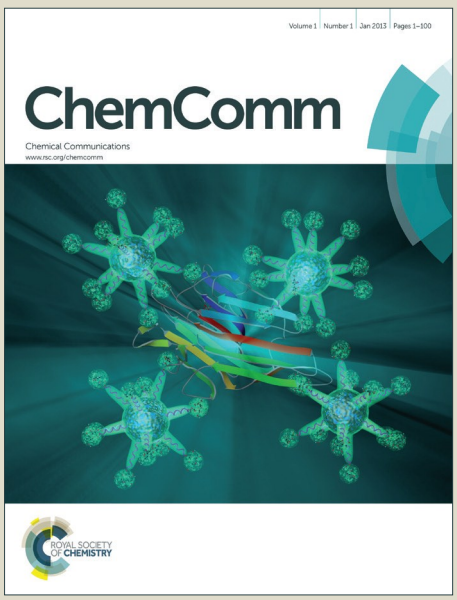
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

A KISS (Keep It Simple, Sensor) Array for Glycosaminoglycans†

Jean-Patrick Francoia and Laurent Vial*

:Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Herein, we demonstrate that it is possible to design a sensor array with one multivalent cationic receptor (a dendrigraft of lysine) and one fluorescent anionic indicator (a fluorescein-labeled peptide). Depending on the loading of the indicator on the receptor, negatively charged glycosaminoglycans (GAGs) induce a positive or negative variation of the fluorescent signal as they displace the indicators from the receptor or they compact the indicators on the receptor's surface, respectively. This unique strategy allows not only the blind identification of pure GAGs with a level of accuracy of 100%, but also the differentiation of mixtures.

Glycosaminoglycans (GAGs) are an important class of exogenous biopolymers that display numerous therapeutic activities.[1] Because GAGs are structurally analogous (Fig. 1), orthogonal and lengthy analytical techniques – involving NMR spectroscopy, HPLC, and capillary electrophoresis – are necessary to identify them.[2] The ability to routinely discriminate and assess the purity of these anionic polysaccharides is mandatory in order to prevent public health disasters. Indeed, the use of contaminated heparin in patients undergoing dialysis resulted in serious acute hypersensitivity reactions (including some resulting in death) during the late 2000s.[3] Contaminants in adulterated lots included chondroitin sulfate A, dermatan sulfate (chondroitin sulfate B), and hyaluronic acid. To address this issue, some research groups developed pattern-based recognition assays for various GAGs by using arrays of nanoparticles, macrocycles, or liposomes.[4] Despite the success of those works, extensive synthetic steps were systematically involved to obtain the multiple receptors and/or indicators that were necessary to construct the arrays; therefore impeding their use for cheap routine screening of GAGs. Herein a much simpler sensor array for GAGs, from a single receptor and a single indicator, is presented.

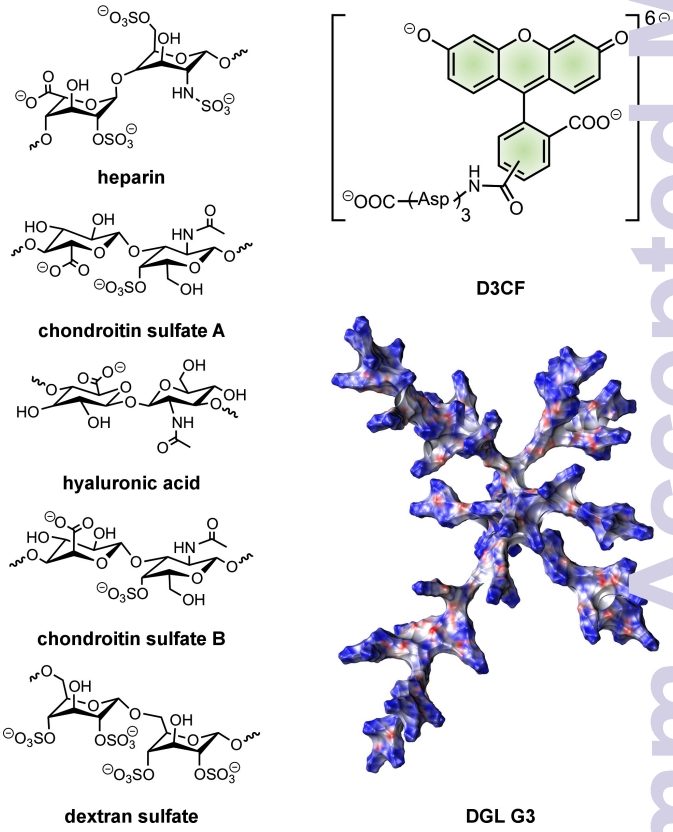


Fig. 1 *Left, from top to bottom*: major repeat unit of selected GAGs. *Top right:* bioconjugate between a homopolymer of aspartic acid and 5(6)-carboxyfluorescein, D3CF. *Bottom right:* minimized structure with electrostatic potential surface of the third-generation dendrigraft poly-*L*-lysine polymer DGL G3.

We recently reported that a cationic dendrigraft poly-*L*-lysine polymer (DGL) was able to form a multi-ligand complex with a conjugate between an anionic homopolymer of aspartic acid and 5(6)-carboxyfluorescein; leading to the extinction of the optical signal as a result of its aggregation on the surface of the receptor.[5] Fluorescence could be restored upon the introduction of heparin in an indicator-displacement process. Although this sensing ensemble allowed us to detect and

Institut des Biomolécules Max Mousseron, UMR 5247 CNRS – Université de Montpellier, Place Eugène Bataillon, cc 1706, 34296 Montpellier cedex 5, France. E-mail: laurent.vial@univ-montp2.fr

† Electronic Supplementary Information (ESI) available: synthetic procedure for D3CF; details for the analytical and computational methods. See DOI: 10.1039/x0xx00000x

quantify – for the first time – heparin in human blood samples at clinically relevant levels, it was not possible to discriminate between GAGs such as heparin and chondroitin sulfate. In this previous study, the fourth-generation dendrigraft poly-*L*-lysine polymer and the seven-membered homopolymer of aspartic acid and 5(6)-carboxyfluorescein bioconjugate were necessary to work in presence of complex biological fluids in order to ensure sufficient binding between the partners. Since the following experiments were all performed in a much less competitive medium (10 mM HEPES buffer, pH 7.8), the more accessible DGL G3 and shorter D3CF were chosen in this study in order to keep these assays for GAGs as simple as possible (Fig. 1).[7] We envisioned that, using an array that was based on the variable loading of the fluorescent indicator on the receptor, each GAG could generate a specific fluorescent response pattern. Indeed, because they display specific electrostatic surfaces, the complexes can be considered as "differential" receptors for GAGs (Fig. 2).[8,9,10]

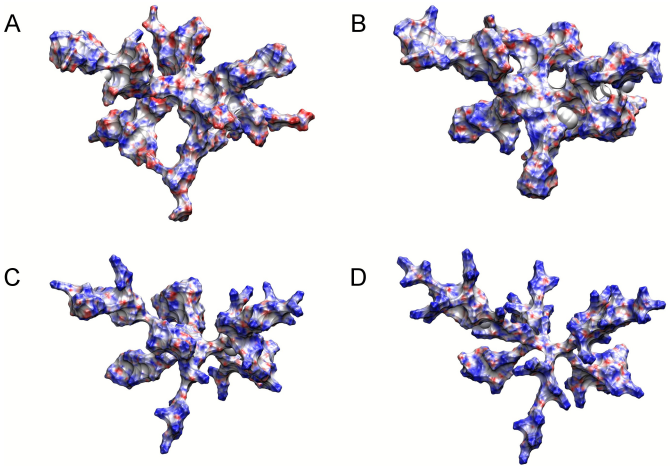


Fig. 2 Minimized structures of G3-D3CF complexes at various loadings, with electrostatic potential surfaces. Loadings: 100% (A), 75% (B), 50% (C), 25% (D).[9] For calculation details, see ESI.†

Upon addition of dextran sulfate to G3-D3CF complexes in water containing 10 mM HEPES and buffered to pH 7.8, distinct optical behaviours were detected depending on the loading of the indicator on the receptor (Fig. 3).[11] At 100% loading, increased levels of dextran sulfate correlated positively with the fluorescence intensity (FI) response (Fig. 3A). At 75% loading, a decrease of the FI occurred upon addition of dextran sulfate until a minimum value FI_{min} (Fig. 3B), followed by the restoration of the optical signal. Similar titration curves could be observed at lower loadings, although reaching FI_{min} required increasing amounts of dextran sulfate (Fig. 3C and 3D). For all these titration curves, it could be noted that i) the initial fluorescence FI_0 is in agreement with the titration curve of G3 with D3CF (see ESI†, Fig. S1), and ii) a final plateau with a specific maximum fluorescence value FI_{max} is reached at a GAG concentration of 4.8 mg.L^{-1}.

A simple model to explain such unexpected behaviors, which depend on the loading of the indicator on the receptor, is proposed (Fig. 4).

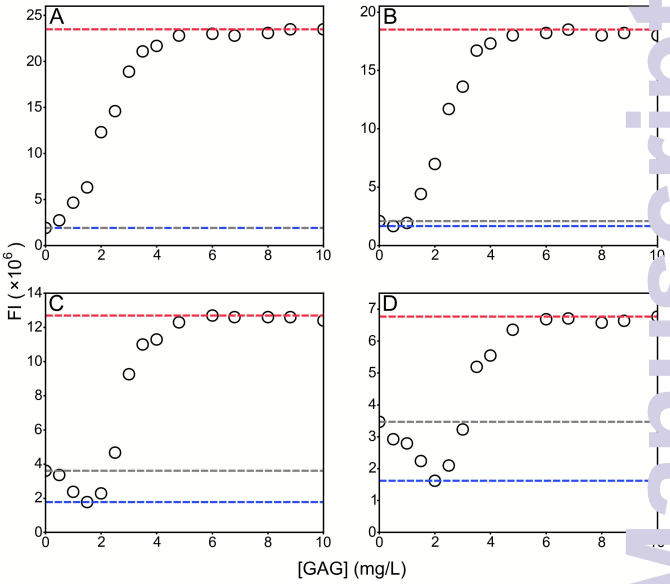


Fig. 3 Fluorescence titrations of the G3-D3CF complexes with dextran sulfate. Loadings: 100% (A), 75% (B), 50% (C), 25% (D).[9] Dashed lines: *grey* = FI_0, *blue* = FI_{min}, *red* = FI_{max}. Conditions: 10 mM HEPES buffer, pH = 7.8, [G3] = 163 ([lysine] = 20 mM), λ_{ex} = 485 nm, λ_{em} = 535 nm.

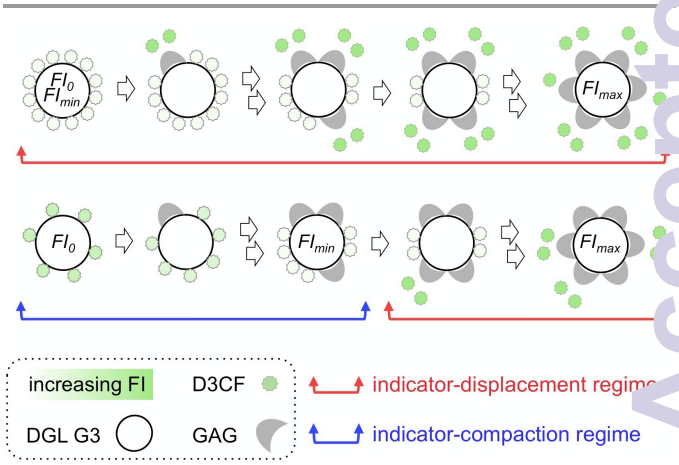


Fig. 4 Proposed model for the various binding behaviours of GAGs to G3-D3CF complexes.

Starting from a fully loaded G3 with D3CF, FI_0 is equal to FI_{min} since the distance between the indicators on the receptor's surface is at its smallest, resulting in maximum fluorescence self-quenching (Fig. 4, top). The subsequent stepwise addition of GAG leads to the continuous displacement of the indicator from the dendrigraft's surface and to the increase of FI, until a plateau that indicates that all the fluorophore is actually free in solution according to the calibration curve of D3CF (see ESI†, Fig. S1). In contrast, a lower loading gives a less compacted starting arrangement of the indicators on the dendrigraft's surface, yielding a higher FI_0 (Fig. 4, bottom). Additional GAG molecules can therefore bind to the receptor without displacing but with compacting the fluorophores on the receptor's surface. This results into the diminution of the optical signal until FI_{min} that is correlated with maximum compaction. Upon this point, additional GAG molec

restitute the indicator-displacement regime. In this model, i) the lower the loading is, the higher the GAG concentration necessary to reach full compaction is, and ii) the same amount of GAG is needed to fully displace the indicator whatever the loading is.

Such a model could be used to virtually explain the similar behaviors that were observed upon the addition of other GAGs, each one displaying specific indicator-displacement or indicator-compaction regimes (see ESI†, Fig. S2). As a consequence, differential optical responses were observed when the selected GAGs were submitted at a concentration of 3.5 mg.L^{-1} to our sensor array (Fig. 5).

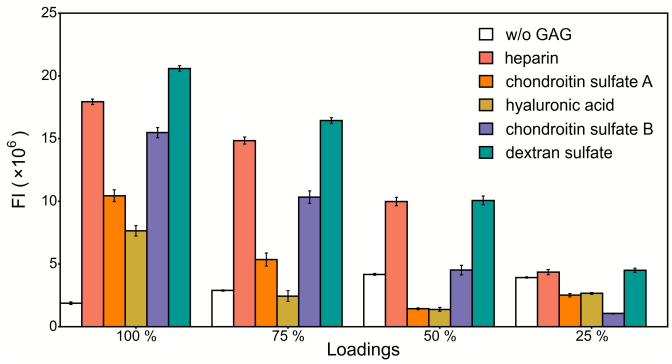


Fig. 5 Fluorescence intensities from the addition of GAGs to G3-D3CF complexes. Loadings: 100% (A), 75% (B), 50% (C), 25% (D). Conditions: 10 mM HEPES buffer, pH = 7.8, [G3] = 163 nM ([lysine] = 20 mM), [GAG] = 3.5 mg.mL^{-1}, λ_{ex} = 485 nm, λ_{em} = 535 nm.

Differentiation was demonstrated and validated by linear discriminant analysis (LDA), which is a supervised procedure that allows the classification of data.[12] In a first place, known samples of GAGs (Fig. 6, triangles) were used as a training input to the procedure, giving a captured variance ratio of 93.4% for the first linear discriminant and 6.1% on the second one. Decision boundaries (Fig. 6, dashed lines) were then generated by using Bayes' rule.[13] These boundaries delimit areas of the reduced space that are exclusively attributable to a single GAG. Finally, unknown samples (Fig. 6, circles) were tested against the training set, achieving 100% of identification with 50 correct cases out of 50, and demonstrating therefore the robustness of this assay.[14] In order to further demonstrate the applicability of our sensor array, we tested its ability to assess the purity of samples of heparin contaminated with 10%, 20%, 30%, and 50% of chondroitin sulfate B, which is the most prevalent impurity in pharmaceutical heparin preparations.[15] From the resulting score plot it was clear that all mixtures were separated with good resolution (Fig. 7).

In conclusion, we reported a simple sensor array that allows not only the unambiguous identification of structurally similar pure glycosaminoglycans, but also the differentiation of mixtures of glycosaminoglycans. The simplicity of this array relies on the use of only one commercial receptor and one highly accessible fluorescent indicator. Furthermore, a new sensing approach is introduced in this study that one could describe as a compaction/displacement-indicator assay. We are confident that such an original approach will find future applications in the supramolecular analytical field.

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and the Université de Montpellier. DGL G3 was supplied by the COLCOM company (Montpellier, France). The authors thank Dr Sophie Sisavath (pharmacist) and Prof. Jean-Noël Bacro (mathematician) for fruitful discussions.

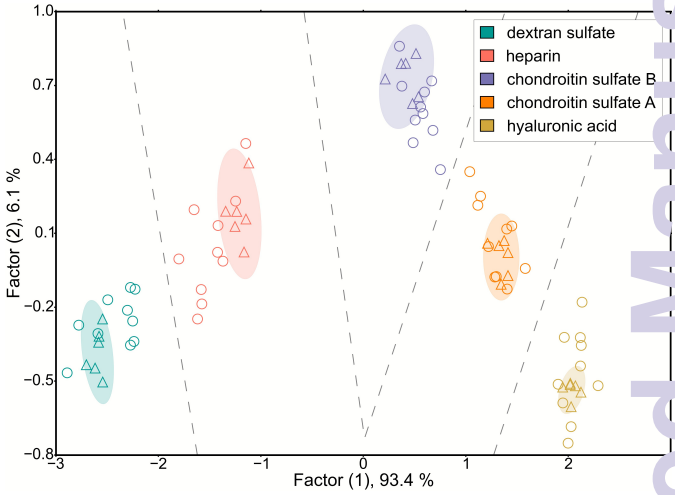


Fig. 6 LDA canonical score plot for the response of the sensor array to GAGs. *Dashed lines*: decision boundaries, *triangles*: training set, *circles*: blind tests, *ellipses*: confidence limits at 99% for the training set.

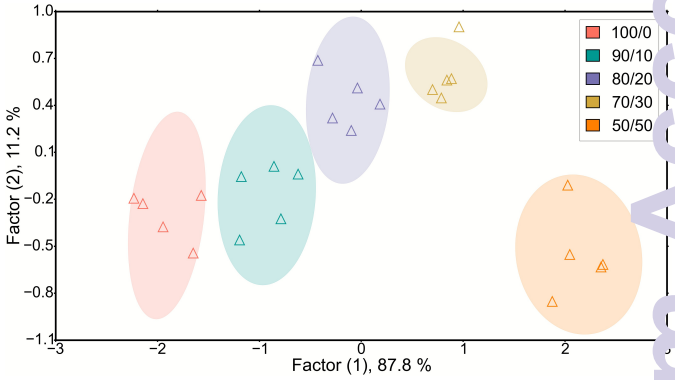


Fig. 7 LDA canonical score plot for the response of the sensor array to heparin/chondroitin sulfate B mixtures. *Ellipses*: confidence limits at 99%.

Notes and references

1 N. Volpi, *Curr. Med. Chem.*, 2006, **13**, 1799.

2 *(a)* S. Beni, J. F. K. Limtiaco and C. K. Larive, *Anal. Bioanal. Chem.*, 2011, **399**, 527; *(b)* M. Guerrini, Z. Zhang, Z. Shriver, A. Naggi, S. Masuko, R. Langer, B. Casu, R. J. Linhardt, G. Torri and R. Sasisekharan, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 16956.

3 D. B. Blossom, A. J. Kallen, P. R. Patel, A. Elward, L. Robinson, G. Gao, R. Langer, K. M. Perkins, J. L. Jaeger, K. M. Kurkjian, M. Jones, S. F. Schillie, N. Shehab, D. Ketterer, G. Venkataraman, T. K. Kishimoto, Z. Shriver, A. W. McMahon, K. F. Austen, S. Kozlowski, A. Srinivasan, G. Turabelidze, C. V. Gould, M. J. Arduino and R. Sasisekharan, *N. Engl. J. Med.*, 2008, **359**, 2674.

4 (*a*) M. Twomey, T. Vokatá, M. R. Kumar and J. H. Moon, *Chem. Commun.*, 2015, **51**, 4065; (*b*) S. G. Elci, D. F. Moyano, S. Rana, G. Y. Tonga, R. L. Phillips, U. H. F. Bunz and V. M. Rotello, *Chem. Sci.*, 2013, **4**, 2076; (*c*) E. K. Nyren-Erickson, M. K. Haldar, Y. Gu, S. Y. Qian, D. L. Friesner and S. Mallik, *Anal. Chem.*, 2011, **83**, 5989; (d) 1 R. B. C. Jagt, R. F. Gómez-Biagi and M. Nitz, *Angew. Chemie - Int. Ed.*, 2009, **48**, 1995.

5 J.-P. Francoia, R. Pascal and L. Vial, *Chem. Commun.*, 2015, **51**, 1953.

6 B. T. Nguyen and E. V. Anslyn, *Coord. Chem. Rev.* 2006, **250**, 3118.

7 Given that the simple carboxyfluorescein was not able to bind efficiently G3 in HEPES buffer (10 mM, pH 7.8),[5] we selected an intermediate peptidic chain length for the indicator. See ESI for the straightforward solid-phase synthesis of D3CF.†

8 A. T. Wright and E. V. Anslyn, *Chem. Soc. Rev.* 2006, **35**, 14.

9 Loadings were calculated from the titration curve of DGL G3 with D3CF, see Fig. S1 in ESI.† For the rationalization of the titration curve's profile, see ref. 5.

10 Upon addition of G3 to D3CF, a red-shift of the absorption spectrum of the dye was observed by UV-visible spectroscopy (λ_{max}=493→501 nm), indicating its stabilization by non covalent interactions (*i.e.*, the reduction of its HOMO-LUMO gap energy) and confirming the formation of a complex between G3 and D3CF.

11 Mass concentrations for GAGs were used in this study in order to: i) be able to blind test the polysaccharides, and ii) get rid of any approximation about their structure that may lead to rough molar concentration calculations.

12 S. Stewart, M. A. Ivy and E. V. Anslyn, *Chem. Soc. Rev.* 2014, **43**, 70.

13 T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning Data, 2nd Ed.*, Springer: New York, 2009.

14 A Python script has been edited in order to automatize this whole analytical procedure, which is based on the LDA class of the scikit-learn Python module (http://scikit-learn.org), see: F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825. The script is available upon simple request.

15 G. A Neville, F. Mori, K. R. Holme and A. S. Perlin, *J. Pharm. Sci.*, 1989, **78**, 101.