PLOS | ONE

RESEARCH ARTICLE

The Calcium Goes Meow: Effects of Ions and Glycosylation on Fel d 1, the Major Cat Allergen

Rodrigo Ligabue-Braun[1☯], Liana Guimarães Sachett[1☯], Laércio Pol-Fachin[2], Hugo Verli[1]*

1 Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil,
2 Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, PE, Brazil

☯ These authors contributed equally to this work.
* hverli@cbiot.ufrgs.br

CrossMark
click for updates

OPEN ACCESS

Citation: Ligabue-Braun R, Sachett LG, Pol-Fachin L, Verli H (2015) The Calcium Goes Meow: Effects of Ions and Glycosylation on Fel d 1, the Major Cat Allergen. PLoS ONE 10(7): e0132311. doi:10.1371/journal.pone.0132311

Editor: Yaakov Koby Levy, Weizmann Institute of Science, ISRAEL

Received: January 1, 2015

Accepted: June 11, 2015

Published: July 2, 2015

Copyright: © 2015 Ligabue-Braun et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. Large simulation trajectories are available upon request.

Funding: This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior), FAPERGS (Fundação de Amparo àPesquisa do Estado do Rio Grande do Sul), and FACEPE (Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, grant number APQ-0398-1.06/13).

Abstract

The major cat allergen, Fel d 1, is a structurally complex protein with two N-glycosylation sites that may be filled by different glycoforms. In addition, the protein contains three putative Ca^{2+} binding sites. Since the impact of these Fel d 1 structure modifications on the protein dynamics, physiology and pathology are not well established, the present work employed computational biology techniques to tackle these issues. While conformational effects brought upon by glycosylation were identified, potentially involved in cavity volume regulation, our results indicate that only the central Ca^{2+}ion remains coordinated to Fel d 1 in biological solutions, impairing its proposed role in modulating phospholipase A_2 activity. As these results increase our understanding of Fel d 1 structural biology, they may offer new support for understanding its physiological role and impact into cat-promoted allergy.

Introduction

Allergic diseases, especially those involving responses mediated by immunoglobulin E (IgE), are increasing in prevalence and becoming major public health issues [1–3]. This increase has been associated with a Westernized life style and urbanization, suggesting that pets could contribute to this scenario as sources of indoor allergens [1, 4]. The most implicated pets in allergic diseases are cats (*Felis domesticus*), which are present in up to 50% of homes [5]. The sensitization prevalence in adults is in the range of 10–15%, with symptoms varying from rhinoconjunctivitis to potentially life-threatening asthmatic exacerbations [6].

Among the few components of cat dander that can elicit IgE response, the Fel d 1 protein is considered the most potent [6]. Fel d 1, the major cat allergen, is a dimer of all-helical heterodimers, included in the secretoglobin family [7, 8]. Each dimer in the native heterotetramer is N-glycosylated [9], and crystallographic analyses of recombinant Fel d 1 revealed that each dimer has a cavity, possibly involved in the transport of an unknown molecule [7, 8]. For crystallization, a fused version of the protein was produced, in which chains 1 and 2 are linked, making the original tetramer (or dimer of dimers), a simple dimer of the fused protein [8] (Fig 1).

Competing Interests: The authors have declared that no competing interests exist.

The crystallographic study of this recombinant Fel d 1 tetramer also revealed three putative calcium-binding sites, one at the tetramerization interface and one on the lateral of each dimer [8]. The lateral binding sites involve O atoms from Asp46 and Met49 (main chain) and three water molecules within coordination distance from the Ca^{2+}, while the binding site at the interface comprises OD1 atoms from Asn89 (from both chains A and B) and Asp130 (chain B), a carbonyl from Ile125 (chain A) and three water molecules [8] (Fig 1). The non-interfacial binding sites are proposed to act on allergic responses by modulating phospholipase A_2 activity via calcium sequestration, a property related to uteroglobins [8, 10]. The interaction of Fel d 1 and calcium is however subject of debate, since for uteroglobin (a member of the secretogolobin protein family) some authors identified calcium binding, while others identified the opposite [10–13].

The importance of Fel d 1 glycosylation in protein structure and immune response has also been somewhat controversial. While some studies show that different glycosylation patterns do not affect IgE production *in vitro* [9, 14, 15], a most recent study demonstrated that the mannose receptor has an essential role in internalizing Fel d 1. Mannose receptor cysteine rich domain recognizes the carbohydrates in Fel d 1 and *in vivo* assays showed that knockout mice for mannose receptor produced lower levels of immunoglobulins E and G [16].

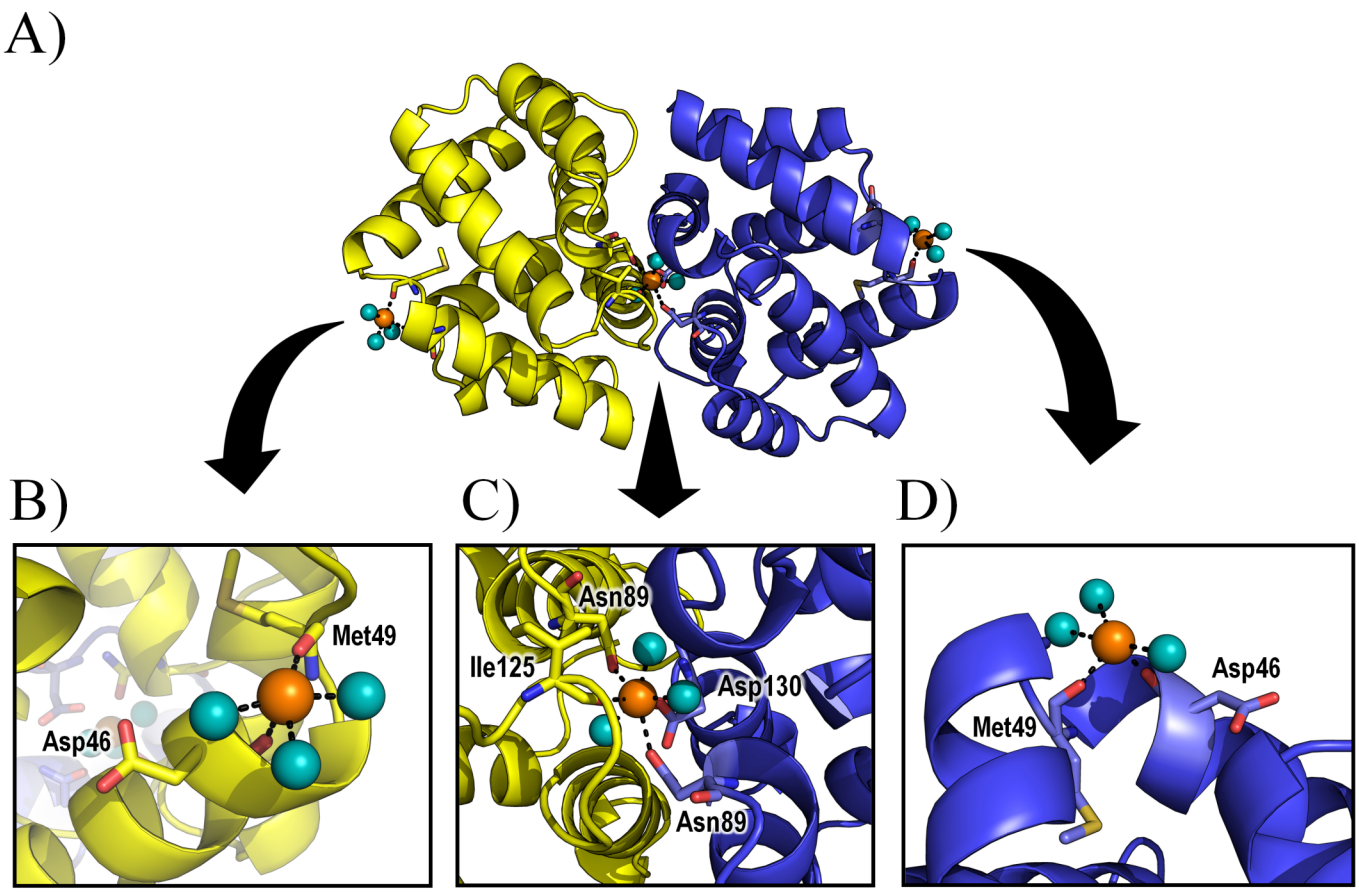

Fig 1. Fel d 1 Ca^{2+} binding sites. (A) Fel d 1 crystallographic structure highlighting the location of calcium ions; (B) Lateral Ca^{2+} binding site A; (C) Interfacial Ca^{2+} binding site; (D) Lateral Ca^{2+} binding site B. Proposed coordination interactions based on the crystallographic structure are shown as dashed lines. Calcium ions are shown as orange spheres, water molecules are shown as cyan spheres, Chain A is shown in yellow, Chain B is shown in blue. (the orientation of the boxes is different from the one in (A) for clarity).

doi:10.1371/journal.pone.0132311.g001

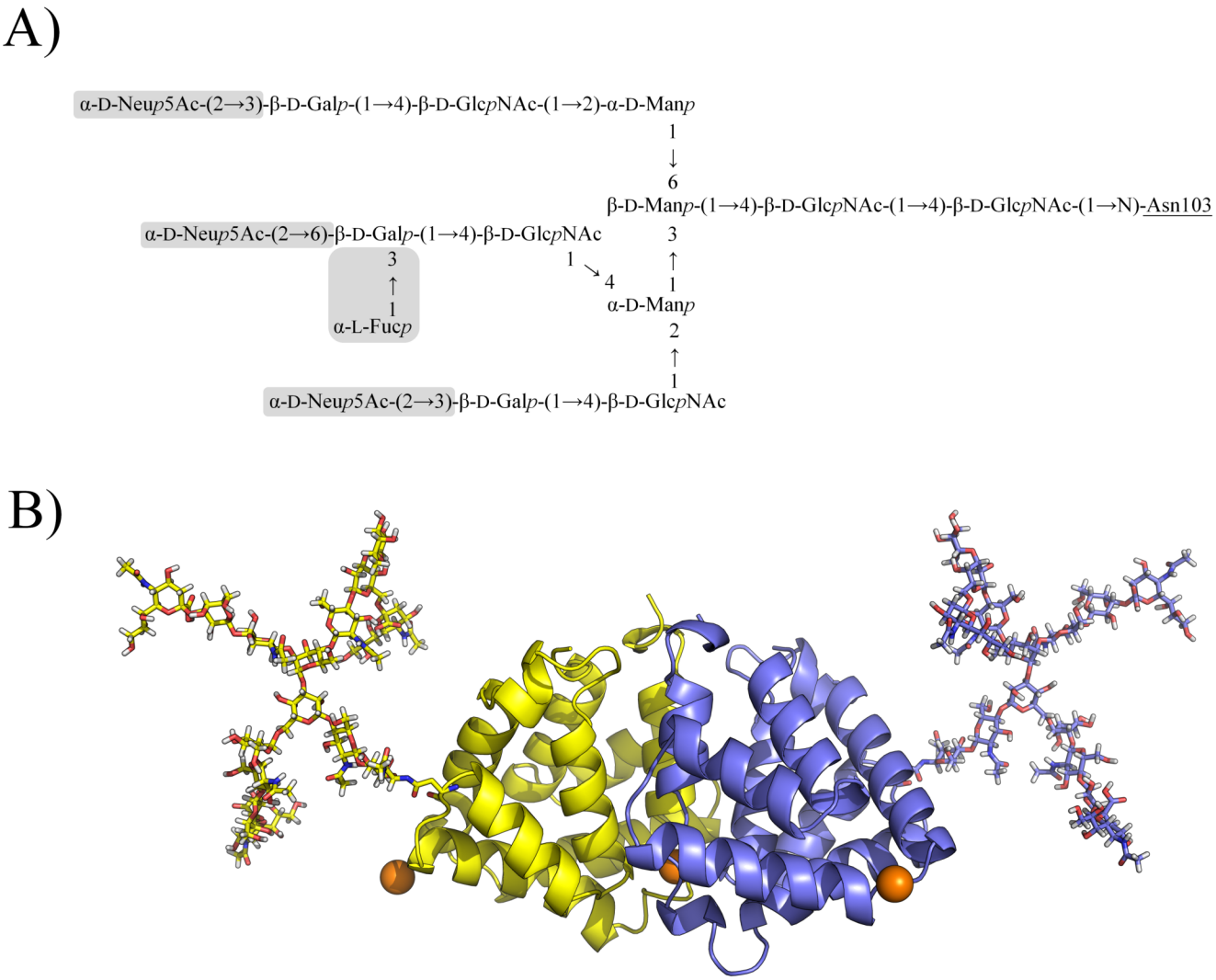

Fig 2. Fel d 1 glycosylation structures. (A) Full (largest) glycosylation structure of Fel d 1, shaded boxes indicate saccharides that are absent in the minimal (smallest) glycosylation structure. (B) Fel d 1 structure with full glycan chains attached. Calcium ions are shown as orange spheres, Chain A is shown in yellow, Chain B is shown in blue.

doi:10.1371/journal.pone.0132311.g002

The glycosylation pattern of Fel d 1 has been determined by mass spectrometry, revealing a series of possible glycoforms bound to the protein [14]. As observed for other glycosylated proteins, there is a minimal structure to which different oligosaccharides may be added, leading to variation in the glycan moiety, which is limited to the largest glycosylation tree identified. For the present work we employed the extremes of the glycosylation structure, analyzing the largest (or "full glycosylation") and the smallest ("minimal glycosylation") structures found by mass spectrometry [14] (Fig 2). Intermediate glycostructures have different numbers of galactose, mannose, and sialic acid residues.

Despite the extensive studies on Fel d 1-induced allergy in humans (as reviewed by [6, 17]), very little information is available regarding its interactions with calcium ions or the effect of glycosylation. Considering that these properties may be important not only for allergy studies but also for understanding the physiological role of this protein in felines, the present work employed computational biology techniques to tackle these issues. Effects brought upon by

glycosylation were identified, while the lateral calcium binding sites proposed by crystallography were not confirmed.

Results

Outline

In this work we performed molecular dynamics (MD) simulations of Fel d 1 using different force fields to evaluate a putative role of parameter differences in describing Ca^{2+} behavior in the protein environment. These molecular mechanics studies were complemented with semi-empirical calculations of the Ca^{2+} interaction with Fel d 1, providing electronic information regarding these ions. In a separate analysis, we inspected the role of two different glycosylation structures on the protein structure and dynamics. These glycoforms are the largest (full glycosylation) and smallest (minimal glycosylation) found on Fel d 1. Protein cavities and ligand-binding possibilities were also studied. The simulated systems and calculations are schematically listed below. Please refer to the Material and Methods section at the end of the article for version information, simulation details, and further references.

By MD we simulated the Fel d 1 dimer of fused dimers under three force fields (AMBER, CHARMM, GROMOS) for 200 ns, in order to guarantee that the Ca^{2+} orientation within Fel d 1 was not due to an specific set of parameters, as well as a simplified structural mockup of the crystallographic environment of Fel d 1 for 25 ns (GROMOS force field) in order to trace potential crystallographic contacts able to lock the Ca^{2+} in position. The role of glycan chains in Fel d 1 dynamics was evaluated through 200 ns simulations using GROMOS force field for two glycosylated systems (Fel d 1 in the largest and smallest glycoforms). After evaluation on the putative Ca^{2+} position dependence on force field parameters, the simulation of the glycoproteins employed GROMOS due to its robust set of parameters for glycoprotein simulations [18, 19], as well as to the lower computational cost associated to united atom force fields in comparison to all atom parameters. All MD simulations were carried out with the GROMACS simulation package. To include direct information on the ions degree of coordination, semi-empirical calculations of the Ca^{2+} binding sites were performed with MOPAC under three scenarios: with implicit solvent, without implicit solvent, and with the inclusion of crystallographic waters.

Implementation and validation of GROMOS 53A6glyc sialic acid parameter

In order to enable Fel d 1 glycoforms to be studied by MD simulations, and considering that parameters for neuraminic (sialic) acid residues were absent in the original GROMOS 53A6glyc implementation, we have constructed a topology for such monosaccharide in the current work (S1 Table). It was built based on ring and hydroxyl parameters for aldohexopyranoses [18] under GROMOS 53A6glyc, as well as N-acetyl and carboxilate parameters previously employed in GROMOS 45A4/53A6glyc force fields [19, 20]. Before employing such parameters for fully glycosylated Fel d 1 MD simulations, we have validated such topology by evaluating isolated sialic acid conformational ring pucker properties, as performed for other monosaccharides [18]. Thus, while metadynamics calculations indicated ${}^{2}C_{5}$ as the preferential ring pucker conformation for sialic acid, unbiased 1000 ns MD simulations confirmed such behavior, in accordance with previous observations (S1 Fig).

Table 1. Distances and bond orders measured by semi-empirical calculation.

Lateral Ca^{2+} (A)		With implicit solvent		Without implicit solvent		With crystallographic waters	
	O atom	Distance (Å)	Bond order	Distance (Å)	Bond order	Distance (Å)	Bond order
Asp46A	Side chain	-	-	2.2	0.274	-	-
	Main chain	2.2	0.234	-	-	-	-
Met49A	Main chain	2.3	0.291	2.4	0.102†	2.4	0.094†
Thr50A	Side chain	2.3	0.230	-	-	-	-
Glu51A	Side chain	-	-	2.3	0.313	2.3	0.279
	Side chain	-	-	-	-	2.3	0.243
	Main chain	-	-	2.3	0.268	-	-
H_2O	Water 1	NA	NA	NA	NA	2.4	0.222
	Water 2	NA	NA	NA	NA	2.4	0.207
	Water 3	NA	NA	NA	NA	2.4	0.179†
Central Ca^{2+}		**With implicit solvent**		**Without implicit solvent**		**With crystallographic waters**	
	O atom	**Distance (Å)**	**Bond order**	**Distance (Å)**	**Bond order**	**Distance (Å)**	**Bond order**
Asn89A	Side chain	2.2	0.255	2.3	0.238	2.3	0.208
Ile125A	Main chain	2.3	0.187†	2.5‡	0.086†	-	-
Asn89B	Side chain	2.3	0.195†	2.3	0.219	2.3	0.227
Ile125B	Main chain	-	-	2.3	0.182†	-	-
Asp130B	Side chain	2.3	0.248	2.3	0.268	2.3	0.278
	Side chain	2.3	0.257	2.3	0.268	2.3	0.227
H_2O	Water 7	NA	NA	NA	NA	2.4	0.197†
	Water 8	NA	NA	NA	NA	2.5‡	0.122†'
Lateral Ca^{2+} (B)		**With implicit solvent**		**Without implicit solvent**		**With crystallographic waters**	
	O atom	**Distance (Å)**	**Bond order**	**Distance (Å)**	**Bond order**	**Distance (Å)**	**Bond order**
Asp46B	Side chain	-	-	2.3	0.192†	2.2	0.226
	Main chain	2.3	0.223	2.3	0.238	-	-
Ala47B	Main chain	2.3	0.211	-	-	-	-
Met49B	Main chain	2.3	0.264	2.4	0.148†	2.3	0.157†
Thr50B	Main chain	-	-	2.5‡	0.102†	-	-
Glu51B	Side chain	2.3	0.250	2.3	0.237	-	-
	Side chain	-	-	2.3	0.296	-	-
H_2O	Water 4	NA	NA	NA	NA	2.3	0.250
	Water 5	NA	NA	NA	NA	2.3	0.301
	Water 6	NA	NA	NA	NA	2.3	0.261

†Bond orders of less than 0.2 are indicative of "no bond" [21]

‡Distances greater than reference distances for calcium coordination in proteins (2.35 Å-2.45 Å) [22].

doi:10.1371/journal.pone.0132311.t001

Analyzes of putative calcium binding sites

The three calcium binding sites proposed for Fel d 1 (Fig 1) were analyzed by semi-empirical calculations and molecular dynamics (MD) simulations. Three conditions were calculated semi empirically: with implicit solvent, without implicit solvent, and including crystallographic waters. The distance and bond orders obtained from these calculations (Table 1) indicate that the central Ca^{2+} is stable even if a little dislocated in its position, while the lateral ions are not bound and move from their proposed locations. MD simulations carried out in three force fields (AMBER, CHARMM, GROMOS) indicate the same dislodgment of the lateral Ca^{2+} ions (Table 2 and S2 Fig). To inspect if crystal packing effects would stabilize these ions, the

Table 2. Average distances measured for molecular dynamics simulations under different force fields (Å).

	*	AMBER	CHARMM	GROMOS	Crystal	Expected*
Lateral Ca^{2+}(A)	Asp46A	37.0±1.0	11.1±1.5	33.7±13.3	13.5±14.8	2.4
	Met49A	36.8±8.9	10.0±1.3	37.3±13.7	14.2±15.4	2.3
Central Ca^{2+}	Asn89A	2.7±0.1	2.3±0.1	2.5±0.3	2.4±0.1	2.4
	Ile125A	2.7±0.1	4.3±0.7	7.2±0.8	4.2±1.0	2.2
	Asn89B	2.7±0.3	5.3±1.0	4.9±0.6	4.2±0.8	2.7
	Asp130B	2.6±0.1	2.3±0.3	2.7±0.2	5.2±0.2	2.2
Lateral Ca^{2+}(B)	Asp46B	4.1±0.9	16.1±1.8	27.3±11.2	7.6±4.5	2.4
	Met49B	5.3±0.6	13.2±2.1	27.0±11.0	6.7±3.7	2.3

*residues making contact with calcium ions and their expected distances were taken from PDB ID 2EJN [8].

doi:10.1371/journal.pone.0132311.t002

neighboring asymmetric units from the crystallographic Fel d 1 were regenerated and simulated (Fig 3). There was no stabilization detected in this condition as well (Table 2).

Calcium effects on protein conformation

Considering that the stable binding of the central Ca^{2+} could be related to stabilization of the protein dimer, the interaction energy of the complex was also measured for the performed MD simulations (Table 3 and S3 Fig). There was a slight decrease in interaction energy (more favored interaction) in the calcium-free conditions when compared to calcium-bound conditions.

The MD simulated systems were generally stable (Fig 4) and had no difference regarding their stability with or without Ca^{2+} for simulations under AMBER and CHARMM force fields. In the simulations performed under GROMOS force field, however, there was a slight difference between metal-bound and metal free conditions combined with noticeable difference in stability in comparison with the other force fields. (Fig 4). Regarding local flexibility, we observed little differences between calcium-free and calcium-bound system (Fig 5).

Glycosylation effects

The minimally and fully glycosylated systems behaved differently. While the fully glycosylated protein had a conformational behavior equivalent to the unglycosylated forms in the same force field (GROMOS), the protein with minimal glycosylation had a greater divergence in terms of conformation (Fig 6).

Moreover, glycosylation seemed to increase flexibility around residue 75, in a loop region of chain A (but not on the equivalent region of chain B) (Fig 6). The minimal saccharide structure also caused an increase in flexibility in the N-terminus region of both chains. The secondary structure content was similar in all simulations, with a noted decrease in helical content for the systems with no Ca^{2+} and with the minimally glycosylated protein (S4 Fig). This decrease was associated with partial unfolding of helices I and II (first and second helices from the N-terminus), reflected in the increased flexibility observed in RMSF analyses (Fig 6). The analysis of secondary structure along the simulated time also indicates structural stability for (at least) the last 20 ns of all simulation with the GROMOS force field. It is interesting to note that the full glycosylation seemed to stabilize the Chain B structure, as revealed by RMSD analyses of each of the protein segments (half-monomers in the fused protein) (S5 Fig). This protein-segment analysis also confirmed the early stabilization of all chains in the Ca^{2+} system.

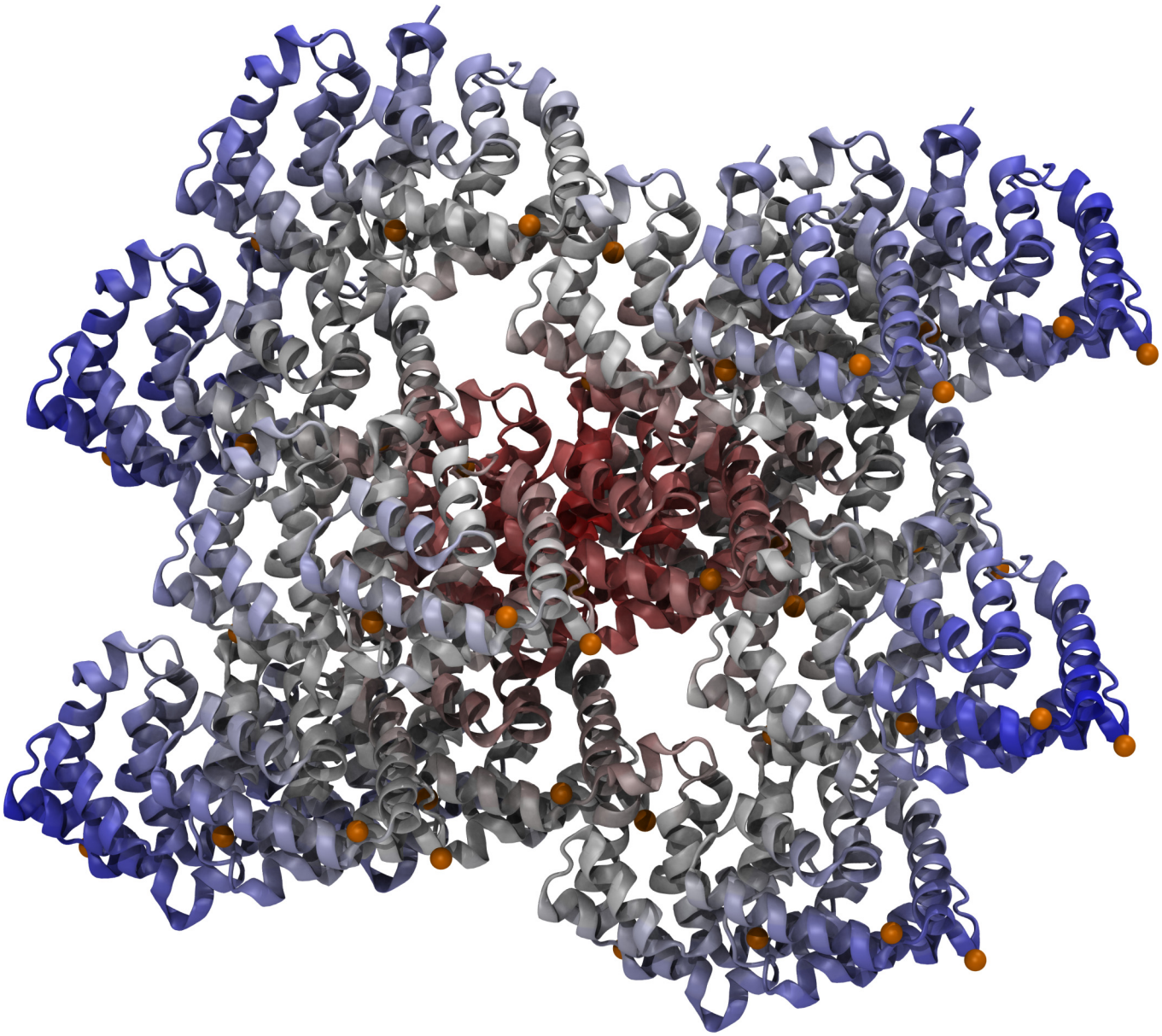

Fig 3. Fel d 1 in its crystallographic environment. Red depicts the central protein, while the white-to-blue gradient indicates surrounding protein molecules. Calcium ions are shown as orange spheres.

doi:10.1371/journal.pone.0132311.g003

Table 3. Interaction energies measured between Fel d 1 monomers during molecular dynamics simulations under different force fields (kJ/mol).

	AMBER	CHARMM	GROMOS
With Ca^{2+} ions	-718.07±65.49	-599.44±69.71	-570.57±67.84
Without Ca^{2+}ions	-728.87±69.52	-772.41±67.47	-617.71±87.07

doi:10.1371/journal.pone.0132311.t003

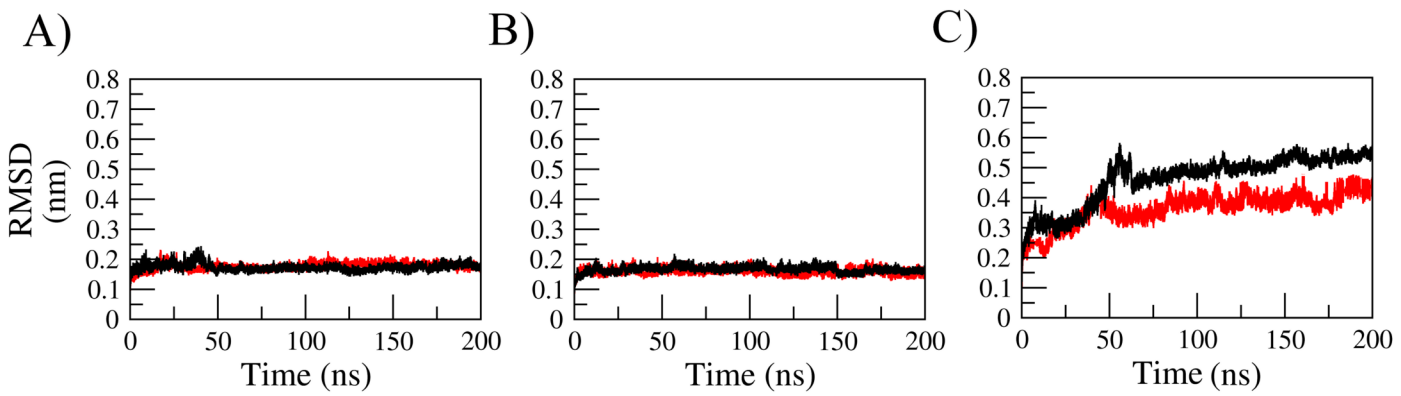

Fig 4. All-atom root mean square deviation (RMSD) measured for different force field simulations of Fel d 1. (A) AMBER, (B) CHARMM, (C) GROMOS. Color coding: Simulations without Ca^{2+}ions (black), simulations with Ca^{2+} ions (red).

doi:10.1371/journal.pone.0132311.g004

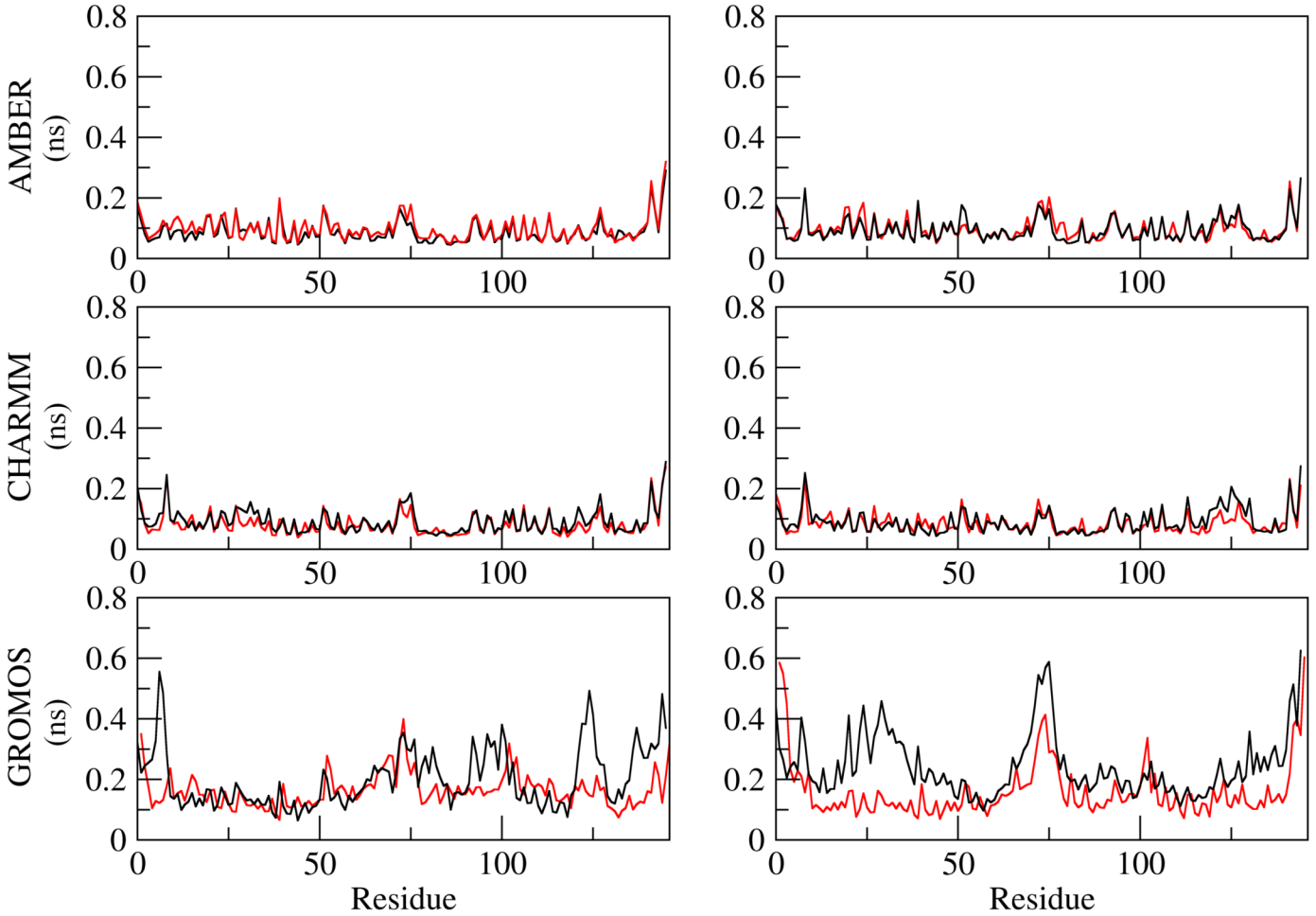

Fig 5. Root mean square fluctuations (RMSF) measured for different force field simulations of Fel d 1. Color coding: Simulations without Ca^{+2} ions (black), simulations with Ca^{+2} ions (red).

doi:10.1371/journal.pone.0132311.g005

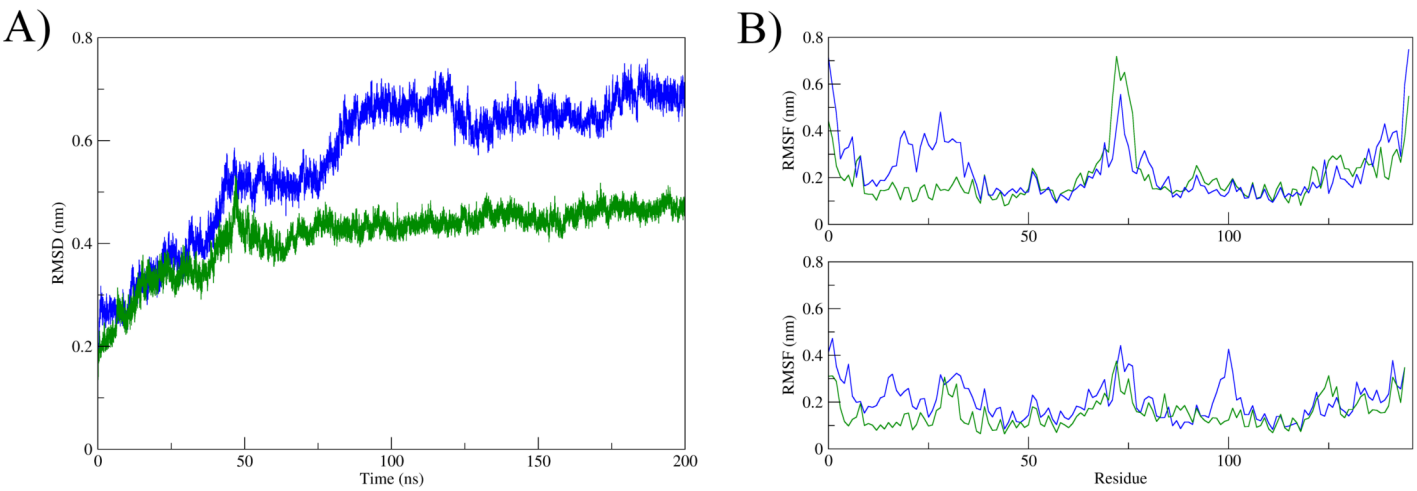

Fig 6. Structural flexibility of glycosylated Fel d 1. (A) All-atom root mean square deviation (RMSD) and (B) Root mean square fluctuations (RMSF) measured for the simulations with minimal (blue) and full glycosylations (green).

doi:10.1371/journal.pone.0132311.g006

Cavity assessment and ligand dockings

The two cavities detected in Fel d 1, which had different volumes in the crystallographic structure [8] were measured for all systems simulated with GROMOS force field. Their volume seems to oscillate periodically along the simulated time (S6 Fig). This is especially clear by analyzing the three main structural clusters found in each simulation (Table 4). The cavities detected in these clusters confirm clearly different cavity sizes for each Fel d 1 chain.

We studied Fel d 1 cavity binding capacity using molecular docking, revealing that cavities of both Fel d 1 chains (A and B) are able to interact with the ligands PCB (an steroid analog), progesterone and testosterone (Fig 7). In chain A, residues Val10, Leu61 and Phe84 presented hydrophobic interactions with all three ligands, and Asp130 was responsible for one hydrogen bond. In chain B, for all ligands, residues Phe13 and Phe84 were the ones making the hydrophobic contacts and Tyr21 was the residue responsible for one hydrogen bond.

Dynamical network and motion analyzes

Inspecting internal motion correlations for the simulated Fel d 1 systems with GROMOS force field, we were able to detect two major trends. First, there is conservation of residues acting as

Table 4. Cavity volume measured for the three main structural clusters from MD simulations under different conditions.

		Volume (Å³)		
		Cluster 1	Cluster 2	Cluster 3
With Ca^{2+}	Chain A	4357	4133	4174
	Chain B	5145	4004	3783
Without Ca^{2+}	Chain A	4397	4590	3208
	Chain B	4035	4383	4454
Glycosylated (minimal)	Chain A	3133	3537	3634
	Chain B	2907	4367	4217
Glycosylated (full)	Chain A	3462	4032	3344
	Chain B	4607	4031	3549

doi:10.1371/journal.pone.0132311.t004

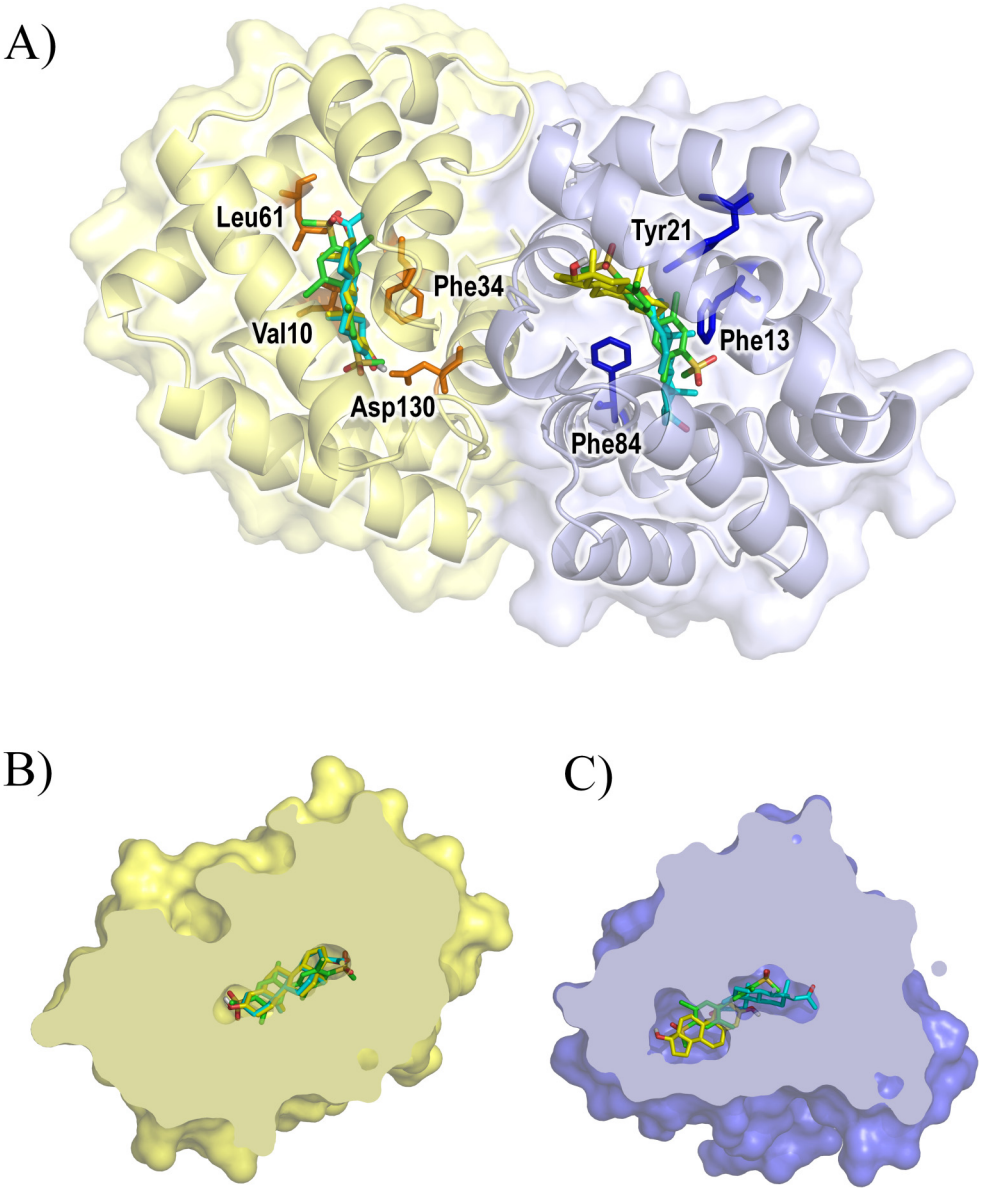

Fig 7. Docking of different ligands to Fel d 1. (A) Docking results overview, highlighting residues involved in binding interactions; Cut-through views of (B) chain A and (C) chain B, highlighting differences in cavity sizes. Cut-through views are rotated in respect to (A) for clarity; chain A shown in yellow and chain B shown in blue, testosterone shown in yellow, progesterone shown in cyan, PCB shown in green.

doi:10.1371/journal.pone.0132311.g007

critical nodes for motion transmission between chains A and B at the interaction interface, regardless of the presence of calcium or glycosylation (S2 Table). Second, there seems to be a hinge-like motion transfer between each original chain (i.e. each half of Fel d 1 crystallographic monomer) in the simulated systems (Fig 8). The observed critical nodes, also conserved, seem to act as unconventional hinges, and an oscillatory motion between these 'flaps' can be observed, especially for Chain A (S7 Fig), although the characteristic disulfide bridges of Fel d 1 may restrict this oscillation.

The motions of each protein chain in respect to each other and within each chain (Fig 9) reveal no clear pattern. The calcium-free system has diffuse motions, whereas the other systems

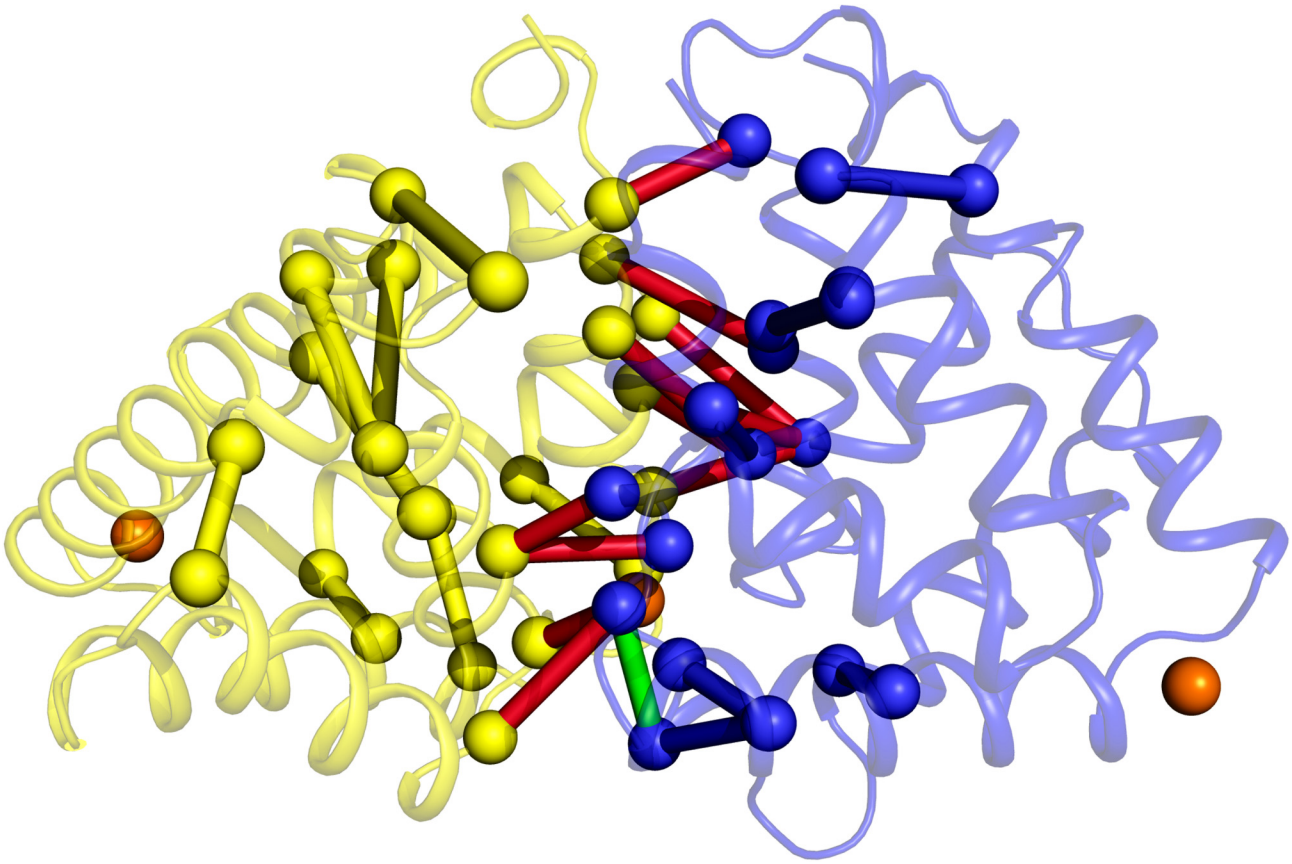

Fig 8. Overview of essential node pairs for motion transmission in Fel d 1 simulations. 'Hinge' pairs in each chain are colored yellow for Chain A and blue for Chain B, while interface pairs are colored red. A connection between interface and intrachain motions (detected in the system simulated with Ca^{2+} and no glycosylation) is shown in green. Calcium ions shown in orange.

doi:10.1371/journal.pone.0132311.g008

seem to have more directed movements. This is especially highlighted in the glycosylated systems, in which the chains of Fel d 1 (monomers in the fused protein) seem to move in the opposite direction of each other. Helices I and II show pronounced mobility (in chain A).

Discussion

Considering the debate on whether Ca^{2+} binds to Fel d 1 or not, for this work we employed different calculation approaches to inspect such interactions. Our results indicate that the central ion is bound to the protein, while the lateral ions are free. This may hinder the proposed action of Fel d 1 in allergy through modulation of phospholipase A_2 activity via calcium sequestration [8, 10]. This property was proposed taking uteroglobin as a model, but even for this protein there is controversy regarding the putative binding of the ion [11–13].

Supporting the concept that Ca^{2+} is not bound to the lateral sites of Fel d 1 is the physiological concentration of calcium in cats. Among the synthesis sites of Fel d 1 (which include perianal, sebaceous, and salivary glands), the calcium concentration in submandibular salivary glands (20 mmol/kg) is considered high for soft tissues in mammals (twice as much as normal human lung tissue) [23, 24], but the calcium concentration in the crystallization medium was 25-fold greater than that [8]. This would lead to unspecific binding of Ca^{2+}, and a fourth ion in the final crystallized structure is considered one of these cases [8].

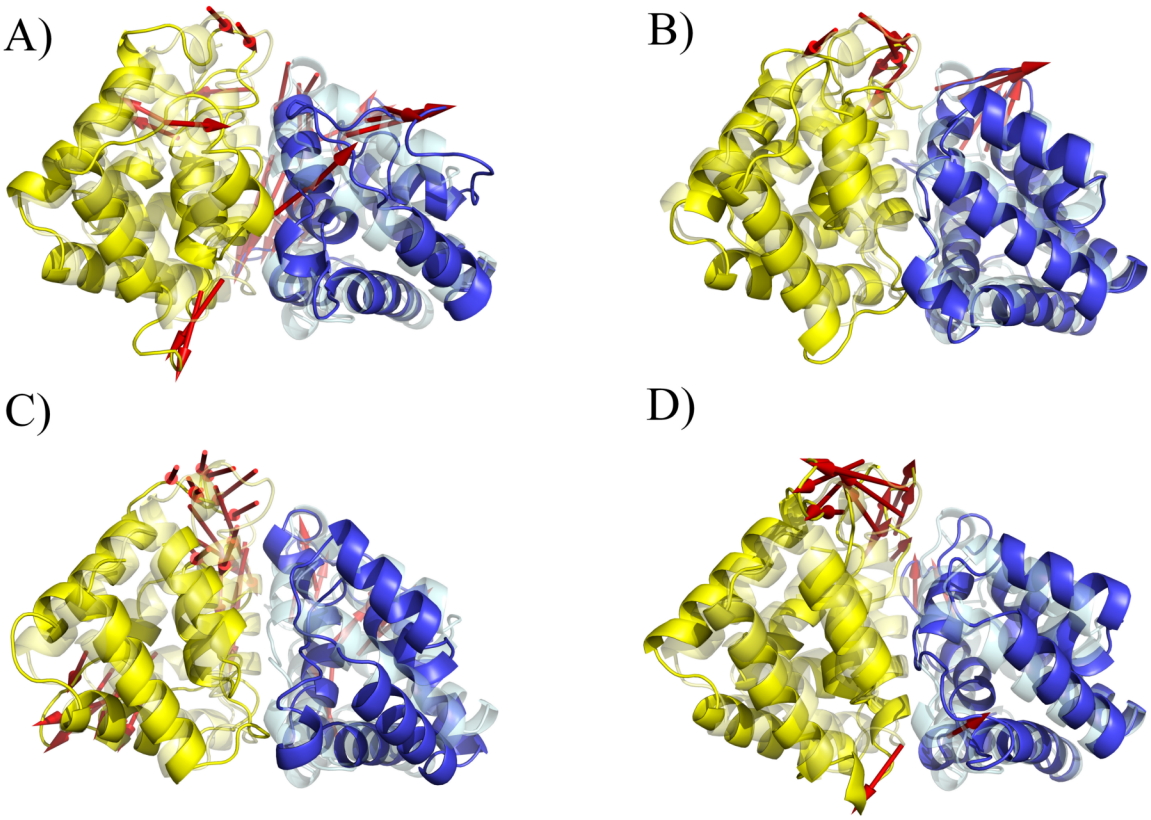

Fig 9. Overview of conformational changes observed for the different GROMOS simulated systems. The initial structures are shown as transparent, with arrows pointing the major motions observed between the initial and final stages of the MD simulations. (A) Calcium-free system; (B) Calcium-bound system; (C) Minimally glycosylated system; (D) Fully glycosylated system. Chain A shown in yellow and chain B shown in blue.

doi:10.1371/journal.pone.0132311.g009

To avoid possible parameter bias in simulating this protein-metal complex, we used three different force fields. There were differences in overall flexibility observed between all-atom force fields (AMBER and CHARMM) and united-atoms force field (GROMOS) in this work. These differences, which may be striking at first, are probably due to faster convergence and enhanced conformational sampling in the united-atoms force field [25]. The alternative explanation to GROMOS apparently not reaching convergence is that AMBER and CHARMM simulations remain stuck in energetic wells in a more rugged energy surface than in GROMOS simulations. However, the concept of convergence involves evaluating if what has been measured is enough to discard the occurrence of anything important which remains unmeasured [26]. There is no guarantee that a stable system (based on the generally used measures, such as RMSD) will remain stable if simulated for some extra nanoseconds. Besides that, convergence varies with every property analyzed and they do not correlate to each other [27]. It has been shown that evaluating convergence is a subjective matter, and that the evaluation of structural features such as RMSD for this purpose is very unreliable [28]. In the context of our work, they show that different levels of detail in the force field parametrization have no difference in describing the Ca^{2+} behavior, allowing us to compare the GROMOS results obtained for the calcium ion studies with the glycosylation studies with the same force-field, for which there is a wealth of saccharide parameters [18,19]. Besides the combination of different simulations, including the crystal environment, the semi-empirical calculations improved the evaluation of such important aspect of Fel d 1 structural biology by exploring the electronic component of

Ca^{2+} ions coordination in this complex. It is possible that Ca2+ ions can bind to different, new sites on Fel d 1, based on the behavior of the GROMOS simulations. This observation reinforces that, apart from the central ion binding site, the lateral binding sites are not stable and are not supported by different inspections.

Another highlighted feature of the major cat allergen is its saccharide content. The active form of Fel d 1 has two N-glycosylation sites (Fig 2), and the attached glycan chains are highly variable [14]. For this work we inspected the largest and the smallest carbohydrate chains observed linked to the protein. Different effects were observed in each case. While the conformational behavior of the fully glycosylated protein did not differ from its unglycosylated counterpart in general, the minimally glycosylated protein behaved differently. There was an increase in flexibility, due to partial unfolding of two helices of one of the Fel subunits. The fully glycosylated protein, however, presented a reduction in its cavity size, also for one of its subunits. This difference could be due to sampling limitations in our simulations, but an extension of differences between subunits detected in the crystallographic structure [8] are not discarded.

A previous study of circular dichroism on Fel d 1observed decrease in its helical content due to removal of glycan portions [9]. In our work, we were able to pinpoint the helices that are more likely to unfold, corresponding to this previous observation. The reduction in cavity volume observed for one subunit of the fully glycosylated protein may hint to Fel d 1glycans acting on molecular transport modulation. It has been shown previously that selective saccharide removal from glycans in human plasmatic transport proteins caused cavity alterations, affecting its binding do hormones and drugs [29].

The putative content of Fel d 1 cavities has not been identified yet [8]. A similar protein, found in the slow loris (*Nycticebus coucang*, a prosimian), is capable of causing anaphylactic shock in humans [30]. These animals are thought to use these proteins defensively, but mainly as a communication tool. Different hydrophobic compounds were identified in association with this molecule, and it is proposed that the protein structure would act as a box, containing and delivering chemical messages [31]. It has been observed that cat females secrete less Fel d 1 than males, and that males that are handling-avoidant (less frightened and more aggressive) have greater release of the allergen than handling-friendly males [32]. This may be an indication of Fel d 1 acting in information transmission in cats.

Considering the synthesis of the protein in perianal glands and its spreading by grooming involving saliva, it is tempting to propose that steroid-like molecules from the perianal glands may be entrapped in Fel d 1 by changes in pH (or other feature) brought about by interaction with saliva. Based on the crystal structure of uteroglobin bound to an steroid analog [33], we performed dockings and confirmed that the Fel d 1 cavities observed in simulation would accommodate progesterone (312 $Å^3$), testosterone (399 $Å^3$), and the steroid analog, PCB (335.6 $Å^3$). A communication role has been proposed for the loris brachial gland protein, (a Fel d 1 fold analog) [27], and has also been proposed for Fel d 1 in the 1990s [34].

Fel d 1 is a dimer of heterodimers in its native form, and was expressed as a dimer of fused monomers for crystallization [8]. For this reason, each monomer in the crystal (originally a dimer of chains α and β) is called a chain (A and B). Despite being a homodimer, the interacting interface of the oligomer is not symmetric. Asymmetry is important for the proper recognition of each monomer (α and β dimer) in terms of immunological response [15]. The relevance of asymmetry for the biological action of the Fel d 1 dimer (or dimer of dimers in its native form) has not been investigated so far, but our results point to differential effects of glycosylation and cavity behavior in each of the asymmetrically positioned Fel d 1 subunits.

In summary, in this work we were able to apply theoretical chemical tools in the study of Fel d 1 molecular behavior and its interaction with calcium ions and glycosylation. Our results not

only corroborate previous observations, but also widen the breadth of action of the major cat allergen. We hope to have highlighted some advantages of the computational structure biology applied to the study of Fel d 1. Future investigations, focused on the influence of variable glycosylation structures on the protein, site-directed mutagenesis of metal-binding and motion-transmitting residues, and especially those focused on the functions carried out by Fel d 1 in feline physiology, may benefit from the observations presented here. For instance, the calcium-binding properties of the protein are less prominent than previously proposed, reducing the ion structural role. The protein binding cavity is compatible with previously proposed ligands, providing structural basis for these proposals. The impact of glycosylation on Fel d 1 seems limited, also reducing its importance for cavity volume control. The models presented here provide structural and rational basis to guide future studies with ligand binding and site-directed mutagenesis. We hope this kind of work may highlight the need for cooperation between laboratory experimentalists and computer modelers with the joint aim of studying structural biology in atomistic scale in both time and space dimensions.

Materials and Methods

Simulated systems

We retrieved the structure for the tetrameric form of Fel d 1 from the RCSB Protein Data Bank under PDB ID 2EJN [8], and used it to prepare nine Fel d 1 systems for MD simulations under different force fields and conditions. The crystal-derived Fel d 1 was simulated with and without crystallographic Ca^{2+} ions employing three force fields (AMBER99SB-ILDN [35], CHARMM27 [36], and GROMOS53a6 [20]). Two glycosylated forms of the protein were also simulated, as derived from MALDI-MS analyses of the native Fel d 1, basically corresponding to a desialylated (minimal or smallest) and a sialylated (full or largest) oligosaccharide structures [14]. We also regenerated the crystalline environment surrounding the tetrameric protein using symmetry tools from PyMol (Schrödinger, LLC) and symmetry group information from the Fel d 1 crystal. This procedure added ten other tetramers to this ninth simulated system.

Glycoprotein construction

Each dimer of Fel d 1 has one N-glycosylation site, at Asn103 [8], in a total of two sites for the biological tetramer [14]. We used the full length [(-Asn-) $GlcNAc_2Man_3GlcNAc_3Gal_3$ (Fuc_1)-$NeuAc_3$] and truncated [(-Asn-) $GlcNAc_2Man_3GlcNAc_3Gal_2$] forms of the glycosidic chains (named 'full' and 'minimal' for reference) previously identified [14] to build the glycosylated proteins. This construction was made via Glycosciences.de modelling tools [37, 38]. The glycosidic linkages composing such glycans had their geometries adjusted to each of their relative abundance as isolated disaccharides in water, thus determined as their main conformational states [39, 40]. Such procedure was successfully applied in previous glycoprotein studies [41, 42]. The oligosaccharide topologies were described under GROMOS 53A6glyc parameters set [18, 19] and the protein moieties under the standard GROMOS96 53A6 force field.

Molecular dynamics studies

We submitted the Fel d 1 systems to MD simulations with the GROMACS 4.5 suite [42] for 200 ns each, with the exception of crystal-environment mimic, which we simulated for 25 ns. The systems were solvated in rectangular boxes using periodic boundary conditions, and subjected to different force fields and water models: GROMOS96 53a6 force field [20] with SPC water models [43], CHARMM27 [32] with TIP3P [44], and AMBER99SB-ILDN [31] also with TIP3P. The Lincs method [45] was applied to constrain covalent bond lengths, allowing an

integration step of 2 fs after an initial energy minimization using Steepest Descents algorithm. Electrostatic interactions were calculated with Particle Mesh Ewald method [46], in which short-range interactions cutoff values were set to 0.9 nm. Temperature and pressure were kept constant by coupling proteins, ions, and solvent to external temperature and pressure baths with coupling constants of $\tau = 0.1$ and 0.5 ps (Berendsen barostat, velocity rescaling thermostat [47, 48], respectively). The dielectric constant was treated as $\varepsilon = 1$, and the reference temperature was adjusted to 300 K to reflect the temperature faced by the protein in cat hair and dandruff. The systems were slowly heated from 50 to 300 K, in steps of 5ps, each step increasing the reference temperature by 50 K, allowing a progressive equilibration of the molecular system. The simulations were performed with no restraint, with a reference value of 3.5 Å between heavy atoms for a hydrogen bond, and a cutoff angle of 30° between hydrogen-donor-acceptor [41]. The so called "crystal-environment mimic" was based on reproducing the crystallographic Fel d 1 neighboring asymmetric units only, not their entire crystallographic chemical environment. The system sizes for all simulations are listed on S3 Table.

Cavities were measured with Mole2 [49, 50]. Essential nodes for motion transmission were inspected as described previously for tRNAs [51, 52]. Ligand volumes were calculated with UCSF Chimera 1.7 [53]. Docking studies of Fel d 1 chains A and B to ligands PCB, progesterone and testosterone were each performed with Autodock 4.2 [54] and lowest energy conformations for each system were analyzed with Poseview [55] and Ligplot [56]. All other analyses were carried out with dedicated tools from the Gromacs suite.

Semiempirical calculations

In order to add information on metal coordination to molecular mechanics calculations we submitted Fel d 1 with all four crystallographic Ca^{2+} ions to semiempirical quantum chemistry calculations using MOPAC [57] with the PM6 Hamiltonian [58] and MOZYME linear scaling method [59]. The following steps, based on a previous study of metal coordination [60], were employed on the calculations: 1) hydrogen atoms optimization to a gradient lower than 5 kJ/mol/Å; 2) global optimization with gradient convergence criterium of 10 kJ/mol/Åand cutoff = 6; 3) global optimization with gradient convergence criterium of 5 kJ/mol/Åand cutoff = 9. We carried out calculations in the absence of crystallographic water molecules, with the COSMO continuum solvation model [61], and with all the crystallographic water molecules within a 5 Å radius from the Ca^{2+} ions. We used the obtained minimum energy conformation was employed to evaluate the bond orders and distances to estimate the possible coordination between amino acid residues and Ca^{2+} ions.

Supporting Information

S1 Fig. Validation of GROMOS 53A6glyc parameters implemented for sialic acid.
(PDF)

S2 Fig. Distances measured for molecular dynamics simulations of Fel d 1 under different force fields. Please refer to Table 1 for details.
(PDF)

S3 Fig. Interaction energies measured between Fel d 1 monomers during molecular dynamics simulations under different force fields. Color coding: Simulations without Ca^{2+} ions (black), simulations with Ca^{2+} ions (red).
(PDF)

S4 Fig. Secondary structure content for the Fel d 1 systems simulated with GROMOS. (A) Fel d 1 without Ca^{2+}; (B) Fel d 1 with Ca^{2+}; (C) Fel d 1 with minimal glycosylation; (D) Fel d 1

with full glycosylation.
(PDF)

S5 Fig. All-atom root mean square deviation (RMSD) measured for Fel d 1 segments corresponding to the original chain regions (fused in the crystallized protein). (A) Calcium-free system; (B) Calcium-bound system; (C) Minimally glycosylated system; (D) Fully glycosylated system. Region A_1 in black, region A_2 in red, region B_1 in blue, region B_2 in green.
(PDF)

S6 Fig. Cavity volume overview (calcium-free system as an example). (A) initial cavity volume, (B) final cavity volume (cavity volumes shown as green surfaces); (C) cavity volume measurements for each Fel d 1 chain in four simulations.
(PDF)

S7 Fig. Structural 'oscillation' observed for (A) chain A, and (B) chain B of Fel d 1 in four different simulations. The distance between each original chain in the native Fel d 1 (half chains of the crystallographic protein) is shown. The oscillation may indicate an opening-closing behavior. (calcium-free system in black, calcium-bound system in red, minimally glycosylated system in blue, fully glycosylated system in green; distances measured for the centers of mass in each structure).
(PDF)

S1 Table. Parameters used in the GROMOS 53A6glyc force field for sialic acid, in a GROMACS-compatible format. These encompass atom types, atomic partial charges, charge-group definition, bond stretching, bond-angle bending, improper dihedral deformation and torsional potential dihedral.
(PDF)

S2 Table. Essential node pairs detected in four Fel d 1 simulated systems. Nodes that appear more than once are grouped by color. Interactions detected between pairs by PDBsum for the crystallographic structure are in parentheses.
(PDF)

S3 Table. Sizes of simulated systems.
(PDF)

Acknowledgments

The authors wish to thank the Centro Nacional de Supercomputação at Universidade Federal do Rio Grande do Sul for the assistance and access to the supercomputer, and the anonymous reviewers whose input greatly improved the final version of the manuscript. This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior), FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul), and FACEPE (Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, grant number APQ-0398-1.06/13).

Author Contributions

Conceived and designed the experiments: RLB LGS LPF HV. Performed the experiments: RLB LGS LPF. Analyzed the data: RLB LGS. Wrote the paper: RLB LGS LPF HV.

References

1. Dharmage SC, Lodge CL, Matheson MC, Campbell B, Lowe AJ. (2012) Exposure to cats: update on risks for sensitization and allergic diseases. Curr Allergy Asthma Rep 12:413–423. doi: 10.1007/s11882-012-0288-x PMID: 22878928
2. Asher MI, Montefort S, Bjorksten B, Lai CK, Strachan DP, Weiland SK, et al. (2006) Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multi country cross-sectional surveys. Lancet 368: 733–743. PMID: 16935684
3. Kay AB. (2000) Overview of 'Allergy and allergic diseases: with a view to the future'. Br Med Bull 56:843–864. PMID: 11359624
4. Carrer P, Maroni M, Alcini D, Cavallo D. (2001) Allergens in indoor air: environmental assessment and health effects. Sci Total Environ 270: 33–42. PMID: 11327396
5. Svanes C, Heinrich J, Jarvis D, Chinn S, Omenaas E, Gulsvik A, et al. (2003) Pet-keeping in childhood and adult asthma and hay fever: European community respiratory health survey. J Allergy Clin Immunol 112: 289–300. PMID: 12897734
6. Grönlund H, Saarne T, Gafvelin G, van Hage M. (2010) The major cat allergen, Fel d 1, in diagnosis and therapy. Int Arch Allergy Immunol 151: 265–274. doi: 10.1159/000250435 PMID: 19844127
7. Kaiser L, Grönlund H, Sandalova T, Ljunggren HG, van Hage-Hamsten M, Achour A, et al. (2003) The crystal structure of the major cat allergen Fel d 1, a member of the secretoglobin family. J Biol Chem 278: 37730–37735. PMID: 12851385
8. Kaiser L, Velickovic TC, Badia-Martinez D, Adedoyin J, Thunberg S, Hallén D, et al. (2007) Structural characterization of the tetrameric form of the major cat allergen Fel d 1. J Mol Biol 370: 714–727. PMID: 17543334
9. Vailes LD, Li Y, Bao Y, DeGroot H, Aalberse RC, Chapman MD. (1994) Fine specificity of B-cell epitopes on *Felis domesticus* allergen I (Fel d I): effect of reduction and alkylation or deglycosylation on Fel d I structure and antibody binding. J Allergy Clin Immunol 93: 22–33. PMID: 7508462
10. Barnes HJ, Nordlund-Moller L, Nord M,Gustafsson J, Lund, Gillner M. (1996) Structural basis for calcium binding by uteroglobins. J Mol Biol 256: 392–404. PMID: 8594205
11. Andersson O, Nordlund-Moller L, Barnes HJ, Lund J. (1994) Heterologous expression of humanuteroglobin/polychlorinated biphenyl-binding protein. Determination of ligand binding parameter sand mechanism of phospholipase A2 inhibition in vitro. J Biol Chem 269: 19081–19087. PMID: 7518449
12. Nord M, Gustafsson JA, Lund J. (1995) Calcium dependent binding of uteroglobin (PCB-BP/CCSP) to negatively charged phospholipid liposomes. FEBS Lett 374: 403–406. PMID: 7589580
13. Chowdhury B, Mantile-Selvaggi G, Miele L,Cordella-Miele E, Zhang Z, Mukherjee AB. (2002) Lys43 and Asp46 in alpha-helix 3 of uteroglobin are essential for its phospholipase A2 inhibitory activity. Biochem Biophys Res Commun 295: 877–883. PMID: 12127976
14. Kristensen AK, Schou C, Roepstorff P. (1997) Determination of isoforms, N-linked glycan structure and disulfide bond linkages of the major cat allergen Fel d 1 by a mass spectrometric approach. Biol Chem 378: 899–908. PMID: 9377487
15. Seppälä U, Hägglund P, Wurtzen PA, Ipsen H, Thorsted P, Lenhard T, et al. (2005) Molecular characterization of major cat allergen Fel d 1: expression of heterodimer by use of a baculovirus expression system. J Biol Chem 280: 3208–3216. PMID: 15546862
16. Emara M, Royer P-J, Abbas Z, Sewell HF, Mohamed GG, Singh S, et al. (2011) Recognition of the Major Cat Allergen Fel d 1 through the Cysteine-rich domain of the Mannose Receptor determines its allergenicity. J. Biol. Chem. 286: 13033–13040. doi: 10.1074/jbc.M111.220657 PMID: 21335554
17. Platts-Mills TA, Woodfolk JA. (2011) Allergens and their role in the allergic immune response. Immunol Rev 242:51–68. doi: 10.1111/j.1600-065X.2011.01021.x PMID: 21682738
18. Pol-Fachin L, Rusu VH, Verli H, Lins RD. (2012) GROMOS 53A6glyc, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. J Chem Theory Comput 8: 4681–4690.
19. Pol-Fachin L, Verli H, Lins RD. (2014) Extension and validation of the GROMOS 53A6glyc parameter set for glycoproteins. J Comput Chem 35: 2087–2095. doi: 10.1002/jcc.23721 PMID: 25196137
20. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 25: 1656–1676. PMID: 15264259
21. MOPAC Manual (2005) Bond order. Available at: http://www.cache.fujitsu.com/mopac/Mopac2002manual/node464.html (Accessed 3 December 2014).
22. Hsin K, Sheng Y, Harding MM, Taylor P, Walkinshaw MD. (2008) MESPEUS: a database of the geometry of metal sites in proteins. J Appl Cryst 41: 963–968.

23. Erjavec F, Stanovnik L. (1987) Histamine and calcium secretion in the submandibular gland of the cat evoked by two types of stimulation. Agents Actions 20: 188–190. PMID: 2440268

24. Larsson L, Öhman S. (1978) Serum ionized calcium and corrected total calcium in borderline hyperparathyroidism. Clin Chem 24: 1962–1965. PMID: 709830

25. Hsieh MJ, Luo R. (2010) Balancing simulation accuracy and efficiency with the Amber united atom force field. J Phys Chem B 114: 2886–2893. doi: 10.1021/jp906701s PMID: 20131885

26. Grossfield A, Feller SE, Pitman MC (2007) Convergence of molecular dynamics simulations of membrane proteins. Proteins 67: 31–40. PMID: 17243153

27. van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, et al. (2006) Biomolecular modeling: Goals, problems, perspectives. Angew Chem Int Ed Engl 45: 4064–4092. PMID: 16761306

28. Knapp B, Frantal S, Cibena M, Schreiner W, Bauer P (2011) Is an intuitive convergence definition of molecular dynamics simulations solely based on the root mean square deviation possible? J Comput Biol 18: 997–1005. doi: 10.1089/cmb.2010.0237 PMID: 21702691

29. Huang RY, Hudgens JW. (2013) Effects of desialylation on human α1-acid glycoprotein-ligand interactions. Biochemistry 52: 7127–7136. doi: 10.1021/bi4011094 PMID: 24041412

30. Krane S, Itagaki Y, Nakanishi K, Weldon PJ. (2003) "Venom" of the slow loris: sequence similarity of prosimian skin gland protein and Fel d 1 cat allergen. Naturwissenschaften 90: 60–62. PMID: 12590298

31. Hagey LR, Fry BG, Fitch-Snyder H. (2007) Talking defensively: a dualuse for the brachial gland exudate of slow and pygmy lorises. In:Gurski S, Nekaris KAI, editors. Primate Anti-predatory Strategies, vol. 2. New York: Springer. pp. 253–272.

32. Bienboire-Frosini C, Cozzi A, Lafont-Lecuelle C, Vervloet D, Ronin C, Pageat P. (2012) Immunological differences in the global release of the major cat allergen Fel d 1 are influenced by sex and behaviour. Vet J 193: 162–167. doi: 10.1016/j.tvjl.2011.09.031 PMID: 22044714

33. Härd T, Barnes HJ, Larsson C, Gustafsson JA, Lund J. (1995) Solution structure of a mammalian PCB-binding protein in complex with a PCB. Nat Struct Biol. 2: 983–989. PMID: 7583672

34. Dornelas de Andrade AD, Birnbaum J, Magalon C, Magnol JP, Lanteaume A, Charpin D, et al. (1996) Fel d I levels in cat anal glands. Clin Exp Allergy 26:178–180. PMID: 8835125

35. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78: 1950–1958. doi: 10.1002/prot.22711 PMID: 20408171

36. MacKerell AD Jr., Banavali N, Foloppe N. (2001) Development and current status of the CHARMM force field for nucleic acids. Biopolymers 56: 257–265.

37. Lütteke T, Bohne-Lang A, Loss A, Goetz T, Frank M, von der Lieth CW. (2006) GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research. Glycobiology 16: 71R–81R. PMID: 16239495

38. Frank M, Schloissnig S. (2010) Bioinformatics and molecular modeling in glycobiology. Cell Mol Life Sci 67: 2749–2772. doi: 10.1007/s00018-010-0352-4 PMID: 20364395

39. Fernandes CL, Sachett LG, Pol-Fachin L, Verli H. (2009) GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. Carbohydr Res 345:663–671. doi: 10.1016/j.carres.2009.12.018 PMID: 20106471

40. Pol-Fachin L, Becker CF, Guimarães JA, Verli H. (2011) Effects of glycosylation on heparin binding and antithrombin activation by heparin. Proteins 29: 2735–2745.

41. Virgens MYF, Pol-Fachin L, Verli H, Saraiva-Pereira ML. (2014) Effects of glycosylation and pH conditions in the dynamics of human arylsulfatase A. J Biomol Struct Dyn 32: 567–579. doi: 10.1080/07391102.2013.780982 PMID: 23581857

42. Hess B, Kutzner C, van der Spoel D, Lindahl E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J Chem Theory Comput 4: 435–447.

43. Berendsen HJC, Grigera JR, Straatsma TP. (1987) The Missing Term in Effective Pair Potentials. J Phys Chem 91: 6269–6271.

44. Jorgensen WL, Madura JD. (1983) Solvation and conformation of methanol in water. J Am Chem Soc 105: 1407–1413.

45. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18: 1463–1472.

46. Darden T, York D, Pedersen L. (1993) Particle Mesh Ewald–an N.log(N) method for Ewald sums in large systems. J Chem Phys 98: 10089–10092.

47. Berendsen HJC, Postma JPM, DiNola A, Haak JR. (1984) Molecular-dynamics with coupling to an external bath. J Chem Phys 81: 3684–3690.

48. Bussi G, Donadio D, Parrinello M. (2007) Canonical sampling through velocity rescaling. J Chem Phys 126:014101. PMID: 17212484

49. Berka K, Hanak O, Sehnal D, Banas P Navratilova V, Jaiswal D, et al. (2012) MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. Nucleic Acids Res 40: W222–W227. doi: 10.1093/nar/gks363 PMID: 22553366

50. Sehnal D, Svobodova Varekova R, Berka K, Pravda L, Navratilova V, Banas P, et al. (2013) MOLE 2.0: Advanced approach for analysis of biomacromolecular channels. J Cheminform 5: 39. doi: 10.1186/1758-2946-5-39 PMID: 23953065

51. Sethi A, Eargle J, Black AA, Luthey-Schulten Z. (2009) Dynamical networks intRNA:protein complexes. Proc Natl Acad Sci U S A. 106:6620–6625. doi: 10.1073/pnas.0810961106 PMID: 19351898

52. Black Pyrkosz A, Eargle J, Sethi A, Luthey-Schulten Z. (2010) Exit strategies for charged tRNA from GluRS. J Mol Biol 16:1350–1371.

53. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. (2004) J Comput Chem 25:1605–1612. PMID: 15264254

54. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS et al. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J. Computational Chemistry 16: 2785–2791.

55. Stierand K, Maaß P, Rarey M. (2006) Molecular Complexes at a Glance: Automated Generation of two-dimensional Complex Diagrams. Bioinformatics, 22: 1710–1716. PMID: 16632493

56. Wallace AC, Laskowski RA, Thornton JM. (1996) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng., 8: 127–134.

57. Stewart J. (2008) MOPAC2009, Stewart Computational Chemistry, Colorado Springs, USA, Retrieved from http://OpenMOPAC.net.

58. Stewart JP. (2007). Optimization of parameters for semi empirical methods V: Modification of nddo approximations and application to 70 elements. J Mol Model 13: 1173–1213. PMID: 17828561

59. Stewart JP. (1996). Application of localized molecular orbitals to the solution of semi empirical self-consistent field equations. Int J Quant Chem 58: 133–146.

60. Fernandes CL, Escouto GB, Verli H. (2014) Structural glycobiology of heparinase II from *Pedobacter-heparinus*. J Biomol Struct Dyn 32:1092–1102. doi: 10.1080/07391102.2013.809604 PMID: 23808670

61. Klamt A, Schüürmann G. (1993) COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. J Chem Soc, Perkin Trans 2: 799–805.

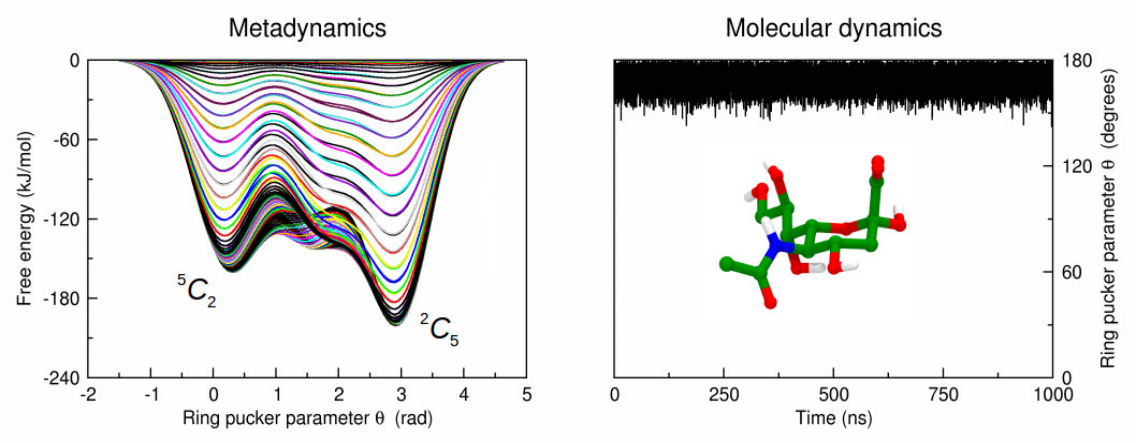

Fig. S1. Validation of GROMOS 53A6glyc parameters implemented for sialic acid.

Table S1. Parameters used in the GROMOS 53A6glyc force field for sialic acid, in a GROMACS-compatible format. These encompass atom types, atomic partial charges, charge-group definition, bond stretching, bond-angle bending, improper dihedral deformation and torsional potential dihedral.

```
[ NEUN ]
 [ atoms ]
   C5   CH1     0.00000     0
   N5     N    -0.28000     0
  HN5     H     0.28000     0
  C5A     C     0.38000     0
  O5A     O    -0.38000     0
  C5B   CH3     0.00000     0
   C4   CH1     0.23200     1
   O4    OA    -0.64200     1
  HO4     H     0.41000     1
   C7   CH1     0.23200     2
   O7    OA    -0.64200     2
  HO7     H     0.41000     2
   C8   CH1     0.23200     3
   O8    OA    -0.64200     3
  HO8     H     0.41000     3
   C9   CH2     0.23200     4
   O9    OA    -0.64200     4
  HO9     H     0.41000     4
   C6   CH1     0.37600     5
   O6    OA    -0.48000     5
   C2   CH1     0.23200     5
   O2    OA    -0.53800     5
  HO2     H     0.41000     5
   C3   CH2     0.00000     5
   C1     C     0.36000     6
  O1A    OM    -0.68000     6
  O1B    OM    -0.68000     6
[ bonds ]
   C5    N5    gb_21
   N5   HN5    gb_2
   N5   C5A    gb_10
  C5A   O5A    gb_5
  C5A   C5B    gb_16
   C5    C4    gb_26
   C5    C6    gb_26
   C4    O4    gb_20
   C4    C3    gb_26
   O4   HO4    gb_1
   C3    C2    gb_26
   C7    C6    gb_26
   C7    O7    gb_20
   O7   HO7    gb_1
   C7    C8    gb_26
   C8    O8    gb_20
   O8   HO8    gb_1
   C9    C8    gb_26
   C9    O9    gb_20
   O9   HO9    gb_1
   C6    O6    gb_20
   O6    C2    gb_20
   C2    O2    gb_20
   O2   HO2    gb_1
   C1   O1A    gb_6
   C1   O1B    gb_6
   C1    C2    gb_26
[ angles ]
;  ai    aj    ak   gromos type
   C6    C5    N5     ga_15
   N5    C5    C4     ga_15
   C5    N5   HN5     ga_18
   C5    N5   C5A     ga_31
  HN5    N5   C5A     ga_32
   N5   C5A   O5A     ga_33
   N5   C5A   C5B     ga_19
  O5A   C5A   C5B     ga_30
   C4    C5    C6     ga_8
   C5    C4    O4     ga_9
   C5    C4    C3     ga_8
   O4    C4    C3     ga_9
   C4    O4   HO4     ga_12
   C4    C3    C2     ga_8
   C5    C6    C7     ga_8
   C5    C6    O6     ga_9
   C7    C6    O6     ga_9
  HO7    O7    C7     ga_12
   O7    C7    C6     ga_9
   O7    C7    C8     ga_9
   C6    C7    C8     ga_8
  HO8    O8    C8     ga_12
   O8    C8    C7     ga_9
   O8    C8    C9     ga_9
   C7    C8    C9     ga_8
   O9    C9    C8     ga_9
  HO9    O9    C9     ga_12
   C6    O6    C2     ga_10
   C3    C2    O6     ga_9
   C3    C2    O2     ga_9
   O6    C2    O2     ga_9
   O6    C2    C1     ga_9
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
  O1A    C1   O1B     ga_38
  O1A    C1    C2     ga_22
  O1B    C1    C2     ga_22
   C3    C2    C1     ga_8
 [ impropers ]
;  ai    aj    ak    al   gromos type
   N5   C5A   C5B   HN5     gi_1
  C5A   C5B    N5   O5A     gi_1
   C5    O6    C7    C6     gi_2
   C5    O4    C3    C4     gi_2
   C5    N5    C4    C6     gi_2
   C2    O6    O2    C3     gi_2
   C3    O6    C1    C2     gi_2
   C1   O1A   O1B    C2     gi_1
[ dihedrals ]
;  ai    aj    ak    al   gromos type
   C3    C4    C5    N5     gd_16
   C7    C6    C5    N5     gd_16
   O4    C4    C5    N5     gd_17
   O6    C6    C5    N5     gd_18
   C4    C5    N5   C5A     gd_39
   C5    N5   C5A   C5B     gd_14
   C6    C5    C4    O4     gd_43
   C6    C5    C4    O4     gd_44
   C6    C5    C4    C3     gd_34
   C4    C5    C6    C7     gd_34
   C4    C5    C6    O6     gd_43
   C4    C5    C6    O6     gd_44
   C3    C4    O4   HO4     gd_30
   C5    C4    C3    C2     gd_34
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C4    C3    C2    O6     gd_43
   C4    C3    C2    O6     gd_44
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C6    C7    O7   HO7     gd_30
   O7    C7    C6    O6     gd_5
   O7    C7    C6    O6     gd_37
   C6    C7    C8    O8     gd_43
   C6    C7    C8    O8     gd_44
   O7    C7    C8    O8     gd_18
   C7    C8    O8   HO8     gd_30
   C7    C8    C9    O9     gd_43
   C7    C8    C9    O9     gd_44
   O8    C8    C9    O9     gd_18
   C8    C9    O9   HO9     gd_30
  O1A    C1    C2    C3     gd_40
   C5    C6    O6    C2     gd_29
   C6    O6    C2    C3     gd_29
   O6    C2    O2   HO2     gd_2
   O6    C2    O2   HO2     gd_32
 [ exclusions ]
  HO2    O6
```

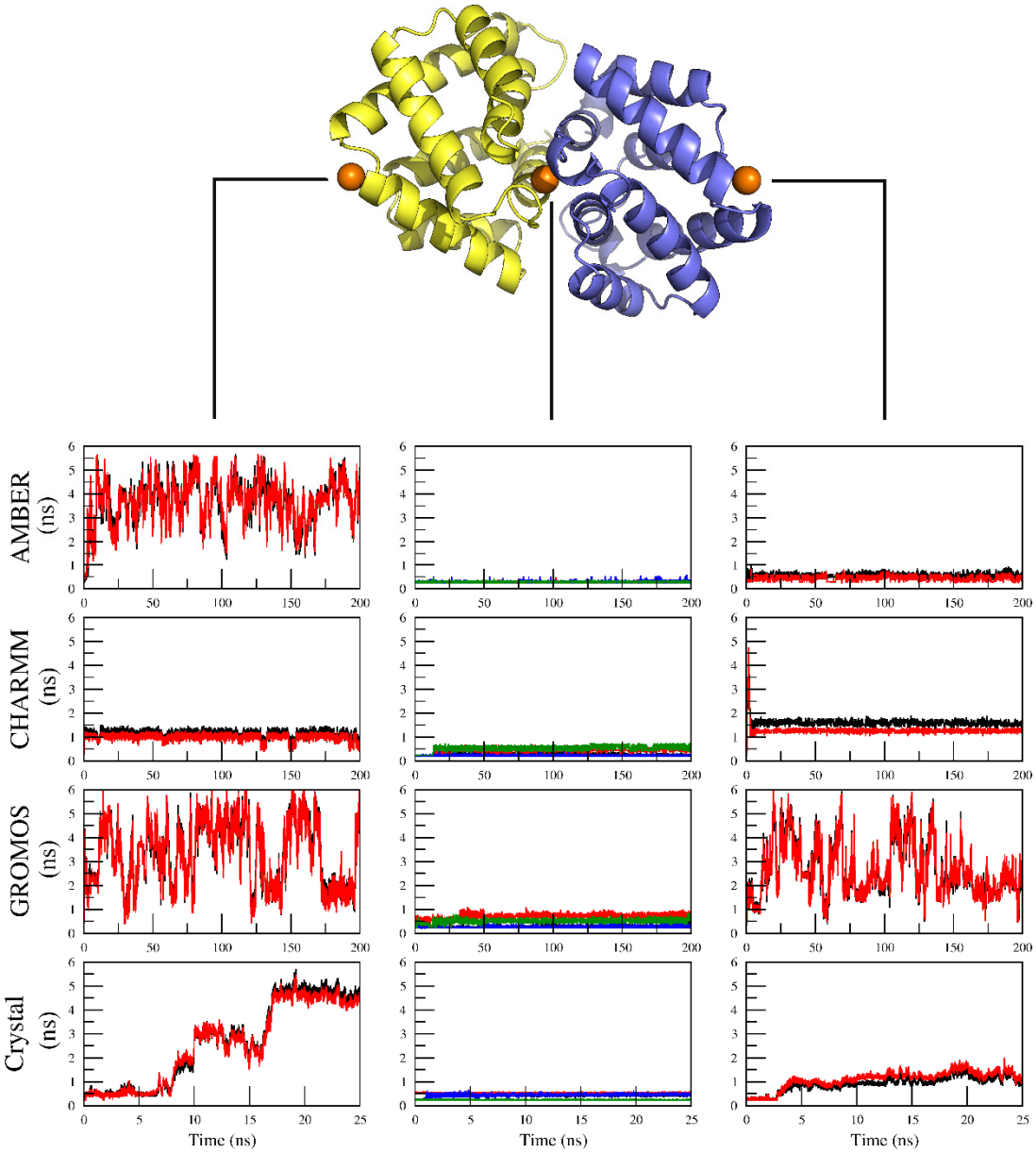

Fig. S2. Distances measured for molecular dynamics simulations of Fel d 1 under different force fields. Please refer to Table 1 for details.

Table S2. Essential node pairs detected in four Fel d 1 simulated systems. Nodes that appear more than once are grouped by color. Interactions detected between pairs by PDBsum for the crystallographic structure are in parentheses.

	Interface (Chain A - Chain B)	**Chain A**	**Chain B**
Without Ca^{2+}	Lys29-Asp82 (1HB, NBC) Leu132-Leu94	Cys44-Cys118 (DS) Ile64-Val102	Leu41-Tyr119 (NBC) Phe85-Met134 (NBC)
With Ca^{2+}	Leu31-Leu94 Asn91-Leu129 (NBC) Leu132-Phe85 (NBC)	Cys3-Ala74 Val10-Phe84	Val6-Tyr81 Asn37-Arg127
Minimal Glycosylation	Phe85-Asp130 (NBC) Val128-Leu94 Ser139-Thr76 (NBC)	Val6-Phe80 Ile40-Asn122 (NBC)	Val6-Phe80 Asn37-Arg127
Full Glycosylation	Phe85-Leu132 (NBC) Asn89-Leu129 (NBC) Thr135-Phe85 (NBC)	Lys7-Phe80 Asn37-Arg127 (2HB)	Cys3-Tyr65 Asn37-Val128

HB: hydrogen bond(s), NBC: non-bonded contacts, DS: disulfide bridge.
Note: There is a pair in the simulation with Ca^{2+} (Leu129B-Arg127B) that connects the interface pair (Asn91-Leu129) to the 'hinge' pair (Arg127-Asn37).

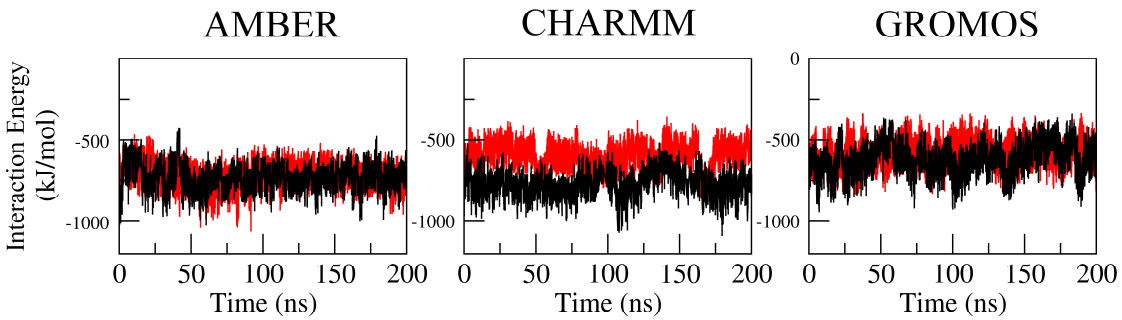

Fig. S3. Interaction energies measured between Fel d 1 monomers during molecular dynamics simulations under different force fields. Color coding: Simulations without Ca^{2+}ions (black), simulations with Ca^{2+} ions (red).

Table S3. Sizes of simulated systems.

System		Number of atoms	Number of counterions (Na^+)
AMBER	With Ca^{2+}	28,951	8
	Without Ca^{2+}	31,382	16
CHARMM	With Ca^{2+}	31,750	8
	Without Ca^{2+}	31,382	16
GROMOS	With Ca^{2+}	34,627	8
	Without Ca^{2+}	28,559	16
Crystal mimic	With Ca^{2+}	586,592	88
Glycosylated	Full glycosylation	88,194	14
	Minimal glycosylation	65,078	8

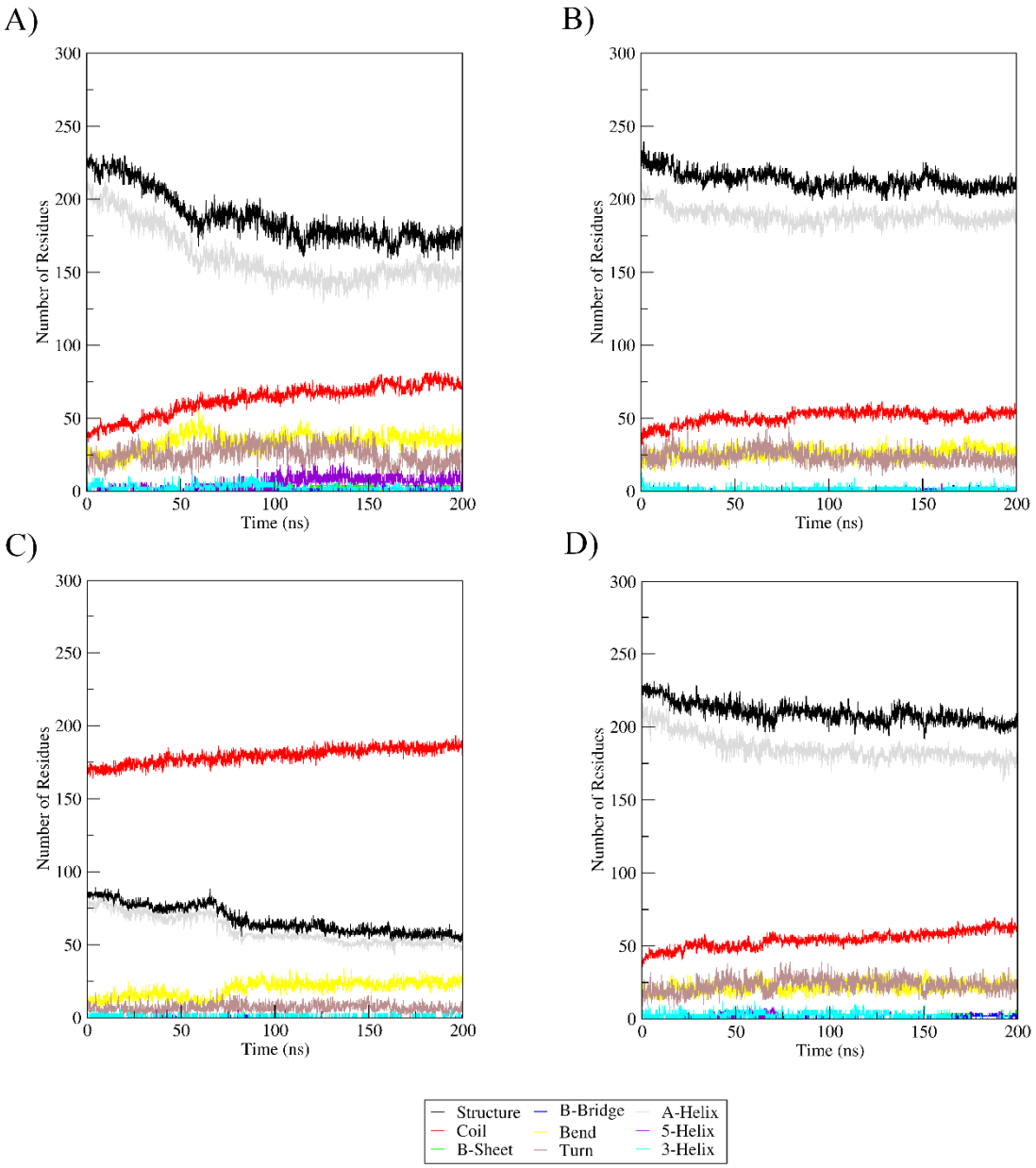

Fig. S4. Secondary structure content for the Fel d 1 systems simulated with GROMOS. (A) Fel d 1 without Ca^{2+}; (B) Fel d 1 with Ca^{2+}; (C) Fel d 1 with minimal glycosylation; (D) Fel d 1 with full glycosylation.

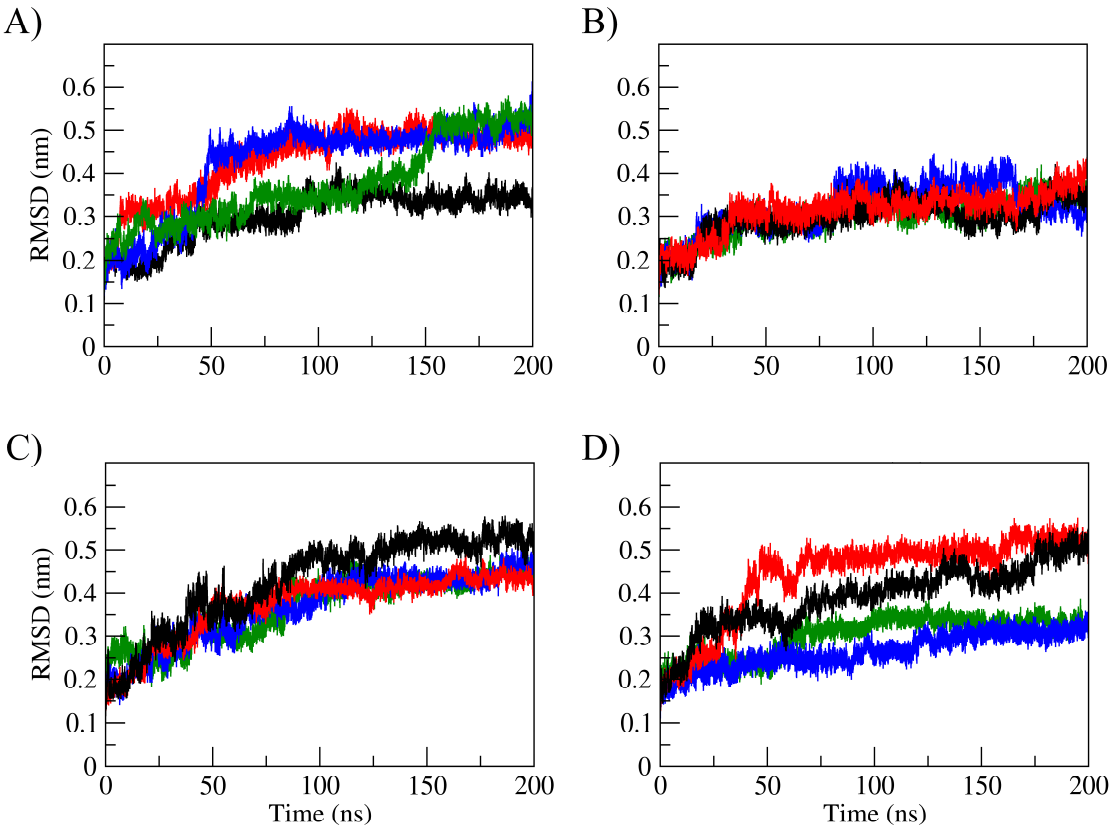

Fig. S5. All-atom root mean square deviation (RMSD) measured for Fel d 1 segments corresponding to the original chain regions (fused in the crystallized protein). (A) Calcium-free system; (B) Calcium-bound system; (C) Minimally glycosylated system; (D) Fully glycosylated system. Region A_1 in black, region A_2 in red, region B_1 in blue, region B_2 in green.

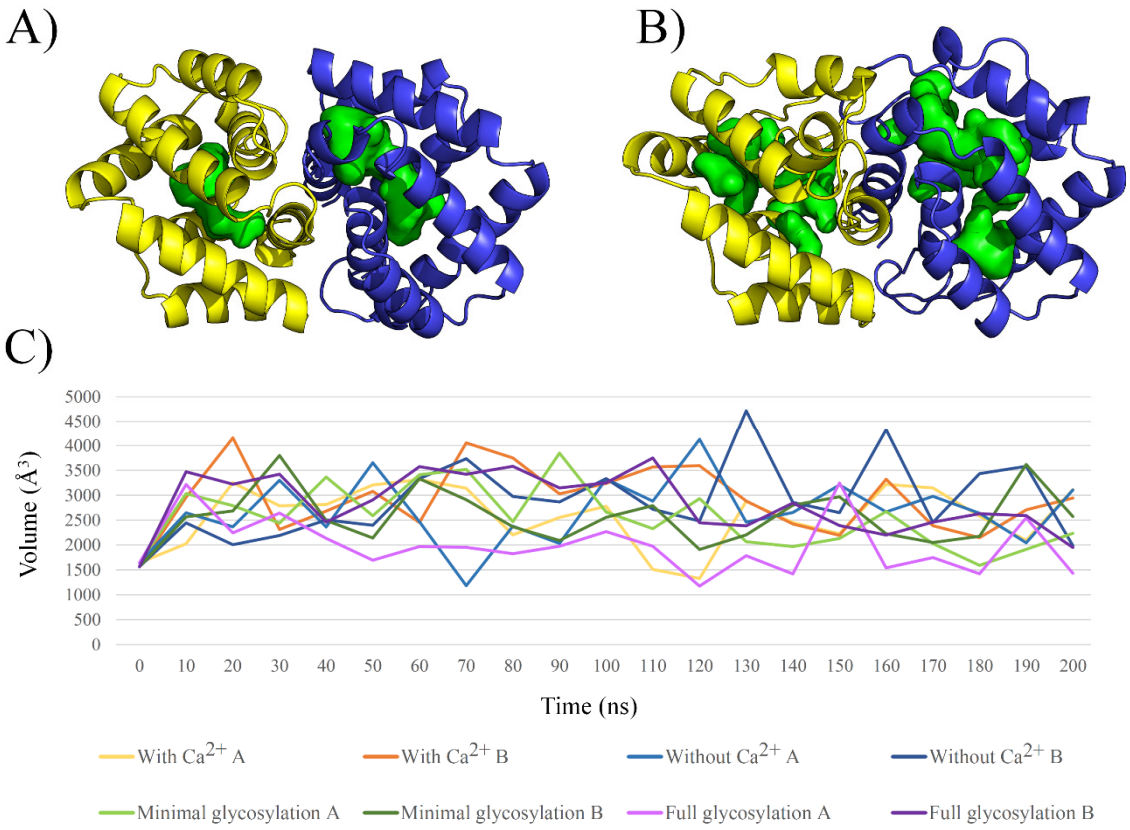

Fig. S6. Cavity volume overview (calcium-free system as an example). (A) initial cavity volume, (B) final cavity volume (cavity volumes shown as green surfaces); (C) cavity volume measurements for each Fel d 1 chain in four simulations.

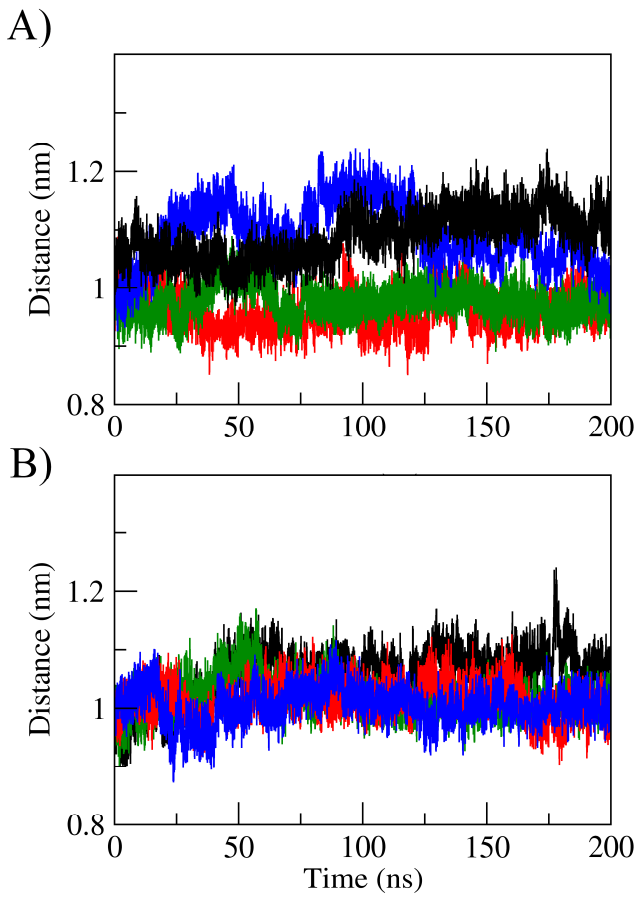

Fig. S7. Structural 'oscillation' observed for (A) chain A, and (B) chain B of Fel d 1 in four different simulations. The distance between each original chain in the native Fel d 1 (half chains of the crystallographic protein) is shown. The oscillation may indicate an opening-closing behavior. (calcium-free system in black, calcium-bound system in red, minimally glycosylated system in blue, fully glycosylated system in green; distances measured for the centers of mass in each structure).