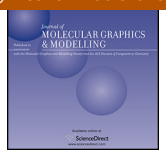
Journal of Molecular Graphics and Modelling 62 (2015) 97–104

Journal of Molecular Graphics and Modelling

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



Identification and characterization of a glycosaminoglycan binding site on interleukin-10 via molecular simulation methods

Jan-Philip Gehrcke, M. Teresa Pisabarro*

Structural Bioinformatics, BIOTEC, TU Dresden, Tatzberg 47-51, 01307 Dresden, Germany

ARTICLE INFO

Article history:
Received 25 June 2015
Received in revised form 13 August 2015
Accepted 1 September 2015
Available online 8 September 2015

Keywords:
Glycosaminoglycan
Interleukin-10
Molecular modeling
Molecular dynamics
GAG binding

ABSTRACT

The biological function of the pleiotropic cytokine interleukin-10 (IL-10), which has an essential role in inflammatory processes, is known to be affected by glycosaminoglycans (GAGs). GAGs are essential constituents of the extracellular matrix with an important role in modulating the biological function of many proteins. The molecular mechanisms governing the IL-10–GAG interaction, though, are unclear so far. In particular, detailed knowledge about GAG binding sites and recognition mode on IL-10 is lacking, despite of its imminent importance for understanding the functional consequences of IL-10–GAG interaction. In the present work, we report a GAG binding site on IL-10 identified by applying computational methods based on Coulomb potential calculations and specialized molecular dynamics simulations. The identified GAG binding site is constituted of several positively charged residues, which are conserved among species. Exhaustive conformational space sampling of a series of GAG ligands binding to IL-10 led to the observation of two GAG binding modes in the predicted binding site, and to the identification of IL-10 residues R104, R106, R107, and K119 as being most important for molecular GAG recognition. *In silico* mutation as well as single-residue energy decomposition and detailed analysis of hydrogen-bonding behavior led to the conclusion that R107 is most essential and assumes a unique role in IL-10–GAG interaction. This structural and dynamic characterization of GAG-binding to IL-10 represents an important step for further understanding the modulation of the biological function of IL-10.

1. Introduction

Interleukin-10 (IL-10) is an immunoregulatory cytokine whose biological relevance has been extensively reviewed [1]. Its most prominent function is to limit and eventually terminate inflammatory responses via its ability to inhibit effector functions of T cells, monocytes, and macrophages, and by inhibiting the synthesis of inflammatory cytokines such as IFN-γ and TNF-α [2–4]. Malfunction of IL-10 leads to improperly regulated immune reactions: IL-10-deficient mice spontaneously develop acute and chronic inflammation [5]. On the macroscopic scale, IL-10 has a crucial impact on tissue repair [6]. Despite the observed anti-inflammatory effects of IL-10, the attempts to use it directly as a therapeutic agent in various inflammatory conditions yielded disappointing results [7]. The IL-10 system turned out to be more complex than initially assumed, and it was found that its functions largely depend on its structural micro-environment [8], and the specific immune environment in which it is released [9]. Furthermore, it was found that IL-10 also has pro-inflammatory effects in certain conditions [10], pointing towards an incredibly multifaceted role of IL-10 in biology. Likewise, IL-10 is often called a *pleiotropic* cytokine.

IL-10 has been shown to bind glycosaminoglycans (GAGs) [11,12]. GAGs are essential building blocks of the extracellular matrix (ECM) and can be characterized as unbranched, negatively charged polysaccharides composed of repeating disaccharide units [13]. Based on the configuration of these disaccharide building blocks, GAGs are grouped into five classes: hyaluronan, chondroitin sulfate, dermatan sulfate, keratan sulfate, and heparan sulfate/heparin. GAGs play a critical role in many biological processes. Their multifarious biological activity arises from their ability to interact with and directly affect the biological activity of many cytokines, chemokines and growth factors [14–18]. Salek-Ardakani et al. demonstrated that GAGs may modulate IL-10 function by showing that soluble GAGs inhibit the IL-10-induced expression of CD16 as well as CD64 on monocytes and macrophages, revealing a dependency of the inhibition strength on the GAG sulfation degree. Interestingly, they also showed that sulfated cell surface-bound GAGs were required for IL-10 to trigger its biological function towards corresponding cells [11]. A recent NMR study showed that GAG sulfation is prerequisite for binding to IL-10: heparin was found to be the strongest binder, and the binding affinity of different GAG types decreased with the sulfation degree. No

* Corresponding author.
E-mail address: mayte@biotec.tu-dresden.de (M.T. Pisabarro).

http://dx.doi.org/10.1016/j.jmgm.2015.09.003

binding was detected for hyaluronan [12]. In essence, the molecular mechanisms of these effects are unclear so far. Thus, insights into the structural principles underlying the interaction between IL-10 and GAGs are required for understanding the biological role of IL-10–GAG interaction.

Here, we report the identification of a GAG binding site on IL-10 and the detailed characterization of the molecular interaction between IL-10 and GAGs in this site based on simulation methods. The calculation and analysis of the Coulomb potential of IL-10 in solution allowed us to locate a GAG binding region on its surface. We then employed Dynamic Molecular Docking (DMD), a recently published specialized method based on molecular dynamics (MD) simulations [20], to investigate the IL-10–GAG interaction in further detail. With DMD we performed exhaustive conformational space sampling while GAG and protein were treated as entirely flexible, and solvent was taken into account explicitly. The DMD calculations yielded two putative GAG binding pose models as well as ensemble-derived single-residue energetics and hydrogen bonding data. These data allowed the identification of those IL-10 residues being most important for binding GAGs. Our results shed light on possible molecular mechanisms governing GAG-mediated IL-10 function modulation.

2. Materials and methods

2.1. IL-10 and GAG structures

The analysis described in the present work was based on the X-ray structure of human IL-10 with PDB ID 2ILK (1.6 Å resolution) [20]. The biologically active unit of IL-10 is known to be a homodimer comprised of two intercalated IL-10 monomers and characterized by a twofold rotational symmetry [21,22]. We have built this homodimer by extending the monomeric 2ILK structure with a copy of itself, rotated by 180° around the crystallographic twofold axis as defined in the PDB entry.

The GAG molecules heparin (HP) dp4, HP dp6, hyaluronan (HA) dp4, HA dp6, chondroitin-4-sulfate (CS4) dp6, chondroitin-6-sulfate (CS6) dp6 ("dp" denoting the degree of polymerization; i.e. the number of sugar rings per molecule) were built with LEAP [23] and parameterized using GLYCAM 06 [24] as described previously [20]. MD snapshots of these molecules are depicted in SI Fig. S1.

2.2. IL-10 Coulomb potential simulation

The electrostatic potential of IL-10 was calculated with a finite-difference numerical solver applied to the linearized Poisson–Boltzmann (PB) equation, using the PBSA program shipped with AmberTools 13 [23]. In the PB model applied here, IL-10 is represented as a dielectric body with vacuum permittivity whose shape is defined by atomic coordinates and radii. The solvent is treated as a continuum with a relative permittivity of 80. The net electrostatic potential of IL-10 was calculated as the sum of its vacuum Coulomb potential and the corresponding solvent reaction field, using PBSA default parameters and a finite element grid spacing of 1 Å. IL-10 atomic coordinates were taken from the IL-10 dimer structure described above. The atomic radii and point charges of the protein were parameterized according to the FF99SB molecular mechanics force field [23]. Source code modification of the PBSA software was required for writing the discretized scalar data in compliance with the OpenDX data format and in units of kcal/(mol e) (whereas 1 kcal/(mol e) $\triangleq$4.18 kJ/(mol e) for appropriate post-processing in VMD [25]. Corresponding software patches were contributed back to the AmberTools project (see SI for further information about these modifications).

2.3. Dynamic Molecular Docking of GAGs to IL-10 and data analysis

The interaction between IL-10 and GAG molecules was investigated via Dynamic Molecular Docking (DMD), a recently published targeted molecular dynamics (tMD)-based docking method especially accounting for the effects of receptor flexibility and system solvation [20]. The DMD principle is schematically visualized in Fig. 1. Its cornerstone is the creation of an ensemble of MD trajectories by repetitively pulling a ligand molecule towards a predefined binding region on a receptor protein. Subsequent analysis of the trajectory data and the ensemble of ligand poses provides atomic information about the ligand–receptor interaction.

The geometrical DMD parameterization requires definition of a so-called *core atom*, an atom in the protein core with as little mobility relative to the bulk of the protein as possible, as well as a *focus point* near the protein surface at the center of the anticipated binding region. The straight line connecting *focus point* and *core atom* determines the directionality of the relative ligand–receptor movement when the ligand is being pulled towards the receptor. The distance between *core atom* and *focus point* corresponds to the so-called *target distance*. For a more complete definition of these terms see SI Fig. S2 and the previously published DMD methodology [20].

In the first stage of a so-called DMD run, the ligand molecule is slowly pulled towards the receptor molecule, starting from a distal position, as schematically visualized in Fig. 1. This pulling process is implemented with a time-dependent harmonic potential applied to the distance between the *core atom* and a central ligand atom. The pulling potential competes with the canonical molecular mechanics potentials and with the thermal fluctuations applied in the simulation. Hence, while being directed, the trajectory of relative movement between receptor and ligand remains random. When the *target distance* is reached, the pulling potential is switched off, and the second stage of the DMD run is performed, which is an unrestrained ("free") molecular dynamics simulation for relaxing the system and for collecting data. A DMD study consists of a large number of DMD run repetitions (see Section 3 for the

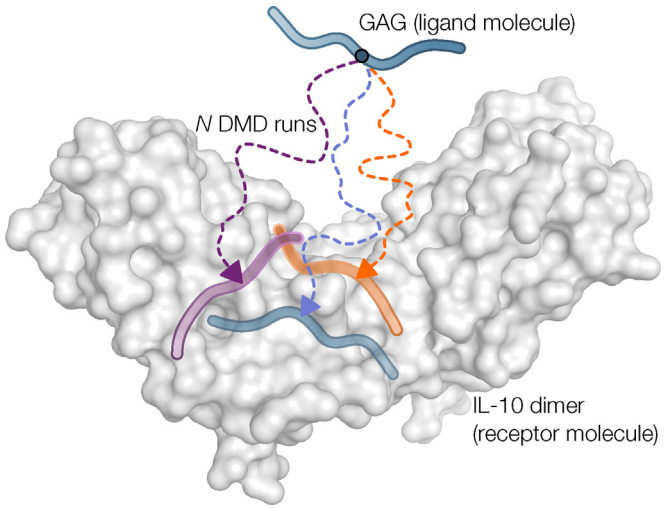


Fig. 1. Schematic representation of a DMD study. A DMD study is comprised of N DMD runs, performed in independent simulations. Each run begins with a pulling process: starting from a distal position, the GAG molecule is pulled towards IL-10 (shown in gray surface representation). During this process, the GAG translates and rotates along a random path and samples its conformational space. The pulling process stops when the GAG is in proximity to the surface of IL-10 and is followed by a long unrestrained MD simulation. This results in an ensemble of MD trajectories as well as in a collection of GAG poses in contact to the receptor surface, as schematically indicated here in purple, blue, orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

exact numbers applied in our studies), independently performed with different random seeds, yielding an ensemble of MD trajectories and ligand docking poses.

The geometrical DMD parameterization (i.e. the selection of *core atom* and *focus point* as introduced above) used for the work presented here was based on the observations made in the Coulomb potential analysis (cf. Section 3). Correspondingly, the DMD *focus point* was centered on the region comprised of residues R102, R104, R106, and R107, as visualized in SI Fig. S2. The C-alpha atom of IL-10 residue M154 was selected as the *core atom*. This choice was mainly driven by the goal to maximize the *target distance*, because a large *target distance* leads to a wide spatial distribution of ligand molecules (among the runs within a DMD study) around the *focus point*, and therefore allows for exhaustive Cartesian space sampling. *Focus point* and *core atom* selection as specified above led to a *target distance* of 20.5 Å, causing DMD to sample a significant fraction of the surface of IL-10 (see SI Fig. S2). This geometrical parameterization complies with the guidelines formulated in the original DMD method description [20].

Independent DMD studies were performed for each of the six different GAG molecules listed in Section 2.1. Molecular dynamics system parameterization, minimization, heat-up, equilibration and production protocols were chosen as described earlier [20]. The tMD simulation time and the duration of the free MD stage were chosen to last 3 ns and 10 ns, respectively. Summed over all DMD studies performed for the presented work, about 30 μs of raw MD trajectory data were created and analyzed. The simulations were performed on a cluster of graphics processing units using Amber's pmemd.cuda [26]. End-point free energy calculations were performed using the MMPBSA.py framework [27] as described earlier [20]. In summary, the MMPBSA method was used for obtaining a free energy of binding estimate ΔG for each docking pose, and the MMGBSA method was used for obtaining a single-residue energy decomposition (SRED) thereof.

Crucial component of the DMD method is the analysis of the ensemble of MD trajectories created in a single DMD study, yielding converged ensemble properties with statistical significance. Here, ensemble-derived single-residue energy decomposition was obtained as described previously [20]. Additionally, we implemented an approach for ensemble-derived hydrogen bonding analysis: for each free MD trajectory within a DMD study, the occupancy of single donor and acceptor atoms was tracked over time, enabling to filter those IL-10 residues being most involved in IL-10–GAG hydrogen bonding. Hydrogen bond detection was based on atom type, distance, and angle criteria as implemented in the MD trajectory analysis software CPPTRAJ [28].

The spatial clustering method applied here for identification of the most frequently occurring GAG binding poses was based on the DBSCAN algorithm [29]. It involved a custom similarity metric optimized for GAG molecules, and was wrapped within a custom clustering parameter optimization procedure for ensuring reproducibility and comparability, as outlined previously [20]. In particular, clustering was performed with the boundary condition that the minimal number of clusters to be found is 1, and that the minimal number of members within each cluster is 5.

3. Results

3.1. IL-10 Coulomb potential analysis

Based on the rational that Coulomb interaction plays a dominant role in protein–GAG interaction, we conducted a Coulomb potential analysis for examining the existence of putative GAG binding regions on IL-10. A numerical Coulomb potential simulation was performed for the IL-10 structure in solution, yielding the spatial distribution of its electrostatic potential in terms of discrete values on a three-dimensional grid. Based on this data, our goal was to identify the main characteristics of the Coulomb potential – not only on the molecular surface of IL-10, but within the entire protein-surrounding volume. To that end, topology and strength of the potential were analyzed with an isosurface visualization while varying the isovalue. Fig. 2 shows the Coulomb potential isosurface of IL-10 for four different positive isovalues Φ.

Considering the lowest isovalue shown (Fig. 2a), the volume surrounding IL-10 can be categorized into different regions: large parts of the molecular surface of IL-10 are not covered by the potential isosurface meaning that they do not exhibit significant electrostatic attraction for GAGs. One noticeable region remains in which the electrostatic potential isosurface is assembled symmetrically, following the twofold rotational symmetry of the IL-10 dimer. When increasing the isovalue (Fig. 2, panels b–d), the Coulomb potential isosurface protrudes less and less into space and splits into two regimes, narrowing down the spatial origins of the attraction of IL-10 for negatively charged molecules: Fig. 2d shows two distinct isosurface patches close to the molecular surface of IL-10. A closer look at the IL-10 structure reveals that the basic residues responsible for creating these two core regions are arginines 102, 104, 106, 107, as well as lysine 119, as depicted in Fig. 2d. Figure panel d shows a special case where the isovalue is chosen so that the Coulomb potential isosurface does almost not reach into space further than the molecular surface. In that case, the isovalue itself characterizes the system by providing an estimation about the strength of the electrostatic interaction between GAG and IL-10 in the bound state. For IL-10, this characteristic isovalue has a magnitude of about 2 kcal/(mol e). Taken literally, this value implicates a Coulomb energy per test charge that is about three times larger than the thermal energy at 300 K. Hence, the identified region (the blue part in Fig. 2d) can be assumed to possibly play a major role in IL-10–GAG interaction, and was treated as putative binding region for the following DMD analysis.

3.2. Dynamic Molecular Docking of GAGs to IL-10

The molecular interaction between GAGs and IL-10 within the identified binding region was characterized via DMD, with the goal to obtain important insights about the dynamics, energetics, and atomic details of this interaction. About 30 μs of MD trajectory data were created and analyzed in the course of performing DMD.

3.2.1. Hydrogen bonding analysis

One of the goals was to identify those IL-10 residues that are important for the establishment of hydrogen bonds to GAGs. To that end, ensemble-derived hydrogen bond information was extracted from a DMD study with IL-10 and HP dp6. For each DMD run in this study, the occupancy of single hydrogen bond donor atoms in IL-10 with GAG acceptor atoms was time-averaged over the last nanosecond of the free MD stage, yielding an average occupancy value per donor atom and DMD run. Subsequently, atoms were grouped by residues (for instance, an arginine can theoretically donate up to four hydrogens at the same time, and its maximum possible average occupancy therefore is 4). For each protein residue, the average occupancy was then normalized over all DMD runs ($N = 300$), providing the mean occupancy by which the residues were ranked thereafter. Fig. 3 shows the top four donating residues, including box plot representations of the occupancy distribution over single DMD runs.

R107 is ranked first by mean as well as median, with a measured mean occupancy of about 1.7. That is, throughout all DMD runs in the study this residue donated on average almost two hydrogen atoms to GAG acceptor atoms, simultaneously. This can be interpreted as a strong indicator for the relevance of R107 in IL-10–GAG

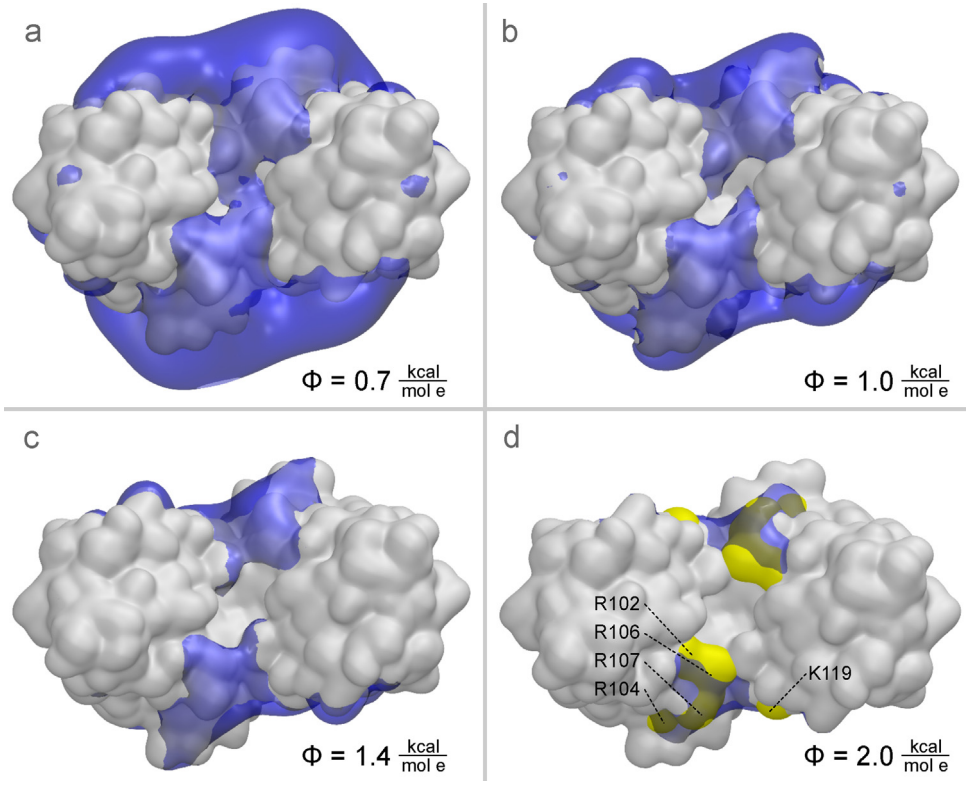


Fig. 2. Isosurface representation of the Coulomb potential of IL-10 (blue) shown for multiple isovalues Φ, attractive for negative probe charges. The molecular surface of the IL-10 dimer (PDB ID 2ILK) is shown in gray. In panel d, the molecular surface of arginines 102, 104, 106, 107 and lysine 119 is highlighted in yellow. There, Φ is chosen so that the isosurface does almost not reach into space further than the molecular surface of IL-10. This visualization therefore narrows down the spatial origins of the electrostatic attraction of IL-10 for negatively charged molecules. Hence, IL-10–GAG interaction most likely takes place within the symmetrically arranged regions indicated in blue in panel d. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interaction. However, the essential observation here is that the occupancy distributions underlying the box plot representation of K119 and R107 have a fundamentally different shape when compared to the occupancy distributions of R104 and R106, and all other relevant residues: the occupancy distributions of R104 and R106 have their maximum near zero and a strong decay towards higher occupancies, yielding a median close to zero. In contrast, in case of R107 and K119, the distributions have their global maximum at about 1.5 and 0.8, respectively, so that these distributions have a median far from zero. That is, the hydrogen bonding properties of K119 and R107 qualitatively stand out compared to the other amino acids in IL-10. The described qualitative difference between the occupancy distribution of R107 and the other amino acid residues can be observed in all IL-10–GAG DMD studies we have performed (data not shown). Hence, according to DMD and the resulting hydrogen bonding data, the special role of R107 is independent of the GAG type under investigation.

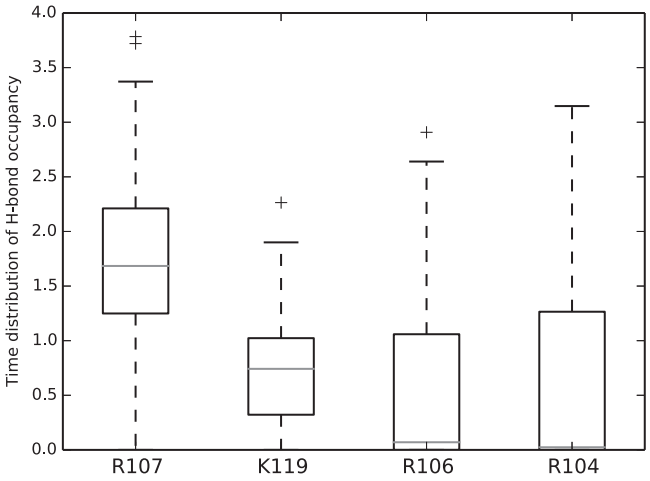


Fig. 3. IL-10 key residues for GAG-interaction, as identified by ensemble-derived hydrogen bonding information obtained from a DMD study with IL-10 and an HP dp6 ligand. The box plot representation is built from $N = 300$ samples per box, whereas each sample is the number of occupied hydrogen bond donor atoms averaged over the data production period of one free MD trajectory for the residue specified in the abscissa label. The residues are sorted by the *mean* (not shown), the gray bars indicate the *median* (close to zero if not visible).

3.2.2. Single-residue energy decomposition (SRED)

Ensemble-derived single-residue energy decomposition was used as a second method for determining those IL-10 residues that are major contributors to IL-10–GAG binding. This analysis was performed on the same DMD study as the preceding hydrogen bonding analysis was based on. In SRED, classical molecular mechanics energy terms as well as solvation energy terms are extracted from an MD trajectory and decomposed into the contribution of single residues to the overall receptor–ligand binding energy. Within each DMD run, this type of data was time-averaged. Subsequently, the time-averaged data were merged among all DMD runs, yielding ensemble statistics. Fig. 4 shows the four top-ranked residues that are, according to SRED, most relevant for the interaction of IL-10 with heparin.

Considering the statistical error, the residues on ranks two to four are not distinguishable, ranging in the interval between −6 kcal/mol and −8 kcal/mol. R107, however, assumes a special role, with a mean interaction energy of -13.5 ± 0.5 kcal/mol (note that the absolute values of these energies as well as the inter-residue energy differences are not evaluated here, we focus on looking at the relative energy differences, which can be assumed to

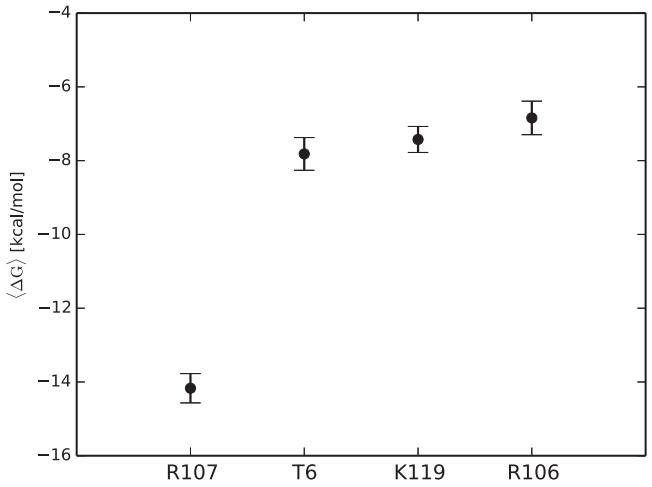


Fig. 4. Top four residues in IL-10 for binding heparin, according to ensemble-merged single-residue energy decomposition (SRED) data from a DMD study with IL-10 and an HP dp6 ligand. The error bars represent the standard error of the mean obtained from $N = 300$ samples.

yield resilient information). Hence, just like the hydrogen bonding analysis, SRED suggests a distinguished behavior of R107 compared to the other amino acid residues.

The ensemble-derived SRED analysis was performed for all GAG types listed in Section (see SI Fig. S3). In all of these analyses, R106, R104, as well as K119 are contained within the top five residues contributing to binding, and R107 ranks first.

3.2.3. Impact of R107A mutation

Three independent DMD studies with IL-10 and a HP dp4 ligand were performed for further characterizing the impact of arginine 107 on IL-10–GAG interaction. Two of the studies were set up equivalently ("setup A") except for an R107A mutation of the IL-10 sequence in one of the two studies. The third study was performed as a control, with the wildtype IL-10 sequence and with a slightly different geometrical DMD parameterization ("setup B") than used in the other two studies. Four different ensemble-averaged quantities were extracted from each of the three studies. They are presented in Table 1. Most notably, the extracted quantities are well-converged, with small values for their standard error of the mean. Despite the marginal differences between setup A and B, both DMD studies with the wildtype sequence yielded the same results, which supports the idea that DMD is a well-reproducible method and tolerant against small variations in the geometrical setup. This is a crucial observation, meaning that significant differences between the results of both setup A studies can be entirely attributed to the R107A mutation.

Compared to the wildtype system, R107A mutation led to a decrease of about 40 % in terms of the protein–GAG binding energy estimate ΔG obtained from MMPBSA analysis. The decomposition of ΔG into single energy term contributions revealed that said decrease stems almost entirely from a significant drop in the protein–GAG Coulomb interaction energy $\Delta E_{Coulomb}$. The phenomenological consequences of the described interaction energy decrease in the R107A system are (i) a significant enhancement of the mobility of the GAG molecule relative to the protein, and (ii) a significant reduction of the number of simultaneously established hydrogen bonds between protein and GAG.

SI Fig. S4 shows a comparison of the hydrogen bonding behavior of the mutated and wildtype IL-10. Interestingly, the box plot representations show similar hydrogen bonding behavior of R104, R106, and K119 in both setup-A-studies. Particularly, the box plot representation of the occupancy of R107 in the wildtype IL-10 finds no resemblance in the DMD study with the mutated protein. That is, no other amino acid residue was able to compensate the loss of the polar and hydrogen bonding properties of R107. The results presented in this section support the view in which R107 assumes a unique role for IL-10–GAG interaction.

3.2.4. Clustering of GAG poses

The final state of each DMD run defines a certain GAG conformation and position relative to IL-10 which is what we call "docking pose". Likewise, a DMD study yields an ensemble of docking poses. The spatial distribution of GAG structures within each docking pose ensemble was grouped into clusters based on spatial similarity, for identifying those GAG poses that occurred with highest probability and therefore are representative. Fig. 5 shows the most populated docking pose cluster for each of the six DMD studies performed for the presented work. The obtained clustering data is best interpreted in a coarse grained fashion with a focus on identifying common as well as divergent properties among the clusters obtained from the different studies: it can be observed that CS4 dp6, HA dp4, HA dp6 as well as HP dp6 adopt a common cluster placement and orientation that differs from the cluster poses presented by CS6 dp6 and HP dp4. This coarse grained interpretation results in a reduction of the raw clustering data towards two principal GAG binding orientations: these two poses, which we name A and B, are schematically represented in Fig. 6. In pose A, the GAG molecule is placed in a surface groove below residue R107. Pose B, a qualitatively different pose, is oriented from the top left to the bottom right, with the GAG placed above R107, and squeezed between R107 and K119. This analysis indicates the possibility of different GAG binding poses in the discussed binding region of IL-10. Notably, all obtained clusters heavily involve R107 in GAG binding.

4. Discussion

We have made use of innovative computational approaches which especially account for the dominance of long-range electrostatic interaction to investigate the molecular details of recognition between IL-10 and a series of GAG ligands. The major reasons for taking these computational approaches were to explore methodological advances for conducting *in silico* research on protein–GAG systems in general, and to reveal structural details about the so far structurally uncharacterized IL-10–GAG system. Certain

Table 1
Impact of R107A mutation on various ensemble properties of IL-10 interacting with a HP dp4 ligand, as obtained from three independently performed DMD studies. Setups A and B differ slightly in their geometrical DMD parameterizations. The two setup-A-studies differ only in terms of the R107A mutation. All properties shown here are time and ensemble averages ± the corresponding standard error of the mean, extracted from N DMD runs. ΔG is a binding energy estimate for the protein–GAG system, obtained via the MMPBSA method. N_{Hbond} is the number of simultaneously established hydrogen bonds between protein and GAG. m_{GAG} reflects the GAG mobility relative to the protein (the measured quantity is the standard deviation of a GAG–GAG root mean square distance time series, with protein–protein coordinate alignment). $\Delta E_{Coulomb}$ is the vacuum Coulomb energy of the protein–GAG interaction in the molecular dynamics simulations.

DMD study	$\langle\Delta G\rangle$(kcal/mol)	$\langle N_{Hbond}\rangle$	$\langle m_{GAG}\rangle$(Å)	$\langle\Delta E_{Coulomb}\rangle$(kcal/mol)
IL-10 / HP dp4 (setup B, $N = 200$)	−44.0 ± 1.1	5.8 ± 0.1	2.7 ± 0.1	−1175 ± 13
IL-10 / HP dp4 (setup A, $N = 100$)	−43.5 ± 1.5	5.6 ± 0.2	2.6 ± 0.1	−1130 ± 20
IL-10_{R107A} / HP dp4 (setup A, $N = 100$)	−27.0 ± 1.1	4.2 ± 0.2	3.4 ± 0.3	−726 ± 21

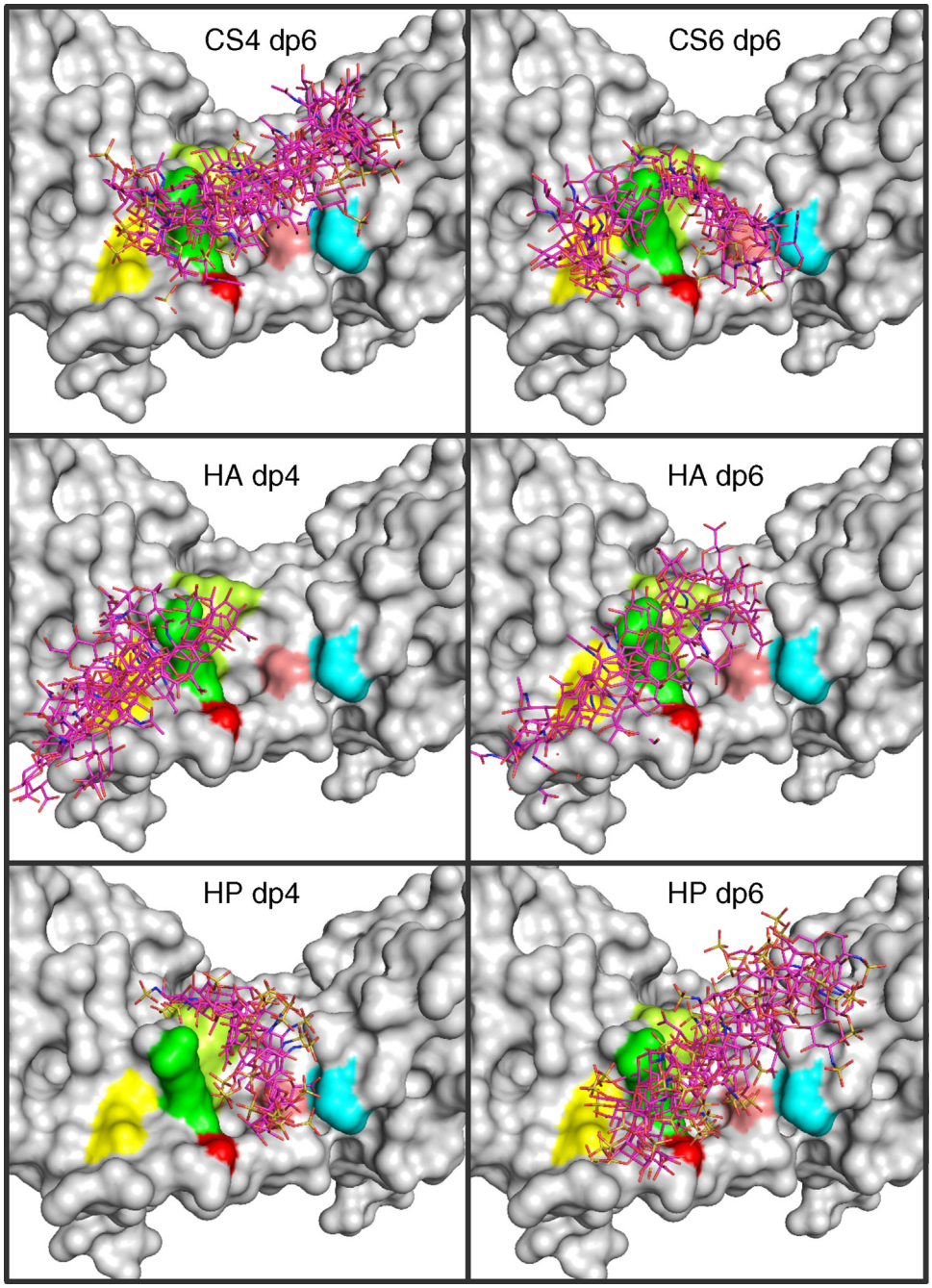


Fig. 5. Most populated docking pose clusters of six DMD studies differing in GAG type and length obtained with a reproducible DBSCAN-based clustering approach. The GAG structures are shown in stick representation with carbon atoms in magenta. The IL-10 structure (PDB ID 2ILK) is shown in gray surface representation. The surface of R107 is colored green, R106 is shown in light green, R104 in yellow, K119 in cyan, N10 in red, and T6 in light red.

experimental challenges exhibited by this macromolecular complex make the use of computational approaches an alternative – and in some cases complementary – choice to investigate the basis of its molecular recognition: most notably, their low binding affinity (which is in the mM to μM K_D range [12]) severely limits the applicability of traditional structure determination approaches such as X-ray crystallography and common NMR methods.

In previous studies, we have investigated exemplary protein–GAG systems with the same Coulomb potential analysis method as described here. We concluded that distinct regions of significant electrostatic attraction are observable on the protein surface at experimentally proven GAG binding sites [20]. Therefore, analysis of the strength and topology of the Coulomb potential surrounding a protein is a simple approach for making a well-founded theoretical prediction about those regions of the protein on which GAG ligands most likely bind. In case of IL-10, one such distinct region was identified (and as of the spatial symmetry of the IL-10 dimer, this region occurs twice). Hence, the distribution of the Coulomb potential of IL-10 in space provides strong evidence that the interaction of IL-10 with GAGs is most likely localized in the region comprised of residues R102, R104, R106, and R107. It is quite important to note that already in one of the very first structural descriptions of IL-10, Zdanov et al. characterized this region as being a special feature of IL-10, as of the dense occurrence of basic residues [30]. Interestingly, this region is conserved among species and unique to IL-10 as it is not observed

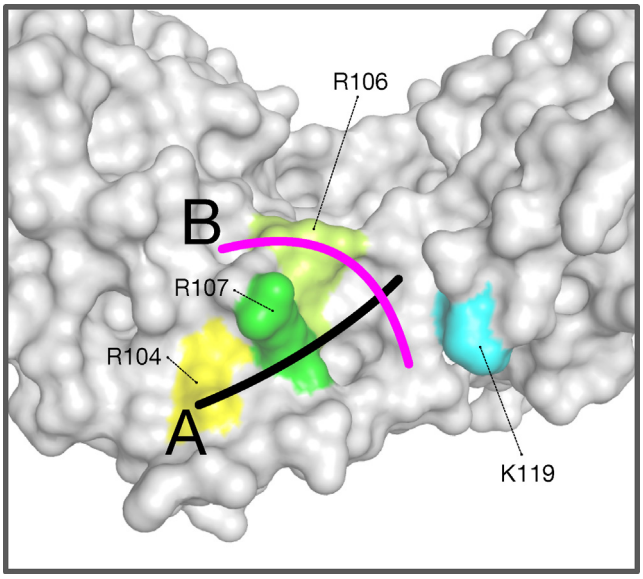


Fig. 6. Schematic visualization of the two principally different GAG binding modes observed via docking ensemble clustering after performing DMD studies with different GAG types. In pose A (black), the GAG is located in an IL-10 surface groove right below R107. In pose B (magenta), the GAG is placed above R107 and within the groove between R107 and K119. The surface of the four residues being most important for GAG binding, according to DMD ensemble-derived hydrogen bonding analysis, is colored and labeled. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for other members of the IL-10 family of cytokines [31,22]. With about 2.0 kcal/(mol *e*), the characteristic isovalue of IL-10 (providing an estimate for the strength of its electrostatic potential near its surface in the proposed binding region) was determined to be significantly larger than the thermal energy at room temperature. However, compared to exemplary protein–GAG systems analyzed with the same method, IL-10 appears to exhibit a weak potential (for instance, FGF2 from PDB entry 1BFB and SDF-1 from PDB entry 2NWG have characteristic isovalues of 6.0 kcal/(mol *e*) and 8.5 kcal/(mol *e*), respectively [20], also see SI Figures S5 and S6).

Exhaustive sampling of the dynamics of a series of IL-10–GAG systems via Dynamic Molecular Docking (DMD) has led to unambiguous data about the importance of one particular amino acid residue in the proposed binding region for IL-10–GAG interaction: R107. Ensemble-merged properties obtained by MMGBSA single-residue energy decomposition as well as hydrogen bonding analysis yielded well-converged numbers and strongly suggest that R107 behaves qualitatively different from and outweighs the remaining residues in the mentioned region. This observation is supported by all performed DMD studies. *In silico* analysis of the impact of R107A mutation corroborates this view by illustrating that R107 assumes a unique role in IL-10–GAG interaction, in the sense that no other residue is able to compensate the loss of R107.

Both of the two principal poses derived from the clustering results (Fig. 6) involve the groove between R107 and K119 and therefore match the fact that SRED and hydrogen bonding analysis identify K119 as being important for IL-10–GAG interaction: aforementioned groove could serve as a trap, potentially anchoring a GAG from two sides simultaneously, with a positively charged amino acid residue on both sides. That groove is just wide enough for a GAG molecule, as can be seen in the HP dp4 panel of Fig. 5.

An important experimental observation supporting the results presented here is that the crystallographic structure of IL-10 (PDB ID 2ILK) contains three sulfate molecules, two of which are crystallized in positions that overlap with the principal binding poses identified here (see SI Fig. S7): pose A overlaps with both sulfate locations, and pose B overlaps with one of them. Furthermore, previously published NMR experiments suggest that the IL-10–GAG binding stoichiometry changes for GAG molecules longer than hexasaccharides, supporting the view in which long GAG molecules tend to simultaneously bind to both IL-10 monomers in the IL-10 dimer, an effect known as positive cooperativity [12]. Our model indicates that a GAG molecule bound in the groove delimited by R107 and K119 has the appropriate orientation for being able – when elongated – to connect the twofold symmetry-related binding regions proposed here, which would fit aforementioned observations.

The binding region discussed in the present work is comprised of residues which are conserved among species, which by itself could be an indicator for a functional role of IL-10–GAG interaction. Salek-Ardakani et al. have shown in *in vitro* experiments that IL-10 function can in fact be modulated by the presence of heparin [11]. One conceivable scenario for this to happen is that GAGs could directly affect signal transmission by hindering the recognition of IL-10 by its receptors. The crystallographic structure of IL-10 in complex with the extracellular domain of its high affinity transmembrane receptor IL-10R1 [32] reveals that the GAG binding region we have predicted here does not geometrically overlap with the IL-10/IL-10R1 interface. The structure of the low affinity receptor IL-10R2 is known [33], but there is so far no experimentally confirmed structure of the ternary IL-10/IL-10R1/IL-10R2 complex available. Before 2012, various different models of the ternary IL-10 signaling complex were published, based on rational positioning along the lines of very limited experimental data, or based on molecular modeling techniques [32,34–36,22,33]. Unfortunately, these models are contradictory. Recently, the structure of the IL-20/IL-20R1/IL-20R2 complex has been published [37], which allowed us to obtain a simple homology-based model: alignment of IL-20R1 (from the ternary IL-20 complex) onto IL-10R1 (from the binary IL-10 complex) was followed by aligning IL-10R2 (resolved individually) onto IL-20R2, yielding a basic IL-10/IL-10R1/IL-10R2 binding model (SI Fig. S8) similar to the one published by Zdanov and Pletnev [22,35]. Based on this model, it can not be excluded that IL-10–GAG interaction in the predicted binding region interferes with IL-10/IL-10R2 interaction, particularly considering the case of long GAG polymers present in the extracellular matrix. Therefore, one possible scenario for the biological role of IL-10–GAG interaction which needs to be addressed in future studies is that GAGs could directly modulate IL-10 function by modulating the IL-10/IL-10R2 interaction. Other than that, it is often speculated that cytokine–GAG interaction in certain cases does not directly affect the signaling cascade, but rather may be a mechanism for cytokine concentration and diffusion control, for instance for retaining a type of cytokine close to its site of secretion in the tissue, which seems to be the case for IL-2 [38,39] and IL-5 [40]. The scenario in which longer GAG chains impair the diffusion of IL-10 without affecting its binding to the receptors is a valid model. In that case, interaction of a polymeric GAG chain with the predicted binding region could be an effective mechanism for hindering the diffusion of the IL-10 molecule. This model would very well fit the idea of impaired cytokine diffusion due to cell surface-attached proteoglycans. The identification of a GAG binding site as presented here serves as crucial starting point to further investigate both outlined scenarios.

5. Conclusions

In conclusion, application of innovative *in silico* methods yielded compelling evidence that binding of GAGs to IL-10 may occur in a well-defined region. Furthermore, the analysis presented here reveals an outstanding role of R107 for GAG binding, and *in silico* mutation of R107 to alanine suggests that its contribution to

binding cannot be resembled by any other residue of IL-10. To further support this hypothesis, future experiments could test the binding affinity of different IL-10 mutants for GAGs, e.g. by means of NMR or calorimetry methods. Furthermore, the effect of R107A mutation on the biological activity of IL-10 could be investigated. Our results provide important insights about the determinants of IL-10–GAG interaction and shed light on the molecular basis of the functional role of GAGs in IL-10 biology.

Acknowledgements

This work was supported by the German Research Foundation (SFB-TRR67, project A7). JPG was supported by a PhD fellowship of the Studienstiftung des deutschen Volkes. The authors thank the ZIH TU Dresden for providing high-performance computational resources. We are also very thankful to Georg Künze for helpful comments and stimulating discussion.

We thank the Dresden International Graduate School for Biomedicine and Bioengineering (DIGS-BB) and the German Research Foundation (DFG) for their support with the publication costs in the context of the Excellence Initiative.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2015.09.003.

References

[1] K.W. Moore, R. de Waal Malefyt, R. L. Coffman, A. O'Garra, Interleukin-10 and the interleukin-10 receptor, Ann. Rev. Immunol. 19 (1) (2001) 683.
[2] D.F. Fiorentino, M.W. Bond, T.R. Mosmann, Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones, J. Exp. Med. 170 (6) (1989) 2081–2095, http://dx.doi.org/10.1084/jem.170.6.2081.
[3] D.F. Fiorentino, A. Zlotnik, P. Vieira, T.R. Mosmann, M. Howard, K.W. Moore, A. O'Garra, IL-10 acts on the antigen-presenting cell to inhibit cytokine production by Th1 cells, J. Immunol. 146 (10) (1991) 3444–3451.
[4] L. Ding, E.M. Shevach, IL-10 inhibits mitogen-induced T cell proliferation by selectively inhibiting macrophage costimulatory function, J. Immunol. 148 (10) (1992) 3133–3139.
[5] R. Kühn, J. Löhler, D. Rennick, K. Rajewsky, W. Müller, Interleukin-10-deficient mice develop chronic enterocolitis, Cell 75 (2) (1993) 263–274, http://dx.doi.org/10.1016/0092-8674(93)80068-P.
[6] S.A. Eming, S. Werner, P. Bugnon, C. Wickenhauser, L. Siewe, O. Utermohlen, J.M. Davidson, T. Krieg, A. Roers, Accelerated wound closure in mice deficient for interleukin-10, Am. J. Pathol. 170 (1) (2007) 188–202, http://dx.doi.org/10.2353/ajpath.2007.060370.
[7] K. Asadullah, W. Sterry, H.D. Volk, Interleukin-10 therapy-review of a new approach, Pharmacol. Rev. 55 (2) (2003) 241–269, http://dx.doi.org/10.1124/pr.55.2.
[8] A. Roers, W. Müller, Distinct functions of interleukin-10 derived from different cellular sources, Curr. Immunol. Rev. 4 (1) (2008) 37–42, http://dx.doi.org/10.2174/157339508783597262.
[9] D.M. Mosser, X. Zhang, Interleukin-10: new perspectives on an old cytokine, Immunol. Rev. 226 (2008) 205–218, http://dx.doi.org/10.1111/j.1600-065X.2008.00706.x.
[10] F.N. Lauw, D. Pajkrt, C.E. Hack, M. Kurimoto, S.J. van Deventer, T. van der Poll, Proinflammatory effects of IL-10 during human endotoxemia, J. Immunol. 165 (5) (2000) 2783–2789, http://dx.doi.org/10.4049/jimmunol.165.5.2783.
[11] S. Salek-Ardakani, J.R. Arrand, D. Shaw, M. Mackett, Heparin and heparan sulfate bind interleukin-10 and modulate its activity, Blood 96 (2000) 1879–1888.
[12] G. Künze, J.-P. Gehrcke, M.T. Pisabarro, D. Huster, NMR characterization of the binding properties and conformation of glycosaminoglycans interacting with interleukin-10, Glycobiology 24 (11) (2014) 1036–1049, http://dx.doi.org/10.1093/glycob/cwu069.
[13] L.U. Esko, J.D. Kimata K, Proteoglycans and sulfated glycosaminoglycans, in: E.J. Varki A, R.D. Cummings (Eds.), Essentials of Glycobiology, 2nd ed., Cold Spring Harbor Laboratory Press, 2009.
[14] L. Pellegrini, D.F. Burke, F.v. Delft, B. Mulloy, T.L. Blundell, Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin, Nature 407 (6807) (2000) 1029–1034, http://dx.doi.org/10.1038/35039551.
[15] J. Schlessinger, A.N. Plotnikov, O.A. Ibrahimi, A.V. Eliseenkova, B.K. Yeh, A. Yayon, R.J. Linhardt, M. Mohammadi, Crystal structure of a ternary FGF–FGFR–heparin complex reveals a dual role for heparin in FGFR binding and dimerization, Mol. Cell 6 (3) (2000) 743–750, http://dx.doi.org/10.1016/S1097-2765(00)00073-3.
[16] T.M. Handel, Z. Johnson, S.E. Crown, E.K. Lau, M. Sweeney, A.E. Proudfoot, Regulation of protein function by glycosaminoglycans – as exemplified by chemokines., Ann. Rev. Biochem. 74 (1) (2005) 385–410, http://dx.doi.org/10.1146/annurev.biochem.72.121801.161747.
[17] A. Imberty, H. Lortat-Jacob, S. Pérez, Structural view of glycosaminoglycan-protein interactions, Carbohydr. Res. 342 (3) (2007) 430–439, http://dx.doi.org/10.1016/j.carres.2006.12.019.
[18] N. Gandhi, R. Mancera, The structure of glycosaminoglycans and their interactions with proteins, Chem. Biol. Drug Des. 72 (6) (2008) 455–482, http://dx.doi.org/10.1111/j.1747-0285.2008.00741.x.
[20] S.A. Samsonov, J.-P. Gehrcke, M.T. Pisabarro, Flexibility and explicit solvent in molecular dynamics-based docking of protein–glycosaminoglycan systems, J. Chem. Inf. Model. 54 (2) (2014) 582–592, http://dx.doi.org/10.1021/ci4006047.
[20] A. Zdanov, C. Schalk-Hihi, A. Wlodawer, Crystal structure of human interleukin-10 at 1.6 Å resolution and a model of a complex with its soluble receptor, Protein Science 5 (10) (1996) 1955–1962, http://dx.doi.org/10.1002/pro.5560051001.
[21] R. Syto, N.J. Murgolo, E.H. Braswell, P. Mui, E. Huang, W.T. Windsor, Structural and biological stability of the human interleukin 10 homodimer, Biochemistry 37 (48) (1998) 16943–16951, http://dx.doi.org/10.1021/bi981555y.
[22] A. Zdanov, Structural analysis of cytokines comprising the IL-10 family, Cytokine Growth Factor Rev. 21 (5) (2010) 325–330.
[23] D. Case, T. Darden, T. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, M. Crowley, R. Walker, W. Zhang, K. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. Wong, F. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. Mathews, M. Seetin, C. Sagui, V. Babin, P. Kollman, Amber 12, University of California, San Francisco, 2012, Also see http://ambermd.org/.
[24] K. Kirschner, A. Yongye, S. Tschampel, J. González-Outeiriño, C. Daniels, L. Foley, R. Woods, GLYCAM06: a generalizable biomolecular force field. carbohydrates, J. Comput. Chem. 29 (4) (2008) 622–655, http://dx.doi.org/10.1002/jcc.20820.
[25] W. Humphrey, A. Dalke, K. Schulten, VMD - Visual molecular dynamics, J. Mol. Graphics 14 (1996) 33–38.
[26] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald, J. Chem. Theory Comput. 9 (9) (2013) 3878–3888, http://dx.doi.org/10.1021/ct400314y.
[27] B.R. Miller, T.D. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBS.A.py. An efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (9) (2012) 3314–3321, http://dx.doi.org/10.1021/ct300418h.
[28] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (7) (2013) 3084–3095, http://dx.doi.org/10.1021/ct400341p.
[29] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: E. Simoudis, J. Han, U.M. Fayyad (Eds.), Second International Conference on Knowledge Discovery and Data Mining, AAAI Press, 1996, pp. 226–231.
[30] A. Zdanov, C. Schalk-Hihi, A. Gustchina, M. Tsang, J. Weatherbee, A. Wlodawer, Crystal structure of interleukin-10 reveals the functional dimer with an unexpected topological similarity to interferon gamma, Structure 3 (6) (1995) 591–601, http://dx.doi.org/10.1016/S0969-2126(01)00193-9.
[31] R. Sabat, IL-10 family of cytokines, Cytokine Growth Factor Rev. 21 (5) (2010) 315–324, http://dx.doi.org/10.1016/j.cytogfr.2010.11.001.
[32] K. Josephson, N.J. Logsdon, M.R. Walter, Crystal structure of the IL-10/IL-10R1 complex reveals a shared receptor binding site, Immunity 15 (1) (2001) 35–46, http://dx.doi.org/10.1016/S1074-7613(01)00169-8.
[33] S.I. Yoon, B.C. Jones, N.J. Logsdon, B.D. Harris, A. Deshpande, S. Radaeva, B.A. Halloran, B. Gao, M.R. Walter, Structure and mechanism of receptor sharing by the IL-10R2 common chain, Structure 18 (5) (2010) 638–648, http://dx.doi.org/10.1016/j.str.2010.02.009.
[34] S.I. Yoon, B.C. Jones, N.J. Logsdon, M.R. Walter, Same structure, different function: crystal structure of the Epstein–Barr virus IL-10 bound to the soluble IL-10R1 chain, Structure 13 (4) (2005) 551–564, doi:10.1016/j.str.2005.01.016.
[35] S. Pletnev, E. Magracheva, A. Wlodawer, A. Zdanov, A model of the ternary complex of interleukin-10 with its soluble receptors, BMC Struct. Biol. 5 (1) (2005) 10, http://dx.doi.org/10.1186/1472-6807-5-10.
[36] S.I. Yoon, N.J. Logsdon, F. Sheikh, R.P. Donnelly, M.R. Walter, Conformational changes mediate interleukin-10 receptor 2 (IL-10R2) binding to IL-10 and assembly of the signaling complex, J. Biol. Chem. 281 (46) (2006) 35088–35096, http://dx.doi.org/10.1074/jbc.M606791200.
[37] N.J. Logsdon, A. Deshpande, B.D. Harris, K.R. Rajashankar, M.R. Walter, Structural basis for receptor sharing and activation by interleukin-20 receptor-2 (IL-20R2) binding cytokines, Proc. Natl. Acad. Sci. U. S. A. 109 (31) (2012) 12704–12709, http://dx.doi.org/10.1073/pnas.1117551109.
[38] S. Najjam, R.V. Gibbs, M.Y. Gordon, C.C. Rider, Characterization of human recombinant interleukin 2 binding to heparin and heparan sulfate using an ELISA approach, Cytokine 9 (12) (1997) 1013–1022, http://dx.doi.org/10.1006/cyto.1997.0246.
[39] S. Najjam, B. Mulloy, J. Theze, M. Gordon, R. Gibbs, C.C. Rider, Further characterization of the binding of human recombinant interleukin 2 to heparin and identification of putative binding sites, Glycobiology 8 (5) (1998) 509–516.
[40] R.J. Lipscombe, A.M. Nakhoul, C.J. Sanderson, D.R. Coombe, Interleukin/5 binds to heparin/heparan sulfate. A model for an interaction with extracellular matrix, J. Leukoc. Biol. 63 (3) (1998) 342–350.