J Comput Aided Mol Des (2014) 28:951–960
DOI 10.1007/s10822-014-9774-7

Engineering the *Pseudomonas aeruginosa* II lectin: designing mutants with changed affinity and specificity

Zdeněk Kříž · Jan Adam · Jana Mrázková · Petros Zotos · Thomais Chatzipavlou · Michaela Wimmerová · Jaroslav Koča

Received: 15 November 2013 / Accepted: 2 July 2014 / Published online: 12 July 2014

Abstract This article focuses on designing mutations of the PA-IIL lectin from *Pseudomonas aeruginosa* that lead to change in specificity. Following the previous results revealing the importance of the amino acid triad 22–23–24 (so-called specificity-binding loop), saturation *in silico* mutagenesis was performed, with the intent of finding mutations that increase the lectin's affinity and modify its specificity. For that purpose, a combination of docking, molecular dynamics and binding free energy calculation was used. The combination of methods revealed mutations that changed the performance of the wild-type lectin and its mutants to their preferred partners. The mutation at position 22 resulted in 85 % in inactivation of the binding site, and the mutation at 23 did not have strong effects thanks to the side chain being pointed away from the binding site. Molecular dynamics simulations followed by binding free energy calculation were performed on mutants with promising results from docking, and also at those where the amino acid at position 24 was replaced for bulkier or longer polar chain. The key mutants were also prepared in vitro and their binding properties determined by isothermal titration calorimetry. Combination of the used methods proved to be able to predict changes in the lectin performance and helped in explaining the data observed experimentally.

Keywords Lectin · Carbohydrate · Mutagenesis · Docking · Molecular dynamics

Zdeněk Kříž, Jan Adam and Jana Mrázková have equally contributed to this work.

Z. Kříž · J. Adam · M. Wimmerová (✉) · J. Koča (✉)
CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic
e-mail: michaw@chemi.muni.cz

J. Koča
e-mail: jaroslav.koca@ceitec.muni.cz

Z. Kříž · J. Adam · J. Mrázková · P. Zotos · T. Chatzipavlou · M. Wimmerová · J. Koča
National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic

J. Mrázková · M. Wimmerová
Department of Biochemistry, Faculty of Science, Masaryk University, Brno, Czech Republic

P. Zotos · T. Chatzipavlou
Division of Pharmaceutical Chemistry, School of Pharmacy, National and Capodistrian University of Athens, Athens, Greece

Introduction

For a long time, the saccharides were considered to be primarily a source of energy for all living systems. Exploring their abilities to form complex and very variable structures at the basis of countless possibilities of glycosidic intermolecular linking came particularly in the focus during the last decades of the twentieth century, emphasizing the role of saccharides in the biological processes that take place in the living systems.

Lectins are proteins that are ubiquitous in the realm of the living organisms [1]. They display a characteristic ability to bind carbohydrates specifically, but they have no catalytic activity overall. The unifying feature of this large family of proteins is the ability to decode information hidden in the saccharide code [2]. It has been shown that the lectin–saccharide interaction plays a very important role in many biological functions, both in normal and pathological situations. The lectin–saccharide interaction serves in processes of cell adhesion, cell recognition,

immune response, tumor genesis and endotoxin binding. As for saccharide binding, various architectures of the saccharide binding site have been reported. The binding sites have been found and described in numerous possible protein folds. One of the most common ways used by the lectins for binding a saccharide ligand is combining and balancing the contribution of hydrogen bonds with stacking of aliphatic rings of the monosaccharides and aromatic rings of side chains of some amino acids [3].

Another major type of lectin–saccharide binding is based on the presence of bridging cations (usually calcium) in the binding site. These so-called C-type lectins were reported to play important roles in the process of cell adhesion [4–6]. Despite the fact that there are variations in the structures of C-type lectins, the amino acids of the binding site are conserved. The binding itself is usually realized by a combination of two hydroxyl oxygen atoms coordinated to the calcium ion, combined with hydrogen bonds between the saccharide and the surrounding amino acids [7]. Interesting calcium dependent lectins were isolated and reported from several opportunistic bacteria, first and foremost the *Pseudomonas aeruginosa* [8]. This bacterium is responsible for causing high mortality and morbidity rates in patients with the cystic fibrosis disease. These lectins often play a crucial role in the host recognition and biofilm forming during the colonization of the host tissue [8–11]. A possible role in pilus biogenesis and proteolysis was also reported [12]. *Pseudomonas aeruginosa* produces two lectins that are related to its virulence, PA-IL [13] and PA-IIL [14]. The PA-IL lectin is a tetrameric, galactose-specific lectin with one calcium ion in the binding site and a binding mode that is analogical to the known C-type lectins. The PA-IIL lectin is also a tetrameric lectin but it exhibits an unusually high (micromolar, which is not common for protein/carbohydrate interactions) affinity to monosaccharides, namely L-fucose, L-galactose, and D-arabinose, and their methylated forms [15], and a moderate affinity to D-fructose or D-mannose. All monosaccharides must have either an L-galacto or a D-manno configuration that contains two equatorial and one axial hydroxyl groups when in the $^{1}C_{4}$ and $^{4}C_{1}$ pyranose conformation, respectively (Fig. 1a, b). This lectin displays a previously unobserved sugar binding site architecture with two close calcium ions (3.7 Å) embedded in an acidic pocket. The binding is realized by the completion of the coordination spheres of the ions by three oxygen atoms from the hydroxyl groups of the saccharide [14]. In this case, both calcium ions participate in the saccharide coordination, a fact that was previously unobserved in the lectin realm. Aside from the binding via calcium coordination, the surrounding amino acids interact with other atoms of the saccharide by hydrogen bonds as well as hydrophilic and hydrophobic interactions (Fig. 1c, d). The

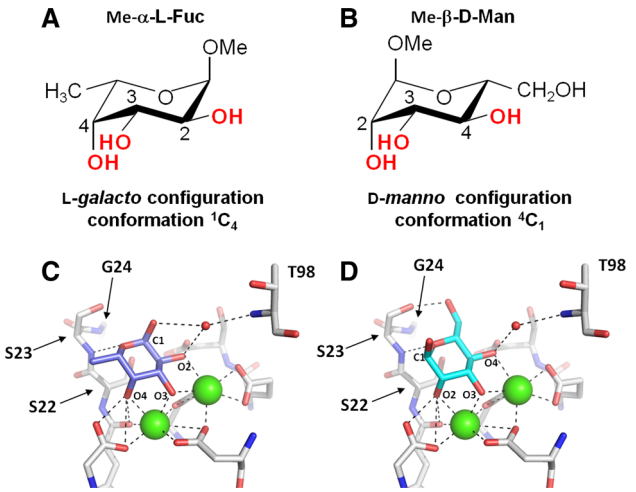

Fig. 1 *Upper panel* The stereochemical representation of methyl-α-L-fucopyranoside in the $^{1}C_{4}$ (**a**) and methyl-α-D-mannopyranoside in the $^{4}C_{1}$ (**b**) conformations required for their interaction with PA-IIL. *Lower panel* Detail of the PA-IIL binding site (in *grey*) with α-L-fucopyranose (**c**) in *blue* (PDB ID: 1GZT) and α-D-mannopyranose (**d**) in cyan (PDB ID: 1OUR). Calcium ions are represented by *green* cpk, the conserved water molecule by a *red ball*

specificity is very closely related to the composition of the site, and a slight change usually leads to remarkable changes in binding, as reported in the studies comparing the PA-IIL with its closest homologues, the CV-IIL lectin [16] from *Chromobacterium violaceum* and RS-IIL lectin [17] from *Ralstonia solanacearum*.

The dependence of the sugar preference in the PA-IIL family was described recently [16, 18, 19]. In the composition of the binding site, the prime importance is assigned to the so-called specificity-binding loop, the triad of amino acids at the positions 22, 23 and 24. The serine-to-alanine mutation at the position 22 was already revealed to be responsible for the switch of the PA-IIL preference from L–fucose to D-mannose, thanks to the fact that the hydroxyl group of serine 22 creates a hydrogen bond with aspartate 96, discriminating the mannose by denying the interaction of its C6-hydroxymethyl group, as described in [18]. The other two mutations at positions 23 (serine-to-alanine) and 24 (glycine-to-asparagine) modified the affinity slightly, but did not result in preference change.

This study is focused on applying docking procedures in combination of molecular dynamics as a tool for *in silico* prediction of specificity changes caused by the mutation of the binding loop amino acids, substantially deepening the previously published analysis [20]. All three amino acids of the specificity-binding loop, i.e., serine 22, serine 23, and glycine 24, were mutated into all 19 other possible amino acids, one at a time. Therefore, the total of fifty-seven single-point mutants were obtained and evaluated. Subtracting the three previously analyzed and described (S22A, S23A and

G24N), fifty-four new structures were prepared for investigation. For the *in silico* mutagenesis, the software MODELLER [21] was used at the basis of previous experience.

For the purpose of evaluation of the changes in binding, the method of molecular docking was chosen as the first step after mutagenesis. Applications of various docking software on recognition processes involving saccharides, such as interaction with lectins [22], enzymes [23], or immuno-based processes [24, 25] were recently described in several studies, as well as analyses of their capabilities and performance [26–28].

The docking software for this study was also chosen on the basis of previously reported performance. AutoDock3 [29] has performed the best of the tested freely available docking programs and has proven to be satisfactory for providing results that are comparable in order to experimental values for this particular type of protein–ligand system. Finally, this combination is proven to work well with the software TRITON [20], an in-house developed graphical tool for protein modelling, which was used as an interface for work with both MODELLER and AutoDock. The TRITON was also able of serving as a molecular editor and an interface for preparation, launching and subsequent analysis of the results of docking runs.

The study explores the effects of possible mutations on the 22–23–24 amino acid triad, with the intention to create mutants that display stronger affinities to their preferred binding partner than the wild-type lectin and its mutant G24N. A particular attention was devoted to two mutants with a large side chain in the position 24, where glycine was mutated into tryptophan, arginine and aspartic acid. These mutants were prepared in vitro and their binding properties were characterized by isothermal titration calorimetry.

Experimental procedures

Computational methods

Source data

The original PDB file used for the mutations was taken from the PDB database, from the structure file containing lectin complex with methyl-α-L-fucoside (PDB code 1GZT [14]). Similarly as in the previous study, only functional half of the molecule was used for the computations. The structures of the ligands were modelled by using Glycam biomolecular builder [30].

Preparation of protein files

Before any computational operations were performed at the structure files, all water molecules, ligands and cocrystallized molecules were removed from the structures. The hydrogen atoms were added by WHAT IF software [31]. The structures of single point mutants were then modeled *in silico* by the homology modeling-using program TRITON interfaced with MODELLER v. 6.2 [21].

Preparation of ligand files

The structures of the ligands were modeled by using Glycam biomolecular builder [30]. The 3-D structures of the saccharides were geometrically optimized using Hartree–Fock method at 6-31G* level, under Gaussian 03 [32]. ESP charges were calculated, and RESP charges were derived using *antechamber* program of the AMBER package [33, 34]. A resulting Sybyl mol2 file of the ligand was obtained, as a readable input file for AutoDock 3.0 under TRITON.

AutoDock v. 3.1

Molecular docking was performed by AutoDock v 3.1, a suite of docking tools optimized for performing docking of small molecules in larger receptors. For the protein, the Kollman united atom charges [35] were used, as recommended by the user manual. All the non-polar hydrogens were merged. The program uses Lamarckian genetic algorithm combined with Solis and Wets local search method for identifying the global minima at the potential hypersurface of possible conformations. The initial population was 50 randomly placed individuals, and the maximum number of energy evaluations was 5 million. Atomic maps were calculated for 48 × 48 × 48 point grid, centered at the binding site with 0.25-Å grid spacing. Docked orientations within a root-mean square deviation of 1.0 Å were clustered together and the most populated cluster was taken for further consideration. Input files for AutoDock3 with parameters mentioned above were prepared by using TRITON.

Molecular dynamics

The coordinates of the complexes obtained from docking procedures were loaded into Leap AMBER program. The Glycam force field parameters were applied on the saccharide atoms and the Duan et al. force field [36] parameters were applied on the protein atoms. The complexes were solvated by TIP3P water molecules using octahedral box with 12 Å thick water layer. The Na^+ and Cl^- ions with concentration of 0.15 M were added into box as a model of intracellular conditions at steady state. The systems were relaxed and slowly heated to simulation temperature before the production phase of the MD simulations. The first step was the minimization and MD simulation of the solvent molecules and ions. The low-

temperature (at 5 and 10 K) MD simulations of the side chains, solvent molecules and ions with decreasing restrain forces applied to backbone atoms were followed by low-temperature (at 5 and 10 K) MD simulations of the whole systems. In the final steps of the relaxation procedure, the simulation systems were slowly heated to 298.16 K for 200 ps followed by a 50 ps long MD simulation under NPT conditions.

The production phase of MD was 100 ns long for each system and was run twice with different velocity distribution using different random seed. The Pmemd program from the Amber software package was used for the simulations. The Particle-Mesh Ewald (PME) method was used for dealing with electrostatic interactions, and all simulations were performed under periodic boundary conditions in the [NpT] ensemble at 298.16 K and 1 atm using a 2 fs integration step. The SHAKE algorithm [37], with a tolerance of 10^{-5} Å, was used to fix the positions of all hydrogen atoms, and a 9.0 Å cutoff was applied to non-bonding interactions. A Berendsen thermostat was used [38].

The binding free energies were calculated using MM/GBSA method as implemented in the AMBER program package. The last 30 ns of the trajectories were used for the MM/GBSA. The salt concentration parameter with value of 0.15 M was used for calculations.

The binding free energy was calculated as average value according to Eq. (1).

$$\Delta E_{bind} = \frac{1}{N}\sum_{1}^{N} E_{complex} - (E_{receptor} + E_{ligand}) \quad (1)$$

where N is a number of snapshots included in the calculation. Energies are calculated as a combination of molecular mechanics energy and solvation energy, which is calculated using the Generalized Born approach [39].

Experimental methods

Construction of the plasmid for expression of PA-IIL G24W, G24D and G24R mutants

Previously prepared plasmid pET-25(b+)*pa2l*, containing the the full-length wild type *P. aeruginosa* PA-IIL encoding gene, was used as a template [15].

Site-directed mutagenesis was performed by PCR, using the QuikChangeTM Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's instructions. The reaction mixture (50 µl in total) consisted of 5 µl 10× *PfuUltra* HF reaction buffer (Stratagene) 1 µl dNTP Mix (10 mM each nucleotide, New England BioLabs), 1 µl of each primer (concentration of each primer: 100 ng μl^{-1}; for G24W: primer *pa2l-g24w* 5′-TTC GCC AAC TCG TCC TGG ACC CAG ACG GTG AAC-3′, primer *pa2l-g24w-anti* 5′-GTT CAC CGT CTG GGT CCA GGA CGA GTT GGC GAA-3′, 33-mer, for G24D: primer *pa2l-g24d* 5′-TCG CCA ACT CGT CCG ATA CCC AGA CGG TGA AC-3′, primer *pa2l-g24d-anti* 5′-GTT CAC CGT CTG GGT ATC GGA CGA GTT GGC GA-3′, 32-mer, and for G24R: primer *pa2l-g24r*: 5′-CGC CAA CTC GTC CCG CAC CCA GAC GGT GA-3′; primer *pa2l-g24r-anti*: 5′-TCA CCG TCT GGG TGC GGG ACG AGT TGG CG-3′; 29-mers), 1 µl plasmid pET-25(b+)*pa2l* (30 ng μl^{-1}; the same template as above) and 1 µl *PfuUltra* HF DNA polymerase (2,5 U μl^{-1}; Stratagene). The reaction conditions were: (1) 95 °C 1 min, (2) 18 cycles: 95 °C 50 s, 65 °C 50 s, 72 °C 12 min, (3) 72 °C 10 min. PCR mixture was treated by *Dpn I* restrictase (New England BioLabs) subsequently. The introduced mutation was confirmed by DNA sequencing and the *E. coli* expression cells were transformed by these new constructs. *E. coli* Tuner (DE3) strain (Novagen) was transformed by pET-25(b+)*pa2l_g24w* and *E. coli* BL21 (DE3) strain (Novagen) was transformed by pET-25(b+)*pa2l_g24r* and pET-25(b+)*pa2l_g24d.*

Expression and purification of PA-IIL mutants

E. coli Tuner (DE3) cells harboring the plasmid pET25-*pa2l_g24w*, *E. coli* BL21 (DE3) cells harboring the plasmid pET25*pa2l_g24r* and *E. coli* BL21 (DE3) cells harboring the plasmid pET25*pa2l_g24d* were cultured in 1 l of Luria broth (LB) at 37 °C. When the culture reached an optical density of 0.5–0.6 at 600 nm, iso-propyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM. Cells were harvested after 3 h incubation at 30 °C, washed and resuspended in 10 ml of the equilibrating buffer (20 mM Tris/HCl, 100 mM NaCl, 100 µM $CaCl_2$, pH 7.5). The cells were disrupted by sonication (Soniprep 150, Schoeller instruments, GB). After centrifugation at 21,000*g* for 1 h, the supernatant was purified on Mannose-Agarose (Sigma-Aldrich, USA). The PA-IIL mutant was eluted by the elution buffer (20 mM Tris/HCl, 100 mM NaCl, 100 µM $CaCl_2$, 0.1 M D-mannose, pH 7.5). The purified protein was intensively dialyzed against distilled water for 1 week, concentrated by lyophilization and kept at −20 °C.

Microcalorimetry

Isothermal titration calorimetry experiments were performed using ITC_{200} microcalorimeter (GE Healthcare). All titrations were performed in 0.1 M Tris/HCl buffer containing 3 mM $CaCl_2$, pH 7.5 at 25 °C. Aliquots of 2 µl of Me-α-L-Fuc (0.6 mM) and Me-α-D-Man (3 mM), respectively, dissolved in the same buffer were added at

3–4 min intervals to the lectin solution (0.06 and 0.3 mM, respectively) present in the calorimeter cell. Three independent titrations were performed for each ligand tested. The temperature of the cell was controlled to 25 ± 0.1 °C. Control experiments performed by injections of buffer in the protein solution and sugar in the buffer, respectively, yielded to insignificant heats of dilution. Integrated heat effects were analysed by non-linear regression using a single site-binding model (Microcal Origin 7). Fitted data yielded the association constant (K_A) and the enthalpy of binding (ΔH). Other thermodynamic parameters, i.e., changes in free energy, ΔG, and entropy, ΔS, were calculated from Eq. (2),

$$\Delta G = \Delta H - T\Delta S = -RT\ln K_A = -RT\ln(1/K_D) \quad (2)$$

where T is the absolute temperature and $R = 8.314$ J mol^{-1} K^{-1}.

Results and discussion

In silico mutagenesis and docking

The mutagenesis and docking studies were performed by TRITON using MODELLER 6v2 and AutoDock 3.1. Three times nineteen mutant structures were constructed, one set of possible nineteen single-point mutants for each of the three positions of amino acid 22, 23 and 24. The resulting mutant files were then subjected to the preparative procedure as described in "Computational methods".

For the docking procedure itself, two ligands were chosen to serve as benchmarks—methyl-α-L-fucopyranoside and methyl-α-L-mannopyranoside. The rationale for this choice came from the fact that the fucose-mannose preference was confirmed to be very sensitive towards the composition of the specificity-binding loop in previous studies, standing at the opposite poles of preference range. Evaluating the shift towards one of the border cases can therefore efficiently serve as the indicator of the impact of the mutation.

The docked structures that came up as the best from the performed docking runs (see later) were analyzed qualitatively (correct/incorrect binding mode, with "correct" representing the mode described in [15]) using the testing set of ligands, and also quantitatively by the energy score values provided by AutoDock. The results are summarized in Table 1. Analysis of the results revealed some important observations. Almost any mutation at the position 22, where the serine side chain is replaced by a longer or a bulkier side chain, results in destruction of the binding ability due to sterical hindrances coming from the side chain, preventing the saccharide from completing the coordination spheres by its hydroxyl oxygen atoms. Exceptions were recorded also for two longer chains, lysine and glutamine, where the main part of the side chain was turned outside and did not hinder the binding, resulting in tilted, but correct position of the ligand.

The mutations in the position 23 generally turned out to have non-detrimental effect to the binding. In this case, only bulky aromatic ligands were outright detrimental to the binding, the others allowed the saccharide to adopt at least partially correct binding mode. In this case, the side chain is almost exclusively directed out of the saccharide-binding site, and no direct interactions of the side chain and the ligand were observed. The more favorable energies for some complexes with methylmannoside can be explained by the type of scoring the AutoDock v. 3.1 uses, where the presence of hydroxymethyl group of C6 together with basic residues like arginine results in better score than the absence of said group in complexes with methylfucoside.

The most interesting results were expected from the mutations at the position 24. First and foremost, the side chains in this position have lesser potential to sterically hinder the binding and greater potential to actually create new interactions with the atoms of the ligand, basically being positioned not directly into the binding site. Our previous mutagenesis studies [18, 20] have shown that substitution of small glycine by asparagine has no impact on the binding site architecture.

It appears that any change on the position 24 seemingly increased the affinity of the mutant to methyl-α-L-mannoside. We have chosen two mutants for the experimental and molecular dynamics study. Based on the overall score, the the G24R mutant with arginine was chosen for closer investigation, in order to reveal possible structural reasoning for the increased affinity. Another mutation chosen for further investigation was the G24W mutation, where the glycine 24 is replaced by bulky aromatic side chain of tryptophan. Displaying also a good energy score, the mutant was investigated for potential occurrence of hydrophobic or stacking interactions. Last but not least, we took into account the fact that that glycine 24 in PA-IIL is replaced by asparagine in its mannose-preferring homologue lectin RS-IIL from *R. solanacearum*. We made a decision to analyze in detail the specificity-modifying potential of dicarboxylic amino acid derivatives—i.e., asparagine and aspartate—thanks to the fact that they have rather long and flexible side chain bearing a charged group that can serve as a potential hydrogen bond acceptor.

Except for significantly worse energies in cases of the outright detrimental mutations, no mutation was found that could be unambiguously identified as mediating more favourable binding than the native structure This is caused mainly by the fact that the side chain of the amino acid 24 usually sticks out of the binding site after MODELLER performs the *in silico* mutagenesis. In order to explore the possibilities of the binding site actually accommodating for

Table 1 Results of *in silico* mutagenesis and docking of methylfucoside and methylmannoside into single-point specificity binding loop (22–23–24) mutants of PA-IIL

	Mode	S22 mutation		Mode	S23 mutation		Mode	G24 mutation-	
		Me-α-L-Fuc	Me-α-D-Man		Me-α-L-Fuc	Me-α-D-Man		Me-α-L-Fuc	Me-α-D-Man
Gly	✔	−9.08	−9.46	✔	−9.25	−9.47	✔	**−10.70**	**−9.23**
Arg	×	−5.11	−5.24	✔	−9.41	−10.04	✔	−9.78	−9.84
Trp	×	−5.09	−5.85	×	−7.22	−7.58	✔	−9.76	−9.77
Phe	×	−4.59	−5.30	✔	−9.51	−9.83	✔	−9.71	−9.61
Pro	✔	−9.44	−9.80	✔	−7.67	−8.09	✔	−9.69	−9.45
Asn	×	−7.44	−7.88	✔	−9.46	−9.62	✔	***−9.67***	***−9.15***
Ile	×	−5.37	−5.36	✔	−9.63	−9.50	✔	−9.62	−9.50
Val	×	−7.30	−7.66	✔	−9.52	−9.59	✔	−9.59	−9.49
Leu	×	−5.52	−5.39	✔	−9.47	−10.20	✔	−9.58	−9.45
Cys	×	−7.01	−7.22	✔	−7.26	−8.04	✔	−9.57	−9.40
Asp	×	−7.91	−7.92	✔	−9.49	−9.77	✔	−9.57	−9.40
Ser	✔	**−10.70**	**−9.23**	✔	**−10.70**	**−9.23**	✔	−9.56	−9.40
Thr	✔	−9.15	−9.43	✔	−9.51	−9.92	✔	−9.56	−9.41
Met	×	−5.18	−5.59	✔	−9.61	−9.91	✔	−9.55	−9.45
Gln	✔[a]	−8.83	−9.36	✔	−9.73	−8.70	✔	−9.55	−9.47
Lys	✔[a]	−9.48	−10.06	✔	−8.69	−10.05	✔	−9.54	−9.21
His	×	−7.61	−8.04	✔	−9.39	−9.98	✔	−9.53	−9.37
Ala	✔	***−9.26***	***−10.47***	✔	***−10.49***	***−9.15***	✔	−9.53	−9.39
Tyr	×	−4.15	−5.04	×	−8.01	−8.06	✔	−9.52	−9.41
Glu	×	−5.42	−6.01	✔	−9.47	−9.75	✔	−9.52	−9.45

Ticks and crosses signs mark correct or incorrect binding mode, respectively. The PA-IIL wild type is highlighted by bold font, the mutants analyzed in the previous study are highlighted by bold italics. The mutation of serine 22 results in majority of cases in the inability of bind the saccharide properly, whereas the mutations at 23 and 24 show more promise. E is the energy score (k cal mol^{-1}) provided by AutoDock

[a] Very tilted position of saccharide, but remaining contact of the hydroxyl groups of the saccharide with the coordination spheres of calcium atoms

the sugar binding and forming new interactions, the molecular dynamics simulations in explicit solvent were run starting from the results provided by the docking programs.

The best docked structures of methylfucoside and methylmannoside were taken as the starting structures for molecular dynamics simulation, following the above described protocol, both in equilibration and simulation. The trajectories were analyzed both from energetic and geometric point of view, searching for correspondence between changes of energy and distances of chosen atoms. The atoms were chosen on the basis of geometrically most favourable position to form additional connections with the saccharide.

Molecular dynamics

G24N mutant

The G24N mutant models one of the mutations of the specificity-binding loop that has a real natural counterpart

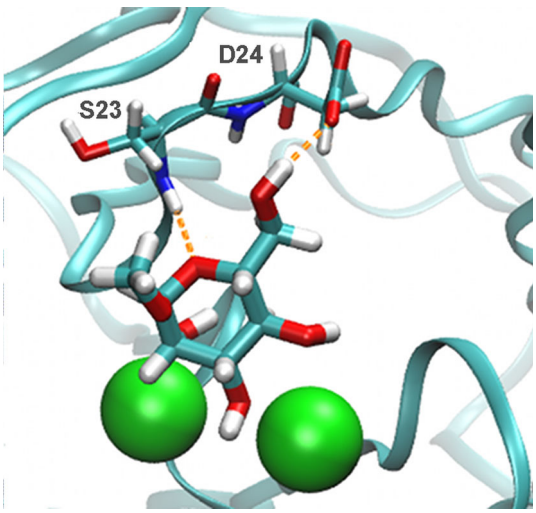

Fig. 2 The structure of G24D mutant with methyl-α-D-mannoside. Two interactions are highlighted by *orange dotted lines*—the amino group of the serine 23 with the ring oxygen of the ligand, and the interaction between hydroxyl group of C6 with the carboxyl group of aspartate

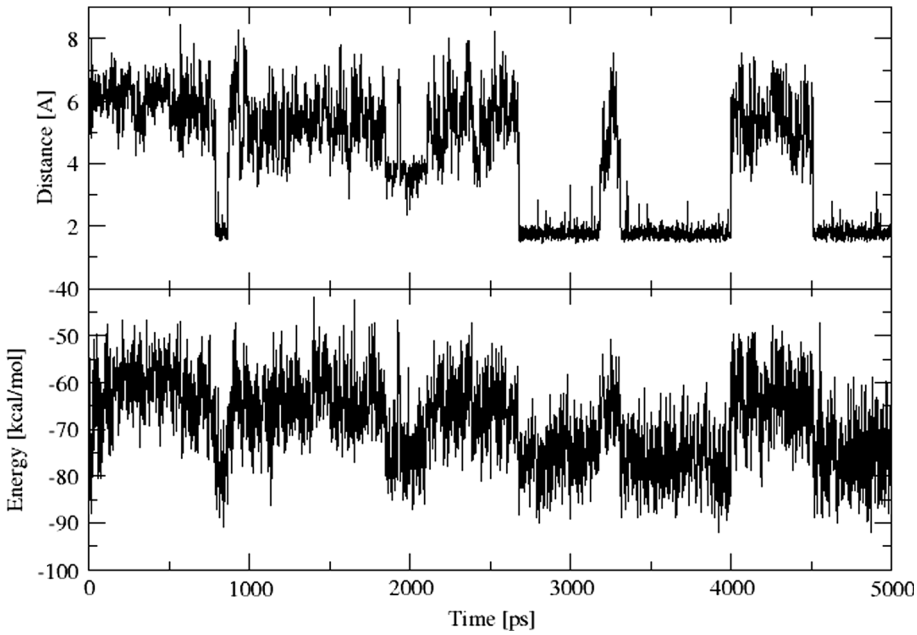

Fig. 3 Plot of interaction energies calculated in 1 ps snapshots within a 5 *ns* window from molecular dynamics simulation trajectory on G24D/Me-α-D-Man complex. Relationship between total interaction energy (above, kcal mol^{-1}) and the distance between C6 (methoxy-carbon) of the ligand and O1 of the aspartate side chain (*below* Å). Upon the proper side chain's change of position, the energy is decreased, and a positive effect at the binding is observed

(asparagine replaces glycine at this position in the RS-IIL lectin). This mutant has shown increased affinity to methyl-fucoside in vitro. The analysis of the trajectory has shown a hydrogen bond between serine 23 and the acetal oxygen of the saccharide. This observation is in accordance with the crystal structures, and it has been seen also in the other investigated mutants.

G24D mutant

The G24D mutant has the glycine in position 24 replaced by aspartate. The side chain was initially placed in direction out of the binding site by the MODELLER software. During the molecular dynamics run with methyl-α-D-mannoside, a hydrogen bond between the backbone amino group of the serine 23 and the acetal oxygen of the saccharide was formed (see Fig. 2). Even more interesting possible interaction was found in the case of carboxyl group of the aspartate and the hydroxymethyl C6 group of the saccharide. By comparing the distance of the carboxyl oxygen and the C6 carbon, a synergy has been found between the distance and an improvement in binding energy during the whole trajectory. This is illustrated in the 5 ns part of the trajectory by Fig. 3, in regions around 850–900, 2,650–3,200 and 3,300–4,000 ps. However, the interaction is unstable in the time.

G24W mutant

In the case of the tryptophan mutant, a possible CH-π interaction of the benzene ring of the indole side chain was expected. The tryptophan mutant indeed displays a possibility of hydrophobic pocket being created in part of the binding site. The presence of such pocket would favor the interaction of alkyl-derivatized saccharide molecules or stacking interactions of the indole benzene ring and other saccharide units in case of oligosaccharides. The closest, as well as most frequently occupied position of the tryptophan side chain during the whole run is depicted on Fig. 4. No particular direct agreement between the position of the tryptophan side chain in terms of improvement of the binding energy was observed. No clear stacking interaction was observed, either.

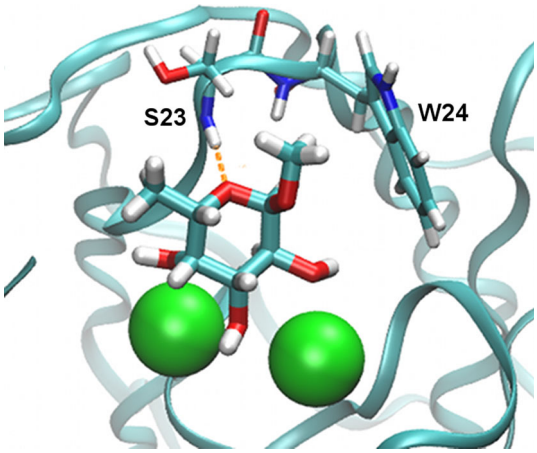

Fig. 4 The most frequently observed structure of the G24W/Me-α-L-Fuc complex during the molecular dynamics run, illustrating a possible formation of a hydrophobic pocket, hinting at the CH-π interaction between the methyl group of the sugar and the aromatic ring of the tryptophan indole side chain. Like in the previous case, there is an optional hydrogen bond formed between the oxygen of the saccharide ring and the amino group of serine 23

Table 2 Binding free energies (kcal mol^{-1}) calculated from MD trajectories using MM/GBSA approach

	Me-α-L-Fuc		Me-α-D-Man	
	$\Delta G_{binding}$	SD	$\Delta G_{binding}$	SD
PA-IILwt	−11.74	2.84	−11.32	5.88
G24D	−5.89	3.61	−9.11	5.34
G24N	−10.71	3.34	−9.58	2.98
G24W	−19.34	3.63	−17.08	3.85
G24R	−18.48	3.28	−13.46	4.57

For more information see "Experimental methods" section

G24R mutant

In the case of the arginine mutant, a search for possible interaction with the large basic side chain of the amino acid was performed. However, no particular relation between the position of the side chain and the ligand with the binding energy was observed in this case.

Binding free energy calculations and solvent behaviour

Based on the above trajectories, binding free energies were calculated using MM/GBSA method. The results are summarized in Table 2.

The G24W and G24R clearly show increasing binding free energy while the binding free energy of G24N remains about the same. The binding free energy is calculated significantly lower for the G24D mutant.

Important role in lectin/carbohydrate binding may be played by the solvent (see, for example, [28, 40–44]). Molecular dynamics provide data that allow for analysis of solvent behaviour. We have performed such analysis on all trajectories and found out that there is only one tightly bound water position bridging the ligand and threonine 98. This water molecule is also present in all crystal structures of PA-IIL with sugar ligands solved until now (for sample structures, see Fig. 1c, d).

In vitro studies

Cloning and production of G24W, G24R and G24D mutants in E. coli

All three mutants were prepared in *E. coli* host cells. The resulting proteins were purified by affinity chromatography on the mannose-agarose column and run as sharp bands of about 11 kDa on SDS-PAGE. A typical yield was about 10 mg of the purified protein per litre of culture.

Thermodynamical characterization of protein/saccharide interactions

Microcalorimetry data were recorded for the interaction of the PA-IIL mutants with methyl-α-L-fucoside and methyl-α-D-mannoside, respectively, and results were compared with newly determined data with previously obtained wild-type PA-IIL and another mutant in the same position, G24N [18]. Three mutants (G24N, G24W and G24R) show higher affinity towards both sugars than the wild type PA-IIL, while the G24D mutation led to decreased affinity. The highest increase in affinity was recorded for the tryptophan mutant. The thermodynamic data of the interaction of PA-IIL and its mutants with both sugars is summarized in Table 3. The comparison of data obtained indicates that all mutant lectins bind preferentially to the methyl derivative of fucose, with about two orders of magnitude higher

Table 3 Thermodynamics of binding for Me-α-L-Fuc and Me-α-D-Man with PA-IIL and its mutants at position G24 as determined by ITC at 25 °C

	K_A ($\times 10^4$ M^{-1})	K_D (μM)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	−TΔS (kcal mol^{-1})
Me-α-L-Fuc					
G24D	87 (±3.7)	1.15 (±0.05)	−8.1 (±0.03)	−8.46 (±0.01)	1.51 (±0.10)
G24R	298 (±7.1)	0.36 (±0.04)	−8.85 (±0.02)	−9.43 (±0.07)	0.57 (±0.10)
PA-IILwt	303 (±21.5)	0.33 (±0.02)	−8.84 (±0.04)	−9.44 (±0.03)	0.60 (±0.06)
G24N	408 (±4.1)	0.25 (±0.02)	−9.01 (±0.01)	−9.54 (±0.05)	0.52 (±0.04)
G24W	771 (±23.3)	0.13 (±0.04)	−9.39 (±0.02)	−9.44 (±0.07)	0.05 (±0.09)
Me-α-D-Man					
G24D	1.18 (±0.03)	89.6 (±2.2)	−3.61 (±0.01)	−5.52 (±0.01)	1.91 (±0.02)
PA-IILwt	1.55 (±0.03)	64.7 (±1.20)	−5.71 (±0.01)	−4.87 (±0.20)	−0.84 (±0.03)
G24N	2.47 (±0.07)	40.5 (±1.08)	−6.00 (±0.02)	−6.81 (±0.02)	0.81 (±0.03)
G24R	2.52 (±0.13)	39.8 (±1.94)	−6.00 (±0.02)	−6.92 (±0.24)	0.92(±0.02)
G24W	3.27 (±0.21)	30.6 (±1.91)	−6.17 (±0.05)	−5.50 (±0.1)	−0.65 (±0.05)

Standard deviations calculated from three independent measurements

affinity compared with methylmannoside similarly to the wild type protein.

The interested and may be not expected feature of PA-IIL/carbohydrate binding is low entropy. The reason is that there are three water molecules tightly bound to the calcium ions. These water molecules are released when the complex is created, which decreases entropy penalization that would normally occur when the complex is created.

Conclusions

It was demonstrated recently that the combination of *in silico* mutagenesis and docking is able to provide results in accordance to experimentally observed data and therefore has a potential for serving as a predictive tool for protein engineering. Also this mutagenesis study performed at the specificity-binding loop has shown that there is a possibility of creating mutants with additional contributions to the interaction with the ligand. However, it is important to note that, thanks to the fact that the active site is localized at the surface of the protein, the structural changes that were intended to observe are very small, and an appropriate level of precision is demanded for increasing the probability of modelling them correctly. Our results prove that docking combined with binding free energy calculation is capable of providing correct insight into the nature of lectin–saccharide interactions. We have demonstrated that even in systems such close to each other in terms of structure, appropriate combination of molecular modelling methods is able to help to identify the prospective protein mutations for tuning interactions in lectin–saccharide systems.

Acknowledgments This work was supported by the Czech Science Foundation (Contract No. 13-25401S), the Ministry of Education, Youth and Sports of the Czech Republic (Contract No. LH13055), The European Community's Seventh Framework Programme under European Regional Development Fund (Contract No. CZ.1.05/1.1.00/02.0068) and under the "Capacities" specific programme (Contract No. 286154—SYLICA). The access to MetaCentrum supercomputing facilities provided under the research intent LM2010005 is highly appreciated.

References

1. Lis H, Sharon N (1998) Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem Rev 98(2):637–674
2. Gabius HJ, Andre S, Kaltner H, Siebert HC (2002) The sugar code: functional lectinomics. Biochim Biophys Acta Gen Subj 1572(2–3):165–177
3. Rini JM (1995) Lectin structure. Ann Rev Biophys Biomol Struct 24:551–577
4. Cambi A, Koopman M, Figdor CG (2005) How C-type lectins detect pathogens. Cell Microbiol 7(4):481–488
5. Drickamer K, Fadden AJ (2002) Genomic analysis of C-type lectins. Glycogenomics Impact Genomics Inform Glycobiol 69:59–72
6. Drickamer K, Taylor ME (1993) Biology of animal lectins. Ann Rev Cell Biol 9:237–264
7. Drickamer K (1996) Ca2+-dependent sugar recognition by animal lectins. Biochem Soc Trans 24(1):146–150
8. Gilboa-Garber N (1982) *Pseudomonas aeruginosa* lectins. Methods Enzymol 83:378–385
9. Rhim AD, Stoykova LI, Trindade AJ, Glick MC, Scanlin TF (2004) Altered terminal glycosylation and the pathophysiology of CF lung disease. J Cyst Fibros 3:95–96
10. Roussel P, Lamblin G (2003) The glycosylation of airway mucins in cystic fibrosis and its relationship with lung infection by *Pseudomonas aeruginosa*. Adv Exp Med Biol 535:17–32
11. Tielker D, Hacker S, Loris R, Strathmann M, Wingender J, Wilhelm S, Rosenau F, Jaeger K-E (2005) *Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation. Microbiology 151:1313–1323
12. Sonawane A, Jyot J, Ramphal R (2006) *Pseudomonas aeruginosa* LecB is involved in pilus biogenesis and protease IV activity but not in adhesion to respiratory mucins. Infect Immun 74(12):7035–7039
13. Cioci G, Mitchell EP, Gautier C, Wimmerova M, Sudakevitz D, Perez S, Gilboa-Garber N, Imberty A (2003) Structural basis of calcium and galactose recognition by the lectin PA-IL of *Pseudomonas aeruginosa*. FEBS Lett 555(2):297–301
14. Mitchell E, Houles C, Sudakevitz D, Wimmerova M, Gautier C, Perez S, Wu AM, Gilboa-Garber N, Imberty A (2002) Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients. Nat Struct Biol 9(12):918–921
15. Mitchell EP, Sabin C, Snajdrova L, Pokorna M, Perret S, Gautier C, Hofr C, Gilboa-Garber N, Koca J, Wimmerova M, Imberty A (2005) High affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 angstrom resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins 58(3):735–746
16. Pokorna M, Cioci G, Perret S, Rebuffet E, Kostlanova N, Adam J, Gilboa-Garber N, Mitchell EP, Imberty A, Wimmerova M (2006) Unusual entropy-driven affinity of *Chromobacterium violaceum* lectin CV-IIL toward fucose and mannose. Biochemistry 45(24):7501–7510
17. Sudakevitz D, Kostlanova N, Blatman-Jan G, Mitchell EP, Lerrer B, Wimmerova M, Katcoff DJ, Imberty A, Gilboa-Garber N (2004) A new *Ralstonia solanacearum* high affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL. Mol Microbiol 52:691–700
18. Adam J, Pokorna M, Sabin C, Mitchell EP, Imberty A, Wimmerova M (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa*—unravelling the role of the specificity loop for sugar preference. BMC Struct Biol 7:36
19. Sabin C, Mitchell EP, Pokorna M, Gautier C, Utille JP, Wimmerova M, Imberty A (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: thermodynamics data correlated with X-ray structures. FEBS Lett 580(3):982–987
20. Adam J, Kriz Z, Prokop M, Wimmerova M, Koca J (2008) *In silico* mutagenesis and docking studies of *Pseudomonas aeruginosa* PA-IIL lectin—predicting binding modes and energies. J Chem Inf Model 48(11):2234–2242
21. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234(3):779–815
22. Nurisso A, Kozmon S, Imberty A (2008) Comparison of docking methods for carbohydrate binding in calcium-dependent lectins and prediction of the carbohydrate binding mode to sea cucumber lectin CEL-III. Mol Simul 34(4):469–479

23. Moura-Tamames S, Ramos MJ, Fernades PA (2009) Modelling beta-1,3-exoglucanase–saccharide interactions: structure of the enzyme-substrate complex and enzyme binding to the cell wall. FEBS J 276:156–157
24. Agostino M, Jene C, Boyle T, Ramsland PA, Yuriev E (2009) Molecular docking of carbohydrate ligands to antibodies: structural validation against crystal structures. J Chem Inf Model 49(12):2749–2760
25. Agostino M, Sandrin MS, Thompson PE, Yuriev E, Ramsland PA (2009) *In silico* analysis of antibody-carbohydrate interactions and its application to xenoreactive antibodies. Mol Immunol 47(2–3):233–246
26. Mishra SK, Adam J, Wimmerova M, Koca J (2012) *In silico* mutagenesis and docking study of *Ralstonia solanacearum* RSL lectin: performance of docking software to predict saccharide binding. J Chem Inf Model 52(5):1250–1261
27. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. J Mol Recognit 24(2):149–164
28. Gauto DF, Petruk AA, Modenutti CP, Blanco JI, Di Lella S, Marti MA (2013) Solvent structure improves docking prediction in lectin–carbohydrate complexes. Glycobiology 23(2):241–258. doi:10.1093/glycob/cws147
29. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comp Chem 19(14):1639–1662
30. Woods RJ, Woods Group (2005–2014) Biomolecule builder, GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia. http://www.glycam.com
31. Vriend G (1990) What if—a molecular modeling and drug design program. J Mol Graph 8(1):52–56
32. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GER, M. A.; Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision C.02. Gaussian, Inc. Wallingford
33. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comp Chem 26(16):1668–1688
34. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P (1995) Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. Comp Phys Commun 91(1–3):1–41
35. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984) A new force-field for molecular mechanical simulation of nucleic-acids and proteins. J Am Chem Soc 106(3):765–784
36. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012
37. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with constraints—molecular-dynamics of N-alkanes. J Comput Phys 23(3):327–341
38. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690
39. Onufriev A, Case DA, Bashford D (2002) Effective born radii in the generalized born approximation: the importance of being perfect. J Comp Chem 23(14):1297–1304
40. Lazaridis T (2000) Solvent reorganization energy and entropy in hydrophobic hydration. J Phys Chem B 104(20):4964–4979
41. Di Lella S, Ma L, Diaz Ricci JC, Rabinovich GA, Asher SA, Alvarez RMS (2009) Critical Role of the solvent environment in galectin-1 binding to the disaccharide lactose. Biochemistry 48(4):786–791
42. Di Lella S, Petruk AA, de Armino DJA, Alvarez RMS (2010) Specific intermolecular interactions of conserved water molecules with amino acids in the galectin-1 carbohydrate recognition domain. J Mol Struct 978(1–3):220–228
43. Kadirvelraj R, Foley BL, Dyekjaer JD, Woods RJ (2008) Involvement of water in carbohydrate-protein binding: concanavalin a revisited. J Am Chem Soc 130(50):16933–16942
44. Pearlstein RA, Sherman W, Abel R (2013) Contributions of water transfer energy to protein-ligand association and dissociation barriers: watermap analysis of a series of p38 MAP kinase inhibitors. Proteins 81(9):1509–1526