J Comput Aided Mol Des (2010) 24:1009–1021
DOI 10.1007/s10822-010-9392-y

Selenoglycosides in silico: ab initio-derived reparameterization of MM4, conformational analysis using histo-blood group ABH antigens and lectin docking as indication for potential of bioactivity

Francesco Strino · Jenn-Huei Lii · Chaitanya A. K. Koppisetty · Per-Georg Nyholm · Hans-Joachim Gabius

Received: 13 April 2010 / Accepted: 6 October 2010 / Published online: 26 October 2010

Abstract The identification of glycan epitopes such as the histo-blood group ABH determinants as docking sites for bacterial/viral infections and signals in growth regulation fuels the interest to develop non-hydrolysable mimetics for therapeutic applications. Inevitably, the required substitution of the linkage oxygen atom will alter the derivative's topology. Our study addresses the question of the impact of substitution of oxygen by selenium. In order to characterize spatial parameters and flexibility of selenoglycosides, we first performed ab initio calculations on model compounds to refine the MM4 force field. The following application of the resulting MM4R version appears to reduce the difference to ab initio data when compared to using the MM4 estimator. Systematic conformational searches on the derivatives of histo-blood group ABH antigens revealed increased flexibility with acquisition of additional low-energy conformer(s), akin to the behavior of S-glycosides. Docking analysis using the Glide program for eight test cases indicated potential for bioactivity, giving further experimental investigation a clear direction to testing Se-glycosides as lectin ligands.

Keywords Blood group antigens · Drug design · Lectins · Molecular mechanics · Selenoglycoside

Introduction

The increasing insights into the functional roles of the sugar part of cellular glycoconjugates attest to the enormous coding capacity of glycans, which underlies the concept of the sugar code [1]. Remarkable progress in carbohydrate chemistry and in defining targets for carbohydrate-based drug design is opening routes to potential medical applications [2–5]. To ensure biostability by precluding hydrolytic cleavage of oligosaccharides as pharmaceutics, the replacement of the oxygen atom in the glycosidic linkage is required. Toward this end, a series of modified glycosides can be prepared, for instance yielding C-, S- or Se-glycosides. Inevitably, the incorporation of linkage atoms other than oxygen will have a bearing on spatial characteristics of the resulting variant (Table 1). In experimental terms, for example, C-lactoside or thiodigalactoside acquire topological parameters and flexibility levels different from the O-glycoside but, remarkably, maintain bioactivity when tested with adhesion/growth-regulatory human lectins [6–8].

Electronic supplementary material The online version of this article (doi:10.1007/s10822-010-9392-y) contains supplementary material, which is available to authorized users.

F. Strino · P.-G. Nyholm
Institute of Biomedicine/Section of Medical Biochemistry, University of Gothenburg, Medicinaregatan 9, 40530 Gothenburg, Sweden
e-mail: nyholm@medkem.gu.se

F. Strino · C. A. K. Koppisetty · P.-G. Nyholm (✉)
Biognos AB, Generatorsgatan 1, 40274 Gothenburg, Sweden
e-mail: per-georg.nyholm@biognos.se

J.-H. Lii
Department of Chemistry, National Changhua University of Education, Changhua 50058, Taiwan

C. A. K. Koppisetty
Department of Computer Science and Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden

H.-J. Gabius
Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität München, Veterinärstr. 13, 80539 München, Germany
e-mail: gabius@tiph.vetmed.uni-muenchen.de

Table 1 Geometric properties of the natural glycosidic linkage and its C-, S- and Se-derivatives

Linkage	vdW radius (Å)	C-X bond (Å)	C-X-C angle (°)	C-X-C distance (Å)
O	1.5	1.4	115	2.4
C	1.7	1.5	110	2.5
S	1.8	1.8	95	2.9
Se	1.9	1.9	95	3.0

For each linkage type, van der Waals radius of the central atom, length of the C-X bond, value of the C-X-C angle and distance between the two ends of the glycosidic linkage are listed

To systematically explore the consequences of substitutions of the oxygen of the glycosidic linkage we have previously studied thioglycosides, using histo-blood group ABH saccharides as a biologically relevant test case [9]. Ab initio calculations led to the addition of parameters for the MM4 force field specifically suited for S-linked saccharides. MM4 calculations with the new parameters revealed increased flexibility for S-linked glycans compared to their parental, O-linked counterparts [9]. A hitherto open question concerns the behavior of Se-glycosides. Their elegant introduction to crystallography in the case of a bacterial adhesin, here the β-methyl derivative of N-acetylglucosamine [10], intimates that they may find more applications in lectin research than merely solving the crystallographic phase problem.

Herein, we first report on ab initio calculations to enable a refinement of MM4-based calculations. Then we document the resulting conformational parameters of Se-glycosides of the histo-blood group epitopes. Finally, as an indication of potential bioactivity, we report docking studies with eight proteins of known reactivity including a presentation of the interaction scores obtained with the GLIDE program.

Methods

Nomenclature definitions and computational processing follow our previous report [9]. In detail, the notation (1→) indicates the natural O-glycosidic linkage (1–S→) a thioglycosidic linkage and (1–Se→) a selenoglycosidic linkage. Torsion angles are defined using heavy atoms, e.g. for the (1–Se→2) linkage, φ = O5-C1-Se-C′2 and ψ = C1-Se-C′2-C′3. The notation H2-X denotes the α-L-Fuc*p*-(1-X→2)-β-D-Gal*p* disaccharide derivatives where X can be O for O-glycosidic, S for thioglycosidic or Se for selenoglycosidic. Similarly, A2-X is used for α-D-Gal*p*NAc-(1-X→3)-β-D-Gal*p* and B2-X stands for α-D-Gal*p*-(1-X→3)-β-D-Gal*p*. Regarding trisaccharides, the notation A3-X/X is used for the histo-blood group A epitope to specify the type of the glycosidic bond in the cases of α-L-Fuc*p*-(1-X→2)-β-D-Gal*p* and α-D-Gal*p*NAc-(1-X→3)-β-D-Gal*p* linkages, respectively. Histo-blood group B epitope and its derivatives are similarly characterized using the B3-X/X notation. When no X is explicitly given, O-glycosidic bonds are present, e.g. A3 = A3-O/O and H2 = H2-O.

Selenium parameterization

Parameter input for selenium atoms to the force field MM4(03) [11–16], commonly used for energy calculations, was derived from ab initio results for the compounds CH_3-Se-CH_3, CH_3-Se-C_2H_5, CH_3-Se-C_3H_7, HO-CH_2-CH_2-Se-CH_3, CH_3-O-CH_2-Se-CH_3 and CH_3-O-CH_2-Se-C_2H_5 selected to resemble relevant fragments of a Se-disaccharide. Initially, the C-Se bond and the C-Se-C angles were parameterized by fitting available experimental data on dimethyl selenide, i.e. 1.943 Å and 96.2°, respectively [17]. Vibrational frequencies were fit to MP2/6-311++G(2d,2p) results for all compounds mentioned above. For the two compounds CH_3-O-CH_2-Se-CH_3 and CH_3-Se-C_2H_5, the relevant torsion angles were sampled using a 30° step size. Each conformation was independently minimized energetically and its bond distances and angles were recorded. A torsional profile for CH_3-Se-C_2H_5 and a two-dimensional map for CH_3-O-CH_2-Se-CH_3 were thereby established. Quantum mechanical calculations were done applying both the Becke3LYP/6-311+G(d) DFT theory and the computationally more demanding MP2/6-311++G(2d,2p) theory using the Gaussian 09A software (Gaussian Inc., Wallingford, CT, USA). Both theories were applied in parallel to test the effects of different treatments of electron correlation considering a previous result that the Becke3LYP DFT theory may result in smaller intramolecular basis set superposition errors (BSSE) than MP2 on carbohydrates [18]. The data sets were combined to yield optimized MM4-based torsional parameters using the least-squares fitting program TORSFIND. The previously published parameter set MM4R [9] was then extended by the resulting parameters enabling their comparison to the parameters automatically estimated by the built-in parameter estimator of MM4. Finally, the ability to model the φ torsion angle of selenoglycosides was tested by comparing the torsional profiles of the pyranoside analogue 2-(methylselanyl)tetrahydro-2H-pyran. For both α and β geometries at the anomeric oxygen atom, energy profiles were obtained by ab initio calculations using both the MP2/6-311++G(2d,2p) and Becke3LYP/6-311+G(d) theories as well as by MM4 calculations using the default and the MM4R parameter sets.

Molecular mechanics calculations

Successive MM4 calculations were performed based on the obtained MM4R parameters for selenoglycosides using a dielectric constant of 80 to simulate an aqueous environment and a tight energy convergence criterion of $0.8 \times 10^{-5} \times$ N kcal/mol (where N is the number of atoms of the molecule) instead of the default of $0.8 \times 10^{-4} \times$ N kcal/mol. Systematic searches were initially performed on all constituent disaccharide moieties of the histo-blood group ABH derivatives using a step size of 15° for the φ and ψ glycosidic torsion angles. Two energetically favored conformations of the C5-C6 bond were investigated for each residue. Clockwise (C) and anti-clockwise (R) arrangements of the secondary hydroxyl groups were also considered. During energy minimizations, the values of the glycosidic torsion angles were constrained, while other parts of the molecule were relaxed. The sampled points were then used to draw adiabatic energy maps, i.e. for each φ/ψ geometry the lowest energy conformation was selected.

The established disaccharide maps served as filters for generating a four-dimensional energy map for each of the studied trisaccharides. All conformations with a more than a 12 kcal/mol energy difference from the global minimum of the tested disaccharide were rigorously excluded. This simple filtering allowed skipping detailed energy assessment for a large number of conformations. During trisaccharide map generation, the optimal hydroxyl conformations of the constituent disaccharide moieties were used. The population density of relative minima was estimated by first computing Boltzmann probabilities of each grid point at 25 °C and then summing up all data points up to the isocontour at 3 kcal/mol above the global minimum. Regions of the conformational space accounting for less than 1% of the population or more than a 2 kcal/mol difference from the global minimum were excluded. Adiabatic potential energy maps were then computed for all possible combinations of O-, S- and Se-linkages of the three hexopyranose rings, excluding those previously calculated [9]. The four-dimensional maps (A3-O/Se, A3-Se/O, A3-Se/Se, A3-S/Se, A3-Se/S, B3-O/Se, B3-Se/O, B3-Se/Se, B3-S/Se, B3-Se/S) were set up with a step size of 15° for the torsion angles of the glycosidic linkages.

The GLYGAL [19, 20] software was applied for running the energy calculations, analyzing the obtained data, and drawing the adiabatic maps. These calculations did not make use of the genetic algorithm search implemented in GLYGAL [20] but took advantage of its systematic search and filtering capabilities. Data from our previous study of thioglycosidic derivatives [9] were used for all moieties not containing selenium atoms. Finally, since MM4 uses a hydrogen bonding term, the so-called charge-transfer term [21], which is independent of the dielectric constant, we performed a series of test calculations, in which this term was scaled down by a factor of 1/80, on the seleno-disaccharide α-L-Fuc*p*-(1–Se→2)-β-D-Gal*p*. The resulting φ/ψ energy maps showed only very minor differences when compared to the maps obtained with the regular MM4 output, strongly suggesting that the charge-transfer term does not play a significant role for the interresidue energetics of these systems.

Dockings

The dockings were performed using software from the Schrödinger suite (Schrödinger LLC, New York, NY, USA). Ligand structures were computationally processed with Ligprep (version 2.3; Schrödinger LLC, New York, NY, USA). For each molecule, all energetically favorable conformers were used as starting structures. As targets, eight carbohydrate-binding proteins were selected due to the availability of structural data and experimentally proven reactivity: FcsSBP (PDB 2W7Y), a protein of *Streptococcus pneumoniae*; the fungal galectin CGL2 (PDB 1ULF); the A/B histo-blood group antigen-specific domain of the β-galactosidase from *Clostridium perfringens* termed GH98CBM51, a bacterial carbohydrate-binding module (PDB 2VNG); the galactoside-specific agglutinin from the European mistletoe *Viscum album L.* (VAA), strongly reacting with H- and B-disaccharides [22], by combining the structural model of VAA with information from experimental input and molecular dynamics runs at neutral pH [23]; the B-determinant-specific *Marasmius oreades* agglutinin (MOA) (PDB 3E2F); the *Pseudomonas aeruginosa* lectin I (PA-IL, PDB 2VXJ) which binds α-galactose through a calcium-mediated interaction; and the fucose-binding proteins *Ulex europaeus* agglutinin I (UEA-I; PDB 1JXN) and *Anguilla anguilla* agglutinin (AAA; PDB 1K12), which have been co-crystallized with fucose. For each protein, hydrogens were added and the structures were minimized in the OPLS2005 force field in order to remove close contacts and clashes using the Impref (version 1.27; Schrödinger LLC) restrained minimization utility. Thereafter, a grid was set up centered on the binding site, and a box with a side length of 18 Å was defined. Finally, dockings were performed using the program Glide XP (version 5.5; Schrödinger LLC) with the OPLS2005 [24] force field. Although Glide is optimized for OPLS2001, the OPLS2005 force field was applied due to the absence of selenium parameters in the 2001 version and superiority of the updated version for running calculations on disaccharides [25].

Since the dockings failed to reach the energetically most favorable bound-state conformation in certain cases, they were also repeated with ligand conformers suitably fitted

into the binding site. In particular, the central galactose moiety was superimposed on the position of this residue in the co-crystallized ligand for FcsSBP, CLG2, CBM51 and MOA. A similar superimposition process was applied to the fucose moiety in the cases of UEA-I and AAA, with the obvious exclusion of A2 and B2 disaccharides. This procedure was not applied to VAA and PA-IL, because fitted conformations suffered from steric clashes with the protein. The best OPLS2005 docking score from the complete ensemble of all different starting conformers was finally recorded for each compound.

Binding affinities were converted to binding energies at room temperature using the formula ΔG (in kcal/mol) = $-1.36 \times \log_{10}(K_a)$. Five complexes with experimental binding data were rescored with Glide XP: for CGL2, the extended disaccharides were directly obtained from the coordinates available in the PDB files 1ULD and 1ULE; for CBM51 the extended trisaccharides were obtained by adding (1→4)-β-D-Glc*p*NAc1CH_2-CH_2-N_3 to the histo-blood group A- and B-epitopes in the crystal structures (2VNG and 2VNO); and for PA-IL, the coordinates of the extended B2 structure α-D-Gal*p*-(1→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p* were taken directly from the PDB file (2VXJ).

The molecular graphics were produced with Sybyl 8.0 (Tripos Inc., St. Louis, MO, USA). The MOLCAD module of Sybyl was used to generate the solvent-accessible surfaces and their lipophilic potentials [26]. Protein–ligand interactions were visualized using edited LIGPLOT [27] diagrams obtained from the PDBsum website [28].

Results and discussion

Selenium parameterization

The results of the ab initio calculations on the selenium-containing compounds given above were included in MM4R, this addition being available as supplementary material. The selenoglycosidic linkage CH_3-O-CH_2-Se-CH_3 was modeled considerably better when the new parameters were used instead of relying on MM4's built-in parameter estimator. Explicitly, the root mean squared error (RMSE) based on the MP2 calculations with the MM4 parameters obtained with the parameter estimator is 1.70 kcal/mol. In comparison, applying the new parameters reduced the error to only 1.01 kcal/mol. Of note, the errors are mainly found in energetically highly unfavorable conformational regions. If the RMSE is weighted by the Boltzmann probability at room temperature, the weighted RMSE decreases from 1.07 to 0.36 kcal/mol for MM4 and MM4R, respectively.

The ab initio energy profiles obtained by fixing one torsion at 60° or 180° are shown in Fig. 1. The profiles, in which the C-O-C-Se dihedral is constrained, mimic the profiles of the φ torsion angles of Se-glycopyranosides; in particular, constraining at 60° mimics the α geometry and at 180° the β geometry. There is only one significant deviation from the O-glycosidic linkage type [12] in the O-C-Se-C profile with the C-O-C-Se torsion fixed at 60°. In this profile, the difference between the energetically favored conformation at 60° and the secondary

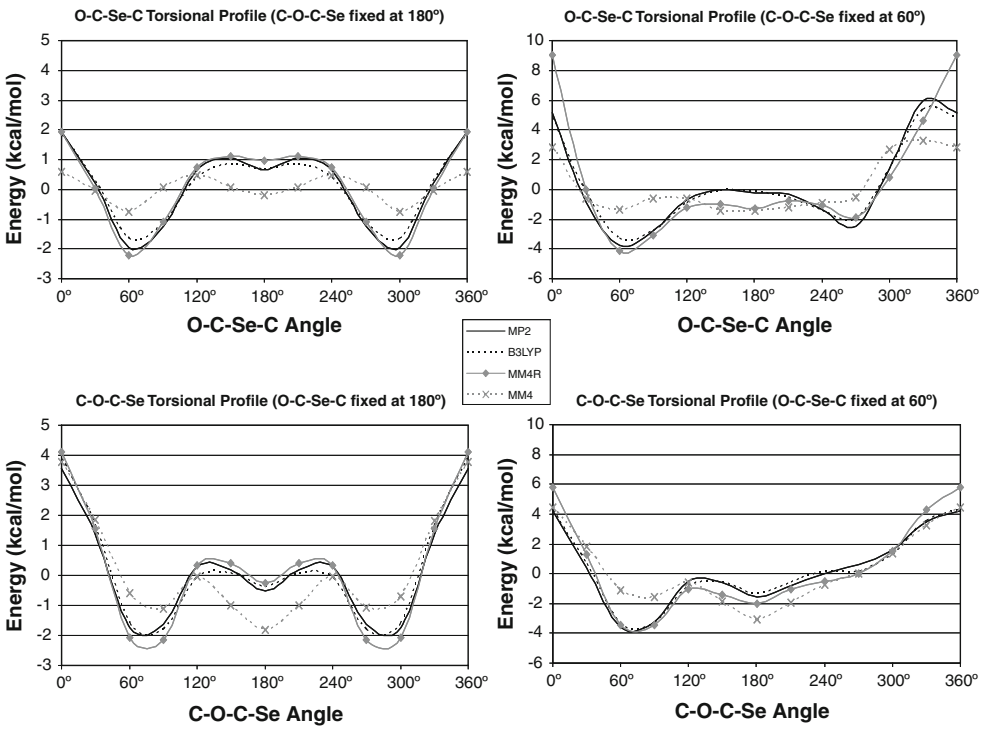

Fig. 1 Energy profiles for the central torsion angles of the fragment CH_3-O-CH_2-Se-CH_3, representing a selenoglycoside. In the graphs the torsional profiles calculated with the standard MM4 force field, MM4 with the new parameter set (MM4R), the B3LYP/6-311+G(d) theory and the MP2/6-311++G(2d,2p) theory are shown (see central inset for assignment of curves)

conformation at −90° is around 2 kcal/mol in the seleno-derivative, whereas it is about 4 kcal/mol in the natural glycosidic linkage. Of note, the profiles for the Se-glycosidic linkage are very similar to those obtained for the thioglycosidic linkage [9]. The crystal structures of the two β-methylselenides of N-acetylgalactosamine [29] and N-acetylglucosamine in complex with a bacterial adhesin [10], offer the opportunity to check the energetically most favored conformations in the free and the bound states. The constrained ring torsion C5-O5-C1-Se assumes the values 176.1° and 172.0°, respectively, whereas the flexible φ glycosidic torsion O5-C1-Se-C is at 61.5° and 61.7°, respectively. These values fall within the global minimum of the calculated profile for O-C-Se-C with C-O-C-Se fixed at 180°. Comparative calculations performed on the φ torsional profiles of β- and α-methylselenide-substituted tetrahydropyran derivatives obtained using both sets of MM4 parameters and two ab initio theories (Fig. 2) were in accord with the results obtained for the CH_3-O-CH_2-Se-CH_3 fragment. This additional result on a model compound indicates that MM4R can model the energy profile of the φ torsion of selenoglycosides significantly more accurately than default parameters obtained by the parameter estimator.

When comparing the geometric data for O-, C-, and S-glycosides to the Se-dependent parameters compiled in Table 1, the difference to O/C-glycosides is underscored and similar properties for S- and Se- glycosides are revealed. The van der Waals radii were obtained from the literature [30], while the geometrical properties were estimated by MP2/6-311++G(2d,2p) calculations.

Disaccharide derivatives

The results of the systematic search on the Se-disaccharides are summarized and compared to the previously published data for the natural disaccharide moieties [9] in Table 2. The adiabatic maps are shown in Fig. 3. The disaccharide H2-Se has two preferred conformers (Fig. 3a′) with similar energy level and population density centered at $\varphi/\psi \approx -75°/-90°$ and $\varphi/\psi \approx -75°/60°$. The latter is not energetically favorable in the O-glycoside (Fig. 3a).

Table 2 Characterization of the energetically favorable conformations calculated for the disaccharide moieties of histo-blood group ABH antigens and their derivatives

Compound	φ/ψ	Energy	Population (%)
H2-O	−90°/−120°		99.0
H2-S	−90°/−90°	+0.03	62.4
	−90°/45°		35.1
H2-Se	−75°/60°	+0.37	45.3
	−75°/−90°		51.9
A2-O	75°/90°		93.4
A2-S	75°/60°	+0.40	77.3
	75°/−75°		14.8
A2-Se	75°/−75°	+0.24	35.5
	120°/60°		58.0
B2-O	75°/90°		93.4
B2-S	90°/75°	+0.43	75.4
	75°/−60°		19.2
B2-Se	75°/−75°	+1.01	49.5
	60°/75°		42.2

Similarly, the maps for the A2-Se (Fig. 3b′) and B2-Se (Fig. 3c′) disaccharides reveal two energetically favored regions: the first, defined by φ varying between 45° and 120° and ψ between 45° and 180°; the second one is a well-defined conformation around $\varphi/\psi \approx 75°/-60°$, which is not populated in the natural disaccharide (Fig. 3b, c). As expected, N-acetylation of the galactosamine moiety does not strongly affect the conformational properties, with the exception of the conformer of A2-Se with $\varphi/\psi \approx 120°/60°$, where the O4 atom of the galactose moiety is in hydrogen bond contact with the hydrogen of the amide.

Regarding ring conformations, the galactose moieties prefer the *gt* configuration of the C5-C6 torsion, while the energy of the *gg* configuration is within 1 kcal/mol. In both the Fucα and the Galα moieties, the secondary hydroxyls preferentially assume the clockwise orientation. Looking at the central Galβ moiety, this pattern does usually not occur,

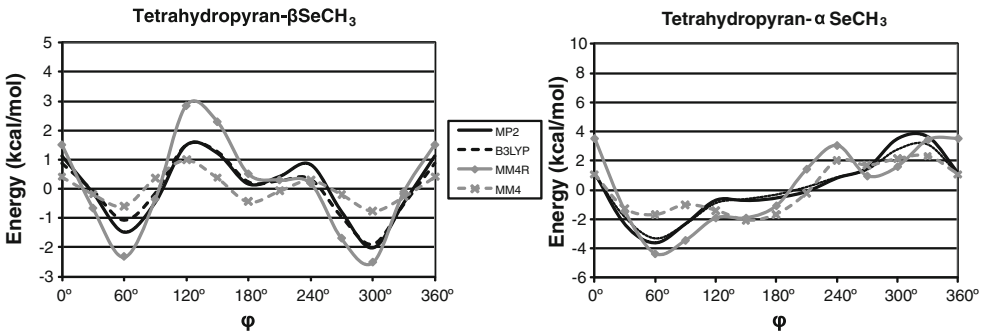

Fig. 2 Energy profiles for the φ torsion of tetrahydro-2H-pyran substituted with Se-CH_3 groups at β and α positions. In the graphs the torsional profiles calculated with the standard MM4 force field, MM4 with the new parameter set (MM4R), the B3LYP/6-311+G(d) theory, and the MP2/6-311++G(2d,2p) theory are shown for direct comparison

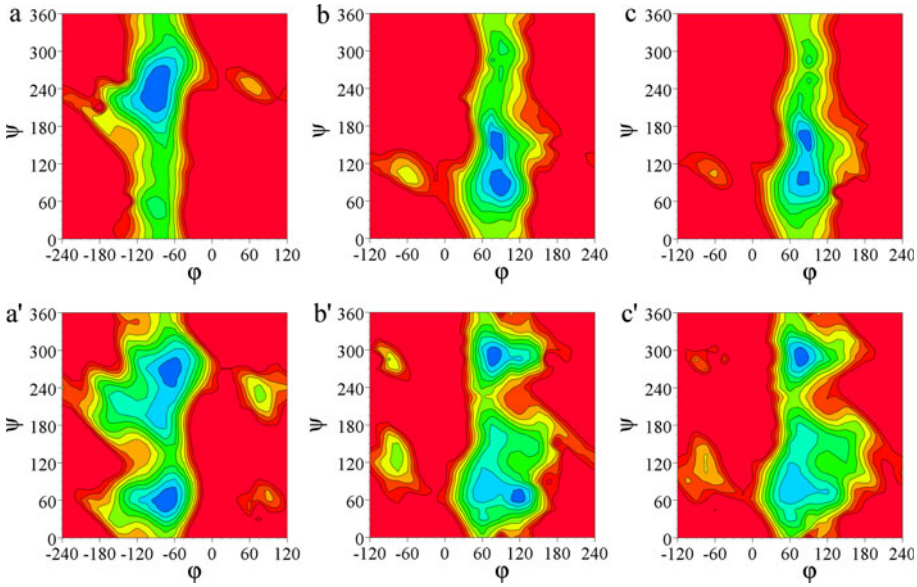

Fig. 3 Adiabatic maps for the disaccharide units of the histo-blood group ABH antigens and derivatives. Contour levels are drawn at steps of 1 kcal/mol with high- to low-energy regions colored from red to blue gradually. **a** α-L-Fuc*p*-(1→2)-*β*-D-Gal*p*; **a′** α-L-Fuc*p*-(1–Se→2)-*β*-D-Gal*p*; **b** α-D-Gal*p*NAc-(1→3)-*β*-D-Gal*p*; **b′** α-D-Gal*p*NAc-(1–Se→3)-*β*-D-Gal*p*; **c** α-D-Gal*p*-(1→3)-*β*-D-Gal*p*; **c′** α-D-Gal*p*-(1–Se→3)-*β*-D-Gal*p*

with hydroxyls not acquiring a well-ordered configuration after minimization.

In aggregate, the Se-substitution increased the overall flexibility and enabled a secondary conformational population in all linkages. Se-substitutions thus appear to have consequences similar to the S-incorporations reported previously [9].

Trisaccharide derivatives

The natural A- and B-type trisaccharides have a single well-defined conformation with α-L-Fuc*p*-(1→2)-*β*-D-Gal*p* at $\varphi/\psi \approx -75°/-105°$ and α-D-Gal*p*[NAc]-(1→3)-*β*-D-Gal*p* at $\varphi/\psi \approx 90°/75°$, whereas S-derivatives show different conformational properties in distinct cases [9]. Adiabatic energy maps were calculated for the Se-derivatives A3-Se/O, A3-O/Se, A3-Se/Se, B3-Se/O, B3-O/Se, B3-Se/Se and for the seleno/thio-derivatives A3-Se/S, A3-S/Se, B3-Se/S, B3-S/Se. The data are summarized in Table 3, the adiabatic energy maps are shown in Fig. 4 for the A-type structures and in Fig. 5 for B-type structures. The most relevant conformers for the A-type trisaccharides are shown in Fig. 6.

The compounds obtained by a single O → Se substitution prefer the conformation of the O-glycosidic compounds, while the glycosidic linkages of the compounds with two heterosubstitutions can adopt both conformations as characteristic of their disaccharide moieties. Interestingly, secondary conformers often have similar—if not more favorable—energy values, but acquire a relatively smaller population density. This is due to the decreased level of flexibility and to steric clashes between the fucose and the terminal galactose moieties that arise after small changes in the glycosidic angles of the minima not present in the natural saccharides.

The main differences between the A- and B-type epitopes are found in the α-D-Gal*p*NAc-(1-S→3)-*β*-D-Gal*p* and α-D-Gal*p*NAc-(1-Se→3)-*β*-D-Gal*p* glycosidic linkages, where conformations with $\varphi/\psi \approx 120°/75°$ show a hydrogen bond contact between the hydrogen of the amide and the O4 oxygen of the central *β*-galactose moiety, and in conformations with $\varphi/\psi \approx 75°/-60°$, where the hydrogen of the amide interacts with the O2 of the fucose residue. These results suggest interesting differences between the conformational properties of A3-Se/Se, A3-Se/S, and their B analogues. Having hereby completed the conformational analysis of the selenides, we next proceeded to study their behavior in docking.

Dockings

An initial comparison of the docking scores of thioglycosides between the OPLS2001 and OPLS2005 force fields showed an RMSE of 0.94 kcal/mol, which decreases to 0.61 kcal/mol when models that did not dock in the same site were excluded. Differences in the relative energies and ranking of the ligands were observed: Pearson's r value is 0.94 and Spearman's rank correlation coefficient is 0.79. The reasonable agreement between OPLS2001 and OPLS2005 was in line with previous findings [25] and justifies the selection of OPLS2005, which was required to process the seleno-compounds.

In this final step, the correlation between the experimental binding energies and the Glide XP scores (herein considered dimensionless although designed to represent kcal/mol) was investigated. The Glide XP scores for the binding of FcsSBP with A3 (−13.46) and B3 (−14.57) appeared to overestimate the experimental binding energies,

Table 3 Characterization of the energetically favorable conformations calculated for the trisaccharide moieties of histo-blood group ABH antigens and their derivatives

Compound	Fuc-(1-X→2)-Gal	GalNAc-(1-X→3)-Gal	Energy	Population (%)
A3-O/O	−75°/−105°	90°/75°		90.9
A3-O/Se	−105°/−165°	120°/150°		29.5
	−75°/−120°	75°/75°	+0.12	56.6
	−90°/−150°	105°/−60°	+1.48	4.3
A3-Se/O	−75°/−120°	90°/90°		90.9
	−60°/75°	105°/75°	+1.66	2.5
A3-Se/Se	−75°/−135°	75°/−60°		40.0
	−60°/75°	120°/75°	+0.95	7.9
	−75°/−135°	75°/90°	+1.07	31.9
	−105°/−60°	45°/60°	+1.90	2.6
A3-S/Se	−75°/75°	120°/75°		57.9
	−120°/−165°	75°/−90°	+1.07	4.7
	−75°/−120°	75°/90°	+1.30	20.5
A3-Se/S	−60°/−120°	75°/−60°		30.5
	−60°/75°	105°/75°	+0.37	38.9
	−75°/−105°	75°/90°	+1.88	11.1
	−75°/60°	90°/−75°	+1.88	2.7
A3-O/S	−90°/−120°	75°/−60°		89.0
	−90°/−120°	105°/75°	+2.38	3.7
A3-S/O	−75°/−135°	90°/90°		94.0
A3-S/S	−75°/−135°	90°/105°		42.3
	−75°/−120°	75°/−60°	+0.31	19.6
	−120°/60°	60°/75°	+0.86	12.8
	−90°/60°	75°/−60°	+0.91	13.0
Compound	Fuc-(1-X→2)-Gal	Gal-(1-X→3)-Gal	Energy	Population (%)
B3-O/O	−75°/−105°	90°/75°		91.4
	−135°/−150°	150°/135°	+1.24	1.6
B3-O/Se	−75°/−120°	60°/75°		65.3
	−105°/−165°	120°/150°	+0.03	27.4
B3-Se/O	−60°/−105°	90°/90°		94.1
	−75°/75°	105°/90°	+1.78	1.6
B3-Se/Se	−75°/−135°	60°/75°		60.3
	−75°/−135°	75°/−60°	+0.21	19.4
	−105°/60°	45°/60°	+0.76	7.4
	−75°/60°	75°/−60°	+1.31	2.9
B3-S/Se	−120°/60°	60°/75°		31.9
	−75°/−120°	60°/90°	+0.52	53.7
	−90°/60°	75°/−75°	+1.87	1.6
B3-Se/S	−60°/75°	105°/75°		31.3
	−75°/−105°	75°/90°	+0.57	41.6
	−75°/−120°	90°/−60°	+1.06	6.9
	−75°/60°	90°/−75°	+1.25	5.3
	−60°/−105°	120°/150°	+1.29	1.9
B3-O/S	−75°/−135°	75°/−60°		60.4
	−75°/−120°	90°/75°	+0.82	33.3
B3-S/O	−75°/−135°	90°/90°		90.4

Table 3 continued

Compound	Fuc-(1-X→2)-Gal	Gal-(1-X→3)-Gal	Energy	Population (%)
B3-S/S	−75°/−135°	90°/105°		53.4
	−90°/−105°	75°/−60°	+0.91	21.9
	−120°/60°	60°/75°	+1.16	6.9
	−90°/60°	75°/−60°	+1.55	4.4

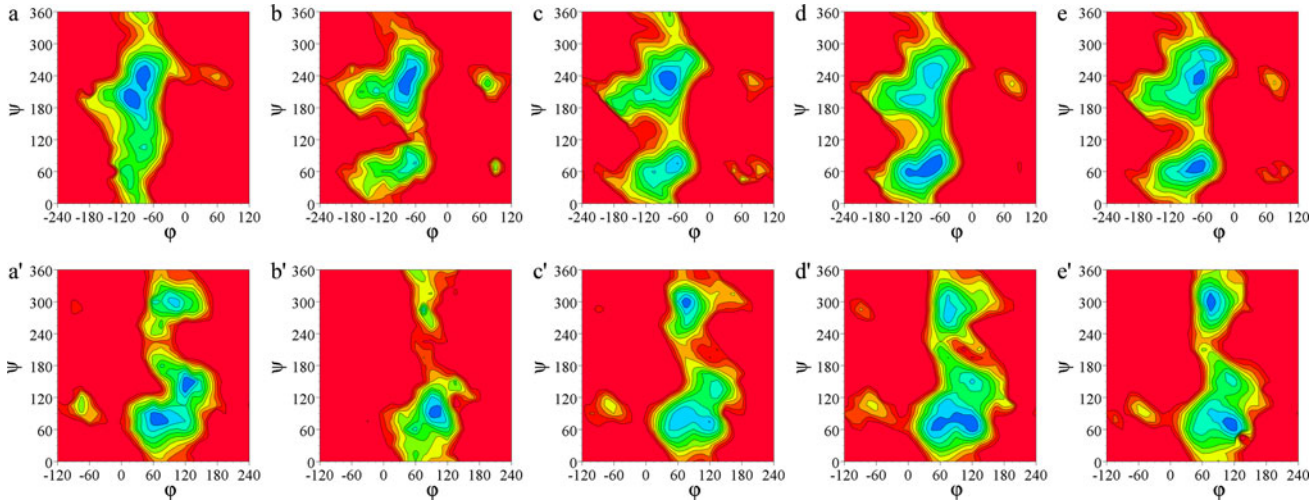

Fig. 4 Adiabatic maps for the glycosidic linkages in the histo-blood group A-trisaccharide selenoglycosides. **a** A3-O/Se α-L-Fuc*p*-(1→2)-β-D-Gal*p*; **a′** A3-O/Se α-D-Gal*p*NAc-(1-Se→3)-β-D-Gal*p*; **b** A3-Se/O α-L-Fuc*p*-(1-Se→2)-β-D-Gal*p*; **b′** A3-Se/O α-D-Gal*p*-NAc-(1→3)-β-D-Gal*p*; **c** A3-Se/Se α-L-Fuc*p*-(1–Se→2)-β-D-Gal*p*; **c′** A3-Se/Se α-D-Gal*p*NAc-(1-Se→3)-β-D-Gal*p*; **d** A3-S/Se α-L-Fuc*p*-(1–S→2)-β-D-Gal*p*; **d′** A3-S/Se α-D-Gal*p*NAc-(1–Se→3)-β-D-Gal*p*; **e** A3-Se/S α-L-Fuc*p*-(1–Se→2)-β-D-Gal*p*; **e′** A3-Se/S α-D-Gal*p*NAc-(1–S→3)-β-D-Gal*p*

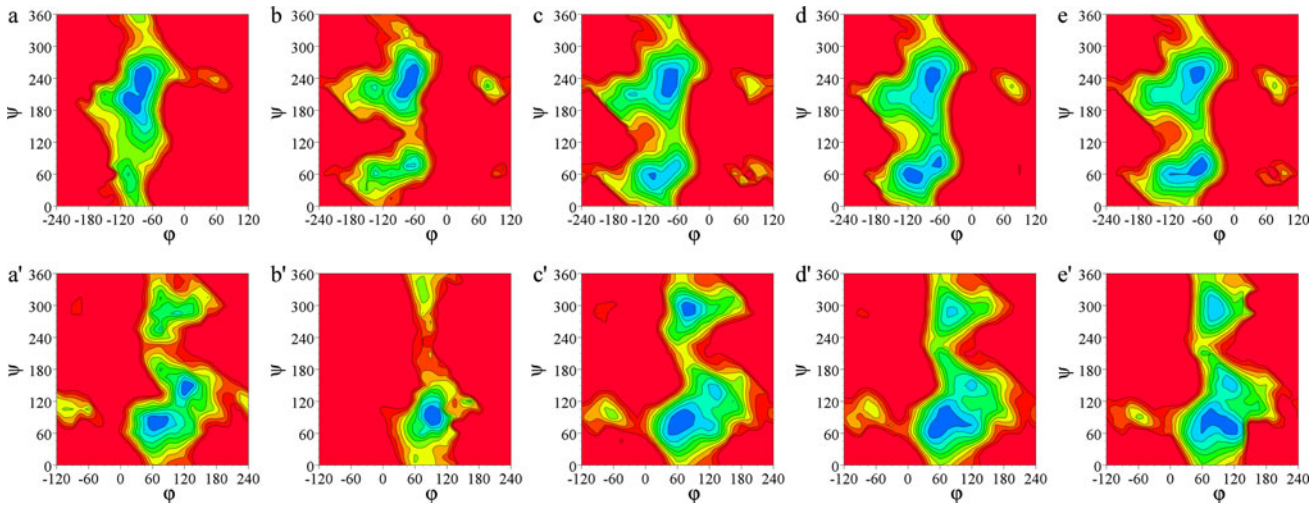

Fig. 5 Adiabatic maps for the glycosidic linkages in the histo-blood group B-trisaccharide selenoglycosides. **a** B3-O/Se α-L-Fuc*p*-(1→2)-β-D-Gal*p*; **a′** B3-O/Se α-D-Gal*p*-(1-Se→3)-β-D-Gal*p*; **b** B3-Se/O α-L-Fuc*p*-(1-Se→2)-β-D-Gal*p*; **b′** B3-Se/O α-D-Gal*p*-(1→3)-β-D-Gal*p*; **c** B3-Se/Se α-L-Fuc*p*-(1–Se→2)-β-D-Gal*p*; **c′** B3-Se/Se α-D-Gal*p*-(1-Se→3)-β-D-Gal*p*; **d** B3-S/Se α-L-Fuc*p*-(1–S→2)-β-D-Gal*p*; **d′** B3-S/Se α-D-Gal*p*-(1–Se→3)-β-D-Gal*p*; **e** B3-Se/S α-L-Fuc*p*-(1–Se→2)-β-D-Gal*p*; **e′** B3-Se/S α-D-Gal*p*-(1–S→3)-β-D-Gal*p*

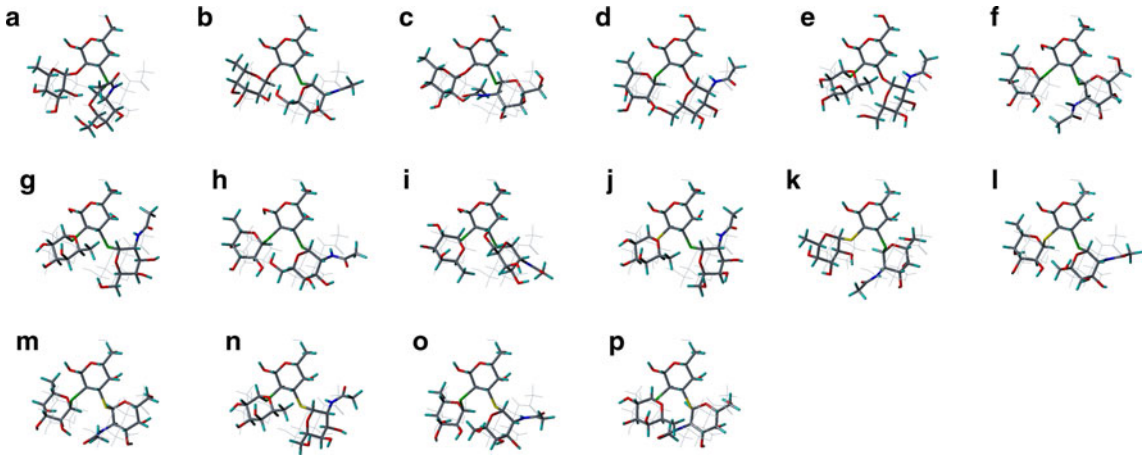

Fig. 6 Superimpositions on the Gal moiety of the energetically most favorable conformations of histo-blood group A derivatives containing Se-glycosidic linkages (depicted in *green*). The preferred conformation of the natural saccharide is shown in *gray lines*. Multiple conformers are sorted by energy, the global minimum listed first. **a–c** A3-O/Se; **d–e** A3-Se/O; **f–i** A3-Se/Se; **j–l** A3-S/Se; **m–p** A3-Se/S

which are −8.24 kcal/mol for A3 and −8.17 kcal/mol for B3 [31]. For the complex of the galectin with the extended H2 (Fucα1-2Galβ1-4GlcNAc) and B2 (Galα1-3Galβ1-4GlcNAc) trisaccharides, the docking scores of −5.97 and −6.33 are in reasonable agreement with the binding energies of −7.27 and −8.05 kcal/mol, respectively [32]. In the case of the GH98CBM51 domain, Glide failed to dock the extended A3 (Fucα1-2[GalNAcα1-3]Galβ1-4GlcNAcβ1CH_2-CH_2-N_3) and B3 (Fucα1-2[Galα1-3]Galβ1-4GlcNAcβ1CH_2-CH_2-N_3) tetrasaccharides, for which experimental binding energies of −4.88 and −4.23 kcal/mol, respectively, are available [33]. However, the trisaccharides B3-O/O and A3-O/O docked successfully with the protein. Although these scores (−9.47 for A3 and −9.05 for B3) obviously overestimated the binding energy, the difference between the A- and B-type epitopes is in agreement with the experimental data. For VAA, the ranking of the Glide XP scores (B2: −8.46, H2: −7. 97, A2: −6.84 and B3: −8.72) is in accord with the IC_{50}-values for blocking binding of the biotinylated VAA to a lactosylated neoglycoconjugate with B2 (0.04 mM), H2 (0.77 mM), A2 (2.6 mM) and B3 (0.7 mM) [22]. Furthermore, the preference for B-type epitopes of MOA [34] is reflected in the Glide scores for A3 and B3. Finally, Glide estimated the binding score between PA-IL and the extended B2 disaccharide to be −5.18, in reasonable agreement with the experimental value of −5.7 kcal/mol [35]. This comparability between results of dockings and experiments suggests that the computational approach is able to yield a satisfactory accordance. Of course, the docking scores should nonetheless be considered as semi-quantitative. From these results, the application to estimate bioactivity of the substituted glycosides appears to be validated at the given level of accordance.

Looking at the results of the OPLS2005 dockings in detail, the substituted histo-blood group ABH antigens and their natural counterparts are presented in Table 4, the three-dimensional structures of the best binders are depicted in Fig. 7. Additionally, the visualization of the patterns of protein–ligand interactions for the natural ligands and the best selenoglycoside binders are provided as supplementary material. In these dockings, the best binders are accommodated at the same site as the natural ligand. With the exception of AAA, the orientation of the key residue is always conserved. In the deep binding pocket of FcsSBP, the conformation of B3-S/O is nearly identical to the co-crystallized trisaccharides A3-O/O and B3-O/O. In CLG2, the glycosidic linkage of the flexible H2-Se moiety is arrested in a conformation that would be very unfavorable in the natural sugar. Energetic compensation is attributable to favorable interactions between the protein and the fucose moiety. In CBM51, the O → Se substitution in A3-O/Se and a small change in the glycosidic linkage of the H2 moiety allow a more favorable position of the fucose moiety. The binding pattern for VAA and the co-crystallized galactose residue is only found with respect to the position of the terminal galactose in the more flexible compounds that have a O → Se substitution in the strongly bioactive B2 moiety. Other compounds bind with a different pattern, in which the Gal moiety is rotated so that O1 of the ligand has roughly the position of C6 of the crystal. Experimental studies mapping the glycan-protein contact, e.g. by crystallography or saturation transfer difference NMR spectroscopy, used in compound library screening for VAA [36], will be needed to establish the basis for further comparison. Looking at the MOA complex, the flexible A3-Se/Se can be docked with favorable positioning of the N-acetylgalactosamine residue, while preserving the optimal position of

Table 4 Docking scores of histo-blood group ABH derivatives obtained with OPLS2005

Compound	FcsSBP	CGL2	CBM51	VAA	MOA	PA-IL	UEA-I	AAA
A3-O/O	−13.46	−7.86	−9.47	−7.18	−7.96	−5.59	−7.98	−6.75
A3-O/S	−15.20	−7.83	−9.97	−7.30	−9.31	−6.64	−8.62	−6.69
A3-O/Se	−13.65	−5.58	−10.65	−7.77	−8.84	−5.89	−8.52	−6.16
A3-Se/O	−13.98	−5.59	−8.04	−7.69	−8.23	−5.19	−7.57	−6.65
A3-Se/S	−14.65	−8.53	−9.27	−7.07	−8.17	−7.06	−8.28	−7.52
A3-Se/Se	−11.95	−6.83	−7.97	−7.31	−10.22	−5.39	−9.71	−7.68
A3-S/O	−14.49	−6.94	−6.95	−7.59	−7.83	−6.37	−8.85	−7.19
A3-S/S	−14.93	−7.12	−9.84	−8.04	−9.02	−6.31	−8.67	−7.76
A3-S/Se	−13.30	−8.38	−8.72	−5.99	−8.49	−5.47	−8.57	−7.69
B3-O/O	−14.57	−8.09	−9.05	−8.72	−8.38	−6.56	−9.30	−8.02
B3-O/S	−15.00	−7.59	−9.28	−8.45	−8.68	−6.12	−9.43	−7.40
B3-O/Se	−13.59	−6.94	−9.19	−8.76	−8.76	−6.22	−9.56	−6.89
B3-Se/O	−15.31	−8.26	−8.12	−8.50	−8.99	−6.88	−9.63	−7.92
B3-Se/S	−15.11	−7.37	−8.87	−8.34	−9.40	−6.97	−8.95	−9.24
B3-Se/Se	−14.61	−6.47	−8.38	−9.71	−8.84	−7.11	−8.63	−8.28
B3-S/O	−15.39	−6.10	−8.47	−9.13	−7.89	−6.80	−8.76	−7.80
B3-S/S	−14.89	−8.44	−8.96	−8.53	−9.02	−7.02	−9.65	−8.71
B3-S/Se	−14.48	−6.68	−7.76	−9.23	−9.54	−7.04	−9.94	−7.17
H2-O	−11.40	−6.18	−6.81	−7.97	−6.91	−6.24	−7.48	−6.82
H2-S	−10.91	−7.29	−7.89	−6.90	−9.77	−5.22	−8.39	−7.91
H2-Se	−11.15	−5.86	−7.38	−6.35	−8.63	−6.15	−8.12	−6.47
A2-O	−11.35	−5.54	−6.39	−6.84	−8.21	−3.57	−5.64	−6.94
A2-S	−12.03	−6.88	−6.53	−6.87	−8.68	−5.81	−7.84	−6.61
A2-Se	−9.90	−7.55	−8.77	−6.47	−8.17	−5.52	−6.99	−6.91
B2-O	−11.24	−7.06	−8.43	−8.46	−6.51	−5.47	−6.97	−7.62
B2-S	−12.44	−7.13	−7.80	−7.54	−7.32	−6.45	−7.85	−8.17
B2-Se	−11.81	−7.38	−7.88	−8.53	−7.05	−6.44	−7.43	−7.15

the central galactose unit of the natural saccharide. In PA-IL, the galactose moieties of most ligands appear to coordinate with the calcium ion as in the crystal, whereas the other sugar moieties generally assume different poses (as does B3-Se/Se in Fig. 7). This reflects the ability of PA-IL to bind a variety of glycans presenting a terminal α-galactose unit [37]. In UEA-I, the conformation of the fucose moiety is well conserved in the deep binding pocket, while the enhanced flexibility of the S- and Se- glycosidic linkages enables the terminal galactose moiety of B3-S/Se to be engaged in more favorable interactions with the protein. On the other hand, the favored docked conformations in AAA present the same orientation of the fucose residue only for the H-disaccharide and its derivative H2-Se, while in the best predicted binder, i.e. B3-Se/S, the terminal galactose moiety interacts with the same residues as the fucose unit in the crystal, but with a different topology (Fig. 7). As noted above, this result gives a clear direction to further experiments on mapping contacts.

As expected, derivatives with conformational properties similar to the natural compounds often have similar docking scores. The docking results indicate bioactivity of selenoglycosides as ligands for lectins. The glycan derivatives with new conformational properties may even exploit different interaction topologies and improve binding affinities.

Conclusions

The parameterization of MM4R for selenoglycosidic linkages was made possible through ab initio calculations. It yielded an obvious increase in agreement to quantum mechanics-based calculations, with both the MP2 and Becke3LYP theories, relative to the default estimated parameters. Thus, we have established a basis for MM4-based computational studies on saccharides with selenium replacing oxygen. The conformational studies on the histo-blood group ABH saccharides indicated that selenodisaccharides behave more like S-substituted compounds than O-glycosides and that trisaccharides with selenium in the glycosidic linkage can have conformational properties similar to the O-glycosides.

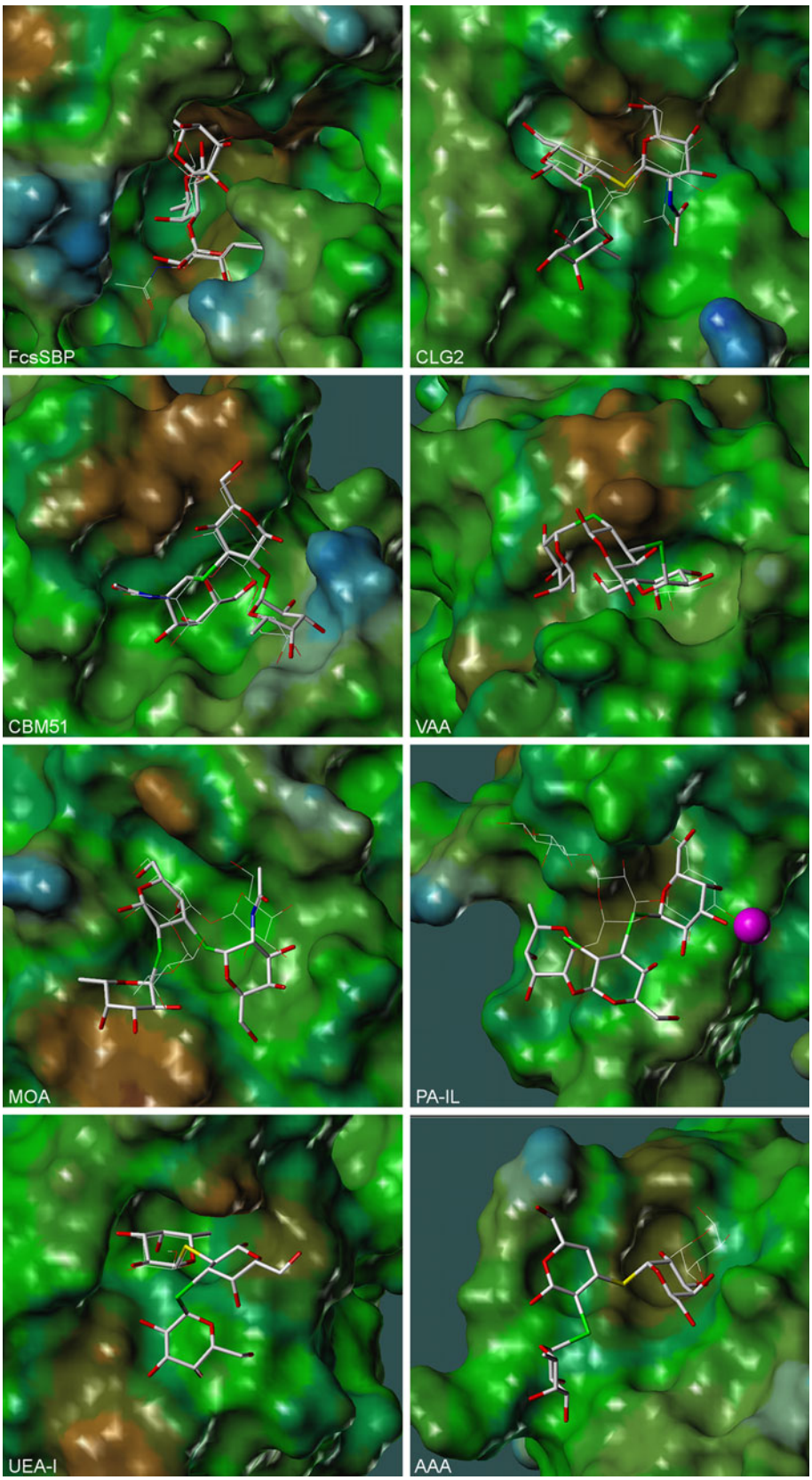

Fig. 7 Docking poses of the best histo-blood group ABH-type ligands (in capped sticks) in the cases of the carbohydrate-binding proteins FcsSBP, CLG2, CBM51, VAA, MOA, PA-IL, UEA-I, and AAA. Co-crystallized natural ABH-type structures are shown in *lines*, the calcium ion in PA-IL as a violet sphere and the solvent-accessible surface of the protein is colored according to lipophilicity (*blue* hydrophilic, *brown* lipophilic)

The docking results indicate that S-, Se- and S-/Se-derivatives can share the binding properties of their natural counterparts. Variations in bond lengths/angles (see Table 1), flexibility, lipophilicity and conformational properties may even enhance affinity to lectins. The study of binding of S-/Se- glycosides to lectins will thus very likely unravel new details on affinity generation [38]. Along this line, it also appears that Se- and/or S-derivatives of the histo-blood group ABH determinants could serve as non-hydrolysable mimetics of the O-glycosides. Therefore, they warrant detailed studies on endogenous adhesion/growth-regulatory effectors and bacterial/viral agglutinins [39–42]. Combined with a clustered display, e.g. on cyclodextrin, calixarene, or glycocyclophane [43–46], and proper linker tailoring [47], these compounds may prove therapeutically active.

Supplementary material

The MM4R parameters have been attached and are formatted according to the syntax of the MM4 file for replacement parameters, MM4.PARA. For each protein, LIGPLOT diagrams of the sets of interactions with the natural ligand and the best binder from the panel of studied glycan derivatives have also been included.

Acknowledgments Financial support from the Swedish Medical Research Council (K2000-03x-00006-36A), the research initiative LMU*excellent,* an EC Marie Curie Research Training Network grant (MRTN-CT-2005-019561) and Biognos AB (Göteborg) is gratefully acknowledged. Also, access to the Linux cluster at the Institute of Biomedicine, Gothenburg University, Sweden is gratefully acknowledged. We are also grateful to the reviewers of the manuscript for their expert input.

References

1. Gabius H-J (ed) (2009) The sugar code: fundamentals of glycosciences. Wiley, Weinheim
2. Gabius H-J (1988) Angew Chem 100:1321–1330 Gabius H-J (1988) Angew Chem Int Ed Engl 27:1267–1276
3. Gabius H-J (2008) Biochem Soc Transact 36:1491–1496
4. Oscarson S (2009) The chemist's way to synthesize glycosides. In: Gabius H-J (ed) The sugar code: fundamentals of glycosciences. Wiley, Weinheim, pp 31–51
5. Osborn HMI, Turkson A (2009) Sugars as pharmaceuticals. In: Gabius H-J (ed) The sugar code: fundamentals of glycosciences. Wiley, Weinheim, pp 469–483
6. Asensio J, Espinosa J, Dietrich H, Schmidt R, Martín-Lomas M, André S, Gabius H-J, Jiménez-Barbero J (1999) J Am Chem Soc 121:8995–9000
7. Ahmad N, Gabius H-J, Kaltner H, André S, Kuwabara I, Liu F-T, Oscarson S, Norberg T, Brewer CF (2002) Can J Chem 80:1096–1104
8. André S, Pei Z, Siebert H-C, Ramström O, Gabius H-J (2006) Bioorg Med Chem 14:6314–6326
9. Strino F, Lii J-H, Gabius H-J, Nyholm P-G (2009) J Comput Aided Mol Des 23:845–852
10. Buts L, Loris R, De Genst E, Oscarson S, Lahmann M, Messens J, Brosens E, Wyns L, De Greve H, Bouckaert J (2003) Acta Crystallogr D Biol Crystallogr 59:1012–1015
11. Allinger NL, Chen K-H, Lii J-H, Durkin KA (2003) J Comput Chem 24:1447–1472
12. Lii J-H, Chen K-H, Durkin KA, Allinger NL (2003) J Comput Chem 24:1473–1489
13. Lii J-H, Chen K-H, Grindley TB, Allinger NL (2003) J Comput Chem 24:1490–1503
14. Lii J-H, Chen K-H, Allinger NL (2003) J Comput Chem 24:1504–1513
15. Lii J-H, Chen K-H, Johnson GP, French AD, Allinger NL (2005) Carbohydr Res 340:853–862
16. Lii J-H, Chen K-H, Allinger NL (2004) J Phys Chem A 108:3006–3015
17. Beecher JF (1966) J Mol Spectrosc 21:414–424
18. Lii J-H, Ma B, Allinger NL (1999) J Comput Chem 20:1593–1603
19. Strino F, Nahmany A, Rosén J, Kemp GJL, Sá-correia I, Nyholm P-G (2005) Carbohydr Res 340:1019–1024
20. Nahmany A, Strino F, Rosén J, Kemp GJL, Nyholm P-G (2005) Carbohydr Res 340:1059–1064
21. Lii J-H, Allinger NL (2008) J Phys Chem A 112:11903–11913
22. Galanina OE, Kaltner H, Khraltsova LS, Bovin NV, Gabius H-J (1997) J Mol Recognit 10:139–147
23. Jiménez M, André S, Barillari C, Romero A, Rognan D, Gabius H-J, Solís D (2008) FEBS Lett 582:2309–2312
24. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) J Phys Chem B 105:6474–6487
25. Stortz CA, Johnson GP, French AD, Csonka GI (2009) Carbohydr Res 344:2217–2228
26. Heiden W, Moeckel G, Brickmann J (1993) J Comput Aided Mol Des 7:503–514
27. Wallace AC, Laskowski RA, Thornton JM (1995) Protein Eng Des Sel 8:127–134
28. Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM (1997) Trends Biochem Sci 22:488–490
29. Traar P, Belaj F, Francesconi KA (2004) Aust J Chem 57:1051–1053
30. Bondi A (1964) J Phys Chem 68:441–451
31. Higgins MA, Abbott DW, Boulanger MJ, Boraston AB (2009) J Mol Biol 388:299–309
32. Walser PJ, Haebel PW, Künzler M, Sargent D, Kües U, Aebi M, Ban N (2004) Structure 12:689–702
33. Gregg KJ, Finn R, Abbott DW, Boraston AB (2008) J Biol Chem 283:12604–12613
34. Teneberg S, Alsén B, Ångström J, Winter HC, Goldstein IJ (2003) Glycobiology 13:479–486
35. Blanchard B, Nurisso A, Hollville E, Tétaud C, Wiels J, Pokorná M, Wimmerová M, Varrot A, Imberty A (2008) J Mol Biol 383:837–853
36. Ribeiro JP, André S, Cañada FJ, Gabius H-J, Butera AP, Alves RJ, Jiménez-Barbero J (2010) ChemMedChem 5:415–419
37. Chen C-P, Song S-C, Gilboa-Garber N, Chang KSS, Wu AM (1998) Glycobiology 8:7–16
38. Solís D, Romero A, Menéndez M, Jiménez-Barbero J (2009) Protein-carbohydrate interactions: basic concepts and methods for analysis. In: Gabius H-J (ed) The sugar code: fundamentals of glycosciences. Wiley, Weinheim, pp 233–245
39. Wu AM, Wu JH, Liu J-H, Singh T, André S, Kaltner H, Gabius H-J (2004) Biochimie 86:317–326
40. Gabius H-J (2006) Crit Rev Immunol 26:43–79
41. Villalobo A, Nogales-González A, Gabius H-J (2006) Trends Glycosci Glycotechnol 18:1–37

42. Holgersson J, Gustafsson A, Gaunitz S (2009) Bacterial and viral lectins. In: Gabius H-J (ed) The sugar code: fundamentals of glycosciences. Wiley, Weinheim, pp 279–300
43. André S, Kaltner H, Furuike T, Nishimura S-I, Gabius H-J (2004) Bioconjug Chem 15:87–98
44. André S, Sansone F, Kaltner H, Casnati A, Kopitz J, Gabius H-J, Ungaro R (2008) ChemBioChem 9:1649–1661
45. Chabre YM, Roy R (2009) The chemist's way to prepare multivalency. In: Gabius H-J (ed) The sugar code: fundamentals of glycosciences. Wiley, Weinheim, pp 53–70
46. Leyden R, Velasco-Torrijos T, André S, Gouin S, Gabius H-J, Murphy PV (2009) J Org Chem 74:9010–9026
47. André S, Giguère D, Dam TK, Brewer CF, Gabius H-J, Roy R (2010) New J Chem 34:2229–2240