and two for (B) independent experiments.

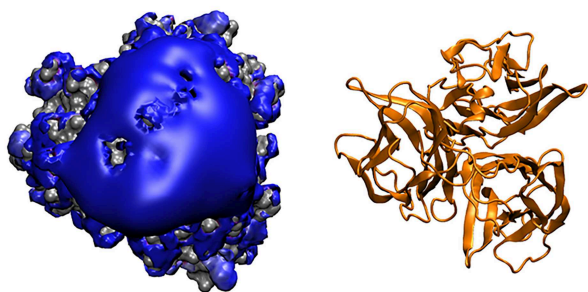
surface of the APRIL was inspected with the PBSA method. Then, using computational approach, the results from ELISA/flow cytometry were confirmed. Furthermore, interactions of APRIL protein with different lengths of HP, CSe and CSc were examined. It was shown which residues promote GAG binding and which obstruct it. Then, the structural properties of N-terminal fragment that truncated APRIL was lacking were analyzed. Afterward, experimental and computational investigations of APRIL's receptors—TACI and BCMA—and their potential binding to HP were carried out. At the end, an in-depth analysis of the performance and prediction power of tools used in this study was carried out.

Binding of H_{115} APRIL/ A_{105} APRIL with HSPG

We assessed the binding of H_{115} APRIL to HSPG by ELISA. We could detect a binding to soluble HSPG when H_{115} APRIL was oligomerized upon coating onto the plastic surface (Figure 1A). This binding was to a lesser extent than for coated A_{105} APRIL. We also detected by flow cytometry that a binding of H_{115} APRIL to HSPG expressed at the surface of HEK cells, but again, only when it was oligomerized by dimerization with an Fc fusion partner (Figure 1B). In this latter experiment, trimeric A_{105} APRIL could bind to the HEK cells. Taken together, these experiments show that H_{115} APRIL lacking its N-terminal tail could still bind to HSPG but with an overall lower affinity since requiring oligomerization.

After the experimental confirmation of binding of APRIL protein variants to HSPG, we performed a series of computational experiments to gain atomistic insights into the APRIL–GAG interactions.

Full length APRIL



Truncated APRIL

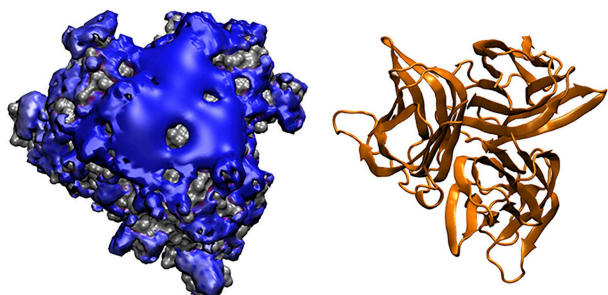


Fig. 2. Positive electrostatic potential isosurfaces (colored in blue; 4 kcal/mol \cdot e $^{-1}$ in case of truncated APRIL and 5 kcal/mol \cdot e $^{-1}$ in case of full-length APRIL) obtained by PBSA approach show potential capability of N- and C-terminal regions of the APRIL protein to bind GAGs.

PBSA-based prediction of GAG-binding region

To predict GAG-binding regions on the APRIL protein surface, PBSA method from AMBER suite was used. It was previously proven that, due to the highly charged nature of these ligands, such methodology is successful in predicting GAG-binding regions on the protein surface (Samsonov and Pisabarro 2016). The results obtained with the PBSA method showed positive electrostatic potential in the N-terminal region of each monomer that is spatially close in the trimeric APRIL (Figure 2). This region with the three lysines, LYS 110, LYS 112 and LYS 113, was previously demonstrated in binding assays with HS-positive cells to be responsible for binding GAGs (Hendriks et al. 2005; Ingold et al. 2005). Nevertheless, we found out that the C-termini may contribute to the positive potential suggesting that APRIL–GAG interactions may be more complex than previously thought.

Computational investigation of APRIL–GAG interactions

HP dp4, HP dp6, CSe dp4, CSe dp6, CSc dp4 and CSc dp6 were analyzed for their binding properties to full-length soluble APRIL (A_{105} APRIL) and a N-term truncated variant (H_{115} APRIL) using MM/GBSA and LIE methods, with protocols calibrated for this system as described at the end of the Results section. To obtain the A_{105} APRIL protein structure, we applied coarse-grained approach to calculate theoretical models for the protein as there is no experimental structure available yet. In order to achieve this, we used H_{115} APRIL variant PDB structure and modeled additional 10 amino

acid residues using UNRES software as described in the Materials and methods section. This full-length protein model was used for further APRIL–GAG interactions analysis along with truncated variant for molecular docking and binding free energy assessment. Both MM/GBSA and LIE were in agreement that HP binds the strongest out of the compared GAGs. The weakest binding was found for CSc, while the CSe bound stronger than CSc but weaker than HP, suggesting not only net electrostatic effect on binding but also the specific role of sulfation pattern for CS.

All the binding energies obtained by LIE and MM/GBSA for both A_{105} APRIL and H_{115} APRIL variants are listed in the Table I. Overall, GAGs bind much stronger to A_{105} APRIL than to H_{115} APRIL. This finding is in agreement with our PBSA electrostatic potential analysis that suggested much higher positive potential in the region of N-terminus. It is also clear that dp6 GAGs are bound stronger than dp4 counterparts, especially for CSe and HP GAGs. The only exception is CSc dp4 in the case of A_{105} APRIL that shows insignificantly lower MM/GBSA and LIE values than for CSc dp6. However, the overall differences in energies are lower for CSc than for other GAGs, which suggests very weak or no binding for this GAG. The differences in the results obtained by MM/GBSA and LIE methods may appear due to few factors. First of all, LIE is a less complex method and may not account for as many energy components, and in particular, in terms of solvent treatment, as the MM/GBSA does. There are several differences in free energy calculation in both methods (Genheden and Ryde 2011): LIE takes into account the Van der Waals (VdW) component and the electrostatic component (ELE) in vacuo scaled by a dielectric constant to consider the interactions between the ligand with the receptor and solvent environments; MM/GBSA calculates free energy using VdW and electrostatics in vacuo energies, polar solvation energy partially including the entropic component of the solvent, nonpolar solvation free energy. Second, LIE is used in our work with standard parameters (dielectric constant of 80), which should be calibrated for each particular molecular system type using experimental data if available.

Significant residues in the APRIL–GAG-binding site

All of the mentioned GAGs were docked with Autodock3 software to the N-/C-terminus region of the APRIL protein (Figure 3). Per-residue free energy decomposition analysis allowed to propose the residues defining the putative binding site. The most important residues for H_{115} APRIL were: HIS 115, ARG 143, ARG 144, GLY 145, ARG 146 and LYS 249 (Supplementary Figures S2 and S3), while the following residues provided the most unfavorable contribution to the binding: ASP 159, GLU 185, GLU 191, ASP 223 and LEU 250 (Supplementary Figure S4). For A_{105} APRIL, the same residues were found to be disruptive (Figure 4). However, when analyzing the most important residues favorable for binding, there was a significant difference because of the additional LYS 110, GLN 111, LYS 112 and 113 that were present in the N-terminus (Figure 5 and Supplementary Figure S2). This confirmed that A_{105} APRIL is more prone to binding GAGs.

Heparin binding to the LYS-rich N-term peptide of APRIL

Additional computational experiments were performed to analyze more comprehensively APRIL–GAGs interactions. We took the first 10 amino-acid residues from the N-terminus of the APRIL protein that are missing in the truncated form and simulated (as described in the Methods and materials section) them for over 12 μ s to assure

Table 1. MM/GBSA and LIE binding free energy analysis for the truncated and full-length APRIL in complex with HP dp4 and CSc dp4

	H115APRIL, ΔG (kcal/mol)		H115APRIL, ΔG (kcal/mol)	
	MM/GBSA	LIE	MM/GBSA	LIE
HP dp4	−56.4 ± 8.3	−38.1 ± 5.8	−119.9 ± 20.8	−61.9 ± 9.2
HP dp6	−88.4 ± 16.7	−50.3 ± 9.5	−141.1 ± 23.5	−69.3 ± 11.7
CSe dp4	−34.1 ± 0.28	−31.9 ± 3.0	−52.1 ± 9.8	−42.6 ± 8.5
CSe dp6	−47.9 ± n/a	−40.7 ± n/a	−90.0 ± 18.0	−62.8 ± 9.9
CSc dp4	−33.7 ± 8.1	−29.3 ± 5.9	−63.7 ± 20.3	−43.8 ± 11.0
CSc dp6	−36.5 ± 1.0	−30.8 ± 12.8	−50.9 ± 13.2	−40.3 ± 9.4

n/a: not available; only two values were obtained for this system.

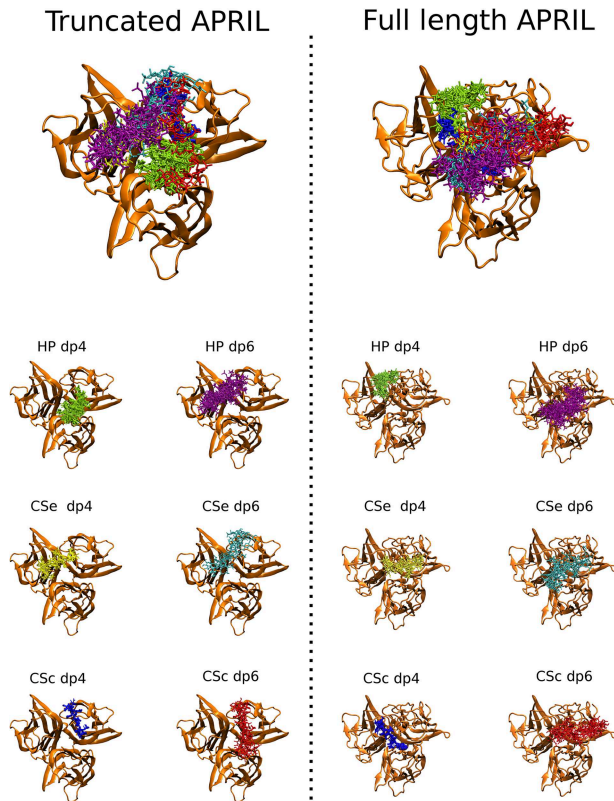


Fig. 3. The representative clusters of GAG solutions (in sticks) obtained for H115APRIL/APRIL (in orange cartoon) by Autodock3 docking. On the top of the panel, all clusters are shown together, while on the bottom side of the panel they are shown separately per each GAG.

proper folding. Then, clustering (using DBSCAN algorithm) with a set of different parameters was used to obtain the representative structures (Figure 6). We have chosen the cluster that appeared in 10% of the frames of the 12 μs MD simulation. The peptide was analyzed in terms of its structural parameters. DSSP algorithm (Kabsch and Sander 1983) was used to assess the secondary structure of the peptide and N-terminus of A105APRIL to compare their properties (Supplementary Figure S5). The most common elements in case of the peptide and two out of three N-termini from A105APRIL were the bend and the turn, while in case of the third N-terminus, a high number of 3₁₀-helix was found. Afterward, using Autodock3, HP dp6 was docked to this 10 amino acid residue peptide. It was followed by 10 ns MD runs that were repeated 50 times (each individual run

for each of the 50 best docking poses). Those runs were analyzed in terms of binding free energy, using both MM/GBSA and LIE approaches. Mean value of −29.5 kcal/mol and −32.9 kcal/mol for the 50 runs were obtained, respectively. Experiment was repeated using DMD method (as described in Methods and materials section) and was analyzed with MM/GBSA. Obtained binding free energy values were very close to those obtained after rigid docking to the most probable peptide conformer with Autodock3 (−30.0 vs. 29.5 kcal/mol). When compared to the values of −141.1 kcal/mol (MM/GBSA) and −69.3 kcal/mol (LIE) in case of binding to APRIL, it is clear that the strength of the binding to the peptide representing N-terminus is far from that of binding to the protein. In fact, in one of the cases, we even observed start of the dissociation of HP from the peptide during the MD run. This may indicate that the N-terminus itself could not be sufficient for proper GAG binding and other peripheral/C-terminus residues in the protein (such as ARG 143, ARG 144, GLY 145, ARG 146 and LYS 249) are essential.

No evidence for an interaction of TACI with HSPG in biological assays

We next assessed the binding of HSPG to TACI in ELISA by coating TACI-Fc. No binding of biotinylated HSPG ranging from 1 ng/mL up to 100 μg/mL was detected (Figure 7A). In this experiment, coated TACI was biochemically active since it could bind to a complex of Fc-A105APRIL/HSPG. In the latter condition, presence of A105APRIL did not potentiate the interaction of TACI with HSPG since the binding recorded was not superior to the conditions with coated BCMA. When APRIL was coated, we still did not observe any potentiation of the TACI/HSPG interaction (Figure 7B). We next assessed the interaction with TACI expressed at the surface of the transfected cells. We again did not detect the binding of biotinylated HSPG at any concentrations tested (Figure 7C). Multiple myeloma cells such as the L363 cell line express the HSPG, syndecan 1. On these cells, A105APRIL could bind in an HSPG-dependent manner once inhibited by heparin (Figure 7D). We could not detect the binding of up to 30 μg/mL of TACI-Fc to L363 cells.

APRIL's receptor TACI and BCMA—their contribution to GAG binding

It was proposed that APRIL's receptor TACI was able to bind to HSPG while BCMA could not (Bischof et al. 2006). Those findings are in contrary to our experiments and to the data published recently by Kowalczyk-Quintas et al. (2019). To further analyze those contra-

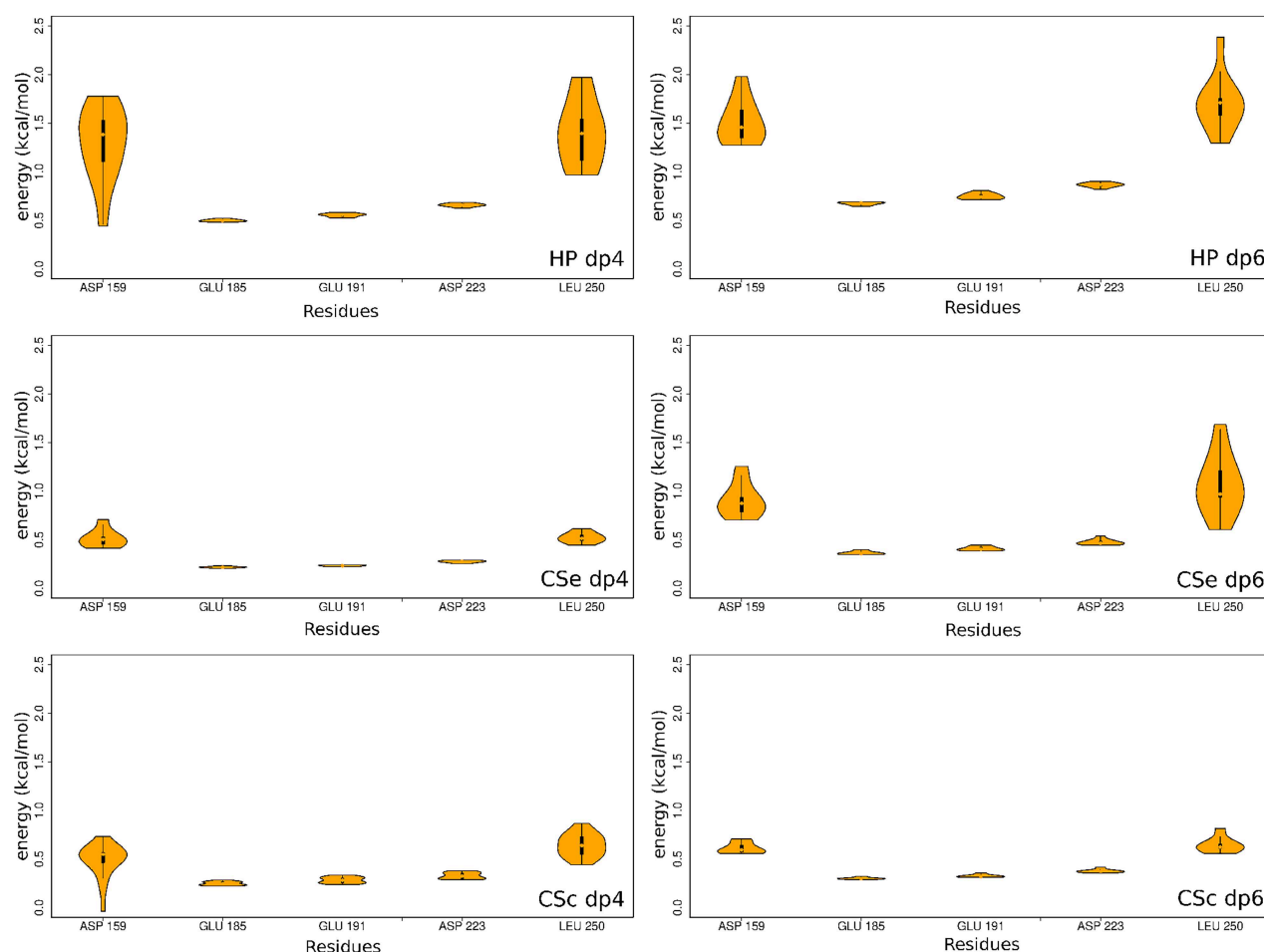


Fig. 4. Per-residue free energy decomposition of the residues that show the most unfavorable GAG-binding impacts in case of the A_{105} APRIL variant. Violin plots represent the energy distribution for particular residues binding different GAGs from all MD simulations with the starting structures obtained by Autodock3 docking.

dictory findings and to understand TACI–GAGs interactions at the molecular level, we first performed electrostatic potential analysis using the PBSA method from AMBER suite. APRIL protein with soluble fragments of TACI or BCMA in complexes (PDB IDs 1XU1 and 1XU2, respectively) was used for this analysis. The obtained data indicate more favorable electrostatic potential in case of complex of APRIL–TACI than in APRIL–BCMA (Supplementary Figure S6). Then, we docked HP dp6 to both TACI- and BCMA-soluble fragments using Autodock3. Obtained binding poses were clustered (Methods and materials section), and poses from the best clusters for TACI (eight poses) and BCMA (10 poses) were used in the MD simulations. Afterward, the MM/GBSA method was applied to assess free binding energies in TACI–HP dp6 and BCMA–HP dp6 complexes. MM/GBSA analysis showed free binding energy values of -35.4 kcal/mol (standard deviation [SD] 7.9) and -15.5 kcal/mol (SD 8.7) for TACI and BCMA complexes, respectively. Those values clearly indicate the distinguishable strength of potential GAG binding to two different APRIL's receptors. Free binding energy of -15.5 kcal/mol suggests that HP binding to BCMA is unlikely. -35.4 kcal/mol value from the MM/GBSA analysis in case of the TACI–HP complex could suggest potential weak binding. LIE analysis has shown the same mean value of -35.4 kcal/mol (SD 6.8). However, the value of -35.4 kcal/mol indicates some uncertainty,

and we are unable to clearly and confidently claim for this particular complex whether there is a binding or not. It is also worth mentioning that HP is the most sulfated form of HS and thus the most charged variant of HS, which consists of mixture of GAGs with different sulfation patterns. Taking this into account, the HP may show slightly better binding properties in this particular complex than the rest of the HS family. Therefore, the borderline free binding energy values for the HP binding to TACI could suggest that potentially weaker binding of HS to TACI would not be sufficient enough for an effective binding that could be observed in the experiment.

DMD in GAG-binding site prediction

In order to further analyze and understand the predictive power of our docking results from Autodock3 and to provide more details on GAGs binding to A_{105} APRIL, we performed DMD for HP dp4 and 6. DMD should allow for more flexibility for the molecules in the docking process, which could suggest that DMD values may be more reliable. Another advantage of DMD is that it allows for the inclusion of explicit solvent in local molecular docking. Previously, it was shown that in comparison to AD3, the DMD method was capable to reliably identify the receptor residues contributing most to binding and it had higher complex structure prediction performance,

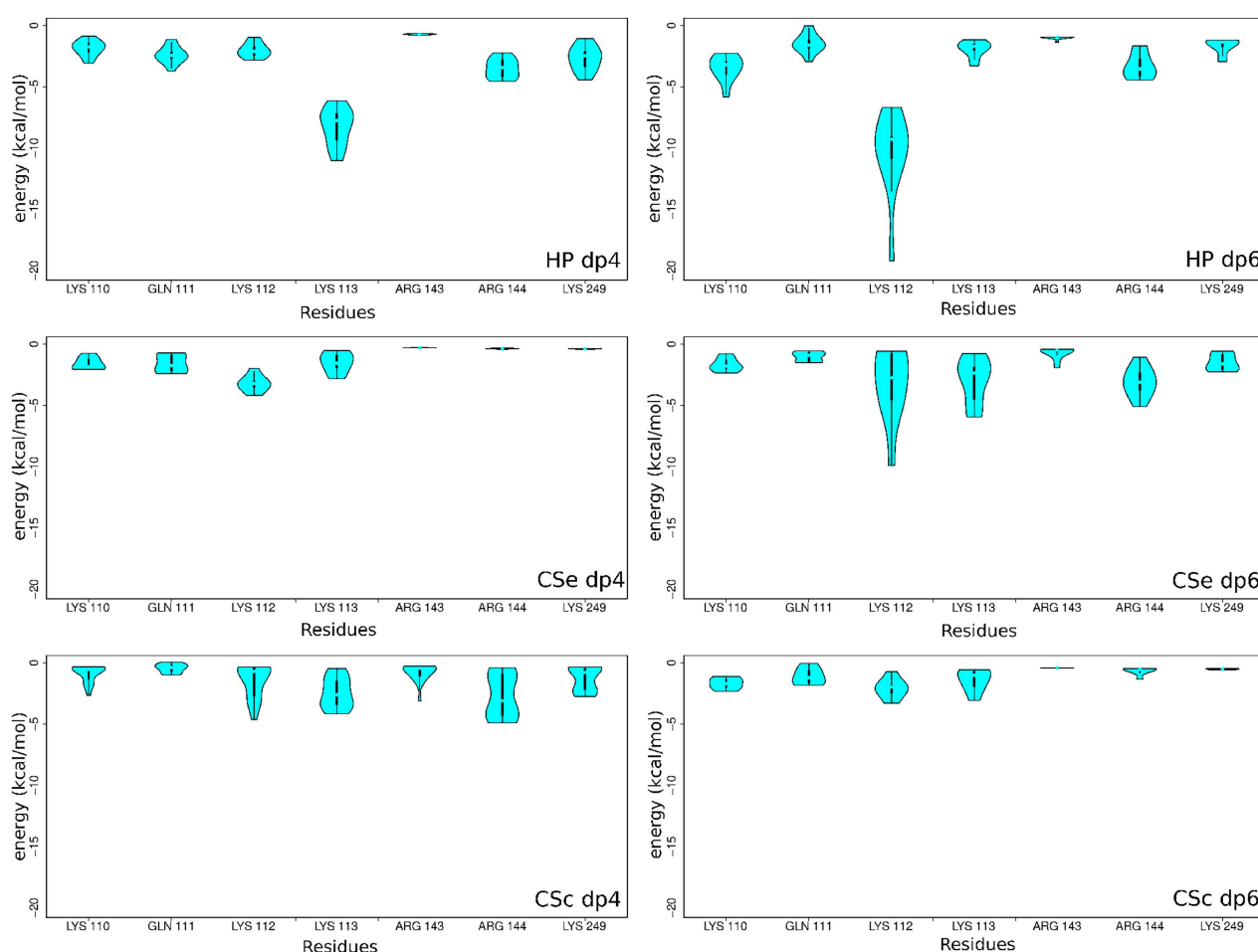


Fig. 5. Per-residue free energy decomposition of the residues that show the most favorable GAG-binding impacts in case of A_{105} APRIL variant. Violin plots represent the energy distribution for particular residues binding different GAGs from all MD simulations with the starting structures obtained by Autodock3 docking.

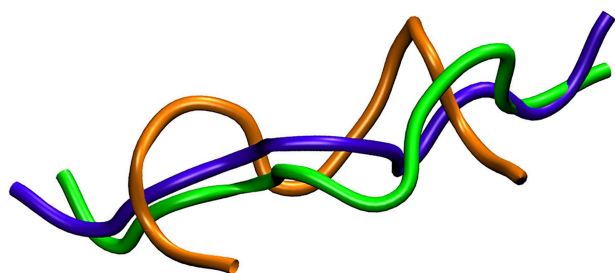


Fig. 6. Three representative structures of the N-terminal fragment models (aa 105–114) from the A_{105} APRIL protein obtained from the 12 μ s MD simulation. The DBSCAN algorithm was used for clustering with following parameters: orange—epsilon = 2 and minpoints = 3, green—epsilon = 2 and minpoints = 5, violet (chosen for the further analysis)—epsilon = 3 and minpoints = 4.

especially for systems with highly flexible and negatively charged ligands like GAGs (Gehrcke and Pisabarro 2015; Salbach-Hirsch et al. 2015; Babik et al. 2017) 100 independent docking runs were obtained for each GAG. MM/GBSA binding energy analysis showed an average energy of -84.8 kJ/mol for the HP dp4 and -93.1 kJ/mol for HP dp6. Therefore, it is shown once more that dp6 GAGs bind

stronger to APRIL protein than dp4 GAGs. Obtained docking poses are similar to those docked with Autodock3 (Figure 8). However, MM/GBSA applied to DMD docking poses yields lower binding energy values than the ones obtained from MD simulations, starting from the binding poses produced by Autodock3. This points out the importance of taking into account the flexibility of the receptor for such calculations. The potentially unstructured nature of A_{105} APRIL N-termini corroborates the fact that up to now no structural data on the N-terminus of the protein were obtained by X-ray/NMR/CryoEM experiments. In our simulations, we observed high mobility of the N-termini of the A_{105} APRIL trimer both at coarse-grained and in all-atom levels.

Multiple GAG binding to APRIL protein

DMD method was adopted to check whether binding of multiple GAG fragments of different types is possible. To our knowledge, this is a novel approach, and no one has docked multiple GAG fragments to any protein yet. We used DMD to dock CSe dp6 to the full-length APRIL with already harbored HP dp6 (Supplementary Figure S7). The free binding energy analysis shows much weaker CSe binding to APRIL in the presence of HP in comparison to its absence. The mean value of 100 analyzed runs with MM/GBSA approach was

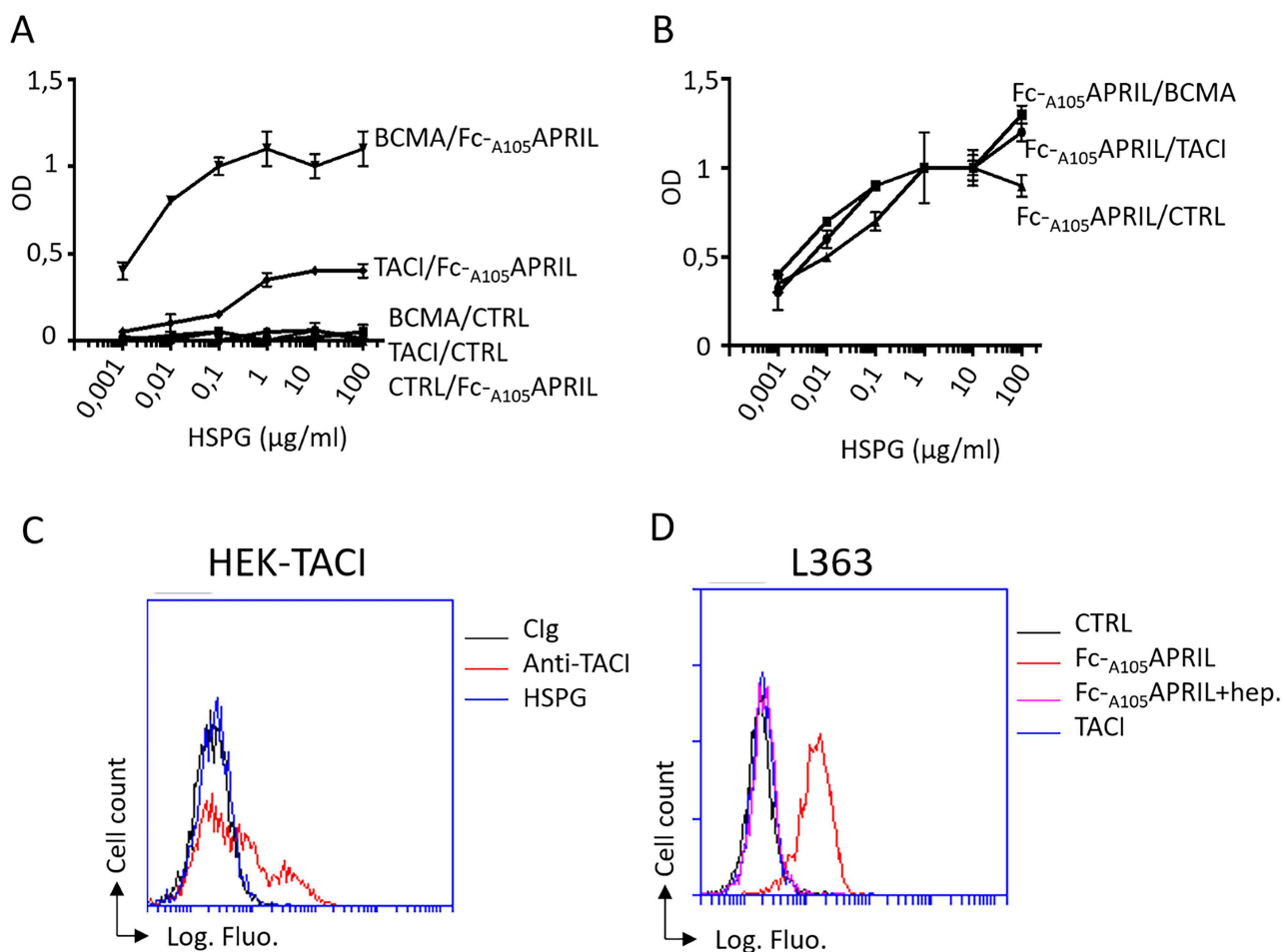


Fig. 7. No evidence for a direct interaction of TACI with HSPG. The binding of increasing concentrations of biotinylated HSPG to TACI was tested in ELISA. **(A)** APRIL receptors TACI and BCMA were coated, and the interaction was tested in the presence or absence of 1 μg/mL of soluble Fc-A105APRIL. **(B)** Fc-A105APRIL was coated and the binding was tested in the presence or absence of 1 μg/mL of the soluble receptors, TACI and BCMA. Results are presented as in Figure 1A. Three independent experiments were performed. **(C)** Binding of biotinylated HSPG ranging from 1 ng/mL to 10 μg/mL was assessed on HEK cells transiently transfected with full length human TACI. Results for only the highest HSPG concentration is shown. Reactivity with an anti-TACI is also shown (Clg = isotype-matched control immunoglobulin). **(D)** Binding of TACI-Fc (from 1 ng/mL to 30 μg/mL) was tested on L363 cells. Results from only the highest concentration of TACI are shown. HSPG expression on L363 cells is shown by the binding of Fc-A105APRIL (1 μg/mL) inhibited by heparin (hep.). Overlaid histograms plots are representative of at least two independent experiments. Thy-1-FC was used as control (CTRL) Fc-fused molecule.

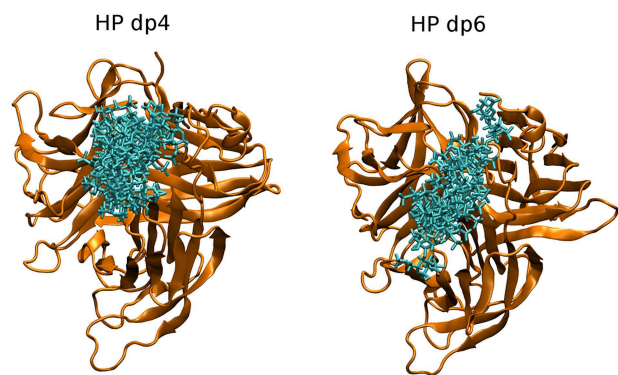


Fig. 8. Cluster representing the best docked structures for the HP dp4 and dp6 (in cyan sticks) after DMD to A105APRIL variant (in orange cartoon).

–27.2 kcal/mol compared to –90 kcal/mol found in the regular MD run analysis with MM/GBSA. LIE calculations also showed drop in the binding strength represented by –29.7 kcal/mol compared to

–62.8 kcal/mol to what we saw in the absence of the prebound HP. When visually analyzing the runs, we observed the dissociation of several CSe monomeric units in two of the runs (out of 100). In those two cases, the MM/GBSA binding analysis yielded values of –5.4 kcal/mol and –5.6 kcal/mol, while the LIE approach claimed –7.5 kcal/mol and –9.2 kcal/mol, respectively. Therefore, it is clear that CSe dp6 binding is much weaker in the presence of HP. When such a multiple GAG binding may not be favorable, it is still feasible. We performed energy decomposition to further analyze those relations and to check for the residues that hinder the interaction of CSe dp6 and the preformed complex with HP the most. It was not surprising that among the first six residues that repelled CSe the most belonged to HP. The other residues that negatively affected CSe–complex binding were the same that were shown in the absence of HP: ASP 159, GLU 185, GLU 191, ASP 223 and LEU 250. Interaction-favoring residues were found to be similar to those from the earlier analysis without HP. However, we found two new additional amino acids that are essential for the strengthening of the interactions, VAL 106 and LEU 107, while LYS 112, ARG 146 and LYS 249 have been shown to be still important.

Table II. MM/GBSA and LIE binding free energy analysis for the truncated APRIL variant in complex with HP dp4 and CSc dp4

	HP dp4, ΔG (kcal/mol)		CSc dp4, ΔG (kcal/mol)	
	MM/GBSA	LIE	MM/GBSA	LIE
First ns	-66.7 ± 12.6	-45.8 ± 5.6	-43.9 ± 15.5	-36.9 ± 13.6
Last ns	-57.1 ± 12.5	-37.5 ± 5.7	-31.4 ± 10.1	-28.5 ± 9.5
First 10 ns	-58.4 ± 13.1	-41.0 ± 7.2	-36.4 ± 12.1	-31.5 ± 6.3
Last 10 ns	-56.6 ± 13.4	-36.9 ± 8.1	-32.7 ± 8.8	-28.4 ± 6.9
First 50 ns	-56.5 ± 9.4	-38.5 ± 6.2	-33.8 ± 9.4	-29.7 ± 5.9
Last 50 ns	-56.4 ± 9.7	-37.6 ± 6.6	-33.3 ± 8.3	-28.8 ± 6.9
Whole simulation	-56.4 ± 8.3	-38.1 ± 5.8	-33.7 ± 8.1	-29.3 ± 5.9

Methodological aspects of binding free energy analysis by MM/GBSA and LIE

For the verification of our results and for the evaluation of the methodology used in our study, we performed extensive investigation regarding the statistical relevance of the data from free energy calculations. Beside such a verification, this analysis was aimed to understand our particular molecular system more comprehensively but also to gain insights into the nature of protein–GAGs interactions in general.

Binding free energy calculations is one of the most important and still challenging parts of computational analysis of protein–ligand interactions (Kollman et al. 2000; Genheden and Ryde 2015). There are several ways to assess binding free energy. The most commonly used ones are molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) and MM/GBSA, where the evaluation of the free energy is described as a sum of in vacuo molecular mechanics energy terms and a solvation free energy term in implicit solvent (Kollman et al. 2000). The obtained MM/PBSA (MM/GBSA) binding energies should be rather understood as enthalpies with the entropy of the solvent implicitly accounted for than the full free binding energies (Genheden and Ryde 2015). Another approach for free energy calculations is LIE, which is in general less accurate but computationally less expensive and is based on VdW and electrostatic energies linear combination obtained directly from each frame of the MD trajectory. In LIE, electrostatics are calculated in vacuo and are simply scaled by a dielectric constant. However, it is hard to evaluate how precise and reliable are the mentioned methods when applied on molecular docking simulations output, in particular, in case of the highly charged systems as ours. In this study, we used both methods in more high-throughput manner to investigate our model comprehensively.

For this methodological part of the study, we have chosen H115APRIL—a truncated version of the protein—as it lacks the first 10 amino acid residues that seem to be unstructured. Those residues could negatively affect our methodological analysis and contribute to the noise in our data, which in such a high-throughput and technique-oriented analysis, we would avoid at all cost. 50 top-scored poses obtained for each of CSc4 and HP dp4 GAGs docked to the truncated variant of APRIL protein with Autodock3 program were analyzed in terms of the binding energy obtained from 100 ns MD simulations by both the MM/GBSA and LIE approaches. First, the differences between the first and the last 1, 10 and 50 ns were compared for MM/GBSA (Table II, Figure 9). For most of the CSc4 dp4 binding poses, energies were higher at the end of the MD runs than at the beginning (Table II). This is probably due to the fact that molecular docking predicts interactions between ligand and protein without taking into account the solvent explicitly which is supposed

to be very important for the protein–GAG interactions (Teyra et al. 2006; Samsonov et al. 2011; Samsonov et al. 2014).

In the course of MD simulations, the interface between the receptor and the ligand is filled with the molecules of water, which in general, weakens the interactions in the complex. This result may look counterintuitive since it is rather expected that the MD simulation would correct the initial structure and, therefore, decrease the total energy of binding. However, the solvent impact is not considered explicitly in the free energy calculations, while the penetration of the water molecules into the protein–GAG interface leads to the generally less favorable MM/GBSA energies of binding in comparison to the case when no water molecules are in the complex interface. In case of CSc dp4, the density of probability maxima of the binding energy spectrum at the last part of the runs were sharper, which suggests tending toward the same energy values in the course of the MD run due to the convergence of the various MD simulations starting from different initial conformations (Supplementary Figure S8). It shows that the longer the run, the more reliable results are, which is, in general, expected. However, it is worth mentioning that extending the length of the MD runs over a certain point is highly cost-inefficient and yields from insignificant to no improvements. Our data clearly show how the uncertainties of the simulation result in such an electrostatically driven system change with the elongation of the simulation. In HP dp4 spectra, there were no differences observed in terms of the width of the peaks (Figure 9). Similarly to CSc dp4, for the HP dp4, there was an increase in binding energy during simulation. Almost no differences in the case of binding energy comparison after the first and second half of the runs for both CSc dp4 and HP dp4 were observed (Figure 9, Table II and Supplementary Figure S8). Moreover, no significant differences were found when the binding energies were compared for the first and the last 10 ns of the run for both CSc and HP. On the other hand, differences were significant and rather high when the first and the last ns of the run were compared. This suggests that the length of the simulation of 10 ns could be enough for converging results from the 50 starting poses obtained by docking for this system. It was also shown that the binding energies of the 50 docking poses differed essentially (Table II and Figure 10). Fortunately, our clustering protocol and the parameter choice for DBSCAN algorithm yielded poses spread normally across the range of energy binding values (Figure 10). This supports the idea that the obtained clusters of structures from 50 AD3 top-scored binding poses represent a structural ensemble properly.

EGB and ESURF are the components of the MM/GBSA free binding energies describing the impact of the solvent in the molecular interactions. We analyzed their effect on the final outcome of the free energy calculations. When comparing the first and the second

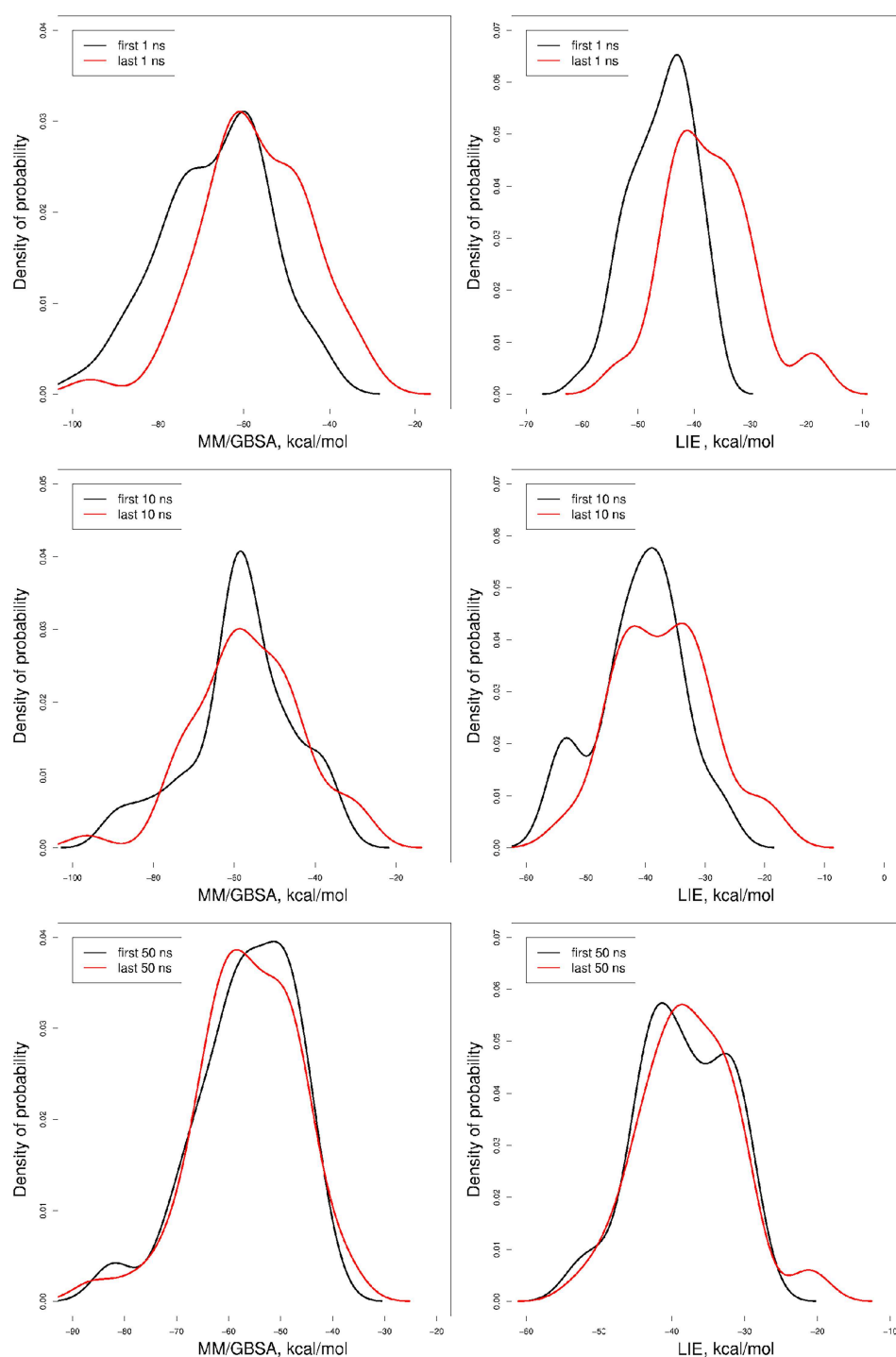


Fig. 9. MM/GBSA and LIE binding free energy analysis of HP dp4 to H115APRIL variant showing differences in total free energies between the first and the last ns (top), the first and the last 10 ns (middle), and the first and the second halves of the MD run (bottom).

halves of the runs—for both CSc dp4 and HP dp4—we detected almost no difference in the shapes of the EGB values distribution (Supplementary Figures S9 and S10). The same applied to the ESURF density of the probability curves (Supplementary Figures S9 and S10). There was a slight curve shift toward the lower values in the last 10 ns when compared to the first 10 ns for CSc dp4 and HP dp4 (Supplementary Figures S9 and S10). For the ESURF component, the

curves were slightly moved toward higher values—again for both CSc dp4 and HP dp4. Looking at the first and the last ns of the runs, trends remained the same in terms of the difference of the curve's shift. There was a much higher curves' shift for the first and the last ns of the runs for CSc dp4 and HP dp4 (Supplementary Figures S9 and S10). Those findings are in agreement with the results for other MM/GBSA components, showing that 1 ns of MD run is not sufficient to yield

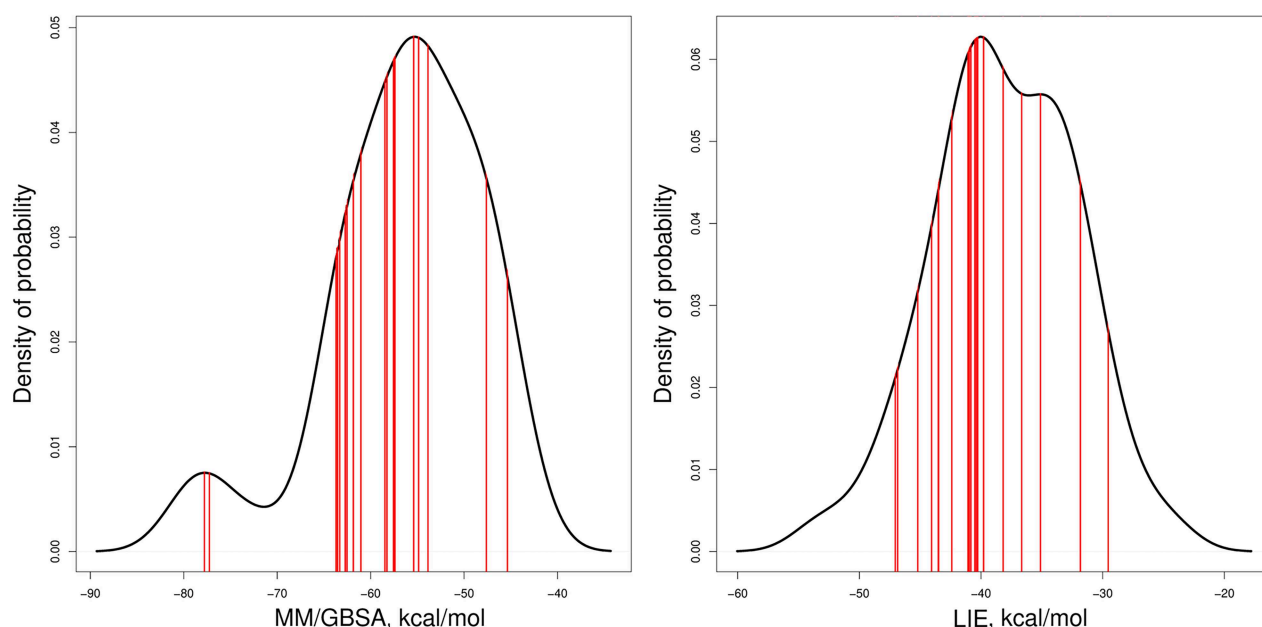


Fig. 10. Total binding free energy distribution analysis for the complex between H_{115} APRIL variant and HP dp4 from 50 independent MD runs. Red vertical lines correspond to binding free energies from particular MD simulations representing the structures belonging to the obtained clusters.

reliable results in terms of free energy of binding by the MM/GBSA approach for this particular molecular system.

LIE analysis yielded qualitatively similar results (Figure 9, Supplementary Figure S8 and Table II). Despite in general lower values of the obtained binding energies, they followed the same trend as in the case of the MM/GBSA analysis. Similarly, LIE showed no significant differences between the first and the last 10 or 50 ns of the MD simulation. LIE method resulted in narrower energy distribution across different binding poses, however, the results themselves look less distinctive and show smaller differences between certain types of GAGs (Supplementary Figures S9 and S10). These observations suggest that MM/GBSA could be considered as a more appropriate method to calculate the binding free energies in protein–GAG systems, which is similar to the data obtained for other biomolecular complexes (Tsui and Case 2000; Genheden and Ryde 2011).

Similarly to the MM/GBSA free energy components, we analyzed both ELE and VdW components of LIE in the course of the MD simulations (Supplementary Figures S9 and S10). For both components, when compared, the first and the second halves of the MD simulations curves looked almost identical in the case of both analyzed GAGs. In comparison of the first and the last 10 ns of MD runs, we observed only small shifts in the distribution curves' shapes. During the run, both VdW and ELE components shifted toward higher values with the exception that the ELE component moved only slightly toward higher values. This applied to both CSc dp4 and HP dp4. Likewise, in the MM/GBSA analysis, the differences were more distinctive when looking at the first and the last ns of the MD simulations. The highest shift was observed in the VdW component for both CSc dp4 and HP dp4.

The conclusion drawn from these free energy calculations is that 10 ns simulation in case of APRIL–GAG or similar systems in terms of size and interaction patterns is sufficient for the analysis of MD trajectories, especially when working on the bigger data sets of 50 repeated MD runs, like it was done in this study. Working only on around 10 MD runs with the length of 10 ns may not always be

enough to get a full insight of the protein–ligand interactions. The gains between 10 ns and 50 ns analyses were little, and thus, this would not be computationally effective to elongate the simulation up to 50 ns. Therefore, we suggest that it is better to enlarge the number of MD runs rather than elongate the runs. Similar statements were previously reported in other study (Genheden and Ryde 2015). The improvements were even more negligible after 50th ns of the MD runs when considering free energy analysis, and further elongation of the simulation is computationally expensive but would not improve the accuracy of the results. It is worth to notice that those insights cannot be extrapolated for all types of molecular systems. Minimal sufficient length of the MD simulation may depend on many variables, such as the system's charge, size of the ligand, flexibility of both ligand and receptor, geometry of the receptor-binding region, dominant interactions' types in the site of binding and most importantly on the goal of the study itself.

Conclusions

We present novel data which indicate that GAGs may also bind to the truncated version of the APRIL, but this binding is definitely less potent than the one to the full-length APRIL. In this study, it was shown that heparin binds the strongest to both variants of APRIL. This binding was confirmed both by computational and wet lab (ELISA/flow cytometry) experiments. CS-4,6-sulfate manifest weaker binding than heparin, while chondroitin-6-sulfate shows very weak-to-no binding properties. All of the mentioned GAGs when bound were located near the N-termini of the monomers that are spatially close in the APRIL trimer and form one common binding site. This region was analyzed with PBSA method and positive electrostatic potential was revealed near the N- and C-termini. Not surprisingly, in energy decomposition analysis, we also found most of the residues that significantly affected binding of the GAGs in this region. However, our study on heparin binding to peptide consisting of the first 10 amino acid residues from

the N-terminus of the APRIL protein indicates that GAG binding to peptide itself is not as strong as the binding to the full protein. The latter represents another argument in favor of a contribution of other peripheral and C-terminus residues in GAG binding. Structurally, the peptide lacked essential propensities for secondary structure features, and in this term, was similar to two out of three N-termini from the model of A₁₀₅APRIL. In this study, we further assessed the multiple GAG binding to single APRIL. Our findings show that it could be possible, but the strength of the binding is much weaker than in case for a single GAG molecule.

The biological assays performed in the present study showed no binding of HSPG to TACI. This is in agreement with the recent findings of Kowalczyk-Quintas et al. (2019) who claimed no binding of TACI to HP, which is the most sulfated compound in the HS family, unless oligomerized. Our computational analysis yielded borderline energy values for TACI–HP binding that suggests no binding or very weak one. Energies obtained for the BCMA–HP complex were not favorable enough and thus indicate no binding in this case.

To analyze our data, we used two methods of free binding energy assessment: MM/GBSA and LIE. While the MM/GBSA yielded more statistically reliable results, the LIE method is much faster and requires very little computational resources. It is also worth mentioning that for LIE to be used effectively, it is best that it would be optimized in advance using experimental data rather than default parameters. We also applied DMD and confirmed our results from docking with Autodock3. This novel GAG-specific docking approach allows for the inclusion of explicit solvent in local molecular docking and for molecular flexibility during docking process. To sum up, we rigorously evaluated the currently used MD-based methodology for protein–GAG complexes' theoretical description. We believe that our APRIL–GAG interactions analysis will be useful for further studies on the APRIL-related molecular mechanisms and the insights we shared in this work will help to facilitate drug development for autoimmune diseases and B-cell malignancies.

Supplementary data

Supplementary data for this article are available online at <http://glycob.oxfordjournals.org/>.

Acknowledgements

We would like to thank Pascal Schneider (University of Lausanne) for providing APRIL constructs. Computational resources were provided by the Polish Grid Infrastructure (PL-GRID, grants gagstr1, gpuaprilgags, aprilgagsinteractions, plgaprilgag2), ZIH at TU Dresden (grant p_gag) and our cluster at the Faculty of Chemistry, University of Gdansk.

Funding

National Science Centre of Poland (UMO-2018/30/E/ST4/00037).

Conflict of interest statement

None declared.

Abbreviations

APRIL, A proliferation-inducing ligand; BCMA, B-cell maturation antigen; CS, Chondroitin sulfate; DMD, Dynamic molecular docking; ELE, Electrostatic component of MM/GBSA, MM/PBSA, LIE free

binding energies; ELISA, Enzyme-linked immunosorbent assay; FGF, Fibroblast growth factor; GAG, Glycosaminoglycan; IgG, Immunoglobulin G; IgG2a, Immunoglobulin G2a; HP, Heparin; HS, Heparan sulfate; HSPG, Heparan sulfate proteoglycan; LIE, Linear interaction energy; MD, Molecular dynamics; MM/GBSA, Molecular mechanics generalized Born surface area; MM/PBSA, Molecular mechanics Poisson–Boltzmann surface area; PBSA, Poisson–Boltzmann Surface Area; TACI, Transmembrane activator and CAML interactor; TGF- β 1, Transforming growth factor beta one; TNF, Tumor necrosis factor; VdW, Van der Waals; VEGF, Vascular endothelial growth factor; WHAM, Weighted histogram analysis method.

References

- Adcock SA, McCammon JA. 2006. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem Rev.* 106:1589–1615.
- Ahmed MC, Huard B. 2021. Inhibition of chondroitin sulfate proteoglycans by APRIL. In: *Methods in Molecular Biology*. Springer, Berlin. Vol. 2248. p. 43–61.
- Babik S, Samsonov SA, Pisabarro MT. 2017. Computational drill down on FGF1-heparin interactions through methodological evaluation. *Glycoconj J.* Springer, Berlin. 34:427–440.
- Baert L, Manfroi B, Casez O, Sturm N, Huard B. 2018. The role of APRIL—A proliferation inducing ligand—in autoimmune diseases and expectations from its targeting. *J Autoimmun.* 95:179–190.
- Baert L, Benkhoucha M, Popa N, Ahmed MC, Manfroi B, Boutonnat J, Sturm N, Raguenez G, Tessier M, Casez O et al. 2019. A proliferation-inducing ligand-mediated anti-inflammatory response of astrocytes in multiple sclerosis. *Ann Neurol.* 85:406–420.
- Bischof D, Elsaawa SF, Mantchev G, Yoon J, Michels GE, Nilson A, Sutor SL, Platt JL, Ansell SM, von Bulow et al. 2006. Selective activation of TACI by syndecan-2. *Blood.* 107:3235–3242.
- Bojarski KK, Sieradzian AK, Samsonov SA. 2019. Molecular dynamics insights into protein-glycosaminoglycan systems from microsecond-scale simulations. *Biopolymers.* 110:e23252.
- Bossen C, Cachero TG, Tardivel A, Ingold K, Willen L, Dobles M, Scott ML, Maquelin A, Belnoue E, Siegrist CA et al. 2008. TACI, unlike BAFF-R, is solely activated by oligomeric BAFF and APRIL to support survival of activated B cells and plasmablasts. *Blood.* 111:1004–1012.
- Crijns H, Vanheule V, Proost P. 2020. Targeting chemokine—Glycosaminoglycan interactions to inhibit inflammation. *Front Immunol.* 11:483.
- Czaplewski C, Kalinowski S, Liwo A, Scheraga HA. 2009. Application of multiplexed replica exchange molecular dynamics to the unres force field: Tests with α and $\alpha+\beta$ proteins. *J Chem Theory Comput.* 5:627–640.
- Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham IIITE, Cruzeiro VDW, Darden TA, Duke RE, Ghoreishi D, Gilson MK et al. 2018. *AMBER16*. San Francisco (CA): University of California.
- Derler R, Gesslbauer B, Weber C, Strutzmann E, Miller I, Kungl A. 2017. Glycosaminoglycan-mediated downstream signaling of CXCL8 binding to endothelial cells. *Int J Mol Sci.* 18:2605.
- Digabriele AD, Lax I, Chen DI, Svahn CM, Jaye M, Schlessinger J, Hendrickson WA. 1998. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature.* 393:812–817.
- Ester M, Kriegl HP, Sander J, Xu X. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *AAAI.* 226–231.
- Faham S, Hileman RE, Fromm JR, Linhardt RJ, Rees DC. 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science.* 271:1116–1120.
- Gandhi NS, Mancera RL. 2011. Molecular dynamics simulations of CXCL-8 and its interactions with a receptor peptide, heparin fragments, and sulfated linked cyclitols. *J Chem Inf Model.* 51:335–358.
- Gehrcke JP, Pisabarro MT. 2015. Identification and characterization of a glycosaminoglycan binding site on interleukin-10 via molecular simulation methods. *J Mol Graph Model.* 62:97–104.

- Genheden S, Ryde U. 2011. A comparison of different initialization protocols to obtain statistically independent molecular dynamics simulations. *J Comput Chem.* 32:187–195.
- Genheden S, Ryde U. 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discovery.* 10:449–461.
- Habuchi H, Habuchi O, Kimata K. 2004. Sulfation pattern in glycosaminoglycan: Does it have a code? *Glycoconj J.* 21:47–52.
- Hahne M, Kataoka T, Schröter M, Hofmann K, Irmeler M, Bodmer JL, Schneider P, Bornand T, Holler N, French LE *et al.* 1998. APRIL, a new ligand of the tumor necrosis factor family, stimulates tumor cell growth. *J Exp Med.* 188:1185–1189.
- Hendriks J, Planelles L, de Jong-Odding J, Hardenberg G, Pals ST, Hahne M, Spaargaren M, Medema JP. 2005 Heparan sulfate proteoglycan binding promotes APRIL-induced tumor cell proliferation. *Cell Death Differ.* 12:637–648.
- Huige CJM, Altona C. 1995. Force field parameters for sulfates and sulfamates based on ab initio calculations: Extensions of AMBER and CHARMM fields. *J Comput Chem.* 16:56–79.
- Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. *J Mol Graph.* 14:33–38.
- Hymowitz SG, Patel DR, Wallweber HJA, Runyon S, Yan M, Yin JP, Shriver SK, Gordon NC, Pan B, Skelton NJ *et al.* 2005. Structures of APRIL-receptor complexes: Like BCMA, TACI employs only a single cysteine-rich domain for high affinity ligand binding. *J Biol Chem.* 280:7218–7227.
- Imberty A, Lortat-Jacob H, Pérez S. 2007. Structural view of glycosaminoglycan–protein interactions. *Carbohydr Res.* 342:430–439.
- Ingold K, Zumsteg A, Tardivel A, Huard B, Steiner QG, Cachero TG, Qiang F, Gorelik L, Kalled SL, Acha-Orbea H *et al.* 2005. Identification of proteoglycans as the APRIL-specific binding partners. *J Exp Med.* 201:1375–1383.
- Jarzynski C. 1997. Nonequilibrium equality for free energy differences. *Phys Rev Lett.* 78:2690.
- Joseph PRB, Mosier PD, Desai UR, Rajarathnam K. 2015. Solution NMR characterization of chemokine CXCL8/IL-8 monomer and dimer binding to glycosaminoglycans: Structural plasticity mediates differential binding interactions. *Biochem J.* 47:121–133.
- Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
- Kimberley FC, Van Bostelen L, Cameron K, Hardenberg G, Marquart JA, Hahne M, Medema JP. 2009. The proteoglycan (heparan sulfate proteoglycan) binding domain of APRIL serves as a platform for ligand multimerization and cross-linking. *FASEB J.* 23:1584–1595.
- Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: A generalizable biomolecular force field. *Carbohydrates.* *J Comput Chem.* 29:622–655.
- Koehler L, Samsonov S, Rother S, Vogel S, Köhling S, Moeller S, Schnabelrauch M, Rademann J, Hempel U, Pisabarro TM *et al.* 2017. Sulfated hyaluronan derivatives modulate TGF- β 1:receptor complex formation: Possible consequences for TGF- β 1 signaling. *Sci Rep.* 7:1210.
- Kokenyesi R, Bernfield M. 1994. Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. *J Biol Chem.* 16:12304–12309.
- Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W *et al.* 2000. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc Chem Res.* 33:889–897.
- Kowalczyk-Quintas C, Willen D, Willen L, Golob M, Schuepbach-Mallepell S, Peter B, Eslami M, Vigolo M, Broly H, Samy E *et al.* 2019. No interactions between heparin and atacept, an antagonist of B cell survival cytokines. *Br J Pharmacol.* 176:4019–4033.
- Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. 1992. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem.* 13:1011–1021.
- Künze G, Köhling S, Vogel A, Rademann J, Huster D. 2016. Identification of the glycosaminoglycan binding site of interleukin-10 by NMR spectroscopy. *J Biol Chem.* 291:3100–3113.
- Künze G, Gehrcke JP, Pisabarro MT, Huster D. 2014. NMR characterization of the binding properties and conformation of glycosaminoglycans interacting with interleukin-10. *Glycobiology.* 24:1036–1049.
- Larsen CG, Anderson AO, Appella E, Oppenheim JJ, Matsushima K. 1989. The neutrophil-activating protein (NAP-1) is also chemotactic for T lymphocytes. *Science.* 243:1464–1466.
- Ling RF, Späth H. 1981. Cluster analysis algorithms for data reduction and classification of objects. *Dent Tech.* 23:417–418.
- Liwo A, Oldziej S, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA. 1997. A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J Comput Chem.* 18:849–873.
- López-Fraga M, Fernández R, Albar JP, Hahne M. 2001. Biologically active APRIL is secreted following intracellular processing in the Golgi apparatus by furin convertase. *EMBO Rep.* 2:945–951.
- Luster AD. 1998. Chemokines—Chemotactic cytokines that mediate inflammation. *N Engl J Med.* 338:436–445.
- Lyon M, Rushon G, Gallagher JT. 1997. The interaction of the transforming growth factor- β s with heparin/heparan sulfate is isoform-specific. *J Biol Chem.* 272:18000–18006.
- Mason IJ. 1994. The ins and outs of fibroblast growth factors. *Cell.* 4:547–552.
- Matthes T, McKee T, Dunand-Sauthier I, Manfroi B, Park S, Passweg J, Huard B. 2015. Myelopoiesis dysregulation associated to sustained APRIL production in multiple myeloma-infiltrated bone marrow. *Leukemia.* 29:1901–1908.
- McCaffrey TA, Falcone DJ, Du B. 1992. Transforming growth factor- β 1 is a heparin-binding protein: Identification of putative heparin-binding regions and isolation of heparins with varying affinity for TGF- β 1. *J Cell Physiol.* 152:430–440.
- Moreaux J, Sprynski AC, Dillon SR, Mahtouk K, Jourdan M, Ythier A, Moine P, Robert N, Jourdan E, Rossi JF *et al.* 2009. APRIL and TACI interact with syndecan-1 on the surface of multiple myeloma cells to form an essential survival loop. *Eur J Haematol.* 83:119–129.
- Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem.* 19:1639–1662.
- Moustakas A, Souchelnytskyi S, Heldin CH. 2001. Smad regulation in TGF- β signal transduction. *J Cell Sci.* 114:4359–4369.
- Nordsieck K, Baumann L, Hintze V, Pisabarro MT, Schnabelrauch M, Beck-Sickingler AG, Samsonov SA. 2018. The effect of interleukin-8 truncations on its interactions with glycosaminoglycans. *Biopolymers.* 109:e23103.
- Onufriev A, Case DA, Bashford D. 2002. Effective Born radii in the generalized Born approximation: The importance of being perfect. *J Comput Chem.* 23:1297–1304.
- Panitz N, Theisgen S, Samsonov SA, Gehrcke JP, Baumann L, Bellmann-Sickert K, Köhling S, Pisabarro RJ, Rademann J, Huster D *et al.* 2016. The structural investigation of glycosaminoglycan binding to CXCL12 displays distinct interaction sites. *Glycobiology.* 26:1209–1221.
- Park S, Schulten K. 2004. Calculating potentials of mean force from steered molecular dynamics simulations. *J Chem Phys.* 120:5946.
- Penk A, Baumann L, Huster D, Samsonov SA. 2019. NMR and molecular modeling reveal specificity of the interactions between CXCL14 and glycosaminoglycans. *Glycobiology.* 10:715–725.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF chimera—A visualization system for exploratory research and analysis. *J Comput Chem.* 25:1605–1612.
- Pichert A, Samsonov SA, Theisgen S, Thomas L, Baumann L, Schiller J, Beck-Sickingler AG, Huster D, Pisabarro MT. 2012. Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling. *Glycobiology.* 22:134–145.
- Pomin VH. 2016. Paradigms in the structural biology of the mitogenic ternary complex FGF:FGFR:heparin. *Biochimie.* 127:214–226.
- Potthoff J, Bojarski KK, Kohut G, Lipska AG, Liwo A, Kessler E, Ricard-Blum S, Samsonov SA. 2019. Analysis of procollagen C-proteinase enhancer-1/glycosaminoglycan binding sites and of the potential role of calcium ions in the interaction. *Int J Mol Sci.* 20:5021.

- Risau W. 1997. Mechanisms of angiogenesis. *Nature*. 386:671–674.
- Sakurai D, Hase H, Kanno Y, Kojima H, Okumura K, Kobata T. 2007. TACI regulates IgA production by APRIL in collaboration with HSPG. *Blood*. 109:2961–2967.
- Salbach-Hirsch J, Samsonov SA, Hintze V, Hofbauer C, Picke AK, Rauner M, Gehrcke JP, Moeller S, Schnabelrauch M, Scharnweber D *et al.* 2015. Structural and functional insights into sclerostin-glycosaminoglycan interactions in bone. *Biomaterials*. 67:335–345.
- Šali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 234:779–815.
- Samsonov SA, Gehrcke JP, Pisabarro MT. 2014. Flexibility and explicit solvent in molecular-dynamics-based docking of protein-glycosaminoglycan systems. *J Chem Inf Model*. 54:582–592.
- Samsonov SA, Pisabarro MT. 2016. Computational analysis of interactions in structurally available protein–glycosaminoglycan complexes. *Glycobiology*. 26:850–861.
- Samsonov SA, Teyra J, Pisabarro MT. 2011. Docking glycosaminoglycans to proteins: Analysis of solvent inclusion. *J Comput Aided Mol Des*. 25:477–489.
- Schuepbach-Malpepp S, Das D, Willen L, Vigolo M, Tardivel A, Lebon L, Kowalczyk-Quintas C, Nys J, Smulski C, Zheng TS *et al.* 2015. Stoichiometry of heteromeric BAFF and APRIL cytokines dictates their receptor binding and signaling properties. *J Biol Chem*. 290:16330–16342.
- Schwaller J, Schneider P, Mhawech-Fauceglia P, McKee T, Myit S, Matthes T, Tschopp J, Donze O, Le Gal FA, Huard B. 2007. Neutrophil-derived APRIL concentrated in tumor lesions by proteoglycans correlates with human B-cell lymphoma aggressiveness. *Blood*. 109:331–338.
- Van Der Smissen A, Samsonov SA, Hintze V, Scharnweber D, Moeller S, Schnabelrauch M, Pisabarro MT, Anderegg U. 2013. Artificial extracellular matrix composed of collagen i and highly sulfated hyaluronan interferes with TGF β 1 signaling and prevents TGF β 1-induced myofibroblast differentiation. *Acta Biomater*. 9:7775–7786.
- Sugita Y, Okamoto Y. 2000. Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. *Chem Phys Lett*. 329:261–270.
- Swee LK, Ingold-Salamin K, Tardivel A, Willen L, Gaide O, Favre M, Demotz S, Mikkola M, Schneider P. 2009. Biological activity of ectodysplasin A is conditioned by its collagen and heparan sulfate proteoglycan-binding domains. *J Biol Chem*. 284:27567–27576.
- Teyra J, Doms A, Schroeder M, Pisabarro MT. 2006. SCOWLP: A web-based database for detailed characterization and visualization of protein interfaces. *BMC Bioinform*. 7:104.
- Tsui V, Case DA. 2000. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*. 56:275–291.
- Uciechowska-Kaczmarzyk U, Babik S, Zsila F, Bojarski KK, Beke-Somfai T, Samsonov SA. 2018. Molecular dynamics-based model of VEGF-A and its heparin interactions. *J Mol Graph Model*. 82:157–166.
- Uciechowska-Kaczmarzyk U, de Beauchene I, Samsonov SA. 2019. Docking software performance in protein-glycosaminoglycan systems. *J Mol Graph Model*. 90:42–50.
- Vallet SD, Miele AE, Uciechowska-Kaczmarzyk U, Liwo A, Duclos B, Samsonov SA, Ricard-Blum S. 2018. Insights into the structure and dynamics of lysyl oxidase propeptide, a flexible protein with numerous partners. *Sci Rep*. 8:11768.
- Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH *et al.* 2015. *Essentials of Glycobiology*. 3rd ed ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. p. 2015–2017.
- Wijdenes J, Vooijs WC, Clément C, Post J, Morard F, Vita N, Laurent P, Sun RX, Klein B, Dore JM. 1996. A plasmacyte selective monoclonal antibody (B-B4) recognizes syndecan-1. *Br J Haematol*. 94:318–323.
- Young MR, Pande VS. 2003. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophys J*. 84:775–786.