J Comput Aided Mol Des (2014) 28:1129–1142
DOI 10.1007/s10822-014-9789-0

Creating and virtually screening databases of fluorescently-labelled compounds for the discovery of target-specific molecular probes

Rhiannon L. Kamstra · Saedeh Dadgar · John Wigg · Morshed A. Chowdhury · Christopher P. Phenix · Wely B. Floriano

Received: 7 March 2014 / Accepted: 14 August 2014 / Published online: 24 August 2014

Abstract Our group has recently demonstrated that virtual screening is a useful technique for the identification of target-specific molecular probes. In this paper, we discuss some of our proof-of-concept results involving two biologically relevant target proteins, and report the development of a computational script to generate large databases of fluorescence-labelled compounds for computer-assisted molecular design. The virtual screening of a small library of 1,153 fluorescently-labelled compounds against two targets, and the experimental testing of selected hits reveal that this approach is efficient at identifying molecular probes, and that the screening of a labelled library is preferred over the screening of base compounds followed by conjugation of confirmed hits. The automated script for library generation explores the known reactivity of commercially available dyes, such as NHS-esters, to create large virtual databases of fluorescence-tagged small molecules that can be easily synthesized in a laboratory. A database of 14,862 compounds, each tagged with the ATTO680 fluorophore was generated with the automated script reported here. This library is available for downloading and it is suitable for virtual ligand screening aiming at the identification of target-specific fluorescent molecular probes.

Keywords Target-specific molecular probe · Fluorescent probe · Database for virtual screening · ATTO680 · E6 · *Botulinum* neurotoxin subtype A

Introduction

Many techniques in molecular biology laboratories demand the use of fluorescent reagents, which often function as optical tracers. These molecules may be intrinsically fluorescent or labelled with a fluorophore moiety. Reactive fluorophores are commonly used for this type of labeling, which is often performed on large biomolecules like antibodies and nucleic acids [1, 2]. The most commonly used reactive fluorophores are designed to covalently label molecules containing a primary amine functional group. *N*-hydroxysuccinimidyl ester and isothiocyanate derivatives are among the most well characterized reactive fluorophores that are commercially available [1]. Relatively early applications for reactive fluorophores include targeting serum components with labelled antibodies [3] and visualizing cell structures (e.g. actin filaments) using indirect immunofluorescent staining [4]. Fluorescently-labelled or conjugated antibodies continue to be used for various direct and indirect immunofluorescence techniques that have applications in fundamental research and diagnostic pathology [2, 5–8]. Although antibodies can be used as selective probes, small molecules are often more desirable for certain applications, including in vivo imaging, where the production of antibodies is not feasible, or where size, immunogenicity, or poor pharmacokinetic properties of antibodies becomes problematic [9]. With the introduction

Electronic supplementary material The online version of this article (doi:10.1007/s10822-014-9789-0) contains supplementary material, which is available to authorized users.

R. L. Kamstra · S. Dadgar · J. Wigg · W. B. Floriano (✉)
Department of Chemistry, Lakehead University, Thunder Bay, ON P7B 5E1, Canada
e-mail: wely.floriano@lakeheadu.ca

R. L. Kamstra · S. Dadgar · J. Wigg · M. A. Chowdhury · C. P. Phenix · W. B. Floriano
Thunder Bay Regional Research Institute, Thunder Bay, ON P7B 5E1, Canada

of fluorophores that are active in the near-infrared spectral region, in vivo fluorescence imaging has become a viable functional technique for visualizing specific molecular characteristics or events on a macroscopic scale [10, 11]. Fluorescent probes can be designed that target specific receptors [12], that respond to environmental changes (e.g. pH) [13], or that are activated by the activity of a specific enzyme [14, 15].

High-throughput virtual screening methods have been used to identify high-affinity ligands for a variety of protein targets [16–21]. Hierarchical virtual ligand screening (HierVLS) [22] is one protocol that uses a structure-based approach to screen large libraries of compounds against biologically relevant target structures, estimating a force-field binding energy for the complex. This method has been validated and used in previous investigations [16, 22, 23]. HierVLS, and other similar virtual screening methodologies have typically been used for identifying "hit" compounds intended for later development into therapeutic agents [22, 24, 25].

In this paper, we demonstrate that the creation and computational screening of fluorescence-labelled small molecule libraries is an efficient strategy for discovery of target-specific fluorescent molecular probes. We first demonstrate that the binding outcomes of virtually screening fluorescently-labelled compounds against selected protein targets agree with experimental data. Then we exploit the known reactivity of commercially available fluorophores, such as NHS-esters, to create large virtual databases of fluorescently-labelled small molecules that can be easily synthesized in a laboratory.

Our experimental test cases use a small compound library and its fluorescently labelled counterpart, which are screened against model structures of human papillomavirus variant 16 (HPV16) protein E6 and *botulinum* neurotoxin subtype A (BoNTA) [23]. Each of these protein targets have publically available experimental structures, are biological relevant and are the subject of ongoing investigation in our lab. E6 is a cancer-associated protein encoded by human papillomavirus (HPV) that causes ubiquitin-mediated degradation of a key tumor suppressor, p53, in infected cells [26–28]. There is evidence that E6, and certain E6 variants, play a role in conferring a high-risk phenotype to certain HPV strains [27, 28]. Specific detection agents for E6 variants could provide valuable insight into tumour cell pathology in vivo, and could also be developed to complement existing diagnostic tools. *Botulinum* neurotoxin (BoNT), produced by *Clostridium botulinum* bacteria, is one of the most potent known toxins, known for its ability to cause food-borne illness (botulism) [29] and for its potential to be used as a biological weapon [30, 31]. BoNT is also the key component of several available therapeutic and cosmetic treatments (e.g. "Botox" by Allergan, Inc.) [32, 33]. While some in vitro BoNT identification platforms exist, the current standard for confirmatory testing remains the "mouse assay", which is inefficient and takes several days to complete [31, 32]. Particularly in light of its risk both in food products and as a bioterrorism threat, rapid in vitro tests for BoNT are currently very in demand [30]. In an earlier investigation [16], a small molecule, paclitaxel was identified as a selective inhibitor of BoNT serotype A using virtual ligand screening (VLS), and then confirmed as a selected inhibitor of the target, giving confidence that BoNTA is a feasible target for future VLS investigations.

The experimental validation of our fluorescent-labelling and screening protocol provides the rationale and framework for developing a high-throughput method for generating virtual fluorophore-ligand conjugates, in order to facilitate the screening of large libraries of such compounds. We describe the development and implementation of a software-based script for parsing databases of small molecules containing appropriate reactive functional groups, and generating structures that are conjugated to user-specified reactive fluorophores. These databases of novel fluorescent conjugates can be virtually screened against protein biomarkers for the rapid identification of target-specific fluorescent molecular probes.

Materials and methods

Preparation of the test libraries

A library of 1,153 small molecules, each containing a single primary amine functional group, was used as a test virtual screening library. Compounds for this primary amines database (PAD) were downloaded from PubChem [34] by drawing a basic primary amine structure under the search options "Structure Search", "Substructure". The structure drawn corresponds to the SMILE string C(*)N([H])[H]. The search was performed using default parameters and Lipinski's "rule of 5" as filter for drug-likeness [35] (molecular weight less than 500 Da, no more than five hydrogen bond donors, no more than 10 hydrogen bond acceptors, octanol–water partition coefficient no greater than 5). The first 2,000 compounds were downloaded in SDF format. Solvent and co-ion molecules were removed during processing, and entries containing atoms not parameterized by common force fields were excluded. The database was built and edited using MOE software [36]. The Gasteiger method [37] was used to assign partial charges and an energy minimization was conducted using MMFF94x force field [38]. The database was then saved in a Tripos MOL2 format, with a total of 1,153 entries.

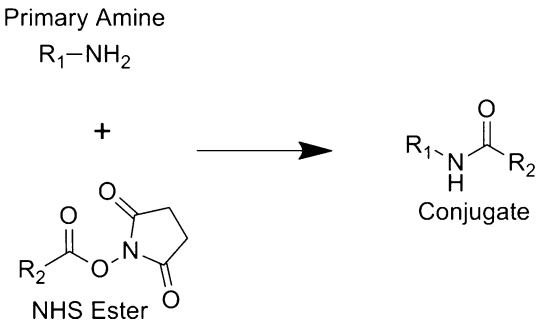

Fig. 1 *N*-hydroxysuccinimidyl ester (NHS ester) reactive fluorophores react with primary amines to form fluorophore/ligand conjugates that are linked via an amide bond

In addition, the PAD was also tagged computationally with a Bodipy fluorophore (3-Bodipy-propanoylaminocaproic acid, *N*-hydroxysuccinimide ester) by manually editing the structures [23]. This library is referred to as FL-PAD. Conjugate structures were created according to the general reactivity scheme between primary amines and *N*-hydroxysuccinimidyl ester fluorophore derivatives that is described both in the literature [39, 40] and in the Molecular Probes Handbook [1]. This scheme is shown in Fig. 1.

Protein structure preparation, VLS, and hit selection

Protein structure preparation

The experimental structure for the C-terminal domain of HPV16 protein E6 (PDB Code: 2FK4) was obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB). The structure was checked for missing atoms, and energy-minimized with MOE [36] using the MMFF94x force field [38]. A model protein structure for BoNTA was prepared previously based on the experimentally determined structure with PDB code 2NYY [16].

Virtual ligand screening

BoNTA and E6 structures were scanned for docking regions using the PASS method [41]. Both structures were virtually screened against the PAD with molecular docking and scoring performed according to the HierVLS approach [22]. Briefly, for each ligand in the database docking and scoring are carried out through three stages of increasing computational demand into each binding site within the target protein. The first stage is a coarse grain generation of at least 10,000 docked conformations of a flexible ligand into a single, fixed conformation of the protein. Incremental torsion angles are used to account for ligand flexibility during docking. A buried surface filter eliminates any docked structure from stage 1 with less than a user-specified percent of the ligands molecular surface buried into the protein's binding site (70 % for small compound libraries and 30 % for large compounds). This filter eliminates complexes that score well is a gas-phase environment, but display very little contact between ligand and protein. The top 10 % by force field energy of all the docked conformers passing the filter are subjected to the 2nd stage, which consists of energy-minimization with fixed protein to optimize the ligands conformation within the binding site. The best five conformers by force field energy from the protein fixed minimization stage are carried to the 3rd stage which is an all-atoms energy minimization that optimizes both protein and ligand structures. The best docked conformer by energy from the all-atoms minimization is used to calculate binding energies taking into account the solvation energies associated to the unbound protein, unbound ligand, and bound complex using an implicit solvation model [22].

Docking and scoring were performed independently for each binding region identified by PASS (33 regions in BoNTA (2NYY) and one region in the C-terminal domain of E6 (2FK4)). Good binding candidates were selected based on whether they had a calculated binding affinity of one standard deviation better than (below) the mean binding score for the entire library screened, at each region of interest. The Bodipy-tagged PAD was docked against BoNTA and E6 using the same approach, with an adjusted buried surface area cutoff of 30 % (from 70 %), to accommodate the increased average size of the molecules being docked. The scores of each selected binding candidate from the unlabelled ligand set were compared with those of their labelled counterparts. For E6, the analysis focused on the single binding region identified. Docking and scoring of BoNTA against PAD and FL-PAD was carried out in a previous investigation, with detailed methodology and results described by Dadgar et al. [16, 23]. Here we show the PAD force field scores for the best binding region for a compound, aspartame (APM), selected for conjugation and experimental testing, and the corresponding scores for the FL-PAD.

Synthesis and purity of fluorescent conjugates

3-Bodipy-propanoylaminocaproic acid, *N*-hydroxysuccinimide ester (BODIPY® FL-X, hereafter abbreviated Bodipy-NHS; CAS 217190-09-5; Cat. B2527-40; Fig. 2) was purchased from US Biological and LKT Laboratories Inc.. *O*-Succinyl-L-homoserine (Fig. 3a; CAS 1492-23-5) was purchased from

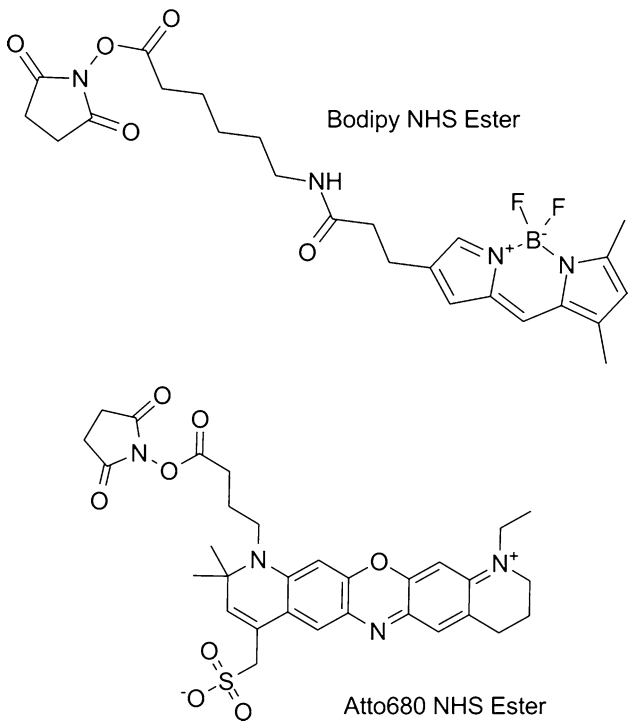

Fig. 2 Bodipy NHS ester was the reactive fluorophore used for the experimental synthesis and testing of fluorescent small-molecule conjugates. ATTO680 NHS ester is the reactive dye chosen for computational generation of tagged compound libraries

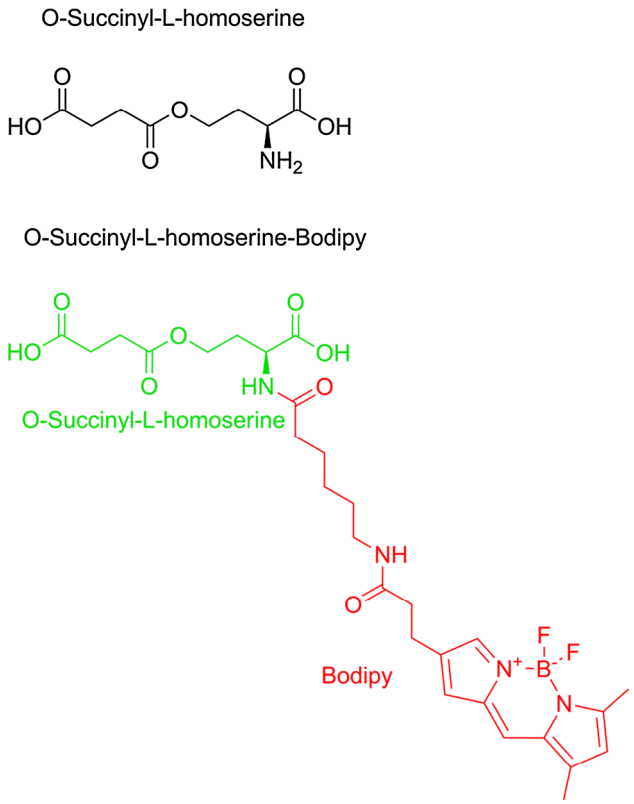

Fig. 3 *O*-succinyl-L-homoserine (*top*), a putative ligand for protein E6, is pictured along with the structure of *O*-succinyl-L-homoserine-Bodipy (*bottom*), the fluorescent conjugate of *O*-succinyl-L-homoserine and 3-Bodipy-propanoylaminocaproic acid, *N*-hydroxy-succinimide ester (Bodipy)

Sigma Aldrich (Cat. S7129-25MG; 98 % purity TLC). *O*-Succinyl-L-homoserine-Bodipy (Fig. 3) was synthesized by dissolving Bodipy-NHS (5 mg), in 500 µl of dimethylformamide (DMF). The reaction vial was placed, with stirring, in a water/ice bath at 0 °C followed by addition of *O*-succinyl-homoserine (2.7 mg) and triethylamine (1.40 µl). The reaction was flushed with argon and left to proceed for 30 min at 0 °C after which was allowed to stand at room temperature overnight. The reaction was quenched by addition of distilled water followed by extraction with ethyl acetate. The aqueous phase was frozen and sublimed using a freeze drier to yield the desired product (3 mg). The purity of the product was confirmed using TLC (5:1 dichloromethane:methanol) and the identity confirmed with liquid chromatography mass spectrometry (LC–MS) and nuclear magnetic resonance (NMR).

Detailed information on the synthesis and the purification of APM's Bodipy conjugate (APMBDP) is presented elsewhere [23]. Identity and purity of the isolated conjugate were confirmed using mass spectrometry and NMR respectively [23].

BoNTA fluorescence polarization assay

APMBDP was tested in a fluorescence polarization assay competing with unlabelled APM against BoNTA light chain (LC), as reported elsewhere [23].

Protein E6 intrinsic tryptophan fluorescence assay

Protein E6 contains only one tryptophan which is partially solvent-accessible in the experimental structure of the unbound protein, making it suitable for assays based on intrinsic tryptophan (Trp) fluorescence. Solvent-exposed Trp residues have a maximum fluorescence emission at 340–350 nm, whereas buried Trp typically exhibit maximum emission at 330 nm. Binding to a ligand may alter the environment of the Trp residues present in a protein, affecting its intrinsic Trp fluorescence spectrum. Preliminary experiments with an E6-specific antibody (6F4) indicated that binding to E6 significantly quenched the emission at 350 nm associated to E6's single Trp. Unlabeled *O*-succinyl-L-homoserine was tested at 8 concentrations (0.01, 0.1, 1, 3, 10, 25, 100, and 300 µM), diluted from a 1 mM stock solution dissolved in deionized water. The concentration of E6 in the assays was fixed at 0.03 mg/ml, diluted from a stock concentration of 0.201 mg/ml. The assay was buffered to a pH of 7.5 using 20 mM HEPES, 0.01 % tween (v/v). Control wells were included for the unquenched protein using E6 (0.03 mg/ml)

in buffer, and for each concentration of *O*-succinyl-L-homoserine. Samples were plated in triplicate and controls in duplicate in a 96-well black bottom microplate. The E6-specific antibody 6F4 was used as a positive control at a concentration of 0.4 ng/ml. E6 was added last to sample or control wells and the plate was incubated in the plate reader until the temperature reached 37 °C. Emission was read 310–700 nm with an excitation of 288 nm, probe height of 4 mm, 150 μs delay between wells, and 20 reads per well, every 30 min for 60 min using a sensitivity value of 100 with Gen5 software (Biotek). The baseline fluorescence of each concentration of *O*-succinyl-L-homoserine was discounted from the samples containing protein prior to calculating percent change relative to unquenched protein at 350 nm.

Protein E6 fluorescence polarization assay

O-Succinyl-L-homoserine-Bodipy was tested in a fluorescence polarization assay at three concentrations (1, 10, and 100 μM) with a fixed concentration of protein E6 (0.04 mg/ml). The assay was buffered to a pH of 7.5 using 20 mM HEPES, 0.01 % tween (v/v) and incubated to a temperature of 37 °C. Control wells were included for the protein without a ligand, and *O*-succinyl-L-homoserine-Bodipy in the absence of protein at each concentration tested. Samples were plated as single wells in a 96-well black bottom microplate. The experiment was run using a Synergy 4 Microplate Reader (Biotek, Inc.). An excitation filter (485 nm/20) and emission filter (528 nm/20) were used with a dichromatic mirror (510 nm). Fluorescence polarization was read using a probe height of 4 mm, a 350 μs delay between wells, and 40 reads per well. Data was collected every 15 min for 75 min using a sensitivity value of 65 using Gen5 software (Biotek, Inc.). Fluorescence polarization values were calculated with blank values discounted. Data analyses was performed using Prism [42].

Fluorophore conjugation script and preparation of a labelled library

We designed a computational script able to virtually conjugate reactive fluorophores with ligands possessing the correct functional group, according to a pre-defined reactivity scheme. Literature review was used to identify common reactivity schemes for commercially available reactive fluorophores. MOE's scientific vector language (SVL) [36] was chosen as the scripting environment due to the large number of available and relevant built-in functions and its suitability for large databases of small organic molecules. A SMARTS-based (http://www.daylight.com/) functional group identification algorithm, included as part of the MOE package, was used to recognize the functional groups involved in the fluorophore-ligand reaction. We designed this first version of the script to identify *N*-hydroxysuccinimidyl ester derivatives, and to label small molecules containing one primary amine functional group. The small molecule becomes labelled with an amide bond according to the scheme shown in Fig. 1 [1]. The script accepts user-specified MOE molecular database (MDB) files as input for both fluorophores and ligands. Molecules (either fluorophore or ligand) that do not contain the appropriate functional groups required for the fluorescent labelling reaction are skipped. Conjugate structures are outputted as entries to a new MOE database and assigned molecule names. The script (provided as supplementary material) was tested using small databases of fluorophores and ligands to ensure that it correctly matched the specified molecular patterns targeted by the script, and skipped molecules containing "decoy" functional groups. Using this script, the same base compound library containing n primary amines can be conjugated with multiple fluorophores (m) to generate a library of $n \times m$ conjugates.

Base compound library for conjugation

The molecular libraries small molecule repository (MLSMR) (https://mli.nih.gov/mli/compound-repository/mlsmr-compounds/) was selected as a source of molecular structures. There are three subsets each containing approximately 100,000 compounds, as well as some additional structures, totalling over 400,000 compounds as of June 2013 [43]. The Molecular Libraries Initiative (MLI) describes the selection criteria for these compounds on their website (https://mli.nih.gov/mli/compound-repository/mlsmr-compounds/). PubChem offers structures for download in both 2D and 3D formats. PubChem generates 3D structures for all compounds matching a specific set of criteria (detailed on PubChem3D) [44]. One 3D conformer (if available) was downloaded in SDF format for each MLSMR compound from PubChem3D (http://pubchem.ncbi.nlm.nih.gov/), totalling 383,987 structures. The database was refined using MOE software (www.chemcomp.com/). SDF structures were loaded into MOE molecular databases and checked for counter ions and solvent molecules. Gasteiger atomic charges [37] were calculated for each molecule. An energy minimization was performed on each database using the MMFF94x force field [45]. Chirality was preserved for the downloaded structures.

Generating a library of fluorescent conjugates

The fluorescent labelling script was used to parse the database of MLSMR compounds from PubChem3D and conjugate all compounds containing one primary amine to one NHS-ester functionalized fluorophore. ATTO680

NHS-ester (PubChemCID 16218508) was used as the amine-reactive fluorophore in this case (Fig. 2). MOE software (www.chemcomp.com) was used for database preparation and refinement. Conjugates were outputted to a molecular database (MDB format) before being assigned Gasteiger atomic charges [37] and energy minimized using MMFF94x force-field [45]. Conjugates were subjected to a conformational search using a built-in MOE function ("Conformational Import"), which is a high-throughput method for searching each molecules conformational space. Only one conformer was selected for each compound. Stochastic search failure and iteration limits of, respectively, 20 and 250 were used. Superposed conformations with an RMSD value of 0.15 or less were considered identical. Default constraints were used.

Results

Virtual screening and hit selection

A library of 1,153 primary amine-containing compounds (PAD) was screened using the HierVLS approach against two biologically relevant protein targets, BoNTA and HPV16 protein E6. The force field-based binding scores from this screen are presented in Fig. 4a, c. The structures in this library were subsequently tagged with Bodipy fluorophore (3-Bodipy-propanoylaminocaproic acid, *N*-hydroxysuccinimide ester) by manually editing the structures to form an amide bond between primary amine nitrogen and the carbonyl carbon of the *N*-hydroxysuccinimidyl ester reactive group from the fluorophore (Fig. 1). The Bodipy-labelled PAD (FL-PAD) was virtually screened against both targets and the resulting binding scores are presented in Fig. 4b, d. The Bodipy fluorophore was chosen for virtual labelling because it is commercially available and suited for use with the available optical instruments in our laboratory. APM and APMBDP were both predicted to bind to BoNTA [23]. Both APM and APMBDP score below the one standard deviation threshold (Fig. 4a, b), however the binding score for APMBDP (−74.90 kcal/mol) is very close to the threshold for the FL-PAD library (−76.84 kcal/mol). APM was selected because it is a well-know, safe to handle, readily available and inexpensive compound and, most importantly, the position of the primary amine in its structure was expected to simplify the fluorescent labelling reaction. A similar analysis of the force field-based binding scores for the non-conjugate and conjugated libraries screened against E6 (Fig. 4c, d) indicates that *O*-succinyl-L-homoserine is expected to bind to E6 while its labelled counterpart, *O*-succinyl-L-homoserine-Bodipy (Fig. 3), is not. The decision threshold of one standard deviation below the mean score for the entire library screened is derived under the assumption that most randomly chosen compounds in a screening library are not expected to bind to a target protein with high affinity, and justified by a fairly normal distribution of the scores (Fig. 5, top). Therefore, any ligand within a 95 % confidence interval from the mean should be expected not to bind. Scores much better (below) than the mean are indicative of meaningful interactions between ligand thus suggesting binding. These test cases were subject to experimental validation to confirm whether predicted binding or non-binding would be replicated in in vitro assays.

A common bias in virtual screening results is the favouring of higher molecular weight compounds [46–49]. In HierVLS the molecular weight bias is expected to be significantly reduced by the buried surface area filter applied in stage 1. This filter eliminates docked conformers of a given ligand that are not deeply buried within a binding site. As a consequence, docked conformers of large compounds that make extensive contacts with the surface of the protein but do not have a significant portion of their structures buried within a cavity are eliminated even if they score well. Thus, the conformer of a large compound that is selected for binding score calculation is not necessarily the one with the lowest (most favourable) energy, but the one with the most "meaningful" contacts with the protein. This effectively reduces the bias towards large compounds scoring higher just because of their size. To investigate the effect of molecular weight on our scores, in Fig. 5 we plotted the force field-based binding scores obtained from virtual screening against the molecular weight of all compounds surviving the 3 stages of HierVLS.

As observed in Fig. 5, the binding scores of PAD compounds docked to E6 show a bias towards larger molecular weights, whereas PAD docked to BoNTA has a more even distribution of molecular weights. The bias towards heavier compounds is not observed among the compounds from the high-molecular weight library FL-PAD docked to E6 and, to a lesser extent, FL-PAD docked to BoNTA. These observations are supported by a Spearman's two-tailed correlation analysis of molecular weight versus binding scores. On average, the binding scores for PAD docked to E6 are clearly worse than FL-PAD, as indicated in Fig. 5 and Table 1 (binding thresholds are, respectively, −33.19 and −66.60 kcal/mol). In contrast, score distribution and the binding thresholds for PAD and the much larger FL-PAD against BoNTA are virtually the same (Fig. 5, Table 1). The preference for large ligands observed for E6 is likely related to the nature of the target rather than to a general bias of the methodology towards high-molecular weight compounds. E6 is a small protein known to bind to multiple proteins and, therefore, expected to have binding sites that are appropriate for contact with larger molecules.

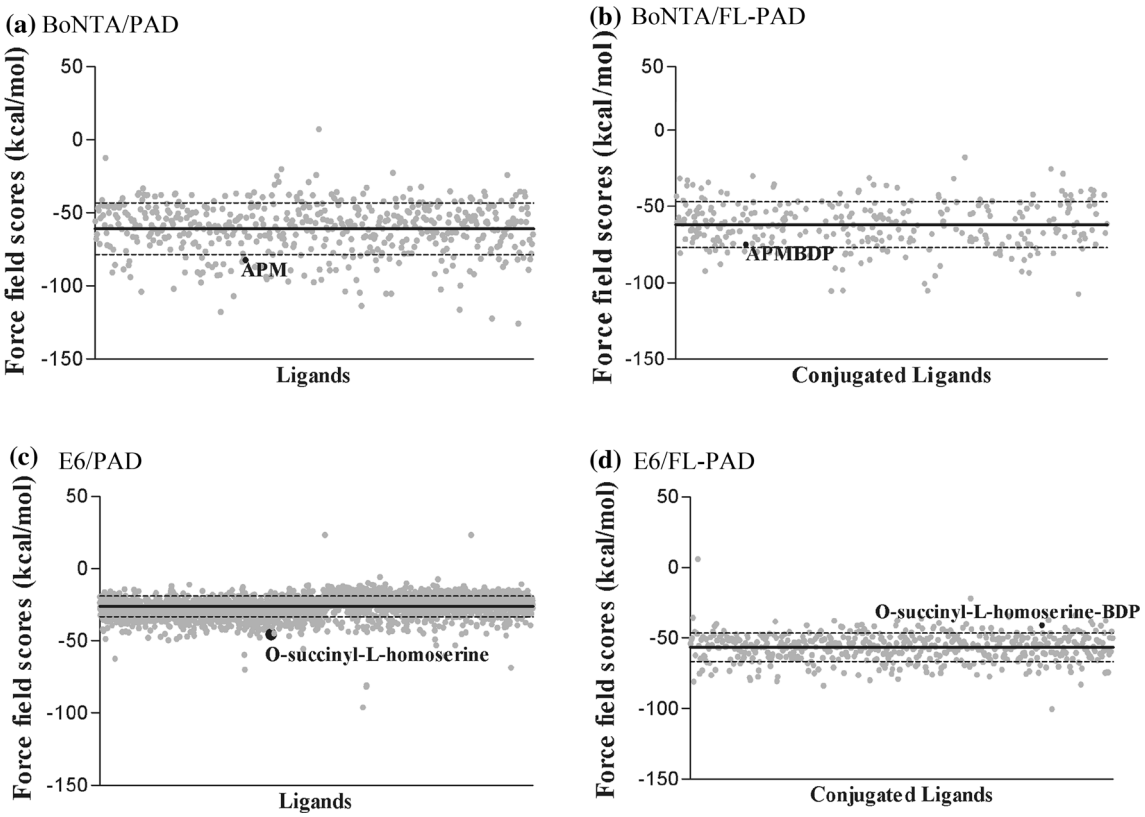

Fig. 4 Force field-based binding scores for non-conjugated (PAD) and conjugated (FL-PAD) library docked to **a** BoNTA and **b** to E6. The *solid and dotted lines* represent mean and 1 standard deviation (1D) above or below the mean, respectively. Ligands at or below the 1D threshold are predicted to bind to their targets. Both APM and APMBDP score well, with a predicted binding score that is more than 1D better than the average, or is just above 1D. *O*-succinyl-L-homoserine scored below the 1D threshold, which indicates high binding affinity for protein E6. In contrast, *O*-succinyl-L-homoserine-Bodipy scored relatively poorly, with a predicted binding affinity almost 2 standard deviations worse than the average docked compound

In vitro testing of BoNTA ligands

The synthesized APMBDP was tested in a fluorescence polarization binding assay to determine its binding affinity experimentally for the target, BoNTA LC [23]. Fluorescence polarization assays rely on the principle that, when excited with polarized light, a static fluorophore will emit light that is polarized along the same plane. Depolarization of the light occurs in proportion to the amount of free rotation of the molecule in solution. Therefore, high polarization values are normally indicative of larger complexes that rotate less rapidly in solution, while low polarization indicates a relatively unrestrained fluorophore. Fluorescence polarization has been examined as a powerful and versatile method that can be used without complex separation or washing, does not require radiolabelled tracers and can be scaled for high-throughput applications [50–52]. Experimentally, both APM and APMBDP show affinity for BoNTA LC [23]. Moreover, unlabelled and labelled APM are competing for the same binding region on the BoNTA LC, as evidenced by the decrease in polarization observed in samples with both ligands shown in Fig. 6. These results are consistent with the virtual screening scores which show both labelled and unlabelled APM above the threshold for predicting binding to BoNTA LC [23]. Unlabelled and labelled APM/BoNTA complexes have very comparable energy components (Table 2), with APM having stronger hydrogen bond interactions but weaker overall Coulombic interactions in comparison to APMBDP.

In vitro testing of E6 ligands

Unlabelled *O*-succinyl-L-homoserine was tested 8 concentrations (0.01, 0.1, 1, 3, 10, 25, 100, and 300 μM) for binding to E6 (0.03 mg/ml) using an intrinsic tryptophan fluorescence assay. These results are shown in Fig. 7a. Unlabelled *O*-succinyl-L-homoserine exhibits a distinctive increase in tryptophan fluorescence relative to baseline control. An increase in signal is consistent with binding, as indicated by the positive (E6 antibody) control. However, no significant dose-dependency is observed at the

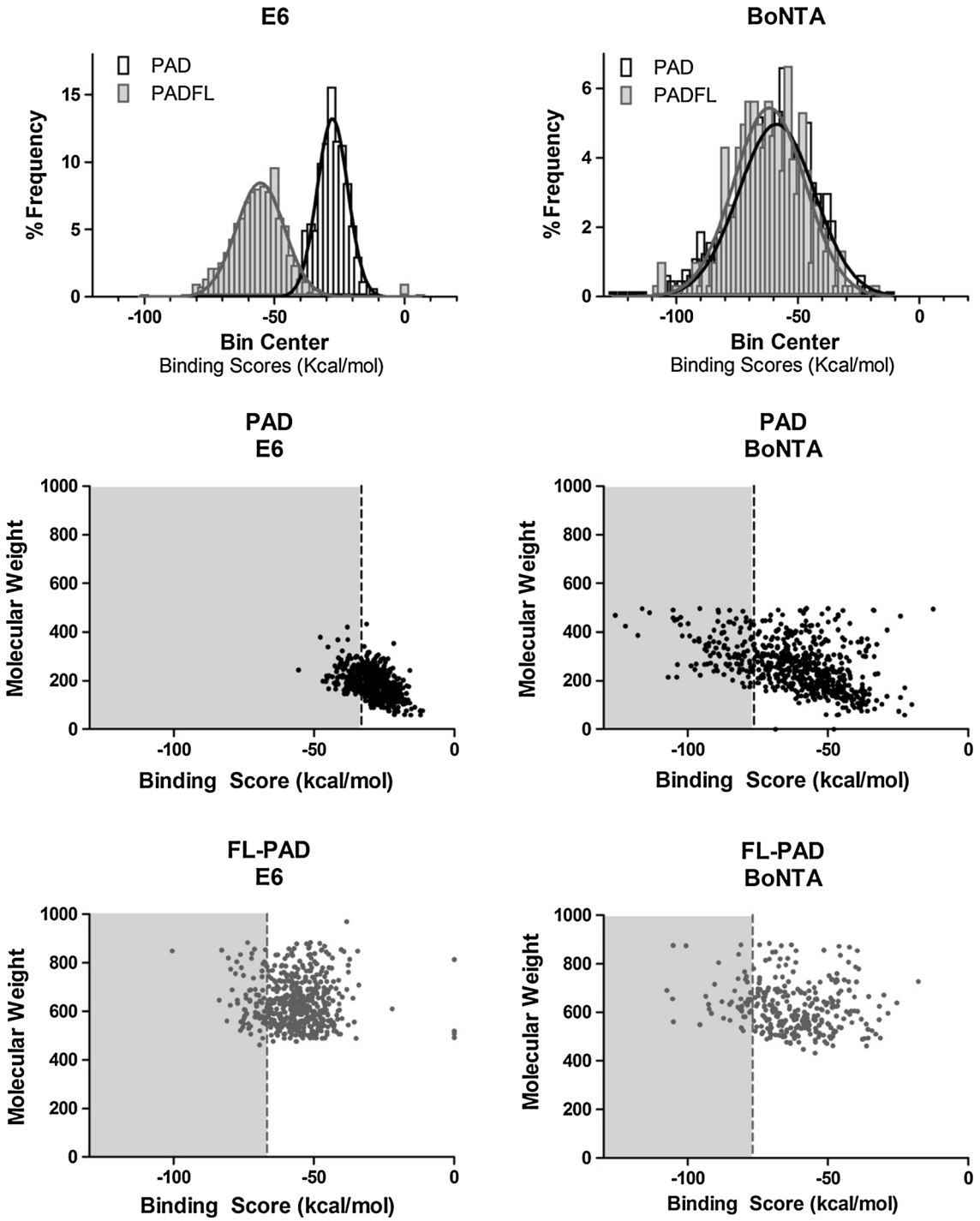

Fig. 5 Binding scores frequency distribution and molecular weight (MW) dependency for the non-conjugated (PAD; *black*) and conjugated (FL-PAD; *gray*) libraries docked to BoNTA or to E6. The *dotted lines* marks the binding threshold used for hit selection, with compounds to the *left* (*shaded area*) being considered potential binders. A total of 126 (PAD docked to E6), 75 (FL-PAD docked to E6), 118 (PAD docked to BoNTA), and 44 (FL-PAD docked to BoNTA) compounds passed the binding thresholds

concentrations tested. *O*-succinyl-L-homoserine-Bodipy was synthesized and tested for binding against protein E6 using a fluorescence polarization-based assay. The ligand was tested in three concentrations (1, 10,100 μM) against a fixed concentration of protein. The results are shown in Fig. 7b. Polarization values after 1 h do not differ substantially from control. This binding assay yielded no evidence that *O*-succinyl-L-homoserine-Bodipy binds to protein E6, predicted by the low binding score of the docked compound relative to the threshold. These results

Table 1 Predicted binding scores, expected experimental result, and whether experimental binding was observed for four selected compounds

Compound	Binding score (kcal/mol)	Binding threshold (kcal/mol)	Predicted binding (yes/no)	Experimental binding (yes/no)
APM	−82.20	−76.38	Yes	Yes
APMBDP	−74.90	−76.84	Yes	Yes
O-succinyl-L-homoserine	−45.62	−33.19	Yes	Yes
O-succinyl-L-homoserine-Bodipy	−40.85	−66.60	No	No

Binding scores are calculated as the difference between the total energy, including solvation, of the complex and unbound ligand and protein

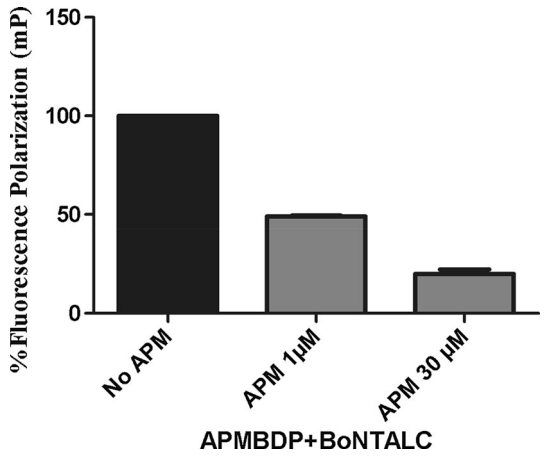

Fig. 6 Fluorescence polarization binding assay using APMBDP against BoNTA LC. Labelled APM is shown as *black bar* in the presence of BoNTA LC. Unlabelled APM at two concentrations (1 and 30 μM) correspond to *gray bars*. The addition of unlabelled APM at 1 and 30 μM reduces the fluorescence polarization associated with the binding of APMBDP to BoNTA, indicating that APM competes with labelled APM for binding to BoNTA. A detailed set of experiments is reported in a separate manuscript [23]

are consistent with the virtual screening scores (Fig. 4b, d) which suggest that unlabelled *O*-succinyl-L-homoserine binds to E6, but *O*-succinyl-L-homoserine-Bodipy does not. For *O*-succinyl-L-homoserine docked to E6, the total energy of the conjugate complex is much worse than of the unlabeled complex (Table 2). The breakdown of the total energy for unlabelled and labelled *O*-succinyl-L-homoserine/E6 complexes indicates that conjugated *O*-succinyl-L-homoserine has less favourable hydrogen bonds and overall Coloumbic interactions when bound to E6 than unlabelled *O*-succinyl-L-homoserine. An analysis of the binding modes obtained from docking (Fig. 8) suggests that *O*-succinyl-L-homoserine moiety of the fluorescent conjugate sits within the same pocket as unlabelled *O*-succinyl-L-homoserine, however the end containing the primary amine group is turned 180̊ to accommodate the bodipy fluorophore away from the cavity. Favourable interactions with Glu121, Arg124, and Arg142 present in the docked unlabelled *O*-succinyl-L-homoserine are lost in the conjugate, partially explaining the lack of binding affinity of the conjugate compared to unlabelled *O*-succinyl-L-homoserine.

Efficient generation of large libraries of fluorescence-labelled compound enables the use of VLS for target-specific molecular probe discovery

A script for conjugating reactive fluorophores to databases of small molecules (supplementary material) was prepared using MOE's SVL. This script was able to distinguish between true candidates for conjugation and decoys, and was used to virtually label a publically available database of base compounds with an ATTO680 (ATTO-TEC, Siegen, Germany; PubChemCID 16218508) according to the reactivity scheme shown in Fig. 1. The resultant database contains 14,862 fluorescently-labelled compounds (supplementary material) that would be suitable for VLS applications aiming at identifying targeted optical probes. The averages of molecular weight, number of hydrogen bond donors and acceptors, and number of rotatable bonds for the labelled library are compared in Fig. 9 to the MLSMR library, which includes the parent primary amine compounds that were conjugated to the ATTO680 dye. The average molecular weight of the ATTO680-labelled library is significantly higher (2.4 times) than the small-compounds MLSMR, as expected. The labelled compounds are also more flexible than the compounds in the MLSMR library, as indicated by the number of rotatable bonds. Both molecular descriptors, molecular weight and number of rotatable bonds, are known to increase the computational time required for molecular docking. The average number of hydrogen donors and acceptors, which are critical for binding to many active sites, are not significantly different between the two libraries.

The script can be easily modified to work with a different amine-reactive fluorophore for labelling. It can also be expanded to include other reactive fluorophores, such as isothiocyanates and maleimides, thus expanding the scope of possible conjugation reactions.

Discussion and conclusions

Significantly more extensive validation procedures have been previously described, given us confidence that Hier-VLS is a valid approach for screening small molecules for

Table 2 Total potential energy and its components for unlabelled and fluorescently-labelled complexes of E6 and BoNTA

Energy (kcal/mol)	*O*-succinyl-L-homoserine/E6 complex	*O*-succinyl-L-homoserine-bodipy/E6 complex	APM/BoNTA complex	APMBDP/BoNTA complex
Total energy	−2,675.85	−1,697.72	−41,477.90	−41,520.00
Bonded	683.37	65.63	9,890.69	9,965.15
Non-bonded				
Coulomb	−1,687.11	−173.34	−30,128.41	−30,242.75
Van der Waals	464.63	27.10	7,101.97	7,118.88
Hydrogen bonds	−574.19	−3.02	−11,599.19	−11,575.50
Solvation	−1,562.59	−1,614.09	−16,742.9	−16,785.73

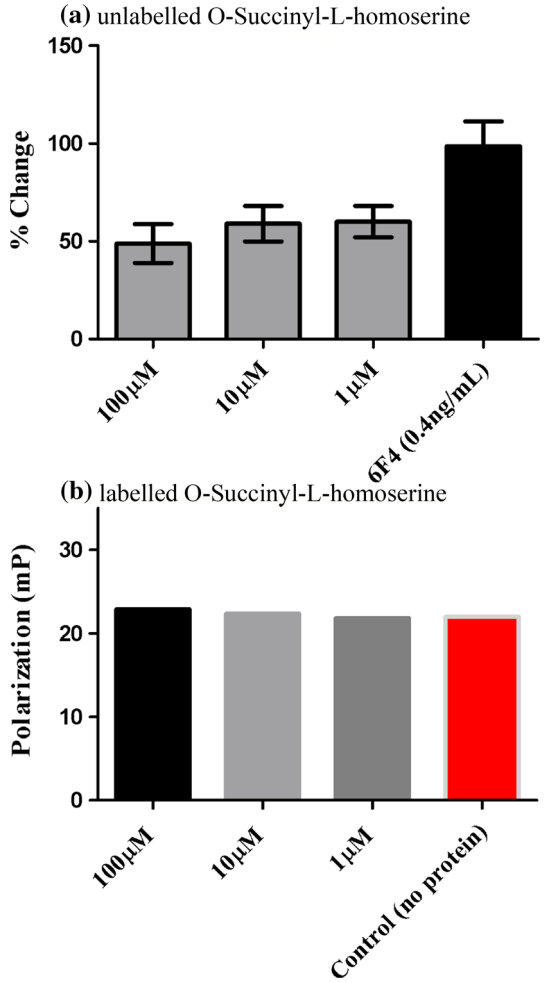

Fig. 7 **a** Percent change in intrinsic tryptophan fluorescence at 350 nm induced by unlabelled *O*-succinyl-L-homoserine at three concentrations (1, 10,100 μM) suggests that this compound does bind to E6 (0.03 mg/ml), in agreement with the computational prediction. **b** Fluorescence polarization values associated to three concentrations (1, 10,100 μM) of *O*-succinyl-L-homoserine-Bodipy was unchanged in the presence of protein E6 (0.04 mg/ml). This indicates that *O*-succinyl-L-homoserine-Bodipy does not bind to E6, in agreement with the computational prediction. An E6-specific antibody (6F4) was used as positive control

binding affinity [20, 24–27]. However, this approach has not been used specifically for estimating the affinity of larger, fluorophores and their small molecule conjugates. Our experimental results are in agreement with our computational predictions, which suggest that both unlabelled and labelled APM have significant binding affinity for BoNTA LC and that *O*-succinyl-L-homoserine and its Bodipy conjugate will and will not bind to E6, respectively (a summary of results is presented in Table 1). This demonstrates that a VLS approach, such as HierVLS, can be an appropriate tool for estimating the binding of conjugates of reactive fluorophores and appropriate small molecules to protein targets. Moreover, our results indicate that virtually screening the much larger conjugates is necessary to identify non-binders before costly synthesis and experimental testing, even though it incurs a much higher computational cost compared to screening a non-conjugated library. We suggest as strategy to reduce computational costs a two-step VLS: (1) virtually screen a non-conjugated library; (2) virtually screen the fluorescent conjugates of the top 10 % non-conjugated hits. The suggested 10 % mark is based on enrichment factors reported for various VLS schemes [53]. Although we have only docked fluorescent-labelled compounds using HierVLS, it is reasonable to assume that other docking methodologies are also capable of handling these large molecules, albeit at a different computational costs.

As these screening approaches are designed to be high-throughput methods, allowing for thousands of molecules to be screened against targets of interest, it is necessary to devise a method for generating fluorophore/ligand conjugates rapidly, from large databases of substructures. To realize this idea, we designed a script that recognizes the appropriate reactive groups associated with conjugating an NHS-ester fluorophore to ligands containing primary amines. The script was written using a vector-based programming language known as SVL used by MOE (www.chemcomp.com) software. This specific language and the

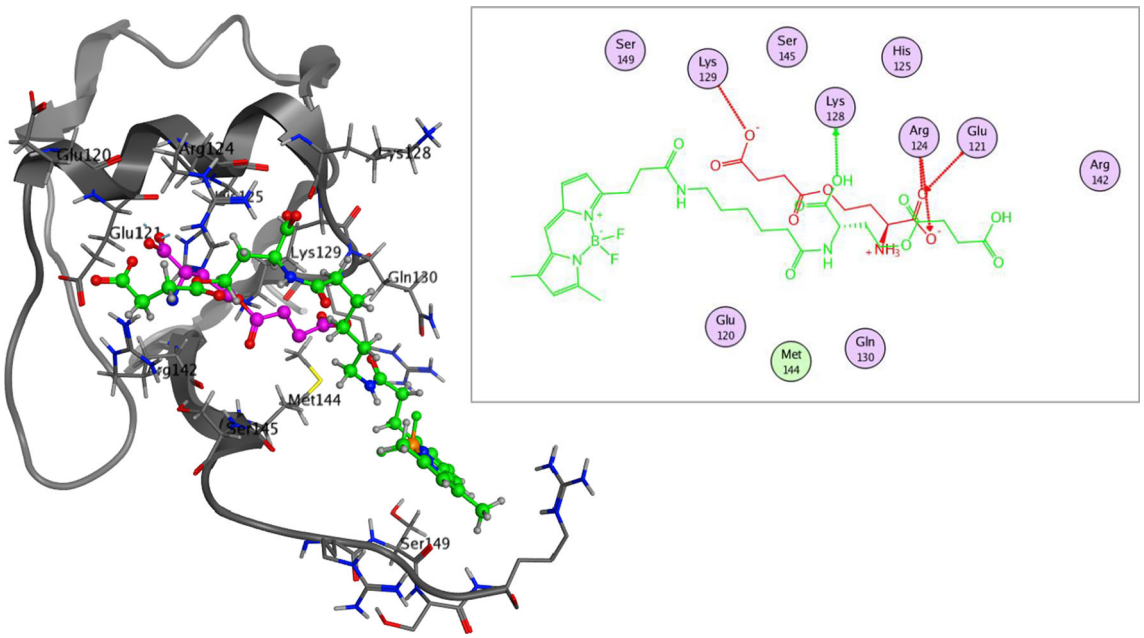

Fig. 8 Overlay of binding modes for *O*-succinyl-L-homoserine (*pink*/*red* carbons) and *O*-succinyl-L-homoserine-Bodipy (*green* carbons) docked to E6, in a 3D (*right*) and simplified 2D (*left*) representation. The much larger fluorescent conjugate has diminished ionic and hydrogen bond interactions relative to unlabelled *O*-succinyl-L-homoserine, the orientation of *O*-succinyl-L-homoserine within the binding site flips *left*-to-*right* in the conjugate, to accommodate the added fluorophore which is attached to the primary amine group of *O*-succinyl-L-homoserine

additional pre-defined applications and functions found in MOE are particularly suited for high-throughput applications involving large molecular databases. The script is capable of handling a fluorophore database containing multiple entries, and will create conjugates for each fluorophore containing an NHS-ester functional group. The script will only create conjugates for NHS-ester/primary amine combinations, and will output all conjugates with a name based on the concatenated ligand and fluorophore names or IDs. In cases where either the fluorophore or ligand is missing the appropriate reactive group, the conjugation is skipped, and a message is written to a log file. A SMARTS-based connectivity algorithm, built into MOE, is used for recognizing the appropriate functional groups. A database of test compounds containing various functional groups was prepared and tested to ensure that functional groups other than primary amines were not being inadvertently recognized by the algorithm used. Particular attention was given to ensure that other nitrogen-containing functional groups, including nitriles, amides, nitrates, and secondary or tertiary amines were being correctly discriminated from primary amines. Compounds containing multiple primary amines were completely avoided by the script, with a message written to a log file to alert the user. While these compounds would likely react with NHS-esters, it would be necessary to determine which amine would be preferentially labelled. Moreover, experimental synthesis of these conjugates would likely require additional purification or separation steps to ensure that the appropriate conjugate was being selected and used for testing. These practical considerations led to the decision to avoid compounds containing multiple reactive groups. This version of the script is designed to be a basic implementation of our idea, which can be expanded in later versions to include other types of reactive fluorophores, and exclude entries containing undesirable molecular features such groups which could confer toxicity or that commonly participate in side reactions with reactive fluorophores.

Once tested, the script was deemed adequate for the labelling of large databases of small molecules with reactive NHS-ester fluorophores. We generated a large labelled database using publically available small molecule structures and one NHS-ester fluorophore, for later use in virtual screening applications. [1-(3-Carboxypropyl)-11-ethyl-2,2-dimethyl-1,8,9,10-tetrahydro-2H-dipyrido[3,2-b:2′,3′-i]phenoxazin-11-ium-4-yl]methanesulfonate, marketed as ATTO680 (ATTO-TEC, Siegen, Germany) is the fluorophore moiety that was chosen for this database. ATTO680 is a relatively hydrophilic commercial fluorophore that fluoresces in the red region [54], which is more appropriate for cell-based in vitro applications, and possibly in vivo applications. In comparisons with other popular fluorophore series including AlexaFluor® (Molecular Probes Inc.) and BODIPY® (Molecular Probes Inc.), ATTO680 exhibits among

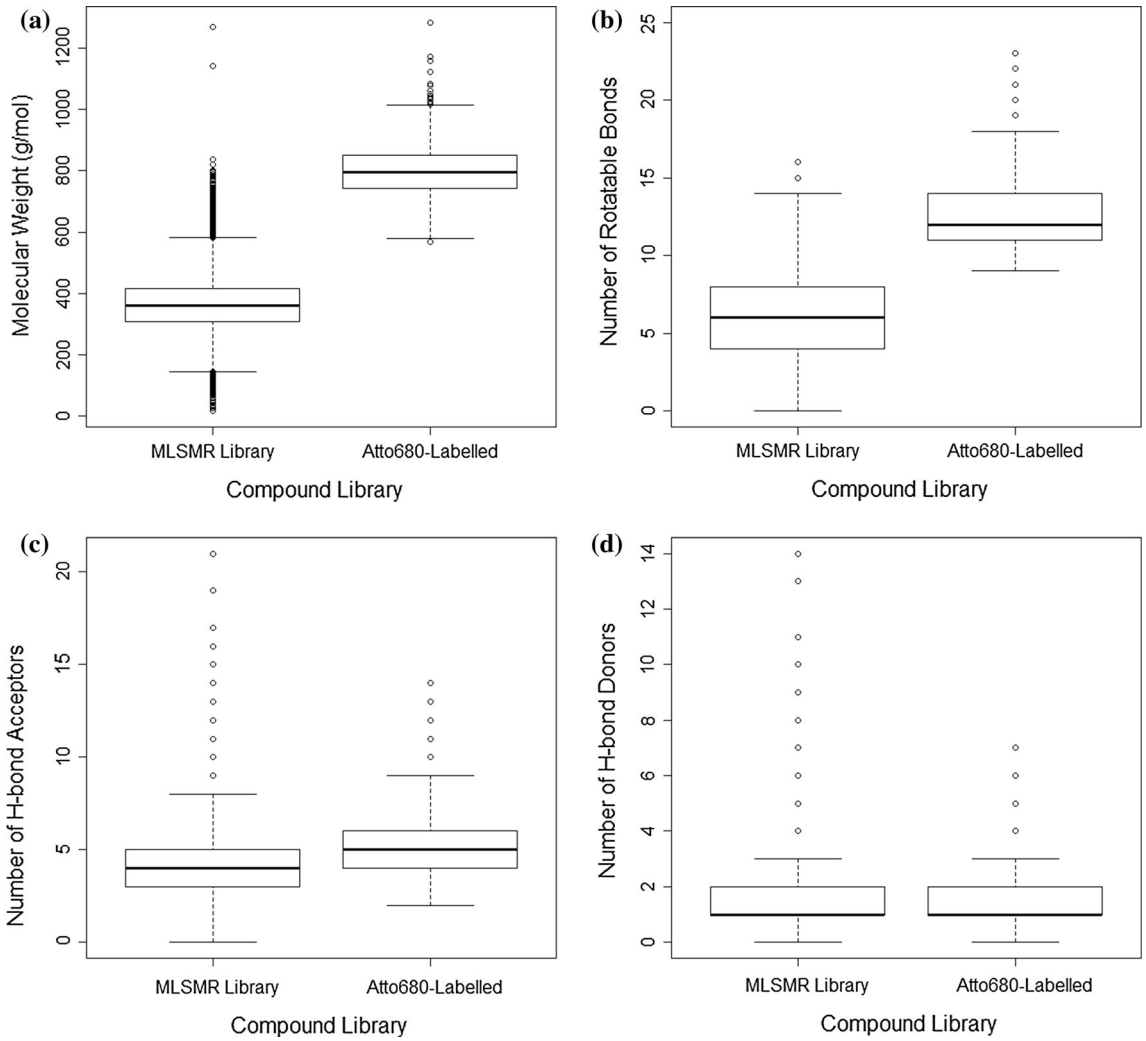

Fig. 9 Distribution of **a** molecular weight, **b** number of rotatable bonds, number of hydrogen donor **c** acceptors and **d** donors for the MLSMR (383,987 compounds) and the ATTO680-labelled (14,862 compounds) libraries. The ATTO680-labelled library is comprised of single-group primary amine compounds from the MLSMR, conjugated to ATTO680. Minimum (*lower bar*), maximum (*upper bar*), median (*dark line inside box*), 25–75 % (*box*), and outlier values are shown. The average molecular weights (g/mol) are 365 and 860 for compounds in the MLSMR and ATTO680-labelled libraries, respectively. For comparison, the average molecular weight for PAD compounds is 243 g/mol, and for FL-PAD compounds is 631 g/mol

the best signal to noise ratios for aqueous applications [54]. ATTO680 has been shown to be an appropriate choice for both steady-state and time-resolved fluorescence applications [54], which would enhance the experimental utility and versatility of any fluorescent conjugates bearing this moiety. ATTO680 NHS-ester (PubChemCID 16218508) was used as the amine-reactive fluorophore (Fig. 2). Its structure is publically available and does not contain exotic atoms that could interfere with VLS. While ATTO680 was chosen for generating an in-house tagged library, the script can easily work other NHS-ester fluorophores.

The Molecular Libraries Small Molecule Repository (MLSMR) (https://mli.nih.gov/mli/compound-repository/mlsmr-compounds/) was selected as a source of small molecule structures due to its public availability and its suitability for probe development. MLSMR is a collection of small molecules accessible for the purposes of high-throughput screening for biomedical applications, maintained as a part of the MLI. The MLSMR restricts entries based on numerous criteria with specific considerations given to the suitability of these compounds for high-throughput experimental screening. For example, minimum water solubility thresholds and minimum vendor amounts and purity were considered. The database is designed to retain a high level of diversity while implementing minimal restrictions on the types of chemical structures [43]. These characteristics may facilitate the translation of future screening efforts to experimental work.

The final fluorescence-labelled database numbers 14,862 compounds, each tagged with the ATTO680 fluorophore.

This library is highly suitable for VLS efforts seeking to identify fluorescent probes for biomarkers. While the use of fluorescent labelling technology has been highlighted by large biomolecule labelling, there is the demand for labelled small molecules for specific in vitro and in vivo applications, the discovery of which can be computationally aided through such methods.

Acknowledgments This research was supported with funds from the Thunder Bay Regional Research Institute, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the RBC Royal Bank's Dr. Mark Poznansky Mentorship Development Award. High-throughput computational work in this project was made possible with resources provided through SHARCNET (www.sharcnet.ca) and Lakehead University's High Performance Computing Centre (LUHPCC). The authors wish to acknowledge Darryl Willick for his help with data transfer and manipulation.

References

1. Johnson I, Michelle TZS (eds) (2010) Molecular probes handbook: a guide to fluorescent probes and labeling technologies, 11th edn. Life Technologies Coorporation, USA
2. Giepmans BN, Adams SR, Ellisman MH, Tsien RY (2006) The fluorescent toolbox for assessing protein location and function. Science (New York, NY) 312(5771):217–224. doi:10.1126/science.1124618
3. Riggs J, Seiwald R, Burckhalter J, Downs CM, Metcalf T (1958) Isothiocyanate compounds as fluorescent labeling agents for immune serum. Am J Pathol 34(6):1081
4. Lazarides E, Weber K (1974) Actin antibody: the specific visualization of actin filaments in non-muscle cells. Proc Natl Acad Sci 71(6):2268–2272
5. Sack U, Conrad K, Csernok E, Frank I, Hiepe F, Krieger T, Kromminga A, Von Landenberg P, Messer G, Witte T, Mierau R (2009) Autoantibody detection using indirect immunofluorescence on HEp-2 cells. Dtsch Med Wochenschr 134(24):1278–1282. doi:10.1055/s-0029-1225278
6. Hoxha E, Harendza S, Zahner G, Panzer U, Steinmetz O, Fechner K, Helmchen U, Stahl RAK (2011) An immunofluorescence test for phospholipase-A2-receptor antibodies and its clinical usefulness in patients with membranous glomerulonephritis. Nephrol Dial Transplant 26(8):2526–2532. doi:10.1093/ndt/gfr247
7. Kinders RJ, Hollingshead M, Lawrence S, Ji J, Tabb B, Bonner WM, Pommier Y, Rubinstein L, Evrard YA, Parchment RE, Tomaszewski J, Doroshow JH (2010) Development of a validated immunofluorescence assay for γH2AX as a pharmacodynamic marker of topoisomerase I inhibitor activity. Clin Cancer Res 16(22):5447–5457. doi:10.1158/1078-0432.ccr-09-3076
8. Schmidt E, Zillikens D (2010) Modern diagnosis of autoimmune blistering skin diseases. Autoimmun Rev 10(2):84–89. doi:10.1016/j.autrev.2010.08.007
9. Chames P, Van Regenmortel M, Weiss E, Baty D (2009) Therapeutic antibodies: successes, limitations and hopes for the future. Br J Pharmacol 157(2):220–233. doi:10.1111/j.1476-5381.2009.00190.x
10. Ntziachristos V, Bremer C, Weissleder R (2003) Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging. Eur Radiol 13(1):195–208. doi:10.1007/s00330-002-1524-x
11. Hilderbrand SA, Weissleder R (2010) Near-infrared fluorescence: application to in vivo molecular imaging. Curr Opin Chem Biol 14(1):71–79. doi:10.1016/j.cbpa.2009.09.029
12. Tung C-H, Lin Y, Moon WK, Weissleder R (2002) A receptor-targeted near-infrared fluorescence probe for in vivo tumor imaging. ChemBioChem 3(8):784–786. doi:10.1002/1439-7633(20020802)3:8<784:AID-CBIC784>3.0.CO;2-X
13. Zhang Z, Achilefu S (2005) Design, synthesis and evaluation of near-infrared fluorescent pH indicators in a physiologically relevant range. Chem Commun 47:5887–5889. doi:10.1039/B512315A
14. Weissleder R, Tung CH, Mahmood U, Bogdanov A Jr (1999) In vivo imaging of tumors with protease-activated near-infrared fluorescent probes. Nat Biotechnol 17(4):375–378. doi:10.1038/7933
15. Messerli SM, Prabhakar S, Tang Y, Shah K, Cortes ML, Murthy V, Weissleder R, Breakefield XO, Tung CH (2004) A novel method for imaging apoptosis using a caspase-1 near-infrared fluorescent probe. Neoplasia 6(2):95–105
16. Dadgar S, Ramjan Z, Floriano WB (2013) Paclitaxel Is an inhibitor and its boron dipyrromethene derivative is a fluorescent recognition agent for botulinum neurotoxin subtype A. J Med Chem 56(7):2791–2803
17. Shah F, Mukherjee P, Gut J, Legac J, Rosenthal PJ, Tekwani BL, Avery MA (2011) Identification of novel malarial cysteine protease inhibitors using structure-based virtual screening of a focused cysteine protease inhibitor library. J Chem Inf Model 51(4):852–864. doi:10.1021/ci200029y
18. Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, Mizote A, Rosen D, Thomas JM, Izumi M, Ganesan A (2009) Identification of a chemical probe for NAADP by virtual screening. Nat Chem Biol 5(4):220–226
19. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, Connolly DT, Shoichet BK (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. J Med Chem 45(11):2213–2221. doi:10.1021/jm010548w
20. Kolb P, Rosenbaum DM, Irwin JJ, Fung JJ, Kobilka BK, Shoichet BK (2009) Structure-based discovery of β2-adrenergic receptor ligands. Proc Natl Acad Sci 106(16):6843–6848. doi:10.1073/pnas.0812657106
21. De Graaf C, Kooistra AJ, Vischer HF, Katritch V, Kuijer M, Shiroishi M, Iwata S, Shimamura T, Stevens RC, De Esch IJP, Leurs R (2011) Crystal structure-based virtual screening for fragment-like ligands of the human histamine H 1 receptor. J Med Chem 54(23):8195–8206
22. Floriano WB, Vaidehi N, Zamanakos G, Goddard WA (2004) HierVLS hierarchical docking protocol for virtual ligand screening of large-molecule databases. J Med Chem 47(1):56–71
23. Dadgar S, Chowdhury M, Phenix C, Floriano WB (2014) Systematic discovery of novel fluorescent molecular probes targeting multiple non-orthosteric spatially distinct sites in BoNTA (in preparation)
24. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3(11):935–949
25. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des 7(2):146
26. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM (1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell 75(3):495–505. doi:10.1016/0092-8674(93)90384-3
27. Huibregtse JM, Scheffner M, Howley PM (1991) A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. EMBO J 10(13):4129

28. Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM (1990) The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 63(6):1129–1136. doi:10.1016/0092-8674(90)90409-8
29. Sobel J, Tucker N, Sulka A, McLaughlin J, Maslanka S (2004) Foodborne botulism in the United States, 1990–2000. Emerg Infect Dis 10(9):1606
30. Cai S, Singh BR, Sharma S (2007) Botulism diagnostics: from clinical symptoms to in vitro assays. Crit Rev Microbiol 33(2):109–125
31. Arnon SS, Schechter R, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Hauer J, Layton M (2001) Botulinum toxin as a biological weapon. JAMA 285(8):1059–1070
32. Jankovic J (2004) Botulinum toxin in clinical practice. J Neurol Neurosurg Psychiatry 75(7):951–957
33. Münchau A, Bhatia K (2000) Regular review: uses of botulinum toxin injection in medicine today. Br Med J 320(7228):161
34. Bolton E, Wang Y, Thiessen PA, Bryant SH (2008) PubChem: integrated platform of small molecules and biological activities. In: Wheeler RA, Spellmeyer DC (eds) Annual reports in computational chemistry, vol 4. American Chemical Society, Washington, DC
35. Lipinski CA (2000) Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 44(1):235–249. doi:10.1016/s1056-8719(00)00107-6
36. Molecular Operating Environment (MOE) (2010) 2010.10 edn. Chemical Computing Group, Inc., Montreal, Quebec, Canada
37. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. Tetrahedron 36:3219–3228. doi:10.1016/0040-4020(80)80168-2
38. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17(5–6):490–519
39. Mujumdar RB, Ernst L, Mujumdar SR, Lewis CJ, Waggoner AS (1993) Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters. Bioconjug Chem 4:105–111
40. Southwick P, Ernst L, Tauriello E (1990) Cyanine dye labeling reagents—carboxymethylindocyanine succinimidyl esters. Cytometry 430:418–430
41. Brady GP, Stouten PF (2000) Fast prediction and visualization of protein binding pockets with PASS. J Comput Aided Mol Des 14:383–401
42. PRISM (2013) 6.UPDATE for Windows* edn. GraphPad Software, La Jolla, California, USA
43. Program ML (2007) MLSMR compounds. https://mli.nih.gov/mli/compound-repository/mlsmr-compounds/. Accessed June 2013 2013
44. PubChem N (2011) PubChem3D release notes. http://pubchem.ncbi.nlm.nih.gov/release3d.html. Accessed June 2013
45. Halgren T (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17:490–519
46. Carta G, Knox AJ, Lloyd DG (2007) Unbiasing scoring functions: a new normalization and rescoring strategy. J Chem Inf Model 47(4):1564–1571. doi:10.1021/ci600471m
47. Konstantinou-Kirtay C, Mitchell JB, Lumley JA (2007) Scoring functions and enrichment: a case study on Hsp90. BMC Bioinform 8:27. doi:10.1186/1471-2105-8-27
48. Jacobsson M, Karlen A (2006) Ligand bias of scoring functions in structure-based virtual screening. J Chem Inf Model 46(3):1334–1343. doi:10.1021/ci050407t
49. Pan Y, Huang N, Cho S, MacKerell AD Jr (2003) Consideration of molecular weight during compound selection in virtual target-based database screening. J Chem Inf Comput Sci 43(1):267–272. doi:10.1021/ci020055f
50. Allen M, Reeves J, Mellor G (2000) High throughput fluorescence polarization: a homogeneous alternative to radioligand binding for cell surface receptors. J Biomol Screen 5(2):63–69. doi:10.1177/108705710000500202
51. Jolley ME (1996) Fluorescence polarization assays for the detection of proteases and their inhibitors. J Biomol Screen 1(1):33–38
52. Checovich WJ, Bolger RE, Burke T (1995) Fluorescence polarization—a new tool for cell and molecular biology. Nature 375(6528):254–256
53. Warren GL, Andrews CW, Capelli A-M, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS (2005) A critical assessment of docking programs and scoring functions. J Med Chem 49(20):5912–5931. doi:10.1021/jm050362n
54. Buschmann V, Weston KD, Sauer M (2003) Spectroscopic study and evaluation of red-absorbing fluorescent dyes. Bioconjug Chem 14:195–204. doi:10.1021/bc025600x