J|A|C|S
COMMUNICATIONS

Published on Web 10/05/2004

Phagemid Encoded Small Molecules for High Throughput Screening of Chemical Libraries

Jun Yin,† Fei Liu,† Martina Schinke,‡ Christian Daly,‡ and Christopher T. Walsh*,†

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, and Bauer Center for Genomics Research, Harvard University, Cambridge, Massachusetts 02138

Received August 12, 2004; E-mail: christopher_walsh@hms.harvard.edu

High throughput screening of combinatorial chemical libraries holds tremendous promise for accelerating the process of drug discovery.[1] To increase the screen throughput, a number of strategies have been developed to encode synthetic molecules with DNA or RNA tags so that the whole library can be selected for target binding in one pool and the structures of the selected compounds can be quickly identified by sequencing the DNA or RNA tags.[2] In another strategy, the DNA "bar code" is not physically linked to the encoded small molecule but is embedded within the phage genomic DNA and encapsulated inside the phage capsid proteins to which the small molecules are attached.[3] The phage displayed small molecule library can then be selected for target binding, and upon amplification of the selected phages, the corresponding phagemid is sequenced to give the identity of the selected molecule. Using the phage platform for small molecule encoding has the advantages that the DNA bar code is shielded by the phage capsid proteins, which eliminates the interference of DNA secondary structures with the binding of small molecules to the target proteins, and at the same time the DNA bar code is protected from degradation so that the phage displayed small molecule libraries can be selected for binding with in vivo targets.[4] Also, since the amplification of the phage is spontaneous and highly efficient, the enrichment of the selected phages provides a sensitive readout for library screen.[3] The main drawback of the current method for small molecule phage display is that tens to hundreds of molecules are attached to each phage particle due to the nonspecific nature of the conjugation reaction. The high density of the small molecule ligands on a single phage particle may lead to high avidity of the phages with the target protein due to the multivalent binding of the small molecules carried by the same phage. This may give high background binding of low affinity ligands with the target protein during selection. Moreover, phages displaying high affinity ligands with the target protein may be impossible to elute due to the multivalent binding and thus cannot be amplified and identified during the selection process.

To circumvent these problems, here we report a strategy for attaching small molecules site specifically to the peptidyl carrier protein (PCP) monovalently displayed on M13 phage in order to achieve one-phage-one-compound display of the small molecules (Figure 1). PCPs are 10 kDa stably folded domains excised from nonribosomal peptide synthetases (NRPSs).[5] Sfp phosphopantetheinyl transferase[6a] was used to covalently link the small molecules to a specific serine residue of the PCP through a phosphopantetheinyl group using various coenzyme A (CoA)−small molecule conjugates as substrates (**1**−**5**).[7] We have shown that the selection of the small molecules monovalently displayed on the phage is highly efficient; typically, a single round of selection gives more than 2000-fold enrichment of the ligands with nM dissociation constants with the target protein. Furthermore, to accelerate the selection process, DNA microarrays were used to decode the identities of the selected small molecules.

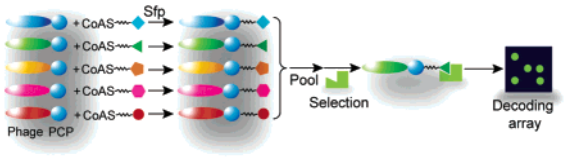

Figure 1. Schematic representation for displaying small molecules on the phage surface by Sfp catalyzed PCP modification with CoA−small molecule conjugates followed by selection for target binding and DNA microarray decoding.

Sfp showed the anticipated impressive substrate promiscuity with small molecule entities attached to CoA,[6] and in this study, molecules with diverse structures such as biotin (**1**), fluorescein (**2**), *N*-methylmesoporphyrin (NMP) (**3**), glutathione (**4**), and α-D-galactose (**5**) were conjugated to CoA (**6**), respectively, in one step with high yield through disulfide formation or thioether formation.[7] Thus, the synthesis of CoA conjugates for other small molecule libraries can be readily achievable once thio or maleimide functionalities are incorporated into the building blocks of the combinatorial library. The CoA−small molecule conjugates **1**−**5** were first tested as substrates of the Sfp catalyzed PCP modification reaction; the conversion of the free apo PCP domain to various small molecule labeled PCPs was complete within 30 min, as validated by matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) analysis.[7]

The PCP domain from NRPS module GrsA (80 amino acids)[8] was cloned in the pComb3H vector[9] for monovalent display of the PCP domain as N-terminal fusions to the phage capsid protein pIII. Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) of the PCP displayed phages and Western blot probed with anti-pIII antibody showed a band of PCP−pIII fusion proteins of the correct size.[7] CoA conjugates **1**−**5** were then attached to the phage by the Sfp catalyzed PCP−pIII modification reaction. SDS PAGE of the phage particles displaying biotin, fluorescein, and NMP followed by Western blots probed with streptavidin, anti-fluorescein antibody, and anti-NMP antibody, respectively, showed only one band corresponding to the size of PCP−pIII fusion, suggesting the exclusive attachment of the small molecules to the PCP−pIII fusion of the phage.[7] Thus, the stoichiometry of the small molecules displayed on each phage particle is controlled by the number of PCP−pIII fusions displayed on the phage and here monovalent display of the small molecules on the phage is achieved using the pComb3H system.

To construct DNA bar codes for small molecule encoding, 20-base-pair (bp) DNA fragments of unique sequences were cloned into the pComb3H vector immediately after the gene of PCP−pIII. The PCP displayed phages with unique bar code sequences were then covalently loaded with small molecules **1**−**5** by Sfp in

† Harvard Medical School.
‡ Harvard University.

10.1021/ja045127t CCC: $27.50 © 2004 American Chemical Society

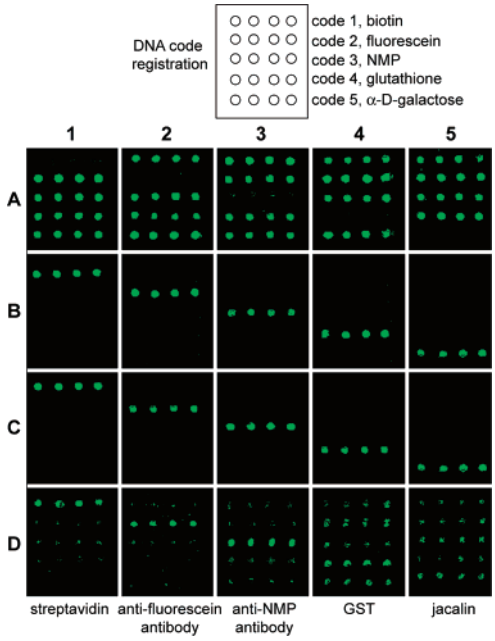

Figure 2. Evaluating the selection of small molecule displayed phage libraries by DNA decoding arrays. The DNA code registration panel shows the positions of the immobilized oligonucleotides on the array encoding each small molecule. (A) Hybridization pattern of the PCR products from the 1/4000 diluted phage pools before selection for binding with streptavidin (1), anti-fluorescein antibody (2), anti-NMP antibody (3), GST (4), and jacalin (5). (B−D) Hybridization patterns of the PCR products from the 1/40, 1/400, and 1/4000 diluted phage pools after the selection. The intensity of the brightest spot on each array was normalized for comparison.

separate tubes so that each small molecule was attached to the phage with a unique bar code for identification. To test the selection efficiency of various phage displayed small molecules with their binding targets, phage loading reactions with **1**−**5** were pooled at different ratios in order to dilute the phages displaying target binding small molecules with nonbinding phages 40-, 400-, and 4000-fold, respectively. The pooled phages with **1**−**5** displayed were then selected for binding with streptavidin, anti-fluorescein antibody, anti-NMP antibody 7G12,[10a] glutathione *S*-transferase (GST), and jacalin, respectively. After selection, phages bound to the target proteins were rescued by infecting *E. coli* cells and the bar code region of the phagemid DNA was amplified by polymerase chain reaction (PCR) using Cy3 labeled primers for subsequent DNA microarray analysis.

Decoding DNA microarrays were prepared by spotting the array surface with oligonucleotides complementary to the DNA bar code sequence amplified by the Cy3 labeled primer. Hybridization of the decoding arrays with PCR products from the selection gave a quick evaluation of the selection efficiency. As shown in Figure 2A, before the selection of the 1/4000 diluted pool, only the DNA sequences encoding nonbinding small molecules gave an intense hybridization signal due to the large excess of the nonbinding phage population in the pool over the target binding phage. However, after just one round of selection for binding with streptavidin, anti-fluorescein antibody, and anti-NMP antibody,[10a] respectively, the DNA sequence encoding biotin, fluorescein, and NMP gave the strongest hybridization intensity to the array, suggesting the specific enrichment of the phages displaying target binding small molecules (Figure 2D). Glutathione and α-D-galactose displayed phages showed less enrichment, as suggested by the weaker intensity for the hybridization of the DNA sequence encoding these molecules to the array after selection for GST or jacalin binding, respectively, presumably due to the lower affinity of these ligands to their target proteins.[10b,c] In comparison, when the target binding phages were diluted only 40- or 400-fold, the decoding array gave a clean hybridization signal of the DNA bar codes encoding the target binding small molecules, suggesting the predominant enrichment of the phages displaying these molecules after the selection (Figure 2B and C). The selection results were also confirmed by colony PCR[11] using DNA primers that bind the bar code sequence.[7] Thus, both the decoding array and the colony PCR experiments suggested that one round of selection of the phage displayed small molecules typically gave more than 2000-fold enrichment of the molecules with nM affinity to the target protein.

In summary, we have developed an efficient method for monovalently displaying small molecules on the phage surface through Sfp catalyzed site specific modification of the phage displayed PCP. CoA−small molecule conjugates as the substrates of Sfp can be easily synthesized by facile disulfide, thioether, or thioester formation with the small molecules. The substrate promiscuity of Sfp with the molecular entities linked to CoA is especially suitable for conjugating a library of small molecules to the phage displayed PCP. Both the conjugation of small molecules to CoA and the subsequent attachment of small molecules to the phage can be carried out in a parallel fashion for high throughput processing of large libraries. The small molecule library monovalently displayed on the phage surface can then be selected in one pool for target binding and achieve more than 2000-fold enrichment in a single round with the small molecule ligands showing nM affinity to the target. The use of DNA decoding arrays for identifying the selected molecules further accelerates the process for screening the phage encoded small molecule libraries.

Acknowledgment. This work was funded by NIH Grant GM2011 (C.T.W.). We thank Peter G. Schultz for providing pComb3H and 7G12 Fab, and Matthew Kanan and Claire Bailey for suggestions.

Supporting Information Available: Structures of CoA conjugates, Western blots, results of MALDI MS and colony PCR, and the experimental procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411. (b) Gray, N. S.; et al. *Science* **1998**, *281*, 533. (c) Shogren-Knaak, M. A.; Alaimo, P. J.; Shokat, K. M. *Annu. Rev. Cell Dev. Biol.* **2001**, *17*, 405. (d) Kuruvilla, F. G.; et al. *Nature* **2002**, *416*, 653. (e) Lin, H.; Cornish, V. W. *Angew Chem., Int. Ed.* **2002**, *41*, 4402. (f) Boger, D. L.; Desharnais, J.; Capps, K. *Angew Chem., Int. Ed.* **2003**, *42*, 4138. (g) Alluri, P. G.; Reddy, M. M.; Bachhawat-Sikder, K.; Olivos, H. J.; Kodadek, T. *J. Am. Chem. Soc.* **2003**, *125*, 13995. (h) Speers, A. E.; Cravatt, B. F. *ChemBioChem* **2004**, *5*, 41.

(2) (a) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381. (b) Frankel, A.; Millward, S. W.; Roberts, R. W. *Chem. Biol.* **2003**, *10*, 1043. (c) Halpin, D. R.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, E174. (d) Li, X.; Gartner, Z. J.; Tse, B. N.; Liu, D. R. *J. Am. Chem. Soc.* **2004**, *126*, 5090. (e) Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D. *Nat. Biotechnol.* **2004**, *22*, 568.

(3) (a) Woiwode, T. F.; et al. *Chem. Biol.* **2003**, *10*, 847. (b) Fellouse, F.; Deshayes, K. *Chem. Biol.* **2003**, *10*, 783.

(4) (a) Arap, W.; et al. *Nat. Med.* **2002**, *8*, 121. (b) Essler, M.; Ruoslahti, E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2252. (c) Laakkonen, P.; Porkka, K.; Hoffman, J. A.; Ruoslahti, E. *Nat. Med.* **2002**, *8*, 751. (d) Kolonin, M. G.; Saha, P. K.; Chan, L.; Pasqualini, R.; Arap, W. *Nat. Med.* **2004**, *10*, 625.

(5) (a)Walsh, C. T.; et al. *Curr. Opin. Chem. Biol.* **1997**, *1*, 309. (b) Weber, T.; et al. *Structure Fold Des.* **2000**, *8*, 407.

(6) (a) Lambalot, R. H.; et al. *Chem. Biol.* **1996**, *3*, 923. (b) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. *Science* **1999**, *284*, 486. (c) La Clair, J. J.; et al. *Chem. Biol.* **2004**, *11*, 195. (d) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 7754.

(7) See Figures S1, S2 and Tables S1−S3, Supporting Information.

(8) Stachelhaus, T.; Walsh, C. T. *Biochemistry* **2000**, *39*, 5775.

(9) Barbas, C. F., III; Kang, A. S.; Lerner, R. A.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7978.

(10) (a) Yin, J.; Andryski, S. E.; Beuscher, A. E. T.; Stevens, R. C.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 856. (b) Gustafsson, A.; Pettersson, P. L.; Grehn, L.; Jemth, P.; Mannervik, B. *Biochemistry* **2001**, *40*, 15835. (c) Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 6140.

(11) Yin, J.; Mills, J. H.; Schultz, P. G. *J. Am. Chem. Soc.* **2004**, *126*, 3006.

JA045127T