Combinatorial Chemistry

Dynamic Combinatorial Chemistry: Lysozyme Selects an Aromatic Motif That Mimics a Carbohydrate Residue**

Sandrine Zameo, Boris Vauzeilles, and Jean-Marie Beau**

Dedicated to Professor Julius Rebek, Jr. on the occasion of his 60th birthday

In a series of recent developments, dynamic combinatorial chemistry (DCC)[1] has proved its capacity to select potent enzyme inhibitors.[2] The success of these preliminary validation studies was mainly based on the use of a previously known high-affinity scaffold[2a,c,d,f] that could reversibly self-assemble with several complementary building blocks. This process results in the formation of a dynamic library of potential ligands in which the target of interest, which acts as a thermodynamic trap, can select its best binder. When the experimental conditions are properly chosen,[3] binding leads to an amplification of this particular compound which can be detected by an appropriate analytical method such as HPLC.

In these early studies, the products that were selected bore very close similarities, both in structure and affinity, to previously known inhibitors of the target enzymes. We were curious to challenge this process with an enzyme for which a scaffold with relatively poor binding properties would hopefully direct members of the dynamic combinatorial library (DCL) towards the active site. Such situations are frequently found in glycobiology, an area of research that is rapidly

[*] S. Zameo, Dr. B. Vauzeilles, Prof. J.-M. Beau
Université Paris-Sud
Laboratoire de Synthèse de Biomolécules associé au CNRS
Institut de Chimie Moléculaire et des Matériaux
91405 Orsay Cedex (France)
Fax: (+33) 1-6985-3715
E-mail: jmbeau@icmo.u-psud.fr

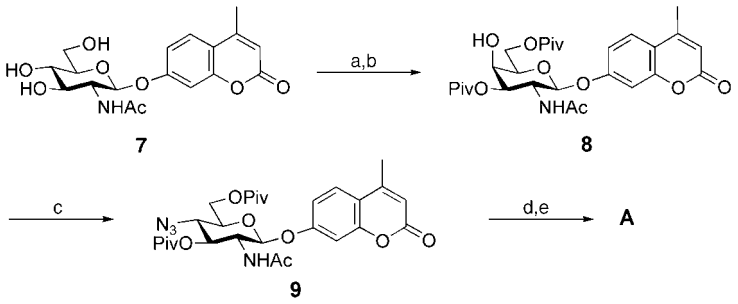
[**] We gratefully acknowledge the Conseil Général de l'Essonne for financial support of this study and GlaxoSmithKline for their interest.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200462150

becoming fundamental in the post-genomic era.[4] Indeed, many carbohydrate-binding proteins including enzymes (termed Group II proteins)[5] comprise shallow carbohydrate-binding sites that result in relatively weak carbohydrate–protein interactions with dissociation constants often in the millimolar range. We therefore decided to assess the capacity of DCC to probe such low energy interactions and selected for an initial proof-of-principle study a well-known and readily available Group II glycosidase, namely hen egg-white lysozyme (HEWL). Involved in peptidoglycan degradation, HEWL is known to cleave *N*-acetyl-glucosamine oligomers (chito-oligosaccharides) into their smaller units, which, up to the tetramer, behave as competitive inhibitors.[6] The affinity of *N*-acetyl-D-glucosamine (D-GlcNAc) for the HEWL-binding site lies in the 20–50 mM range, whereas chitobiose and chitotriose display higher affinities.[7] On the basis of this knowledge, we designed a DCL of potential HEWL binders starting with a D-GlcNAc (or D-Glc) motif as a scaffold and by using the generation of imines to introduce diversity (Scheme 1). This reversible connecting system, which functions under mild conditions, has previously been reported to be suitable for the formation of DCLs.[2a,c,d,8]

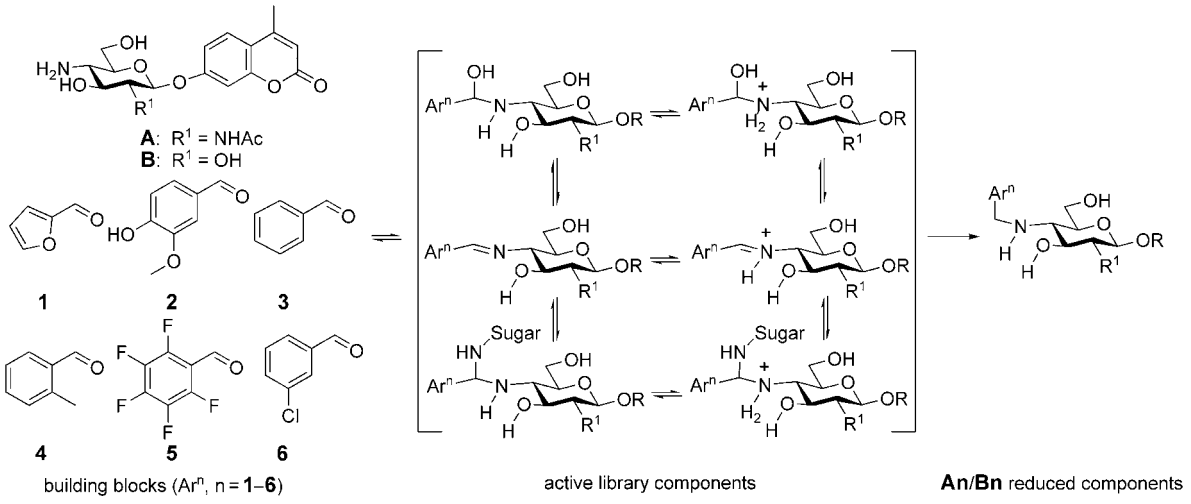
Two amino-derived carbohydrate compounds (amines **A** and **B** in the D-GlcNAc and D-Glc series, respectively) and six differently substituted aromatic aldehydes (**1**–**6**) were selected.[9] A 4-methylumbelliferyl chromophore was introduced at the anomeric position of the carbohydrate building blocks **A** and **B** to allow equal HPLC detection of the different library products at a specific near-UV wavelength ($\lambda = 322$ nm). The aromatic motifs may "mimic", after imine formation, a second carbohydrate ring and potentially interact with complementary aromatic residues of the binding site. Stacking interactions between aromatic residues of the amino acid side chains and the hydrophobic faces of carbohydrates are a common feature of almost every carbohydrate–protein complex studied until present.[5a,10]



Scheme 2. Synthesis of carbohydrate scaffold **A**. a) PivCl, pyridine, CH_2Cl_2, 0 °C, 2 h (78%); b) 1) Tf_2O, pyridine, CH_2Cl_2, −15 °C, 1 h; 2) $NaNO_2$, DMF, 20 °C, 16 h; 3) $NaHCO_3$, H_2O, 20 °C, 5 min (64%); c) 1) Tf_2O, pyridine, CH_2Cl_2, −15 °C, 1 h; 2) NaN_3, Bu_4NHSO_4, DMF, 20 °C, 16 h (61%); d) MeONa, MeOH, 20 °C, 5 d (84%); e) H_2, Pd/C (5%), MeOH, 20 °C, 4 h (90%). Piv = pivaloyl, Tf = trifluoromethylsulfonate, DMF = *N*,*N*-dimethylformamide.

Amine **A** was synthesized from *N*-acetyl-D-glucosamine following the sequence presented in Scheme 2. The 4-methylumbelliferyl glycoside **7** was obtained by Roy's procedure[11] and by using Horton's chloride.[12] Regioselective pivaloylation of hydroxyl groups at the 3 and 6 positions,[13] followed by epimerization at C-4[14] furnished the D-galacto derivative **8**. Triflate formation and nucleophilic substitution by sodium azide provided the azido intermediate **9**, which after deprotection of the hydroxyl groups and catalytic hydrogenation afforded amine **A**. A similar sequence was applied to the preparation of amine **B** from D-galactose (see Supporting Information).

When a mixture of amines **A** and **B** (0.4 mM) were equilibrated with aldehydes **1**–**6** (0.4 mM each) at room temperature in an aqueous phosphate buffer solution (pH 6.2) in the presence of sodium cyanoborohydride (3.6 mM), slow formation of the amines through reduction of the iminium species was observed. After 24 h, the twelve expected products could be detected by reverse-phase HPLC (Figure 1a). Equilibration of the same library in the presence of HEWL (0.4 mM) induced a detectable change in the distribution of the amines in the mixture with enrichment of



Scheme 1. Structures of the building blocks and components of the dynamic combinatorial library (DCL).

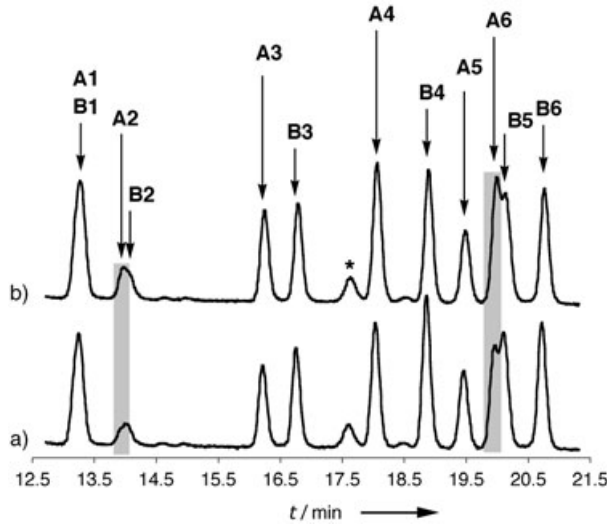


Figure 1. Selected portion of HPLC chromatograms of the DCL made from amines **A** and **B**, and aldehydes **1**–**6** over 24 h: a) in the absence of HEWL and b) in the presence of HEWL (1 equiv). (* indicates residual 4-methylumbelliferone).

amines **A2** and, to a lesser extent, **A6** (resulting from reductive alkylation of amine **A** with aldehydes **2** and **6**, respectively; see Figure 1b). Although HPLC–MS analysis easily confirmed the structures, amplification of the active components becomes difficult to evaluate when a strong overlap occurs (see, for instance, amines **A1** and **B1** in Figure 1), and this difficulty increases with the size of the library.[9,15] We therefore constructed and analyzed sublibraries with amines **A** and **B**. As expected from the previous experiments, amplification of amines **A2** and **A6** was observed when comparing libraries synthesized from **A** in the presence (Figure 2b) and in the absence (Figure 2a) of HEWL. No amplification was detected in the presence of HEWL in the sublibrary that was based on amine **B** (see Supporting Information). This clearly indicates that the enzyme selects amine **A** as a scaffold rather than amine **B**. When the reaction was performed at a higher concentration of HEWL (1.2 mM), the amplification was significantly enhanced (Figure 2c). Also, the addition of chitotriose, a good HEWL inhibitor, led to a reduction (0.4 mM chitotriose, not shown) or disappearance (1.2 mM chitotriose, Figure 2d) of the amplification effect.

All these simple experiments clearly indicate an amplification of the library members selected through interaction with the HEWL active site. Interaction of scaffold **A** with the active site rather than scaffold **B** further indicates that the 2-acetamido group is an essential feature for recognition of the library members by HEWL.[7b,16] The active site of the enzyme is a cleft composed of six carbohydrate-binding subsites (A to F[17] or −4 to +2[18]). Subsites −4, −2, and +1 are D-GlcNAc-binding sites which are sensitive to the acetamido substituent by strict hydrogen-bonding and hydrophobic interactions, whereas the three other sites are not.[17] As chitotriose is known to tightly bind to the −4 to −2 subsites,[17] the extra stabilization of the precursors of amines **A2** and **A6** above

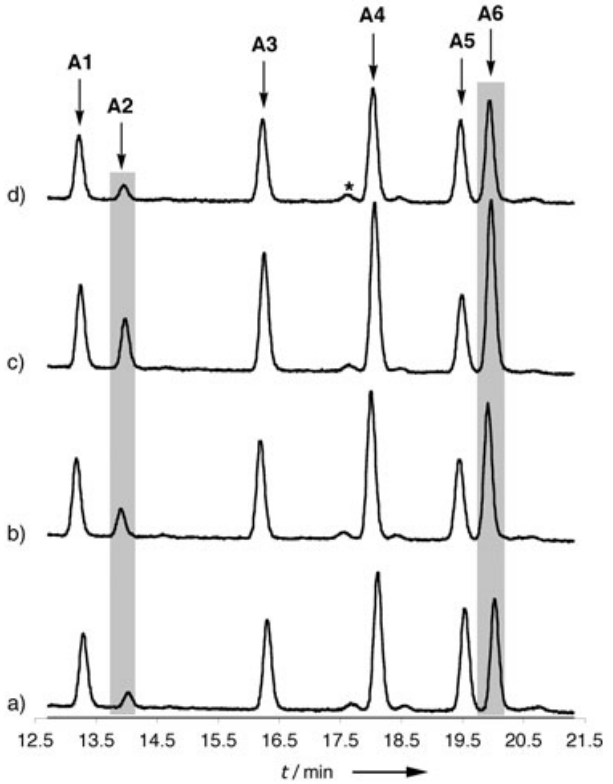
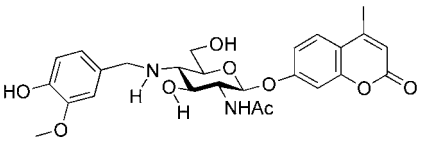


Figure 2. Selected portion of HPLC chromatograms showing the products of the DCL prepared from amine **A** and aldehydes **1**–**6** over 24 h: a) in the absence of HEWL, b) in the presence of 1 equivalent of HEWL, c) in the presence of 3 equivalents of HEWL, and d) in the presence of 1 equivalent of HEWL and 3 equivalents of chitotriose. (* indicates residual 4-methylumbelliferone).

suggests that the D-GlcNAc scaffold may occupy subsite −2 with the selected aromatic motif occupying subsite −3.

If this reasoning is correct, amines **A2** or **A6** may display inhibitory activities which are comparable to that of chitobiose.[19] To test this hypothesis, the most amplified compound (amine **A2**, Scheme 3) was synthesized separately,[20] and its



A2

Scheme 3. Amine **A2**, which results from reductive alkylation of amine **A** with aldehyde **2**.

inhibitory activity was measured and compared with the chito-oligomers. We used the lysis rate of *Micrococcus lysodeikticus*, a Gram-positive bacteria that is very sensitive to HEWL,[21] as a measure of the enzyme activity. Lineweaver–Burk analysis is presented in Figure 3. The inhibition constants (K_i) measured for *N*-acetyl-D-glucosamine and chitotriose were the same as those reported (not shown).[7] The value found for chitobiose was in the range reported (0.6 mM instead of 0.2 mM reported, see Supporting Informa-

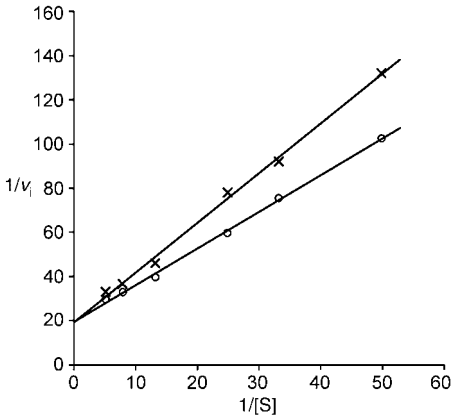


Figure 3. Lineweaver–Burk analysis (double reciprocal plot) of the observed initial lysis rate (V_i) as a function of the concentration of *Micrococcus lysodeikticus* ([S]) in the absence (○) and in the presence (×) of amine **A2** (214 μM).

tion). The inhibition assays carried out with amine **A2** (Figure 3) yielded a competitive type of inhibition with a K_i value that was similar to the value observed for chitobiose ($K_i = 0.6$ mM).

These results indicate an additional binding effect from the aromatic motifs, with amine **A2** being around 100-fold more active than the starting *N*-acetyl-D-glucosamine. With the peptidoglycan substrate, HEWL subsite −3 hosts a carbohydrate moiety (a *N*-acetylmuramic acid unit) which is known to interact with complementary aromatic residues through stacking interactions.[17] The K_i values, which are similar for **A2** and chitobiose, suggest that the enzyme selects from the DCL an aromatic motif with a precise substitution pattern that best fits in subsite −3.

We have shown that hen egg-white lysozyme selects from a dynamic library of potential active-site ligands an optimal binder, which comprises an aryl group that mimics a carbohydrate unit. In such a dynamic combinatorial assay, the system has shown its capacity to discriminate subtle affinity variations induced by the different substitutions on the aromatic moiety of the aldehydes. Further structural studies to confirm this hypothesis and experiments to develop and extend this approach are currently in progress in our group.

Experimental Section

Preparation of a DCL from amines **A** and **B** and aldehydes **1**–**6**: Stock solutions of amines **A** (12 mM) and **B** (2.4 mM) were prepared in distilled water. Solutions of aldehydes (0.6 mM) in phosphate buffer (30 mM, pH 6.2) and sodium cyanoborohydride (36 mM) in distilled water were prepared just prior to use. Solutions of amine **A** (16.7 μl), amine **B** (83.5 μl), and the aldehydes (333 μl) were introduced into an Eppendorf tube, which was equipped with a small magnetic stirrer bar, and the volume was adjusted to 450 μl with distilled water. After stirring, a 225-μl aliquot of the mixture was removed and introduced into an Eppendorf tube that contained HEWL (1.5 mg, 1 equiv). Both Eppendorf tubes were stirred for 30 min before solutions of sodium cyanoborohydride (25 μl) were added. The resulting mixtures (250 μl) were equilibrated at room temperature for 24 h. 20-μl aliquots were removed from these solutions and guanidine thiocyanate (4.7 mg), a good denaturant of HEWL, was added prior to HPLC analysis to ensure the release of any bound ligand from the active site. This treatment was also applied in the control experiment without HEWL. After 1 h, acetic acid (20 μl) was added, and the solution was diluted with methanol (120 μl). Analytical chromatography was performed on a JASCO LC-1500 system equipped with a Phenomenex LUNA C18 (2) 5μ reversed-phase HPLC column (150 × 4.60 mm) with UV/Vis detection at $\lambda = 322$ nm. A binary solvent gradient (solvent A: 0.1% trifluoroacetic acid in 95:5 H_2O/CH_3CN, solvent B: 0.1% trifluoroacetic acid in 95:5 CH_3CN/H_2O) was optimized to separate most of the DCL compounds: from 90:10 A/B to 70:30 A/B over 20 min, with a flow rate of 0.8 mL min^{-1}. HPLC peaks were assigned by means of LC–MS analysis. Some products were synthesized and characterized separately to validate their assignments on the chromatograms.

Received: September 29, 2004
Published online: January 11, 2005

Keywords: carbohydrates · combinatorial chemistry · enzyme inhibitors · molecular recognition · reduction

[1] J.-M. Lehn, *Chem. Eur. J.* **1999**, *5*, 2455; I. Huc, R. Nguyen, *Comb. Chem. High Throughput Screening* **2001**, *4*, 109; S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders, J. F. Stoddart, *Angew. Chem.* **2002**, *114*, 938; *Angew. Chem. Int. Ed.* **2002**, *41*, 898.

[2] a) I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2106; b) T. Bunyapaiboonsri, O. Ramström, S. Lohman, J.-M. Lehn, L. Peng, M. Goeldner, *ChemBioChem* **2001**, *2*, 438; c) M. Hochgürtel, H. Kroth, D. Piecha, M. W. Hofmann, C. Nicolau, S. Krause, O. Schaaf, G. Sonnenmoser, A. V. Eliseev, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3382; d) M. Hochgürtel, R. Biesinger, H. Kroth, D. Piecha, M. W. Hofmann, S. Krause, O. Schaaf, C. Nicolau, A. V. Eliseev, *J. Med. Chem.* **2003**, *46*, 356; e) T. Bunyapaiboonsri, H. Ramström, O. Ramström, J. Haiech, J.-M. Lehn, *J. Med. Chem.* **2003**, *46*, 5803; f) M. S. Congreve, D. J. Davis, L. Devine, C. Granata, M. O'Reilly, P. G. Wyatt, H. Jhoti, *Angew. Chem.* **2003**, *115*, 4617; *Angew. Chem. Int. Ed.* **2003**, *42*, 4479.

[3] Z. Grote, R. Scopelliti, K. Severin, *Angew. Chem.* **2003**, *115*, 3951; *Angew. Chem. Int. Ed.* **2003**, *42*, 3821; P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* **2004**, *10*, 3139.

[4] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373; R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683; C. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357; P. M. Rudd, R. A. Dwek, *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 1; H. J. Gabius, *Biochim. Biophys. Acta* **2002**, *1572*, 165; H. Kogelberg, D. Solis, J. Jiménez-Barbero, *Curr. Opin. Struct. Biol.* **2003**, *13*, 646.

[5] a) F. A. Quiocho, N. K. Vyas in *Bioorganic Chemistry: Carbohydrates* (Ed.: S. M. Hecht), Oxford University Press, Oxford, **1999**, pp. 441–457; b) P. Sears, C.-H. Wong, *Angew. Chem.* **1999**, *111*, 2446; *Angew. Chem. Int. Ed.* **1999**, *38*, 2300.

[6] J. A. Rupley, *Proc. R. Soc. London, Ser. B* **1967**, *167*, 416.

[7] a) F. W. Dahlquist, L. Jao, M. Raftery, *Proc. Natl. Acad. Sci. USA* **1966**, *63*, 26; b) D. M. Chipman, N. Sharon, *Science* **1969**, *165*, 454.

[8] H. Li, P. Williams, J. Micklefield, J. M. Gardiner, G. Stephens, *Tetrahedron* **2004**, *60* , 753.

[9] For clarity, only six out of the thirty-six aromatic aldehydes that were tested are shown in this work (see Supporting Information).

[10] C.-H. Wong, *Acc. Chem. Res.* **1999**, *32*, 376.

[11] D. Carrière, S. J. Meunier, F. D. Tropper, S. Cao, R. Roy, *J. Mol. Catal. A* **2000**, *154*, 9.

[12] D. Horton, *Org. Synth.* **1966**, *46*, 1.
[13] L. Lay, F. Nicotra, L. Panza, G. Russo, E. Adobati, *Helv. Chim. Acta* **1994**, *77*, 509.
[14] F. Santoyo-Gonzalez, C. Uriel, J. A. Calvo-Asin, *Synthesis* **1998**, 1787.
[15] Out of thirty-six aromatic aldehydes tested (see Supporting Information), the product derived from aldehyde **2** was the most amplified by the enzyme.
[16] A. Neuberger, B. M. Wilson, *Nature* **1967**, *215*, 524.
[17] C. C. F. Blake, D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips, V. R. Sarma, *Nature* **1965**, *206*, 757; L. N. Johnson, D. C. Phillips, *Nature* **1965**, *206*, 761; C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips, V. R. Sarma, *Proc. R. Soc. London Ser. B* **1967**, *167*, 378; K. Maenaka, M. Matsushima, H. Song, F. Sunada, K. Watanabe, I. Kumagai, *J. Mol. Biol.* **1995**, *247*, 281.
[18] G. J. Davies, K. S. Wilson, B. Henrissat, *Biochem. J.* **1997**, *321*, 557; D. J. Vocadlo, G. J. Davies, R. Laine, S. G. Withers, *Nature* **2001**, *412*, 835.
[19] It should, however, be noted that the true binders are one of the two hemiaminals, the aminal, the imine, or their protonated species and that the corresponding reduced compound may have much poorer binding properties.
[20] The imine was preformed in methanol (amine **A**, aldehyde **2**, MeOH, 3 Å molecular sieves, 48 h) and then reduced ($NaBH_4$, 16 h, 83%).
[21] D. Shugar, *Biochim. Biophys. Acta* **1952**, *8*, 302; A. L. N. Prasad, G. Litwack, *Anal. Biochem.* **1963**, *6*, 328.

