Bioorganic & Medicinal Chemistry 17 (2009) 4825–4832

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Phage display screening against a set of targets to establish peptide-based sugar mimetics and molecular docking to predict binding site

Limin Yu [a], Peggy Shuang Yu [b], Elizabeth Yee Yen Mui [b], Jennifer C. McKelvie [c], Thi Phuong Tu Pham [a], Yan Wen Yap [d], Wan Qing Wong [d], Jiawen Wu [a], Weiqiao Deng [a,*], Brendan P. Orner [a,*]

[a] Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
[b] School of Life and Chemical Technology, Ngee Ann Polytechnic, Singapore 599489, Singapore
[c] School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[d] School of Biological Sciences, Nanyang Technological University, Singapore 637371, Singapore

ARTICLE INFO

Article history:
Received 31 December 2008
Revised 24 March 2009
Accepted 25 March 2009
Available online 1 April 2009

Keywords:
Phage display
Docking
Glyomimetic
ConA
Peptide

ABSTRACT

A novel selection approach is presented to screen phage display peptide libraries against sets of receptors that share specificity for the same ligand. This strategy was applied to the discovery of glycomimetic peptides. Through these screens, a number of peptide clones were discovered that bind the lectins used in the screen, in a sugar competitive manner. In addition, the majority of the selected peptides demonstrate sugar type mimicry consistent with lectin specificity. Docking studies were conducted to establish whether the mimetic peptides bind to the lectin ConA at the sugar binding site or to a nearby, alternative site shown to bind to YPY-containing peptides previously discovered from single-target screens. Of the three cyclic peptides subjected to computational docking, CNTPLTSRC had the highest predicted affinity and CSRILTAAC demonstrated specificity for the sugar binding site comparable to the natural ligand itself.

1. Introduction

Discovering drugs with multiple activities is an emerging goal of medicinal chemistry. For example, a drug that can bind a set of homologous targets from multiple species could be potently anti-fungal or anti-biotic, and one that inhibits multiple enzymes in a viral replicative cycle could help to bypass difficult-to-administer, 'cocktail'-style drug regimens. In addition, designing or screening for a molecule that has activity against a set of probable mutants of a protein target could be a first step in developing a therapy that is less susceptible to resistance.[1] In addition, there is precedent suggesting that molecules with complex chemical-genetic phenotypes may require activities against multiple targets.[2,3] However the development of high throughput affinity screens to select in favor of molecules that can bind to a number of specific targets remains largely unexplored mainly because the difference between a non-specific molecule and one that has multiple specific activities is somewhat obscured. As part of a larger project to rationally design general screening strategies for the discovery of specific ligands sharing a desired level of multi-receptor binding activity, we revisited the generation of sugar-mimicking peptides.

Although many high-yielding and specific methodologies have been developed,[4–6] carbohydrate-focused synthetic chemistry is still the realm of specialists. Difficulties include tedious but essential protection/deprotection schemes, characterization and control of complex stereochemistry, and purification complications caused by the polar and non-chromagenic nature of the products. Multiple solid-phase solutions have been developed,[7,8] however, they have not proven to be completely universal for the synthesis of all the stereo- and regio-chemical isomers of a complex carbohydrate.

The lack of robust methods for the routine synthesis of carbohydrates is unfortunate in light of the fact that they are obvious targets for medicinal advancement. Interactions involving carbohydrates play roles in fertilization, viral and bacterial infection, and cancer metastasis, among other important biological processes potentially treated with therapeutic agents. Moreover, the fact that unique carbohydrates are expressed on the surface of pathogens would make them ideal starting points for the generation of vaccines; unfortunately carbohydrates are rarely immunogenic.[9,10]

The fundamental drawbacks of carbohydrates as therapeutic agents have led to the realization that molecules which mimic their activity need to be established. Numerous small molecule solutions to this problem have been developed, and most provide a cyclic scaffold for the vectoral projection of hydroxyl functionality.[11,12] An advantage of these strategies is that due to their ra-

* Corresponding authors. Tel.: +65 6316 8796 (W.D.); tel.: +65 6316 8757; fax: +65 6791 1961 (B.P.O.).
E-mail addresses: wqdeng@ntu.edu.sg (W. Deng), orner@ntu.edu.sg (B.P. Orner).

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.03.054

tional design and structural resemblance to the carbohydrates they are intended to mimic, many of these molecules can be modified for the application to multiple systems.

A limitation of these strategies is shared by most other rationally designed small molecule approaches when they are expanded into libraries of compounds. Synthetic small molecule libraries based on single scaffolds often explore limited chemical and spatial diversity due to synthetic constraints and the restricted number of conformers energetically permitted to medium sized rings. Moreover, the overall size of small molecule libraries is often limited in response to challenges associated with the identification and characterization of resulting hits. Screening strategies focused on sugar structure mimics have a further drawback in that the serendipitous discovery of non-structural, but functionally active compounds, based on unexpected structures, rarely occurs. This is most likely due to limitations inherent in the design of the strategy itself.

Molecular biology-based combinatorial approaches can compliment small molecule library tactics by circumventing a number of their shortcomings. Libraries with multiple orders of magnitude more library members are possible and, although they are often limited to nucleic acid or peptide-based molecules, the physical size of the molecules can be easily varied as can the wide range of chemical functionalities explored. In addition, the libraries can be generated and expanded using very routine techniques without the need for tedious chemical synthesis and purification. Moreover, molecular biology-based combinatorial techniques link the phenotype (presented functionality which directly affects their activity) of the library members with their genotype, allowing simple and straightforward tagging and deconvolution, thereby permitting much larger diversities than would be possible for small molecule libraries. Furthermore, molecular biology techniques can be amplified between rounds of screening, effectively permitting extremely small sub-populations to be selected out of a pool of candidates. Finally, due to the large functional diversities of molecular biology-generated libraries, they can be used to discover classes of functional mimics that would not immediately be predicted for activity by rational design.

One molecular biology-based combinatorial technique that has been widely adopted for the screening of peptide libraries is phage display.[13–15] This approach screens libraries of randomized peptides that have been fused to coat proteins on bacteriophage particles. Each phage encapsulates the genome that encodes for the peptide sequence displayed on its surface thereby linking an identification tag to the peptide. Affinity selection (or 'panning') can be performed to select for binders that can then be amplified following their infection of bacteria. Generally this is repeated for three to six rounds after which the genomes of the winning phage are sequenced to determine the identity of the peptide hits. Phage display can be performed in any lab with basic microbiology facilities and does not require any out of the ordinary expertise.

The screening of phage-displayed combinatorial peptide libraries has been used previously to discover peptides that functionally mimic small molecules. For example, peptides have been discovered that bind to the biotin binding site of variants of the protein streptavidin[16,17] and therefore serve as a functional mimetic of the biotin. Although their binding can be rationalized after sufficient structural information is obtained,[18] these sequences would never have been predicted rationally nor would molecules with similar modes of binding have been discovered by the screening of small molecule libraries based on designed scaffolds.

Following this strategy, phage display has been used to discover functional mimics of sugars.[19–28] This is an enticing approach, in that glycomimetic peptides could be used to inhibit sugar–receptor interactions, or as part of a vaccine strategy.[29,30] Contributing to this attractiveness is that the rational design of small molecule peptidomimetics based on peptide sequences is a mature field and has had great success in medicinal chemistry.[31] In other words, if a peptide sequence can be established that truly mimics a sugar, developing small molecule glycomimetics should be relatively routine (Fig. 1).

Structural studies, however, have raised the question as to whether some of these peptides are indeed true sugar mimetics. For example, it has been established that a number of phage-selected peptides bind to a distinct site near the sugar binding site

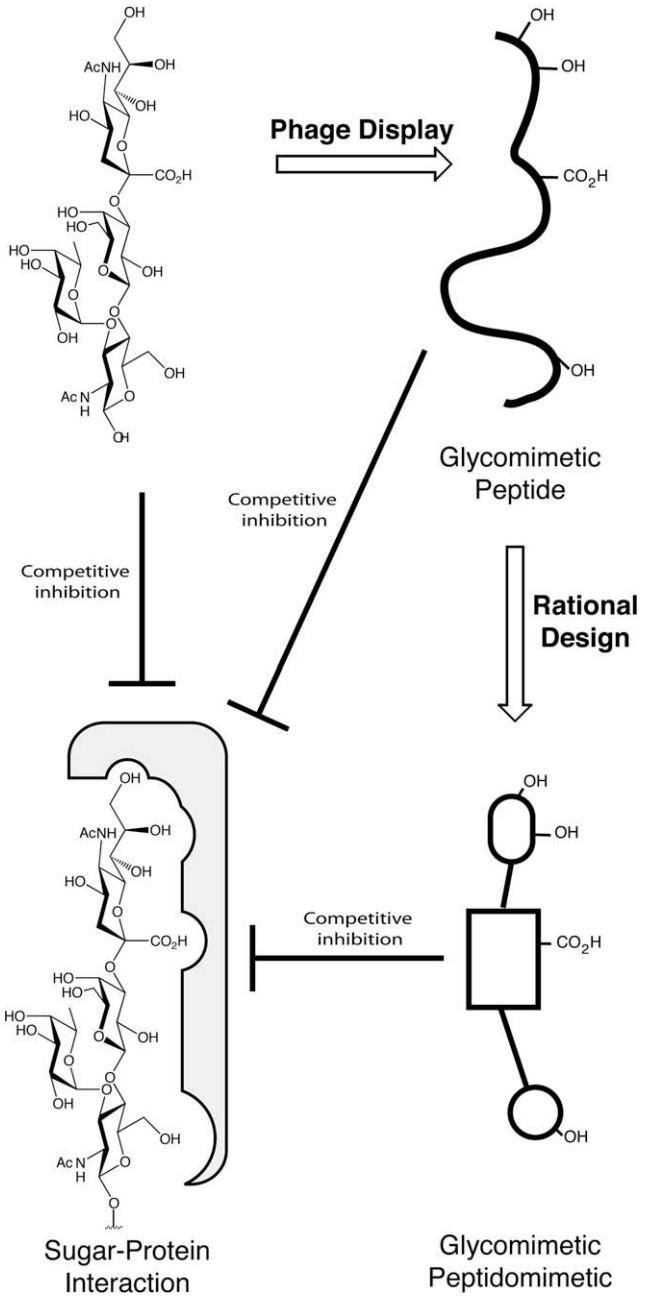

Figure 1. Sugar–protein interactions are attractive drug targets in that they are involved in many biological processes. They could be rationally inhibited by the administration of structurally related carbohydrates, but this approach is limited by synthetic intractability. An alternative approach is to use a non-carbohydrate-based scaffold to vectorally project essential functionality. Peptides are an attractive scaffold in that they can be rationally advanced to small molecule peptidomimetics and can be discovered using molecular biology-based combinatorial techniques. One currently unexplored problem is how to screen for glycomimetic peptides that bind to the multiple targets of the original mimicked sugar.

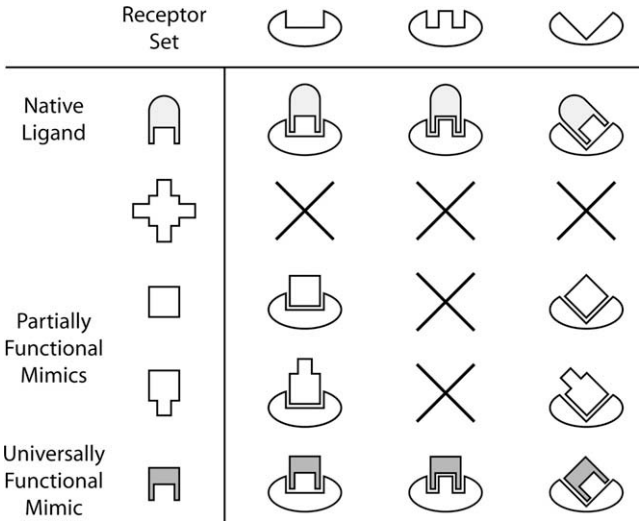

Figure 2. A universally functional mimic of a native ligand should bind all members of the receptor set that interacts with the native ligand otherwise it should be considered to be only partially functional.

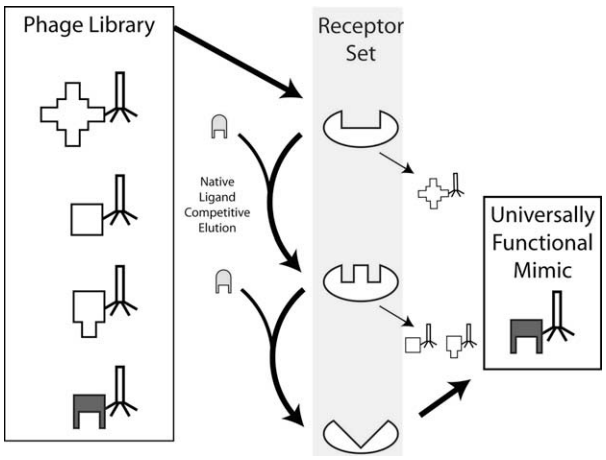

Figure 3. A single round of a phage display screening strategy to discover universally functional mimics of a native ligand that can bind to multiple receptors. This strategy involves screening against a set of targets and eluting with a native ligand to release bound phage clones.

of the lectin Concannavalin A, but not directly to the sugar binding site itself. These peptides are therefore not acting as true mimics of the sugar substrate and are most likely imitating some of its activity via allosteric communication between the two binding sites.[32,33] It is possible that this binding mode is a result of how the screens were conducted. Following this logic it may therefore be possible to perform a screen that can select a 'truer' sugar mimetic for this system. We postulated that the problems with the resulting sugar mimetic peptides were rooted in the fact that the initial screens were performed against a single target although the sugars themselves bind to multiple receptors.

As part of a larger project to establish peptides as mimics of ligands that bind to multiple proteins, we set out to explore ways of targeting a class of receptors through screening. Most affinity screens are designed so that the library member with the highest affinity is selected. As a result, screening against a single target will result in a molecule that binds that target well although it may not bind in the mode desired. We therefore designed a screening strategy that does not necessarily select for peptides that are the best binders to a single target but rather for peptides that are the best binders to multiple targets (Fig. 2). Our selection strategy is designed to screen against a set of multiple targets within the same round so that only phage clones that bind to all targets within that set are selected. In addition, the successful clones are eluted from the bound target with the native ligand to ensure that the peptides are binding to the native active site. Due to the substantial size difference between a typical native ligand and the M13 phage viron, the excess eluting ligand can be removed via dialysis in between targets as part of a round of screening. This strategy can therefore select for phage that are binders to multiple targets while reducing the number of hits that bind at non-specific binding sites (Fig. 3).

2. Results and discussion

2.1. Phage display

Although this strategy could be applied to many systems, the first application that was attempted was the discovery of truly glycomimetic peptides. Due to the ready commercial availability of sets of lectins with shared selectivities, we set out to establish peptide mimics of their sugar ligands. Two peptide libraries were used in the screening. One, 'X12', displays 12 random residues in a linear format and the other, 'C–X7–C', displays seven random amino acids flanked by two cysteine residues that cyclize the peptide through the formation of a disulfide bond. Using these libraries, phage display screening was performed against the glucose/mannose-binding lectins Concanavalin A (ConA), Lens Culinaris Agglutinin (LCA), and Pisum Sativum Agglutin (PSA) which are derived from jack bean, lentil, and pea respectively. Phage display screens in general must balance stringent selective pressure with permissiveness. This is especially important in early rounds where the number of copies of each clone is small. As the goals of these screens are unique, the optimal balance was not completely understood. Multiple strategies that are stringent and permissive to different degrees were envisioned to select for multiple targets. Three such screens were performed. Strategy A expands the number of targets in each round as the screen progresses, repeating previously screened targets. Strategy B screens against only one target in the first round and an expanded set in the second; in the third round it only repeats the second of the previously screened targets and adds another target. Strategy C is Strategy B with the first round removed (Fig. 4).

To discover glycomimetic peptides these three strategies were employed using the lectins as targets. Strategies A and B both used ConA, the least expensive lectin, as the first target, LCA as the second and PSA as the third whereas all possible combinations of targets were employed in the implementation of Strategy C. Strategy A resulted in no recovered phage, suggesting that it was too stringent. This stringency may be a result of the third round involving three targets. It is possible that inserting another two-target round into the middle of this strategy might sufficiently maturate the library population to prepare it for the three-target round. Of course, the problem could arise not with the strategy itself but with relative and absolute affinities and the ordering of the targets themselves. Strategies B and C, however resulted in the isolation of eleven phage which converged strongly but to varying degrees with screens employing the C–X7–C library, on the whole converging more tightly than the X12 library (Table 1).

The observed strong convergence within each screen but lack of clear consensus between the different screens may be a result of screening against sets of targets instead of a single one combined with rapid convergence (due to over-stringent conditions), statistical phenomena (brought about by the incomplete diversity coverage of the library), or relative binding avidities between the phage and the various lectins (caused, in part, by the multivalent nature of the phage).

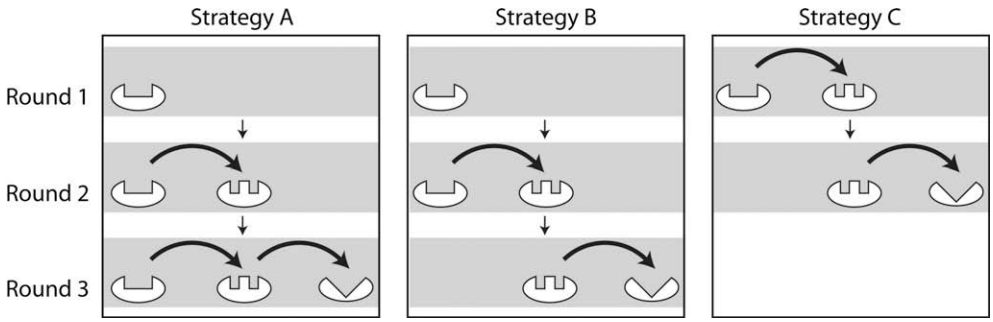

Figure 4. Many strategies could be employed to screen for sets of multiple targets to find peptides that have broad based activity. The strategies must balance stringency with permissiveness to discover not the best binder for one target but the best binder that can bind all targets.

Table 1
Convergent sequences after phage display with C–X7–C[a] and X12 libraries employing different screening strategies and lectin combinations.

Screen[b]	Order of lectins screened	Convergent sequences (occurrence)[c,d]
A	Round 1: ConA	Nil
	Round 2: ConA, LCA	
	Round 3: ConA, LCA, PSA	
B	Round 1: ConA	**CNTPLTSRC** (100%)
	Round 2: ConA, LCA	**STLHALDSHLAL** (36%)
	Round 3: LCA, PSA	HWDPFSLSAYFP (18%)
C1	Round 1: ConA, LCA	Nil
	Round 2: LCA, PSA	
C2	Round 1: LCA, ConA	**CKPHASSMC** (100%)
	Round 2: ConA, PSA	
C3	Round 1: LCA, PSA	**CSRILTAAC** (100%)
	Round 2: PSA, ConA	
C4	Round 1: ConA, PSA	**CSPIYKDTC** (100%)
	Round 2: PSA, LCA	**YQTSSPAKQSVG** (100%)
C5	Round 1: PSA, ConA	**CKQMGTLKC** (100%)
	Round 2: ConA, LCA	**ADPQFSGHTPPQ** (75%)
		VPPTRPATSTQL (25%)
C6	Round 1: PSA, LCA	CNNPRAINC (100%)
	Round 2: LCA, ConA	

[a] The peptides are cyclized through disulfide bonds between the two cysteine residues.
[b] Letters correlate to the strategies depicted graphically in Figure 4.
[c] Percentages represent the percent that sequence appeared out of the total number of clones sequenced. In some instances the numbers do not total 100% either because different libraries were used or because some sequences only appeared once and were therefore not determined to be a convergent sequence. 'Nil' indicates no convergent sequences.
[d] Sequences in bold represent clones that were carried forward for the binding experiments.

Many of the previously reported glycomimetic peptides contain the consensus sequence Tyr-Pro-Tyr (YPY). It has been speculated that the YPY motif mimics the sugar hydroxyls by binding through the phenol groups of the tyrosine side chains.[16,17,21] Interestingly, none of our convergent clones contained the YPY motif. This is not altogether surprising considering the fundamental selection pressure differences between screening against single and multiple targets. Although the resulting sequences from our screens did not converge on a YPY consensus, they are rich in threonine (T) and serine (S) residues providing possible hydrogen bonding interactions. Proline (P) residues, which appear in many of the convergent sequences from the screens, may serve an important structural role by spatially orienting key side chains.

Nine of the eleven convergent phage clones, due to either their strong convergence or the desire to include representatives from different screens and/or libraries, were selected for further analysis. Initial assay optimization studies suggested that the majority of the clones from the C–X7–C library demonstrated higher specificity than the X12 library (data not shown). An exception to this trend was CNNPRAINC, which therefore was dropped from further study.

2.2. Binding analysis

To determine the ability of the phage display derived sequences to bind to the lectins, a phage based ELISA experiment was carried out (Fig. 5). Microtiter wells were incubated with lectins, blocked, incubated with a normalized titer of phage, washed, and then probed for phage binding by addition of an anti-M13 HRP/antibody conjugate. A negative control consisted of using wells that were blocked but with no lectin. Positive control wells were incubated

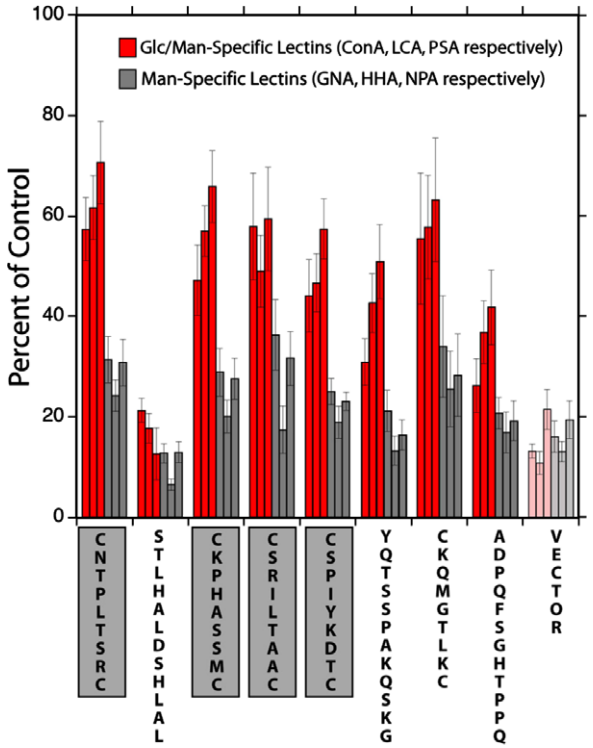

Figure 5. Binding of phage display derived sequences to glucose/mannose selective lectins (ConA, LCA, PSA respectively, in red) compared to mannose-only selective lectins (GNA, HHA, NPA respectively, in gray). The boxed sequences were carried on for competition experiments. The sequences from the C–X7–C library were cyclized through a disulfide bond between the two cysteine side chains. 'Vector' refers to a phage with no peptide fusion displayed on its coat which acts as a negative control. Error bars represent standard error for at least three replicates. '100% of control' reflects the phage solution binding to an anti-phage antibody whereas '0% of control' is derived from the same phage solution binding to a blocked well with no target (no antibody or lectin).

with an anti-M13 antibody (with no conjugated HRP) replacing the lectins. An additional negative control consisted of the use of phage with no peptide fusion, that is, phage encapsulating only cloning vector. In order to determine the specificity of the interactions, as ConA, LCA, and PSA bind both glucose and mannose, three mannose-only specific lectins (GNA, HHA, and NPA) were used as further negative controls. As expected, the phage with no fusion exhibited very little binding to the lectins when compared to the absence of target negative control and the antibody positive control. In addition, it showed no appreciable difference in binding to the lectins with different sugar selectivity. With the exception of STLHALDSHLAL, which showed a similar binding profile to that of the negative control phage, all the selected clones bound strongly to ConA, LCA, and PSA. In addition, all these sequences demonstrated marked selectivity by exhibiting weaker affinity toward the three mannose specific lectins. The clones derived from the C–X7–C library demonstrated stronger binding to the mannose/glucose specific lectins than did those derived from the X12 library, most likely due to their pre-organization in a binding conformation. While this experiment does not definitively prove that these peptides are indeed glucose mimics, the fact that they show diminished binding to a set of lectins that share mannose-only specificity is implicative. Further experiments were deemed necessary, however, to strengthen this implication. Four clones that exhibited the highest lectin selectivity were carried forward to determine whether their binding is sugar-competitive.

With the knowledge that the peptides bind to the desired lectins and that these interactions are lectin specific, insight into their mechanism of binding was sought. To further confirm that the peptides behave in a sugar-like manner by binding in or near the sugar binding site, the ELISA was modified so that the lectin-phage complexes were incubated solutions of various sugars (Fig. 6). For all peptides and all lectins, the binding was competitive with both glucose and mannose, as would be expected for these lectins, indicating that the phage-display derived peptides compete with the sugar binding site, further supporting the notion that they are indeed sugar mimetic.

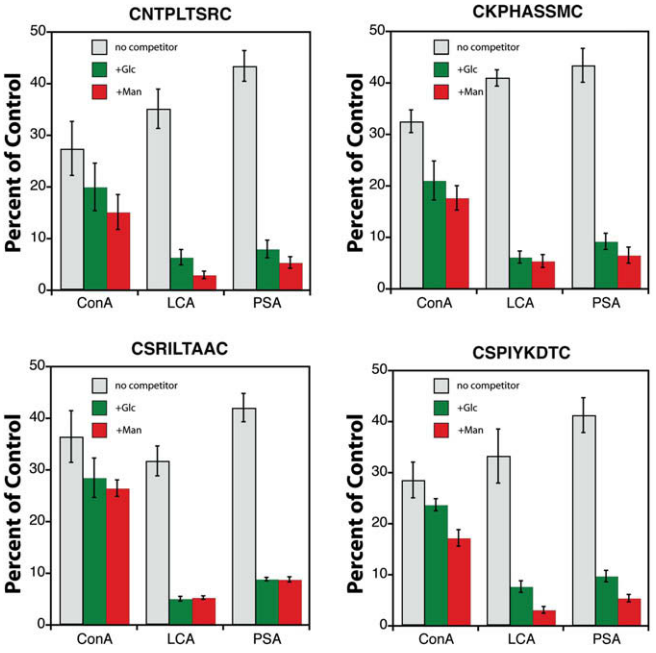

Figure 6. The binding of the phage-display derived sequences is sugar competitive for the lectins ConA, LCA, and PSA suggesting that they are binding at or near the sugar binding site. Addition of 250 mM solutions of either mannose (red) or glucose (green) releases the peptides, indicating competition for the sugar binding site. Gray indicates addition of buffer with no sugar. All four clones are cyclized through disulfide bonds through the two cysteine residues. Error bars represent standard error of at least three replicates. '100% of control' reflects the phage solution binding to an anti-phage antibody whereas '0% of control' is derived from the same phage solution binding to a blocked well with no target (no antibody or lectin).

2.3. Computational docking

To refine the characterization of the binding mode of the phage-display selected peptides, and to further determine how sugar-like their binding is, computational docking studies were conducted. Both the sugar binding site and the alternative site that is bound by the YPY peptides were sampled to determine the specificity of our peptides. The sugar binding site was defined by Ala207, Arg228, Asn14, Asp16, Asp208, Gly98, Gly224, Gly227, Leu99, Thr 226, Tyr12, and Tyr100 from the structure of ConA bound to a complex carbohydrate (1CVN),[34] and the alternative binding site included Asn41, Asn44, Asp203, Gln43, His205, Met42, Pro23, Pro206, Ser21, Ser201, Ser204, and Tyr22 from the co-crystal structure of ConA bound to the peptide MYWYPYASGS (1JUI).[29] As allostery might cause sugar-competitive binding even if a ligand is bound to the alternative binding site we sought to determine not only the strength and geometry of binding for our phage library-selected peptides, but also the preference for these two sites within the target.

Computational docking was performed using the Glide docking software within the Schrodinger Suite 2007 package in conjugation with the OPLS 2007 force field. (Fig. 8) To verify the docking parameters, the trimannose ligand of the 1CVN structure was removed from the binding site and re-docked into the sugar binding site. The root mean square deviation between the best docking conformation of the trimannose and the X-ray crystallographic conformation was 0.7 Å. This result demonstrates that the parameters are reliable and can reproduce the experimental binding mode for a ligand, and that the chosen parameters are capable of reproducing experimental binding geometries. 'G-score' functions are calculated from the docked structures (a more negative G-score indicates stronger binding) and this affinity-related value is broken into component contributions enabling relative comparisons between ligands. (Tables S1 and S2 and Fig. 7) The G-score for trimannose binding to the sugar binding site is −6.81 and −4.15 for the

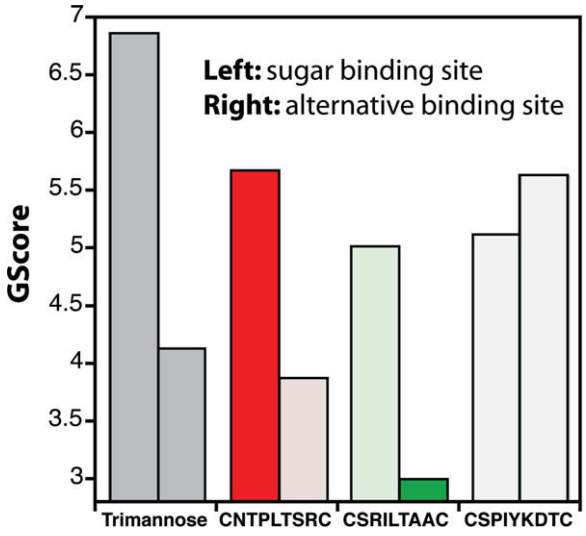

Figure 7. Relative docking energies of carbohydrate ligand and various phage-display derived ligands for the sugar binding site (left) and the alternative, peptide binding site (right) of ConA. Red highlights the highest peptide binding energy for the sugar binding site and green the lowest peptide binding energy for the alternative site. (Note: GScores are depicted as positive values).

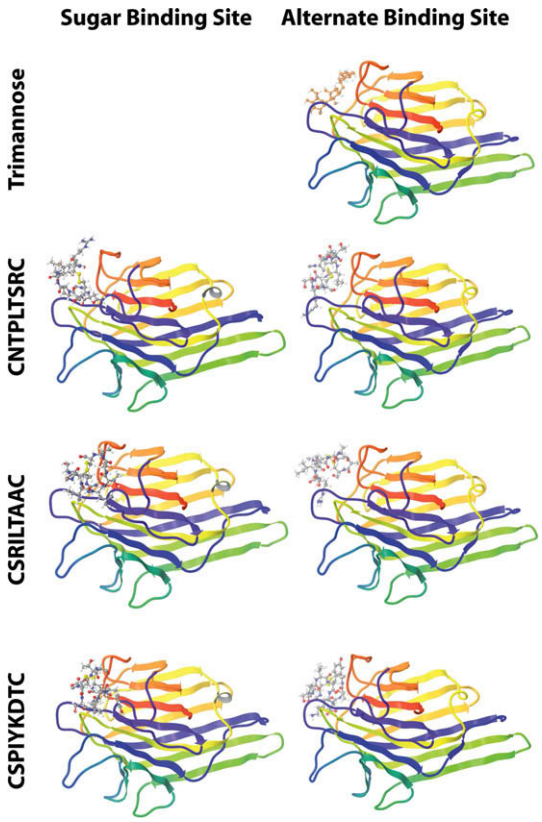

Figure 8. Docked structures of trimannose and phage-display derived peptides bound to the sugar binding site and alternative binding site of ConA. The peptides are cyclized through a disulfide bond between their terminal cysteine residues. In addition they are drawn as the C-terminal primary amide and N-terminal acetate. Structure files derived from these docking calculations are included as part of the supporting information.

peptide binding site, suggesting that, as expected, the carbohydrate has a preference for binding the sugar binding site. However, these data also suggest that the peptide binding site has some sugar binding character which would explain why traditional combinatorial screens and non-computational binding studies might have difficulties establishing the site specificities of YPY peptides. Comparison of the component interaction energies between the two docked structures indicates that the sugar is binding the two sites in different modes. Van der Waals and Coulombic interactions provide near equal contributions to the binding affinity of trimannose in the peptide binding site and hydrogen bonds are negligible. However, when bound to the sugar binding site, hydrogen bonds play a larger role and energetic contributions due to Coulombic contacts dominate over Van der Waals.

Docking between the peptide CSPIYKDTC and ConA indicates an affinity preference disfavoring the sugar binding site in favor of the peptide binding site (−5.13 vs −5.64 G-scores, respectively) (Tables S1 and S2 and Fig. 7), which parallels the structural data for the YPY-containing peptide. Within the carbohydrate binding site, the peptides form hydrogen bonds with residues Arg228, Asp208, Gly224, and Tyr12 whereas in the alternative binding site the peptide interacts with residues Asn41, Asp203, Met42 and Ser201.

Of the three phage display-derived peptides subjected to computational docking CNTPLTSRC demonstrates the highest affinity for the carbohydrate binding site. In addition, it exhibits a preference for this site over that of the peptide binding site (−5.68 and −3.89 G-scores, respectively) (Tables S1 and S2 and Fig. 7). When bound to the sugar binding site, the peptide forms hydrogen bonds with residues Asp16, Gly224, Thr15, Tyr12, and Tyr100.

The peptide CSRILTAAC docks to the sugar binding site by making hydrogen bonds with the residues Arg228, Asp16, and Asp208. Although it has a lower affinity than CNTPLTSRC for the sugar binding site, it also has less binding ability for the alternative binding site (−5.03 and −3.01 G-scores, respectively). In fact, taking the ratio of the predicted affinities for the two sites indicates that it has specificity on par with the natural ligand (Tables S1 and S2 and Fig. 7). In terms of specificity for the two binding sites, this peptide is the most glycomimetic.

Combined, these results suggest that through screening phage libraries against a set of lectins that bind a shared sugar, it is possible to discover peptides that mimic the binding properties of that sugar. This method was even able to combinatorially find ligands that distinguish the sugar binding site from an alternative binding site. The subtle differences between these binding sites are evidenced by the fact that even the natural ligand shows affinity (albeit poorer) for the alternate site. The screening technique was able to discover a peptide with nearly equal site specificity as the natural ligand. As YPY-containing peptides are known to have specificity for the peptide binding site, it may be possible to use these peptides to further modify our screening strategies to include a negative selection step to remove binders competitive for the alternate site and thereby discover peptides that are even more selective for the sugar binding site. Although no clear consensus sequences immediately stand out, perhaps with increased binding specificity a deeper understanding of the structure–activity relationships will be developed and rules for the eventual rational design of sugar mimicking-peptidomimetics can be established. In addition, this study emphasizes the role computational docking can play to help peel apart the delicate affinity balance between binding sites of subtle specificity. Comparisons of this sort would be difficult, if not impossible, to perform cleanly purely by experimentation. With strong evidence that screens of this type can achieve improved mimicry of natural ligands compared to other screening strategies, we are applying them to other systems.

3. Conclusions

A new selection approach was developed to screen phage display peptide libraries against a set of lectins that share specificity for the same sugar ligand. Through these screens, a number of peptide clones were discovered that bind the lectins used in the screen. In addition, most showed significantly weaker binding to lectins that have selectivity for other sugars. This binding was shown to be competitive with the natural ligand, suggesting that they share a binding site. Exploiting the ability of computational analyses to peel apart binding subtleties not readily investigated through experimentation alone, docking studies were conducted to establish whether the mimetic peptides were binding to the sugar binding site or a nearby, alternative site. Of the three sequences subjected to computational docking, CNTPLTSRC had the highest predicted affinity and CSRILTAAC demonstrated a specificity for the sugar binding site comparable to the natural ligand itself.

4. Experimental

Concanavalin A (ConA), Pisum Sativum Agglutinin (PSA), Lens Culinaris Agglutinin (LCA), Galanthus Nivalis Agglutinin (GNA), Hippeastrum Hybrid Agglutinin (HHA), and Narcissus Pseudonarcissus Agglutinin (NPA) were obtained from Vector Laboratories (Burlingame, CA). Phage displayed random peptide libraries ('Ph.D.-C7C' and 'Ph.D.-12') and M13KE cloning vector (for 'vector phage') were from New England Biolabs. Anti-M13 monoclonal antibody (Cat. Number: 27942001) and Anti-M13-HRP monoclonal conjugate (Cat. Number: 27942101) were from GE Healthcare (Sin-

gapore). All dialysis between lectins in a round was performed with Slide-A-Lyzer MINI dialysis units (3500 MWCO) from Pierce. The e. coli strain ER2738 was obtained from New England Biolabs and was used for all tittering and amplification steps. All dideoxy sequencing was performed by 1st Base Pte. Ltd. (Singapore).

4.1. General screening procedure

Each lectin (1 mg) was dissolved in TBSC (500 μl, 50 mM Tris, 150 mM NaCl, 0.1 mM $CaCl_2$, pH 7.5) and dialysis was performed against TBSC (500 ml × 2, 30 min) to remove excess salt. Microtiter wells from a 96-well plate were incubated with diluted lectin solution (100 μg/ml, 100 μl/well) at 4 °C overnight. The wells were washed with TBSCT (120 μl/well, 4×, TBSC containing 0.1 % (v/v) Tween-20, pH 7.5). The plates were blocked with 1% BSA/TBSCT (100 μl/well, 37 °C). Phage solutions diluted in 1% BSA/TBSCT ($\sim 10^{10}$ pfu/ml, 100 μl/well) were added to the wells and incubated (37 °C, 1 h). The wells were then washed (TBSCT, 120 μl/well, 8×) to remove the unbound phage. Then bound phage were eluted by addition of glucose solution (500 mM glucose in TBST, 75 μl/well, incubation 10 min, 37 °C, 2×).

In the case where screening against another lectin was to be performed before amplification, glucose was removed by dialyzing the two separate elution solutions (75 μl/each) against TBSC (200 ml, 15 min, 4×).

After a round of screening against single or multiple lectins, the eluted phage (150 μl) was amplified and the concentration of the eluted phage was determined by tittering following standard procedures.[11,12] At the end of each screen, serial dilutions of phage solutions were plated, individual plaques were amplified, and their DNA was isolated and submitted for sequencing.

4.2. Binding ELISA

Either lectins (100 μg/ml, 50 μl/well), or anti-M13 monoclonal antibody (10 μg/ml, 100 μl/well) were incubated in a microtiter well of a 96 well plate (4 °C, overnight). The wells were then washed (TBSCT, 120 μl/well, 4×). These wells and wells without protein (acting as negative controls) were then blocked with 0.1% BSA/TBSCT (100 μl/well, 37 °C, 1 h).

Phage solutions (10^{10} pfu/ml in 1% BSA/TBST) were added to the wells and incubated (37 °C, 1 h). The wells were then washed (TBSCT, 120 μl/well, 4×). Anti-M13-HRP (10 μg/ml, 100 μl/well) was added and incubated (37 °C, 1 h), and the wells were washed (TBSCT, 120 μl/well, 4×). The substrate ABTS (0.22 mg/ml, 100 μl/well) was added and incubated (rt, 40 min). The optical density (O.D.) was measured at a wavelength of 410 nm. The wells coated with anti-M13 monoclonal antibody and blocked wells without target proteins were used as a positive control and a negative control, respectively. Standard error was calculated from at least three replicates.

4.3. Competition ELISA to confirm binding to the sugar binding site

Competition ELISA was performed identically to the ELISA except that the phage solutions contained either 250 mM glucose or 250 mM mannose. O.D. readings were compared to those of wells with no added sugar. Standard error was calculated from at least three replicates.

4.4. Computational docking

All calculations were performed using the Schrodinger Suite 2007 (Schrödinger Inc.). Ligands were generated using the Maestro 8.0 fragment directory. The peptides were drawn as the C-terminal amide and N-terminal acetate so as to more closely mimic a protein fusion. Additionally they were cyclized through disulfide bonds between their terminal disulfides. They were first soaked in water and then minimized (Impact 4.5). Partial charges were then assigned using the OPLS_2005 force field.

As ConA has several binding sites, two structures with ligands bound to different sites were selected for a comparison. The two protein data bank files that were used were 1CVN (trimannose bound to ConA)[31] and 1JUI (10-mer peptide MYWYPYASGS binding with ConA).[29] Two binding sites were defined as the locations of the trimannose and 10-mer peptide binding sites in the receptor. For the carbohydrate binding pocket, the amino acid residues involved are Ala207, Arg228, Asn14, Asp16, Asp208, Gly98, Gly224, Gly227, Leu99, Thr 226, Tyr12, and Tyr100. For the 10-mer peptide's binding site, the amino acid residues involved are Asn41, Asn44, Asp203, Gln43, His205, Met42, Pro23, Pro206, Ser21, Ser201, Ser204, and Tyr22. Two concentric boxes were defined for these two sites: the bounding box and the enclosing box. The bounding box contains the center of any acceptable ligand pose while the enclosing box contains all the ligand atoms of an acceptable pose. Energy grids were calculated and stored for each of these sites.

After the preparations for the ligands and receptors were conducted and the receptor grids determined, docking was carried out using the normal precision mode of Glide. The scale factor for van der Waals radii was applied to atoms with absolute partial charges less than or equal to 0.15 and 0.25 electrons for the ligands and the proteins, respectively. When each calculation was complete, many poses per ligand were generated, and the best docked structure was chosen using a Glidescore ('G-Score') function. A more negative G-Score value indicates stronger binding. This value is defined by the following relationship: G-Score = a^* vdW + b^* Coul + Hbond + Metal + BuryP + RotB + Site where $a = 0.065$; $b = 0.130$; vdW = van der Waals; Coul = Coulomb; Lipo = lipophilic; Hbond = hydrogen bonding; metal = metal binding; BuryP = penalty for buried polar groups; RotB = penalty for freezing rotable bonds; site = polar interaction in the active site; CvdW = Coul + vdW is considered the non-bonded interaction energy; Emodel describes the internal torsional energy of a specific conformer of the ligand and combines the Gscore, CvdW, and Intern.

Acknowledgments

We thank Alison Lester, Alyssa Mahaffey, and the members of the Orner lab for critical reading of the manuscript. L.Y. was supported by an SPMS graduate scholarship. E.Y.Y.M. and P.S.Y. thank the NTU/NAP exchange program. J.M. thanks the NTU/Southampton exchange program. P.T.P.T., Y.Y.W., and W.W.Q. thank the URECA program for undergraduate research, and J.W. thanks the CBC undergraduate summer research fund. This research was supported by an SPMS start up grant and a 'Singapore Ministry of Education Academic Research Fund Tier 1 Grant' (RG 53/06).

A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2009.03.054.

References and notes

1. Cao, Z. W.; Han, Y. L.; Zheng, C. J.; Ji, Z. L.; Chen, X.; Lin, H. H.; Chen, Y. Z. *Drug Discovery Today* **2005**, *10*, 521.
2. Chen, S.; Takanashi, S.; Zhang, Q.; Xiong, W.; Zhu, S.; Peters, E. C.; Ding, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10482.
3. Fang, Y. Q.; Wong, W. Q.; Yap, Y. W.; Orner, B. P. *Comb. Chem. High Throughput Screening* **2007**, *10*, 635.

4. Paulsen, H. *Angew. Chem., Int. Ed. Engl* **1982**, *21*, 155.
5. Nicolaou, K. C.; Mitchell, H. J. *Angew. Chem., Int. Ed.* **2001**, *40*, 1576.
6. Carmona, A. T.; Moreno-Vargas, A. J.; Robina, I. *Curr. Org. Synth.* **2008**, *5*, 33.
7. Seeberger, P. H.; Haase, W.-C. *Chem. Rev.* **2000**, *100*, 4349.
8. Seeberger, P. H.; Werz, D. B. *Nat. Rev. Drug. Discovery* **2005**, *4*, 751.
9. Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357.
10. Monzavi, B.; Cunto-Amesty, G.; Luo, P.; Kieber-Emmons, T. *Trends Biotechnol.* **2002**, *20*, 207.
11. Sears, P.; Wong, C-H. *Angew. Chem., Int. Ed.* **1999**, *38*, 2300.
12. DeMatteo, M. P.; Snyder, N. L.; Morton, M.; Baldisseri, D. M.; Hadad, C. M.; Peczuh, M. W. *J. Org. Chem.* **2005**, *70*, 24.
13. Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391.
14. *Phage Display of Peptides and Proteins: A Laboratory Manual*; Kay, B. K., Winter, J., McCafferty, J., Eds.; Academic Press: San Diego, 1996.
15. *Phage Display: A Laboratory Maual*; Barbas, C. F., Burton, D. R., Scott, J. K., Silverman, G. J., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2001.
16. Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. *Science* **1990**, *249*, 404.
17. Meyer, S. C.; Gaj, T.; Ghosh, I. *Chem. Biol. Drug. Des.* **2006**, *68*, 3.
18. Weber, P. C.; Pantoliano, M. W.; Thompson, L. D. *Biochemistry* **1992**, *31*, 9350.
19. Scott, J. K.; Longanathan, D.; Easley, R. B.; Gong, X. F.; Goldstein, I. J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5398.
20. Oldenburg, K. R.; Longanathan, D.; Goldstein, I. J.; Schultz, P. G.; Gallop, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5393.
21. Hoess, R.; Brinkmann, U.; Handel, T.; Pastan, I. *Gene* **1993**, *128*, 43.
22. Shickman, A. R.; Greenspan, N. S.; Cunningham, M. W. *J. Immol.* **1994**, *153*, 5593.
23. Martens, C. L.; Cwirla, S. E.; Lee, R. Y. W.; Whitehorn, E.; Chen, E. Y. F.; Bakker, A.; Martin, E. L.; Wagstrom, C.; Gopalan, P.; Smith, C. W.; Tate, E.; Koller, K. J.; Schatz, P. J.; Dower, W. J.; Barrett, R. W. *J. Biol. Chem.* **1995**, *270*, 21129.
24. Kaur, K. J.; Khurana, S.; Salunke, D. M. *J. Biol. Chem.* **1997**, *272*, 5539.
25. Vaughan, H. A.; Oldenburg, K. P.; Gallop, M. A.; Atkin, J. D.; McKenzie, I. F. C.; Sandrin, M. S. *Xenotranspl* **1996**, *3*, 18.
26. Miura, Y.; Sakaki, A.; Kamihira, M.; Iijima, S.; Kobayashi, K. *Biochim. Biophys. Acta* **2006**, *1760*, 883.
27. Apostolopoulos, V.; Sandrin, M. S.; McKenzie, I. F. C. *J. Mol. Med.* **1999**, *77*, 427.
28. Bachle, D.; Loers, G.; Guthohrlein, E. W.; Schachner, M.; Sewald, N. *Angew. Chem., Int. Ed.* **2006**, *45*, 6582.
29. Kaur, K. J.; Jain, D.; Goel, M.; Salunke, D. M. *Vaccine* **2001**, *19*, 3124.
30. Hossany, R. B.; Johnson, M. A.; Eniade, A. A.; Pinto, B. M. *Bioorg. Med. Chem.* **2004**, *12*, 3743.
31. Ladner, R. C.; Sato, A. K.; Gorzelany, J.; de Souza, M. D. *Drug Discovery Today* **2004**, *9*, 525.
32. Janin, D.; Kaur, K. J.; Salunke, D. M. *Biophys. J.* **2001**, *80*, 2912.
33. Jain, D.; Kaur, K. J.; Goel, M.; Salunke, D. M. *Biochem. Biopys. Res. Commun.* **2000**, *272*, 843.
34. Naismith, J. H.; Field, R. A. *J. Biol. Chem.* **1996**, *271*, 972.