RSC Advances

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Synthetic Oligodeoxynucleotide Purification by Capping Failure Sequences with a Methacrylamide Phosphoramidite Followed by Polymerization†

Durga Pokharel,[a] Yinan Yuan,[b] Suntara Fueangfung[a] and Shiyue Fang*[a]

[a] Department of Chemistry, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan 49931, USA

[b] School of Forest Resources and Environmental Science, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan 49931, USA

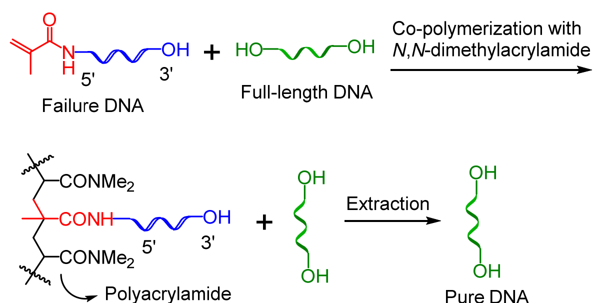
E-mail: shifang@mtu.edu

† Electronic supplementary information (ESI) available: General methods; effects of old capping agent, radical scavenger, harsher deprotection conditions, low concentration capping solution, and air on purification; ^{1}H, ^{13}C and ^{31}P NMR spectra; mass spectra; HPLC profiles; peak areas in HPLC profiles of nucleosides from digested ODN; recipes of polymerization solutions; and ODN synthesis cycles. See DOI: 10.1039/

RSC Advances Accepted Manuscript

Table of contents entry

Synthetic oligodeoxynucleotides are simply purified by capping failure sequences with a methacrylamide phosphoramidite, co-polymerization with *N*,*N*-dimethylacrylamide and extraction with water.



Abstract

Oligodeoxynucleotide (ODN) purification was achieved by capping failure sequences with a polymerizable methacrylamide phosphoramidite during automated synthesis, polymerizing the failure sequences into an acrylamide gel after cleavage and deprotection, and extraction of full-length sequences with water. The details about the technology including the capping efficiency of four polymerizable phosphoramidites, optimal capping time, diffusion speeds of ODN from gels with different cross-linking ratios to solution, and the efficiency of ODN extraction from gel were investigated. In addition, the technology was tested for purification of a long sequence and purification on larger scales. We also found that polymerization of failure sequences in a centrifuge tube in air did not affect purification results. Finally, we provided additional evidences that ODNs are stable under radical polymerization conditions by complete digestion of ODN followed by reversed-phase HPLC analysis of nucleosides.

RSC Advances Accepted Manuscript

Introduction

Applications such as total gene synthesis require the production of large numbers of short ODNs within a short time with high efficiency.[1] An ideal method to achieve this is through high throughput production. The high interest in using oligonucleotides (ONs) as therapeutic agents to cure human diseases generates a need for large quantities of ONs for preclinical research, clinical trial and patient use. This requires to develop technologies for large scale production of ON.[2] Currently, high throughput and large scale syntheses of ON are possible due to the commercialization of oligonucleotide microarray technologies,[3] high throughput automated solid phase synthesizer,[1c,4] and large scale synthesizer.[5] However, the products contain impurities, which must be removed before the applications. The impurities can be classified into three types. The first type is the truncated failure sequences generated in each synthetic cycle due to incomplete coupling of nucleoside phosphoramidite monomer. These failure sequences are capped usually with acetic anhydride to prevent them from growing in subsequent synthetic cycles and becoming the more difficult-to-remove deletion sequences. For a typical 20-mer ODN synthesis, they usually consist of 30-60% ODN content depending on the scale of synthesis, quality of reagents and solvents, and carefulness of the synthesizer operator to minimize exposure of reagents and solvents to air.[6] Because failure sequences have very similar physical properties as the desired full-length sequences, this class of impurities is challenging to remove. The second type of impurities is small organic molecules from protecting groups. These molecules are neutral, and therefore are easy to remove from the negatively charged full-length sequences. The third type impurities include deletion sequences, addition sequences, and sequences resulted from depurination as well as other modifications during synthesis, cleavage and deprotection. Deletion sequences can be resulted from incomplete capping, and incomplete

RSC Advances Accepted Manuscript

de-tritylation. Addition sequences can be resulted from premature de-tritylation. Deletion and addition sequences are most difficult to remove. Fortunately, they only exist in minute quantities in crude ODN. In ODN production, it is best to tune synthesis, cleavage and deprotection conditions to avoid the third type impurities. As a result, the major task of ODN purification usually is the removal of failure sequences from the full-length sequences.

Current methods for ODN purification include gel electrophoresis, ion exchange HPLC, trityl-on reversed-phase (RP) HPLC, RP cartridge chromatography, hydrophobic RP chromatography,[7] fluorous affinity purification,[6,8] biotin-avidin based affinity purification[9] and reaction-based solid phase extraction.[10] The limitations of these methods in large scale ODN purification have been discussed earlier.[11] For applications in high throughput purification, gel electrophoresis and all chromatographic methods are not practical or highly expensive, affinity purification methods may be expensive, and to our best knowledge, reaction-based solid phase extraction methods have not been well developed.

Recently, we communicated our preliminary results on the development of a new ODN purification technology based on catching failure sequences by polymerization.[11b] Using this method, the failure sequences in crude ODN could be simply removed by co-polymerizing them into an insoluble polymer, and pure ODN could be obtained by simple extraction followed by *n*-BuOH precipitation.[12] The method does not require any chromatography, and purification is achieved with simple manipulations such as mixing, shaking and extraction. As a result, it is potentially useful for large scale and high throughput ODN purification. In this article, we further report our detailed studies of the technology.

RSC Advances Accepted Manuscript

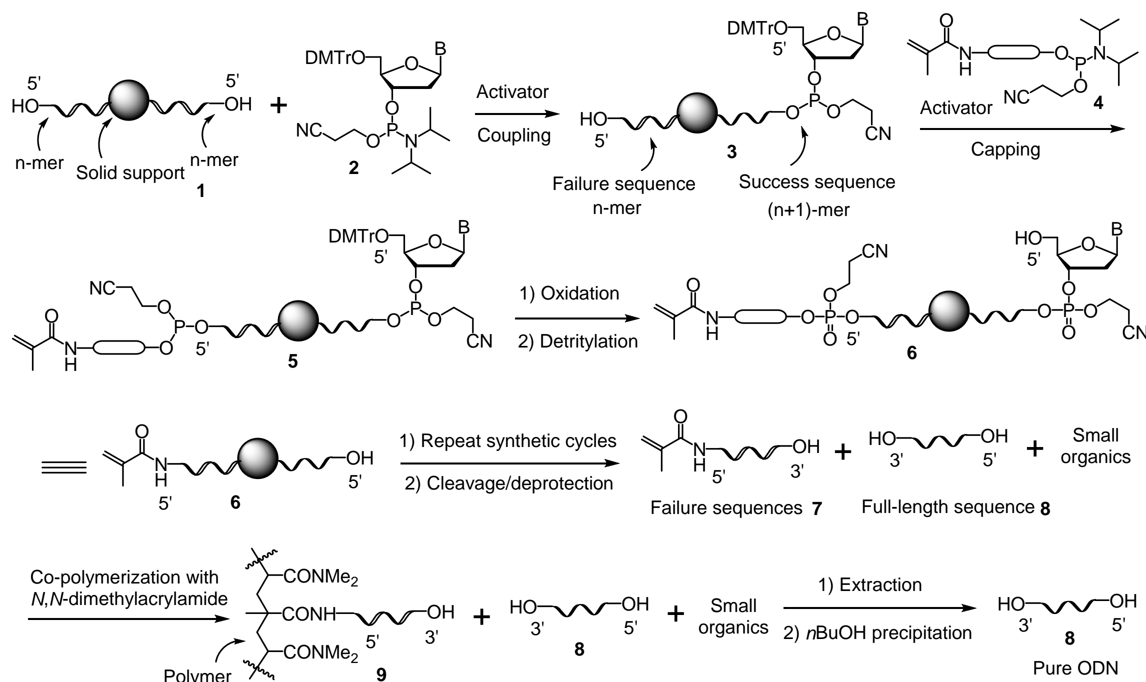
Results and discussion

The working principle of the catching failure sequences by polymerization ODN purification method is detailed in Scheme 1. The entity **1** represents two of many n-mer ODNs on solid support (usually controlled pore glass, CPG) at the end of nth synthetic cycle, at which time the 5′-end DMTr (4,4′-dimethoxytrityl) group is removed (each cycle contains four steps: coupling, capping, oxidation and detritylation). Coupling **1** with excess phosphoramidite monomer **2** in the presence of an activator to generate (n+1)-mer is usually highly efficient, but it is inevitable for small amount of n-mer to fail the coupling and therefore generating failure sequences. The entity **3** represents one failure and one success sequences on solid support. Repeated coupling of **3** with **2** to complete the reaction is not a good option due to the expense of **2** and potential side reactions such as premature detritylation. However, if the failure sequences were carried on to the next synthetic cycles, they would become deletion sequences, which are highly challenging to remove during purification. As a result, they are usually capped with acetic anhydride. For using our catching failure sequences by polymerization ODN purification technology, however, the capping agent is replaced with the polymerizable methacrylamide phosphoramidite **4**, and the resulting material is depicted with **5**. After oxidation and detritylation steps, **5** becomes **6**. The synthesis is continued and the capping failure sequences with **4** is performed in each cycle.

After cleavage and deprotection, there are mainly three entities in the crude product, the failure sequences that contain a methacrylamide group (**7**), the full-length sequences that do not contain a methacrylamide group (**8**), and small organics from protecting groups. The crude product is subjected into radical polymerization conditions in the presence of additional polymerization monomer (e.g. *N*,*N*-dimethylacrylamide) and a cross-linker (e.g. *N*,*N*′-methylenebis(acrylamide)). The failure sequences are incorporated into an insoluble polymer (**9**).

RSC Advances Accepted Manuscript

The full-length sequences (**8**) and small organics remain in solution and are extracted with water. The small organics are neutral and soluble in *n*BuOH, while the full-length sequences are anionic and insoluble in *n*BuOH. The two are separated by precipitation of ODN **8** with *n*BuOH from a solution in concentrated NH_4OH (Scheme 1).



Scheme 1. The catching failure sequences by polymerization ODN purification method

Capping efficiency of four methacrylamide phosphoramidites and optimal capping time

Previously, we used the methacrylamide phosphoramidite **10** as the reagent for capping failure sequences.[11b] In the present study, we synthesized three more (**11-13**, Fig. 1). The capping efficiency of them and **10** were evaluated. Compared with **10**, compound **11** has an ether oxygen atom in the linker, which may increase its solubility in polar organic solvents such as acetonitrile and have better reaction kinetics during capping. Compound **12** has two ether oxygen atoms in the linker and the linker is longer. These may offer benefits for reaction kinetics in both capping

RSC Advances Accepted Manuscript

and polymerization steps. Compound **13** has an even longer linker, and in the linker, there is an amide bond. It was expected that the compound exist as a solid at room temperature and therefore its purification and storage be more convenient. Compound **11** was simply synthesized from amino alcohol **14** in two steps in excellent yields (Scheme 2). The intermediate compound **15** is known,[13] but we used a different procedure for its preparation (see experimental section). Capping phosphoramidites **12-13** were prepared from **16**[14] and **18**,[15] respectively using the same procedures for preparing **11**. The intermediate compound **17** was known,[16] and **19** is new and was fully characterized.

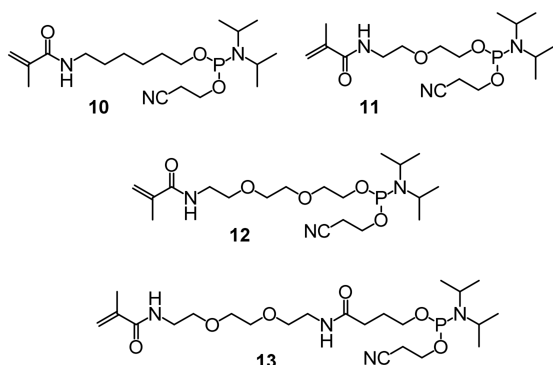
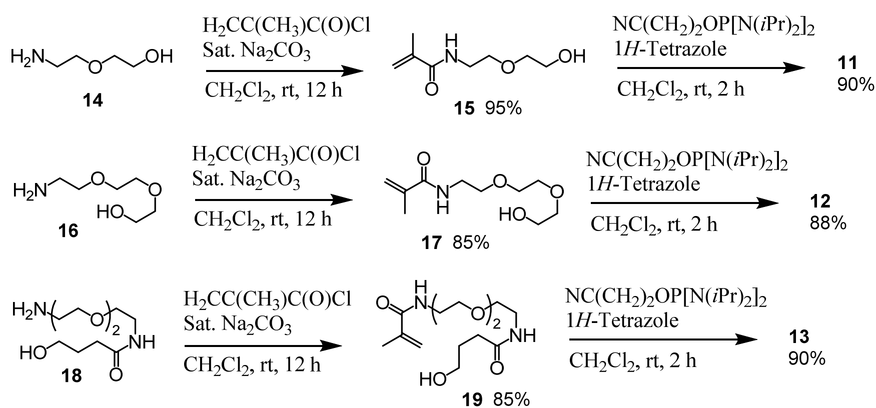


Fig. 1. Structures of polymerizable capping phosphoramidites



Scheme 2. Synthesis of polymerizable capping phosphoramidites

RSC Advances Accepted Manuscript

In our initial studies, during automated ODN synthesis, capping failure sequences was performed after the oxidation step that oxidizes the newly formed phosphite triester bonds in the success sequences.[11b] As a result, after capping, another oxidation step was needed to oxidize failure sequences. In the present study, we planned to simplify the process by omitting the first oxidation step and to oxidize the success and failure sequences in one single step. In addition, reduction of capping delivery time and the concentration of capping solutions were also planned. To test feasibility, the 20-mer ODN **20** was synthesized using a 0.15 M acetonitrile solution of **11** for capping. After the coupling steps, without oxidation, the capping solution was delivered to the synthesis column along with the activator 1*H*-tetrazole. The delivery time and manner were

20: 5' HO—TCA TTG CTG CTT AGA CCG CT—OH 3'
20(c**10**l): **20** with **10** for *c*apping using *l*onger capping time
20(c**11**l): **20** with **11** for *c*apping using *l*onger capping time
20(c**12**l): **20** with **12** for *c*apping using *l*onger capping time
20(c**13**l): **20** with **13** for *c*apping using *l*onger capping time
20(c**10**s): **20** with **10** for *c*apping using *s*horter capping time
20(c**11**s): **20** with **11** for *c*apping using *s*horter capping time
20(c**12**s): **20** with **12** for *c*apping using *s*horter capping time
20(c**13**s): **20** with **13** for *c*apping using *s*horter capping time
20(c**11**ls): **20** with **11** for *c*apping on a *l*arger *s*cale
20(c**11**dci): **20** with **11** for *c*apping and *dci* as activator
23: 5' HO—TAA AGC TAT AGG TAC AGT ATT AGT AGG ACC
3' HO—T GGC TGG ACC TGG TAA TAG AAC TGT CCA CAT
24: 5' HO—TCA TTG CT(8-oxo-dG) CTT AGA CCG CT—OH 3'
25: mixture of 20-mer ODNs with T at 3'-end and A, T, G, C on all other 19 positions, which therefore has 4^{19} theoretical sequences

Fig. 2 Labeling of ODNs

about the same as in the coupling step except that the reagents were delivered for two more times, and after each delivery a waiting step was added. More specifically, the reagents were delivered to the synthesis column for a total of four times (2.0 seconds, 1.5 seconds × 3). After each delivery the reagents remained in the column for 90 seconds (Cycle 1 in ESI). These capping steps consumed less capping phosphoramidite compared with those used in our initial

RSC Advances Accepted Manuscript

studies, which were 2.5 seconds, 1.5 seconds × 3 with 0.2 M solution.[11b] After the capping steps, the column was washed with acetonitrile, and the synthesis was allowed to proceed as normal. The final detritylation, cleavage and deprotection were also carried out under commonly used conditions. The crude product, which was labeled as **20**(c**11**l) (ODN labeling is summarized in Fig. 2), was analyzed with RP HPLC (trace a, Fig. 3).

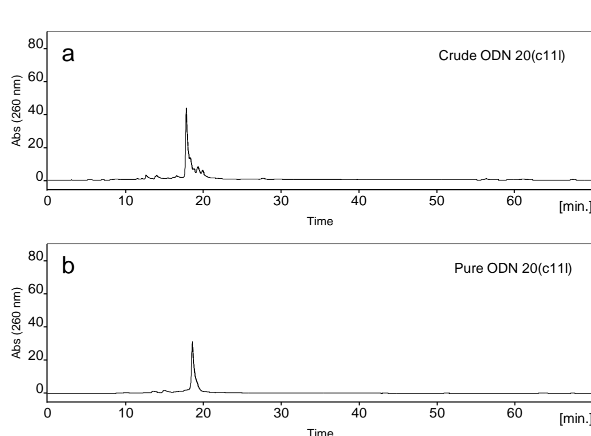
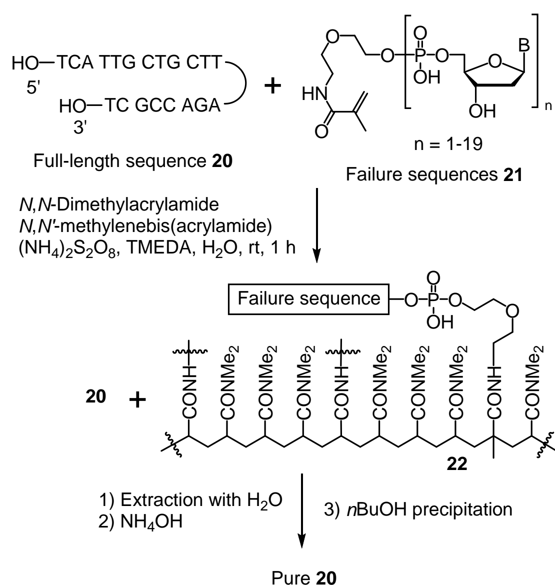


Fig. 3 RP HPLC profiles of ODN **20**(c**11**l). (a) Crude ODN. (b) ODN purified with catching by polymerization.

The remaining crude **20**(c**11**l) was subjected to radical acrylamide polymerization conditions widely used for the preparation of gels in electrophoresis experiments (Scheme 3). Briefly, the crude ODN, which mainly contained the full-length sequences **20**, the failure sequences **21** and small organic molecules from protecting groups, was mixed with a pre-prepared polymerization solution containing the polymerization monomer *N*,*N*-dimethylacrylamide and the cross-linker *N*,*N'*-methylenebis(acrylamide). Polymerization was then initiated under a nitrogen atmosphere with ammonium persulfate and *N*,*N*,*N'*,*N'*-tetramethylethylenediamine (TMEDA). The failure sequences were copolymerized into the acrylamide gel **22**. The full-length sequences (**20**) and small organics remained in solution, and were extracted with water. After evaporation of volatiles and heating with concentrated NH_4OH

RSC Advances Accepted Manuscript



Scheme 3. Purification of ODN **20** by polymerization of failure sequences

briefly, the ODN was precipitated with *n*-BuOH.[11b] The purified **20**(c**11**l) was analyzed with RP HPLC (trace b, Fig. 3) and the purity was found to be 93%. The recovery yield of the purification process was estimated to be 93% by dividing the area of the peak at 19 minutes in trace b by that in trace a in Fig. 3. The results indicate that the new agent **11** with the simplified capping procedure is capable of capping failure sequences completely. We also tested the capping agents **10**, **12** and **13** under the same new ODN synthesis and catching by polymerization purification conditions. All of them were found equally suitable for the application. The RP HPLC profiles for ODNs **20**(c**10**l), **20**(c**12**l) and **20**(c**13**l) are shown in ESI. Their recovery yields and purity are shown in Table 1.

In order to see if capping time could be further shortened and at the same time to distinguish the efficiency of the four capping reagents, the above ODN synthesis procedure was repeated with the capping conditions being changed from phosphoramidite and activator delivering time of 2.0 seconds, 1.5 seconds × 3, and waiting time of 90 seconds after each delivery to delivering

RSC Advances Accepted Manuscript

time of 2.0 seconds, 1.5 seconds × 2, and waiting time of 30 seconds (Cycle 2 in ESI). After cleavage and deprotection, the four crude samples **20**(c**11**s), **20**(c**10**s), **20**(c**12**s) and **20**(c**13**s) were purified using the same procedure as described above. The purified ODNs were analyzed with RP HPLC, and good purity and recovery yields were found (Table 1). Because the delivery time of capping reagents was almost the same as that of nucleoside phosphoramidite monomers in typical coupling steps (2.0 seconds, 1.5 seconds), further shortening of capping time was not pursued. All the four capping phosphoramidites (**10-13**) performed equally well, and therefore any of them could be used for the catching by polymerization technology. However, compounds **12-13** are relatively more expensive to make. Compound **13** did not exist as a solid or foam at room temperature as we hoped, and moreover, it was found more prone to oxidation during purification. As a result, compounds **10-11** are preferred capping reagents for the catching failure sequences by polymerization purification technology.

Table 1 ODN purity and recovery yield[a]

ODN	Purity	Recovery yield
20(c**10**l)	89%	92%
20(c**11**l)	93%	93%
20(c**12**l)	99%	72%
20(c**13**l)	100%	81%
20(c**10**s)	100%	82%
20(c**11**s)	100%	86%
20(c**12**s)	98%	95%
20(c**13**s)	94%	70%
23	92%	67%
20(c**11**ls)	94%	73%
20(c**11**dci)	100%	88%

[a] Purity was calculated from the areas of ODN and impurity peaks in HPLC profile. Yield was calculated by dividing the area of ODN peak in HPLC profile of purified sample by that in the profile of crude sample.

Diffusion speed of ODN from gel and extraction efficiency

The diffusion speed of ODN from the acrylamide gel to solution is an important factor to determine the efficiency of ODN extraction. The purification technology is most useful if the diffusion speed is high and ODN can be extracted with high efficiency and recovery yield. Our

RSC Advances Accepted Manuscript

common sense was that gels with higher cross-linking ratio [the mole ratio of *N*,*N′*-methylenebis(acrylamide) over *N*,*N*-dimethylacrylamide] would result in lower diffusion speed and extraction efficiency. Gels with lower cross-linking ratio were expected to give better results. However, if the cross-linking ratio is too low, the gel may be too soft and therefore difficult to handle. We were interested to find a compromised point where the most common 20-mer ODNs could be extracted with acceptable efficiency while the gel is hard enough to be easily handled. From experiments on ODN purification with polyacrylamide gel electrophoresis, it is known that gels with a cross-linking ratio of 1:25 are reasonably hard and 20-mer ODNs can be extracted from them with acceptable efficiency and yields. We therefore decided to start our experiments with this cross-linking ratio.

Table 2 Monitoring ODN diffusion from gels with different cross-linking ratios to supernatant[a]

Entry	Time (min)	ODN Con. (ng/μl)[b]	% of ODN[b,g]	ODN Con. (ng/μl)[c]	% of ODN[c,g]	ODN Con. (ng/μl)[d]	% of ODN[d,g]	ODN Con. (ng/μl)[e]	% of ODN[e,g]	ODN Con. (ng/μl)[f]	% of ODN[f,g]
1	10	47.7	66.0	29.9	47.5	32.7	45.0	27.1	39.9	25.8	39.3
2	20	53.3	73.7	37.8	60.1	40.9	56.3	36.0	53.0	34.7	52.9
3	30	54.7	75.7	43.4	69.0	49.5	68.2	42.0	61.9	37.7	57.5
4	40	58.6	81.1	48.6	77.3	53.4	73.6	46.9	69.1	42.8	65.2
5	50	64.6	89.3	50.9	80.9	56.7	78.1	50.5	74.4	47.0	71.7
6	60	65.1	90.0	54.6	86.8	59.2	81.5	53.3	78.5	50.0	76.2
7	70	66.4	91.8	56.0	89.0	64.2	88.4	56.1	82.6	53.3	81.3
8	80	67.6	93.5	58.0	92.2	64.6	89.0	59.7	87.9	56.0	85.4
9	90	68.6	94.9	59.6	94.8	67.2	92.6	59.9	88.2	58.6	89.3
10	100	70.9	98.1	61.2	97.3	68.0	93.7	64.0	94.3	59.8	91.2
11	110	71.7	99.2	62.0	98.6	70.6	97.3	66.2	97.5	65.4	99.7
12	120	72.3	100	62.9	100	72.6	100	67.9	100	65.6	100

[a] Crude ODN containing **20** and **21** was copolymerized into gel. Water was added to extract **20**. The concentration of **20** in the supernatant was monitored with UV over a two-hour period.
[b-f] In supernatants of gel with cross-linking ratio of 1:2, 1:7, 1:15, 1:25, 1:50, respectively.
[g] Calculated by dividing ODN concentration at a specific time by that after 120 min.

A portion of crude ODN **20**(c**11**s), which contained **21**, was subjected to polymerization using a polymerization solution that had a mole ratio of cross-linker over monomer of 1:25 (see details in Experimental section). Water was added to the gel, and small aliquots of supernatant were taken out every 10 min and the concentrations of ODN were determined with UV

RSC Advances Accepted Manuscript

measurement. As shown in Table 2, after 10 min, the concentration of ODN in the supernatant had reached about 40% (entry 1) of the value at 2 h, at which time the diffusion was close to equilibrium. For the concentration to reach 90% of equilibrium value, about 90 min were required (entry 9). To find out how many times were needed to extract all or most ODN from gel, after 2 h, the supernatant was removed completely (entry 12, Table 2 and entry 1, Table 3). To the gel, water was added again, and after 2 h, the supernatant was removed and its concentration was determined with UV measurement (entry 2, Table 3). The extraction and UV measurement were repeated for three additional times. As shown in Table 3, after two extractions, only about 3% extractable ODN remained in the gel (entry 3). After three extractions, the value lowered to 0.4% (entry 4). The results in Tables 2 and 3 indicate that ODN **20**(c**11**s) can be efficiently extracted from gel that has a cross-linking ratio of 1:25 with water for three times in a total of 6 h.

Table 3 Extraction efficiency of ODN from gel[a]

Entry	Time (h)	ODN Con. (ng/μl)[b]	% of ODN[b,g]	ODN Con. (ng/μl)[c]	% of ODN[c,g]	ODN Con. (ng/μl)[d]	% of ODN[d,g]	ODN Con. (ng/μl)[e]	% of ODN[e,g]	ODN Con. (ng/μl)[f]	% of ODN[f,g]
1	2	72.3	85.5	62.9	82.4	72.6	77.2	67.9	71.7	65.6	63.3
2	4	9.5	11.2	10.3	13.5	18.1	19.3	23.5	24.8	22.3	21.5
3	6	2.8	3.3	3.1	4.1	3.3	3.51	2.6	2.8	13.2	12.7
4	8	0.0	0.00	0.0	0.0	0.0	0.0	0.4	0.4	2.1	2.0
5	10	0.0	0.00	0.0	0.0	0.0	0.0	0.3	0.3	0.4	0.4

[a] The experiments in Table 1 were continued by removing the supernatant after 2 h. Water was added to the gels again, and supernatants were removed after 2 h. The extractions were repeated every 2 h for three additional times.
[b-f] In supernatants of gel with cross-linking ratio of 1:2, 1:7, 1:15, 1:25, 1:50, respectively.
[g] Calculated by dividing ODN concentration in each extract by that of the total amount of ODN extracted from the gel in the same volume of water.

To find out if a gel with lower cross-linking ratio could offer higher diffusion speed and better extraction efficiency, the above experiments were repeated with the exception of using a polymerization solution that had a cross-linking ratio of 1:50. As shown in Tables 2 and 3, the diffusion speed and extraction efficiency were not better, which was contrary to our prediction.

RSC Advances Accepted Manuscript

Instead, the data showed a slightly slower diffusion speed and lower extraction efficiency although the differences were small (e.g. 39.3% vs. 39.9% at 10 min entry 1, and 91.2% vs. 94.3% at 100 min entry 10, Table 2). Because gels with a higher cross-linking ratio gave slightly better results, we repeated the experiments with a cross-linking ratio of 1:15 (Tables 2 and 3). According to the trend, the diffusion speed and extraction efficiency should become better, which was indeed the case although the difference was small again. For example, within 10 min, the concentration of ODN in the supernatant had reached about 45% of equilibrium value (entry 1). For the concentration to reach 90% of equilibrium value, about 80 min were required (entry 8). For extraction efficiency, after three extractions, no more ODN was detectable in additional extracts (entry 4, Table 3). We then further increased the cross-linking ratio to 1:7. The trends of diffusion speed and extraction efficiency changes were the same. Within 10 min, the concentration of ODN in the supernatant had reached 48% of equilibrium value (entry 1). For the concentration to reach 90% of equilibrium value, about 70 min were needed (entry 7). When the cross-linking ratio was increased to 1:2, within 10 min, the concentration reached 66% (entry 1), and the time for reaching 90% was reduced to 60 min (entry 6). These results consistently show that the diffusion speed and ODN extraction efficiency increase with increased cross-linking ratio of gel. This is significant for the technology because gels with higher cross-linking ratios are harder and therefore easier to handle.

It was quite surprising to find that the ODN diffusion speed and extraction efficiency increase with cross-linking ratio. We had expected the contrary because gels with higher cross-linking ratio have smaller pores and ODN molecules would be more difficult to travel through them. Our explanation of the observed results is that the ODN molecules were not totally in the gel even though this seemed impossible because the gel with other contents including water and

RSC Advances Accepted Manuscript

ODN looked like homogenous. Instead, when the polymers grew, ODN molecules were expelled or partially expelled out from the polymer matrix. With lower cross-linking ratio when the pores in polymer were larger, a higher portion of ODN molecules was in the polymer matrix, or all ODN molecules were not in the polymer matrix, but they could diffuse into it with relatively high rates. These resulted in seemingly lower diffusion speeds of ODN from gels to solution and lower extraction efficiency. With higher cross-linking ratio when the pores in the polymer were smaller, a lower portion of ODN molecules was in the polymer matrix, or all ODN molecules were not in the polymer matrix, and they could only diffuse into it with low rates. These resulted in seemingly higher diffusion speeds of ODN from gels to solution and higher extraction efficiency.

Long sequence purification

For long sequence ODN purification, the catching failure sequences by polymerization method requires significant amount of polymerizable capping phosphoramidite, which is more expensive than acetic anhydride or phenoxyacetic anhydride capping agents used in catching full-length sequences by polymerization method.[11a] We therefore recommended the catching full-length sequences by polymerization method for long sequence purification.[11b,17] However, it is still interesting to test the suitability of the catching failure sequences method for ODN purification because it may be more convenient in certain applications such as high throughput purification. The 61-mer ODN **23**, which is a portion of the HIV protease gene, was selected as the target. The synthesis was carried out on a 0.2 µmol scale using the procedure with shortened capping time. Polymerizable phosphoramidite **11** was used as the capping agent. In the coupling step, after finishing phosphoramidite and activator delivery, an additional waiting of 30 seconds was added to increase yields (Cycle 2 in ESI with additional waiting after coupling). The crude

RSC Advances Accepted Manuscript

ODN was purified by catching failure sequences by polymerization as we did for shorter sequences. RP HPLC analysis showed that the ODN had good purity and recovery yield (Table 1 and ESI). The identity of purified ODN was confirmed with ESI-MS (ESI).

Purification at larger scales

After investigation of the effect of several additional factors such as capping agent storage time, radical scavenger, harsher deprotecting conditions, lower concentration of capping solution, polymerization in air (see ESI), we decided to test the technology for larger scale purification. For this goal, ODN **20** was synthesized on a 1 μmol scale. The synthesis conditions were only slightly different from those using shortened capping time described earlier, which was used for 0.2 μmol synthesis. The standard 1 μmol synthesis cycle provided by manufacturer of the synthesizer was modified. The steps for capping with acetic anhydride were replaced with those for capping with polymerizable phosphoramidite. The latter was achieved by delivering the solution of **11** in acetonitrile and activator to the synthesis column in the same manner as in the coupling steps except that an additional delivery was added and a waiting of 30 seconds was added after each delivery (Cycle 3 in ESI). At the end of synthesis, DMTr group was removed, and cleavage and deprotection were performed. The crude ODN was purified by catching failure sequences by polymerization using the same procedure described earlier. The only differences were that the 1 μmol ODN was purified in one batch, the volume of the polymerization solution was increased to 10 times of that used in the general 0.05 μmol purification procedure, and the volume of extraction water was slightly larger. The ODN, which is labeled as **20**(c**11**ls), was found to have good purity and recovery yield with RP HPLC (Table 1 and ESI). It is remarkable that even though the purification scale was increased from 0.05 μmol to 1 μmol, which was 20

RSC Advances Accepted Manuscript

times, increasing the volume of polymerization solution to 10 times was sufficient to catch all the failure sequences.

We further performed three 1 μmol syntheses simutaneously under conditions that are more close to actual large scale ODN production. Specifically, 1*H*-tetrazole is considered as an explosive and not recommended for large scale synthesis. Therefore, 4,5-dicyanoimidazole (DCI) was used as activator. To save phosphoramidite monomers, their concentration was lowered from 0.1 M to 0.05 M. A 0.1 M solution of **11** in acetonitrile was used for capping. With these modifications, ODN **20** was synthesized under otherwise the same conditions for the synthesis of **20**(c**11**ls). According to trityl assay and RP HPLC analysis of crude ODN, the synthesis yields were not significantly different from synthesis under conditions without the above modifications. The 3 μmol ODN was purified in one batch under the same conditions described for purification of **20**(c**11**ls) and is labeled as **20**(c**11**dci). RP HPLC analysis indicated that the ODN was highly pure and had excellent recovery yield (Table 1 and ESI).

ODN stability under radical polymerization conditions

A concern on the catching by polymerization purification technologies, which involves copolymerization of ODN into polymer under free radical conditions, is the potential damage of ODN under such conditions. Polymerization of ODN into polyacrylamide gels under radical conditions is known in the literature.[18] However, to our best knowledge, detailed study of ODN stability in the presence of free radicals has not been reported. Certainly, RP HPLC analysis of ODNs purified with the catching by polymerization technologies by comparing with authentic ODN strongly supports that ODNs can survive such conditions. Another evidence of stability comes from mass spectrometry, which consistently gave molecular mass consistent with calculated values.[11] Further, we also subjected the nucleosides guanosine, adenosine, thymidine

RSC Advances Accepted Manuscript

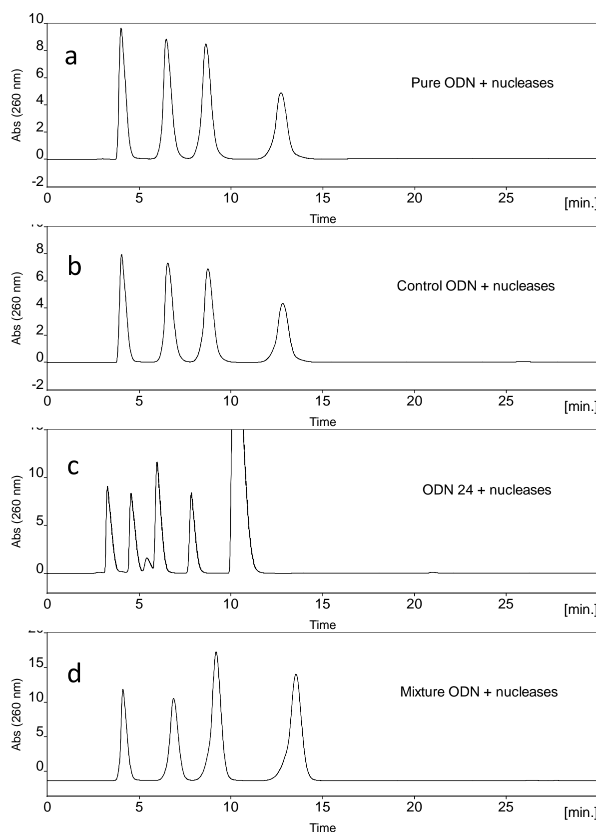


Fig. 4 RP HPLC profiles of nucleosides from ODN enzymatic digestion essay. (a) Nucleosides from ODN **20**(c**11**s), which was purified with the catching by polymerization technique. (b) Nucleosides from authentic ODN **20**. (c) Nucleosides from ODN **24**. (d) Nucleosides from the mixture ODN **25**. The peaks in a-b and d correspond to dC, dG, dT and dA, respectively. Those in c correspond to dC, dG, 8-oxo-dG, dT and dA, and 2-mercaptoethanol, respectively. The inconsistency of retention times in c with those in a-b and d is caused by the use of a different column.

and cytidine to free radical acrylamide polymerization conditions, extracted them from the polymer, and analyzed them with RP HPLC. No damaged nucleoside was detected. Among the four nucleosides, guanosine oxidation is more common, and the oxidation product is 8-oxo-guanosine. We compared the guanosine from the polymer with authentic 8-oxo-guanosine by co-injection into RP HPLC, the two were found to have different retention times, and no 8-oxo-guanosine was detected in the sample from polymer.[11a] However, concerns on ODN stability may still remain because this is so important an issue for purification, especially for ODN drug

RSC Advances Accepted Manuscript

purification. For RP HPLC analysis, if only one of the 20 bases of a 20-mer ODN is damaged, and the physical properties of the damaged nucleobase are not changed significantly, the analysis may not be able to differentiate damaged ODN from authentic ODN. For mass spectrometry analysis, if only a minute portion of ODN is damaged, the method may not be sensitive enough for detection. For the experiments to test stability of nucleosides under radical conditions, the reactivity of nucleosides in an ODN may not be the same as individual nucleosides. As a result, all the methods we used to support ODN stability under radical conditions are not conclusive.

To obtain additional evidences on the stability of ODN under radical acrylamide polymerization conditions, we used ion exchange HPLC to analyze one of our samples purified with the catching failure sequences by polymerization process. It was expected that ion exchange HPLC have different resolution on potentially damaged ODNs than RP HPLC. To our satisfaction, we still did not found any damaged ODN, and only one peak was observed (ESI). We also tried to find any impurity using RP HPLC by UV detecting at 210 nm instead of the usually used 260 nm; again, only one peak was observed (ESI). Further, we performed an ODN enzymatic digestion essay to detect any damaged nucleosides.[19] A portion of the ODN sample **20**(c**11**s) purified with the catching failure sequences by polymerization method was digested with Snake Venom Phosphodiesterase (SVP) into nucleotides, which were dephosphorylated to give nucleosides by Bacterial Alkaline Phosphatase (BAP). The neutral nucleosides were then separated from the highly polar species including inorganic salts and enzymes in the digestion mixture by precipitation of the latter materials by ethanol from a sodium acetate buffer. The nucleosides were soluble in ethanol and remained in the supernatant. The polar species were not soluble and were precipitated. The nucleosides were analyzed with RP HPLC. As shown in Fig. 4, only four peaks, which correspond to the four natural nucleosides, were observed (trace a). No

RSC Advances Accepted Manuscript

additional peaks that correspond to any damaged nucleosides were detected. An authentic ODN **20**, which was from a commercial source and was purified with preparative RP HPLC, was also subjected into the same digestion process. The resulting HPLC profile of nucleosides (trace b) was exactly the same as that from **20**(c**11**s). Another ODN (**24**), which is **20** with one dG being replaced with 8-oxo-dG (Fig. 2), was synthesized and purified by DMTr-on preparative RP HPLC (see ESI for crude and pure HPLC profiles and MALDI-TOF MS), and subjected to digestion.[20] The HPLC profile (trace c) showed five peaks besides a peak from 2-mercaptoethanol indicating that the essay could detect potentially damaged nucleosides. The relative retention times of the nucleosides including 8-oxo-dG were consistent with literature values.[19-20] To address the concern that ODN damage may be dependent on the secondary structures resulted from different sequences, a complex mixture of 20-mer ODN (**25**), which contained 4^{19} theoretical sequences, were synthesized using a mixture of the four nucleoside phosphoramidite monomers. The synthesis procedure with a shortened capping time was used (Cycle 2 in ESI), and the failure sequences were capped with methacrylamide phosphoramidite **11**. The crude ODN was subjected to the general catching failure sequences by polymerization process (ESI). The resulting mixture of full-length sequences was subjected to the same SVP digestion, BAP dephosphorylation, and RP HPLC analysis process (trace d, Fig. 4). Only four peaks were observed. To confirm that the peaks were from ODN and not from SVP and BAP, the essay process was repeated without using any ODN. No nucleoside peaks were observed (ESI). The four peaks in traces a-b and d correspond to dC, dG, dT and dA, respectively. The areas of them are consistent with calculated values (ESI).[19] These results provide strong evidence that ODNs with different sequences are stable under the radical acrylamide polymerization conditions. Further analysis to detect any minute quantities of potentially damaged ODNs under

RSC Advances Accepted Manuscript

the radical conditions is beyond the scope of this study. Future directions would be to use the purified ODNs in molecular biology experiments.

Conclusions

Several details of the catching failure sequences by polymerization ODN purification technology were investigated. We found that the four methacrylamide phosphoramidites **10-13** were similarly effective for capping failure sequences in the purification technique. However, because **10** and **11** are easier to prepare, they are the best choices. The studies on the diffusion speeds of ODN from polyacrylamide gel to solution indicate that extraction of ODN from gel is efficient. Normally, three extractions will recover more than 90% ODN. Contrary to our prediction, ODN extraction efficiency is higher in the case of gels with higher cross-linking. This is a favorable finding for the technology because gels with higher cross-linking are harder and easier to work with. The technology was successfully applied for the purification of a long sequence and purification at a larger scale. Capping failure sequences with older methacrylamide phosphoramidite and phosphoramidite solution containing a radical scavenger were also studied and were found that they have little adverse effects on the purification results. The more commonly used *exo*-amino nucleobases protecting groups that require harsher conditions to remove were tested in the ODN purification method. They were found equally compatible with the technology as the more labile phenoxyacetyl protecting groups. Polymerization in air is more convenient than under a nitrogen atmosphere. This is especially important for high throughput purification. We demonstrated that it does not affect purification results. Finally and most importantly, we provided further evidences to support that ODNs are stable under free radical

RSC Advances Accepted Manuscript

acrylamide polymerization conditions. This study is most important in the context of ODN drug purification.

Experimental

Methacrylamide 15

The compound is known,[13] but a different procedure was used for its preparation. To a round-bottomed flask under a nitrogen atmosphere was added **14** (2.0 g, 19.0 mmol, 1.0 equiv), freshly distilled CH_2Cl_2 (60 mL), and saturated Na_2CO_3 (60 mL). The mixture was cooled to 0 ^oC, and methacrolyl chloride (1.86 mL, 19.0 mmol, 1.0 equiv) was added via a syringe slowly with vigorous stirring. After addition, the reaction was allowed to proceed overnight while warming to rt gradually. The contents were then transferred into a separatory funnel. After removing the organic layer, the aqueous layer was extracted with CH_2Cl_2 (50 mL × 3). The combined organic phase was dried over anhydrous Na_2SO_4, and filtered. Volatiles were removed under reduced pressure. The crude product was purified with flash column chromatography (SiO_2, CH_2Cl_2/CH_3OH, 9:1) giving **15** as a colorless oil (3.12 g, 18.0 mmol, 95%): R_f = 0.7 (SiO_2, CH_2Cl_2/CH_3OH, 9:1).

Phosphoramidite 11

An oven-dried round-bottomed flask containing **15** (0.5 g, 2.89 mmol, 1.0 equiv) and a magnetic stirring bar was evacuated and then refilled with nitrogen. The evacuation and nitrogen-filling cycle was repeated for three times. Freshly distilled CH_2Cl_2 (5 mL) was added via a syringe forming a light yellow solution. 2-Cyanoethyl-*N*,*N*,*N’*,*N’*-tetraisopropylphosphoramidite (1.01 mL, 3.18 mmol, 1.1 equiv) and 1*H*-tetrazole (0.45 M solution in CH_3CN, 7.07 mL, 3.18 mmol, 1.1 equiv) were added slowly via syringes

RSC Advances Accepted Manuscript

sequentially. After stirring at rt for 2 h, the reaction mixture was concentrated to dryness using a flow of nitrogen. The residue was purified with flash column chromatography (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2) to give **11** as a colorless oil (0.98 g, 2.62 mmol, 90%): R_f = 0.7 (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.29 (br s, 1H), 5.61 (s, 1H), 5.24 (s, 1H), 3.81-3.60 (m, 4H), 3.57-3.50 (m, 6H), 3.45-3.40 (m, 2H), 2.56 (t, J = 6.4 Hz, 2H), 1.88 (s, 3H), 1.18-1.04 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 168.6, 140.2, 119.6, 117.8, 71.2 (d, J = 7.6 Hz), 69.8, 62.9 (d, J = 16.7 Hz), 58.6 (d, J = 19.7 Hz), 43.2 (d, J = 12.1 Hz), 39.6, 24.8, 24.7 × 2, 24.66, 20.55, 20.47, 18.8; ^{31}P NMR (162 MHz, $CDCl_3$) δ 148.3.

Methacrylamide 17

The compound is known,[16] but a different procedure, which was the same as that for the synthesis of compound **15**, was used for its preparation. Reaction of compound **16**[14] (1.3 g, 8.71 mmol, 1.0 equiv) with methacrolyl chloride (851 µL, 8.71 mmol, 1.0 equiv) in saturated Na_2CO_3 (40 mL) and CH_2Cl_2 (40 mL) gave **17** as a colorless oil after flash column chromatography purification (SiO_2, CH_2Cl_2/CH_3OH, 9:1; 1.6 g, 7.37 mmol, 85%): R_f = 0.6 (SiO_2, CH_2Cl_2/CH_3OH, 9:1).

Phosphoramidite 12

The same procedure for the synthesis of phosphoramidite **11** was followed. Thus compound **17** (500 mg, 2.30 mmol, 1.0 equiv), 2-cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphoramidite (0.80 mL, 2.53 mmol, 1.1 equiv) and 1*H*-tetrazole (0.45 M in CH_3CN, 5.60 mL, 2.53 mmol, 1.1 equiv) in CH_2Cl_2 (5 mL) gave compound **12** as a colorless oil after flash column chromatography purification (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2; 0.85 g, 2.04 mmol, 88%): R_f = 0.7 (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.33 (br s, 1H), 5.61 (s, 1H), 5.22 (s, 1H), 3.76-3.70 (m, 4H), 3.58-3.49 (m, 10H), 3.43-3.38 (m, 2H), 2.56 (t, J = 6.4 Hz, 2H),

RSC Advances Accepted Manuscript

1.86 (s, 3H), 1.10-1.07 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 168.5, 140.2, 119.6, 117.9, 71.4 (d, J = 7.6 Hz), 70.7, 70.4, 69.8, 62.8 (d, J = 18.2 Hz), 58.6 (d, J = 18.2 Hz), 43.2 (d, J = 12.1 Hz), 39.5, 24.8, 24.7 × 2, 24.66, 20.5, 20.4, 18.8; ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.5.

Methacrylamide 19

The same procedure for the synthesis of compound **15** was followed. Reaction of **18**[15] (2.0 g, 8.54 mmol, 1.0 equiv) with methacrolyl chloride (0.84 mL, 8.54 mmol, 1.0 equiv) in CH_2Cl_2 (80 mL) and saturated Na_2CO_3 (60 mL) gave **19** as a colorless oil (2.2 g, 7.28 mmol, 85%) after flash column chromatography purification (SiO_2, CH_2Cl_2/CH_3OH, 9:1): R_f = 0.4 (SiO_2, CH_2Cl_2/CH_3OH, 9:1); ^1H NMR (400 MHz, $CDCl_3$) δ 6.94 (br s, 1H), 6.80 (br s, 1H) 5.58 (s, 1H), 5.19 (s, 1H), 3.53-3.44 (m, 6H), 3.43-3.32 (m, 4H), 3.30-3.24 (m, 4H), 2.19 (t, J = 7.0 Hz, 2H), 1.81 (s, 3H), 1.74-1.66 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.1, 169.1, 139.9, 120.1, 70.3, 69.92, 69.86, 61.7, 50.3, 39.6, 39.4, 33.4, 28.5, 18.7. HRMS (ESI, $[M+H]^+$) m/z calcd for $C_{14}H_{27}N_2O_5$ 303.1914, found 303.1918.

Phosphoramidite 13

The same procedure for the synthesis of phosphoramidite **11** was followed. Thus compound **19** (0.5 g 1.65 mmol, 1.0 equiv), 2-cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphoramidite (0.58 mL, 1.82 mmol, 1.1 equiv) and 1*H*-tetrazole (0.45 M in CH_3CN, 4.05 mL, 1.82 mmol, 1.1 equiv) in CH_2Cl_2 (5 mL) gave compound **13** as a colorless oil after flash column chromatography purification (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2; 0.75 g, 1.49 mmol, 90%): R_f = 0.65 (SiO_2, hexanes/ethyl acetate/Et_3N, 1:3:0.2); ^1H NMR (400 MHz, $CDCl_3$) δ 6.40 (br s, 1H), 6.22 (br s, 1H), 5.64 (s, 1H), 5.27 (s, 1H), 3.84-3.69 (m, 2H), 3.65-3.41 (m, 14H), 3.40-3.32 (m, 2H), 2.58 (t, J = 6.0 Hz, 2H), 2.24 (t, J = 7.4 Hz, 2H), 1.95-1.87 (m, 5H), 1.16-1.08 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.8, 168.6, 140.2, 119.7, 118.0, 70.4, 69.9 (d, J = 18.3 Hz), 62.9

RSC Advances Accepted Manuscript

(d, J = 16.7 Hz), 58.4 (d, J = 19.8), 58.5, 45.5 (d, J = 6.0 Hz), 43.1 (d, J = 10.6), 39.5, 39.3, 33.1, 27.3, 24.8 × 2, 23.11, 23.04, 20.6, 20.3, 18.8; ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.5.

Synthesis of ODN 20 using polymerizable phosphoramidites 10-13 to cap failure sequences – The general ODN synthesis procedure

The synthesis using **11** as the capping agent is described as an example. The phosphoramidite chemistry was used. CPG (1000 Å pore size) with a LCAA-succinyl ester linkage was used as the solid support. The phosphoramidite monomers used were Pac-dA-CE, Ac-dC-CE, *i*-Pr-Pac-dG-CE and dT-CE. For capping failure sequences, a 0.15 M solution of **11** in dry acetonitrile (distilled under nitrogen or from a commercial source) was placed on the 5th phosphoramidite position. The two bottles normally used for storing Ac_2O capping reagents were left empty. A new synthetic cycle was created by copying the standard 0.2 µmol synthesis cycle into a new file. The steps for capping failure sequences with Ac_2O were deleted. After the coupling steps, new steps for delivering **11** from the 5th bottle position and 1*H*-tetrazole solution from the 9th position (i.e. the same activator for the coupling steps) were added. These steps were the same as those in the coupling steps except that the base was from the 5th bottle, the reagents were delivered for two additional times, and after each delivery a waiting of 90 seconds was added. More specifically, the solutions of **11** and 1*H*-tetrazole were delivered to column for 2.0 seconds, 1.5 seconds × 3, and after each delivery, the reagents remained in the column for 90 seconds. After capping, steps for washing the column with acetonitrile and oxidation were added (Cycle 1 in ESI). The synthesis was initiated using the above reagents and cycle on a 0.2 µmol scale. In the last synthetic cycle, the 5′-DMTr group was removed. Cleavage and deprotection were carried out on the synthesizer with concentrated NH_4OH (900 min × 4) at rt. The synthesized ODN is labeled as **20**(c**11**l). The same sequence was synthesized under the same

RSC Advances Accepted Manuscript

conditions using polymerizable phosphoramidites **10** and **12-13** as capping reagent giving **20**(c**10**l), **20**(c**12**l) and **20**(c**13**l), respectively.

Purification of ODN 20 by polymerization of failure sequences – The general ODN purification procedure

Purification of **20**(c**11**l) is described as an example. The ODN solution from the synthesizer was divided equally into four portions and dried in centrifuge tubes in a centrifugal vacuum concentrator. One portion was dissolved in 150 µL water, 20 µL was injected into RP HPLC to generate trace a (Fig. 3). The remaining 130 µL solution was transferred into a 25 mL 2-necked round-bottomed flask containing a magnetic stirring bar. The centrifuge tube was washed with water (40 µL × 3), and the washes were transferred to the same flask. A polymerization solution (250 µL, cross-linking ratio 1:25) was added via a pipette under positive nitrogen pressure. The solution was gently stirred under a nitrogen flow at rt. After 5 min, the nitrogen flow was stopped but stirring was continued under a nitrogen atmosphere. The polymerization reaction was then initiated with $(NH_4)_2S_2O_8$ (10%, 5 µL) and *N*,*N*,*N'*,*N'*-tetramethylethylenediamine (TMEDA, 5 µL), which were added sequentially via a pipette and a syringe, respectively. A clear gel was formed within 5 min and stirring stopped automatically. The gel was allowed to stand for 1 h to ensure complete polymerization.

The gel was broken into about 4 pieces with a clean spatula or preferably a knife with a sharp blade that does not grind the gel into fine pieces. To the flask, sufficient water that could cover all the gel (~3 mL) was added. The mixture was stirred or shaken gently at rt for 3 h. The supernatant was separated from the gel. The gel was further extracted with water (3 mL × 3) under the same conditions. After the volume was reduced in a centrifugal vacuum concentrator, the extracts were filtered into one 1.5 mL centrifuge tube that has a spin filtration function. The

RSC Advances Accepted Manuscript

extracts were evaporated into dryness. To the residue, concentrated NH_4OH (100 μL) was added. The tube was closed. After vortexing and centrifuging briefly, it was kept for 25 min on a heating block that was pre-heated to 80 ^{o}C. The solution was then cooled to rt, and *n*-BuOH (900 μL) was added. The tube was vortexed for 30 sec and centrifuged for 8 min at 14.1 relative centrifugal force. The supernatant was removed carefully with a pipette. The residue was dried in a centrifugal vacuum concentrator shortly, and then dissolved in 130 μL water. After vortexing and centrifuging shortly, 20 μL was injected into RP HPLC to generate trace b (Fig. 3). The recovery yield of the catching by polymerization process was estimated to be 93% by dividing the area of the peak at 19 min in trace b by that in trace a (Fig. 3). Crude ODNs **20**(c**10**l), **20**(c**12**l) and **20**(c**13**l) were purified using the same procedure, and their recovery yields and purity are in Table 1. RP HPLC profiles are in ESI.

Synthesis of ODN 20 using polymerizable phosphoramidites 10-13 to cap failure sequences with shorter capping time

ODN **20** was synthesized again following the general ODN synthesis procedure described earlier on a 0.2 μmol scale except that the capping time was decreased from 2.0 seconds, 1.5 seconds × 3 deliveries and 90 seconds waiting after each delivery to 2.0 seconds, 1.5 seconds × 2 deliveries and 30 seconds waiting after each delivery (Cycle 2 in ESI). The synthesized ODNs are labeled as **20**(c**10**s), **20**(c**11**s), **20**(c**12**s) and **20**(c**13**s).

Purification of ODN 20 synthesized using shorter capping time by polymerization of failure sequences

The general ODN purification procedure described earlier was used. The RP HPLC profiles are in ESI. The recovery yields and purity are in Table 1.

RSC Advances Accepted Manuscript

Diffusion speed of ODN from acrylamide gel and extraction efficiency

One fourth of the crude ODN **20**(c**11**s) was subjected into the radical acrylamide polymerization conditions in a 25 mL 2-necked round-bottomed flask as described in the general ODN purification procedure. The polymerization solution used in the procedure had a 1:25 cross-linking ratio between the cross-linker *N*,*N′*-methylenebis(acrylamide) and the polymerization monomer *N*,*N*-dimethylacrylamide. After complete polymerization, the gel was broken into several pieces with a clean spatula. Water (3 mL) was added to the gel, and the mixture was stirred gently with magnetic stirring bar. The aliquots of supernatant (5 μL each) were taken out via a pipette in every 10 min until 2 h. The ODN concentrations in the aliquots were determined using a UV-Vis spectrophotometer at 260 nm (Table 2). At 2 h, all the remaining supernatant were taken out and transferred into clean centrifuge tubes. The ODN concentration in the supernatant was determined with UV. To the gel was added water (3 mL) again and the mixture was gently stirred. After 2 h, the supernatant was removed and the ODN concentration was determined with UV. The extraction and ODN quantification process was performed for three more times (Table 3). The ODN diffusion and extraction studies were repeated four more times using polymerization solutions with a different cross-linking ratio. One with a ratio of 1:50 between *N*,*N′*-methylenebis(acrylamide) and *N*,*N*-dimethylacrylamide, and the others 1:15, 1:7 and 1:2. The data were recorded in Tables 2 and 3.

Synthesis of long sequence ODN 23

The 61-mer ODN **23** was synthesized on a 0.2 μmol scale using phosphoramidite **11** for capping failure sequences. The general synthesis procedure with shorter capping time (Cycle 2 in ESI) was used except that a 30 seconds additional waiting step was added after the coupling steps to increase coupling yields.

RSC Advances Accepted Manuscript

Purification of ODN 23 by polymerization of failure sequences

The general ODN purification procedure was used. The RP HPLC profiles from crude and purified ODN **23** were generated using a different buffer A that contained 10% urea under otherwise the same conditions described in the general methods section (ESI). Recovery yield and purity are in Table 1. The purified **23** was analyzed with negative ESI MS using Waters' Tofspec-2E, calcd for $C_{599}H_{749}N_{235}O_{361}P_{60}$ 18875.1, found 18875.1. Image of the MS is included in ESI.

Synthesis of ODN 20 on a larger scale

The general ODN synthesis procedure was followed except that a different synthesis cycle was used. To create the new cycle (Cycle 3 in ESI), the standard 1 μmol synthesis cycle provided by the manufacturer of the synthesizer was copied into a new file. The steps for capping failure sequences with Ac_2O were deleted. After the coupling steps, new steps for delivering polymerizable phosphoramidite **11** (0.15 M acetonitrile solution) from the 5th phosphoramidite bottle position and 1*H*-tetrazole solution from the 9th position (i.e. the same activator for the coupling steps) were added. These steps were the same as those in the coupling steps except that the base was from the 5th bottle, one additional delivery of reagents was added, and after each delivery a waiting of 30 seconds was added. More specifically, the solutions of **11** and 1*H*-tetrazole were delivered to column for 2.5 seconds × 3, and after each delivery, the reagents remained in the column for 30 seconds. Other modifications of the cycle were the same as detailed in the general synthesis procedure. The synthesis was performed using this new cycle at 1 μmol scale following the conditions described in the general synthesis procedure. The ODN is labeled as **20**(c**11**ls).

RSC Advances Accepted Manuscript

Purification of 20(c11ls) by polymerization of failure sequences

The general ODN purification procedure was followed. The 1 μmol ODN was purified in one batch. The volume of the polymerization solution was 10 times of that used in 0.05 μmol scale purification described in the general procedure. Extraction of full-length sequences from gel was achieved by water (4 mL × 4). HPLC profiles are in ESI. Recovery yield and purity are in Table 1.

Synthesis of ODN 20 with DCI as activator

ODN **20** was synthesized under the same conditions described for the synthesis of **20**(c**11**ls) except for the following modifications. DCI (0.25 M in acetonitrile) instead of 1*H*-tetrazole was used as the activator for coupling and capping. The concentration of phosphoramidite monomers was lowered from 1.0 M to 0.05 M. Capping was performed using a 0.1 M acetonitrile solution of **11**. The synthesis was performed in three 1 μmol columns simultaneously. The ODN is labeled as **20**(c**11**dci).

Purification of 20(c11dci) by polymerization of failure sequences

The 3 μmol ODN was purified in one batch following the procedure for the purification of **20**(c**11**ls). The volumes of the polymerization solution and extraction solutions were increased proportionally (i.e. 3 times of those for the 1 μmol **20**(c**11**ls) purification). The volumes of the polymerization initiators and the reagents for ODN precipitation from NH_4OH solution by *n*-BuOH were not changed. The RP HPLC profiles from crude and purified ODN **20**(c**11**dci) including one that was generated by UV detection at 210 nm from the purified ODN are included in ESI. Recovery yields and purity are in Table 1.

RSC Advances Accepted Manuscript

Synthesis of mixture ODN 25

The general ODN synthesis procedure with a shorter capping time was followed (Cycle 2 in ESI). The synthesis was carried out on a 0.2 μmol scale using polymerizable phosphoramidite **11** for capping failure sequences. To synthesize the mixture ODN, the solution of Pac-dA-CE, Ac-dC-CE, *i*-Pr-Pac-dG-CE and dT-CE in 1:1:1:1 mole ratio was used as the phosphoramidite monomers in the coupling steps in all synthesis cycles. The cycle was executed 20 times to give the 20-mer ODN mixture **25**. At the end of synthesis, the DMTr group was removed. The ODN were cleaved and deprotected with concentrated NH_4OH at rt on synthesizer.

Purification of ODN mixture 25 by polymerization of failure sequences

The general ODN purification procedure was used to remove failure sequences. The RP HPLC profiles from crude and purified ODN **25** (only failure sequences were removed; theoretically it still contained 4^{19} sequences) are included in ESI. The recovery yield of the purification process was estimated to be 78%.

ODN stability studies through nuclease digestion followed by RP HPLC analysis

To a 1.5 mL centrifuge tube, which contained 18 μg ODN **20** purified by catching failure sequences by polymerization, was added 29 μL master mixture solution containing 16.45 unit SVP (Worthington Biochemical Corporation), 0.4 unit BAP (TaKaRa Bio Inc.), and 1× of BAP buffer. The tube was vortexed and centrifuged briefly to mix thoroughly and to bring the master mixture to the bottom of the tube. The mixture was then incubated at 37 ºC for 24 h. The digested ODN was then prepared for RP HPLC analysis following reported procedure.[19] Briefly, sodium acetate buffer (3 M, 4 μL) and ethanol (100 μL) were added into the tube. The tube was then vortexed shortly and chilled on dry ice for 10 min. After centrifuging for 5 min at 14.1 relative centrifugal force, the supernatant was carefully transferred into another 1.5 mL

RSC Advances Accepted Manuscript

centrifuge tube. To the tube was added ethanol (300 μL), and the contents were vortexed and chilled on dry ice again for 10 min. The tube was centrifuged for 5 min, and the supernatant was transferred into a new 1.5 mL centrifuge tube. After drying in a centrifugal vacuum concentrator, the nucleosides were dissolved in water (30 μL). The sample (20 μL) was analyzed with RP HPLC under the conditions described in the General section except that the gradient system was changed to: solvent A (100%) for 5 min, solvent B (0%-10%) in solvent A over 30 min, followed by solvent B (10%-100%) in solvent A over 25 min. The resulting profile is shown in Fig. 4 (trace a). The same ODN digestion and RP HPLC analysis procedure was applied to a control ODN (18 μg) from a commercial source purified with preparative RP HPLC (trace b), the mixture ODN **25** (33 μg, trace d), and a blank control (ESI). The control ODN **24** (9 μg) was synthesized and digested following known procedure (trace c).[20]

Acknowledgements

Financial support from US NSF (CHE-0647129 and CHE-1111192), Michigan Universities Commercialization Initiative, MTU Research Excellence Fund (REF-TC), MTU Chemistry Department, MTU Biotech Research Center, and The Royal Thai Government Scholarship (S. F.); the assistance from Mr. Dean W. Seppala (electronics) and Mr. Jerry L. Lutz (NMR); and NSF equipment grants (CHE-9512445 and CH-1048655) are all gratefully acknowledged.

Notes and references

1 (a) T. Yu, X. Bao, W. Piao, J. Peng, W. Li, C. Yang, M. Xing, Y. Zhang, J. Qi, L. Xu, L. Xu and Q. Liu, *Recent Pat. DNA Gene Seq.,* 2012, **6**, 10; (b) S. Y. Ma, N. Tang and J. D. Tian, *Curr. Opin. Chem. Biol.,* 2012, **16**, 260; (c) R. A. Hughes, A. E. Miklos and A. D. Ellington in *Gene*

RSC Advances Accepted Manuscript

Synthesis: Methods and Applications, *Vol. 498* (Ed. C. Voigt), Elsevier Academic Press Inc, San Diego, **2011**, pp. 277.

2 M. Schulte, N. Luhring, A. Keil and Y. S. Sanghvi, *Org. Process. Res. Dev.,* 2005, **9**, 212.

3 C. Agbavwe, C. Kim, D. Hong, K. Heinrich, T. Wang and M. M. Somoza, *J. Nanobiotechnol.,* 2011, 1.

4 (a) J. D. Tian, K. S. Ma and I. Saaem, *Mol. Biosyst.,* 2009, **5**, 714; (b) J. Y. Cheng, H. H. Chen, Y. S. Kao, W. C. Kao and K. Peck, *Nucleic Acids Res.,* 2002, **30**, e93.

5 (a) S. Aitken and E. Anderson, *Nucleos. Nucleot. Nucl.,* 2007, **26**, 931; (b) Q. L. Song, Z. W. Wang and Y. S. Sanghvi, *Nucleos. Nucleot. Nucl.,* 2003, **22**, 629.

6 W. H. Pearson, D. A. Berry, P. Stoy, K. Y. Jung and A. D. Sercel, *J. Org. Chem.,* 2005, **70**, 7114.

7 B. S. Sproat, T. Rupp, N. Menhardt, D. Keane and B. Beijer, *Nucleic Acids Res.,* 1999, **27**, 1950.

8 C. Beller and W. Bannwarth, *Helv. Chim. Acta,* 2005, **88**, 171.

9 (a) S. Fang and D. E. Bergstrom, *Tetrahedron Lett.,* 2004, **45**, 7987; (b) S. Fang and D. E. Bergstrom, *Nucleic Acids Res.,* 2003, **31**, 708; (c) S. Fang and D. E. Bergstrom, *Bioconjugate Chem.,* 2003, **14**, 80.

10 W. Pieken, A. Wolter and M. Leuck, *U.S. Patent US20030195351*, **2003**.

11 (a) S. Fang and S. Fueangfung, *Org. Lett.,* 2010, **12**, 3720; (b) S. Fang, S. Fueangfung, X. Lin, X. Zhang, W. Mai, L. Bi and S. A. Green, *Chem. Commun.,* 2011, **47**, 1345; (c) Y. Yuan, S. Fueangfung, X. Lin, D. Pokharel and S. Fang, *RSC Adv.,* 2012, **2**, 2803.

12 M. Sawadogo and M. W. Vandyke, *Nucleic Acids Res.,* 1991, **19**, 674.

RSC Advances Accepted Manuscript

13 A. Gomez-Valdemoro, M. Trigo, S. Ibeas, F. C. Garcia, F. Serna and J. M. Garcia, *J. Polym. Sci. Pol. Chem.,* 2011, **49**, 3817.

14 L. H. Liu, H. Dietsch, P. Schurtenberger and M. D. Yan, *Bioconjugate Chem.,* 2009, **20**, 1349.

15 W. Tang and S. Fang, *Tetrahedron Lett.,* 2008, **49**, 6003.

16 P. I. Dalven, J. R. Hildebrandt, A. Shamir, A. J. Laccetti, L. T. Hodgins and H. P. Gregor, *J. Appl. Polym. Sci.,* 1985, **30**, 1113.

17 S. Fang, S. Fueangfung and Y. Yuan, *Purification of Synthetic Oligonucleotides via Catching by Polymerization*. In *Current Protocols in Nucleic Acid Chemistry*, Eds.: M. Egli, P. Herdewijn, A. Matusda and S. Y. Sanghvi, John Wiley and Sons, New York, **2012**, Unit 10.14.

18 (a) Y. Helwa, N. Dave, R. Froidevaux, A. Samadi and J. W. Liu, *ACS Appl. Mater. Interfaces,* 2012, **4**, 2228; (b) A. Baeissa, N. Dave, B. D. Smith and J. W. Liu, *ACS Appl. Mater. Interfaces,* 2010, **2**, 3594; (c) N. M. Toriello, C. N. Liu, R. G. Blazej, N. Thaitrong and R. A. Mathies, *Anal. Chem.,* 2007, **79**, 8549.

19 A. Andrus and R. G. Kuimelis, *Base Composition Analysis of Nucleosides Using HPLC*. In *Current Protocols in Nucleic Acid Chemistry*, Eds.: M. Egli, P. Herdewijn, A. Matusda and Y. S. Sanghvi, John Wiley and Sons, New York, **2001**, UNIT 10.6.

20 V. Bodepudi, S. Shibutani and F. Johnson, *Chem. Res. Toxicol.,* 1992, **5**, 608.

RSC Advances Accepted Manuscript