ChemComm

Coupling oligonucleotides possessing a poly-cytosine tag with magnetic ionic liquids for sequence-specific DNA analysis

Journal:	*ChemComm*
Manuscript ID	CC-COM-07-2018-005954.R1
Article Type:	Communication

ChemComm

COMMUNICATION

Coupling oligonucleotides possessing a poly-cytosine tag with magnetic ionic liquids for sequence-specific DNA analysis†

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

DOI: 10.1039/x0xx00000x

www.rsc.org/chemcomm

Xitian Peng [ab], Kevin D. Clark [b], Xiong Ding [b], Chenghui Zhu [b], Marcelino Varona [b], Miranda N. Emaus [b], Jiwoo An [b], Jared L. Anderson *[b]

Oligonucleotide probes were designed with a poly-cytosine region that facilitates stable anchoring to a magnetic ionic liquid support. By tethering a recognition sequence to the poly-C tag, the resulting diblock oligonucleotides distinguished single-nucleotide variants and captured DNA targets from interfering genomic DNA and cell lysate for qPCR amplification.

Molecular recognition chemistry plays a key role in biotechnology.[1, 2] Single-stranded DNA (ss-DNA) has the unique ability to pair with their complementary sequence by Watson-Crick hydrogen bonding.[3] Based on this specific molecular recognition mechanism, DNA has been widely exploited to construct DNA-material conjugates for chemical, biological, and medical applications.[4-6] One core challenge for such applications lies in designing simple and cost-effective methods to stabilize and attach DNA to substrates without destabilizing base-pairing interactions.[7] Thiol or amino modified DNA is commonly used to form stable conjugates on gold or carboxyl-terminated surfaces, respectively.[8,9] However, chemical modification of DNA may be time intensive and require the use of expensive reagents, leading to tedious separation processes and high cost.[5]

As an alternative, physisorption may provide another simple and cost-effective method for immobilization of DNA on the surface of a support material. For example, the affinity of consecutive adenine (poly-A) oligonucleotides for gold nanoparticles (AuNPs) has been studied systematically and utilized to design DNA-AuNP conjugates.[7, 10-13] Recently, consecutive cytosine (poly-C) oligonucleotides were found to serve as strong ligands to certain inorganic nanomaterials.[14, 15] A major advantage of employing DNA homopolymers as ligands for inorganic materials is the low cost of synthesizing oligonucleotides with high purity, which largely avoids the complex/expensive chemical modifications of oligonucleotide.[7]

A variety of solid materials such as magnetite, silica, and gold are commonly used as supports for DNA immobilization. Among them, magnetite-based materials are extremely popular in separation science due to their ease of manipulation by an external magnetic field. Using this platform, specific DNA sequences can be extracted by ss-DNA probes immobilized on the surface of magnetic active beads to facilitate rapid isolation from complex biological samples.[16] However, these substrates may suffer from aggregation over time resulting in lower capture efficiency and obstruction of liquid handing devices.[17, 18] Furthermore, the hybridization dynamics of DNA on the solid–liquid interface may be slower than in solution.[19] As a result, long incubation/agitation times are required to extract a sufficient quantity and quality of nucleic acid. The use of a liquid support has the potential to address the aforementioned limitations of solid support substrates. Magnetic ionic liquids (MILs) have been applied as alternatives to magnetic bead-based approaches for the extraction of nucleic acids from biological samples.[20, 21] Similar to traditional ionic liquids, MILs possess unique physicochemical properties that can be tuned by changing the combination of cations and anions while also exhibiting susceptibility to applied magnetic fields. These features render MILs useful solvents for magnet-based separations while avoiding the time-consuming centrifugation processes in traditional LLE or particle aggregation in magnetic bead/particle systems.[22-24]

Herein, we report the discovery of a MIL support that exhibits strong affinity for poly-C sequences compared to random DNA sequences. Diblock DNA oligonucleotides were designed with a poly-C block for anchoring the DNA probe to the MIL and a sequence recognition block to hybridize with target DNA sequences. The diblock DNA-MIL conjugate has the potential to distinguish between complementary DNA and single nucleotide variants. Furthermore, the poly-C DNA-MIL method was employed for the sequence-specific extraction of target DNA from a solution of interfering genomic DNA and cell lysate with real-time quantitative polymerase chain

[a] Institute of Agricultural Quality Standards and Testing Technology Research, Hubei Academy of Agricultural Science, Wuhan, 430064, China
[b] Department of Chemistry, Iowa State University, Ames, IA 50011, USA. E-mail: andersoj@iastate.edu; Tel: +1 515-294-8356
†Electronic Supplementary Information (ESI) available: Regents and instruments, detailed experimental methods and additional data. See DOI: 10.1039/x0xx00000x

reaction (qPCR) amplification for detection, demonstrating practical application of the poly-C DNA-MIL approach for targeted DNA analysis.

Due to their high hydrophobicity and low viscosity, two metal-based MIL supports, named as trihexyl(tetradecyl)phosphonium cobalt(II) hexafluoroacetylacetonate ($[P_{66614}^+][Co(hfacac)_3^-]$) and trihexyl(tetradecyl)phosphonium nickel(II) hexafluoroacetylacetonate ($[P_{66614}^+][Ni(hfacac)_3^-]$) (Fig. 1a),[25] were selected to extract a series of random ss-DNA molecules with different GC content (20%, 80%) and secondary structures (linear or hairpin) (see Table S1 in ESI). The extraction efficiency of ss-DNA sequences from aqueous solution into the MIL was evaluated by HPLC-UV detection, in which the amount of DNA remaining after extraction was compared to the initial amount of DNA in solution. No unique selectivity of the Ni(II)-based MIL for DNA was observed as demonstrated by the high extraction efficiencies for all DNA sequences, except Oligo 6 (ESI, Fig. S1). On the other hand, the Co(II)-based MIL exhibited extraction efficiencies lower than 50% and 30% for DNA sequences with 20% GC content (Oligo 1, 2, 3) and hairpin sequences with 80% GC content (Oligo 7, 8, 9), respectively (ESI, Fig. S1). Interestingly, greater than 80% extraction efficiency of linear DNA sequences with 80% GC content (Oligo 4, 5) was observed, compared to relatively poorer extraction efficiency of Oligo 6 (about 45%). Comparing the structures of the linear ss-DNA sequences with 80% GC content revealed that Oligo 6 possessed just four consecutive cytosine nucleotides while Oligo 4 and 5 contained nine consecutive cytosines. These results suggested that the Co(II)-based MIL exhibited stronger affinity for sequences with extended poly-C domains.

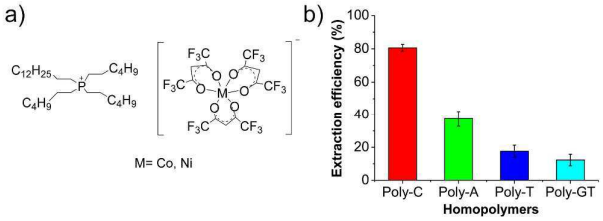


Fig. 1a) Chemical structure of the hydrophobic Co(II) and Ni(II)-based MILs examined in this study; b) Extraction efficiencies of four 20-mer homopolymer DNA sequences using the Co(II)-based MIL. A solution containing 0.44 µM ss-DNA in 50 µL DI H_2O was extracted using 1 µL of the $[P_{66614}^+][Co(hfacac)_3^-]$ MIL with a 30 s extraction time and a vortex rate of 2000 rpm.

To further investigate the high affinity of the Co(II)-based MIL for poly-C sequences, the extraction of 20-mer ss-DNA homopolymers poly-C, poly-A, poly-T, and poly-GT (Table S1) was performed. Here, poly-G was not examined due to the easy formation of G-quadruplex, which may influence the analytical results. As shown in Fig. 1b, the Co(II)-based MIL possessed higher affinity for 20-mer poly-C DNA with more than 80% extraction efficiency compared to less than 40% extraction efficiencies for the other three 20-mer sequences. These results further confirmed that poly-C exhibits the highest affinity for the Co(II)-based MIL.

DNA can associate with certain inorganic materials through hydrogen bonding, π–π stacking and electrostatic interactions. However, little is known about the interactions between DNA and MILs. A series of experiments were performed to study the interaction between the Co(II)-based MIL and poly-C DNA. Firstly, 20 bp double-stranded DNA (ds-DNA) with 80% GC content was extracted using the MIL solvent. The result showed that less than 15% of the DNA duplex was extracted by the Co(II)-based MIL (ESI, Fig. S2), which was much lower than their ss-DNA counterparts (Oligo 4, 5 in Table S1) that were extracted with more than 80% efficiency. This observation was conceivably due to the inaccessibility of the nucleobases in duplex DNA, decreasing π-π stacking and hydrophobic interactions between DNA and MIL, which has also been observed for graphene oxide (GO) based materials.[26] Next, aqueous solutions containing urea and/or NaCl were used to study the recovery of poly-C DNA from the MIL phase. The results indicated that 8 M urea combined with 1 M NaCl could elute more than 70% poly-C DNA from the MIL (ESI, Fig. S3). While these data suggest that hydrogen bonding and ionic interactions also play a role in the extraction of poly-C DNA by the MIL, they do not explain the unusual affinity of poly-C DNA for the Co(II)-based MIL. In a previous report, the affinity of poly-C to carbon-based materials (i.e., GO, SWNT) was mediated by the formation of the i-motif DNA structure, which increased π-π stacking between the nucleic acid and carbon material.[15, 27, 28] In order to determine whether this unique tertiary structure influenced affinity for the MIL phase, we tested the extraction of free human telomeric DNA (ESI, Table S1) that can form the i-motif structure in the presence of carbon-based materials. Only approximately 40% of human telomeric DNA could be extracted (ESI, Fig. S4), indicating that the preferential extraction of poly-C by MIL was likely not due to i-motif formation. Overall, the affinity of poly-C sequences for the Co(II)-based MIL solvent may involve hydrogen bonding, π–π stacking, and ionic interactions facilitated by sequential cytosine nucleobases.

The high selectivity of the Co(II)-based MIL for poly-C DNA suggested that a poly-C block could be incorporated within a probe sequence to serve as an anchor to the hydrophobic MIL and facilitate sequence-specific DNA extraction. Diblock DNA probes composed of a poly-C sequence (Oligo 4, 5, and C_{20}, C_{10}, C_0) for surface anchoring onto the MIL and a 20-mer probe sequence for hybridization with complementary target DNA were designed (Table 1 and S1). As shown in ESI (Fig. S5), DNA probes with poly-C blocks can be extracted by the MIL with extraction efficiencies greater than 80%. The poly-C sequence with 20 cytosine nucleotides provided the highest extraction efficiency. To further study the length of poly-C sequence on the affinity for the MIL, a fluorophore-labeled DNA sequence (FAM-DNA) was extracted by the MIL. Non-complementary poly-C DNA sequences with 0, 10 and 20 cytosine nucleobases as well as a sequence complementary (cDNA) to the pre-loaded FAM-DNA were then added to solution to study the desorption of the extracted FAM-DNA. A scheme for this process is shown in the ESI (Fig. S6). The fluorescence signal

was detected after incubation with one of the aforementioned sequences for 30 min. As shown in Fig. 2a, the fluorescence intensity increased after incubation with complementary DNA, probably due to the formation of duplex DNA that has lower affinity for the MIL than ss-DNA. For the poly-C DNA sequences with 0, 10 and 20 cytosines, the fluorescence intensity increased as the DNA length increased. The fluorescence intensity with C_{20} sequence was surprisingly higher than the fluorescence obtained when using complementary DNA to desorb the FAM-labeled DNA. These results showed that a longer poly-C DNA was highly efficient in displacing DNA from the MIL, demonstrating a strong affinity of poly-C sequences for the MIL. Therefore, the poly-C probe with 20 cytosines was chosen as the probe for sequence-specific DNA capture.

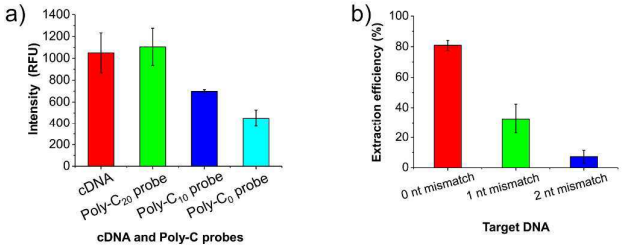


Fig. 2a) Fluorescence intensity of FAM-labeled DNA after incubation with MIL and subsequent desorption by the addition of complemetary (cDNA) and poly-C sequences with different lengths (non-complementary). A 0.1 μM solution of FAM-labeled DNA was incubated with 1 μL of MIL in 20 mM Tris-HCl, pH 8.0, with 25 mM NaCl. A molar equivalent of cDNA or non-cDNA was added to desorb FAM-DNA for 30 min. b) Extraction efficiencies of 20-mer complementary target DNA and its 1 or 2 nt mismatched DNA using poly-C DNA-MIL conjugate by the load first procedure. Extraction conditions: DNA concentration: 0.44 μM; mole ratio of poly-C probe to target: 1:1; total solution volume: 50 μL; time: 1 min; MIL volume: 1 μL; rotation rate: 2000 rpm.

Two different approaches for sequence-specific DNA extraction were employed in this study, as shown in Figure 3: (1) the probe sequence was first loaded onto the MIL and then used to capture a complementary target sequence (Target 1) or (2) the probe and target were hybridized first, followed by addition of the MIL to bind the probe–target duplex.[21] The load first procedure extracted more than 80% of target 1 DNA compared to the hybridize first procedure which yielded an approximate 50% extraction efficiency (ESI, Fig. S7). In order to determine the amount of non-specific DNA extraction, ss-DNA target and its corresponding duplex were subjected to MIL-based extraction without the use of poly-C probes. Only about 25% target ss-DNA and less than 5% duplex DNA partitioned to the MIL (ESI, Fig. S8),

Table 1 Poly-C probe sequences and target sequences used in this study[a]

Name	Sequence
Poly-C_{20} probe	$(C)_{20}$ CAC GCT TAC ATT CAC GCC CT
Poly-C_{10} probe	$(C)_{10}$CAC GCT TAC ATT CAC GCC CT
Poly-C_0 probe	CAC GCT TAC ATT CAC GCC CT
Target 1	AGG GCG TGA ATG TAA GCG TG
Target 2[b]	AGG GCG TCA ATG TAA GCG TG
Target 3[b]	AGG GCG TCA ATG TAA GCC TG

[a] All the sequences are listed from the 5′ to 3′-end.

[b] The underlined nucleobases represent mismatches to probe sequence.

which further confirmed that the extraction of target DNA was greatly enhanced by the use of the poly-C diblock DNA probes. Since the hybridize first procedure forms partially duplexed DNA before partitioning into the MIL support, lower extraction efficiencies were observed due to the lower affinity of the MIL support for duplex DNA.

In order to investigate the selectivity of poly-C DNA-MIL conjugate for nucleotide mismatches, we tested the extraction of four other non-complementary random DNA sequences with GC contents ranging from 20% to 80% (Oligo 1, Oligo 6, Oligo 9 and poly-C_0 probe). Less than 5% of these non-complementary random DNA molecules could be extracted, indicating that the diblock poly-C DNA-MIL conjugate has low non-specific extraction of other random non-complementary DNA sequences. To simulate a real sample that often contains sequences highly similar to the DNA target, we investigated the extraction of oligonucleotides with one or two nucleobase variations of the target sequence (Target 2 and 3 in Table 1). As shown in Fig. 2b, the DNA–MIL conjugate extracted over 80% of a complementary sequence, while the one-base and two-base mismatched DNA were extracted with 33% and 7% efficiency, respectively (ESI, Fig. S9). These results show that the diblock DNA-MIL conjugate possesses the necessary selectivity to distinguish between oligonucleotides that differ by one nucleobase.

Since the poly-C probe sequences were anchored within the MIL phase via non-covalent interactions, the anchoring DNAs may be susceptible to nonspecific displacement by other molecules.[14] To test the adsorption stability of the poly-C DNA-MIL conjugate, several competing macromolecules such as proteins (bovine serum albumin, albumin from chicken egg white) and non-ionic surfactants (Tween 20, Pluronic F-108) were incubated with poly-C DNA-MIL conjugates in aqueous solution for 30 min and the fraction of poly-C DNA released from the MIL was evaluated. As shown in Fig. S10, less than 1% poly-C probe DNA was released from the MIL when treated with solutions containing a 1% concentration (w/v) of those interfering macromolecules. These results indicate good stability of the present poly-C DNA-MIL conjugates for practical application in biological samples.

In order to interface the poly-C DNA-MIL based extraction method with qPCR amplification, a 261 bp target DNA with a terminal segment fully complementary to the 20-mer recognition block of the poly-C probe was selected as a model sequence. The poly-C DNA probe was first hybridized with the target DNA and then MIL was added to extract the target DNA (Fig. 3). According to the quantification cycle (Cq) values, the amplification of target DNA was significantly enhanced by the poly-C DNA-MIL method (Cq=24.7±1.0, n=4) compared to a MIL-based extraction method without a poly-C probe (Cq=32.9+±0.8, n=4). The Cq value of poly-C DNA-MIL extraction was 8.2 cycles lower than the direct extraction method, indicating that an approximate 300-fold greater amount of target DNA was extracted with the poly-C DNA-MIL method. To evaluate the effect of matrix components on DNA

extraction using the poly-C DNA-MIL approach, the extraction of 261 bp target DNA from *E. coli* cell lysate or DNA from salmon testes was performed. According to the Cq values in the presence of *E. coli* cell lysate (Cq=20.8±1.0, n=4) or DNA from salmon testes (Cq=21.0±0.5, n=4, ESI, Fig. S11 and S12), a matrix-enhancement effect was observed, meaning that more target DNA was extracted in the presence of the sample matrix components. The results show great practical application of the poly-C DNA-MIL approach for targeted DNA analysis.

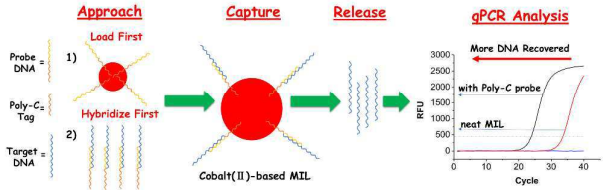


Fig. 3 Schematic illustration describing the load first and hybridize first approaches examined in this study. To achieve sequence selective extraction of DNA, the poly-C DNA probe was hybridized with target DNA and then captured by the MIL. DNA was then desorbed from the MIL and compared to a direct extraction with neat MIL (containing no poly-C probe) using qPCR amplification. Target DNA: 1.69 fmol; poly-C_{20} probe: 16.9 pmol; extraction time: 1 min; desorption time: 10 min; desorption method: 25 mM NaCl at 60 °C for 10 min; desorption solvent volume: 50 µL.

In conclusion, we discovered that poly-C DNA sequences exhibit much stronger affinity for a Co(II)-based MIL than other random DNA sequences, providing the basis for a sequence-specific DNA extraction method that does not rely on complex/expensive oligonucleotide modifications. The diblock poly-C DNA-MIL conjugates demonstrated excellent hybridization efficiency for target DNA and high selectivity to distinguish DNA targets from sequences possessing single nucleotide mismatches. Furthermore, the poly-C DNA-MIL conjugate exhibited good stability when treated with several interfering macromolecules such as proteins and non-ionic surfactants. According to qPCR results, nearly a 300-fold greater amount of target DNA could be selectively extracted from aqueous solution using the poly-C DNA-MIL method compared to a direct extraction method without the poly-C probe. The poly-C DNA-MIL approach was also successfully employed for the extraction of target DNA from cell lysate and a solution of DNA from salmon testes. Our study establishes a simple, cost-effective strategy for binding a DNA probe sequence to a MIL support, thereby avoiding the complex chemical DNA modification process and making this approach particularly attractive for sequence-specific DNA analysis.

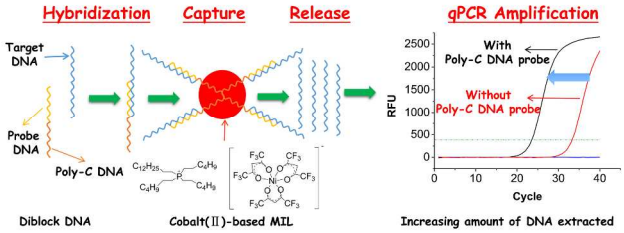
The authors acknowledge funding from the Chemical Measurement and Imaging Program at the National Science Foundation (Grant No. CHE-1709372) and Hubei Academy of Agricultural Science Visiting Scholar Program.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. H. Pei, X. Zuo, D. Zhu, Q. Huang and C. Fan, *Acc. Chem. Res.*, 2014, **47**, 550-559.
2. P. Kohli, C. C. Harrell, Z. H. Cao, R. Gasparac, W. H. Tan and C. R. Martin, *Science*, 2004, **305**, 984-986.
3. J. L. Sessler, C. M. Lawrence and J. Jayawickramarajah, *Chem. Soc. Rev.*, 2007, **36**, 314-325.
4. L. Ding, C. L. Xiang and G. Zhou, *Talanta*, 2018, **181**, 65-72.
5. A. D. Chowdhury, N. Agnihotri, R. A. Doong and A. De, *Anal. Chem.*, 2017, **89**, 12244-12251.
6. M. R. Jones, N. C. Seeman and C. A. Mirkin, *Science*, 2015, **347**.
7. H. Pei, F. Li, Y. Wan, M. Wei, H. Liu, Y. Su, N. Chen, Q. Huang and C. Fan, *J. Am. Chem. Soc.*, 2012, **134**, 11876-11879.
8. J. M. Carnerero, A. Jimenez-Ruiz, P. M. Castillo and R. Prado-Gotor, *ChemPhysChem*, 2017, **18**, 17-33.
9. N. Mohanty and V. Berry, *Nano Letters*, 2008, **8**, 4469-4476.
10. A. Opdahl, D. Y. Petrovykh, H. Kimura-Suda, M. J. Tarlov and L. J. Whitman, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 9-14.
11. S. M. Schreiner, D. F. Shudy, A. L. Hatch, A. Opdahl, L. J. Whitman and D. Y. Petrovykh, *Anal. Chem.*, 2010, **82**, 2803-2810.
12. X. Zhang, B. Liu, N. Dave, M. R. Servos and J. Liu, *Langmuir*, 2012, **28**, 17053-17060.
13. D. Zhu, P. Song, J. Shen, S. Su, J. Chao, A. Aldalbahi, Z. Zhou, S. Song, C. Fan, X. Zuo, Y. Tian, L. Wang and H. Pei, *Anal. Chem.*, 2016, **88**, 4949-4954.
14. Z. Huang and J. Liu, *Langmuir*, 2018, **34**, 1171-1177.
15. C. Lu, Z. Huang, B. Liu, Y. Liu, Y. Ying and J. Liu, *Angew. Chem. Int. Ed.*, 2017, 56, 6208-6212.
16. K. D. Clark, C. Zhang and J. L. Anderson, *Anal. Chem.*, 2016, **88**, 11262-11270.
17. D. Liu, G. Liang, Q. Zhang and B. Chen, *Anal. Chem.*, 2013, **85**, 4698-4704.
18. J. A. Davis, D. W. Inglis, K. J. Morton, D. A. Lawrence, L. R. Huang, S. Y. Chou, J. C. Sturm and R. H. Austin, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 14779-14784.
19. M. R. Henry, P. Wilkins Stevens, J. Sun and D. M. Kelso, *Anal. Biochem.*, 1999, **276**, 204-214.
20. M. N. Emaus, K. D. Clark, P. Hinners and J. L. Anderson, *Anal. Bioanal. Chem*, 2018, **410**, 4135-4144.
21. K. D. Clark, M. Varona and J. L. Anderson, *Angew. Chem. Int. Ed.*, 2017, **56**, 7630-7633.
22. K. D. Clark, J. A. Purslow, S. A. Pierson, O. Nacham and J. L. Anderson, *Anal. Bioanal. Chem*, 2017, **409**, 4983-4991.
23. K. D. Clark, M. M. Yamsek, O. Nacham and J. L. Anderson, *Chem. Commun.*, 2015, **51**, 16771-16773.
24. K. D. Clark, O. Nacham, H. L. Yu, T. H. Li, M. M. Yamsek, D. R. Ronning and J. L. Anderson, *Anal. Chem.*, 2015, **87**, 1552-1559.
25. S. A. Pierson, O. Nacham, K. D. Clark, H. Nan, Y. Mudryk and J. L. Anderson, *New J. Chem.*, 2017, **41**, 5498-5505.
26. P. J. J. Huang and J. W. Liu, *Nanomaterials*, 2013, **3**, 221-228.
27. X. Li, Y. Peng, J. Ren and X. Qu, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 19658-19663.
28. Y. Peng, X. Wang, Y. Xiao, L. Feng, C. Zhao, J. Ren and X. Qu, *J. Am. Chem. Soc.*, 2009, **131**, 13813-13818.



Oligonucleotide probes containing a poly-cytosine region have remarkable affinity to a magnetic ionic liquid support and are used to distinguish single-nucleotide variants and capture DNA targets for qPCR amplification.