

# NIH Public Access

**Author Manuscript**

*Chemistry*. Author manuscript; available in PMC 2014 December 16.

### Published in final edited form as:

*Chemistry*. 2013 December 16; 19(51): 17291–17295. doi:10.1002/chem.201303848.

# **Directed Evolution of 2G12-Targeted Nonamannose Glycoclusters by SELMA**

**J. Sebastian Temme**, **Michael Drzyzga**, **Iain S. MacPherson**, and **Isaac J. Krauss**\* Department of Chemistry, Brandeis University, 415 South St., Waltham, MA 02454 (USA)

# **Abstract**

**Optimizing multivalency:** Clusters of Man<sub>9</sub> glycans which are recognized by broadlyneutralizing anti-HIV antibody 2G12 have potential as HIV vaccines. However, optimal recognition by 2G12 requires optimal clustering of glycans. Using our recently described SELMA technique (SELection with Modified Aptamers), we have developed Man<sub>9</sub> clusters in which glycans are supported by DNA sequences selected from among  $2 \times 10^{13}$  variants.

#### **Keywords**

carbohydrates; cluster effect; directed evolution; glycoconjugates; multivalency

Carbohydrate-protein interactions mediate numerous important biological events, such as cell signalling, recognition and adhesion, and host-pathogen recognition.[1] The affinity of individual glycans for individual protein binding sites is normally very low, with  $K_d$ 's in the mM range. High avidity binding is achieved by multivalency, wherein multiple glycans within a cluster bind cooperatively to multiple sites in a single protein or in several clustered proteins.[2] This concept has been extensively applied to the design of synthetic glycan clusters that bind to interesting biological targets.[3]

However, the number of glycans (valency) is not the only important feature in multivalent ligand design. For applications in which the carbohydrate binding sites are separated by a fixed distance, avidity is also greatly affected by the spacing and orientation of glycans within the cluster. Although glycans attached by long flexible linkers can easily span target binding sites, the resulting avidity is suboptimal due to entropic costs, as the unbound ligand has many more possible conformations than the bound one.<sup>[2b]</sup> Moreover, flexible glycan clusters may not be specific for a single target. Design of more conformationally-defined glycocluster ligands is desirable due to the potential to achieve high avidity and specificity, but is challenging, as small deviations from ideal glycan arrangement may completely prevent the desired binding interaction. For these rigid glycoclusters, rational design of the optimal structure may require extensive trial and error as well as structural insight.

As an alternative to rational glycocluster design, we have developed SELMA (SELection with Modified Aptamers, Scheme 1).<sup>[4]</sup> This technique enables evolution of DNA aptamers bearing large covalent modifications such as oligosaccharides.<sup>[5]</sup> The process begins with a library template of synthetic DNA hairpins, containing a single-stranded random sequence region and a 3'-terminal self-priming patch of self-complementary sequence. Polymerase extension with EdU (alkynyl U) substituted for T installs alkynes at every position in which

<sup>\*</sup>Fax: (+) 1-781-736-2516 kraussi@brandeis.edu Homepage: [http://www.brandeis.edu/departments/chemistry/faculty/krauss.html.](http://www.brandeis.edu/departments/chemistry/faculty/krauss.html) Supporting information for this article is available on the WWW under<http://www.angewandte.org> or from the

Temme et al. Page 2

the single stranded template contains an A, and then CuAAAC chemistry<sup>[6]</sup> attaches glycans (or potentially other large modifications) to each of these alkynes. The glycosylated strand is then displaced by extension of a small primer inside the hairpin loop, using all-natural dNTPs. After this step, the single-stranded glycosylated DNA can adopt a sequencedependent fold, whereas the covalently-attached dsDNA contains a copy of the same sequence, though with all natural bases. This "genotype" sequence can readily be amplified by PCR if the "phenotype" portion enables survival during selection. In our original report, we applied SELMA to the evolution of decavalent clusters of Man<sub>4</sub> oligosaccharides, which bound to HIV broadly-neutralizing antibody 2G12 with  $K_d$ 's of ~250–500 nM. In the present paper, we describe further developments of this method to afford Man<sub>9</sub> glycoclusters of higher avidity, smaller size, and lower valency.

 $2G12$  is a target of interest because it neutralizes a broad range of HIV strains<sup>[7]</sup> and confers sterilizing immunity in rhesus macaque models of HIV infection.[8] As 2G12 binds to several Man<sub>9</sub> glycans on the HIV envelope protein gp120,<sup>[9]</sup> there has been considerable effort to design Man<sub>9</sub> glycoclusters which mimic the 2G12 epitope and might re-elicit  $2G12$ -like antibodies.<sup>[10]</sup> Immunogenicity studies with Man<sub>9</sub> glycoclusters designed so far have been unsuccessful at eliciting a 2G12-like antibody response, so design of new constructs which might mimic the true Man<sub>9</sub> presentation on gp120 are desirable.

For the present study (Figure 1), we designed our SELMA selection to differ from the original study<sup>[4]</sup> in 2 respects: 1) glycoclusters would contain the full Man<sub>9</sub> oligosaccharide rather than the Man<sub>4</sub> fragment which was used previously, [11] and 2) we would begin selection with libraries biased to contain species with fewer glycans. As depicted in Scheme 1, the positions of glycans in library sequences are dictated their CuAAAC attachment to "alkynyl-U" (EdU) bases, which are in turn located opposite A bases in the template strand.<sup>[6]</sup> Thus, to create starting libraries with less glycosylation, we made two modifications which reduce the number of A's in the original library template sequences (Figure 1a). Whereas the original library contained two fixed A's in a primer region and 25% A content in the random sequence region, our two new libraries contained no A's in the constant primer regions, and a lower percentage of A's (7% or 15%) in the random sequence region. Figure 1b shows the statistically predicted percentages of different valencies within each starting library; whereas species containing 6–10 glycans were prevalent in the original library, the new starting libraries contain larger percentages of species with fewer glycans.

To enable selection with Man<sub>9</sub> oligosaccharides, we prepared Man<sub>9</sub> azide by the convergent block assembly method shown in Scheme 2. Man4 tetrasaccharide **1**, containing a reducingterminal β-mannose and cyclohexyl linker was prepared as in our previous study.<sup>[4]</sup> Deprotection of the benzylidene 6-hydroxyl, followed by Sinay glycosylation<sup>[12]</sup> with Man<sub>5</sub> donor  $2^{[13]}$  and partial deprotection afforded Man<sub>9</sub> derivative 3 in 52 % yield over three steps. Global deprotection,[14] followed by Wong's azidation protocol,[15] afforded **4** in 63 % yield (nearly 100 mg) over two steps.

With Man<sub>9</sub> derivative 4 in hand, we then initiated SELMA selections with 40 pmol each (2)  $\times$  10<sup>13</sup> sequences) of the 7 % and 15 % libraries shown in Figure 1. Briefly, libraries generated as outlined in Scheme 1 (see Supporting Information for details) were incubated with mAb 2G12, and the bound complexes were captured with protein A magnetic beads. A counterselection with only protein A beads was run in rounds 2, 4 and 6 to remove potential protein A binders from the library. Stringency was increased by decreasing 2G12 concentration from 50 nM in rounds  $1-4$  to 10, 5 and 1.4 nM in rounds 5, 6 and 7, respectively. Library enrichment was monitored by observing the decrease in the number of PCR cycles required to regenerate the library after selection. After seven rounds, enrichment

levelled off and the libraries were cloned, yielding 31 unique sequences (among 93 sequence reads, see SI Table 3).

These sequences contained from 3 to 12 glycosylation sites, with an average of 7.9, as compared to 10.5 in the sequences obtained from our original 25 % glycosylated library selection. Filter binding studies on nine representative sequences (Table 1) showed that these glycoclusters were recognized by 2G12 with  $K_d$ 's ranging from 310 nM to > 1 $\mu$ M. Although these sequences exhibited moderate reductions in  $K_d$  and valency compared with those from our original library (compare with entry 9), CLUSTAL sequence alignment<sup>[16]</sup> (SI Table 3) revealed several interesting features. Most sequences resulting from the new selections contained a consensus motif (Table 1, in yellow), and strikingly, this sequence was present in clones from both libraries. Moreover, the motif was located at various positions, either early or late, within different sequences. This suggests that the multiple clones containing this motif resulted from convergent evolution, and were not derived from mutations of a single clone or cross-contamination between libraries.

The existence of the consensus motif in most selection winners suggested this region as a common site of 2G12 recognition. To investigate this further, we performed Mfold secondary structure analyses<sup>[17]</sup> on representative sequences (see Supporting Information). As Mfold does not take into account the potential major effects of our large unnatural base modifications we cautiously regarded the computational output only as a starting point to suggest further experiments. For most sequences, Mfold predicted most stable secondary structures in which the "consensus" motif was present at the end of a stem-loop region. In the case of clone 7A8-8 from the 7 % glycosylated library, we noted that a large region of sequence outside the consensus motif (Figure 2a, regions outside the red boxes) was not predicted to have any predominant structure. We thus prepared truncated versions of 7A8-8, removing regions at the 3' and 5' ends. In contrast to our previous work with the 25 % glycosylated library, we found in this case that large truncations were not only tolerated, but beneficial.

Truncation at the 5' end (7A-8-T1, Figure 2b, entry 2) improved 2G12 recognition slightly, from 320 to 270 nM, even though the number of glycosylation sites decreased from 7 to 5. Additional truncation at the 3' end (7A-8-T2, entry 3) further improved binding to 150 nM. Although removal of two more bases from the 3'end appeared to be allowable on the basis of the M-fold structure, this truncation (7A-8-T3, entry 4) eliminated recognition by 2G12, showing that glycosylation at this 5<sup>th</sup> site in the 7A-8-T2 sequence is important. As expected, binding of these more extreme 3' truncations was not rescued by lesser truncations at the 5' end (entries 5 and 8). Finally, truncation *only* at the 3' end resulted in  $K_d$ 's similar to truncation *only* at the 5' end (entries 6 and 7).

We were interested in whether all five of the glycosylation sites in 7A-8-T2 were necessary for 2G12 recognition, or whether further reductions in valency would be tolerated. To this end, we prepared a series of (S->C) mutants, corresponding to deletion of the glycan at each of the five remaining positions (Figure 2b, entries 9–13). We were surprised to find that deletion of glycan at any of these positions resulted in major loss of binding to 2G12. Based on published 2G12 structural studies, it is unlikely that all five glycans bind to 2G12 simultaneously. However, each glycan likely plays a major structural role, considering that glycans comprise 50% of the molecular weight of this small aptamer.

In conclusion, we have found that SELMA selection of Man9-cluster libraries biased toward low carbohydrate valency leads to selection winners with significantly lower valency than our previous library and modestly better recognition by  $2G12$ .<sup>[4]</sup> Interestingly, these selection winners contained a conserved 2G12 recognition motif; moreover, it was found

that the region outside this motif could be significantly truncated, which actually led to improved 2G12 recognition with fewer glycans. Compared with monovalent Man<sub>9</sub> glycan ( $K_d$  = 180 µM), these pentavalent glycoclusters exhibit up to a ~1000-fold binding enhancement ( $K_d = 150$  nM), though they are still 30-fold less antigenic than Wong's nonavalent Man<sub>9</sub> dendrimer.<sup>[10i, 18]</sup> Our truncated oligo, 7A8-8-T2, is relatively short and should be amenable to further structural study by NMR and X-ray crystallography. Most importantly, this construct is a candidate for immunogenicity studies to re-elicit 2G12-like antibodies.

## **Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

#### **Acknowledgments**

I.J.K. gratefully acknowledges Brandeis University and the NIH (R01 AI090745). We are grateful to Marcus Long for insightful discussions, and Polymun Scientific for a gift of mAb 2G12.

#### **References**

- 1. "Part Four Protein-Carbohydrate Interactions" Ch 13–20, in Gabius, Hans-JoachimThe Sugar Code. 2009WeinheimWiley-Blackwell, Wiley-VCH Verlag GmbH & Co. KGaA:233–346.
- 2. a) Mammen M, Choi S-K, Whitesides GM. Angew. Chem. Int. Ed. 1998; 37:2754–2794.b) Fasting C, Schalley CA, Weber M, Seitz O, Hecht S, Koksch B, Dernedde J, Graf C, Knapp EW, Haag R. Angew. Chem. Int. Ed. 2012; 51:10472–10498.
- 3. a) Levine PM, Carberry TP, Holub JM, Kirshenbaum K. MedChemComm. 2013; 4:493–509.b) Wittmann V, Pieters RJ. Chem. Soc. Rev. 2013; 42:4492–4503. [PubMed: 23598793]
- 4. MacPherson IS, Temme JS, Habeshian S, Felczak K, Pankiewicz K, Hedstrom L, Krauss IJ. Angew. Chem. Int. Ed. 2011; 50:11238–11242.
- 5. SELEX has previously been performed with dNTP's bearing modifications which are tolerated by polymerases. For selected examples, see: Keefe AD, Cload ST. Curr. Opin. Chem. Biol. 2008; 12:448–456. [PubMed: 18644461] Li MY, Lin N, Huang Z, Du LP, Altier C, Fang H, Wang BH. J. Am. Chem. Soc. 2008; 130:12636–12638. [PubMed: 18763762] Hollenstein M, Hipolito CJ, Lam CH, Perrin DM. ChemBioChem. 2009; 10:1988–1992. [PubMed: 19591189] Vaught JD, Bock C, Carter J, Fitzwater T, Otis M, Schneider D, Rolando J, Waugh S, Wilcox SK, Eaton BE. J. Am. Chem. Soc. 2010; 132:4141–4151. [PubMed: 20201573] for selected examples of carbohydrates attached to nucleic acids, see: Matsuura K, Hibino M, Yamada Y, Kobayashi K. J. Am. Chem. Soc. 2000; 123:357–358. [PubMed: 11456533] Adinolfi M, De Napoli L, Di Fabio G, Iadonisi A, Montesarchio D. Org. Biomol. Chem. 2004; 2:1879–1886. [PubMed: 15227540] Matsuura K, Hibino M, Ikeda T, Yamada Y, Kobayashi K. Chem. Eur. J. 2004; 10:352–359. [PubMed: 14735503] Yamada Y, Matsuura K, Kobayashi K. Bioorg. Med. Chem. 2005; 13:1913–1922. [PubMed: 15727847] Matsui M, Nishiyama Y, Ueji SI, Ebara Y. Bioorg. Med. Chem. Lett. 2007; 17:456–460. [PubMed: 17064895] D'Onofrio J, Petraccone L, Martino L, Di Fabio G, Iadonisi A, Balzarini J, Giancola C, Montesarchio D. Bioconjugate Chem. 2008; 19:607–616. Gorska K, Huang K-T, Chaloin O, Winssinger N. Angew. Chem. Int. Ed. 2009; 48:7695–7700. Moni L, Pourceau G, Zhang J, Meyer A, Vidal S, Souteyrand E, Dondoni A, Morvan F, Chevolot Y, Vasseur JJ, Marra A. ChemBioChem. 2009; 10:1369–1378. [PubMed: 19405074] Kiviniemi A, Virta P, Lonnberg H. Bioconjugate Chem. 2010; 21:1890–1901. Ciobanu M, Huang KT, Daguer JP, Barluenga S, Chaloin O, Schaeffer E, Mueller CG, Mitchell DA, Winssinger N. Chem. Commun. 2011; 47:9321– 9323. Scheibe C, Bujotzek A, Dernedde J, Weber M, Seitz O. Chem. Sci. 2011; 2:770–775. Matsui M, Ebara Y. Bioorg. Med. Chem. Lett. 2012; 22:6139–6143. [PubMed: 22940447] Scheibe C, Wedepohl S, Riese SB, Dernedde J, Seitz O. ChemBioChem. 2013; 14:236–250. [PubMed: 23292704]
- 6. a) Kolb HC, Finn MG, Sharpless KB. Angew. Chem.-Int. Edit. 2001; 40:2004–2021.b) Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. Angew. Chem.-Int. Edit. 2002; 41:2596–2599.c) Gierlich

J, Burley GA, Gramlich PM, Hammond DM, Carell T. Org. Lett. 2006; 8:3639–3642. [PubMed: 16898780] d) Gierlich J, Gutsmiedl K, Gramlich PM, Schmidt A, Burley GA, Carell T. Chem. Eur. J. 2007; 13:9486–9494. [PubMed: 17868170]

- 7. Binley JM, Wrin T, Korber B, Zwick MB, Wang M, Chappey C, Stiegler G, Kunert R, Zolla-Pazner S, Katinger H, Petropoulos CJ, Burton DR. J. Virol. 2004; 78:13232–13252. [PubMed: 15542675]
- 8. a) Mascola JR, Stiegler G, VanCott TC, Katinger H, Carpenter CB, Hanson CE, Beary H, Hayes D, Frankel SS, Birx DL, Lewis MG. Nat. Med. 2000; 6:207–210. [PubMed: 10655111] b) Hessell AJ, Rakasz EG, Poignard P, Hangartner L, Landucci G, Forthal DN, Koff WC, Watkins DI, Burton DR. PLoS Pathog. 2009; 5:e1000433. [PubMed: 19436712]
- 9. a) Trkola A, Purtscher M, Muster T, Ballaun C, Buchacher A, Sullivan N, Srinivasan K, Sodroski J, Moore JP, Katinger H. J. Virol. 1996; 70:1100–1108. [PubMed: 8551569] b) Sanders RW, Venturi M, Schiffner L, Kalyanaraman R, Katinger H, Lloyd KO, Kwong PD, Moore JP. J. Virol. 2002; 76:7293–7305. [PubMed: 12072528] c) Scanlan CN, Pantophlet R, Wormald MR, Ollmann Saphire E, Stanfield R, Wilson IA, Katinger H, Dwek RA, Rudd PM, Burton DR. J. Virol. 2002; 76:7306– 7321. [PubMed: 12072529] d) Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, Zhu P, Wormald MR, Stanfield RL, Roux KH, Kelly JW, Rudd PM, Dwek RA, Katinger H, Burton DR, Wilson IA. Science. 2003; 300:2065–2071. [PubMed: 12829775] e) Calarese DA, Lee H-K, Huang C-Y, Best MD, Astronomo RD, Stanfield RL, Katinger H, Burton DR, Wong C-H, Wilson IA. Proc. Natl. Acad. Sci. U. S. A. 2005; 102:13372–13377. [PubMed: 16174734]
- 10. For selected approaches, see: Singh S, Ni JH, Wang LX. Bioorg. Med. Chem. Lett. 2003; 13:327– 330. [PubMed: 12565922] Li HG, Wang LX. Org. Biomol. Chem. 2004; 2:483–488. [PubMed: 14770226] Wang LX, Ni J, Singh S, Li H. Chem. Biol. 2004; 11:127–134. [PubMed: 15113002] Ni JH, Song HJ, Wang YD, Stamatos NM, Wang LX. Bioconjugate Chem. 2006; 17:493–500. Wang J, Li H, Zou G, Wang LX. Org. Biomol. Chem. 2007; 5:1529–1540. [PubMed: 17571181] Krauss IJ, Joyce JG, Finnefrock AC, Song HC, Dudkin VY, Geng X, Warren JD, Chastain M, Shiver JW, Danishefsky SJ. J. Am. Chem. Soc. 2007; 129:11042–11044. [PubMed: 17711286] Joyce JG, Krauss IJ, Song HC, Opalka DW, Grimm KM, Nahas DD, Esser MT, Hrin R, Feng MZ, Dudkin VY, Chastain M, Shiver JW, Danishefsky SJ. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:15684–15689. [PubMed: 18838688] Luallen RJ, Lin JQ, Fu H, Cai KK, Agrawal C, Mboudjeka I, Lee FH, Montefiori D, Smith DF, Doms RW, Geng Y. J. Virol. 2008; 82:6447– 6457. [PubMed: 18434410] Wang SK, Liang PH, Astronomo RD, Hsu TL, Hsieh SL, Burton DR, Wong CH. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:3690–3695. [PubMed: 18310320] Astronomo RD, Kaltgrad E, Udit AK, Wang SK, Doores KJ, Huang CY, Pantophlet R, Paulson JC, Wong CH, Finn MG, Burton DR. Chem. Biol. 2010; 17:357–370. [PubMed: 20416507] Doores KJ, Fulton Z, Hong V, Patel MK, Scanlan CN, Wormald MR, Finn MG, Burton DR, Wilson IA, Davis BG. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:17107–17112. [PubMed: 20852065] Marradi M, Di Gianvincenzo P, Enriquez-Navas PM, Martinez-Avila OM, Chiodo F, Yuste E, Angulo J, Penades S. J. Mol. Biol. 2011; 410:798–810. [PubMed: 21440555] Clark BE, Auyeung K, Fregolino E, Parrilli M, Lanzetta R, De Castro C, Pantophlet R. Chem. Biol. 2012; 19:254–263. [PubMed: 22365608]
- 11. The Man4 fragment used in our previous approach (Ref. 4) is known to be responsible for most of the contacts with 2G12, according to X-ray crystallographic data (Ref. 10d,e)
- 12. Sinay P. Pure and Applied Chemistry. 1991; 63:519–528.
- 13. Geng X, Dudkin VY, Mandal M, Danishefsky SJ. Angew. Chem. Int. Ed. 2004; 43:2562–2565.
- 14. Iserloh U, Dudkin V, Wang Z-G, Danishefsky SJ. Tetrahedron Lett. 2002; 43:7027–7030.
- 15. Nyffeler PT, Liang C-H, Koeller KM, Wong C-H. J. Am. Chem. Soc. 2002; 124:10773–10778. [PubMed: 12207533]
- 16. Thompson JD, Higgins DG, Gibson TJ. Nucleic Acids Res. 1994; 22:4673–4680. [PubMed: 7984417]
- 17. Zuker M. Nucl. Acids Res. 2003; 31:3406–3415. [PubMed: 12824337]
- 18. The assay used in Wong's study (ref. 10i), based on competition binding to Man<sub>9</sub>-printed glass slides, was sensitive to glycan printing density. Given the large differences between our solutionphase assay format and their surface-based assay, caution should be exercised in comparison of the two types of data





#### **Figure 1.**

(a) new library templates, containing reduced A's, leading to reduced glycosylation sites (b) predicted multivalency profile of each library (% of sequences with 0, 1, 2, 3, …, 15 glycans)

Temme et al. Page 7



#### **Figure 2.**

a) M-fold structure predictions of clone 7A8-8 and suggested truncation. b) Effect of truncations/deletions on  $K_d$  of 2G12 recognition. [a] no binding



**scheme 1.** Overview of SELMA (SELection with Modified Aptamers)

Temme et al. Page 9



#### **scheme 2.**

Synthesis of Man<sub>9</sub> azide. a) PhBCl<sub>2</sub>, TES-H, 4Å MS, 80 %; b) 2, Sinaÿ Reagent [(p-BrC<sub>6</sub>H<sub>4)3</sub>N<sup>+</sup> SbCl<sub>6</sub><sup>d</sup>], 4Å MS, 0 °C, MeCN, 81 %; c) NaOMe, MeOH/THF, 81 %; d) Na metal, THF, -78 °C, NH<sub>3</sub>(l); e) TfN<sub>3</sub>, K<sub>2</sub>CO<sub>3</sub>, CuSO<sub>4</sub>, DCM/MeOH/H<sub>2</sub>O, 63 % (2 steps).

#### **Table 1**

Representative 5'->3' sequences after 7 rounds of SELMA selection,  $K_d$  for binding to 2G12, and consensus sequence.



*a*<br>
Not determined; binding curve was linear up to highest 2G12 concentration tested.

*b* no binding.

 $c$ <sup>10</sup>-Man4 winner from previously published selection with 25 % glycosylated library (Ref 4).