



**HAL**  
open science

## Virtual screening against carbohydrate-binding proteins: evaluation and application to bacterial *Burkholderia ambifaria* lectin

Tamir Dingjan, Emilie Gillon, Anne Imberty, Serge Perez, Elizabeth Yuriev,  
Alexander Titz, Paul A Ramsland

### ► To cite this version:

Tamir Dingjan, Emilie Gillon, Anne Imberty, Serge Perez, Elizabeth Yuriev, et al.. Virtual screening against carbohydrate-binding proteins: evaluation and application to bacterial *Burkholderia ambifaria* lectin. *Journal of Chemical Information and Modeling*, 2018, 58 (9), pp.1976-1989. 10.1021/acs.jcim.8b00185 . hal-02106059

**HAL Id: hal-02106059**

**<https://hal.science/hal-02106059>**

Submitted on 22 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Virtual screening against carbohydrate-binding proteins: evaluation and application to bacterial *Burkholderia ambifaria* lectin**

Tamir Dingjan<sup>1</sup>, Émilie Gillon<sup>2</sup>, Anne Imberty<sup>2</sup>, Serge Pérez<sup>3</sup>, Alexander Titz<sup>4,5,6</sup>, Paul A. Ramsland<sup>7,8,9,10\*</sup>, Elizabeth Yuriev<sup>1\*</sup>

<sup>1</sup>Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

<sup>2</sup>University Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France

<sup>3</sup>University Grenoble Alpes, CNRS, DPM, 38000 Grenoble, France

<sup>4</sup>Chemical Biology of Carbohydrates, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, D-66123 Saarbrücken, Germany

<sup>5</sup>Deutsches Zentrum für Infektionsforschung (DZIF), Standort Hannover-Braunschweig, Germany

<sup>6</sup>Department of Pharmacy, Saarland University, Saarbrücken, Germany

<sup>7</sup>School of Science, RMIT University, Bundoora, Victoria, 3083, Australia

<sup>8</sup>Department of Surgery Austin Health, University of Melbourne, Heidelberg, Victoria, 3084, Australia

<sup>9</sup>Department of Immunology, Monash University, Alfred Medical Research and Education Precinct, Melbourne, Victoria, 3004, Australia

<sup>10</sup>Burnet Institute, Melbourne, Victoria 3004, Australia

\*Correspondence to: Dr Elizabeth Yuriev, Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia. Tel: +61 9903 9611; E-mail: [elizabeth.yuriev@monash.edu](mailto:elizabeth.yuriev@monash.edu), Dr Paul A Ramsland, School of Science, RMIT University, Plenty Road, Bundoora, VIC 3083, Australia. Tel: +61 9225 7024; E-mail: [paul.ramsland@rmit.edu.au](mailto:paul.ramsland@rmit.edu.au)

## **Abstract**

Bacterial adhesion to human epithelia via lectins constitutes a therapeutic opportunity to prevent infection. To enable discovery of compounds able to inhibit bacterial lectins, a virtual screening protocol has been iteratively developed via 194 retrospective screening protocols against four bacterial lectins (BambL, BC2L-A, FimH and LecA) with known ligands. Specific attention was given to the rigorous evaluation of retrospective screening, including calculation of analytical errors for enrichment metrics. The developed virtual screening workflow was applied to BambL from the lung pathogen *Burkholderia ambifaria*, predicting 15 active compounds from virtual libraries of approximately 7 million compounds. Experimental validation using fluorescence polarization confirmed the inhibitory activity for two compounds, which were further characterized by titration calorimetry and surface plasmon resonance. This report demonstrates the utility of virtual screening protocols, integrating ligand-based pharmacophore filtering and structure-based constraints, in the search for bacterial lectin inhibitors.

## **Keywords**

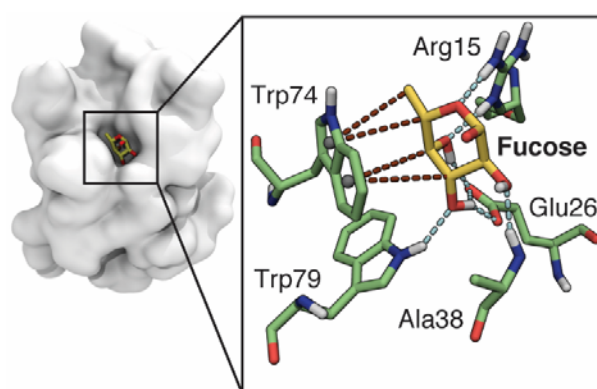
*Burkholderia ambifaria*, carbohydrate, fluorescence polarization, isothermal titration calorimetry, lectin, structure-based virtual screening, surface plasmon resonance, validation

## Introduction

Target dependence of docking protocols and scoring functions is a known obstacle to out-of-the-box application of structure-based virtual screening.<sup>1</sup> Carbohydrate-binding proteins present a particularly challenging target type as they exhibit a range of unique ligand-binding features, such as formation of hydrogen-bonding networks.<sup>2</sup> Therefore, appropriate methodologies need to be tailored for these targets, which are involved in many disease processes. For example, in epithelial infection, adhesion to human tissues is commonly achieved via binding of cell surface carbohydrates by bacterial carbohydrate-binding proteins, termed lectins.<sup>3-4</sup> Interfering with this binding has been pursued as a new mode of antibacterial therapy, mainly due to the slow selective pressure exerted by anti-adhesive agents.<sup>5-6</sup>

Inhibition of lectin-carbohydrate binding by glyco-derived compounds has met with promising clinical success in human urinary<sup>7-8</sup> and respiratory<sup>9-10</sup> infections, albeit hampered by the weak affinity of lectins for carbohydrates (typically in the milli- to micromolar range). Approaches to design efficient lectin inhibitors have explored two main directions: presentation of carbohydrate epitopes on multivalent scaffolds (glycodendrimers,<sup>11</sup> glycopeptides,<sup>12-13</sup> glycoclusters,<sup>14-15</sup> and glycofullerenes<sup>16</sup>) and design of glycomimetics, compounds combining carbohydrate character with functional groups able to form specific binding interactions or provide avenues for further optimization.<sup>17</sup> Indeed, small glycomimetics with high affinity are more likely to lead to drug development, owing to their more favorable pharmacokinetic properties.<sup>18</sup> Such high affinity glycomimetic inhibitors have been reported for multiple lectins, including FimH,<sup>19-20</sup> LecA,<sup>21-23</sup> LecB,<sup>24-26</sup> and BambL.<sup>27</sup>

BambL (the *Burkholderia ambifaria* lectin) displays affinity for fucosylated human blood group determinants.<sup>28</sup> *B. ambifaria*, a bacterial member of the *Burkholderia cepacia* complex (BCC), was identified in clinical isolates from cystic fibrosis patients,<sup>29</sup> a disease in which lectin-mediated bacterial adhesion to fucosylated lung epithelia is suspected to play an important role.<sup>30-31</sup> Compounds able to inhibit the BambL-saccharide interaction may have therapeutic potential as anti-adhesives. The present study applies our structural knowledge of the BambL binding site based on the protein's crystal structure<sup>28</sup> and on our previous simulations of the protein interacting with fucosylated oligosaccharides.<sup>32</sup> In the latter study, we have determined the atomic interactions made by fucosylated carbohydrates recognized by BambL. Each monomer of BambL has two independent but very similar binding sites that feature a network of hydrogen bonds within the fucose-binding pocket, complemented by interactions with solvent-exposed residues outside the binding cleft. Hydrogen-bonding interactions formed by the intramonomeric binding site Arg15, Glu26, and Trp79 were most significant for carbohydrate recognition, accompanied by hydrophobic stacking against Trp74 and hydrogen-bonding to Ala38 (Figure 1).



**Figure 1.** BambL-fucose crystallographic complex showing binding interactions. A single monomer of the protein is shown in white, carbohydrate shown in yellow with coloring by atom type. Hydrogen bonds shown as green dashes. Hydrophobic interactions shown as brown dashes. Non-polar hydrogen atoms omitted for clarity.

One widely used tool for discovery of novel inhibitors is structure-based virtual screening, an *in silico* approach that rapidly assesses the potential for molecules to interact with a specific binding site using molecular docking techniques.<sup>1, 33-34</sup> However, reports describing screening against carbohydrate-binding proteins often apply virtual screening methods as a “black box”, with neither investigations of how well the method would be expected to work (efforts termed retrospective screening or validation), nor attempts to improve the screening method. As Grant and Woods<sup>35</sup> pointed out, the docking methods used for carbohydrate virtual screening are typically evaluated without negative controls, precluding accurate measurement of the method’s ability to discriminate between binders and non-binders.

In the present work, we set out to iteratively develop and validate a workflow for virtual screening against bacterial lectins. Several questions were asked to guide the development of the virtual screening workflow: to what degree is carbohydrate screening enrichment artificially enhanced by poor camouflage (mismatch in molecular properties between actives and decoys)?; can combinations of ligand-based and structure-based screening improve enrichment?; and finally, can an enriching screening method unearth novel lectin ligand scaffolds in a prospective search? Through extensive method development and evaluation, 194 screening protocols combining ligand-based and structure-based screening approaches have been tested using a panel of four bacterial lectins with multiple known ligands (BambL, BC2L-A, FimH and LecA). We have paid specific attention to the validation aspect, using property-matched carbohydrate decoys and calculating analytical errors attending enrichment metrics. Experimental validation of the screening approach was carried out via a prospective search for novel inhibitors against BambL. Confirmatory bioassays using fluorescence polarization-based competitive binding, surface

plasmon resonance (SPR), and isothermal titration calorimetry (ITC) identified two hit compounds showing lectin inhibition.

## Methods

### Virtual screening method development

#### *Selection of actives and decoys*

Target bacterial lectins were selected based on the availability of crystallographic structures, glycan array data, and known binding activity for greater than 10 actives. Four lectins were chosen: BamBL,<sup>28</sup> BC2L-A from *Burkholderia cenocepacia*,<sup>36-37</sup> FimH from uropathogenic *Escherichia coli*,<sup>38</sup> and LecA from *Pseudomonas aeruginosa*.<sup>39-40</sup> Known actives for these lectins were sourced from published literature<sup>19-20, 22, 27, 41-53</sup> and from glycan array data deposited in the Consortium for Functional Glycomics Gateway database,<sup>54</sup> from which unique saccharide actives of tetrasaccharide size and smaller were selected. 3D structures of actives were manually constructed using Maestro.<sup>55</sup> Complete sets of actives are shown in Supplementary Table S1.

Small molecule decoys were selected from the ZINC database<sup>56-58</sup> using the Silico toolkit's molecular property-matching module.<sup>59</sup> Decoy selection was based on similarity to the actives in the following molecular properties: number of heavy atoms, number of hydrogen bond acceptors and donors, number of rotatable bonds, and LogD (predicted using QikProp<sup>60</sup>). A total of 20 decoys were selected per active. The resulting 2D structures were elaborated to corresponding 3D structures using LigPrep<sup>61</sup> without enumeration of protomers, tautomers, or conformers.

Carbohydrate decoys were generated separately for each target lectin by combinatorial assembly. For each target, component monosaccharides were chosen to avoid known binding epitopes, then assembled in a combinatorial fashion to form di-, tri- and tetrasaccharides using Instant JChem reactor modules.<sup>62</sup> For example, the target lectin LecA displays affinity for D-galactose epitopes; therefore, carbohydrate decoys were assembled from L-fucose, D-glucose, D-mannose, D-*N*-acetylgalactosamine, and D-*N*-acetylglucosamine monosaccharides. Assembly involved constructing all glycosidic linkages at all positions on all substrate saccharides, forming all possible oligosaccharides. The required number of decoys were randomly selected from each size category (i.e., di-, tri-, and tetrasaccharide sizes) to match the distribution of glycan actives from glycan arrays. A total of 40 decoys per active were chosen.

#### *Library preparation*

Three different library compositions were constructed: all actives with ZINC-derived decoys, all actives with carbohydrate decoys, and carbohydrate actives (i.e., those derived from glycan arrays) with carbohydrate decoys. In each case, Phase<sup>63</sup> was used to prepare the library. Compound states were generated at pH  $7.0 \pm 2.0$  using Epik.<sup>64</sup> Stereochemistry was determined from 3D geometry; unspecified stereocenters were sampled, retaining four stereoisomers. Rings were sampled with a single conformer retained. High-energy states were removed. For each compound, up to 100 conformers were kept, with a maximum of 10 conformers retained per rotatable bond.

#### *Receptor structure preparation*

Crystallographic structures used for structure-based screening are shown bolded in Supplementary Table S2. Each receptor structure was truncated to a single chain and prepared using the Protein Preparation Wizard implemented within Maestro.<sup>64</sup> Bond orders and hydrogen atoms were added,



metal coordination bonds and disulfide bonds created, and water molecules removed. Resultant models were refined by a restrained minimization of hydrogen atoms only. Receptor grids were generated by centering the cubic docking box on the crystallographic bound ligand, with a docking box side length of 20 Å and a diameter midpoint box side length of 15 Å.

#### *Binding pocket constraints*

A single constraint was defined in each lectin binding site as a 1 Å radius sphere centered on the most deeply buried ligand atom in the crystallographic complex. In all cases, this corresponded to a saccharide hydroxyl oxygen atom. Matching features were set to SMARTS patterns for a single carbon, nitrogen, oxygen, or sulfur atom.<sup>55</sup>

#### *Pharmacophore hypothesis*

A binding pharmacophore was assembled for each lectin by tallying hydrogen bonding and hydrophobic interactions occurring in crystallographic structures (Supplementary Table S3). Hydrogen bonds were defined using the Baker-Hubbard method implemented in MDTraj.<sup>65</sup> Hydrophobic interactions were defined as pairs of aliphatic carbon atoms within a 5 Å maximum distance, where each carbon atom was bonded only to carbon or hydrogen atoms. Frequently-interacting ligand features were incorporated into the pharmacophore hypothesis using the Develop Hypothesis Model tool in Phase.<sup>63</sup> The Advanced Pharmacophore Screening tool was used to find pharmacophore matches prior to structure-based screening: database keys were used for pre-screening, with tolerance of inter-site matching distances set to 2.0 Å.

#### *Interaction constraints*

Using the interaction tallies from crystallographic structures, the four most frequently hydrogen-bonding receptor atoms were used as interaction constraints (Supplementary Table S3). In addition,

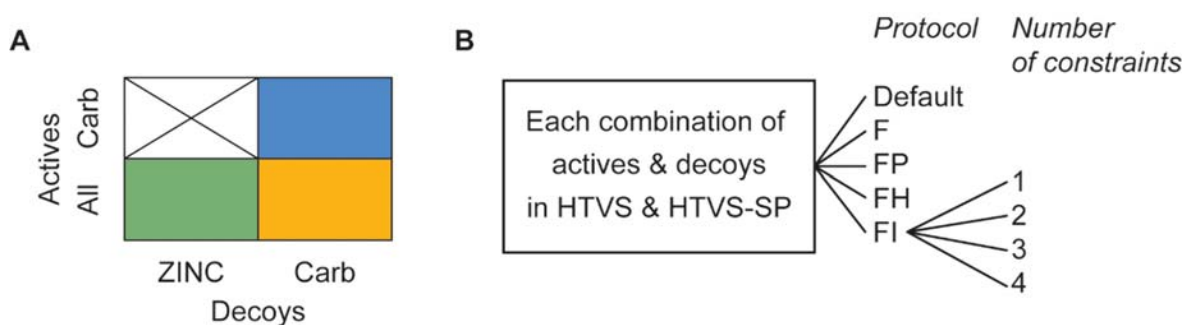
a single hydrophobic interaction constraint was defined as a 1 Å sphere centered on the crystallographic ligand atom most involved in hydrophobic interactions.

Interaction constraints were applied in multiple configurations: first, a hit was required to participate in simultaneous hydrogen bonding interaction with all four atoms, satisfying all four interaction constraints (FH method); second, replacement of the least dominant hydrogen bonding constraint with a hydrophobic constraint (FI method); third, using all hydrogen bonding and hydrophobic constraints while varying the number of required constraint matches from one to four (FI-1, FI-2, FI-3, and FI-4), such that hits would be accepted if able to simultaneously satisfy the required number of constraints. For BC2L-A, no hydrophobic interactions were detected in crystallographic complexes. For both BC2L-A and LecA, binding-site calcium ions are known to be significant for ligand binding; these were defined as an interaction constraint alongside hydrogen bonding interactions. FimH and LecA complexes feature a broad hydrophobic patch at the entrance to the mannose-binding/galactose-binding site, respectively; hydrophobic interaction constraints capturing these interactions took the form of a 3.6 Å spherical constraint positioned between multiple binding site residues (Supplementary Table S3). During screening, hydrophobic constraints were specified to require a bonded carbon and hydrogen atom within the constraint radius using the SMARTS pattern C[H].

### *Virtual screening*

Docking calculations were conducted via the Virtual Screening Workflow available within Maestro.<sup>55</sup> Receptor grids were generated within the screening workflow. To investigate the effect of different active-decoy set combinations on screening enrichment, three library compositions were investigated for each lectin: all actives – ZINC-derived decoys, all actives – carbohydrate decoys, and carbohydrate actives – carbohydrate decoys (Figure 2A). Screening was performed on

each input library using the HTVS algorithm. Additionally, each screen was also performed using the Glide HTVS algorithm followed by a re-screening step using the Glide SP algorithm;<sup>66-67</sup> in these, the top-ranked 10% of hits were re-docked using SP mode, with enhanced sampling enabled and retaining only the highest-scoring state of each compound. The following protocols, summarized in Figure 2B, were tested in this investigation. *Default*: structure-based screening conducted using entirely default settings. *F*: input library filtering using ligand pharmacophore matching as described above followed by structure-based screening using default settings. *FP*: input library filtering followed by structure-based screening using a binding pocket constraint. *FH*: input library filtering followed by structure-based screening using four hydrogen bonding and metal coordination interaction constraints. *FI*: input library filtering followed by structure-based screening with the least-dominant hydrogen bonding constraint replaced by a hydrophobic constraint, maintaining a total of four required constraints. Finally, this screen was repeated with all hydrogen bonding and hydrophobic constraints, varying the number of required constraints from one to four. In total, 194 screening protocols were tested in an iterative fashion (Supplementary Tables S4 and S5).



**Figure 2.** Schematic depicting (A) composition of active and decoy sets and (B) retrospective screening protocols evaluated.

Enrichment metrics were calculated using the Silico toolkit.<sup>59</sup> For both ROC AUC and ROC-based Enrichment Factor, analytical errors were calculated according to the method communicated by McGann et al.<sup>68</sup>

### Prospective virtual screening library preparation

Two large molecular databases were screened prospectively: the eMolecules Screening Library (approximately 7 million compounds) and the Otava Ltd. Catalogue and Glycomimetic Collection (containing 246 compounds). Web-based substructure search tools were used to extract compounds containing a SMILES string CXC(C)C(O)CX, where X = O for the eMolecules Screening Library and O/N for the Otava Ltd. catalogue. The total number of input compounds collected from eMolecules and Otava databases was 11,705. The Generate Phase Database module of Phase<sup>63</sup> was used to prepare the virtual library compounds by adding hydrogens and generating conformers (up to a maximum of 100 conformers per compound). Tautomers were enumerated at pH 7.0 using Epik.<sup>64</sup> Stereochemistry was determined from 2D parities where present; where absent, a maximum of four stereoisomers were retained per compound. Rings of five and six members were sampled, retaining one ring conformation per compound. Virtual screening was performed using the FI method as described above: after applying the ligand-based pharmacophore filter, resulting matches were screened using the HTVS algorithm. The top-ranked 10% of hits were then re-screened using the SP algorithm with enhanced sampling enabled, retaining only the highest-scoring state of each compound. Finally, the top 10% of hits were manually inspected, removing undesirable compounds (simple monosaccharides, compounds of molecular weight higher than 600 g/mol, compounds containing free thiols, and commercially unavailable compounds).

## Isothermal titration calorimetry

Recombinant BambL was produced and purified as previously described by Audfray et al.<sup>28</sup> The protein was dissolved in buffer (20 mM Tris/HCl pH 7.5, NaCl 100 mM with CaCl<sub>2</sub> 100 μM). The protein concentration was checked by measuring A<sub>280</sub> using a theoretical molar extinction coefficient of 40,450 M<sup>-1</sup> cm<sup>-1</sup>. Isothermal titration calorimetry (ITC) was performed with an ITC200 microcalorimeter (MicroCal, Inc). The BambL solution (50 μM) was placed in the 200 μL sample cell at 25 °C. Ligands were dissolved in the same buffer to a concentration of 1 mM, and loaded in the injection syringe. Titration was performed using 20 injections of 1 μL volume of ligand solution, which were spaced at 120 second intervals and injected for a duration of 2 seconds each. Where binding isotherms were observed, data were fitted with MicroCal Origin 7 software, according to standard procedures, using a 1:1 binding site model. Fitted data yielded the stoichiometry (n), the association constant (K<sub>a</sub>) and the enthalpy of binding (ΔH). Other thermodynamic parameters (i.e. changes in free energy ΔG, and entropy ΔS) were calculated from the equation  $\Delta G = \Delta H - T\Delta S = RT\ln K_a$  in which T is the absolute temperature and R = 8.314 J mol<sup>-1</sup> K<sup>-1</sup>. Three independent titrations were performed for each ligand tested.

## Surface plasmon resonance

Surface plasmon resonance (SPR) experiments were performed on a Biacore X100 instrument (GE Healthcare) at 25°C in HBS (10 mM HEPES/NaOH, pH 7.4, 150 mM NaCl, 0.05% Tween 20 and 3 mM EDTA) at a flow rate of 30 μL min<sup>-1</sup>. Streptavidin (100 μg mL<sup>-1</sup>) was immobilized on a research grade CM5 dextran chip using amine coupling (3029.6 resonance units of streptavidin were fixed in channel 1, 2939.5 resonance units in channel 2). Biotin capture was then conducted to affix biotinylated polyacrylamide (PAA) probes (Lectinity, 200 μg mL<sup>-1</sup>) bearing galactose or

fucose to channels 1 (PAA-galactose, 375 resonance units) and 2 (solution containing 90% PAA-galactose and 10% PAA-fucose, total of 447.5 resonance units). Inhibition experiments were performed with the fucosylated channel 2, and plots represent the subtracted data (channel 2 – channel 1). Inhibition studies consisted of the injection (association 180 s, dissociation 180 s) of incubated (>30 min at room temperature) mixtures of BambL (0.8  $\mu$ M) and various concentrations of inhibitor (2-fold cascade dilutions from 200 to 0.2  $\mu$ M). For each inhibition assay, BambL was injected to observe the full adhesion of the lectin onto the sugar-coated surface (0% inhibition). The chip was regenerated by injection of 1 M fucose in running buffer for 80 s. Binding was measured as resonance units over time after blank subtraction, and data were then evaluated by using the Biacore X100 evaluation software, version 2.0. For IC<sub>50</sub> evaluation, the response was taken as the amount of lectin bound to the sugar-coated surface at equilibrium in the presence of a defined concentration of inhibitor. Inhibition curves were obtained by plotting the percentage of inhibition against the inhibitor concentration.

### Competitive binding using fluorescence polarization

Assays were conducted as previously established for LecB<sup>25, 69</sup> and other bacterial lectins.<sup>41, 70-71</sup> A 60  $\mu$ L solution containing lectin (225 nM for LecB, 150 nM for BambL) and reporter ligand FITC-fucoside<sup>25</sup> (1.5 nM for LecB, 75 nM for BambL) in TBS buffer was added to 30  $\mu$ L serial dilutions of testing compounds in TBS (in duplicates for LecB, single serial dilution for BambL) in black 96-well Costar plates (Fisher Scientific, France, cat. no.: 06-443-2). Concentration ranges were selected for each compound to cover the range over which inhibitory activity was visible. For LecB,  $\alpha$ -methyl fucoside,  $\alpha$ -methyl mannoside and L-Fuc $\alpha$ 1-6[D-GlcNAc $\beta$ 1-4]D-GlcNAc $\beta$ 1-OMe were prepared in 5-fold dilutions from 33.3 mM to 400 nM, and 2-deoxy-2-fluoro-L-fucose was prepared using 2-fold dilutions from 25 mM to 0.195 mM. For BambL, concentration ranges

were selected for each compound using 5-fold dilutions for:  $\alpha$ -methyl fucoside (3.33 mM to 213 nM), L-Fuc $\alpha$ 1-6[D-GlcNAc $\beta$ 1-4]D-GlcNAc $\beta$ 1-OMe (3.33 mM to 42.6 nM), and 2-deoxy-2-fluoro-L-fucose (16.6 mM to 213 nM). After addition of the reagents, the plates were sealed with film and shaken for 4 hours in darkness at room temperature. Fluorescence polarization was measured on a BMG LABTECH CLARIOstar multi-mode microplate reader (BMG LABTECH GmbH) with excitation filters at 482-16 nm and emission filters at 530-40 nm. The data were analyzed with Graphpad Prism 6 and fitted according to the four-parameter variable slope model.

## Results

### Virtual screening optimization for bacterial lectins

#### *Selection of target lectins, actives, and decoys*

The BambL, BC2L-A, FimH, and LecA lectins were chosen for virtual screening optimization. For each lectin, decoy compounds were selected from the ZINC database to approximate the molecular properties of known actives as described in the Methods. The molecular properties of ZINC decoys proved a poor match for the molecular properties of active compounds (shown for BambL in Table 1, for all four lectins in Supplementary Figure S1). A second set of decoy compounds was therefore assembled, which exhibited closer property-matching, except for the number of rotatable bonds. Both decoy sets were used for retrospective screening to compare the effect of decoy property-matching on retrospective enrichment.

**Table 1.** Average molecular properties for BambL actives and decoys.

Compound set	Heavy		Rotatable bonds	H-bond acceptor	H-bond donor
	atom count	LogD			
Actives	31.7	-4.2	5.8	12.8	8.3
Actives (Carbohydrate only)	32.5	-4.6	6.1	13.4	8.8
ZINC	25.8	-1.2	5.6	6.9	3.8
Carbohydrate decoys	37.6	-6.2	8.2	16.0	11.0

#### *Virtual screening method development*

The docking software Glide predicts ligand placements within protein binding sites. When docking calculations are performed in a high-throughput manner, Glide may be used to predict binding activity for large collections of molecules. Glide has demonstrated appreciable docking accuracy for carbohydrate-binding proteins in general<sup>2, 72-73</sup> and for bacterial lectins specifically.<sup>74-75</sup>

A virtual screening workflow was developed using Glide<sup>66-67</sup> over a total of 194 retrospective virtual screens spanning all target and active/decoy sets. The multiple methods explored during the iterative workflow development are summarized in Figure 2. The left panel (Figure 2A) indicates the three compositions of active and decoy sets trialled for each lectin (note that carbohydrate actives were not combined with ZINC decoys). The various screening protocols performed are summarized in the right panel (Figure 2B).

Iterative development of the virtual screening method explored combinations of pharmacophore filters, binding site constraints and protein-ligand interaction constraints over a total of 194 screening experiments. Filters were applied in a ligand-based manner prior to structure-based

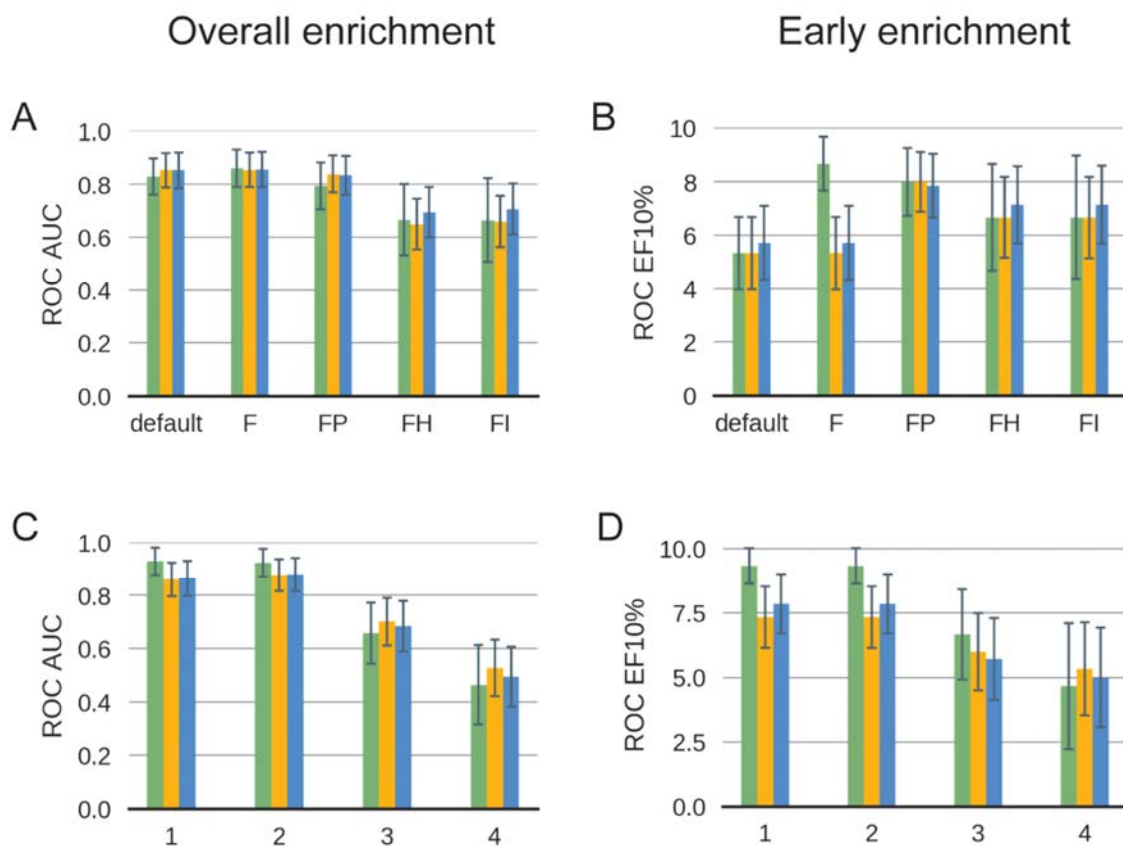


screening, while constraints were applied during structure-based screening to enforce the placement of ligand atoms either within the general binding site, or in positions able to form selected interactions with the lectin. Development of a screening method began with an entirely ‘default’ approach, to which pharmacophore filters were subsequently added (‘F’). Binding pocket constraints were applied (‘FP’), and then exchanged for hydrogen-bonding interaction constraints (‘FH’) and for general protein-ligand interaction constraints (‘FI’). The above screening methods were each conducted using two scoring protocols implemented within Glide: High-Throughput Virtual Screening (HTVS) alone, and in conjunction with Standard Precision scoring (HTVS-SP). Overall enrichment was measured using linear ROC AUC (Supplementary Table S4), which expresses the probability that an active compound will be ranked higher than a decoy (henceforth referred to as “AUC”). An AUC of 0.5 represents essentially random ranking of actives, while an AUC of 1.0 corresponds to perfect ranking of all actives above all decoys. AUC accords the same weight to rankings of actives throughout the hitlist, making the metric equally sensitive to actives present at the start and end of the hitlist. Since the early portion of the hitlist is of greater interest for prospective virtual screening, a metric more sensitive to early predictive ability was also employed. Specifically, to measure early enrichment, the ROC-based enrichment factor EF (Supplementary Table S5) was used, which describes the fraction of actives found once a fixed, chosen fraction of inactives have been found (here, set to 10%). By selecting inactive fractions in the early regions of the ROC curve, a method’s ability to rank actives above decoys in the early portions of the hitlist can be measured. Both scoring protocols HTVS and HTVS-SP achieved very similar enrichments in most experiments (Supplementary Tables S4 and S5). Below quoted results were determined using the HTVS-SP protocol. Analytical errors were calculated for both AUC

and EF, and were taken into account in the comparisons between screening protocols. Where necessary, error values are quoted.

### *BambL*

When property-matched ZINC-derived decoys were spiked with all actives (Figure 3, green), structure-based screening with default parameters achieved high overall enrichment (AUC: 0.83; Figure 3A). Addition of a ligand pharmacophore filter (F) and a binding pocket constraint (FP) enhanced early enrichment (EF: 5.33 → 8.67 [F], 8.00 [FP]; Figure 3B), but not overall enrichment (AUC: 0.86 [F], 0.80 [FP]). Enrichment was decreased by the use of four hydrogen-bonding (FH) and interaction (FI) constraints, both early (EF: 6.67 [FH, FI]; Figure 3B) and overall (AUC: 0.67 [FH, FI]; Figure 3A). Inspection of the ROC profiles for these experiments (data not shown) revealed truncated curves, indicating that interaction constraints were overly restrictive. Methods relaxing the number of required interaction constraints were therefore evaluated (FI-1, FI-2, FI-3, FI-4). By reducing the number of required constraint matches from the initial four to one or two, overall enrichment was substantially improved (AUC: 0.93 [FI-1], 0.92 [FI-2]; Figure 3C).



**Figure 3.** Overall and early enrichment represented by ROC curve metrics for Bambl HTVS-SP screening. (A) Overall enrichment for default, F, FP, FH, and FI protocols. (B) Early enrichment for default, F, FP, FH, and FI protocols. (C) Overall enrichment for FI-1, FI-2, FI-3, and FI-4 protocols. (D) Early enrichment for FI-1, FI-2, FI-3, and FI-4 protocols. Green: ZINC-derived decoys with all actives; yellow: carbohydrate decoys with all actives; blue: carbohydrate decoys with carbohydrate actives only. Error bars show the analytically calculated standard deviation.

Early enrichment was also significantly higher when using fewer required constraint matches (EF: 9.33 [FI-1, FI-2]; Figure 3D). Carbohydrate decoys, with either all or carbohydrate actives (yellow and blue bars in Figure 3, respectively) yielded similar enrichment results for most protocols. However, when using the optimal protocols FI-1 or FI-2, the early enrichment was lower with carbohydrate decoys, compared to ZINC-derived ones (Figure 3D). This finding demonstrated a

more discriminating character of carbohydrate decoys, which would result in a more robust method validation via retrospective screening.

### *BC2L-A*

Overall enrichment for BC2L-A was not significantly altered by the use of a pharmacophore filter (F) or binding pocket constraints (FP) (Supplementary Table S4). Early enrichment measures likewise remained within range of analytical error between experiments (Supplementary Table S5). In screening of all actives against ZINC decoys, the introduction of four simultaneously applied hydrogen-bonding constraints drastically reduced enrichment (AUC:  $0.58 \pm 0.07$  [default]  $\rightarrow$  0.12 [FH]). Relaxing the number of required constraints to one was able to restore the overall enrichment (AUC:  $0.65 \pm 0.07$  [FI-1]). Screening all actives against carbohydrate decoys performed similarly to screening against ZINC decoys. However, screening using only carbohydrate actives, with any number of interaction constraints (FI-1, FI-2, FI-3, and FI-4), returned no active compounds (Supplementary Tables S4 and S5). For this target lectin, enrichment relied strongly on the presence and ranking of non-carbohydrate actives.

### *FimH*

Both overall and early enrichment against ZINC decoys was significantly improved by the use of a pharmacophore filter (AUC: 0.72 [default]  $\rightarrow$  0.82 [F], EF: 4.41 [default]  $\rightarrow$  7.94 [F]; Supplementary Tables S4 and S5). Use of a binding pocket constraint (FP) and hydrogen-bonding constraints (FH) maintained the boost to early enrichment while decreasing overall enrichment. Introduction of a hydrophobic constraint severely reduced both overall and early enrichment (AUC: 0.18 [FI]; EF: 1.76 [FI]). Relaxing the number of required constraint matches rescued the enrichment. Specifically, a single required constraint matched the highest overall and early

enrichment to within the analytical error (AUC: 0.81 [FI-1]; EF: 7.65 [FI-1]). Carbohydrate decoys reduced overall enrichment using this protocol, more so for input containing only carbohydrate actives (AUC: 0.67). Early enrichment dropped from 7.65 to 5.00 and 2.22, for all and carbohydrate actives, respectively (Supplementary Table S5). Again, similar to BambL, these decreases demonstrated the greater discriminating power of carbohydrate decoys, compared to non-carbohydrate ones.

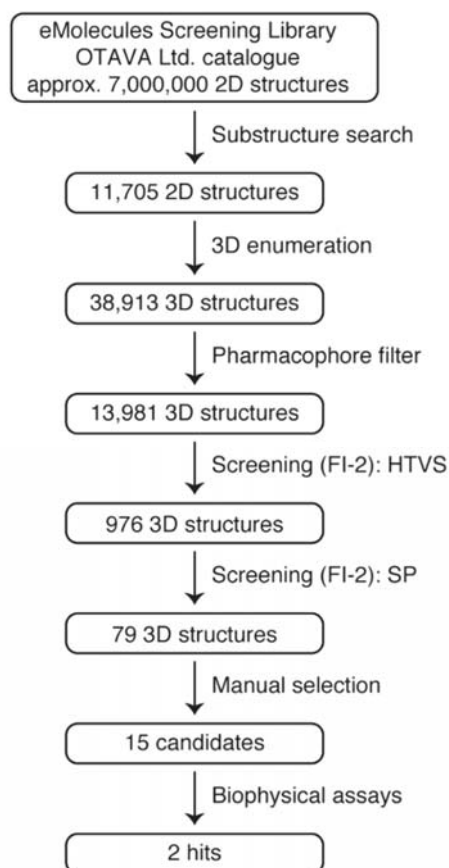
### *LecA*

In general, screening performance for LecA was poorer than for other target lectins (Supplementary Tables S4 and S5). Using ZINC decoys, default screening settings discriminated against actives, performing worse than random selection (AUC: 0.49 [default]). Use of a pharmacophore filter improved overall enrichment slightly (AUC: 0.67 [F]), while the addition of any style of constraints returned both overall and early enrichment to random (AUC: 0.49 [FP]) or null (FH, FI) values. In particular, using any number of interaction constraints returned no active compounds. Screening using carbohydrate decoys performed similarly to ZINC decoys in all respects, yielding at best only random selection of actives.

In summary, retrospective screening enrichment at three out of four bacterial lectins benefited from the use of pharmacophore filters and application of only one or two interaction constraints, with greater numbers of constraints proving detrimental to enrichment. Significantly, the use of carbohydrate decoys was demonstrated to be most appropriate in order to evaluate a virtual screening methodology for application to carbohydrate-binding targets.

## Prospective screening against *BambL*

After the developed virtual screening method demonstrated consistent enrichment at *BambL*, a prospective screen using the same method was conducted in search of novel inhibitory chemistry. The workflow shown in Figure 4 was used to select compounds, from the eMolecules Screening Compounds library and OTAVA Ltd. catalogue, with potential to inhibit the *BambL* protein. The workflow corresponds to the FI-2 screening method developed above, featuring a ligand-based pharmacophore filter to rapidly remove compounds lacking the atomic features required for binding, followed by a structure-based docking screen with binding interaction constraints to remove compounds unable to complement the *BambL* binding site architecture (see Supplementary Table S3 for constraint details). Manual selection according to the criteria outlined above identified 15 candidate compounds for experimental evaluation against *BambL* (Supplementary Figure S2).



**Figure 4.** Workflow of the virtual screening protocol used for prospective screening at BambL.

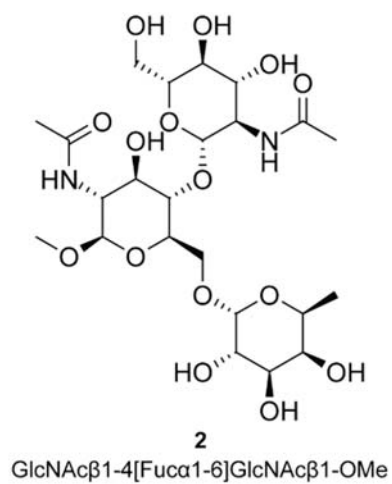
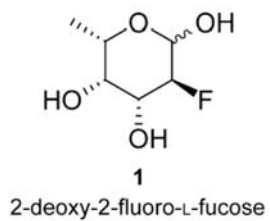
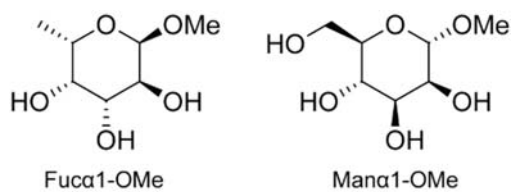
#### *Experimental validation of predicted compounds against BambL*

Experimental activity of the 15 selected compounds towards BambL was first evaluated using fluorescence polarization assays developed for this purpose. The fluorescence polarization assay method, originally developed by Hauck et al.<sup>25</sup> for studying binding to LecB (another fucose-binding bacterial lectin) and other lectins,<sup>41, 70-71</sup> has been applied to BambL. The assay measures displacement of a fluorescein-labeled reporter ligand FITC-fucose (N-(fluorescein-5-yl)-N'-( $\alpha$ -L-fucopyranosyl ethylen)-thiocarbamide) from the lectin binding site by the inhibitor under

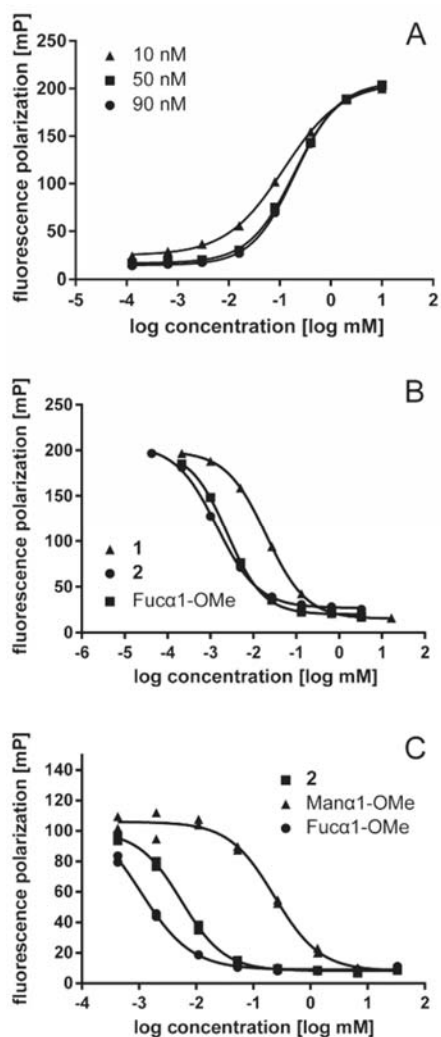
consideration. Here, this assay was used to evaluate the manually-selected virtual screening hits for inhibition of BambL and LecB.

Fuc $\alpha$ 1-OMe and Man $\alpha$ 1-OMe (Figure 5) were also evaluated by fluorescence polarization, for comparison. Two screening hits showed ability to bind to BambL (Figure 5). Compounds **1** and **2** are entirely composed of carbohydrate rings, and both contain a fucose saccharide. Compound **1** showed lower inhibition of BambL-probe binding (Figure 6; IC<sub>50</sub> 19.9  $\mu$ M), compared to Fuc $\alpha$ 1-OMe (IC<sub>50</sub> 2.4  $\mu$ M). Inhibition by compound **2** was similar (Figure 6; IC<sub>50</sub> 1.4  $\mu$ M) to Fuc $\alpha$ 1-OMe. Since compound **2** contains an unmodified fucose saccharide, similar inhibitory activity was expected. Inhibition of LecB-FITC-fucose binding by compound **2** was intermediate between Fuc $\alpha$ 1-OMe and Man $\alpha$ 1-OMe, while compound **1** was less active than either control saccharide (IC<sub>50</sub> >> 1 mM).





**Figure 5.** Fucose monosaccharide, mannose monosaccharide, and virtual screening hit compounds showing inhibitory activity towards BamBL.

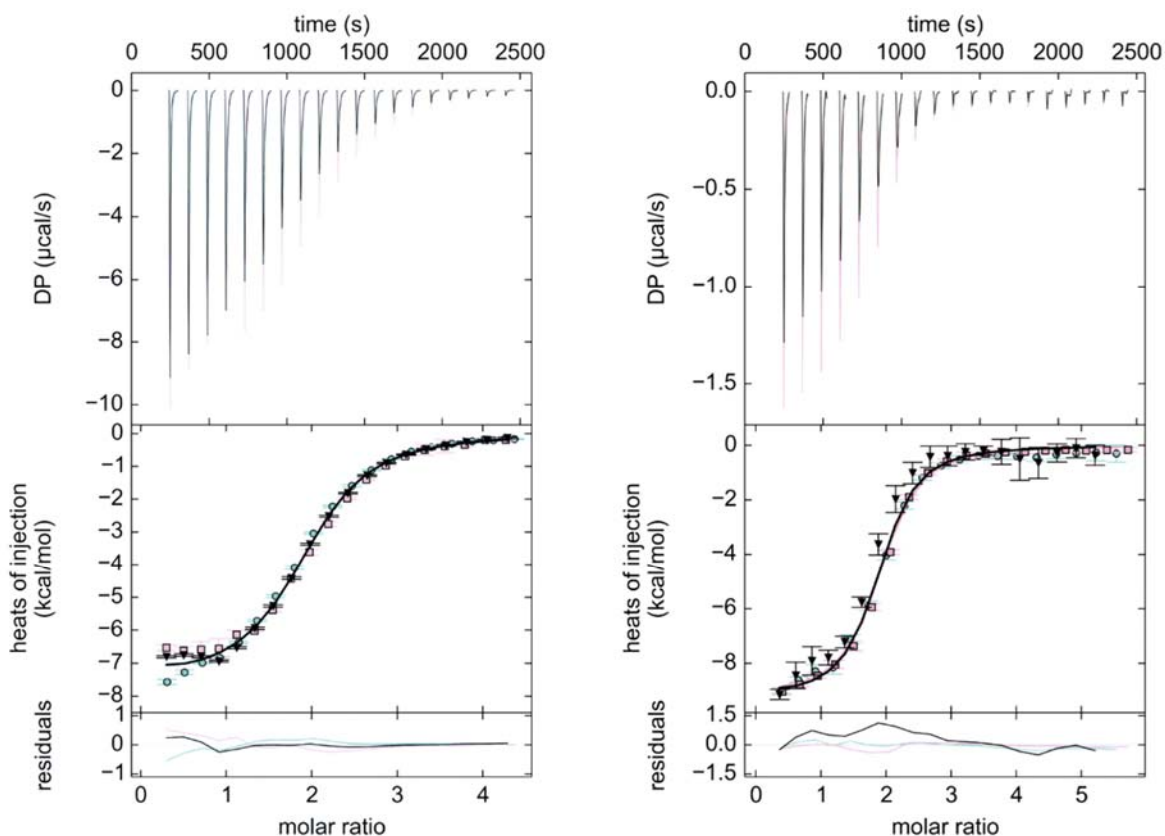


**Figure 6.** Fluorescence polarization evaluation of competitive inhibition of BambL and LecB by virtual screening hit compounds. All replicate measurement values are shown (single measurement for BambL, duplicate measurement for LecB). (A) Titration of FITC-fucose with BambL reveals specific binding at probe compound concentrations of 10 nM (triangles, EC<sub>50</sub> 120 μM), 50 nM (squares, EC<sub>50</sub> 195 μM), and 90 nM (circles, EC<sub>50</sub> 193 μM). (B) Compared to Fuca1-OMe (squares, IC<sub>50</sub> 2.4 μM), hit compounds inhibit the BambL-probe association with similar affinity (**2**, circles, IC<sub>50</sub> 1.4 μM) or poorer affinity (**1**, triangles, IC<sub>50</sub> 19.9 μM). (C) Inhibition of the LecB-probe association by **2** (squares, IC<sub>50</sub> 5.5 μM) is intermediate in affinity between Fuca1-OMe (circles, IC<sub>50</sub> 1.1 μM) and Manα1-OMe (triangles, IC<sub>50</sub> 230 μM). **1** displays poorer affinity for LecB, with IC<sub>50</sub> values greater than 1 mM (data not shown).

Isothermal titration calorimetry confirmed direct binding of both compounds to BamBL in a one-site fitting model (Table 2, Figure 7), revealing essentially identical interactions with both the intra- and inter-monomeric binding sites. Both compounds exhibited dissociation constants in the micromolar range (**1**:  $K_D$  of  $18.8 \pm 2.3 \mu\text{M}$ ; **2**:  $K_D$  of  $1.0 \pm 0.1 \mu\text{M}$ ), with lectin binding characterized by a greater enthalpic, rather than entropic component of binding. Despite the large difference in molecular size, the magnitude of the entropic component of binding is similar between **1** and **2**. BamBL inhibition data from these two solution-based assay methods (fluorescence polarization and ITC) display excellent agreement.

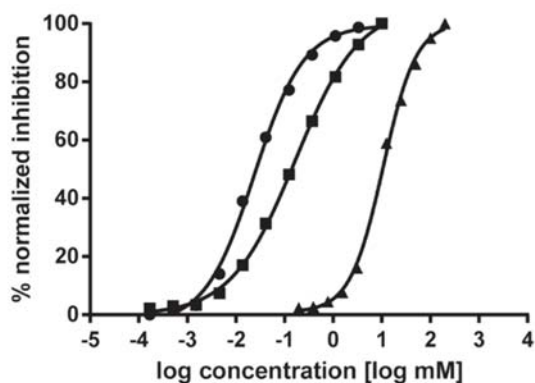
**Table 2.** Titration microcalorimetry data for binding to BamBL. Experiments were performed in triplicate; values shown are mean  $\pm$  SD.

Ligand	$n$	$K_D$ ( $\mu\text{M}$ )	$\Delta G$ ( $\text{kJ}\cdot\text{mol}^{-1}$ )	$\Delta H$ ( $\text{kJ}\cdot\text{mol}^{-1}$ )	$T\Delta S$ ( $\text{kJ}\cdot\text{mol}^{-1}$ )
Fuc $\alpha$ 1-OMe	$1.90 \pm 0.01$	$1.5 \pm 0.2$	$-33.2 \pm 0.3$	$-41.9 \pm 0.5$	$-8.7 \pm 0.7$
<b>1</b>	$1.92 \pm 0.09$	$18.8 \pm 2.3$	$-27.0 \pm 0.3$	$-31.3 \pm 1.6$	$-4.2 \pm 1.9$
<b>2</b>	$1.81 \pm 0.12$	$1.0 \pm 0.1$	$-34.2 \pm 0.3$	$-38.2 \pm 0.4$	$-4.0 \pm 0.8$



**Figure 7.** Thermograms in the upper panels were obtained by titration of BambL ( $50 \mu\text{M}$ ) with virtual screening hit compounds (**1**, left; **2**, right) at  $1 \text{ mM}$ . Binding isotherms in the middle panels show the heat released as a function of total ligand concentration (shown as a molar ratio to the protein concentration). The solid lines indicate the best least-squares regression fit line to the experimental data, using a one-site binding model. Each titration was conducted in triplicate, with each replicate fitted individually (shown as black triangles, blue squares, and grey circles). Error bars show the variation in integrated heat peak area encountered during peak-shape analysis. Residuals in the bottom panels indicate the quality of the regression fit for the correspondingly-coloured replicate.

In surface plasmon resonance assays (Figure 8), BambL inhibition by virtual screening hits was less detectable (**1**: IC<sub>50</sub> of 166 μM; **2**: IC<sub>50</sub> of 10.7 mM). The inhibition efficiency (IC<sub>50</sub>) of these hit compounds was lower than apparent affinity in solution assays. This could be explained by the strong avidity of BambL for multivalent surfaces, which is difficult to compete with when using monovalent ligands.



**Figure 8.** Normalized inhibition by virtual screening hit compounds of BambL-fucosylated polyacrylamide association measured by SPR. Fucα1-OMe showed the greatest inhibitory ability (circles, IC<sub>50</sub> 24.1 μM), followed by **1** (squares, IC<sub>50</sub> 166 μM) and **2** (triangles, IC<sub>50</sub> 10.7 mM).

## Discussion

<sup>767778</sup>The majority of virtual screening studies against carbohydrate-binding proteins do not evaluate the applied method nor use negative controls/decoys to establish retrospective enrichment.<sup>79-84</sup> Validation is critical for the improvement of screening methods against these targets, and allows users to judge the transferability of reported successes. Where retrospective screening has been performed, enrichment indicators used include the raw number of hits returned,<sup>85</sup> the ratio of retrieved actives to total actives,<sup>86-87</sup> and the widely-used 1% Enrichment Factor<sup>88</sup> and ROC AUC.<sup>8987-8889</sup> [EY1]

In this work, a virtual screening protocol for bacterial lectins was iteratively developed and evaluated using the ROC Enrichment Factor and ROC AUC metrics, with calculation of standard errors. The influence of decoy types on retrospective virtual screening enrichment was also examined. The virtual screening workflow was applied prospectively to BambL for identification of novel inhibitors; two of the identified hits consistently showed lectin inhibition in biophysical assays.

Decoy selection in this study was guided by two related considerations: avoiding known epitopes in decoys while matching decoys to actives in physical property space. Specifically, carbohydrate decoys were built from monosaccharides, excluding those with known specificity for a given lectin, as shown by experimental glycan array data. In that sense, these decoys are confirmed non-binders. Non-carbohydrate decoys were drawn from ZINC, and are assumed to be non-binders based on the absence of carbohydrate epitopes in their structures. This type of assumption is a known caveat in virtual screening studies.<sup>90</sup> With respect to the second consideration, many have noted that retrospective screening enrichment depends on the resemblance between actives and decoys.<sup>91-92</sup> Our results are consistent with this observation, showing that enrichment in retrospective screens against bacterial lectins was influenced by the nature of the decoy compounds. Enrichment obtained using all actives may be artificially inflated by poor physical property-matching (either between ZINC-derived decoys and carbohydrate actives, or between carbohydrate decoys and non-carbohydrate actives). In contrast, the poorer enrichment obtained using carbohydrate decoys with carbohydrate actives in these screens provides a more genuine measure of the expected enrichment in a prospective virtual screening against carbohydrate-binding proteins.

Iterative development of a virtual screening workflow explored combinations of ligand-based and structure-based screening techniques. As noted above, the use of binding pocket constraints and

interaction-based constraints have not been previously applied to carbohydrate-binding proteins. In general, pharmacophore-based filtering improved enrichment compared to the entirely default screening setup. Subsequent addition of a binding pocket constraint also provided a small increase to enrichment. Interaction-derived constraints were initially applied in full force (i.e., requiring simultaneous formation of four binding interactions), causing a decrease in both early and overall enrichments. However, once the number of simultaneously-required constraints was reduced, enrichment measures increased again, affording the greatest enrichment for BambL and FimH. These retrospective screening experiments demonstrate that a combination of ligand-based and structure-based techniques enhance enrichment against against carbohydrate-binding proteins.

In our previous efforts,<sup>32</sup> we have determined the atomic interactions made by fucosylated carbohydrates recognized by BambL. BambL-carbohydrate binding involved a network of hydrogen bonds within the fucose-binding pocket as well as interactions with satellite solvent-exposed residues outside of the binding cleft. The interacting residues identified in our earlier work were now used to define interaction constraints for prospective virtual screening. For BambL, the two-match interaction constraint screening method (FI-2) was genuinely able to prioritize binders over non-binders. Application of the FI-2 screening workflow to BambL identified 15 compounds with predicted binding affinity for BambL. Lectin inhibition was confirmed for two of these using fluorescence polarization, SPR, and ITC assays.

Compound **2** is a component of naturally occurring N-glycans such as Nod factors (signalling molecules produced by symbiotic bacteria at legume roots).<sup>93</sup> N-glycans carrying the fucosylated chitobiose epitope, present in compound **2**, display binding affinity toward BambL in glycan array data from the Consortium for Functional Glycomics. In solution-based assays (fluorescence polarization and ITC), compound **1** (2-deoxy-2-fluoro-L-fucose) displays excellent affinity for

BambL in the same range as natural fucose. Substitution of carbohydrate hydroxyl groups by fluorine has been demonstrated in some cases to enhance affinity for protein receptors (e.g., *Toxoplasma gondii* TgMIC1,<sup>94</sup> *Mycobacterium tuberculosis* UDB-galactopyranose<sup>95</sup>). Since fluorine substitution can give rise to lectin specificity, we tested the binding of compound **1** to a second fucose-binding bacterial lectin, LecB from *P. aeruginosa*.

The fucomimetic compound **1** displayed a difference in inhibitory activity between lectins BambL and LecB. In fluorescence polarization assays (Figure 6), compound **1** inhibited BambL-probe binding with an IC<sub>50</sub> of 19.9 μM, while barely inhibiting LecB-probe binding (IC<sub>50</sub> >> 1 mM), an observation that confirms previous experiments with related 1,2-dideoxy-2-fluoro-L-fucose and LecB.<sup>41</sup> A structural rationale for this observation can be suggested. In the LecB-fucose complex,<sup>96-98</sup> the fucose 2-position hydroxyl oxygen directly coordinates one of the two binding site calcium ions and forms a hydrogen bond to Asp99. Thus, the substitution of hydroxyl by fluorine is not favorable for binding. The BambL-fucose complex features a hydrogen bond donated by the backbone amide nitrogen of Asp38 to the fucose 2-position hydroxyl.<sup>28</sup> Therefore, in this case, fluorine can act as a weak hydrogen bond acceptor, making the substitution more acceptable. Compound **1** is therefore an excellent hit compound for the development of a specific inhibitor for BambL and related lectins.

## Conclusion

Through iterative development and retrospective evaluation, a virtual screening workflow for bacterial lectins was developed. For two target systems, BambL and FimH, molecular properties of decoy compounds were found to exert influence upon retrospective enrichment. This finding



showed that enrichment in virtual screening against carbohydrate-binding proteins could be artificially enhanced by poor decoy camouflage. In addition, the presence of non-carbohydrate actives was also conducive to over-estimated enrichment. Combined ligand-based and structure-based screening techniques were demonstrated to improve retrospective enrichment via the use of interaction-based constraints. For BambL, enforcing two interaction constraint matches provided the greatest overall enrichment in retrospective screening. A prospective virtual screen for compounds able to bind to BambL yielded two inhibitory molecules. Compound **1** (2-deoxy-2-fluoro-L-fucose) displayed slightly reduced affinity for BambL compared to L-fucose, but multiple orders-of-magnitude reduction in affinity for LecB. The fucose 2-position is therefore suggested as an opportunity for designing selectivity for BambL in preference to other fucose-binding lectins.

## **Acknowledgments**

The authors acknowledge excellent technical assistance from Dirk Hauck (HIPS). TD acknowledges support received from the Monash University Postgraduate Publications Award. AI and SP acknowledge support from the French ANR through Glyco@Alps (ANR-15-IDEX-02) and Labex ARCANE (ANR-11-LABX-003). AT acknowledges funding from the Helmholtz Association of German Research Centers (grant no. VH-NG-934) and the Deutsche Forschungsgemeinschaft (grant no. Ti756/2-1). PR acknowledges an RMIT University Vice Chancellor's Senior Research Fellowship and the Victorian State Government Operational Infrastructure Support grant to the Burnet Institute.

## References

1. Yuriev, E.; Holien, J.; Ramsland, P. A. Improvements, Trends, and New Ideas in Molecular Docking: 2012–2013 in Review. *J. Mol. Recognit.* **2015**, *28* (10), 581-604. DOI: 10.1002/jmr.2471.
2. Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P. A.; Yuriev, E. Molecular Docking of Carbohydrate Ligands to Antibodies: Structural Validation against Crystal Structures. *J. Chem. Inf. Model.* **2009**, *49* (12), 2749-2760. DOI: 10.1021/ci900388a.
3. Klemm, P.; Schembri, M. A. Bacterial Adhesins: Function and Structure. *Int. J. Med. Microbiol.* **2000**, *290* (1), 27-35. DOI: 10.1016/S1438-4221(00)80102-2.
4. Kline, K. A.; Fälker, S.; Dahlberg, S.; Normark, S.; Henriques-Normark, B. Bacterial Adhesins in Host-Microbe Interactions. *Cell Host Microbe* **2009**, *5* (6), 580-592. DOI: 10.1016/j.chom.2009.05.011.
5. Dickey, S. W.; Cheung, G. Y. C.; Otto, M. Different Drugs for Bad Bugs: Antivirulence Strategies in the Age of Antibiotic Resistance. *Nat. Rev. Drug Discov.* **2017**, *16* (7), 457-471. DOI: 10.1038/nrd.2017.23.
6. Sattin, S.; Bernardi, A. Glycoconjugates and Glycomimetics as Microbial Anti-Adhesives. *Trends Biotechnol.* **2016**, *34* (6), 483-495. DOI: 10.1016/j.tibtech.2016.01.004.
7. Kranjčec, B.; Papeš, D.; Altarac, S. D-Mannose Powder for Prophylaxis of Recurrent Urinary Tract Infections in Women: A Randomized Clinical Trial. *World. J. Urol.* **2014**, *32* (1), 79-84. DOI: 10.1007/s00345-013-1091-6.
8. Porru, D.; Parmigiani, D.; Tinelli, D.; Barletta, D.; Choussos, D.; Di Franco, C.; Bobbi, V.; Bassi, S.; Miller, O.; Gardella, B.; Nappi, R. E.; Spinillo, A.; Rovereto, B. Oral D-Mannose in Recurrent Urinary Tract Infections in Women: A Pilot Study. *J. Clin. Urol.* **2014**, *7* (3), 208-213. DOI: 10.1177/2051415813518332.
9. von Bismarck, P.; Schneppenheim, R.; Schumacher, U. Successful Treatment of Pseudomonas Aeruginosa Respiratory Tract Infection with a Sugar Solution--a Case Report on a Lectin Based Therapeutic Principle. *Klin. Padiatr.* **2001**, *213* (5), 285-7. DOI: 10.1055/s-2001-17220.
10. Hauber, H. P.; Schulz, M.; Pforte, A.; Mack, D.; Zabel, P.; Schumacher, U. Inhalation with Fucose and Galactose for Treatment of Pseudomonas Aeruginosa in Cystic Fibrosis Patients. *Int. J. Med. Sci.* **2008**, *5* (6), 371-6. DOI: 10.7150/ijms.5.371.
11. Touaibia, M.; Wellens, A.; Shiao, T. C.; Wang, Q.; Sirois, S.; Bouckaert, J.; Roy, R. Mannosylated G(0) Dendrimers with Nanomolar Affinities to Escherichia Coli Fimh. *ChemMedChem* **2007**, *2* (8), 1190-1201. DOI: 10.1002/cmdc.200700063.
12. Visini, R.; Jin, X.; Bergmann, M.; Michaud, G.; Pertici, F.; Fu, O.; Pukin, A.; Branson, T. R.; Thies-Weesie, D. M. E.; Kemmink, J.; Gillon, E.; Imberty, A.; Stocker, A.; Darbre, T.; Pieters, R. J.; Reymond, J.-L. Structural Insight into Multivalent Galactoside Binding to Pseudomonas Aeruginosa Lectin Leca. *ACS Chem. Biol.* **2015**, 2455-2462. DOI: 10.1021/acschembio.5b00302.
13. Caumes, C.; Gillon, E.; Legeret, B.; TAILLEFUMIER, C.; Imberty, A.; Faure, S. Multivalent Thioglycopeptoids Via Photoclick Chemistry: Potent Affinities Towards Leca and Bc21-a Lectins. *Chem. Commun.* **2015**, 12301-12304. DOI: 10.1039/C5CC04646G.
14. Boukerb, A. M.; Rousset, A.; Galanos, N.; Méar, J.-B.; Thepaut, M.; Grandjean, T.; Gillon, E.; Cecioni, S.; Abderrahmen, C.; Faure, K.; Redelberger, D.; Kipnis, E.; Dessein, R.; Havet, S.; Darblade, B.; Matthews, S. E.; de Bentzmann, S.; Guéry, B.; Cournoyer, B.; Imberty, A.; Vidal, S. Anti-Adhesive Properties of Glycoclusters against Pseudomonas Aeruginosa Lung Infection. *J. Med. Chem.* **2014**, 10275–10289. DOI: 10.1021/jm500038p.
15. Ligeour, C.; Audfray, A.; Gillon, E.; Meyer, A.; Galanos, N.; Vidal, S.; Vasseur, J.-J.; Imberty, A.; Morvan, F. Synthesis of Branched-Phosphodiester and Mannose-Centered Fucosylated Glycoclusters and Their Binding Studies with Burkholderia Ambifaria Lectin (Bambl). *RSC Adv.* **2013**, *3* (42), 19515-19524. DOI: 10.1039/C3RA43807D.

16. Buffet, K.; Gillon, E.; Holler, M.; Nierengarten, J. F.; Imberty, A.; Vincent, S. P. Fucosylated Derivatives of Rsl and Lecb, Two Bacterial Lectins. *Org. Biomol. Chem.* **2015**, *13* (23), 6482-6492. DOI: 10.1039/c5ob00689a.
17. Cecioni, S.; Imberty, A.; Vidal, S. Glycomimetics Versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands. *Chem. Rev.* **2015**, *115* (1), 525-561. DOI: 10.1021/cr500303t.
18. Wagner, S.; Sommer, R.; Hinsberger, S.; Lu, C.; Hartmann, R. W.; Empting, M.; Titz, A. Novel Strategies for the Treatment of Pseudomonas Aeruginosa Infections. *J. Med. Chem.* **2016**, *59* (13), 5929-5969. DOI: 10.1021/acs.jmedchem.5b01698.
19. Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezençon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. FimH Antagonists for the Oral Treatment of Urinary Tract Infections: From Design and Synthesis to in Vitro and in Vivo Evaluation. *J. Med. Chem.* **2010**, *53* (24), 8627-8641. DOI: 10.1021/jm101011y.
20. Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. Treatment and Prevention of Urinary Tract Infection with Orally Active FimH Inhibitors. *Sci. Transl. Med.* **2011**, *3* (109), 109ra115-109ra115. DOI: 10.1126/scitranslmed.3003021.
21. Rodrigue, J.; Ganne, G.; Blanchard, B.; Saucier, C.; Giguere, D.; Shiao, T. C.; Varrot, A.; Imberty, A.; Roy, R. Aromatic Thioglycoside Inhibitors against the Virulence Factor Leca from Pseudomonas Aeruginosa. *Org. Biomol. Chem.* **2013**, *11* (40), 6906-18. DOI: 10.1039/c3ob41422a.
22. Kadam, R. U.; Garg, D.; Schwartz, J.; Visini, R.; Sattler, M.; Stocker, A.; Darbre, T.; Reymond, J. L. Ch-Pi "T-Shape" Interaction with Histidine Explains Binding of Aromatic Galactosides to Pseudomonas Aeruginosa Lectin Leca. *ACS Chem. Biol.* **2013**, *8* (9), 1925-30. DOI: 10.1021/cb400303w.
23. Wagner, S.; Hauck, D.; Hoffmann, M.; Sommer, R.; Joachim, I.; Müller, R.; Imberty, A.; Varrot, A.; Titz, A. Covalent Lectin Inhibition and Application in Bacterial Biofilm Imaging. *Angew. Chem. Int. Ed. Engl.* **2017**, *56* (52), 16559-16564. DOI: 10.1002/anie.201709368.
24. Sommer, R.; Wagner, S.; Rox, K.; Varrot, A.; Hauck, D.; Wamhoff, E.-C.; Schreiber, J.; Ryckmans, T.; Brunner, T.; Rademacher, C.; Hartmann, R. W.; Brönstrup, M.; Imberty, A.; Titz, A. Glycomimetic, Orally Bioavailable Lecb Inhibitors Block Biofilm Formation of Pseudomonas Aeruginosa. *J. Am. Chem. Soc.* **2017**, *139*, 2537-2545. DOI: 10.1021/jacs.7b11133.
25. Hauck, D.; Joachim, I.; Frommeyer, B.; Varrot, A.; Philipp, B.; Moeller, H. M.; Imberty, A.; Exner, T. E.; Titz, A. Discovery of Two Classes of Potent Glycomimetic Inhibitors of Pseudomonas Aeruginosa Lecb with Distinct Binding Modes. *ACS Chem. Biol.* **2013**, *8* (8), 1775-1784. DOI: 10.1021/cb400371r.
26. Sommer, R.; Hauck, D.; Varrot, A.; Wagner, S.; Audfray, A.; Prestel, A.; Moeller, H. M.; Imberty, A.; Titz, A. Cinnamide Derivatives of D-Mannose as Inhibitors of the Bacterial Virulence Factor Lecb from Pseudomonas Aeruginosa. *ChemistryOpen* **2015**, *4* (6), 756-767. DOI: 10.1002/open.201500162.
27. Richichi, B.; Imberty, A.; Gillon, E.; Bosco, R.; Sutkeviciute, I.; Fieschi, F.; Nativi, C. Synthesis of a Selective Inhibitor of a Fucose Binding Bacterial Lectin from Burkholderia Ambifaria. *Org. Biomol. Chem.* **2013**, *11* (24), 4086-4094. DOI: 10.1039/c3ob40520f.
28. Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoen-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerova, M.; Le Pendu, J.; Romer, W.; Varrot, A.; Imberty, A. Fucose-Binding Lectin from Opportunistic Pathogen Burkholderia Ambifaria Binds to Both Plant and Human Oligosaccharidic Epitopes. *J. Biol. Chem.* **2012**, *287* (6), 4335-47. DOI: 10.1074/jbc.M111.314831.
29. Coenye, T.; Mahenthiralingam, E.; Henry, D.; LiPuma, J. J.; Laevens, S.; Gillis, M.; Speert, D. P.; Vandamme, P. Burkholderia Ambifaria Sp Nov., a Novel Member of the Burkholderia Cepacia Complex Including Biocontrol and Cystic Fibrosis-Related Isolates. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 1481-1490. DOI: 10.1099/00207713-51-4-1481.
30. Glick, M. C.; Kothari, V. A.; Liu, A. H.; Stoykova, L. I.; Scanlin, T. F. Activity of Fucosyltransferases and Altered Glycosylation in Cystic Fibrosis Airway Epithelial Cells. *Biochimie* **2001**, *83* (8), 743-747. DOI: 10.1016/s0300-9084(01)01323-2.

31. Venkatakrisnan, V.; Packer, N. H.; Thaysen-Andersen, M. Host Mucin Glycosylation Plays a Role in Bacterial Adhesion in Lungs of Individuals with Cystic Fibrosis. *Expert Rev. Respir. Med.* **2013**, *7* (5), 553-576. DOI: 10.1586/17476348.2013.837752.
32. Dingjan, T.; Imberty, A.; Pérez, S.; Yuriev, E.; Ramsland, P. A. Molecular Simulations of Carbohydrates with a Fucose-Binding Burkholderia Ambifaria Lectin Suggest Modulation by Surface Residues Outside the Fucose-Binding Pocket. *Front. Pharmacol.* **2017**, *8* (393). DOI: 10.3389/fphar.2017.00393.
33. Yuriev, E.; Agostino, M.; Ramsland, P. A. Challenges and Advances in Computational Docking: 2009 in Review. *J. Mol. Recognit.* **2011**, *24* (2), 149-164. DOI: 10.1002/jmr.1077.
34. Yuriev, E.; Ramsland, P. A. Latest Developments in Molecular Docking: 2010-2011 in Review. *J. Mol. Recognit.* **2013**, *26*, 215-39. DOI: 10.1002/jmr.2266.
35. Grant, O. C.; Woods, R. J. Recent Advances in Employing Molecular Modelling to Determine the Specificity of Glycan-Binding Proteins. *Curr. Opin. Struct. Biol.* **2014**, *28* (0), 47-55. DOI: 10.1016/j.sbi.2014.07.001.
36. Lameignere, E.; Malinowska, L.; Slavikova, M.; Duchaud, E.; Mitchell, E. P.; Varrot, A.; Sedo, O.; Imberty, A.; Wimmerova, M. Structural Basis for Mannose Recognition by a Lectin from Opportunistic Bacteria Burkholderia Cenocepacia. *Biochem. J.* **2008**, *411*, 307-318. DOI: 10.1042/bj20071276.
37. Sulak, O.; Cioci, G.; Delia, M.; Lahmann, M.; Varrot, A.; Imberty, A.; Wimmerova, M. A Tnf-Like Trimeric Lectin Domain from Burkholderia Cenocepacia with Specificity for Fucosylated Human Histo-Blood Group Antigens. *Structure* **2010**, *18* (1), 59-72. DOI: 10.1016/j.str.2009.10.021.
38. Hung, C.-S.; Bouckaert, J.; Hung, D.; Pinkner, J.; Widberg, C.; DeFusco, A.; Auguste, C. G.; Strouse, R.; Langermann, S.; Waksman, G.; Hultgren, S. J. Structural Basis of Tropism of Escherichia Coli to the Bladder During Urinary Tract Infection. *Mol. Microbiol.* **2002**, *44* (4), 903-915. DOI: 10.1046/j.1365-2958.2002.02915.x.
39. Cioci, G.; Mitchell, E. P.; Gautier, C.; Wimmerova, M.; Sudakevitz, D.; Perez, S.; Gilboa-Garber, N.; Imberty, A. Structural Basis of Calcium and Galactose Recognition by the Lectin Pa-II of Pseudomonas Aeruginosa. *FEBS Lett.* **2003**, *555* (2), 297-301. DOI: 10.1016/s0014-5793(03)01249-3.
40. Garber, N.; Guempel, U.; Belz, A.; Gilboagarber, N.; Doyle, R. J. On the Specificity of the D-Galactose-Binding Lectin (Pa-I) of Pseudomonas Aeruginosa and Its Strong Binding to Hydrophobic Derivatives of D-Galactose and Thiogalactose. *Biochim. Biophys. Acta* **1992**, *1116* (3), 331-333. DOI: 10.1016/0304-4165(92)90048-y.
41. Beshr, G.; Sommer, R.; Hauck, D.; Siebert, D. C. B.; Hofmann, A.; Imberty, A.; Titz, A. Development of a Competitive Binding Assay for the Burkholderia Cenocepacia Lectin Bc21-a and Structure Activity Relationship of Natural and Synthetic Inhibitors. *MedChemComm* **2016**, *7* (3), 519-530. DOI: 10.1039/c5md00557d.
42. Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhner, M.; Hung, C. S.; Pinkner, J.; Slattegard, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. Receptor Binding Studies Disclose a Novel Class of High-Affinity Inhibitors of the Escherichia Coli Fimh Adhesin. *Mol. Microbiol.* **2005**, *55* (2), 441-55. DOI: 10.1111/j.1365-2958.2004.04415.x.
43. Pang, L.; Kleeb, S.; Lemme, K.; Rabbani, S.; Scharenberg, M.; Zalewski, A.; Schadler, F.; Schwaradt, O.; Ernst, B. Fimh Antagonists: Structure-Activity and Structure-Property Relationships for Biphenyl Alpha-D-Mannopyranosides. *ChemMedChem* **2012**, *7* (8), 1404-22. DOI: 10.1002/cmdc.201200125.
44. Han, Z.; Pinkner, J. S.; Ford, B.; Chorell, E.; Crowley, J. M.; Cusumano, C. K.; Campbell, S.; Henderson, J. P.; Hultgren, S. J.; Janetka, J. W. Lead Optimization Studies on Fimh Antagonists: Discovery of Potent and Orally Bioavailable Ortho-Substituted Biphenyl Mannosides. *J. Med. Chem.* **2012**, *55* (8), 3945-3959. DOI: 10.1021/jm300165m.
45. Firon, N.; Ashkenazi, S.; Mirelman, D.; Ofek, I.; Sharon, N. Aromatic Alpha-Glycosides of Mannose Are Powerful Inhibitors of the Adherence of Type 1 Fimbriated Escherichia Coli to Yeast and Intestinal Epithelial Cells. *Infect. Immun.* **1987**, *55* (2), 472-476. DOI: NONE.

46. Sperling, O.; Fuchs, A.; Lindhorst, T. K. Evaluation of the Carbohydrate Recognition Domain of the Bacterial Adhesin Fimh: Design, Synthesis and Binding Properties of Mannoside Ligands. *Org. Biomol. Chem.* **2006**, *4* (21), 3913-3922. DOI: 10.1039/b610745a.
47. Firon, N.; Ofek, I.; Sharon, N. Carbohydrate Specificity of the Surface Lectins of Escherichia Coli, Klebsiella Pneumoniae, and Salmonella Typhimurium. *Carbohydr. Res.* **1983**, *120*, 235-249. DOI: 10.1016/0008-6215(83)88019-7.
48. Han, Z.; Pinkner, J. S.; Ford, B.; Obermann, R.; Nolan, W.; Wildman, S. A.; Hobbs, D.; Ellenberger, T.; Cusumano, C. K.; Hultgren, S. J.; Janetka, J. W. Structure-Based Drug Design and Optimization of Mannoside Bacterial Fimh Antagonists. *J. Med. Chem.* **2010**, *53* (12), 4779-4792. DOI: 10.1021/jm100438s.
49. Jiang, X.; Abgottspon, D.; Kleeb, S.; Rabbani, S.; Scharenberg, M.; Wittwer, M.; Haug, M.; Schwardt, O.; Ernst, B. Antiadhesion Therapy for Urinary Tract Infections—a Balanced Pk/Pd Profile Proved to Be Key for Success. *J. Med. Chem.* **2012**, *55* (10), 4700-4713. DOI: 10.1021/jm300192x.
50. Abgottspon, D.; Rölli, G.; Hosch, L.; Steinhuber, A.; Jiang, X.; Schwardt, O.; Cutting, B.; Smiesko, M.; Jenal, U.; Ernst, B.; Trampuz, A. Development of an Aggregation Assay to Screen Fimh Antagonists. *J. Microbiol. Methods* **2010**, *82* (3), 249-255. DOI: 10.1016/j.mimet.2010.06.015.
51. Reymond, J.-L.; Bergmann, M.; Darbre, T. Glycopeptide Dendrimers as Pseudomonas Aeruginosa Biofilm Inhibitors. *Chem. Soc. Rev.* **2013**, *42* (11), 4814-4822. DOI: 10.1039/C3CS35504G.
52. Cecioni, S.; Praly, J.-P.; Matthews, S. E.; Wimmerová, M.; Imberty, A.; Vidal, S. Rational Design and Synthesis of Optimized Glycoclusters for Multivalent Lectin–Carbohydrate Interactions: Influence of the Linker Arm. *Chem. - Eur. J.* **2012**, *18* (20), 6250-6263. DOI: 10.1002/chem.201200010.
53. Kadam, R. U.; Bergmann, M.; Hurley, M.; Garg, D.; Cacciarini, M.; Swiderska, M. A.; Nativi, C.; Sattler, M.; Smyth, A. R.; Williams, P.; Camara, M.; Stocker, A.; Darbre, T.; Reymond, J.-L. A Glycopeptide Dendrimer Inhibitor of the Galactose-Specific Lectin Leca and of Pseudomonas Aeruginosa Biofilms. *Angew. Chem. Int. Ed. Engl.* **2011**, *50* (45), 10631-10635. DOI: 10.1002/anie.201104342.
54. Consortium for Functional Glycomics. Functional Glycomics Gateway. [www.functionalglycomics.org](http://www.functionalglycomics.org) (accessed 11/07/2016).
55. Schrödinger Release 2015-3: Maestro, version 10.3, Schrödinger, LLC, New York, NY, 2014.
56. Sterling, T.; Irwin, J. J. Zinc 15 – Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55* (11), 2324-2337. DOI: 10.1021/acs.jcim.5b00559.
57. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. Zinc: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757-1768. DOI: 10.1021/ci3001277.
58. Irwin, J. J.; Shoichet, B. K. Zinc – a Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177-182. DOI: 10.1021/ci049714+.
59. Silico - a Perl Molecular Modelling Toolkit, version Monash University, Melbourne, 2011.
60. Schrödinger Release 2015-3: Qikprop, version 4.5, Schrödinger, LLC, New York, NY, 2015.
61. Schrödinger Release 2015-3: Ligprep, version 3.5, Schrödinger, LLC, New York, NY, 2015.
62. Instant Jchem, version 6.0.4, ChemAxon, 2013.
63. Schrödinger Release 2015-3: Phase, version 4.4, Schrödinger, LLC, New York, NY, 2015.
64. Schrödinger Release 2014-1: Protein Preparation Wizard; Epik, version 3.6, Schrödinger, LLC, New York, NY, 2014.
65. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. Mdtraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528-1532. DOI: 10.1016/j.bpj.2015.08.015.
66. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739-1749. DOI: 10.1021/jm0306430.

67. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750-1759. DOI: 10.1021/jm030644s.
68. McGann, M.; Nicholls, A.; Enyedy, I. The Statistics of Virtual Screening and Lead Optimization. *J. Comput.-Aided Mol. Des.* **2015**, *29* (10), 923-936. DOI: 10.1007/s10822-015-9861-4.
69. Sommer, R.; Wagner, S.; Varrot, A.; Nycholat, C. M.; Khaledi, A.; Haussler, S.; Paulson, J. C.; Imberty, A.; Titz, A. The Virulence Factor Lecb Varies in Clinical Isolates: Consequences for Ligand Binding and Drug Discovery. *Chem. Sci.* **2016**, *7* (8), 4990-5001. DOI: 10.1039/C6SC00696E.
70. Joachim, I.; Rikker, S.; Hauck, D.; Ponader, D.; Boden, S.; Sommer, R.; Hartmann, L.; Titz, A. Development and Optimization of a Competitive Binding Assay for the Galactophilic Low Affinity Lectin Leca from *Pseudomonas Aeruginosa*. *Org. Biomol. Chem.* **2016**, *14* (33), 7933-7948. DOI: 10.1039/C6OB01313A.
71. Beshr, G.; Sikandar, A.; Jemiller, E.-M.; Klymiuk, N.; Hauck, D.; Wagner, S.; Wolf, E.; Koehnke, J.; Titz, A. Photorhabdus Luminescens Lectin a (Plla) - a New Probe for Detecting A-Galactoside-Terminating Glycoconjugates. *J. Biol. Chem.* **2017**. DOI: 10.1074/jbc.M117.812792.
72. Agostino, M.; Yuriev, E.; Ramsland, P. A. A Computational Approach for Exploring Carbohydrate Recognition by Lectins in Innate Immunity. *Front. Immunol.* **2011**, *2*, 23. DOI: 10.3389/fimmu.2011.00023.
73. Nurisso, A.; Kozmon, S.; Imberty, A. 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.* **2008**, *34* (4), 469-479. DOI: 10.1080/08927020701697709.
74. Topin, J.; Arnaud, J.; Sarkar, A.; Audfray, A.; Gillon, E.; Perez, S.; Jamet, H.; Varrot, A.; Imberty, A.; Thomas, A. Deciphering the Glycan Preference of Bacterial Lectins by Glycan Array and Molecular Docking with Validation by Microcalorimetry and Crystallography. *PLoS One* **2013**, *8*, e71149. DOI: 10.1371/journal.pone.0071149.
75. Kadam, R. U.; Bergmann, M.; Garg, D.; Gabrieli, G.; Stocker, A.; Darbre, T.; Reymond, J.-L. Structure-Based Optimization of the Terminal Tripeptide in Glycopeptide Dendrimer Inhibitors of *Pseudomonas Aeruginosa* Biofilms Targeting Leca. *Chem. - Eur. J.* **2013**, *19* (50), 17054-17063. DOI: 10.1002/chem.201302587.
76. Xuan, P.; Zhang, Y.; Tzeng, T.-r. J.; Wan, X.-F.; Luo, F. A Quantitative Structure-Activity Relationship (Qsar) Study on Glycan Array Data to Determine the Specificities of Glycan-Binding Proteins. *Glycobiology* **2012**, *22* (4), 552-560. DOI: 10.1093/glycob/cwr163.
77. Imberty, A.; Mollicone, R.; Mikros, E.; Carrupt, P.-A.; Pérez, S.; Oriol, R. How Do Antibodies and Lectins Recognize Histo-Blood Group Antigens? A 3d-Qsar Study by Comparative Molecular Field Analysis (Comfa). *Bioorg. Med. Chem.* **1996**, *4* (11), 1979-1988. DOI: 10.1016/S0968-0896(96)00179-4.
78. Kletter, D.; Singh, S.; Bern, M.; Haab, B. B. Global Comparisons of Lectin-Glycan Interactions Using a Database of Analyzed Glycan Array Data. *Mol. Cell. Proteomics* **2013**, *12* (4), 1026-1035. DOI: 10.1074/mcp.M112.026641.
79. Park, H.; Hwang, K. Y.; Kim, Y. H.; Oh, K. H.; Lee, J. Y.; Kim, K. Discovery and Biological Evaluation of Novel A-Glucosidase Inhibitors with in Vivo Antidiabetic Effect. *Bioorg. Med. Chem. Lett.* **2008**, *18* (13), 3711-3715. DOI: 10.1016/j.bmcl.2008.05.056.
80. Park, H.; Hwang, K. Y.; Oh, K. H.; Kim, Y. H.; Lee, J. Y.; Kim, K. Discovery of Novel A-Glucosidase Inhibitors Based on the Virtual Screening with the Homology-Modeled Protein Structure. *Bioorg. Med. Chem.* **2008**, *16* (1), 284-292. DOI: 10.1016/j.bmc.2007.09.036.
81. Hakamata, W.; Ishikawa, R.; Ushijima, Y.; Tsukagoshi, T.; Tamura, S.; Hirano, T.; Nishio, T. Virtual Ligand Screening of Alpha-Glucosidase: Identification of a Novel Potent Noncarbohydrate Mimetic Inhibitor. *Bioorg. Med. Chem. Lett.* **2012**, *22* (1), 62-64. DOI: 10.1016/j.bmcl.2011.11.084.
82. Bai, Y.; Watt, B.; Wahome, P. G.; Mantis, N. J.; Robertus, J. D. Identification of New Classes of Ricin Toxin Inhibitors by Virtual Screening. *Toxicon.* **2010**, *56* (4), 526-534. DOI: 10.1016/j.toxicon.2010.05.009.

83. Barra, P. A.; Jiménez, V. A.; Gavin, J. A.; Daranas, A. H.; Alderete, J. B. Discovery of New E-Selectin Inhibitors by Virtual Screening, Fluorescence Binding Assays, and Std Nmr Experiments. *ChemMedChem* **2016**, *11* (9), 1008-1014. DOI: 10.1002/cmdc.201600058.
84. Kalas, V.; Hibbing, M. E.; Maddirala, A. R.; Chugani, R.; Pinkner, J. S.; Mydock-McGrane, L. K.; Conover, M. S.; Janetka, J. W.; Hultgren, S. J. Structure-Based Discovery of Glycomimetic Fmlh Ligands as Inhibitors of Bacterial Adhesion During Urinary Tract Infection. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (12), E2819-E2828. DOI: 10.1073/pnas.1720140115.
85. Westerfors, M.; Tedebark, U.; Andersson, H. O.; Ohrman, S.; Choudhury, D.; Ersoy, O.; Shinohara, Y.; Axen, A.; Carredano, E.; Baumann, H. Structure-Based Discovery of a New Affinity Ligand to Pancreatic Alpha-Amylase. *J. Mol. Recognit.* **2003**, *16* (6), 396-405. DOI: 10.1002/jmr.626.
86. Ananthula, R. S.; Ravikumar, M.; Pramod, A. B.; Madala, K. K.; Mahmood, S. K. Strategies for Generating Less Toxic P-Selectin Inhibitors: Pharmacophore Modeling, Virtual Screening and Counter Pharmacophore Screening to Remove Toxic Hits. *J. Mol. Graphics Modell.* **2008**, *27* (4), 546-557. DOI: 10.1016/j.jmgm.2008.09.007.
87. Huang, K.; Wu, X. W.; Jiang, Z. Y.; Sun, H. P.; You, Q. D. Novel Dual-Site-Binding Neuraminidase Inhibitor from Virtual Screening by Pharmacophore and Molecular Dynamics Methods. *Chin. J. Chem.* **2012**, *30* (8), 1735-1740. DOI: 10.1002/cjoc.201200313.
88. Al-Asri, J.; Fazekas, E.; Lehoczki, G.; Perdih, A.; Gorick, C.; Melzig, M. F.; Gyemant, G.; Wolber, G.; Mortier, J. From Carbohydrates to Drug-Like Fragments: Rational Development of Novel Alpha-Amylase Inhibitors. *Bioorg. Med. Chem.* **2015**, *23* (20), 6725-6732. DOI: 10.1016/j.bmc.2015.09.007.
89. Al-Asri, J.; Gyemant, G.; Fazekas, E.; Lehoczki, G.; Melzig, M. F.; Wolber, G.; Mortier, J. A-Amylase Modulation: Discovery of Inhibitors Using a Multi-Pharmacophore Approach for Virtual Screening. *ChemMedChem* **2016**, *11* (21), 2372-2377. DOI: 10.1002/cmdc.201600427.
90. Irwin, J. J. Community Benchmarks for Virtual Screening. *J. Comput.-Aided Mol. Des.* **2008**, *22* (3), 193-199. DOI: 10.1007/s10822-008-9189-4.
91. Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T. M.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual Screening Using Protein-Ligand Docking: Avoiding Artificial Enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 793-806. DOI: 10.1021/ci034289q.
92. Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49* (23), 6789-6801. DOI: 10.1021/jm0608356.
93. Geurts, R.; Bisseling, T. Rhizobium Nod Factor Perception and Signalling. *Plant Cell* **2002**, *14* (Suppl), s239-s249. DOI: 10.1105/tpc.002451.
94. Allman, S. A.; Jensen, H. H.; Vijaykrishnan, B.; Garnett, J. A.; Leon, E.; Liu, Y.; Anthony, D. C.; Sibson, N. R.; Feizi, T.; Matthews, S.; Davis, B. G. Potent Fluoro-Oligosaccharide Probes of Adhesion in Toxoplasmosis. *ChemBioChem* **2009**, *10* (15), 2522-2529. DOI: 10.1002/cbic.200900425.
95. N'Go, I.; Golten, S.; Ardá, A.; Cañada, J.; Jiménez-Barbero, J.; Linclau, B.; Vincent, S. P. Tetrafluorination of Sugars as Strategy for Enhancing Protein-Carbohydrate Affinity: Application to Udp-Galp Mutase Inhibition. *Chem. - Eur. J.* **2014**, *20* (1), 106-112. DOI: 10.1002/chem.201303693.
96. Loris, R.; Tielker, D.; Jaeger, K. E.; Wyns, L. Structural Basis of Carbohydrate Recognition by the Lectin Lecb from Pseudomonas Aeruginosa. *J. Mol. Biol.* **2003**, *331* (4), 861-870. DOI: 10.1016/s0022-2836(03)00754-x.
97. Mitchell, E.; Houles, C.; Sudakevitz, D.; Wimmerova, M.; Gautier, C.; Perez, S.; Wu, A. M.; Gilboa-Garber, N.; Imberty, A. Structural Basis for Oligosaccharide-Mediated Adhesion of Pseudomonas Aeruginosa in the Lungs of Cystic Fibrosis Patients. *Nat. Struct. Biol.* **2002**, *9* (12), 918-921. DOI: 10.1038/nsb865.
98. Marotte, K.; Sabin, C.; Preville, C.; Moume-Pymbock, M.; Wimmerova, M.; Mitchell, E. P.; Imberty, A.; Roy, R. X-Ray Structures and Thermodynamics of the Interaction of Pa-Iil from Pseudomonas Aeruginosa with Disaccharide Derivatives. *ChemMedChem* **2007**, *2* (9), 1328-38. DOI: 10.1002/cmdc.200700100.





FOR TABLE OF CONTENTS USE ONLY

**Virtual screening against carbohydrate-binding proteins: evaluation and application to bacterial *Burkholderia ambifaria* lectin**

Tamir Dingjan<sup>1</sup>, Émilie Gillon<sup>2</sup>, Anne Imberty<sup>2</sup>, Serge Pérez<sup>3</sup>, Alexander Titz<sup>4,5,6</sup>, Paul A. Ramsland<sup>7,8,9,10\*</sup>, Elizabeth Yuriev<sup>1\*</sup>

