Research Article

The potential role of glycosaminoglycans in serum amyloid A fibril formation by *in silico* approaches

Martyna Maszota-Zieleniak * **Annemarie Danielsson and Sergey A. Samsonov** *

Faculty of Chemistry, *University of Gdańsk, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland*

Correspondence to Martyna Maszota-Zieleniak and Sergey A. Samsonov: *sergey.samsonov@ug.edu.pl (S.A. Samsonov)*
https://doi.org/10.1016/j.mbplus.2021.100080

Abstract

Serum amyloid A (SAA) is actively involved in such pathological processes as atherosclerosis, rheumatoid arthritis, cancer and Alzheimer's disease by its aggregation. One of the factors that can attenuate its aggregation and so affects its physiological role is its interactions with glycosminoglycans (GAGs), linear anionic periodic polysaccharides. These molecules located in the extracellular matrix of the cell are highly variable in their chemical composition and sulfation patterns. Despite the available experimental evidence of SAA-GAG interactions, no mechanistic details at atomic level have been reported for these systems so far. In our work we aimed to apply diverse computational tools to characterize SAA-GAG complexes formation and to answer questions about their potential specificity, energetic patterns, particular SAA residues involved in these interactions, favourable oligomeric state of the protein and the potential influence of GAGs on SAA aggregation. Molecular docking, conventional and replica exchange molecular dynamics approaches were applied to corroborate the experimental knowledge and to propose the corresponding molecular models. SAA-GAG complex formation was found to be electrostatics-driven and rather unspecific of a GAG sulfation pattern, more favorable for the dimer than for the monomer when binding to a short GAG oligosaccharide through its N-terminal helix, potentially contributing to the unfolding of this helix, which could lead to the promotion of the protein aggregation. The data obtained add to the specific knowledge on SAA-GAG systems and deepen the general understanding of protein-GAG interactions that is of a considerable value for the development of GAG-based approaches in a broad theurapeutic context.

Introduction

Serum amyloid A (SAA) is a highly conserved protein that belongs to the apolipoprotein family and is mainly produced by the liver. Its expression is stimulated by interleukin-1 (IL-1), interleukin-6 (IL-6) and tumor necrosis factor (TNF) [1]. To date, three isoforms of SAA have been described in mouse and in human: SAA1, SAA2 and SAA3. Additionally, a fourth isoform SAA4 has been identified to be produced steadily in the liver [2]. SAA is involved in the metabolism of high-density lipoprotein (HDL), the body's defense against pathogens and in cholesterol transport [3]. It is also presented during diseases such as atherosclerosis, rheumatoid arthritis, cancer and Alzheimer's disease [4]. It plays a key role in the life-threatening complications of rheumatoid arthritis caused by deposition of insoluble amyloids arising from SAA [5–8]. In addition, it is also involved in the body's defense during the acute phase of inflammation (such as infection or tissue trauma), in which its concentration in blood can increase 1000-fold within 24 h [9,10]. Sustained high levels of SAA during

the acute phase of inflammation may lead to aggregation and formation of amyloid deposits by the SAA protein [11].

SAA1 and SAA2 isoforms are dominant in the plasma within the acute phase of inflammation. Each of these isoforms consists of 104 amino acid residues, which corresponds to the molecular weight of 12–14 kDa. They differ from each other by the amino acid residues at positions 60, 68, 69, and 90 [12]. Due to its ability to aggregate, the structure of SAA has been unknown for many years. Amino acid sequence analysis, circular dichroism studies [13], and homologous modeling [14] provided an overall picture of the structure as a typical globular protein. Based on the homology of the SAA protein to the N-terminal domain of hemocyanin, Stevens [14] suggested that the protein contains about 80% of helical structure. He also indicated that the C-terminus of the SAA forms a random coil structure and is characterized by a high mobility. The C-terminus of the SAA protein is rich in proline residues, and most likely the presence of these amino acid residues is the main reason for its disordered structure. The presence of the proline residues, due to the cyclic structure of their side chains, introduces steric hindrances that does not allow the formation of stable secondary structural elements [15]. The first crystal structure of the SAA protein was published in 2014 by Lu et al. [16]. Two different oligomeric states were proposed based on it: a tetramer and a hexamer. The oligomers are formed by monomeric subunits, where each monomer is made of four α-helices arranged in an antiparallel way to each other. The hexametric structure of the SAA consisting of the two homotrimers is stabilized by hydrophobic interactions. Mainly the amino acid residues of the first α-helix of each subunit are involved in stabilizing the structure of the whole oligomer, which indicates that the N-terminal fragment of the protein is responsible for the aggregation of the SAA protein.

Secondary amyloidosis, a complication of rheumatoid arthritis, begins when serum levels of the SAA increase as a result of chronic inflammation. Probably one of the factors contributing to the development of the disease is genetic predisposition [17]. It is also believed that there is an uncharacterized amyloid formation enhancing protein that may have an impact on the onset of the disease [18]. The full molecular mechanism leading to the formation of amyloid deposits is still unknown. There is a hypothesis that SAA when bound to its receptor undergoes proteolysis, because the amyloid plaques contain various lengths fragments of the SAA in addition to its full-chain sequences. This is most likely the result of the incomplete protein degradation [19]. In addition, amyloid deposits accumulating in the parenchymal organs are also rich in non-fibrous components: glycosaminoglycans (GAGs), glycosylated plasma protein, components of the extracellular space (perkelan, laminin, entakin, collagen type IV), apolipoproteins E and J [20,21]. The role of the non-fibrous components has not been understood so far, however, it is believed that GAGs may influence the formation of protofilaments, while glycosylated plasma protein may stabilize amyloid fibrils [18].

Due to the high aggregation capacity of the SAA protein, it is important to find suitable inhibitors that would prevent the formation of toxic SAA aggregates. Bokareva *et al.* revealed the possible interaction between SAA and human cystatin C (hCC) [22]. Their data point to a direct interaction between SAA and hCC, leading to functional consequences to serum amyloid A, and thus to the loss of its ability to accumulate in the human body. Those results indicate that hCC can inhibit the oligomerization of SAA. The *in vitro* studies performed by Spodzieja *et al.* provided direct information about the binding sites of the hCC/SAA complex [23,24]. The epitope extraction/excision mass spectrometry study showed that the binding sites are located at the C-terminal, 19-amino-acid fragment (86–104) of SAA, and the C-terminal 28-amino-acid sequence (93–120) of hCC. Those findings were supplemented by molecular dynamics simulations for the complexes formed by hCC and the SAA (86–104) fragment that resulted in identification of the crucial interaction sites [25].

The studies focused on the participation of GAGs in SAA aggregation could be promising for gaining the knowledge needed to prevent the protein aggregation. It was found that one of GAGs, heparan sulfate (HS), is present in the amyloid deposits [26]. This suggests that GAGs are a key players in the formation of amyloid fibrils. Bazar *et al.* have proposed that heparin (HP), which is HS class member with particular sulfation pattern, acts as a scaffold for SAA fibrils via electrostatic interactions [27]. The HS molecule contains highly sulfated domains separated by elastic fragments with low sulfation [28]. It has been suggested that the sulfate groups in GAGs are important to facilitate the formation of fibrils of amyloidogenic proteins [29]. A study performed by Lu *et al.* identified two HS binding sites in SAA protein: first consisting of Arg 15, Arg 19 and Arg 47 from each monomer, the second consisting of Arg 1, Arg 62 and His 71, which is also the binding site of HDL [16]. The HDL scavenging by HS has been identified as promoting the formation of SAA aggregates. The GAG length required for SAA aggregation is dp14 (dp stand for the degree of polymerization) and above [30]. Egashira *et al.* showed that HP facilitates the formation of SAA fibrils [6]. The fluorescence and circular dichroism measurements showed that SAA(1–27) peptide is involved in HP-induced amyloidogenesis. Relatively small changes in fluorescence and a completely different pattern in the CD spectrum were observed for the SAA(43–63) peptide, whereas SAA(77–104) peptide showed no

HP-induced changes. Electron microscopy showed that the SAA(1–27) peptide forms short and rigid fibrils, while the SAA(43–63) peptide forms much longer and flexible fibrils. Those studies suggest that the N-terminal region of the SAA protein plays a key role as a rigid core of the fibrils and the middle region facilitates fiber elongation in HP-induced SAA amyloidogenesis. Continuation of this research conducted by Takase *et al*. [31] has shown that the GAGs sulphate groups represent key structural requirements for the SAA fibril formation. They also performed microscopic analyzes which showed that HS, which is on average less sulfated than HP, but containing highly sulfated domains, has a relatively high potential to facilitate fibril formation compared to other GAGs. Based on these experimental findings, in our present work we aimed to characterize how GAGs interact with the SAA protein at the molecular level applying rigorous computational approaches, some of which are particularly developed to effectively deal with protein-GAG systems. In this manuscript, we present atomistic models of SAA-GAG systems and propose a potential molecular mechanism underlying the intermolecular interactions that could be key for understanding how SAA fibrils are formed, and which role GAGs can play in this process.

Results and discussion

PBSA electrostatic potential analysis

The SAA protein is rich in negatively charged amino acid residues and has a total net charge of −2. Only the N-terminus of the protein contains positively charged residues that can potentially bind anionic GAGs.

Electrostatic potential isosurfaces for the dimer of SAA are shown in Fig. 1 (right panel). The highest positive potential, the most attractive for negatively charged GAGS, is located in the proximity of the first N-terminal helix. This SAA fragment corresponds to the protein sequence containing positively charged amino acid residues: Arg 1, Arg 15, Arg 19 and Arg 25. There is also some positive potential close to the C-terminal part of the protein. It is induced by Arg 96 and Lys 103 and therefore this fragment has also been taken into account when sizing the grid box in the docking procedure.

Molecular docking and molecular dynamics simulations

Five different GAGs of four different lengths have been docked to the SAA monomer and dimer. The results obtained by molecular docking were sorted according to the best energy and divided into clusters (clustering parameters are listed in Supplementary Table 1). The Figs. 2 and 3 presents 50 most favourable structures and the structures from the most numerous clusters for all docked combinations.

When analyzing the data shown in Fig. 2, it is observed that the GAG ligands, regardless of their length, docked mainly to the N-terminal fragment of the SAA monomer. The only exception is HS with the GlcNS-GlcUA periodic unit, which also can be found near the C-terminal fragment of the protein (for dp2, dp6 and dp8). It can also be seen that the longer the GAG chain is, more probably it interacts with the C-terminus, mainly i case of HS with GlcNS-GlcUA, GlcNS-IdoU(2S) and GlcNS (6S)-GlcUA periodic units. This is especially evident for GlcNS-GlcUA structures where the GAG interacts with the SAA monomer through its N- and C-terminus simultaneously.

Essentially similar docking results were obtained for the SAA dimer – GAG complexes (Fig. 3). As in the complexes with the SAA monomer GAGs are mainly docked at the top of the N-terminal helix in the complexes with the SAA dimer. In addition, GAGs are also partially docked in the middle parts of the first helix in the case of the dimer. For most of the studied GAGs, they docked both on the surface of the dimer as well as within the dimer interface established by the two helices from two different SAA subunits. A significant difference for the monomer and the dimer is observed only for HP dp2 and GlcNS-GlcUA dp2 potentially due to their small size. For the first of them there is a tendency to dock inside the dimer, while for the second GAG, the ligands, in contrast, are mainly docked outside the dimeric interface.

Free energy MM-GBSA analysis and per residue decomposition

We performed free energy analysis based on the MD trajectories obtained. Table 1 consists of the results of the MM-GBSA analysis for the SAA monomer/GAG complexes. The most stable interactions between the protein and a GAG, reflected in ΔG values below −20.0 kcal/mol are observed for cluster 3 of SAA monomer/HP dp2 (–22.3 kcal/mol), cluster 1 of SAA monomer/HP dp6 (–22.4 kcal/mol), cluster 2 of SAA monomer/GlcNS-GlcUA dp8 (-20.6 kcal/mol), cluster 2 of SAA monomer/GlcNS-IdoU(2S) dp6 (–22.8 kcal/mol), cluster 1 of SAA monomer/GlcNS-IdoU(2S) dp8 (–22.8 kcal/mol) and cluster 2 of SAA monomer/GlcNS(6S)-GlcUA dp4 (-20.3 kcal/mol). It is worth noting that the binding strength of HP decreased with its length, which is probably due to the fact that this GAG is the most charged one used in the analysis, while the protein has a net negative charge. In some SAA monomer/GAG complexes ΔG values decrease with the length of the GAG (SAA monomer/GlcNS-GlcUA), or they decrease to their minimum at dp4 (SAA monomer/GlcNS6S-GlcUA) or dp6 (SAA monomer/CS6) and then increase. In case of CS6 dp8, in two out of ten MD simulations, ligand dissociation events

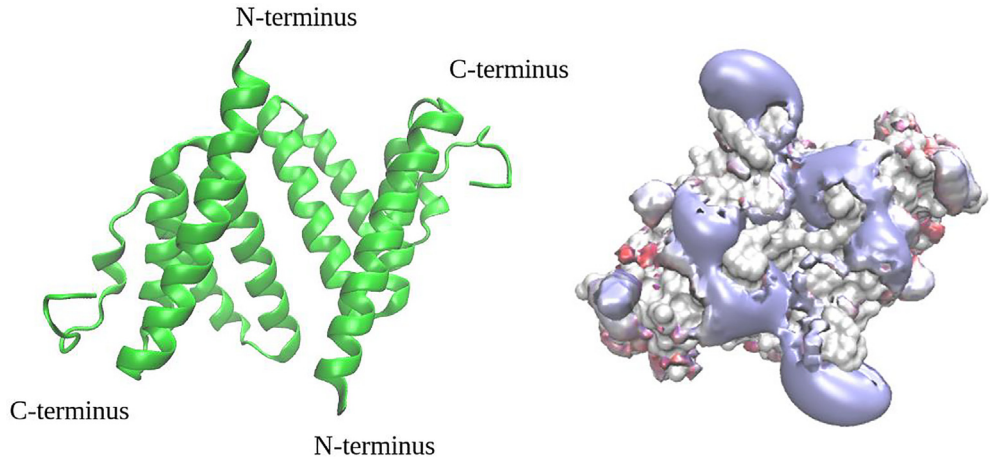

Fig. 1. X-ray structure of the SAA dimer (PDB ID: 4IP9) in cartoon representation (left); electrostatic potential isosurfaces for the SAA dimer in surface representation (red, −2 kcal/mol/e; blue, +2 kcal/mol/e, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

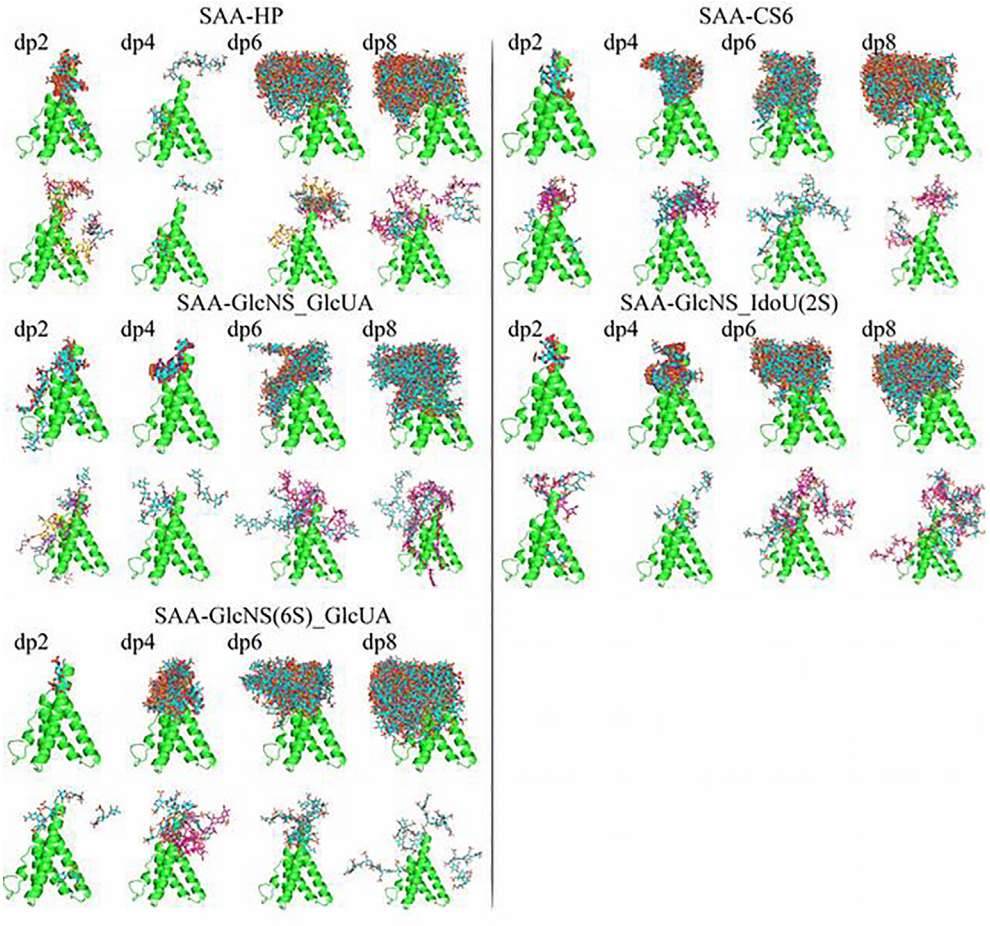

Fig. 2. Top 50 structures after docking procedure (top panels) and 5 structures from each analyzed cluster after MD simulations for SAA monomer – GAG complexes (protein in cartoon presentation, green color; GAG molecules in sticks presentation; light blue color – carbon atoms in top panels, blue/magenta/yellow color – carbon atoms, red color – oxygen atoms, blue color – nitrogen atoms, yellow color – sulfur atoms). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

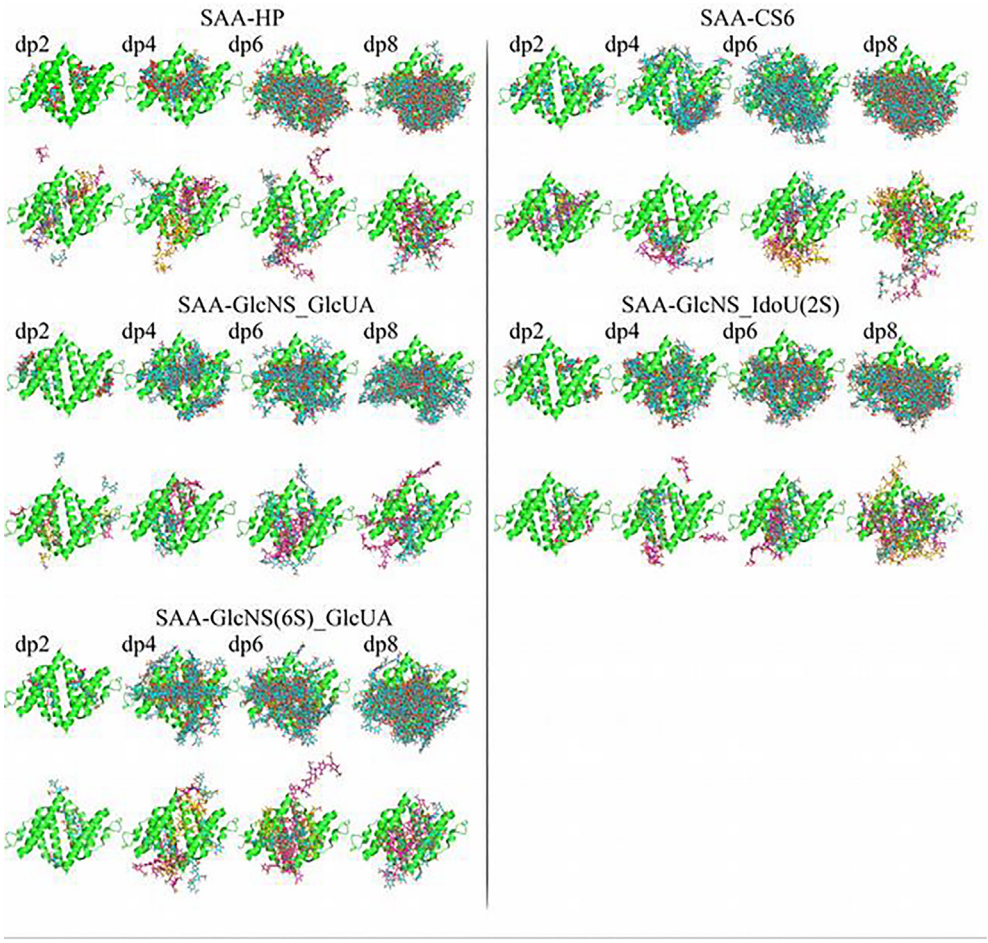

Fig. 3. Top 50 structures after docking procedure (top panels) and 5 structures from each analyzed cluster after MD simulations for SAA dimer – GAG complexes (protein in cartoon presentation, green color; GAG molecules in sticks presentation; light blue color – carbon atoms in top panels, blue/magenta/yellow color – carbon atoms, red color – oxygen atoms, blue color – nitrogen atoms, yellow color – sulfur atoms). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

were observed, which explains the unfavourable mean values of the binding energies. Extrapolating these data obtained for the short GAG oligosaccharides, one can hypothesize that *in vivo* a long and heterogeneous GAG as HS, which is composed by domains with different net sulfation and, therefore, different local charge density, would prefer to interact with SAA by its parts that are not fully sulfated to provide favourable and electrostatically balanced interactions.

Results of the MM-GBSA analysis for the SAA dimer/GAG complexes are summarized in Table 2. In case of complexes formed by the dimer of the SAA protein, there is a much greater variety of free energy values than for SAA monomer/GAG complexes. In contrast to the monomer, the energy of the SAA dimer/HP complexes increases as the degree of HP polymerization increases. On the other hand, for the complexes with CS6, the energy decreases to dp6 reaching the lowest observed free energy of binding value for the SAA dimer/CS6 dp6 complex (-47.3 kcal/mol). The only similar trend to the complexes formed by the monomeric protein can be noted for the SAA dimer/GlcNS-GlcUA where the ΔG values decrease with the length of the GAG. Because in general, the dimer binds GAGs stronger than the monomer, here we list the complexes with ΔG values less than -25.0 kcal/mol, corresponding to the strongest binding observed: cluster 1 and 2 SAA dimer/CS6 dp6 (-47.3 kcal/mol and -26.8 kcal/mol respectively), cluster 2 SAA dimer/CS6 dp8 (-32.3 kcal/mol), cluster 2 SAA dimer/GlcNS-GlcUA dp6 (-30.6 kcal/mol), cluster 1 SAA dimer/GlcNS-GlcUA dp8 (-30.1 kcal/mol), cluster 2 SAA dimer/GlcNS-IdoU(2S) dp6 (-29.8 kcal/mol) and cluster 3 SAA dimer/GlcNS(6S)-GlcUA dp6 (-29.4 kcal/mol).

When analyzing the binding energies obtained in these calculations normalized by the length of

Table 1 MM-GBSA analysis of SAA monomer/GAG complexes.

SAA monomer/GAG complex	Cluster number	ΔG per cluster [kcal/mol]	Average ΔG [kcal/mol]
SAA/HP dp2	1	−10.8	−16.0
	2	−14.9	
	3	−22.3	
SAA/HP dp4	1	−15.6	−15.6
SAA/HP dp6	1	−22.4	−14.6
	2	−16.8	
	3	−4.6	
SAA/HP dp8	1	−10.2	−12.9
	2	−14.9	
SAA/CS6 dp2	1	−7.6	−10.9
	2	−14.2	
SAA/CS6 dp4	1	−14.3	−15.3
	2	−16.3	
SAA/CS6 dp6	1	−17.6	−17.6
SAA/CS6 dp8	1	0.0	−2.7
	2	−6.2	
SAA/GlcNS_GlcUA dp2	1	−5.3	−6.9
	2	−9.4	
	3	−3.9	
	4	−7.5	
	5	−8.2	
SAA/GlcNS_GlcUA dp4	1	−9.9	−9.9
SAA/GlcNS_GlcUA dp6	1	−12.0	−13.2
	2	−14.5	
SAA/GlcNS_GlcUA dp8	1	−10.6	−15.6
	2	−20.6	
SAA/GlcNS_IdoU2S dp2	1	−12.8	−11.8
	2	−10.7	
SAA/GlcNS_IdoU2S dp4	1	−9.9	−9.9
SAA/GlcNS_IdoU2S dp6	1	−15.2	−19.0
	2	−22.8	
SAA/GlcNS_IdoU2S dp8	1	−24.9	−18.7
	2	−12.5	
SAA/GlcNS6S_GlcUA dp2	1	−10.1	−10.1
SAA/GlcNS6S_GlcUA dp4	1	−17.8	−19.1
	2	−20.3	
SAA/GlcNS6S_GlcUA dp6	1	−14.1	−14.1
SAA/GlcNS6S_GlcUA dp8	1	−9.2	−9.2

the simulated GAGs, one can get qualitatively different insights into the role of electrostatics and potential specificity of the binding in this system. If the electrostatics would have been the only substantial component determining binding energy, there would be equal values of binding energies obtained after such normalization for a GAG of the same type. However, the obtained differences of the values suggest that for all analyzed GAG types there are also other factors as the entropic contribution and van der Waals binding free energy component that are not to be disregarded for the description of these interactions. For HP, the most charged GAG from the series, there is a clear dominance of the electrostatics, which is manifested by the stronger repulsion between the receptor and HP upon its elongation. However, for other GAGs the relation between the length of GAGs and their binding strength is more complex which would suggest that GAG interactions with SAA, although dominated by electrostatics, are partially GAG sulfation pattern specific.

Table 2 MM-GBSA analysis of SAA dimer/GAGs complexes.

SAA dimer/GAG complex	Cluster number	ΔG per cluster [kcal/mol]	Average ΔG [kcal/mol]
SAA/HP dp2	1	−17.8	−17.2
	2	−11.9	
	3	−21.5	
	4	−18.4	
SAA/HP dp4	1	−12.0	−17.3
	2	−19.1	
	3	−18.5	
SAA/HP dp6	1	−6.1	−10.1
	2	−14.1	
SAA/HP dp8	1	−7.7	−12.8
	2	−17.0	
SAA/CS6 dp2	1	−12.6	−15.0
	2	−13.5	
	3	−20.2	
	4	−13.7	
SAA/CS6 dp4	1	−24.8	−20.4
	2	−15.9	
SAA/CS6 dp6	1	−47.3	−31.1
	2	−26.8	
	3	−19.3	
SAA/CS6 dp8	1	−12.4	−20.0
	2	−32.2	
	3	−15.2	
SAA/GlcNS_GlcUA dp2	1	−1.4	−4.0
	2	−5.4	
	3	−5.2	
SAA/GlcNS_GlcUA dp4	1	−14.0	−15.2
	2	−16.3	
SAA/GlcNS_GlcUA dp6	1	−13.3	−22.0
	2	−30.6	
SAA/GlcNS_GlcUA dp8	1	−30.1	−24.6
	2	−17.7	
SAA/GlcNS_IdoU2S dp2	1	−14.6	−10.5
	2	−6.4	
SAA/GlcNS_IdoU2S dp4	1	−23.5	−26.3
	2	−29.8	
SAA/GlcNS_IdoU2S dp6	1	−22.7	−23.4
	2	−24.2	
SAA/GlcNS_IdoU2S dp8	1	−14.7	−16.4
	2	−21.3	
	3	−14.2	
	4	−13.9	
SAA/GlcNS6S_GlcUA dp2	1	−10.5	−10.5
SAA/GlcNS6S_GlcUA dp4	1	−10.3	−19.2
	2	−22.3	
	3	−25.0	
SAA/GlcNS6S_GlcUA dp6	1	−17.0	−17.3
	2	−10.3	
	3	−29.4	
SAA/GlcNS6S_GlcUA dp8	1	−19.0	−19.4
	2	−19.7	

Selectivity

To further analyze results of the MD simulation, selectivity analysis for the simulated complexes was performed. Complexes of the monomer of the SAA protein with GAGs dp6 were selected as the most representative. For each type of the analyzed complex the probabilities for the RMSatd value and the free energy of binding value were used. The results are summarized in Supplementary Fig. 1. In case of RMSatd values, which were calculated for each structure obtained in the MD simulation in relation to all other structures, it can be seen that for each analyzed complex, two maxima are visible. For the SAA monomer/GlcNS(6S)-GlcUA dp6 these maxima are most separated. The density of probability plots for the ΔG values are different: the maxima of probability are less significantly separated as in case of RMSatd. The most visible two maxima are observed for the SAA monomer/CS6 dp6 complex, which, combined with the two maxima for the RMSatd plots, may indicate the selectivity of this GAG binding by the SAA protein. Similarly, certain selectivity can be observed in the probability plots for SAA monomer/GlcNS(6S)-GlcUA dp6 and SAA monomer/GlcNS-GlcUA dp6 complexes.

The performed analysis of the density of probability for RMSatd and ΔG values is speculative, and it is impossible to determine which GAGs display higher or lower selectivity of binding. Therefore, a quantitative analysis was also performed, the results of which are summarized in Supplementary Table 2. According to the Siebenmorgen *et al*. the closer to 1 or -1 the S'i value is, the greater the specificity for studied GAG. As can be seen in the table below, the S'i values are not close to 1 or -1 for any of the examined complexes. The highest/lowest specificity values were found for SAA monomer/HP dp6 (-0.5 and 0.5) and SAA monomer/CS6 dp6 (0.5).

Hydrogen bond analysis

For the obtained MD results for the SAA monomer/GAGs and SAA dimer/GAGs complexes, the analysis of hydrogen bonds was performed. The occurrence of hydrogen bonds was averaged for the duration of each simulation, then it was also averaged for individual clusters and types of complexes as well as the average of the two subunits in the SAA dimer was calculated. The obtained results are summarized in Figs. 4, 5 and Supplementary Figs. 3-10.

When analyzing the maps hydrogen bonds occurence, it is easy to notice that mainly N-terminal amino acid residues (1–5 of the SAA fragment) are involved in the stabilization of the SAA monomer/GAGs complexes. For the SAA monomer/HP and SAA monomer/GlcNS(6S)-GlcUA complexes hydrogen bonds are also formed between GAGs and Arg 15 and Arg 47 residues, the frequency of these hydrogen bonds increases with increasing degree of GAG polymerization. The same hydrogen bonds are observed for the SAA monomer/CS6 and SAA monomer/GlcNS-IdoU(2S) complexes, but here they most often occur in complexes where the ligand has a dp6 length. A noticeably lower frequency of stabilizing hydrogen bonds can be observed for SAA monomer/GlcNS-GlcUA, only a few of them have a frequency of $\geq$0.3.

When comparing the results of the hydrogen bond analysis for the SAA monomer and dimer complexes with GAGs, an increase in the importance of amino acid residues belonging to the middle part of the protein is noticeable for the SAA dimer. For those complexes both the N-terminal residues and the SAA(15–25) fragment are involved in stabilizing the complexes, with the frequency of occurrence of the GAG-Arg15 binding being much higher than for the SAA monomer/GAGs complexes. In addition to the above-mentioned fragments, also the Arg 47 and Arg 61 residues are involved in the hydrogen bonds formation. As for the SAA monomer, for the SAA dimer, the GlcNS-GlcUA ligand forms the least frequent hydrogen bonds with the protein. These residues calculated to be most important in terms of H-bonding were previously mentioned in the literature as the ones participating in the fibril formation [6,16].

RS-REMD

Furthermore, we aimed to verify the structures obtained by the molecular docking and conventional MD by the RS-REMD approach, which previously was shown to perform very well for protein-GAG complexes [32]. This method allows to correctly predict a binding site and produces an ensemble of GAG-protein complex structures using an implicit solvent model, while its performance is independent of the GAG length. We chose both the SAA monomer and dimers as receptors and HP dp6 and dp24 as ligands. The HP dp24 was selected to see if a significant change in GAG length would affect the GAG-protein docking site, or connect two alternative binding sites (N- and C-termini of SAA). During RS-REMD simulation for SAA monomer/HP dp6 the ligand found the binding site at the N-terminus of the protein very quickly (0.2 ns) and remained there for the rest of the simulation time. It took longer (30 ns) to find the same binding site by the HP dp24 ligand. In this case, the ligand was also docked at the N-terminus of the SAA monomer and remained there through the rest of the RS-REMD simulation. No complex structure was found in which the ligand would interact with the N- and the C-termini of the SAA protein simultaneously.

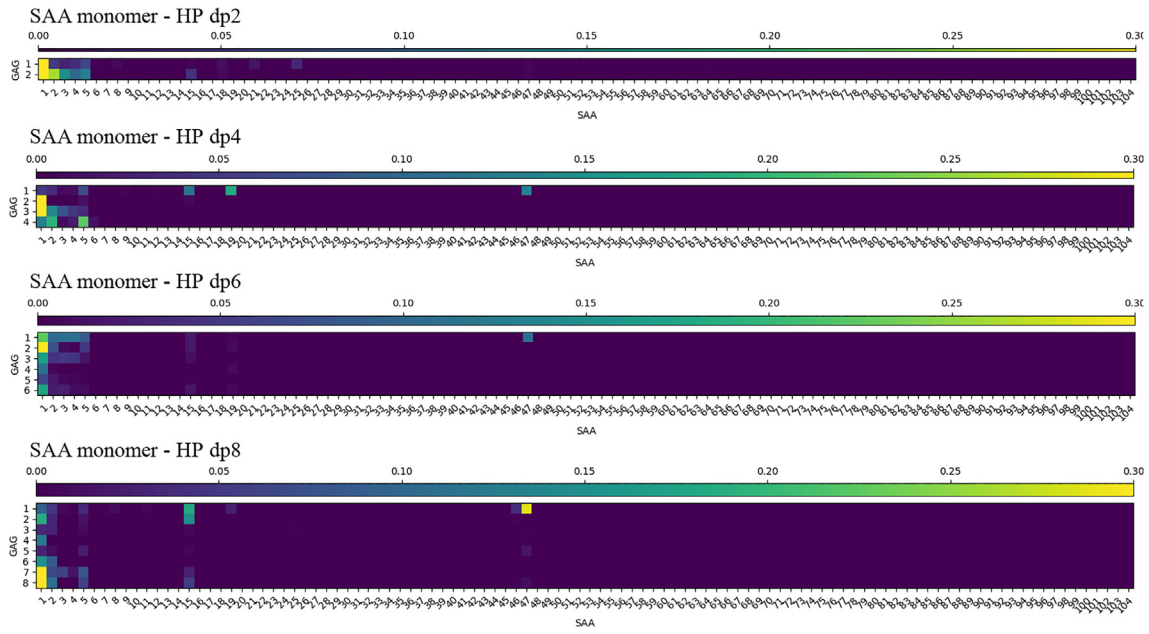

Fig. 4. Heatmap displaying the mean occupancy of hydrogen bonds formed during molecular dynamics (MD) simulations between donor amino acid residues of SAA (in columns) and acceptor residues of HP (in rows). The occupancy is computed as the fraction of MD trajectory frames in which a hydrogen bond was formed as reported by CPPTRAJ between the given residues. The increase in change of fraction values between residue pairs are represented from dark blue to yellow. A hydrogen bond was defined by a distance cutoff of 3.0 Å and a hydrogen bond angle cutoff of 135°. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

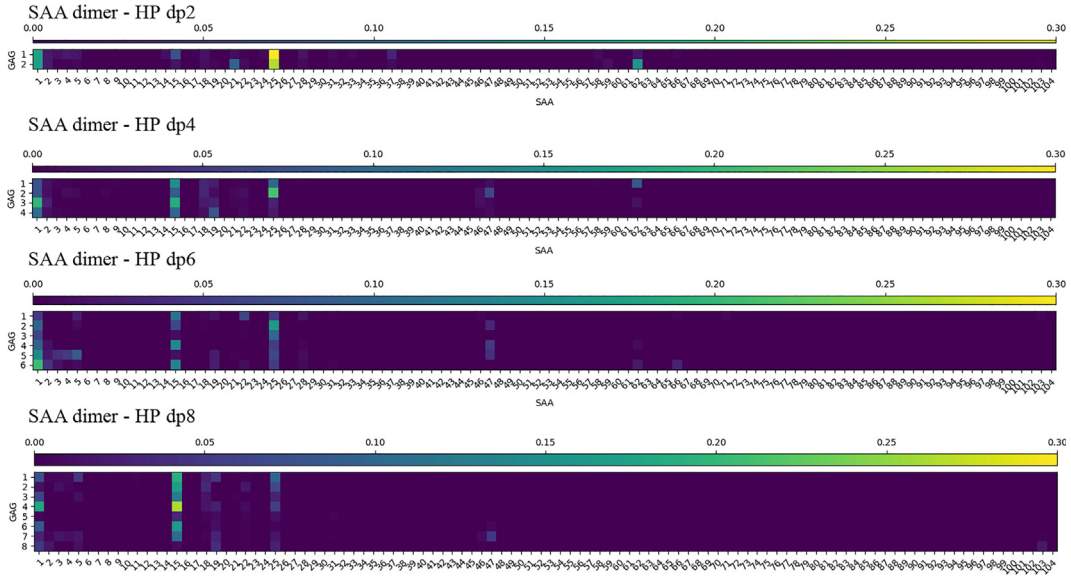

Fig. 5. Heatmap displaying the mean occupancy of hydrogen bonds formed during molecular dynamics (MD) simulations between donor residues of HP (in columns) and acceptor amino acid residues of SAA (in rows). The occupancy is computed as the fraction of MD trajectory frames in which a hydrogen bond was formed as reported by CPPTRAJ between the given residues. The increase in change of fraction values between residue pairs are represented from dark blue to yellow. A hydrogen bond was defined by a distance cutoff of 3.0 Å and a hydrogen bond angle cutoff of 135°. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

As a next step, we refined 100 structures obtained by the RS-REMD approach (Fig. 6). Interestingly, for the SAA dimer/HP dp6 complex the ligand docked to both first and second N-terminal fragments, but did not dock in the middle of the helices as it did when we performed docking by Autodock. This is related to the fact that RS-REMD approach experience difficulties when a binding site is not exposed on the protein surface but rather has a cavity topology [32]. 100 best SAA dimer/HP dp6 complex structures in Fig. 6 are clearly divided into two groups, quite different from what is shown in Fig. 3 for the results obtained by Autodock. Simulations performed for the dimer of the SAA protein and HP dp24 provided qualitatively different results. The ligand did not dock between or near the two first helices, but was searching for a favourable binding site near the C-terminus of one of the monomeric subunits. Perhaps due to its size, HP dp24 could not properly align with the binding sites in the center of the structure. The SAA is a negatively charged protein which surface has mostly negative potential (Fig. 1) and, therefore, can be repulsive for extensively negatively charged long HP it is indicated by the observed dissociation of the GAG after the refinement of RS-REMD docked solutions (Fig. 6). *In vivo,* long HS molecules, therefore, most probably would prefer to interact with SAA by less sulfated domains than HP. Supplementary Table 3 presents the values of the free energy of binding for the 100 complexes after RS-REMD simulations and refinement of the structures. The SAA monomer/HP dp6 complexes have the lowest ΔG values, while the SAA monomer/HP dp24 complexes are unfavourable, which explains the observed dissociation upon the refinement. For the simulations with the SAA dimer, the difference for the dp6 and dp24 ligands is clearly visible. The latter did not bind to the protein, which is also reflected in the positive values of the free binding energy.

To find out which amino acid residues are the most important for the formation of the complexes between the SAA protein and HP dp6, we performed per residue decomposition analysis for the refined complexes obtained from the RS-REMD cimulations (Supplementary Table 4). Arg 1 residue has the highest energetic contribution to the binding (below −15 kcal/mol) and therefore is a key responsible residue for the binding of the HP to the protein. Also the N-terminal residues Ser 2, Phe 3, Phe 4 substantially participate in the formation of SAA monomer/HP dp6 complexes. In contrast, no residues from the C-terminal protein fragment contributed significantly. Arg 1 from both the first and second monomeric subunits for the SAA dimer/HP dp6 complexes also corresponded the lowest value of free energy binding (below −11 kcal/mol). Interestingly, if Arg 1 of one of the subunits contributed the most to HP binding, the next highest contribution belonged to the residue Arg 25 on other monomeric protein subunit (structure no. 83, 331, 490 and 180). There is also less significant commitment of the next N-terminal amino acid residues in favor of the residues in the middle of the sequence (Arg 47, Arg 62) when

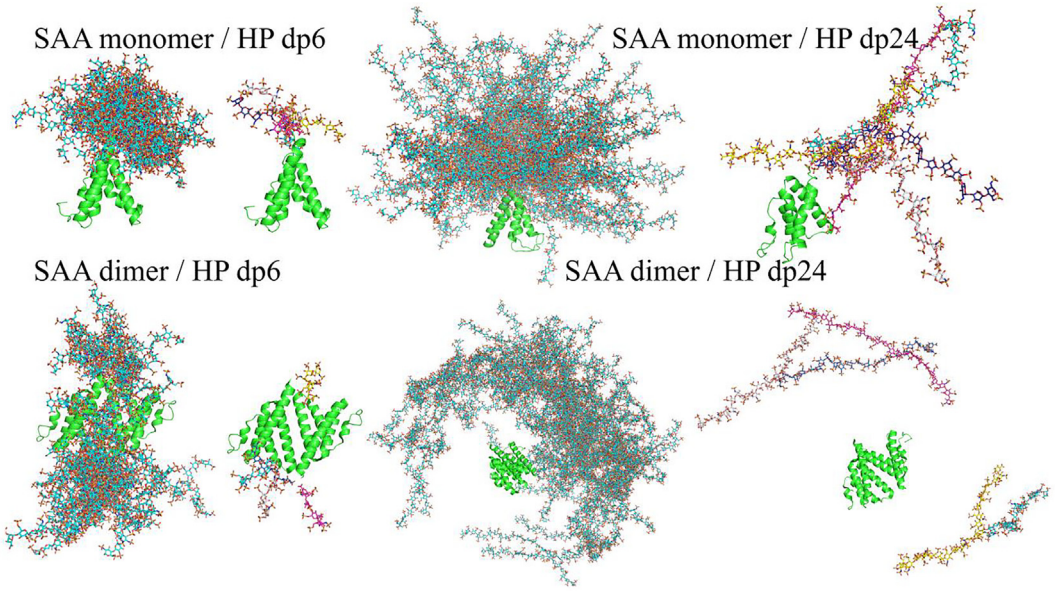

Fig. 6. Results of the RS-REMD simulations for the SAA monomer/dimer with HP dp6 and HP dp24 ligands – 100 most favourable (the lowest E_{ele} values) structures (left) and 5 most favourable structures after refinement (SAA protein green colour, cartoon representation; cyan, yellow, magenta, dark blue, grey colours HP dp6/HP dp24, licorice representation). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

compared to the monomeric complexes. Notably, in the case of complexes in which a SAA dimer is involved, Lys 103 at C-terminus can also be observed as a weak binder (ΔG about -1 kcal/mol).

DSSP analysis

When analyzing the trajectories obtained in the MD simulations, we noticed that the N-terminus of the protein involved in GAGs binding unfolds. Therefore, we aimed to analyze whether the presence of the ligand influences this process. We performed the DSSP secondary structure analysis for the trajectories both obtained in the MD simulations starting from AD docked structures and RS-REMD docked structures (RS-REMD refinement by conventional MD in the explicit solvent) for SAA monomer/HP complexes that corresponded to 7000 ns in total. As a reference, we also simulated the SAA monomer alone using the same protocol as applied for the complexes. The results of the DSSP analysis are summarized in Supplementary Table 5 and Fig. 7, the contents of the 3–10 helix and alpha helix as well as their sum were compared. The analysis was performed for the first 10 amino acid residues corresponding to the first helix in the SAA structure. The helical content for the residues 8–10 is lower for the unbound protein, while it is always higher for the beginning of the N-terminal helix in comparison to the unbound protein in case of the HP dp8 and HP dp6 docked by the RS-REMD approach. This suggests that binding of HP could promote the unfolding of the N-terminal helix. Taking into account the experimental data suggesting that the GAG-induced changes in the SAA (1–27) structure [6] could be crucial for the process of fibrillization, the structural changes computationally observed in the N-terminal helix upon binding GAGs could serve as a potential structural basis of this event (Fig. 8).

Summary

In this computational study, SAA interactions with GAGs were rigorously characterized for 32 systems containing complexes between the monomeric and the dimeric SAA protein with four types of GAGs oligosaccharides with the lengths of dp2, dp4, dp6, dp8. Molecular docking and MD-based analysis of the predicted docking poses suggest the key role of the N-terminal helix for binding GAGs for both oligomeric forms of SAA. The dimer binds GAGs stronger than the monomer, and the residues from both of monomeric units of the protein participate in this binding. The interactions in the analyzed systems are predominantly electrostatics-driven, and are not highly specific, which is also confirmed by the calculated H-bonding patterns. The application of the RS-REMD supports the results of the conventional MD analysis and points out that only short oligomeric GAG sequences can be favourably bound directly to SAA due to its high net negative charge. Finally, the analysis of the SAA secondary structure content in the presence and in the absence of GAGs indicates the potential role of GAGs in the unfolding of the protein, which could serve as a reason for amyloid A fibril formation which is a crucial molecular process for the onset of the SAA-related neurogenerative diseases. Our work contributes to the specific knowledge on the molecular details of SAA-GAG interactions and demonstrates the predictive power of the state-of-art *in silico* approaches applicable for the analysis of protein-GAG systems in general. In the present study, we focused on the monomeric and dimeric forms of SAA which serves as a first step towards understanding higher order oligomers, which are going to be the next objects of our analysis.

Materials and methods

Poisson-Boltzmann accessible surface area (PBSA) electrostatic potential calculations

Electrostatic potential calculations for SAA protein were performed in AMBER16 [33] with default parameters to predict a potential GAG binding site [34].

Molecular docking

SAA1 (named SAA throughout the manuscript) X-ray structure was used (PDB ID: 4IP9) [16] to build the receptors for the molecular docking. The dimeric and monomeric forms were extracted from the hexameric SAA in the experimental structure and minimized in AMBER16 [33] prior to be used as receptors (the protocols are described below as steps 1–3 in the molecular dynamics procedure). Five different GAG types were used as ligands:

1) four structures of HP (GlcNS(6S)-IdoA(2S) dimeric unit) dp2, dp4, dp6 and dp8 with $^1C_4$ ring conformation for IdoA(2S) built from the structure of the unbound HP (PDB ID: 1HPN);

2) four structures of the chondroitin sulfate (CS6, GlcA-GalNAc(6S) dimeric unit) with dp2, dp4, dp6 and dp8;

3) three variants of HS (GlcNS-IdoU(2S), GlcNS-GlcUA and GlcNS(6S)-GlcUA dimeric units, abbreviated as GlcNS_IdoU(2S), GlcNS_GlcUA and GlcNS(6S)_GlcUA, respectively) with dp2, dp4, dp6 and dp8.

The GAGs were built using GLYCAM06 [35] compatible libraries [36] for AMBER [33].

For docking simulations Autodock 3 software was used. A grid box with dimensions of 122 Å $\times$ 80 Å $\times$ 118 Å contained SAA N-terminal fragment. This part of the protein was chosen according to the PBSA analysis results suggesting it to be a potential GAG binding site was used for GAGs docking. The 100 runs of

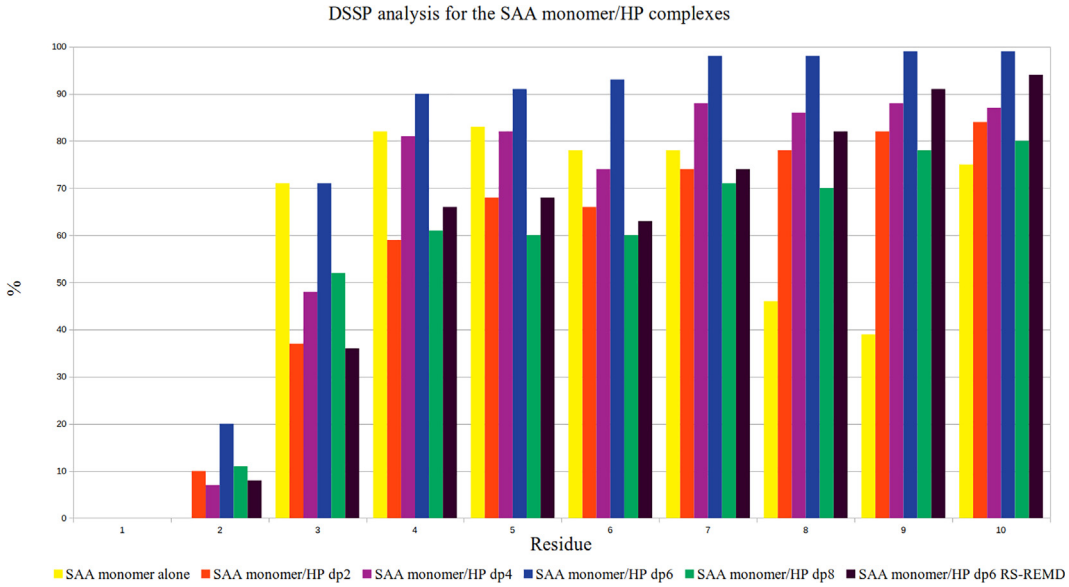

Fig. 7. DSSP analysis for the SAA monomer/HP complexes – percentage of the helical structure in the N-terminal fragment of the SAA monomer per residue.

the Lamarckian genetic algorithm with an initial population size of 300 and a termination condition of 10^5 generations and 9995×10^5 energy evaluations were carried out. Next, the top energetically favorable docking results were clustered with the DBSCAN algorithm [37] using RMSatd (root mean square deviation for atomic types) metric [38]. The parameters for clustering were adjusted for each particular system and are provided in the results section.

Molecular dynamics simulations

Five representative structures, which was found to be an appropriate number to avoid undersampling of docked solutions [39], were chosen for each cluster of the docked structures obtained by molecular docking, and they were used as starting structures for molecular dynamics (MD) simulations. Additionally, simulations of the unbound SAA dimer and monomer were performed to elucidate a potential effect of GAG binding on the

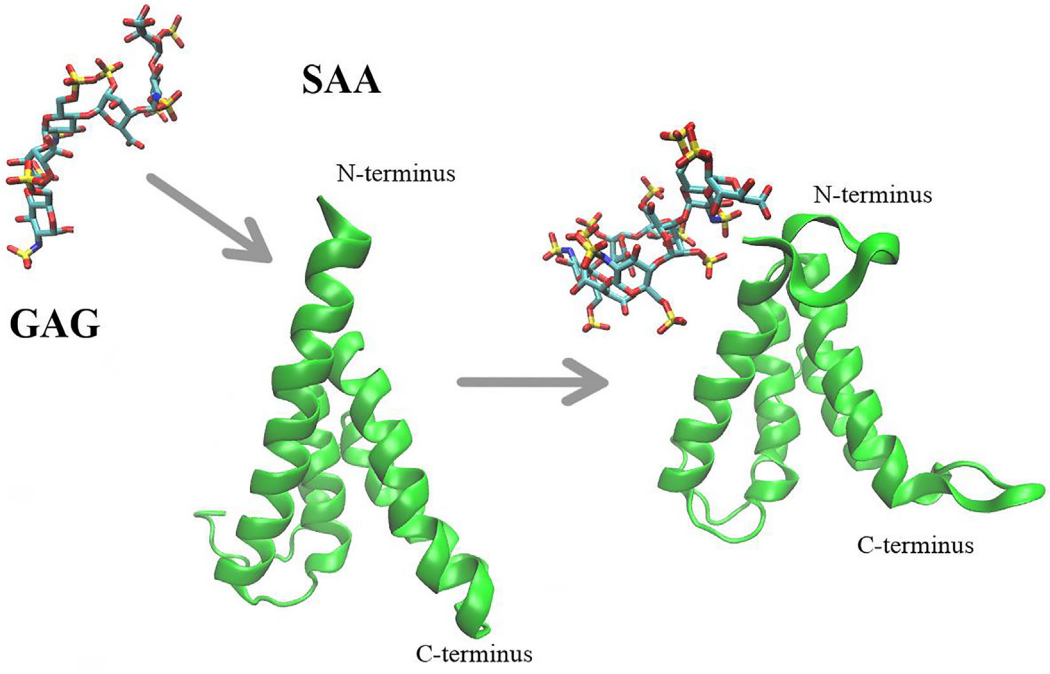

Fig. 8. Schematic representation of GAG binding effect on SAA N-terminal helix unfolding leading to the potential increase of its propensity for fibrillization.

protein unfolding process. All MD simulations were performed in AMBER16 [33] with ff99SBonlysc parameters for proteins and GLYCAM06 [35] parameters for GAGs, respectively. The total length of each MD simulation was 25 ns, which is appropriate for this type of system according to our previous analysis [39]. For the SAA monomer alone and SAA monomer complexed with HP dp2, dp4, dp6 and dp8 complexes MD simulations have been further extended to 100 ns for the DSSP analysis. The MD simulations were performed under periodic boundary conditions with TIP3P [40] cubic water box with 10 Å distance from any complex atom in each direction to the box wall. All initial structures were neutralized by counter ions. The MD protocol was divided into 4 steps: 1) energy minimization to remove close contacts between atoms was performed first with 500 steepest descent cycles and 1000 conjugate-gradient cycles with harmonic force restraints on solute (10 kcal/mol/Å^2) and then with 3000 steepest-descent cycles and 3000 conjugate-gradient cycles without restraints; 2) heated up the systems to 300 K for 10 ps with harmonic force restraints on solute (10 kcal/mol/Å^2); 3) simulation in constant temperature and pressure (NTP ensemble) until the density has stabilized at around 1 g/ml; 4) 25 ns (100 ns for SAA monomer-HP) simulation was carried out at constant pressure using NTP ensemble (with 2 fs time step, the cutoff for non bonded interactions 8 Å and the Particle Mesh Ewald [41] procedure).

The CPPTRAJ program from AMBER Tools 17 was used for the analysis of the obtained MD trajectories.

Free energy MM-GBSA analysis

For the energetic post-processing of the trajectories and per residue decomposition Molecular Mechanics Generalized Born Surface Area (MM-GBSA) was used as implemented in AMBER16 [33]. The MM-GBSA analysis was done for all simulations for the frames corresponding to the fourth step of the MD simulations (as described above). Obtained values for the binding free energy are made up of explicit enthalpy and implicit solvent entropy, and therefore should be understood so rather than as strict values of the free energy of binding.

The selectivity of the binding poses was calculated using method proposed by Siebenmorgen *et al.* [42] and described in details in our previous work [43]. The only modification of this approach applied here consisted of considering best scored pose as a reference for "correct" docking results since there are no reference experimental structures available. In brief, the scores for the docked poses obtained from the MM-GBSA calculations are normalized: the difference between the score and the mean value were divided by the minimum value. Then, all the binding poses were classified as "correct" and "incorrect" depending if they fall within the 1.5 Å cutoff of RMSatd from the reference pose. The normalized scores for both groups were further compared.

Repulsive scaling replica exchange molecular dynamics

For the Repulsive Scaling Replica Exchange Molecular Dynamics (RS-REMD) [44] simulations the protocol as presented in our previous work on protein-GAG complexes was applied [32]. The HP dp6 and dp24 and the SAA dimer and monomer were used as ligands and receptors, respectively. Every ligand was placed at the opposite side of the protein in respect to the putative binding site suggested by the PBSA data. The MD simulations were performed using implicit solvent with igb = 8 with an infinite cutoff for non-bonded interactions. First, the 3000 steps of steepest descent and 3000 steps of conjugate gradient minimization was performed. Next, it was followed heating to 300 K for 10 ns with a a Langevin thermostat (γ = 5 ps^{-1}). In the production run the harmonic restraints of 0.05 kcal/mol/Å^2 were applied on all heavy atoms of protein. The positional restraint of 1.0 kcal/mol/Å^2 between the center of mass (COM) of the receptor and ligand was applied to avoid ligand dissociation too far away from the receptor. For all four systems, 16 replicas were used with different Lennard-Jones (LJ) parameters for atomic pairs from both receptor and ligand molecules: parameters *d* adjusting the effective van der Waals radius and a factor *e* changing the LJ potential well depths were assigned to 0.00 Å, 0.01 Å, 0.02 Å, 0.04 Å, 0.08 Å, 0.12 Å, 0.16 Å, 0.20 Å, 0.24 Å, 0.28 Å, 0.32 Å, 0.38 Å, 0.44 Å, 0.50 Å, 0.58 Å, 0.68 Å and 0.000, 0.015, 0.030, 0.0045, 0.060, 0.075, 0.090, 0.120, respectively as previously calibrated [44]. During production run, 25,000 MD exchange steps between adjacent replicas were produced, obtaining a total of 250 ns per replica. In the production run every 10^3 MD steps an exchange between neighboring replicas was attempted as it is implemented in AMBER MD package [33] by a default replica exchange procedure.

RS-REMD refinement

The refinement of the structures obtained in the RS-REMD step was performed based on electrostatic energy values obtained from the MM-GBSA calculations as described in our previous work [32]. First, the binding free energy was calculated in AMBER16 [33] program. The trajectory for the first replica (with unmodified LJ parameters) was used. Then, the 100 best frames with the lowest value of electrostatic energy were selected for refinement. These structures were refined by carrying out four steps of MD simulation in the explicit as described in the MD simulations section.

Hydrogen bond analysis

Hydrogen bonds between the SAA protein and GAGs were calculated from the MD trajectories using CPPTRAJ [45], defining a hydrogen bond distance cutoff of 3.0 Å and a hydrogen bond angle cutoff of 135°. The number of reported hydrogen bonds was averaged over the entire MD trajectories and across replicates, and subsequently summarized and visualized using in-house scripts coded in Python 3.8.5 using the numpy 1.19.2 [46], pandas 1.1.3 [47] and matplotlib 3.3.2 [48] libraries.

Data analysis

Statistical analysis and graphical presentation of the obtained data were performed by R-package [49].

Visualization

Each trajectory was visualized in VMD [50] and Pymol [51]. The Pymol was also used for the production of figures.

Funding

This research was funded by National Science Centre of Poland, grant number UMO-2018/30/E/ST4/00037.

Acknowledgements

The computational resources were provided by the Polish Grid Infrastructure (PL-GRID, grant plggagstr2gpu) and the cluster at the Faculty of Chemistry, University of Gdańsk. The molecular docking calculations were performed on the ZIH cluster at TU Dresden (grant p_gag).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mbplus.2021.100080.

Received 21 April 2021;
Accepted 9 July 2021;
Available online 22 July 2021

Keywords:
Serum amyloid A;
Glycosaminoglycans;
Fibril formation;
Molecular docking;
Molecular dynamics;
Hydrogen bonds

† Present address: Helmholtz Institute for pharmaceutical sciences Campus E8 1, 66123 Saarbrücken, Germany.

Abbreviations:
ADAMTS, a disintegrin and metalloproteinase with thrombospondin motifs; AS, aortic valve stenosis; BMP, bone morphogenetic protein; CVD, cardiovascular disease; CKD, chronic kidney disease; CP, C-propeptide; CUB, complement, Uegf, BMP-1; DMD, Duchenne muscular dystrophy; ECM, extracellular matrix; EGF, epidermal growth factor; eGFR, estimated glomerular filtration rate; ELISA, enzyme-linked immunosorbent assay; HDL, high-density lipoprotein; HSC, hepatic stellate cell; HTS, hypertrophic scar; IPF, idiopathic pulmonary fibrosis; LDL, low-density lipoprotein; MI, myocardial infarction; MMP, matrix metalloproteinase; mTLD, mammalian tolloid; mTLL, mammalian tolloid-like; NASH, nonalcoholic steatohepatitis; NTR, netrin; PABPN1, poly(A)-binding protein nuclear 1; OPMD, oculopharyngeal muscular dystrophy; PCP, procollagen C-proteinase; PCPE, procollagen C-proteinase enhancer; PNP, procollagen N-proteinase; SPC, subtilisin proprotein convertase; TIMP, tissue inhibitor of metalloproteinases; TGF-β, transforming growth-factor β; TSPN, thrombospondin-like N-terminal

References

[1]. Magnus, J.H., Stenstad, T., Husby, G., (1994). Proteoglycans, glycosaminoglycans and amyloid deposition. *Baillieres Clin. Rheumatol.*, **8** (3), 575–597.
[2]. Steel, D.M., Sellar, G.C., Uhlar, C.M., Simon, S., DeBeer, F.C., Whitehead, A.S., (1993). A constitutively expressed serum amyloid A protein gene (SAA4) is closely linked to, and shares structural similarities with, an acute-phase serum amyloid A protein gene (SAA2). *Genomics*, **16** (2), 447–454.
[3]. Benditt, E.P., Eriksen, N., (1977). Amyloid protein SAA is associated with high density lipoprotein from human serum. *Proc. Natl. Acad. Sci.*, **74** (9), 4025–4028.
[4]. Uhlar, C.M., Whitehead, A.S., (1999). Serum amyloid A, the major vertebrate acute-phase reactant. *Eur. J. Biochem.*, **265** (2), 501–523.
[5]. Husby, G., Husebekk, A., Skogen, B., Sletten, K., Marhaug, G., Magnus, J., Syversen, V., (1988). Serum amyloid A (SAA)—the precursor of protein AA in secondary amyloidosis, Eicosanoids Apolipoproteins Lipoprotein Part. *Atheroscler.*,, 185–192.
[6]. Egashira, M., Takase, H., Yamamoto, I., Tanaka, M., Saito, H., (2011). Identification of regions responsible for heparin-induced amyloidogenesis of human serum amyloid A using its fragment peptides. *Arch. Biochem. Biophys.*, **511** (1-2), 101–106.
[7]. Husby, G., (1985). Amyloidosis and rheumatoid arthritis. *Clin. Exp. Rheumatol.*, **3**, 173–180.
[8]. Cunnane, G., Whitehead, A.S., (1999). Amyloid precursors and amyloidosis in rheumatoid arthritis. *Best Pract. Res. Clin. Rheumatol.*, **13** (4), 615–628.
[9]. Gollaher, C.J., Bausserman, L.L., (1990). Hepatic catabolism of serum amyloid A during an acute phase response and chronic inflammation. *Proc. Soc. Exp. Biol. Med.*, **194** (3), 245–250.
[10]. McAdam, K.P., Sipe, J.D., (1976). Murine model for human secondary amyloidosis: genetic variability of the

acute-phase serum protein SAA response to endotoxins and casein. *J. Exp. Med.*, **144**, 1121–1127.
[11]. Kluve-Beckerman, B., Manaloor, J., Liepnieks, J.J., (2001). Binding, trafficking and accumulation of serum amyloid A in peritoneal macrophages. *Scand. J. Immunol.*, **53**, 393–400.
[12]. Uhlar, C.M., Burgess, C.J., Sharp, P.M., Whitehead, A.S., (1994). Evolution of the serum amyloid A (SAA) protein superfamily. *Genomics*, **19** (2), 228–235.
[13]. Turnell, W., Sarra, R., Glover, I.D., Baum, J.O., Caspi, D., Baltz, M.L., Pepys, M.B., (1986). Secondary structure prediction of human SAA1. Presumptive identification of calcium and lipid binding sites. *Mol. Biol. Med.*, **3**, 387–407.
[14]. Stevens, F.J., (2004). Hypothetical structure of human serum amyloid A protein. *Amyloid.*, **11** (2), 71–80.
[15]. Liepnieks, J.J., Kluve-Beckerman, B., Benson, M.D., (1995). Characterization of amyloid A protein in human secondary amyloidosis: the predominant deposition of serum amyloid A1. *Biochim. Biophys. Acta BBA-Mol. Basis Dis.*, **1270** (1), 81–86.
[16]. Lu, J., Yu, Y., Zhu, I., Cheng, Y., Sun, P.D., (2014). Structural mechanism of serum amyloid A-mediated inflammatory amyloidosis. *Proc. Natl. Acad. Sci.*, **111** (14), 5189–5194.
[17]. Nakamura, T., (2008). Clinical strategies for amyloid A amyloidosis secondary to rheumatoid arthritis. *Mod. Rheumatol.*, **18** (2), 109–118.
[18]. M. Justin, R. Asherson, C. Pusey, The Kidney in Systemic Autoimmune Diseases, Volume, 2008.
[19]. Röcken, C., Shakespeare, A., (2002). Pathology, diagnosis and pathogenesis of AA amyloidosis. *Virchows Arch.*, **440** (2), 111–122.
[20]. Hirschfield, G.M., Hawkins, P.N., (2003). Amyloidosis: new strategies for treatment. *Int. J. Biochem. Cell Biol.*, **35** (12), 1608–1613.
[21]. Alexandrescu, A.T., (2005). Amyloid accomplices and enforcers. *Protein Sci.*, **14** (1), 1–12.
[22]. Bokarewa, M., Abrahamson, M., Levshin, N., Egesten, A., Grubb, A., Dahlberg, L., Tarkowski, A., (2007). Cystatin C binds serum amyloid A, downregulating its cytokine-generating properties. *J. Rheumatol.*, **34**, 1293–1301.
[23]. Spodzieja, M., Rafalik, M., Szymańska, A., Kołodziejczyk, A.S., Czaplewska, P., (2013). Interaction of serum amyloid A with human cystatin C—assessment of amino acid residues crucial for hCC–SAA formation (part II). *J. Mol. Recognit.*, **26** (9), 415–425.
[24]. Spodzieja, M., Szymańska, A., Kołodziejczyk, A., Prądzińska, M., Maszota, M., Stefanowicz, P., Szewczuk, Z., Grubb, A., Czaplewska, P., (2012). Interaction of serum amyloid A with human cystatin C—identification of binding sites. *J. Mol. Recognit.*, **25** (10), 513–524.
[25]. Maszota, M., Karska, N., Spodzieja, M., Ciarkowski, J., Kołodziejczyk, A.S., Rodziewicz-Motowidło, S., Czaplewska, P., (2015). Structural studies of the C-terminal 19-peptide of serum amyloid A and its Pro→Ala variants interacting with human cystatin C. *J. Mol. Recognit.*, **28** (7), 413–426.
[26]. Zhang, X., Li, J.-P., (2010). Heparan sulfate proteoglycans in amyloidosis. *Prog. Mol. Biol. Transl. Sci.*, **93**, 309–334.
[27]. Bazar, E., Jelinek, R., (2010). Divergent heparin-induced fibrillation pathways of a prion amyloidogenic determinant. *ChemBioChem.*, **11** (14), 1997–2002.
[28]. Gallagher, J.T., (2001). Heparan sulfate: growth control with a restricted sequence menu. *J. Clin. Invest.*, **108**, 357–361.
[29]. Li, J.-P., Galvis, M.L.E., Gong, F., Zhang, X., Zcharia, E., Metzger, S., Vlodavsky, I., Kisilevsky, R., Lindahl, U., (2005). In vivo fragmentation of heparan sulfate by heparanase overexpression renders mice resistant to amyloid protein A amyloidosis. *Proc. Natl. Acad. Sci.*, **102** (18), 6473–6477.
[30]. Noborn, F., Ancsin, J.B., Ubhayasekera, W., Kisilevsky, R., Li, J.-P., (2012). Heparan sulfate dissociates serum amyloid A (SAA) from acute-phase high-density lipoprotein, promoting SAA aggregation. *J. Biol. Chem.*, **287** (30), 25669–25677.
[31]. Takase, H., Tanaka, M., Yamamoto, A., Watanabe, S., Takahashi, S., Nadanaka, S., Kitagawa, H., Yamada, T., Mukai, T., (2016). Structural requirements of glycosaminoglycans for facilitating amyloid fibril formation of human serum amyloid A. *Amyloid.*, **23** (2), 67–75.
[32]. Maszota-Zieleniak, M., Marcisz, M., Kogut, M.M., Siebenmorgen, T., Zacharias, M., Samsonov, S.A., (2021). Evaluation of replica exchange with repulsive scaling approach for docking glycosaminoglycans. *J. Comput. Chem.*, **42** (15), 1040–1053.
[33]. Case, D.A., Betz, R.M., Cerutti, D.S., Cheatham III, T.E., Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T.S., LeGrand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, B., Mermelstein, D., Merz, K. M., Monard, G., Nguyen, H., Nguyen, H.T., Omelyan, I., Onufriev, A., Roe, D.R., Roitberg, A., Sagui, C., Simmerling, C.L., Botello-Smith, W.M., Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu, X., Xiao, L., Kollman, P.A., (2016). AMBER 2016. University of California, San Francisco.
[34]. Samsonov, S.A., Pisabarro, M.T., (2016). Computational analysis of interactions in structurally available protein-glycosaminoglycan complexes. *Glycobiology*, **26**, 850–861.
[35]. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L., Woods, R.J., (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **29**, 622–655.
[36]. A. Pichert, S.A. Samsonov, S. Theisgen, L. Thomas, L. Baumann, J. Schiller, A.G. Beck-Sickinger, D. Huster, M. T. Pisabarro, Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling, Glycobiology. 22 (2012) 134–145.
[37]. M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise., in: Kdd, 1996: pp. 226–231.
[38]. Samsonov, S.A., Gehrcke, J.-P., Pisabarro, M.T., (2014). Flexibility and explicit solvent in molecular-dynamics-based docking of protein-glycosaminoglycan systems. *J. Chem. Inf. Model.*, **54** (2), 582–592. https://doi.org/10.1021/ci4006047.

[39]. Marcisz, M., Huard, B., Lipska, A.G., Samsonov, S.A., (2021). Further analyses of APRIL/APRIL- Receptor/ Glycosaminoglycan interactions by biochemical assays linked to computational studies. *Glycobiology*,, cwab016.

[40]. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79** (2), 926–935.

[41]. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103** (19), 8577–8593.

[42]. Siebenmorgen, T., Zacharias, M., (2019). Evaluation of predicted protein–protein complexes by binding free energy simulations. *J. Chem. Theory Comput.*, **15** (3), 2071–2086.

[43]. Kogut, M.M., Maszota-Zieleniak, M., Marcisz, M., Samsonov, S.A., (2021). Computational insights into the role of calcium ions in protein–glycosaminoglycan systems. *Phys. Chem. Chem. Phys.*, **23** (5), 3519–3530.

[44]. Siebenmorgen, T., Engelhard, M., Zacharias, M., (2020). Prediction of protein-protein complexes using replica exchange with repulsive scaling. *J. Comput. Chem.*, **41** (15), 1436–1447.

[45]. Roe, D.R., Cheatham III, T.E., (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, **9**, 3084–3095.

[46]. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., (2020). Array programming with NumPy. *Nature*, **585** (7825), 357–362.

[47]. W. McKinney, Data structures for statistical computing in python, in: Proc. 9th Python Sci. Conf., Austin, TX, (2010): pp. 51–56.

[48]. Hunter, J.D., (2007). Matplotlib: A 2D graphics environment. *IEEE Ann. Hist. Comput.*, **9**, 90–95.

[49]. R.C. Team, R: A language and environment for statistical computing, (2013).

[50]. Humphrey, W., Dalke, A., Schulten, K., (1996). VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

[51]. L.L.C. Schrödinger, The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC (2017), Google Sch. There No Corresp. Rec. This Ref. (n.d.).