ChemComm

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: X. Zhou, B. Fu, Y. Park, K. T. Kim, K. Chen, G. Zou, Q. Wei, S. Peng, Y. Chen and B. H. Kim, *Chem. Commun.*, 2018, DOI: 10.1039/C8CC05197F.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/C8CC05197F

Journal Name

COMMUNICATION

A Novel Nucleic Acid Aptamer Tag: A Rapid Fluorescent Strategy Using A Self-constructing G-quadruplex From AGG Trinucleotide Repeats

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Boshi Fu[a], Yoojin Park[b], Ki Tae Kim[b], Kun Chen[a], Guangrong Zou[a], Qi Wei[a], Shuang Peng[a], Yuqi chen[a], Byeang Hyean Kim[b]* and Xiang Zhou[a]*

Herein, We have developed a novel fluorescence labeling strategy for nucleic acid aptamers based on self-assembling between AGG tri-nucleotide repeats and a pyrene-modified oligonucleotide. This strategy could be an effective tool for developing targeting-imaging systems and biosensor systems to detect the target molecule.

Nucleic acid aptamers are functional oligonucleotides capable of recognizing specific targets such as small molecules, proteins, and nucleic acids, as well as cells, tissues and organisms[1, 2]. The molecular basis for identification and regulation by the aptamers is usually noncanonical spatial structures formed by their oligonucleotides[3]. Depending on the structures, nucleic acid aptamers can even mimic the properties of antibodies and rapidly penetrate tissue to accurately bind to targets[4]. Compared with natural antibodies, they have the advantages of low molecular weights and lacking toxicity or immunogenicity. Owing to such characteristics, nucleic acid aptamers have been used widely as recognition elements in diagnostic detection[5], cell capture[6], drug delivery[7] and targeted therapies[8].

The aptamers are created through SELEX[9]. It is an iterative selection procedure used to find particular oligonucleotides with desired properties, most often the ability to bind to specific targets. As with normal nucleic acids, the monomers of nucleic acid aptamers are nucleotides. Therefore, it is convenient to artificially modify nucleic acid aptamers with multiple functional groups. By means of chemical modification, aptasensors can be used to quantify or image targets using various kinds of signals[10-12]. With the help of nucleic acid aptamers, various functional materials could be endowed with more extensive bioanalytical and biomedical applications[13,14]. Thus, aptamers are powerful tools that could help scientists to solve problems in scientific research[15], industry and clinical medicine[16]. It is worthwhile to develop more and better application for aptamers.

Distinguished from other nucleic acids, aptamers are capable of responding to specific targets or changes in the microenvironment. Often, external stimuli could induce a folding process allowing for the construction of functional noncanonical secondary structures such as hairpins[17] and G-quadruplexes[18], because such structures can offer binding sites for the recognition of targets. Differing from antibodies based on their major structure being a large, Y-shaped protein, aptasensors and aptadevices based on structure switching have been developed[19]. The commonly used fluorophore-quencher-aptamer hairpin system arose based on this mechanism. Such engineered aptamers are impressive in terms of their favourable effectiveness. Aptamer-DNAzyme hairpins[20] have been reported as useful constructs for the detection of their respective aptamer-substrate complexes. After this featured work, more similar aptasensors based on nucleic acid structure switching were developed.

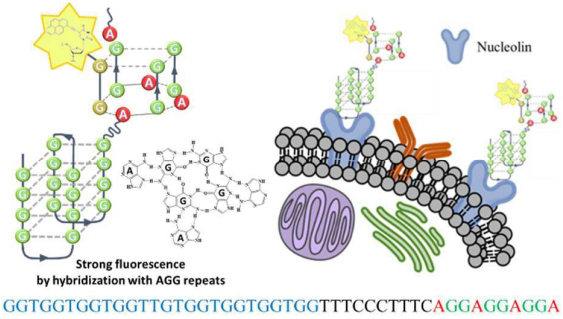
Trinucleotide repeats exist abundantly in prokaryotic and eukaryotic genomes[21]. AGG trinucleotide repeats, one of most common triplet repeats, are associated with some diseases[22]. *In vitro*, oligonucleotides consisting of such sequences can fold into G-quadruplexes[23,24] with hexad or heptad planes in their stem structures[25]. Based on such characteristic, Kim`s group had developed a fluorescent detection method for recognition of AGG trinucleotide repeats. One of the probe strand which they had synthesized was named as PYG3 because the sequence contained three guanines and a 5-(1-pyrenylethynyl)-2`-deoxyuridine (^{Py}U) at its 5`-ends. It could recognize AGG trinucleotide repeats and assemble into (3+1) intermolecular G-quadruplexes[26]. After that, a strong fluorescence could indicate the existence of AGG trinucleotide repeats. In this report, we used the AGG trinucleotide repeats which was the target of this probe strand to develop a novel

[a] College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers of Ministry of Education, Wuhan University, Wuhan, Hubei, 430072, P. R. China. E-mail: xzhou@whu.edu.cn; Fax: +86-27-68756663; Tel: +86-27-68756663
[b] Department of Chemistry, BK School of Molecular Science, Pohang University of Science and Technology, Pohang 790-784, South Korea. E-mail: bhkim@postech.ac.kr; Fax: +82 54-279-3399; Tel: +82 54-279-2115.

†Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here].
See DOI: 10.1039/x0xx00000x

View Article Online
DOI: 10.1039/C8CC05197F

aptamer Tag and built targeting-imaging systems and biosensor systems based on this method (Scheme 1&2).



Scheme 1 The fluorescent signal from the self-assembly between PYG3 and AGG trinucleotide repeats and targeted imaging of cancer cells by the TG-AS1411/PYG3 system

First, we extended the nucleolin targeting aptamer with AGG trinucleotide repeats and named the resulting oligonucleotide "TG-AS1411" (Table S1, ESI†). There were extra ten nucleotides (black) between the AS1411 (Blue) and AGG trinucleotide repeats (red/green). Such interval could avoid the interference from AGG Taq during folding process of the aptamer part. The circular dichroism could verify our design. For TG-AS1411 (Fig. 1 A, red), the positive peak at approximately 265 nm and the negative peak at approximately 240 nm were same as the AS1411 in same condition (Fig. S2, ESI†). These characteristic peaks indicated that the TG-AS1411 and AS1411 folded into the same parallel-strand quadruplex formation in the presence of Na^+/K^+. Aside from such characteristic peaks, there was also a transition band in the 270-300 nm region. This smooth band might be caused by the AGG repeats in the tail of TG-AS1411. The relatively unordered stacking of guanines contained both heteropolar and homopolar formations, similar to the flexible single-strand DNA in this ionic environment. These AGG trinucleotide repeats will self-assemble with the probe strand into a (3+1) intermolecular G-quadruplex structures and produce a strong fluorescent signal. Through this process, we provided the aptamer AS1411 with a fluorescent tracer tag under the premise of maintaining its original construction and function.

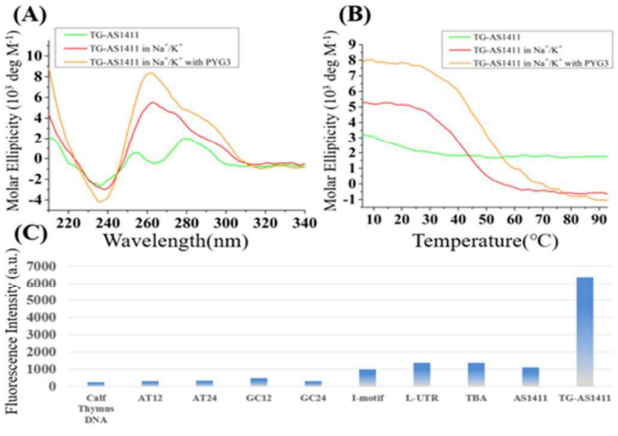


Fig. 1 (A) Circular dichroism spectra of TG-AS1411 (10 μM) in unwound single-strand/G-quadrupex conformation and with PYG3 (10 μM). (B) Thermal curves for melting temperature of TG-AS1411 (10 μM) in unwound single-strand/G-quadrupex conformation and with PYG3 (10 μM). (C) Fluorescence responses of PYG3 (1 μM) after incubation with a variety of sequences (1 μM).

We had demonstrated the AGG repeat tag did not change the stacked formation in TG-AS1411 compared to the normal AS1411. After adding PYG3, the spectrum changed (Fig. 1 A). The increase in the positive peaks at 265 nm and 290 nm indicated that more G-quartets appeared. We could infer that the AGG repeat tag and PYG3 had assembled into a (3+1) intermolecular G-quadruplex. The results (Fig. 1 B) of thermodynamic melting studies showed that such (3+1) intermolecular G-quadruplex structures would cause an increase in the melting temperature (from 43.62 °C to 49.70 °C). There were more G-quartets in the G-quadruplex formed by TG-AS1411 and PYG3 than in the G-quadruplex formed from only TS-AS1411. These G-quartets could make the G-quadruplex more difficult to unwind during the increase in temperature. All the above results indicated that the AGG-trinucleotide tag in the nucleic acid aptamer could provide a binding site for PYG3 to form a (3+1) intermolecular G-quadruplex without impacting the original secondary structures of the aptamers.

As reported, such assembly method made PYG3 become a sensitive and rapid fluorescent sensor for recognizing AGG repeats[27]. By the same way, PYG3 could distinguish the aptamer with the AGG repeat tag from other oligonucleotides. Other G-rich oligonucleotides could not disturb the ability of PYG3 to recognize the AGG repeat tag (Fig. 1 C). The specific fluorescent signal means that the AGG repeat tag was an effective strategy with high selectivity for labelling nucleic acid aptamers.

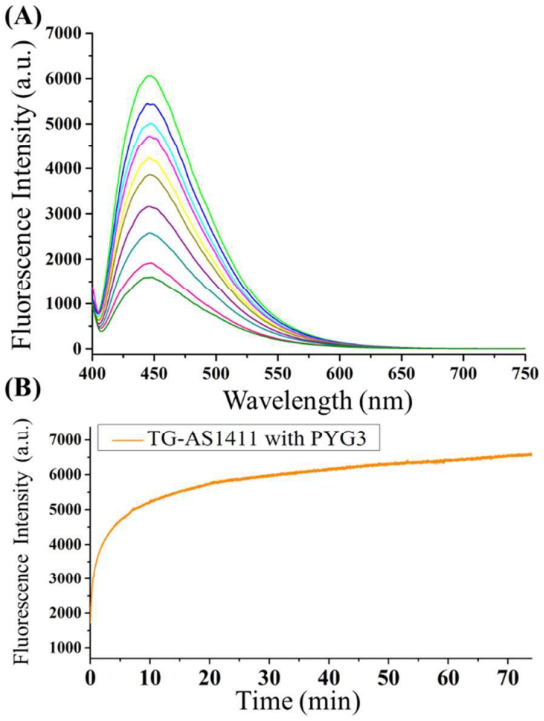


Fig. 2 (A) Fluorescence graphs of PYG3 with different concentrations of TG-AS1411 (10 nM, 25 nM, 50 nM, 75 nm, 0.1 μM, 0.2 μM, 0.4 μM, 0.6 μM, 0.8 μM, 1 μM). (B) Real-time fluorescence monitoring during the self-assembly process between PYG3 (1 μM) and TG-AS1411 (1 μM).

View Article Online
DOI: 10.1039/C8CC05197F

Further investigation demonstrated that our strategy was capable of accurately quantifying nucleic acid aptamers with AGG repeat tags. In the present of PYG3, the fluorescence signal enhancement was observed with the concentration increase of TG-AS1411 (Fig. 2 A). Logistic fitting proved the consistency of the fluorescence intensity and concentration of TG-AS1411 (Fig. S13, ESI†). The results of the kinetic research (Fig. 2 B) showed that the fluorescence signal caused by the construction of the (3+1) intermolecular G-quadruplex plateaued in 1 hour. It is a rapid fluorescent strategy for the labelling of nucleic acid aptamers.

The performance of our novel tag *in vitro* showed its potential in intracellular imaging applications. AS1411, the nucleolin-targeted DNA aptamer[27], has been widely used as a recognition tool to develop biodetection systems and treatment methods[28-29]. Nucleolin (NCL) is always highly expressed on the surface membrane of continuously proliferating cells, including cancer cells[30].

Through the combination of this cancer-specific aptamer with the AGG trinucleotide tag and PYG3, we developed a novel imaging system for targeting cancer cells (scheme 1). Human breast cancer (MCF-7) cells were selected as the target cells of our novel imaging system. Using a confocal fluorescence microscopy imaging system, MCF-7 cells were visualized though the strong fluorescence signal. However, there was no such fluorescence signal that was observable from healthy Chinese hamster ovary (CHO) cells. This result demonstrated that by selectively binding with nucleolin, our system could distinguish MCF-7 breast cancer cells from normal cells though fluorescence imaging. Furthermore, this AGG repeat tag could be combined with all kinds of nucleic acid aptamers to develop bioimaging systems for various targets.

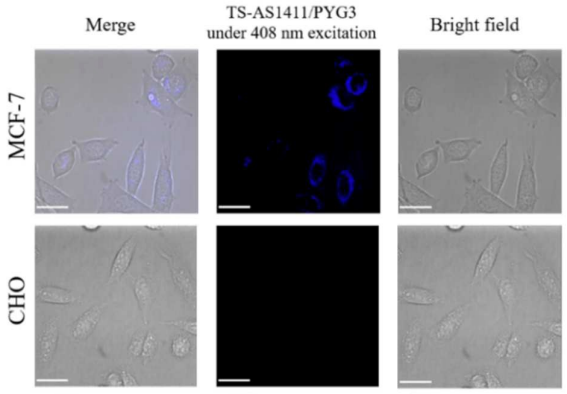


Fig. 3 Confocal fluorescent images of MCF-7 cells versus CHO cells (TG-AS1411/PYG3 (150 nM), Scale bar = 24 μm, experimental details could be read in the ESI.).

Based on the intracellular imaging results (Fig. 3), we tested our system in formalin-fixed, paraffin-embedded (FFPE) tissue sections. From the confocal fluorescence microscopy images, we clearly found that our method effectively imaged the breast cancer tissues and not benign breast tissue (Fig. 4). Our method could provide visual graphics as a reference for cancer diagnosis.

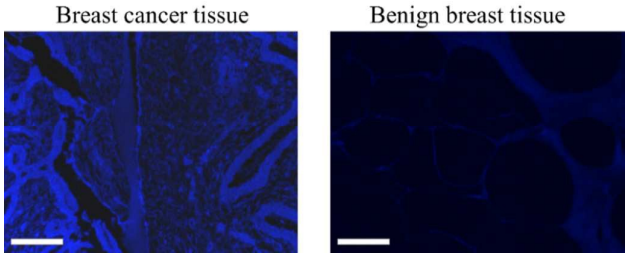


Fig. 4 Confocal fluorescence images of breast cancer tissues versus benign breast tissue (TG-AS1411/PYG3 (150 nM), Scale bar = 60 μm, experimental details could be read in the ESI.).

Compared to other tag methods, the AGG trinucleotide tag has a unique advantage; the homologous nucleobases in the AGG-trinucleotide tag make it possible to construct complex nucleic acid topologies in switching conformations. To implement this idea, we selected the DNA aptamer for adenosine monophosphate (AMP) as a transforming foundation. After simulative evaluation and semi-empirical design (Fig. S6, ESI†), we obtained a series of hairpins containing the AMP aptamer and AGG trinucleotide tag. These hairpins were named AMP-TG-hairpins. The ancillary sequences in the AMP-TG-hairpin series could match the sequences in the other terminus via complementary pairing. Based on such AMP-TG-hairpins, we built a biosensor for AMP by intruding PYG3. When there was no AMP in the solution, the AGG-trinucleotide tag would be caged in the hairpin structure of the stem with parts of the AMP aptamer. The PYG3 could not assemble the (3+1) intermolecular G-quadruplex with the AGG-trinucleotide tag. The fluorescence intensity was weak. However, if AMP was present, the structural switching would occur in AMP-TG-hairpins. To offer the binding site for AMP, the aptamer sequence (blue) would fold spontaneously. This would open the original hairpin structure, releasing the AGG-trinucleotide tag. Under these circumstances, the tag would self-assemble with PYG3 to form the (3+1) intermolecular G-quadruplex, while the fluorescence intensity increased (Fig. 5).

The fluorescence intensity from all biosensors increased when AMP molecules were introduced into the solution (Fig. 5 B, Fig. S9 ESI†). The best one of them, the bio-sensor based on AMP-TG-hairpin_2 was selected to further investigation. As a bio-sensor based on aptamer without amplified process, our method could quantify the adenosine monophosphate (AMP) with a direct detection limit of 2.5 mM. (Fig. 5 B, Fig. S10 ESI†) Compare to other bio-sensor by similar design, our strategy was effective tool as the signal source like DNAzyme[20] or nanoparticles[31]. However, the AGG trinucleotide repeats also might reduce the affinity of original aptamers slightly. The aptamers still had high selectivity for AMP. AMP analogues couldn't increase the fluorescence intensity (Fig. S11 ESI†).

These results demonstrated that our biosensors could be used for the detection of AMP within the solution environment and this strategy could be extended for various low molecular weight molecules or proteins.

View Article Online
DOI: 10.1039/C8CC05197F

Circular dichroism studies provided evidence for the (3+1) intermolecular G-quadruplex formed by the AGG repeat tag and PYG3 after structure switching caused by AMP (Fig. S7, S8 ESI†).

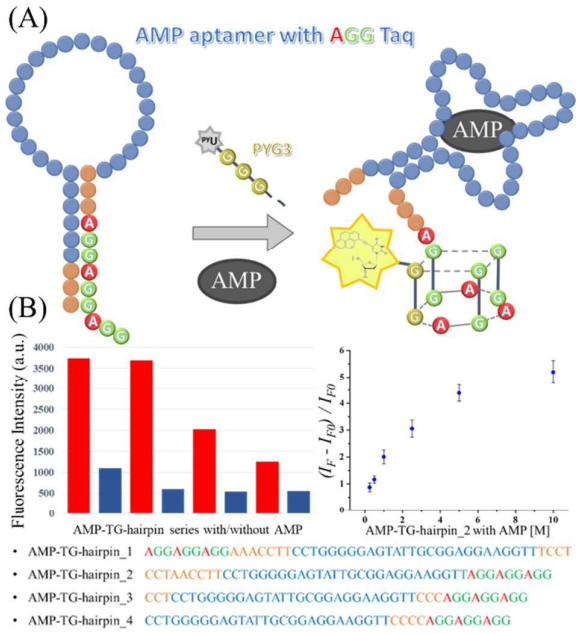


Fig. 5 (A) Bio-sensors for AMP based on switching hairpin. Each switching hairpin was consisted of three part: the AMP aptamer sequence (blue), AGG repeat Taq (red/green) and the ancillary sequence for maintaining the hairpin structure (orange). (B) Fluorescence intensity of biosensor systems in response to AMP. Each biosensor system contained AMP-TG-hairpin (1 μM) and PYG3 (1 μM) with/without AMP (5 μM) in a Na^+/K^+ solution (NaCl 100 mM, KCl 10 mM) and quantification of AMP concentration by the fluorescence signal form the bio-sensors based on AMP-TG-hairpin_2. $((I_F - I_{F0})/I_{F0,}$ where I_F and I_{F0} are the fluorescence intensity at 450 nm in the Na^+/K^+ solution with/without AMP)

In summary, we used the AGG-trinucleotide repeats as a novel tag for nucleic acid aptamers. With the combination of a short G-rich oligodeoxynucleotide modified with a 5-(1-pyrenylethynyl)-2`-deoxyuridine ($^{Py}U$) at their 5`-ends, we developed a rapid fluorescent signalling system. This is an example of an effective fluorescent labelling method for various nucleic acid aptamers with high selectivity. We demonstrated that our strategy could be applied in bioimaging and biosensors. Moreover, such homologous nucleic acid tag could have broader development prospects for research and applications.

This project was funded by the National Natural Science Foundation of China (21432008, 91753201 and 21721005 to X. Z.) and National Research Foundation of Korea (2015M3A9B8029067, 2017R1A2A1A18071086 to B. H. K.).

Conflicts of interest

The authors declare no conflicts of interest.

Notes and references

1. L. Cerchia and V. de Franciscis, *Trends Biotechnol.*, 2010, **28**, 517-525.
2. O. A. Alsager, S. Kumar, B. Zhu, J. Travas-Sejdic, K. P. McNatty and J. M. Hodgkiss, *Anal. Chem.*, 2015, **87**, 4201-4209.
3. Y. Jiang, C. Zhu, L. Ling, L. Wan, X. Fang and C. Bai, *Anal. Chem.*, 2003, **75**, 2112-2116.
4. H. Sun, X. Zhu, P.Y. Lu, R.R. Rosato and W, Tan, Mol. Ther. Nucleic Acids, 2014, **3**, e182.
5. Y. Du, B. Li and E. Wang, Accounts Chem Res, 2012, **46**, 203-213.
6. Mahin Shahdo Rdizadeh, A Rezvan Yazdian-Robati, A Mohammad Ramezani, K. Abnous and S. M. Taghdisi, J. Mater. Chem. B, 2016, **4**, 7766-7778.
7. M. R. Gotrik, T. A. Feagin, A. T. Csordas, M. A. Nakamoto and H. T. Soh, Accounts Chem Res, 2016, **49**, 1903-1910.
8. S. E. Osborne, I. Matsumura and A. D. Ellington, *Curr. Opin. Chem. Biol.*, 1997, **1**, 5-9.
9. C. Tuerk and L. Gold, *Science*, 1990, **249**, 505-510.
10. Y.T. Tang, B. Ge, D. Sen and H. Yu, *Chem. Soc. Rev.*, 2014, **43**, 518-529.
11. M. N. Stojanovic, P. de Prada and D. W. Landry, *J. Am. Chem. Soc.*, 2001, **123**, 4928-4931.
12. S. J. Lee, B. Adler, S. Ekström, M. Rezeli, Á. Végvári, J. Park, J. Malm and T. Laurell, *Anal. Chem.*, 2014, **86**, 7627-7634.
13. J. Li, L. Mo, C. H. Lu, T. Fu, H. H. Yang and W. Tan, *Chem. Soc. Rev.*, 2016, **45**, 1410-1431.
14. R. Savla, O. Taratula, O. Garbuzenko and T. Minko, J. *Control Release*, 2011, **153**, 16-22.
15. E. N. Brody and L. Gold, J. *Biotechnol*, 2000, **74**, 5-13.
16. H. Ma, J. Liu, M. M. Ali, M. A. Mahmood, L. Labanieh, M. Lu, S. M. Iqbal, Q. Zhang, W. Zhao and Y. Wan, *Chem. Soc. Rev.*, 2015, **44**, 1240-1256.
17. L. Jiang and D. J. Patel, *Nat. Struct. Biol.*, 1998, **5**, 769-774.
18. R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe and J. Feigon, *Proc. Natl. Acad. Sci.*, 1993, **90**, 3745-3749.
19. D. Li, S. Song and C. Fan, *Accounts Chem. Res.*, 2010, **43**, 631-641.
20. C. Teller, S. Shimron and I. Willner, *Anal. Chem.*, 2009, **81**, 9114-9119.
21. A. Marquis Gacy, Geoffrey Goellner, Nenad Juranic, Slobodan Macura and C. T. McMurray, *Cell*, 1995, **81**, 533-540.
22. D. C. Crawford, Juan M. Acuña and S. L. Sherman, *Genet. Med.*, 2001, **3**, 359-371.
23. P. Galka-Marciniak, M. O. Urbanek and W. J. Krzyzosiak, *Biol. Chem.*, 2012, **393**, 1299-1315.
24. K. Sobczak, G. Michlewski, M. de Mezer, E. Kierzek, J. Krol, M. Olejniczak, R. Kierzek and W. J. Krzyzosiak, *J. Biol. Chem.*, 2010, **285**, 12755-12764.
25. M. Malgowska, D. Gudanis, R. Kierzek, E. Wyszko, V. Gabelica and Z. Gdaniec, *Nucleic Acids Res.*, 2014, **42**, 10196-10207.
26. Y. Park, K. T. Kim and B. H. Kim, *Chem. Commun.*, 2016, **52**, 12757-12760.
27. A. C. Girvan, *Mol. Cancer Ther.*, 2006, **5**, 1790-1799.
28. P. J. Bates, D.A. Laber, D. M. Miller, S. D.Thomas and J. O.Trent, *Exp. Mol. Pathol.*, 2009, **3**, 151-164.
29. J. Li, H. Zheng, P. J. Bates, T. Malik, X. Li, J. O. Trent and C. K. Ng, *Nucl. Med. Biol.*, 2014, **41**, 179-185.
30. E.J. Joo, G.B. Dam, T. H. van Kuppevelt, T. Toida, R. J. Linhardt and Y. S. Kim, *Glycobiology*, 2005, **15**, 1-9.
31. J. W. Liu and Y. Lu, *Angew.* Chem. Int. Ed., 2006, **118**, 96-100.