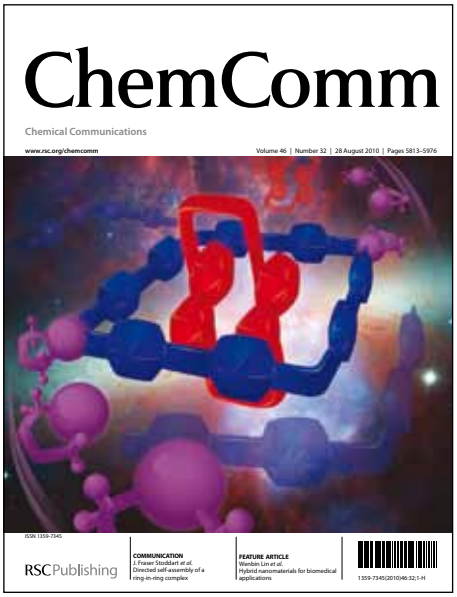
ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the RSC Publishing peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, which is prior to technical editing, formatting and proof reading. This free service from RSC Publishing allows authors to make their results available to the community, in citable form, before publication of the edited article. This *Accepted Manuscript* will be replaced by the edited and formatted *Advance Article* as soon as this is available.

To cite this manuscript please use its permanent Digital Object Identifier (DOI®), which is identical for all formats of publication.

More information about *Accepted Manuscripts* can be found in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics contained in the manuscript submitted by the author(s) which may alter content, and that the standard **Terms & Conditions** and the **ethical guidelines** that apply to the journal are still applicable. In no event shall the RSC be held responsible for any errors or omissions in these *Accepted Manuscript* manuscripts or any consequences arising from the use of any information contained in them.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Quencher-Free Molecular Beacons as Probes for Oligonucleotides Containing CAG Repeat Sequences

Ki Tae Kim,[a] Rakesh N. Veedu,[b] Young Jun Seo[c] and Byeang Hyean Kim*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

We have identified quencher-free molecular beacons that allow the sensitive probing of CAG repeat oligonucleotides, including mRNA fragments of trinucleotide repeat diseases, with significant changes in fluorescence intensity mediated by disruption of the stacking of their ^{Py}U units.

Expansion of the CAG/CTG trinucleotide repeat (TNR) and polyglutamine can result in serious hereditary diseases, including spinocerebellar ataxia, spinal and bulbar muscular atrophy, dentatorubropallidoluysian atrophy, and Huntington's disease.[1] For over a decade, the diagnoses of these genetic disorders have been based on detection tools for CAG/CTG repeat sequences and their expansion. For example, Southern blotting of polymerase chain reaction products,[2a] fluorescence *in situ* hybridization,[2b] and electrochemical sensors[2c] have all been used for the detection of TNRs. Notably, fluorescent probes have been employed widely to detect CAG/CTG repeats and to observe TNR-containing nuclear loci[2b,3]; furthermore, they have recently been used to study structural conversions of TNR sequences based on hybridization between the probe and target TNR sequences.[4] Accordingly, the development of more efficient and sensitive fluorescent probes will encourage further studies into the detection of TNR sequences.

Molecular beacons are very powerful systems for the detection of specific DNA sequences.[5] In a previous study, we developed a quencher-free molecular beacon (Qf-MB) containing only two ^{Py}U units, the fluorescence of which was quenched through self-stacking.[6] This Qf-MB exhibited signal-on fluorescence, with an accompanying hairpin–duplex change, in the presence of a target sequence. Accordingly, we expected this Qf-MB system to also function as an efficient fluorescent probe for TNR sequences. Highly efficient detection might be possible through appropriate ^{Py}U modifications on hairpin structures featuring TNR sequences.

Herein, we report the synthesis of six different types of quencher-free molecular beacons for the CAG trinucleotide repeat. We evaluated their discrimination efficiencies and used them to detect DNA and RNA CAG repeat oligonucleotide, and also mRNA sequences.

First, we set a stem–loop hairpin structure of the $(CTG)_{10}$ repeat as the basic structure for our Qf-MBs to probe the CAG repeat. We designed the Qf-MBs to feature two ^{Py}U units in their sequences; that is, to take advantage of the significant quenching phenomena of pairs of stacked ^{Py}U units. We anticipated that the presence of two ^{Py}U units in symmetric positions in $(CTG)_{10}$ would be preferable because $(CTG)_{10}$ usually forms a hairpin structure (Fig. 1). Based on this hypothesis, we synthesized five Qf-MBs (**P1**–**P5**) having their two ^{Py}U units at different positions within the $(CTG)_{10}$ framework (Table 1). MALDI-TOF mass spectroscopy (Fig. S1, ESI†) confirmed these Qf-MBs and natural oligonucleotides (ODNs).

Next, we investigated the structure of these Qf-MBs by circular dichroism (CD) in the absence and presence of the target sequences **ODN-A** and **ON-A**. The single-strand sequences **P1**–**P5** and **ODN-T** all formed B-form hairpin structures exhibiting positive and negative peaks centered at 285 and 255 nm,

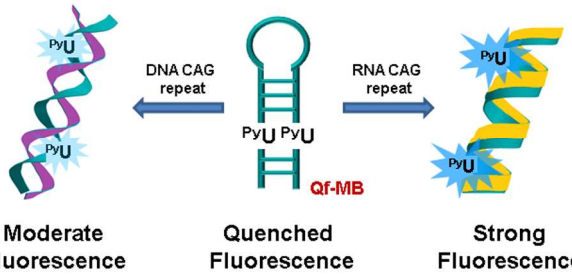
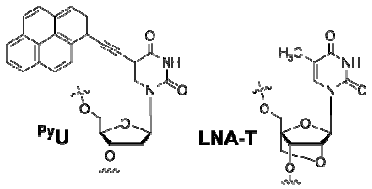


Fig. 1 Conceptual representation of a Qf-MB for the detection of CAG trinucleotide repeats

Table 1 ODN sequences investigated in this study

Name	Sequence (5′ to 3′)
P1	C^{Py}UG CTG CTG CTG CTG CTG CTG CTG C^{Py}UG
P2	CTG C^{Py}UG CTG CTG CTG CTG CTG CTG C^{Py}UG CTG
P3	CTG CTG C^{Py}UG CTG CTG CTG CTG C^{Py}UG CTG CTG
P4	CTG CTG CTG C^{Py}UG CTG CTG C^{Py}UG CTG CTG CTG
P5	CTG CTG CTG CTG C^{Py}UG C^{Py}UG CTG CTG CTG CTG
P3L[a]	CLG CLG C^{Py}UG CLG CLG CLG CLG C^{Py}UG CLG CLG
P3X	CTG CTG C^{Py}UG CTG CTG CTG C^{Py}UG CTG CTG CTG
ODN-A	CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG
ODN-T	CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG
ON-A[b]	cag cag cag cag cag cag cag cag cag cag
ON-U[b]	cug cug cug cug cug cug cug cug cug cug



[a] L = **LNA-T**; [b] RNA sequence

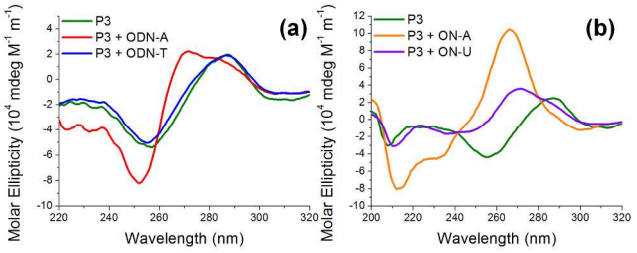


Fig. 2 CD spectra of **P3** in the presence of (a) DNA and (b) RNA CAG or CTG (CUG) repeat sequences [1.5 μM of annealed samples, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C].

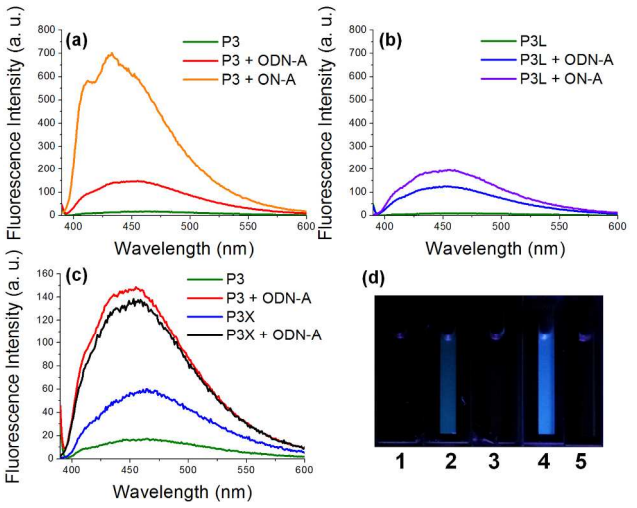


Fig. 3 Fluorescence spectra of (a) **P3**, (b) **P3L**, and (c) **P3X** in the presence of DNA and RNA CAG repeat sequences [1.5 μM of annealed samples, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C]. (d) Fluorescence image of solutions of (1) **P3**, (2) **P3** and **ODN-A**, (3) **P3** and **ODN-T**, (4) **P3** and **ON-A**, and (5) **P3** and **ON-U**).

respectively (Fig. 2 and Fig. S2, ESI†). In the presence of **ODN-A**, we observed a CD signature centered near 272 nm with a shoulder at 290 nm, representing the CAG/CTG duplex, for all of the sequences **P1**–**P5**.[7a] In addition, in the presence of **ON-A**, the CD spectrum was characterized by a dominant positive band at 265 nm and a negative band at 210 nm, corresponding to the CD signature of the A-form duplex.[7b] Furthermore, PAGE experiments provided evidence for the hairpin and duplex structures of these Qf-MBs (Fig. S3, ESI†).

Next, we measured the fluorescence emissions of **P1**–**P5** to evaluate their potential for use as efficient Qf-MBs. All samples were annealed before the measurements. In the presence of **ODN-A**, sequences **P1**–**P5** exhibited 5.7-, 9.4-, 10.6-, 10.5-, and 7.3-fold increases in fluorescence, respectively, but only minor enhancements in the presence of **ODN-T** (Table 2 and Fig. S4, ESI†). In contrast, we observed very strong fluorescence enhancements when these Qf-MBs formed duplexes with the RNA target sequence **ON-A**: 13.5-, 34.7-, 43.5-, 42.6-, and 23.2-fold increases in fluorescence for **P1**–**P5**, respectively. As we had expected, the fluorescence of each of these Qf-MBs was recovered in the full duplex state. We believe that these strong enhancements in fluorescence originated from the low background signals arising from the pairs of well-stacked ^{Py}**U** units at symmetrical positions in the sequences; as evidence, we present the fact that the background signal of **P3X**, the asymmetrically substituted sequence, was higher than that of **P3**

Table 2 Photophysical data of ODNs

ODN[a]	ε_{377}[b]	ε_{399}[b]	QY[c]	B[d]	FE[e]	T_m
P1	6.04	4.72	0.01	746		69
P1+ODN-A	6.34	6.68	0.05	3238	5.7	81
P1+ODN-T	6.24	4.97	0.01	829	1.1	
P1+ON-A	6.08	6.82	0.12	7579	13.5	87
P1+ON-U	5.45	4.33	0.02	945	1.1	
P2	4.91	4.12	0.02	851		70
P2+ODN-A	5.45	5.94	0.11	5933	9.4	80
P2+ODN-T	5.37	4.51	0.02	976	1.2	
P2+ON-A	5.76	7.78	0.34	19,711	34.7	82
P2+ON-U	4.49	3.87	0.02	989	1.1	
P3	4.93	3.99	0.01	682		70
P3+ODN-A	5.55	6.14	0.11	6105	10.6	79
P3+ODN-T	4.51	3.6	0.01	613	0.9	
P3+ON-A	5.87	8.27	0.36	21,291	43.5	81
P3+ON-U	4.45	3.67	0.02	779	1	
P4	4.96	3.99	0.01	701		69
P4+ODN-A	5.91	6.49	0.09	5555	10.5	77
P4+ODN-T	4.96	4.06	0.01	706	1	
P4+ON-A	6.27	8.63	0.32	20,177	42.6	79
P4+ON-U	4.8	3.91	0.02	811	1	
P5	5.45	4.5	0.01	681		55
P5+ODN-A	5.51	5.34	0.07	3763	7.3	78
P5+ODN-T	5.19	4.23	0.01	680	1	
P5+ON-A	6.05	7.86	0.2	12,235	23.2	79
P5+ON-U	4.62	3.92	0.02	809	1	
P3L	4.22	3.33	0.01	384		>83
P3L+ODN-A	5.32	5.82	0.09	4530	16	>87
P3L+ODN-T	4.58	3.63	0.01	628	1.9	
P3L+ON-A	4.47	5.44	0.16	7146	21.6	>93
P3L+ON-U	4.13	3.24	0.01	470	1.1	
ODN-T						60
ODN-T+ODN-A						81
ODN-T+ON-A						84

[a] annealed samples [b] Extinction coefficient (10^4 M^{-1} cm^{-1}). [b] Quantum yield (0.546 for quinine sulfate in 0.5 M H_2SO_4). [d] Brightness (QY × ε_{377}). [e] Fluorescence enhancement (x-fold increase in fluorescence intensity in the presence of the target).

(Fig. 3c). Efficient stacking of the ^{Py}**U** units resulted in the stabilities of the hairpin structures (69–70 °C) of **P1**–**P4** being higher than those of **ODN-T** (60 °C, Table 2, and Fig. S5, ESI†) and in increases in the extinction coefficients of sequences **P1**–**P5** at 399 nm (Table 2 and Fig. S4).

Interestingly, the FEs of the probes **P1**–**P5** were higher in the presence of the RNA sequence **ON-A** than in the presence of the DNA sequence **ODN-A**. Most notably, the FE of **P3** was 43.5 in the presence of **ON-A**, making it a particularly sensitive Qf-MB probe for the RNA CAG repeat. Remarkably, this enhancement in fluorescence was quite visible to the naked eye (Fig. 3d). We suspect that structural distinctions between A- and B-form duplexes were responsible for this difference in fluorescence intensity.[8] For both duplex structures (i.e., with DNA and RNA partners), the ^{Py}**U** units were located in the major groove, the polarity of which in the DNA duplex was higher than that in the DNA/RNA hybrid.[9] We expected the route of weak fluorescence emission via the exciplex of ^{Py}**U** ($Py^{\bullet+}$···$dU^{\bullet-}$) to be more favored in the DNA duplex because the exciplex form is stabilized in polar protic conditions.[10,11] As evidence, we observed different emission peak patterns for **P3** in its DNA duplex (450 nm) and in its DNA/RNA hybrid (407 and 428 nm, Fig. S6a, ESI†). Red-shifted emissions at 450 nm or greater mainly result from exciplex signals,[10] similar to that of monomeric ^{Py}**U** in methanol, a polar protic solvent (Fig. S6b, ESI†). Moreover, we anticipated

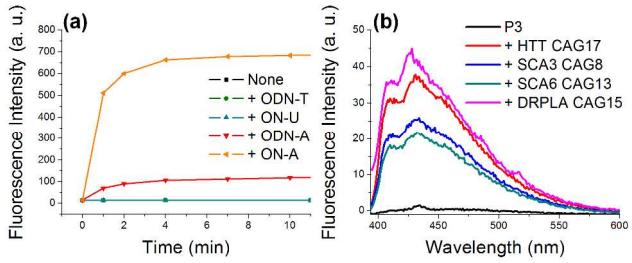


Fig. 4 (a) Time-dependent fluorescence enhancement of the **P3** in the presence of target oligonucleotides [1.5 μM of samples, 1 mL; 25 mM trizma buffer (pH 7.2), 50 mM NaCl, 10 mM $MgCl_2$, 25 °C] and (b) Fluorescence intensity of the **P3** after 1 h. incubation with mRNA sequences from trinucleotide repeat diseases [100 nM of samples, 1 mL; 20 mM trizma buffer (pH 7.2), 80 mM NaCl, 20 mM $MgCl_2$, 37 °C].

that structural restrictions in the DNA/RNA hybrid would provide unsuitable conditions for the formation of stable exciplexes.[11b,12]

In the case of both **ODN-A** and **ON-A**, the probe **P3**, featuring its two ^{Py}**U** units in the middle of the sequence, was the Qf-MB exhibiting the highest efficiency (Table 2), in terms of its good fluorescence quantum yields (QYs; 0.11 and 0.36, respectively), brightness (6105 and 21,291, respectively), and fluorescence enhancement (FEs; 10.6 and 43.5, respectively) while **P1** exhibited the lowest efficiency (Fig. S7, ESI†). Based on **P3**, we synthesized the LNA probe **P3L** containing LNA-T (Table 1) because it has potential to increase the efficiency of the probe due to rigidity and high-affinity hybridization with RNA or DNA.[13,14] As a result, compared with **P3**, **P3L** exhibited a lower FE (decreasing from 43.5 to 21.6) in the presence of **ON-A** while it showed the increased FE (from 10.6 to 16) in the presence of **ODN-A** (Table 2, Fig. 3a and 3b, and Fig. S8, ESI†).

Finally, we evaluated the potential of the Qf-MB under *in vivo*-like conditions (without annealing and non-denatured mRNA). At ambient conditions, **P3** probe detected DNA and RNA CAG target sequences with short time scale (8.4- and 42.1-fold increase in fluorescence within 5 minutes, respectively, Fig. 4a). **P3** also exhibited fluorescence enhancement in the presence of various lengths (longer than 10 repeats) of DNA CAG sequences in the same conditions (Fig. S9, ESI†). Furthermore, we observed significant fluorescence enhancements of the **P3** in the presence of non-denatured mRNA fragments of Huntington disease, Spinocerebellar ataxia type-3, type-6, and Dentatorubropallidoluysian atrophy (37.4, 19.5, 16.2, and 39.7-fold of fluorescence enhancement, respectively, Fig. 4b and Fig. S10, ESI†) under *in vivo*-like conditions.

In conclusion, we have designed and investigated properties of Qf-MBs as probes for the CAG trinucleotide repeat. Among our tested systems, the sequences **P3L** (containing LNA-T residues) and **P3** (containing T residues) were the best at detecting DNA and RNA CAG repeats, respectively. Moreover, **P3** was utilized to detect mRNA fragments of trinucleotide repeat diseases under *in vivo*-like conditions. Thus, we expect that these Qf-MBs can be useful system for detection and study of TNR sequences such as mRNA sequences and longer TNR sequences that induce genetic disorders.

We thank the EPB center program of MEST/NRF (20120000528) and the NRF project of Korea (2012R1A2A2A01047069) for financial support. RNV has been supported by the UQ Research Fellowship Scheme; YJS has been supported by the "Leaders Industry–University Cooperation" Project, itself supported by the Ministry of Education, Science, & Technology (MEST).

Notes and references

[a] *Department of Chemistry, BK School of Molecular Science, Pohang University of Science and Technology, Pohang 790-784, South Korea. Fax: +82 54-279-3399; Tel.: +82 54-279-5856; E-mail: bhkim@postech.ac.kr*

[b] *School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD 4072, Australia. E-mail: rakesh@uq.edu.au; Tel.: +61 7 3365 4611*

[c] *Department of Chemistry, Chonbuk National University, Jeonju 561-756, South Korea. Fax: +82 63-270-3408; Tel.: +82-54-279-2115; E-mail: yseo @ jbnu.ac.kr*

† Electronic Supplementary Information (ESI) available: MALDI-TOF MS data; photophysical properties; melting temperatures of the ODNs.

1 (a) H. T. Orr, M.-Y. Chung, S. Banfi, T. J. Kwiatkowski Jr., A. Servadio, A. L. Beaudet, A. E. McCall, L. A. Duvick, L. P. W. Ranum and H. Y. Zoghbi, *Nature Genet.*, 1993, **4**, 221–226; (b) G. Imbert, F. Saudou, G. Yvert, D. Devys, Y. Trottier, J.-M. Garnier, C. Weber, J.-L. Mandel, G. Cancel, N. Abbas, A. Dürr, O. Didierjean, G. Stevanin, Y. Agid and A. Brice, *Nature Genet.*, 1996, **13**, 285–291; (c) A. R. L. Spada, E. M. Wilson, D. B. Lubahn, A. E. Harding and K. H. Fischbeck, *Nature*, 1991, **352**, 77–79.

2 (a) J. Poirier, K. Ohshima and M. Pandolfo, *Hum. Mutat.*, 1999, **13**, 328–330; (b) C. L. Liquori, K. Ricker, M. L. Moseley, J. F. Jacobsen, W. Kress, S. L. Naylor, J. W. Day and L. P. W. Ranum, *Science*, 2001, **293**, 864–867; (c) M. Fojta, L. Havran, M. Vojtiskova and E. Palecek, *J. Am. Chem. Soc.*, 2004, **126**, 6532–6533.

3 (a) J. Fujimoto, T. Bando, M. Minoshima, S. Uchida, M. Iwasaki, K. Shinohara and H. Sugiyama, *Bioorg. Med. Chem.*, 2008, **16**, 5899–5907; (b) R. Schlapak, H. Kinns, C. Wechselberger, J. Hesse and S. Howorka, *ChemPhysChem* 2007, **8**, 1618-1621; (c) A. Petronis, H. H. Q. Heng, Y. Tatuch, X.-M. Shi, T. A. Klempan, L.-C. Tsui, T. Ashizawa, L. C. Surh, J. J.A. Holden and J. L. Kennedy, *Am. J. Med. Genet.*, 1996, **67**, 85-91.

4 (a) A. Á. Figueroa and S. Delaney, *J. Biol. Chem.*, 2010, **285**, 14648–14657; (b) A. Á. Figueroa, D. Cattie and S. Delaney, *Biochem. Biophys. Res. Commun.*, 2011, **413**, 532–536.

5 S. Tyagi and F. R. Kramer, *Nat. Biotechnol.,* 1996, **14**, 303-308.

6 (a) N. Venkatesan, Y. J. Seo and B. H. Kim, *Chem. Soc. Rev.*, 2008, **37**, 648-663; (b) Y. J. Seo, G. T. Hwang and B. H. Kim, *Tetrahedron Lett.*, 2006, **47**, 4037–4039; (c) G. T. Hwang, Y. J. Seo and B. H. Kim, *J. Am. Chem. Soc.* 2004, **126**, 6528-6529.

7 (a) A. M. Paiva and R. D. Sheardy, *Biochemistry*, 2004, **43**, 14218-14227; (b) J. Kypr, I. Kejnovská, D. Renčiuk and M. Vorlíčková, *Nucleic Acids Res.*, 2009, **37**, 1713–1725.

8 M. A. O'Neill and J. K. Barton, *J. Am. Chem. Soc.*, 2002, **124**, 13053-13066.

9 R. W. Sinkeldam, N. J. Greco and Y. Tor, *ChemBioChem*, 2008, **9**, 706–709.

10 S. T. Gaballah, Y. H. A. Hussein, N. Anderson, T. T. Lian and T. L. Netzel, *J. Phys. Chem. A*, 2005, **109**, 10832–10845.

11 (a) T. L. Netzel, M. Zhao, K. Nafisi, J. Headrick, M. S. Sigman and B. E. Eaton, *J. Am. Chem. Soc.*, 1995, **117**, 9119–9128; (b) A. Trifonov, I. Buchvarov, H.-A. Wagenknecht and T. Fiebig, *Chem. Phys. Lett.*, 2005, **409**, 277–280.

12 C. Boonlua, C. Vilaivan, H.-A. Wagenknecht and T. Vilaivan, *Chem. Asian J.*, 2011, **6**, 3251–3259.

13 H. Kaur, B. R. Babu and S. Maiti, *Chem. Rev.*, 2007, **107**, 4672–4697.

14 (a) U. Wenge, J. Wengel and H.-A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2012, **51**, 10026–10029; (b) K. Kawai, M. Hayashi and T. Majima, *J. Am. Chem. Soc.*, 2012, **134**, 9406–9409.