ChemComm

Messenger RNA Enrichment Using Synthetic Oligo(T) Click Nucleic Acids

Journal:	*ChemComm*
Manuscript ID	CC-COM-08-2020-005815.R1
Article Type:	Communication

COMMUNICATION

Messenger RNA Enrichment Using Synthetic Oligo(T) Click Nucleic Acids

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

DOI: 10.1039/x0xx00000x

Alex J. Anderson[1], Heidi R. Culver[1], Tania R. Prieto[1], Payton J. Martinez[1], Jasmine Sinha[1], Stephanie J. Bryant[1,2,3], Christopher N. Bowman[1,2,3]

Enrichment of mRNA is a key step in a number of molecular biology techniques, particularly in the rapidly growing field of transcriptomics. Currently, mRNA is isolated using oligo(thymine) DNA (oligo(dT)) immobilized on solid supports, which binds to the poly(A) tail of mRNA to pull the mRNA out of solution through the use of magnets or centrifugal filters. Here, a simple method to isolate mRNA by complexing it with synthetic click nucleic acids (CNAs) is described. Oligo(T) CNA bound efficiently to mRNA, and because of the insolubility of CNA in water, >90% of mRNA was readily removed from solution using this method. Simple washing, buffer exchange, and heating steps enabled mRNA's enrichment from total RNA, with a yield of 3.1 ± 1.5% of the input total RNA by mass, comparable to the yield from commercially available mRNA enrichment beads. Further, the integrity and activity of mRNA after CNA-facilitated pulldown and release was evaluated through two assays. *In vitro* translation of EGFP mRNA confirmed the translatability of mRNA into functional protein and RT-qPCR was used to amplify enriched mRNA from total RNA extracts and compare gene expression to results obtained using commercially available products.

Several molecular biology techniques benefit from or require enrichment of messenger RNA (mRNA), including quantitative polymerase chain reaction (qPCR), *in vitro translation* experiments, and next-generation sequencing of RNA (RNA-seq). In such assays, mRNA provides information about gene expression and abnormal gene fusions to better understand different cellular processes and disease states.[1–7] In particular, mRNA provides information about the protein-coding regions of the transcriptome.[8–10] Amplification of mRNAs in qPCR is also a useful diagnostic technique, especially for identifying overexpressed proteins in breast, prostate, and other types of cancers.[6,11–13] Enrichment of mRNA is achieved using commercial kits containing oligo(dT) immobilized on solid supports (e.g., magnetic beads or cellulose) to concentrate poly(A) mRNA selectively, which makes up just 1-5% of total RNA in the cell.[14–18] While this method is effective, it was anticipated that click nucleic acid (CNA) oligonucleotides, a newly developed type of xenonucleic acid (XNA), could be exploited in an alternative, simple strategy for mRNA isolation.

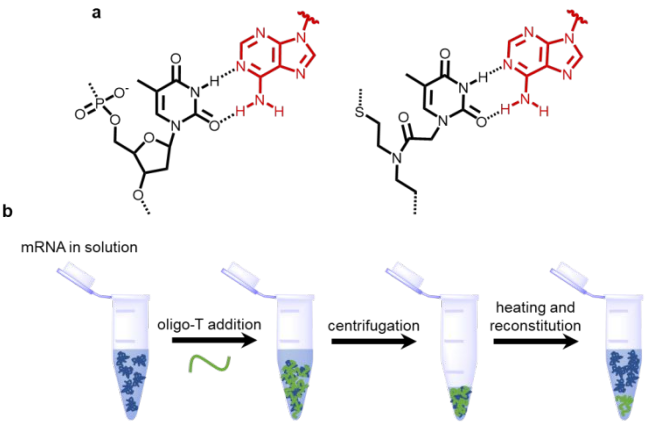


Figure 1 – (a) Comparison of a natural DNA repeat unit to the CNA repeat unit. The thioether backbone removes backbone charge, but the 6-atom spacing allows for binding of complementary nucleic acids. (b) General process of mRNA isolation procedure. The oligo(T) CNA binds to and helps precipitate mRNA in solution and the mRNA can be released through a heating and reconstitution step.

[1]*Department of Chemical and Biological Engineering, University of Colorado, Boulder, CO 80303*
[2]*Material Science and Engineering Program, University of Colorado, Boulder, CO 80303*
[3]*BioFrontiers Institute, University of Colorado, Boulder, CO 80303*
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

The CNAs used in this work (**Figure 1a**) have a six atom spacing per repeat unit similar to DNA/RNA, facilitating binding to native DNA/RNA.[19–21] While there are reports of using other types of modified nucleic acids to selectively isolate mRNA,[22,23] two key attributes give CNA oligonucleotides a distinct advantage over these materials for such applications. Primarily, the thiol-ene "click" reaction used to polymerize CNAs is ideal for producing mononucleotide repeat sequences, particularly oligo(thymine) (oligo(T)), which is synthesized and purified at the hundreds-of-milligram to gram scale in a few hours. Conversely, solid phase synthesis of DNA, RNA, or other xeno nucleic acid (XNA) oligonucleotides suffers from low yields, increased reaction times, and step-wise synthetic approaches that fail to take advantage of the repeating character of the poly(A) sequence.

In addition, CNA oligonucleotides are attractive for mRNA enrichment due to their nonpolar, thioether backbone. Unlike native nucleic acids, which have good solubility in water because of their negatively charged phosphodiester backbone, non-functionalized CNAs are water insoluble. Despite this insolubility, stable CNA-DNA or CNA-RNA hybrids form if CNA oligonucleotides are dissolved in DMSO first and then added to an aqueous solution containing native nucleic acids.[20] Another benefit of the non-ionic CNA backbone is improved hybrid stability, even at low ionic strength,[20] relative to native DNA-DNA or DNA-RNA duplexes, which must overcome electrostatic repulsion between strands to hybridize. Therefore, it was hypothesized that mRNA enrichment could be achieved by a simple process exploiting the insolubility of CNA along with its ability to bind complementary RNA (**Figure 1b**). Herein, conditions under which mRNA is pulled out of solution and subsequently released from oligo(T) CNA are identified and the functionality of mRNA enriched via this method is demonstrated in two molecular assays.

Oligo(T) CNA was synthesized as previously described.[19,20] The number average degree of polymerization for the CNAs used in this work was 16 ± 3 with a polydispersity of 1.5 ± 0.2 (**Figure S1**). Because CNA solubility depends on length, it was anticipated that CNA precipitation and, consequently, mRNA pulldown efficiency would be affected by buffer composition. To test this hypothesis, the ratio of DMSO to water was adjusted to balance CNA solubilization and RNA binding. Pulldown efficiency for an A_{20} RNA oligonucleotide was found to be highest (95 ± 1%) at the lowest DMSO concentrations tested (<20 vol%) (**Figure S2**). This finding is favorable for downstream applications of the isolated mRNA, such as RT-PCR, because DMSO is well-tolerated, even favorable, when included at low concentrations (<10 vol%).[24,25] As such, all subsequent studies were performed using 5 vol% DMSO.

Next, the effects of salt concentration were examined. Methods relying on immobilized oligo(dT), which are densely packed on solid supports, require optimized ionic strength to maximize pulldown efficiency.[26] Sufficiently high cation concentrations are needed to enable hybridization between the oligo(dT) and poly(A).[26–28] Further, salt content has been shown to influence the structure and stability of target RNA in solution.[29] Compared to other monovalent cations, Li+ ions are less effective at precipitating DNA,[30] allowing for selective precipitation of RNA over DNA, and have better solubility in organic cosolvents, including DMSO. In a buffer comparable to that used in a commercially available mRNA isolation kit (Dynabeads® mRNA DIRECT™ kit, 10 mM Tris pH 7.5, 1M LiCl, 1 mM EDTA), 93% of mRNA was pulled out of solution. Unlike methods relying on oligo(dT), it is noteworthy that mRNA pulldown with CNA was also possible in the absence of salts, albeit with lower efficiency (70%) (**Figure S3**). All subsequent pulldown studies were carried out using this LiCl containing buffer (Buffer B, **Table S1**).

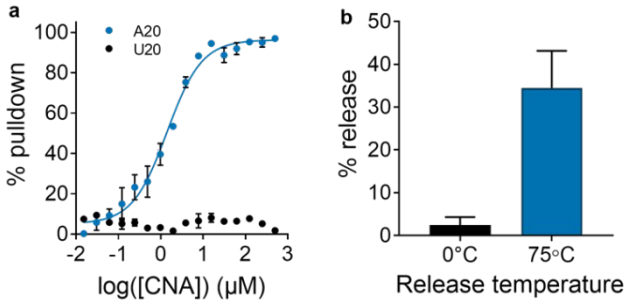


Figure 2 – Pulldown of A_{20} RNA as a function of oligo(T) CNA concentration. Oligo(T) CNA at sufficiently high concentrations achieved >90% pulldown of complementary RNA while effectively no pulldown was observed regardless of concentration for non-complementary sequences. (b) Release of RNA is achieved by heating samples to 75°C to dissociate the hybridization between CNA and RNA. Data is represented as the mean of at least 3 replicates and error bars represent standard deviations.

To begin optimizing the CNA concentration required to isolate RNA, the concentration of CNA was varied while keeping an A_{20} RNA concentration fixed at 125 nM. Greater than 75% pulldown efficiency was achieved using oligo(T) CNA concentrations as low as 7.5 μM (30 μg/mL) (**Figure 2a**). When a non-complementary RNA oligonucleotide, U_{20}, was used in place of the complementary A_{20} oligonucleotide, no pulldown was observed (**Figure 2a**), demonstrating that CNA-mediated pulldown of RNA oligonucleotides was a result of sequence-specific hybridization rather than non-specific, hydrophobic interactions. As with traditional mRNA isolation procedures, RNA was released by heating (**Figure 2b**).

Higher CNA concentrations were required to precipitate EGFP mRNA compared to short RNA oligonucleotides used in the prior study, which is attributed to its larger size (996 nucleotides) and increased hydrophilicity. Still, it was found that greater than 75% of mRNA was removed using a CNA concentration of at least 125 μM (500 μg/mL) (**Figure 3a**). For even higher CNA concentrations (150-250 μM), between 90-92% pulldown efficiency was consistently achieved for EGFP mRNA concentrations over the range of 30-90 ng/mL (**Figure 3b**). Furthermore, even when relatively high concentrations of mRNA were tested (2 – 64 μg/mL), the pulldown efficiency remained above 90% (**Figure S4**).

Having demonstrated efficient pulldown, it was necessary to demonstrate that the released mRNA remained functional, ultimately confirming the utility of CNA as an mRNA enrichment tool. To do so, an *in vitro* translation (IVT) kit was used to translate EGFP mRNA recovered after pulldown and release

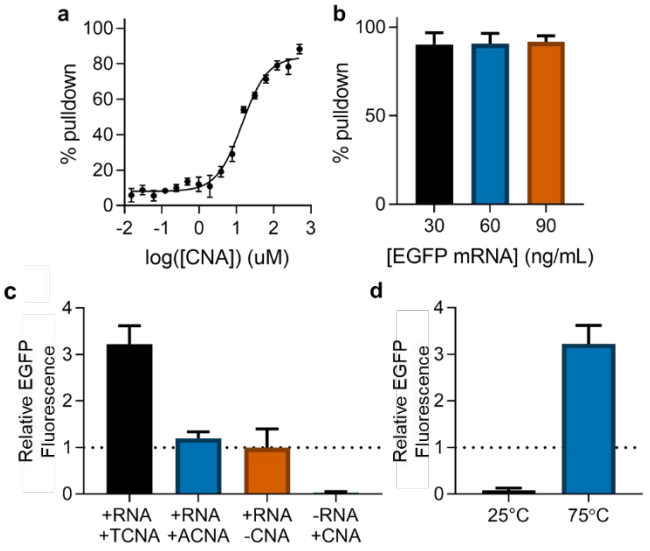


Figure 3 – (a) Effective pulldown of fluorescent EGFP mRNA was achieved at CNA concentrations of 125 µM and higher. (b) Under optimized conditions, greater than 90% pulldown efficiency could be achieved at biologically relevant mRNA concentrations. (c) In vitro translation of enriched mRNA reveals that complementary CNA is needed for isolation to occur. (d) Enrichment of mRNA is only possible after the heating and reconstitution step. Data is represented as the mean of at least 3 replicates and error bars represent standard deviations.

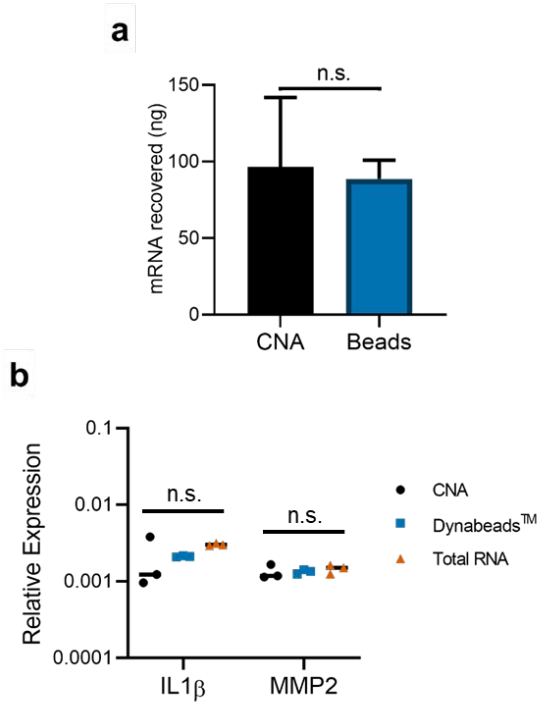


Figure 4 –(a) Using optimized buffer conditions, the performance of oligo(T) CNA compared to Dynabeads showed no statistical difference in mass yield. (b) There was also no statistical difference in relative expression levels measured using mRNA input from different isolation procedures. Data is represented as the mean of at least 3 replicates and error bars represent standard deviations.

from oligo(T) CNA, non-complementary oligo(A) CNA, or no CNA. After washing and release, the concentration of functional EGFP mRNA, or mRNA that could be translated into fluorescent protein, was measured via IVT, where higher EGFP fluorescence indicated a higher concentration of functional mRNA (**Figure S5**). The ratio of the fluorescence after pulldown and release to the fluorescence where no pulldown occurred ("relative EGFP fluorescence") was taken as a measure of the enrichment of specifically functional mRNA (see SI for more details). A relative EGFP fluorescence of 1 would indicate no pulldown or release. For mRNA mixed with complementary oligo(T) CNA and released at 75°C, relative EGFP fluorescence was 3.2 ± 0.4 (**Figure 3C**), while for both negative controls (i.e., oligo(A) CNA or no CNA), relative EGFP fluorescence was close to 1, (1.2 ± 0.1 and 1.0 ± 0.4, respectively). In the absence of heating, EGFP fluorescence for the oligo(T) CNA samples was negligible, comparable to levels observed for controls with no mRNA (**Figure 3d**). These results support the hypothesis that mRNA enriched through the oligo(T) CNA isolation method is readily translated into protein.

In prior studies, it was observed that while RNA pulldown was consistently above 90%, the release efficiency was more variable and sometimes quite poor (e.g. ~34% recovery, **Figure 2B**). After additional optimization of release conditions, the monovalent salt:mRNA concentration ratio in the release buffer was found to be the largest driver of variability due to salt-induced precipitation of mRNA (**Figure S6**). The washing and release procedure which resulted in the highest and most consistent recovery was one that gradually reduced the monovalent salt concentration (Buffers W1, W2, and R2, **Table S1**). This method was used for subsequent experiments.

For applications in molecular biology, it was useful to compare the CNA isolation method with commercially available protocols for enriching mRNA. First, oligo(T)'s ability to selectively precipitate mRNA was confirmed by comparing the pulldown efficiency to that with rRNA, which accounts for ~80% of total RNA (**Figure S7**). Next, using the optimized procedure detailed above, side-by-side isolations of mRNA from cell-isolated total RNA were conducted using oligo(T) CNA and the Dynabeads®, which are magnetic beads functionalized with oligo $(dT)_{25}$. After two washes and heat mediated release, the total mass of enriched mRNA was 100 ± 50 ng and 90 ± 20 ng for CNA and Dynabeads®, respectively (**Figure 4a**). These concentrations accounted for 3.1 ± 1.5% and 2.9 ± 0.8% of the input total RNA for the CNA and Dynabeads® methods, respectively, which fall within the expected range of mRNA abundance in total RNA (1-5%). Overall, the yield of mRNA after isolation with CNA is comparable to the yield using the commercially available Dynabeads®.

RT-qPCR was then performed on isolated mRNA to confirm that it would remain useful for downstream bioanalysis. Results were assessed for accuracy by comparing relative gene expressions to a sample of freshly isolated total RNA and mRNA isolated with Dynabeads®. Complementary DNA was synthesized using an oligo(dT_{12-18}) primer rather than a collection of random primers to ensure that only mRNA was reverse transcribed. Two genes of interest were chosen, *IL1B* (coding for interleukin-1 beta) and *MMP2* (coding for matrix metalloproteinase 2) (**Table S2**), due to their relevance to the

cell line used (MCF10a cells, non-tumorigenic mammary gland cells). *GAPDH* was used as the housekeeping gene. **Figure 4b** shows the relative expression (RE) levels of *IL1B* and *MMP2* compared to the housekeeping gene. For both genes examined, RE levels of mRNA isolated with CNA were not statistically different than either mRNA isolated with Dynabeads® or total RNA samples. These results support the hypothesis that mRNA isolated using CNA maintains its integrity. Further, the isolation process does not bias isolation of specific mRNAs by length, at least for the two genes investigated herein. If the CNA-mediated isolation method did experience length dependent isolation efficiencies, significant decreases in RE would be expected for larger mRNAs. In addition to the fact that this was not observed in this experiment, it is worth noting that the mRNA for *MMP2* is nearly twice the size of the mRNA for *IL1B*, depending on the transcript.

This study reports an alternative strategy of mRNA isolation using hydrophobic oligo(T) CNAs, which have the ability to pull mRNA out of solution via binding to the poly(A) tail. Through the optimization of buffer conditions and material concentrations, mRNA yields as high as 94% were observed. Given the relative cost and scalability of CNA compared to traditional mRNA enrichment methods, CNAs represent a favorable method for situations when a large amount of reagent is necessary. In comparison to the Dynabeads® mRNA DIRECT™ kit, the CNA-mediated isolation method described herein resulted in similar overall yield and quality of mRNA as determined by *in vitro* translation and RT-qPCR analysis. These results indicate that mRNA isolation using hydrophobic CNA offers a competitive alternative to traditional, DNA dependent strategies.

Conflicts of interest

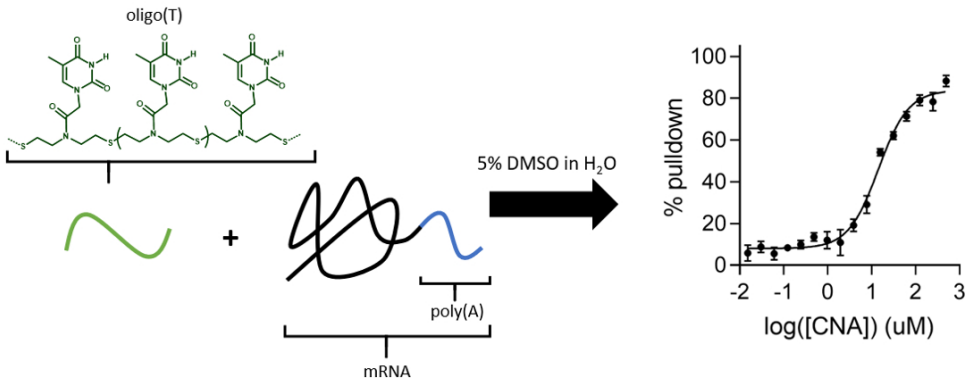
The authors declare no conflicts of interest.

Acknowledgements

This work was completed with support from an NSF MRSEC grant (DMR 1420736), the Colorado Office of Economic Development and International Trade (CTGG12016-2273), and from a US Department of Education GAANN Fellowship to Alex Anderson. MST experiments were performed in the Shared Instruments Pool of the Department of Biochemistry at the University of Colorado Boulder. The MST instrument was funded by NIH Shared Instrumentation Grant S10OD21603.

Notes and references

1 R. Stark, M. Grzelak and J. Hadfield, *Nature Reviews Genetics*, 2019, **20**, 631–656.
2 I. Chepelev, G. Wei, Q. Tang and K. Zhao, *Nucleic Acids Res.*, 2009, **37**, e106.
3 E. T. Cirulli, A. Singh, K. V. Shianna, D. Ge, J. P. Smith, J. M. Maia, E. L. Heinzen, J. J. Goedert, D. B. Goldstein and Center for HIV/AIDS Vaccine Immunology (CHAVI), *Genome Biol.*, 2010, **11**, R57.
4 E. V. Todd, M. A. Black and N. J. Gemmell, *Mol Ecol*, 2016, **25**, 1224–1241.
5 S. A. Bustin, *Journal of Molecular Endocrinology*, 2000, **25**, 169–193.
6 A. S. Devonshire, R. Sanders, T. M. Wilkes, M. S. Taylor, C. A. Foy and J. F. Huggett, *Methods*, 2013, **59**, 89–100.
7 M. L. Wong and J. F. Medrano, *BioTechniques*, 2005, **39**, 75–85.
8 F. Tang, C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Lao and M. A. Surani, *Nature Methods*, 2009, **6**, 377–382.
9 N. Cloonan, A. R. R. Forrest, G. Kolle, B. B. A. Gardiner, G. J. Faulkner, M. K. Brown, D. F. Taylor, A. L. Steptoe, S. Wani, G. Bethel, A. J. Robertson, A. C. Perkins, S. J. Bruce, C. C. Lee, S. S. Ranade, H. E. Peckham, J. M. Manning, K. J. McKernan and S. M. Grimmond, *Nature Methods*, 2008, **5**, 613–619.
10 A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer and B. Wold, *Nature Methods*, 2008, **5**, 621–628.
11 Z. Wang, M. Cai, Y. Weng, F. Zhang, D. Meng, J. Song, H. Zhou and Z. Xie, *J Cancer Res Clin Oncol*, 2015, **141**, 1353–1361.
12 N. C. Wu, W. Wong, K. E. Ho, V. C. Chu, A. Rizo, S. Davenport, D. Kelly, R. Makar, J. Jassem, R. Duchnowska, W. Biernat, B. Radecka, T. Fujita, J. L. Klein, M. Stonecypher, S. Ohta, H. Juhl, J. M. Weidler, M. Bates and M. F. Press, *Breast Cancer Res Treat*, 2018, **172**, 327–338.
13 M. F. de Souza, H. Kuasne, M. de C. Barros-Filho, H. L. Cilião, F. A. Marchi, P. E. Fuganti, A. R. Paschoal, S. R. Rogatto and I. M. de S. Cólus, *PLOS ONE*, 2017, **12**, e0184094.
14 E. Hornes and L. Korsnes, *Gene Analysis Techniques*, 1990, **7**, 145–150.
15 C. Albretsen, K.-H. Kalland, B.-I. Haukanes, L.-S. Håvarstein and K. Kleppe, *Analytical Biochemistry*, 1990, **189**, 40–50.
16 R. E. Pemberton, P. Liberti and C. Baglioni, *Analytical Biochemistry*, 1975, **66**, 18–28.
17 C. Rosenow, R. M. Saxena, M. Durst and T. R. Gingeras, *Nucleic Acids Res*, 2001, **29**, e112.
18 J. Wu, J. Xiao, Z. Zhang, X. Wang, S. Hu and J. Yu, *Genomics, Proteomics & Bioinformatics*, 2014, **12**, 57–63.
19 X. Han, D. W. Domaille, B. D. Fairbanks, L. He, H. R. Culver, X. Zhang, J. N. Cha and C. N. Bowman, *Biomacromolecules*, 2018, **19**, 4139–4146.
20 H. R. Culver, J. Sinha, T. R. Prieto, C. J. Calo, B. D. Fairbanks and C. N. Bowman, *Biomacromolecules*, 2020, **21**, 4205–4211.
21 A. J. Anderson, H. R. Culver, S. J. Bryant and C. N. Bowman, *Polym. Chem.*, 2020, **11**, 2959–2968.
22 N. Jacobsen, P. S. Nielsen, D. C. Jeffares, J. Eriksen, H. Ohlsson, P. Arctander and S. Kauppinen, *Nucleic Acids Res*, 2004, **32**, e64–e64.
23 D. Phelan, K. Hondorp, M. Choob, V. Efimov and J. Fernandez, *Nucleosides, Nucleotides & Nucleic Acids*, 2001, **20**, 1107–1111.
24 J. Strien, J. Sanft and G. Mall, *Mol Biotechnol*, 2013, **54**, 1048–1054.
25 M. A. Jensen, M. Fukushima and R. W. Davis, *PLOS ONE*, 2010, **5**, e11024.
26 P. Gong and R. Levicky, *PNAS*, 2008, **105**, 5301–5306.
27 I. Y. Wong and N. A. Melosh, *Biophysical Journal*, 2010, **98**, 2954–2963.
28 D. Irving, P. Gong and R. Levicky, *J. Phys. Chem. B*, 2010, **114**, 7631–7640.
29 D. E. Draper, *RNA*, 2004, **10**, 335–343.
30 A. A. Zinchenko and K. Yoshikawa, *Biophysical Journal*, 2005, **88**, 4118–4123.


oligo(T)
5% DMSO in H_2O
poly(A)
mRNA
% pulldown
100
80
60
40
20
0
-2
-1
0
1
2
3
log([CNA]) (uM)