Green Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/greenchem

Polymerizable Phosphoramidites with Acid-Cleavable Linker for Eco-Friendly Synthetic Oligodeoxynucleotide Purification

Durga Pokharel, Shiyue Fang*

Department of Chemistry, Michigan Technological University, 1400 Townsend Drive, Houghton, MI 49931 USA

Email: shifang@mtu.edu

Abstract: Methacrylation phosphoramidites containing a linker cleavable with acetic acid were synthesized, and used for synthetic oligodeoxynucleotide (ODN) purification. During automated synthesis, the full-length ODN was tagged with the phosphoramidite. The failure sequences were not. In purification, the full-length ODN was co-polymerized into a polyacrylamide gel, and the failure sequences and other impurities were removed by washing. Pure ODN was cleaved from the gel with 80% acetic acid. Using the method, purification of sequences as long as 197-mer, which are from the phi29 DNA polymerase gene, and at scales as large as 50 μmol was demonstrated. The products have good purity and the recovery yields are high. The method does not use any type of chromatography and purification is achieved through simple manipulations such as shaking and filtration. Compared with gel electrophoresis and HPLC purification methods, the new technology is less labor-demanding and more amendable for automation, consumes less amount of environmentally harmful organic solvents and requires little energy for solvent evaporation. Therefore, it is ideal for high throughput purification and large scale ODN-based drug purification as well as small scale purification.

Introduction

Synthetic oligodeoxynucleotides (ODNs) have found wide applications.[1-5] Important recent examples include total gene synthesis for genome construction in synthetic biology projects,[6-11] and oligonucleotide-based drug development.[12-16] Gene synthesis requires hundreds of ODNs. In order for the synthetic genes to be affordable, the ODNs should be produced in a high throughput fashion. Due to automated ODN synthesis either using modern column-based synthesizers or microarray technologies, large numbers of ODNs can be synthesized simultaneously.[6-11] Before used in gene assembly, it is crucial that the ODNs are purified to remove any failure sequences and other impurities. Currently, ODN purification are mostly achieved using gel electrophoresis or HPLC. However, with these technologies purification in a highly parallel fashion is challenging, and high throughput ODN purification remains a bottle-neck for gene synthesis and synthetic genome construction. For oligonucleotide-based drug development, kilograms to metric tons of oligos are needed for clinical trials and patient use. This requires large scale oligo production, which includes synthesis and purification. Most of the challenges for large scale oligo synthesis have been overcome,[12-13] but large scale purification remains a bottle-neck.[13-16] Currently, the most widely used technology for large scale purification of oligonucleotide-based drugs is reversed-phase (RP) HPLC. The method is expensive to scale up, consumes large volumes

of environmentally harmful organic solvents, and takes significant energy for solvent evaporation. The waste to product mass ratio can be as high as 10^5:1.

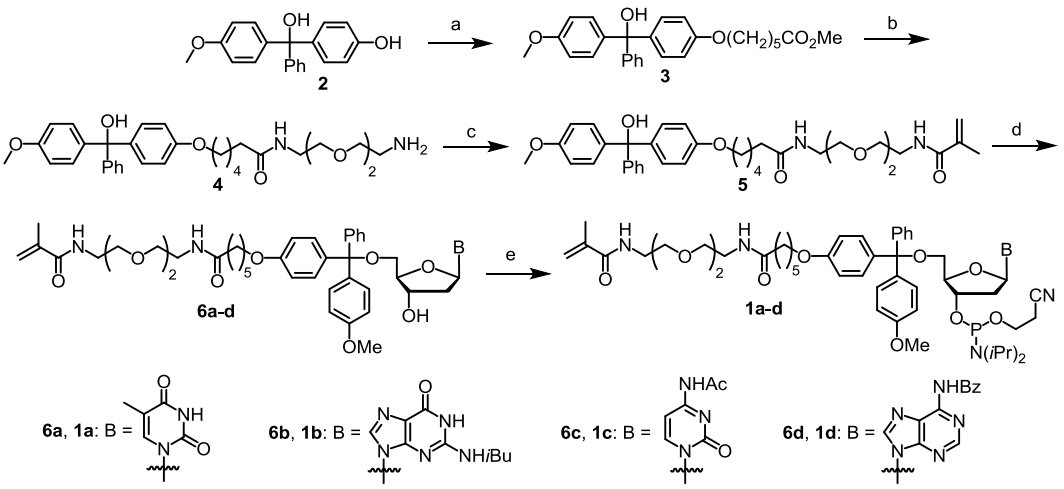
In order to overcome the challenges currently existed in high throughput and large scale ODN purifications, we recently reported a new technology using the concept of catching by polymerization.[17-21] In one instance, during automated synthesis, the failure sequences were capped with acetic anhydride in each synthetic cycle, and the full-length sequences were tagged with a polymerizable methacrylamide group using a methacrylation phosphoramidite that contained a cleavable linker. At purification, the crude product was simply subjected to acrylamide polymerization conditions. The full-length sequences were co-polymerized into the polyacrylamide gel, leaving the failure sequences and other impurities in solution, which were removed by washing. The full-length ODN was then cleaved from the polymer, and extracted with water. Highly pure ODN was obtained with excellent recovery yields.[18] The new purification technology does not involve any types of chromatography, and purification is achieved with simple manipulations such as shaking and filtration. The waste to product mass ratio can be as low as 1:1. As a result, the technology is highly suitable for high throughput and large scale purification. However, in our earlier studies, the linker we used in our methacrylation phosphoramidite for tagging full-length sequences was a diisopropylsilyl acetal function, and the toxic and corrosive hydrogen fluoride had to be used for cleaving the full-length sequences from the polyacrylamide gel. In addition, before cleavage the gel had to be dried under vacuum, which was energy-demanding and time-consuming. Furthermore, the cleaving reaction was carried out in DMF, which is expensive for large scale purification and difficult to evaporate due to its high boiling point.[18] In this Paper, we report the synthesis of four new methacrylation phosphoramidites, which contain a linker that can be readily cleaved with the more environmentally benign 80% acetic acid, and its application in ODN purification using the catching full-length sequences by polymerization approach. We demonstrate that even for purification of sequences as long as 197-mer and at scales as high as 50 μmol, the purification procedure requires little modification, and the technology is still highly efficient.

Results and Discussion

Synthesis of methacrylation phosphoramidites 1a-d

The phosphoramidites (**1a-d**), which contain a polymerizable methacrylamide group and an acid-labile trityl ether linker, was synthesized according to Scheme 1. The known compound **2**, which can be readily prepared from the commercially available 4-hydroxybenzophenone and 4-methoxyphenylmagnesium bromide,[22] was alkylated with methyl 6-bromohexanoate to give **3**. Compound **3** was conveniently converted to **4** by simply heating with 2,2′-(ethylenedioxy)bis(ethylamine) in water. Under these conditions, mono-acylation of the symmetric diamine was favored over di-acylation, which is a common side reaction observed in mono-acylation of diamines.[23-25] Compound **4** was acylated with methacrolyl chloride using DIEA as base to give **5**. Converting **5** to **6a-d** was achieved by first treating **5** with excess acetyl chloride. After removing volatiles, without purification, the intermediate trityl chloride derivative was used directly to react with dT, dG, dC and dA nucleosides, respectively in pyridine. Products **6a-d** were obtained in good isolated yields.[26] Phosphinylation of **6a-d** using 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropylphosphordiamidite with diisopropylammonium tetrazolide as activator gave the target polymerizable phosphoramidite tagging agents **1a-d** in good to excellent yields. All the

reactions are simple, and the reagents are inexpensive. We expect that the synthesis could be easily scaled up to kilograms.



Scheme 1 Synthesis of methacrylation phosphoramidite **1a-d**. Reagents and conditions: (a) $Br(CH_2)_5CO_2CH_3$ (1.0 equiv), K_2CO_3 (4.0 equiv), acetone, reflux, 12 h, 83%. (b) $[CH_2O(CH_2)_2NH_2]_2$ (3.0 equiv), H_2O, 85 °C, 12 h, 75%. (c) Methacrolyl chloride (1.0 equiv), DIEA (3.0 equiv), CH_2Cl_2, 0 °C to rt, 12 h, 87%. (d) Acetyl chloride, rt, 2 h; then nucleoside (1.0 equiv), pyridine, rt, 18 h, 77% (**6a**), 76% (**6b**), 76% (**6c**), 75% (**6d**); (e) $(iPr_2N)_2PO(CH_2)_2CN$ (1.1 equiv), diisopropylammonium tetrazolide (1.0 equiv), rt, 5 h, 87% (**1a**), 86% (**1b**), 81% (**1c**), 77% (**1d**).

Purification at small scales

The 20-mer ODN **7** (see Table 1 for ODN sequences) was used to demonstrate the convenience of the new technology for the purification of short ODNs at small scales. The ODN was synthesized on a 1 μmol scale under standard conditions using phosphoramidite chemistry. The synthesizer manufacturer suggested cycle was slightly modified so that the methacrylation

Table 1 ODN sequences

ODN	Length	Sequence
7	20-mer	5′-TCA TTG CTG CTT AGA CCG CT-3′
8	31-mer	5′-TCG GAT TAT GTC GAT ATT AGG AGA ATG GTA T-3′
9	37-mer	5′-TAT ACC ATT CTC CTA ATA TCG ACA TAA TCC GTC GAT C-3′
10	43-mer	5′-ACG GCT CAA CTC AAA CCT ATC AAA CTT GTA ACC CCT CGG CGC T-3′
11	61-mer	5'-TAA AGC TAT AGG TAC AGT ATT AGT AGG ACC TAC ACC TGT CAA GAT AAT GGT CCA GGT CGG T-3'
12	81-mer	5'-GTA CTC ATA ATA CTG TTT ACC GTC ATC ATC TTG AAG CAA CAT TGT CAC ATC GTA TGA GTC AAC AAA ATC ATT TTG CAC CAT-3′
13	151-mer	5'-TCA ACA AAA TCA TTT TGC ACC ATG TGG AGC ACC TCC AAA TAA CAC CTT TAT AAC CCA TGT GGC GTA ATC ATT GTT TTC CAT CCT AGA AAG CTC ATA CAA TGC GTT TTT CAT GAG TTT ATT TTC ATG CTC TAG TTT AGT CAT CTT CTT TTC T-3'
14	196-mer	5'-AAT ATG TTC TTC TAA CTC TGC ACG CTT CAT GAT ATA ACC CTC CTT ATC AGA TGT CAA ATA TAG TTT TCT CAC GGC TCA ACT CAA ACC TAT CAA ACT TGT AAC CCC TCG GCG CTT TAT ACT TAA TCA ACA CAC TAG CCG CAT CCG TTC TAC TAC GTG CTG TCG TAT CAT AGC GCC TAA CTA CCC CAT TGT CGT TAA T-3'
15	197-mer	5'-CCC CAA CAT ACA CAT GAC AAT GGA AGT ACC GTA CCA TTG ATA CTA CCA TTA TAG CAT ACT CTG TAA TAG TTG TCA AAC ATT AAT AGT TAA TTC TCC CAT TCT AAA ACG ATT TGA TCG TTT ATT TCT ACA ATT AGA TGG ACT CCA TTT GTA CCG AAT GGA TGG ACT TGT AAC TTT ATC GTA CCA TCT TTA AAC ATA TT-3'

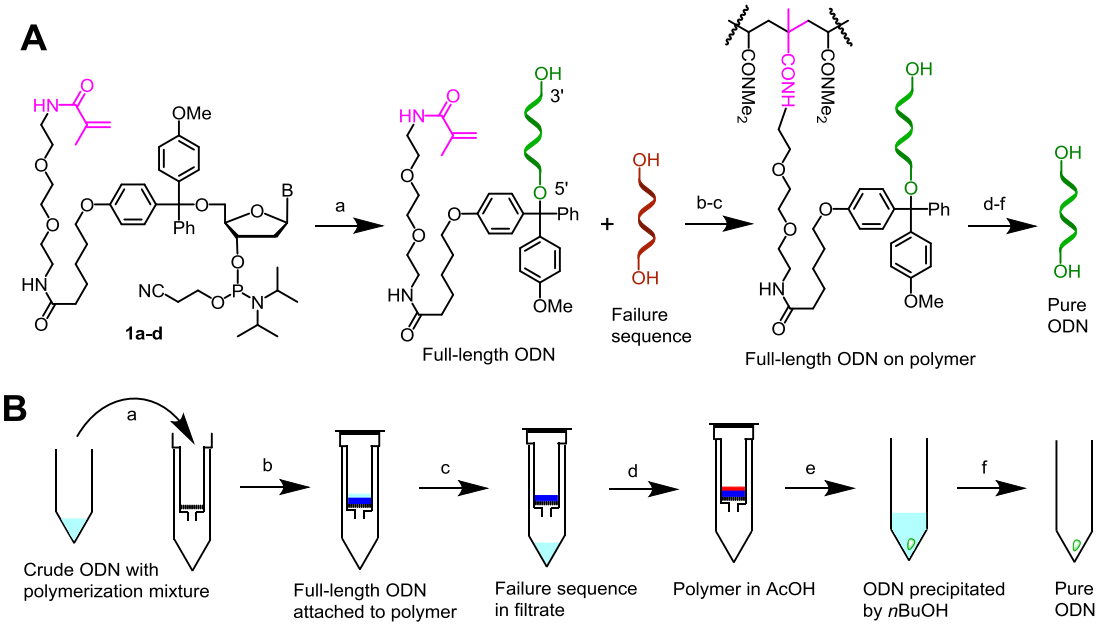


Fig. 1 The catching full-length sequences by polymerization ODN purification technology. (A) Principle: (a) Synthesize ODN using Ac_2O to cap failure sequences and **1a-d** to tag full-length sequences. (b) Polymerize full-length sequences. (c) Wash away failure sequences. (d) Cleave full-length sequences. (e) Extract. (f) Precipitate. (B) Workflow: (a) Add polymerization monomer and initiators to crude ODN, and transfer to a centrifugal filter unit. (b) Wait for polymerization. (c) Spin, add water, and spin to remove failure sequences. (d) Add 80% AcOH to cleave full-length ODN. (e) Spin, add water, and spin to extract full-length ODN; evaporate filtrate, and precipitate ODN with *n*BuOH. (f) Remove supernatant.

phosphoramidite **1a** could be coupled to the 5′-end of full-length sequences on the synthesizer automatically (see Figure 1). If one does not prefer to spend time on cycle modification, manual coupling can also be used. For automated coupling, the standard cycle was copied to a new file. After the coupling steps and before the capping steps, a 180 second waiting step was inserted, and was set to be active only for base 5. The solution of **1a** in acetonitrile at the same concentration (0.1 M) as other standard phosphoramidite monomers was attached to this bottle position. In the sequence, the nucleotide at the 5′-end was edited to be 5, which in the example here incorporated the nucleotide dT while tagging the ODN with a methacrylamide group. Except for these modifications, all other synthesis, deprotection and cleavage conditions were standard. Briefly, CPG with a succinyl ester linker was used as solid support. The phosphoramidite monomers were 5′-DMTr 2-cyanoethyl dT, isobutyryl-dG, acetyl-dC, and benzoyl-dA. Their 0.1 M solutions in acetonitrile were used for coupling with DCI as the activator. Failure sequences were capped with acetic anhydride in each synthetic cycle. At the end of synthesis, detritylation was not carried out as this would remove the methacrylamide tag. Cleavage was carried out on the synthesizer with concentrated ammonium hydroxide at room temperature. Deprotection was achieved by heating the ammonium hydroxide solution to 55 °C for 15 hours. To minimize the possibility of losing the acid sensitive tag, we added ~200 µl DIEA, which has a higher boiling point than ammonia, to the solution. The solution was then divided into five equal portions, and evaporated to dryness under vacuum. One portion was dissolved in 300 µl water, and 20 µl was injected into RP HPLC to give trace a (Figure 2).

Once the full-length sequences were tagged with a methacrylamide group, purification was convenient (Figure 1). To the second portion of the crude ODN (~0.2 µmol), small amounts of water (50 µl), polymerization solution containing *N*,*N*-dimethylacrylamide and *N*,*N*′-

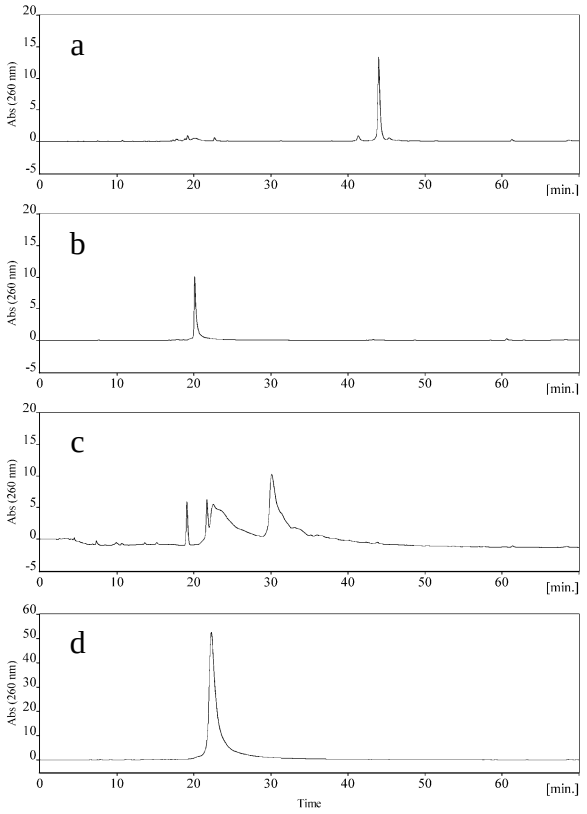


Fig. 2 RP HPLC profiles of ODNs: (a) crude 20-mer **7**; (b) pure **7**; (c) crude 197-mer **15**; (d) pure **15**.

methylenebis(acrylamide) (12 µl), 5% ammonium peroxide (5 µl) and 0.66 M TMEDA solutions (5 µl) were added sequentially (see experimental and protocol in ESI for details). For convenient filtration during downstream washing, cleavage and extraction in the process, before polymerization the solution was transferred into the top compartment of a centrifugal filtering unit using a pipette. The solution could stay above the filter disc even without stopping the bottom of the compartment. The polymerization occurred quickly within 15 minutes, but it was allowed to proceed for 1 hour to ensure completion (Figure 1). The full-length sequences were covalently linked to polymer, and failure sequences remained in solution. The polyacrylamide gel was loosened with a clean spatula, and water containing 5% Et_3N was added to wash away impurities including failure sequences and small molecules. The wash was very convenient because the supernatant could be removed by a simple spin in a centrifuge. So, it was carried out for six times, which was believed to be more than enough to remove all impurities. A final wash with pure water was performed to rinse away residue Et_3N. To cleave the full-length sequences from the polymer, minimum amount of 80% acetic acid that could cover the gel (120 µl) was added, and the mixture was incubated for 5 minutes. The supernatant was removed by spin, and the cleavage was repeated two times. To extract remaining ODN, minimum amount of water that could cover the gel (120 µl) was added, and the supernatant was collected by spin. This was repeated four times. The acetic acid and water extracts were combined and evaporated to dryness. The residue was dissolved in 100 µl concentrated ammonium hydroxide and precipitated with 900 µl *n*-BuOH. Pure ODN **7** was dissolved in 300 µl, and 20 µl was injected into RP HPLC to give trace b (Figure 2). According to the analysis, the ODN was 100% pure. Because the theoretical amount of ODN for generating traces a and b were equal, the recovery yield of the purification process was easily determined to be 68% by dividing the area of the peak in trace b at 19 minutes by that in trace a at 45 minutes. This method of recovery yield determination is applied to all other examples in this paper. Because the overall yield from trityl assay was 86%, and HPLC of the crude ODN indicated ignorable amount of 19-mer, the 68% recovery yield corresponds to a 59% synthesis and purification combined yield.

To evaluate the effectiveness of the cleavage procedure, we repeated the polymerization and washing procedures using the third portion of crude ODN **7**. We monitored the cleavage procedure with RP HPLC. Specifically, the gel was treated with 80% AcOH for 5 minutes for one time and extracted with water for two times. The acid and extracts were combined, and evaporated. The residue was precipitated with *n*-BuOH as described above. The ODN was analyzed with RP HPLC. The cleavage, extraction, precipitation and analysis procedure was repeated on the gel for

two more times. HPLC showed that ~95% ODN was obtained from the first cleavage. The second cleavage gave only ~5% ODN. We did not see any ODN in the third cleavage.

To demonstrate the simplicity and effectiveness of the purification procedure with more examples, the slightly longer ODNs **8** (31-mer), and **10** (43-mer) were synthesized under similar conditions for **7** using **1a** or **1d** as the tagging agent. For cleavage and deprotection, we tested different conditions. The mixture of concentrated ammonium hydroxide and methylamine solutions at 1:1 volume ratio was used in these cases. This reagent cleaved the ODN from CPG and removed all protecting groups at room temperature in three hours, and therefore the procedure was more convenient. The purification procedures were similar to that for purifying **7** with only slight modifications. Both ODNs were obtained in 100% purity according to RP HPLC analysis, and the recovery yields were high. RP HPLC profiles are in ESI. When 80% AcOH is used for cleaving ODN from polyacrylamide gel, there is a possibility that the acid concentration increases during evaporation, which may cause ODN damage. To address this concern, we also tested using the buffers 0.1 mM Et_3NHCl and 10 mM $Et_3NHO_2CCF_3$ at pH 3.0 for the cleavage using the purification of **10** as example. These buffers contain a more volatile acid component, which is expected to alleviate the problem. RP HPLC analyses of the purified samples (ESI) showed that these buffers could also cleave the ODN from gel efficiently and the ODN product was pure.

Purification at larger scales

The technology is expected to have significant advantages over known methods for large-scale ODN purification, which is required for ODN-based drug production. To demonstrate this, we decided to carry out a purification at a 50 μmol scale. The ODN **9** was synthesized in five 10 μmol columns with 1000 Å CPG. The same procedure and reagents for synthesizing **7** were used except that the synthesis cycle was modified from 10 μmol cycle. According to trityl essay before the cycle for attaching **1a**, the average yield of 36-mer for the five synthesis was 50%. Cleavage and deprotection were achieved using the mixture of concentrated ammonium hydroxide and methylamine. RP HPLC analysis of the crude indicated that only about 50% of the 36-mer intermediate was successfully methacrylated with **1a** (ESI). As a result, the synthesis yield of the 37-mer was about 25%. In the future, the synthesis should be improved by adjusting coupling and capping times, and by coupling the polymerizable tagging phosphoramidite (**1a-c** or **1d**) for one more time or allowing it to proceed longer. These modifications are expected to give higher synthesis yield. For purification, the crude ODN was dissolved in 1.25 ml water in a 25 ml round-bottomed flask, and polymerization was carried out using as little as 1.25 ml polymerization solution (Figure 3). The gel was transferred into a 15 ml Büchner vacuum funnel with a sintered glass disc and broken into several pieces with a spatula. Washing away failure sequences and other impurities were simply achieved using water containing Et_3N followed by pure water. The wash was convenient because the solutions could be easily removed by applying vacuum. Cleaving the full-length sequences from gel was achieved under the same conditions in small scale purifications with only a slightly larger volume of 80% AcOH (~4 ml). Extraction of full-length sequences and precipitation with *n*-BuOH were also very convenient (see experimental section). ODN **9** was obtained as nice white cotton-like fluffy fibers

Fig. 3 The 37-mer ODN **9** purified with the catching by polymerization technology: Left, the gel with ODN; middle, gel with ODN extracted; right, pure ODN from the gel.

(Figure 3). The weight was 98 mg, which corresponds to a 70% recovery yield of the purification process and an overall synthesis and purification yield of 18%. RP HPLC analysis showed that the ODN had good purity (ESI). In the future, the quality of the ODN may be further improved by adjusting conditions for washing away failure sequences. Removing 2-cyanoethyl groups before cleaving ODN from CPG may also help.[27]

Long sequence purification

Another important application of the technique is long ODN purification, which is especially important for total gene synthesis.[6-11] Because long ODNs can usually assume secondary structures and the hydrophobility of DMTr group becomes less distinctive as the length increases, they are challenging to purify with known methods. Because our catching by polymerization technology attaches the full-length sequences to a polymer gel covalently, in principle, there is no limit on the length of ODN that can be purified. To demonstrate the power of the technique, the 61-mer **11**, which is a portion of HIV protease gene, and the 81-mer **12**, 151-mer **13**, 196-mer **14** and 197-mer **15**, which are portions of phi29 DNA polymerase gene, were synthesized using **1a-c** or **1d** as polymerizable tagging agent. ODN **11** and **12** were synthesized on 1000 Å CPG, and **13**-**15** were synthesized on 2000 Å CPG. The synthesis conditions were similar as those for **7** except that the CPG was pre-capped with acetic anhydride manually before synthesis. The synthesis yields as indicated by trityl assay before the last cycle for tagging with **1a-d** were 83% (**11**), 48% (**12**), 22% (**13**), 35% (**14**) and 9% (**15**). Cleavage and deprotection were achieved using the mixture of concentrated ammonium hydroxide and methylamine. For polymerization, we added urea to disrupt any potential secondary structures for the long ODNs. We found that even 7.0 M urea did not have any adverse effects on the purification procedure. Except for adding urea, polymerization conditions including the amount of polymerization reagents were the same as described for purifying **7** even though the purification scales and sequence lengths were different. For washing failure sequences from gel, the same conditions for **7** were used for **11**. For the longer ODNs **13**-**15**, in addition to the same conditions, we also tested 3.0 M NaOAc, 10% piperidine in DMF, 10% piperidine in water and 1.0 M NaOH. The residue buffers were removed by washing with 5% Et_3N and water sequentially. We found that 3.0 M NaOAc gave better results in terms of recovery yield and appearance of HPLC peak. Cleavage of ODN from gel, extraction and *n*-BuOH precipitation were carried out under the same conditions described for **7**. Recovery yields were estimated to be 83% (**11**), 80% (**12**), 50% (**13**), 82% (**14**) and 78% (**15**) by comparing areas of pure and crude HPLC peaks. All ODNs were very pure. The crude and pure HPLC profiles for the 197-mer **15** are shown in Figure 2. Others are in ESI. ODN **12** was analyzed with ESI-MS, correct molecular mass was obtained (ESI).

Reagent mass considerations

A major advantage of the catching by polymerization purification technology compared to the most widely used RP HPLC is the low waste to product ratio. For RP HPLC, large volumes of buffers that contain harmful organic solvents are needed. For example, to purify 5 mg of ODN on a semi-preparative column, about 500 L buffer could be used. This corresponds to a waste to product ratio of roughly 10^5:1. In contrast, using our catching by polymerization technique, the waste to product ratio can be drastically reduced. In the examples in this paper, for small scale purifications including 0.2 μmol scale purification of 20-mer **7**, 31-mer **8**, 43-mer **10**, 61-mer **11** and 81-mer **12**, and 0.05 μmol scale purification of 151-mer **13**, 196-mer **14**, and 197-mer **15**, we consistently used 12 μl polymerization solution [3.7 M *N*,*N*-dimethylacrylamide and 0.18 M *N*,*N*′-methylenebis(acrylamide)]. The molar ratio of polymerization monomers [moles of *N*,*N*-

dimethylacrylamide and *N*,*N*′-methylenebis(acrylamide)] over the nucleotides in the ODNs (moles of ODN times the number of nucleosides) were 11.6 (0.2 μmol **7**), 7.5 (0.2 μmol **8**), 5.4 (0.2 μmol **10**), 3.8 (0.2 μmol **11**), 2.9 (0.2 μmol **12**), 6.2 (0.05 μmol **13**), 4.8 (0.05 μmol **14**), and 4.7 (0.05 μmol **15**). For the 50 μmol purification of 37-mer **9**, 1.25 ml polymerization solution was used, and the molar ratio was 2.6. Our guideline is that the molar ratio of polymerization monomer over nucleotides in the ODN should be equal to or slightly higher than 2.5. If we assume that an ODN contains equal numbers of the four nucleotides, the mass ratio of polymerization monomers over ODN is around 0.83, which means that purification of one gram of ODN requires less than one gram of polymerization monomers. Molar ratio lower than 2.5 is not suggested because the full-length sequences may not be caught completely and the polymerization may not be efficient due to diluted polymerization monomer concentration. For small scale purification of short sequences, when the molar ratio of 2.5 is used, the volume of polymerization solution may be too tiny for convenient manipulation. As a result, a larger ratio is suggested. In our examples, 12 μl polymerization solution was consistently used in all the small scale purifications. For larger scale purification, molar ratio much higher than 2.5 is not recommended. This is not from a cost consideration. Compared to ODN, the expense of the polymerization monomers is too low to be a concern. Moreover, the polyacrylamide waste can be potentially used as an excellent slow releasing fertilizer although its environmental effects have to be evaluated first. The recommendation is instead a result of consideration of ODN extraction efficiency. With large amount of gel, extraction of ODN after washing failure sequences requires more water, longer time and more times. The larger volume of extracts also consumes more energy for evaporation.

Besides the mass ratio of polymerization monomers over ODN, there are other mass ratios that need to be decided although these ratios are less critical for success. In all the above purification examples, the molar ratio of *N*,*N*-dimethylacrylamide over *N*,*N*′-methylenebis(acrylamide) was set to be 20:1. Other ratios could also be used, but ratios higher than 40:1 give gels that are too soft for convenient handling and ratios lower than 5:1 may reduce ODN extraction efficiency. For the volume of water used to dissolve ODN before adding polymerization solution, it can also be varied but should not be too far away from our examples. The initiators are usually more than enough even though only 5 μl of each were used in most of our examples. For the 50 μmol purification of **9**, we used 10 μl of each to ensure efficient polymerization. The molar ratio of TMEDA over ammonium persulfate was set to be 3:1 in our examples, but this may not be a requirement. For washing failure sequences from gel, we used water containing bases such as Et_3N, piperidine, NaOH and NaOAc, followed by pure water. To our impression, NaOAc solution consistently gave better results. Adding bases is important because the trityl ether linkage between ODN and gel is very labile even in the presence of trace acid. Base may also increase washing efficiency. The volume of the washing solutions can be large because the solutions are inexpensive and not harmful to environment, and do not require evaporation. For cleaving and extracting full-length sequences from gel, minimum amount of 80% AcOH or other acidic buffers and water that can cover gel should be used. Efficiency should not be achieved using large volume of solutions, instead it should be achieved through repetitions. Minimizing the total volume is important for saving time and energy for evaporation.

Conclusions

Four new methacrylation phosphoramidites that contain a linker cleavable with acetic acid solution have been synthesized, and used for the purification of ODN using the catching full-length

sequences by polymerization approach. The simplicity, convenience, and suitability of the technology with the new linker for small scale purification, large scale purification and long sequence purification have been demonstrated. Compared with the previously used hydrogen fluoride,[18] acetic acid solution is much more environmentally benign, less toxic and less expensive. Using the technology, we successfully purified a 37-mer ODN at 50 µmol scale. The procedure is as simple as purifying at sub-µmol scales. In addition, several long sequences including a 196-mer and a 197-mer were also purified. These long sequences are generally considered difficult to purify with RP HPLC. Using the new method, there is no difference between purifying the long sequences and purifying 20-mer sequences. In all the examples, ODN products were obtained in good purity and recovery yield. We believe that the work described in this paper is important to promote wide adoptions of the purification technology in academic labs, biotechnology companies and pharmaceutical industry.

Experimental Section

General procedures: All reagents and solvents were used from commercial sources as received unless otherwise indicated. Acetonitrile, CH_2Cl_2 and hexanes were dried using an Innovative Technology's PureSolv™ system. Pyridine was distilled over CaH_2 under nitrogen. All reactions were carried out in oven-dried glassware under nitrogen. ^1H, ^{13}C and ^{31}P NMR spectra were measured on a Varian UNITY INOVA spectrometer at 400, 100 and 162 MHz, respectively. High resolution mass spectra were obtained on an Agilent Q-TOF system using a dual electrospray ionization source. MALDI-TOF mass spectrometry experiments were performed on Waters' Tofspec-2E. ODNs were synthesized on an ABI 394 solid phase synthesizer. RP HPLC was carried out using the JASCO LC-2000Plus System with the pump PU-2089Plus Quaternary Gradient; detector UV-2075Plus; column C-18 analytical (5 µm diameter, 100 Å, 250 × 4.6 mm); solvent A, 0.1 M triethylammonium acetate, 5% acetonitrile; and solvent B, 90 % acetonitrile. Profiles were generated by detection of absorbance of ODN at 260 nm using the gradient system, solvent B (0%-45%) in solvent A over 60 min followed by solvent B (45%-100%) in solvent A over 20 min at a flow rate of 1 ml/min unless otherwise noted. In all purification examples, the solution that contains 3.7 M *N*,*N*-dimethylacrylamide and 0.18 M *N*,*N′*-methylenebis(acrylamide), which had a cross-linking ratio of 20:1, was used as the polymerization solution for catching full-length sequences. Solutions of 5% (0.22 M) $(NH_4)_2S_2O_8$ and 0.66 M TMEDA were used as polymerization initiators.

Compound 3: Compound **2**[22] (5.0 g, 16.33 mmol, 1.0 equiv), methyl 6-bromohexanoate (2.6 mL, 16.33 mmol, 1.0 equiv) and K_2CO_3 (9.0 g, 65.32 mmol, 4.0 equiv) were charged to a round-bottomed flask. Acetone (80 mL), which was dried simply by shaking with anhydrous sodium sulfate for 30 min, was added. The mixture was refluxed overnight. Volatiles were removed under reduced pressure. The residue was transferred into a separatory funnel and partitioned between ethyl acetate and water. The organic phase was dried over anhydrous Na_2SO_4, and filtered. The filtrate was concentrated to dryness under reduced pressure. The product was precipitated from methylene chloride with hexane giving **3** as a light yellow thick oil (5.9 g, 13.59 mmol, 83%; later, the reaction was performed on a scale of 13.5 g **2** and the yield was 84%): R_f = 0.4 (hexanes/EtOAc/Et_3N, 3:1:0.4); ^1H NMR (400 MHz, $CDCl_3$): δ 7.39-7.10 (m, 9H), 6.88-6.75 (m, 4H) 3.91 (t, J = 6.4 Hz, 2H), 3.71 (s, 3H), 3.61 (s, 3H), 3.49 (br s, 1H), 2.30 (t, J = 7.2 Hz, 2H), 1.84-1.63 (m, 4H), 1.57-1.41 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 174.3, 158.8, 158.3,

147.9 139.9, 139.8, 129.5, 128.1, 128.0, 127.2, 113.9, 113.3, 81.6, 67.8, 55.4, 51.7, 34.1, 29.2, 25.9, 24.9; HRMS (ESI) *m/z* calcd for $[M+Na]^+$ $C_{27}H_{30}NaO_5$ 457.1991, found 457.1984.

Compound 4: Compound **3** (3.0 g, 6.91 mmol, 1.0 equiv), 2,2′-(ethylenedioxy)bis(ethylamine) (3.1 mL, 20.73 mmol, 3.0 equiv) and water (0.7 mL) were charged to a round-bottomed flask. The mixture was stirred under nitrogen at 85 °C overnight. After cooling to rt, the mixture was concentrated and the residue was purified by flash column chromatography (SiO_2, Et_2O/MeOH/Et_3N, 3:2:0.2) giving **5** as a thick oil (2.85 g, 5.18 mmol, 75%; later, the reaction was performed on a scale of 15 g **3** and the yield was 76%): R_f = 0.5 (Et_2O/MeCN/MeOH/Et_3N, 5:2:2:1); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.26-7.08 (m, 9H), 6.80-6.70 (m, 4H), 6.39 (br t, *J* = 5.2 Hz, 1H), 3.86 (t, *J* = 6.4 Hz, 2H), 3.71 (s, 3H), 3.55-3.45 (m, 6H), 3.41-3.32 (m, 4H), 2.70 (t, *J* = 5.2 Hz, 2H), 2.11 (t, *J* = 7.2 Hz, 2H), 1.79-1.56 (m, 4H), 1.48-1.36 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 173.3, 158.6, 158.1, 147.9 140.1, 139.9, 128.4, 128.0, 127.9, 126.9, 113.7, 113.2, 81.3, 72.8, 70.3, 70.2, 70.1, 67.8, 55.4, 41.5, 39.3, 36.5, 29.2, 25.9, 25.6; HRMS (ESI) *m/z* calcd for $[M+H]^+$ $C_{32}H_{43}N_2O_6$ 551.3121, found 551.3126.

Compound 5: To an oven-dried round-bottomed flask under nitrogen was added **4** (2.7 g, 4.91 mmol, 1.0 equiv), dry CH_2Cl_2 (40 mL) and DIEA (2.6 mL, 14.73 mmol, 3.0 equiv). After cooling to 0 °C, the solution of methacrolyl chloride (479 µL, 4.91 mmol, 1.0 equiv) in 10 mL CH_2Cl_2 was added via an addition funnel slowly with vigorous stirring. After addition, the reaction flask was disconnected from the nitrogen and connected to air through a Drierite tube. The mixture was stirred at rt overnight. The contents were transferred into a separatory funnel and partitioned between 10% Na_2CO_3 and CH_2Cl_2. The organic phase was dried over anhydrous Na_2SO_4, filtered and concentrated to dryness under reduced pressure. The crude product was purified by flash column chromatography (SiO_2, hexanes/EtOAc/Et_3N, 1:5:0.2) giving **5** as a colorless thick oil (2.65 g, 4.29 mmol, 87%; later, the reaction was performed on a scale of 10 g **4** and the yield was the same): R_f = 0.3 (hexanes/EtOAc/Et_3N, 1:5:0.6); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.26-7.21 (m, 5H), 7.12 (t, *J* = 8.8 Hz, 4H), 6.76 (t, *J* = 8.8 Hz, 2H), 6.34 (br s, 1H), 6.12 (br s, 1H), 5.64 (s, 1H), 5.27 (s, 1H), 3.88 (t, *J* = 6.4 Hz, 2H), 3.74 (s, 3H), 3.54 (s, 4H), 3.52 (t, *J* = 5.2 Hz, 2H), 3.48 (t, *J* = 5.2 Hz, 2H), 3.44 (t, *J* = 5.2 Hz, 2H), 3.38 (pentet, *J* = 5.2 Hz, 2H), 2.14 (t, *J* = 6.4 Hz, 2H), 1.90 (s, 3H), 1.78-1.58 (m, 4H), 1.50-1.38 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 173.2, 168.7, 158.7, 158.2, 147.7 140.1, 139.8, 139.7, 129.4, 128.0, 127.9, 127.1, 119.8, 113.8, 113.3, 81.5, 70.4, 70.3, 70.1, 69.9, 67.8, 55.3, 39.5, 39.3, 36.7, 29.2, 25.9, 25.6, 18.8; HRMS (ESI) *m/z* calcd for $[M+Na]^+$ $C_{36}H_{46}N_2NaO_7$ 641.3203, found 641.3209.

Compound 6a: Compound **5** (2.0 g, 3.23 mmol, 1.0 equiv) and thymidine (782.8 mg, 3.23 mmol, 1.0 equiv) in two round-bottomed separate flasks were dried in a dessicator over Drierite under vacuum overnight. The flasks were connected to a Schlenk manifold and the contents were further dried using multiple cycles of evacuation and nitrogen filling. To the flask containing **5**, acetyl chloride (4 mL) was added, and the mixture was stirred at rt for 2 h. Most acetyl chloride was evaporated by a nitrogen flow or under vacuum. Dry hexanes were added, which after a brief shaking were removed using a syringe or cannula. The residue was washed with hexanes one more time, and then dried under vacuum for 1 h. Freshly distilled pyridine (10 mL) was added, and the solution was transferred via a cannula to the flask containing thymidine in pyridine (5 mL). The reaction mixture was stirred at rt overnight. Volatiles were evaporated under reduced pressure. The residue was transferred into a separatory funnel and partitioned between ethyl acetate and 10% Na_2CO_3. The organic phase was dried over anhydrous Na_2SO_4 and filtered. Volatiles were removed under reduced pressure. The product was purified with flash column chromatography

(SiO_2, EtOAc/MeOH/Et_3N, 9.5:0.5:0.2) giving **6a** as a white foam (2.1 g, 2.50 mmol, 77%): R_f = 0.4 (EtOAc/MeOH/Et_3N, 9:1:0.1); ^{1}H NMR (400 MHz, acetone-d_6): δ 7.62 (s, 1H), 7.52-7.46 (m, 2H), 7.40-7.16 (m, 5H), 6.98-6.80 (m, 4H), 6.36 (t, J = 6.0 Hz, 1H), 5.69 (s, 1H), 5.29 (t, J = 1.6 Hz, 1H), 4.62-4.56 (m, 1H), 4.09-4.01 (m, 2H), 3.96 (t, J = 6.0 Hz, 2H), 3.76 (s, 3H), 3.60-3.30 (m, 14H), 2.42-2.28 (m, 2H), 2.18 (t, J = 7.2 Hz, 2H), 1.95 (s, 3H), 1.89 (br s, 3H), 1.88-1.80 (m, 2H), 1.80-1.71 (m, 2H), 1.71-1.60 (m, 2H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.6, 168.0, 163.8, 159.0, 158.7, 158.5, 150.7, 145.3, 140.8, 135.9 (3C), 130.4 (4C), 128.4, 128.1, 127.1, 118.6, 113.9 (2C), 113.3 (2C), 110.3, 86.7, 86.5, 84.6, 71.8, 70.3, 70.2, 69.9, 69.5, 67.6, 64.2, 54.9, 40.5, 39.4, 39.2, 35.9, 25.7, 25.5, 18.3, 11.6; HRMS (ESI) *m/z* calcd for $[M+Na]^+$ $C_{46}H_{59}N_4NaO_{11}$ 865.4000, found 865.3998.

Compound 6b: Compound **6b** was synthesized similarly as **6a** using **5** (218 mg, 0.35 mmol, 1.0 equiv) and N^2-isobutyryl-2'-deoxyguanosine (118.9 mg, 0.35 mmol, 1.0 equiv). The product was purified with flash column chromatography (SiO_2, EtOAc/MeOH/Et_3N, 8:2:0.5) giving **6b** as a white foam (250 mg, 0.27 mmol, 76%): R_f = 0.4 (EtOAc/MeOH/Et_3N, 8:2:0.5); ^{1}H NMR (400 MHz, acetone-d_6): δ 8.08 (s, 1H), 8.05 (s, 1H), 7.58-7.10 (m, 9H), 6.82-6.68 (m, 4H), 6.33-6.25 (m, 1H), 5.72 (s, 1H), 5.59 (s, 1H), 5.30 (s, 1H), 4.72 (s, 1H), 4.70 (s, 1H), 4.30-4.10 (m, 2H), 3.98-3.86 (m, 1H), 3.71 (s, 3H), 3.70-3.15 (m, 10H), 2.96 (pentet, J = 6.4 Hz, 2H), 2.81-2.70 (m, 2H), 2.54-2.45 (m, 2H), 2.24 (t, J = 6.4 Hz, 2H), 2.06-2.01 (m, 1H), 1.90 (s, 3H), 1.80-1.60 (m, 4H), 1.52-1.40 (m, 2H), 1.26-1.17 (m, 6H); ^{13}C NMR (100 MHz, acetone-d_6): δ 180.5, 173.4, 168.4, 158.8, 158.3, 155.8, 148.8, 148.6, 145.5, 140.6, 137.9, 136.2, 135.9, 130.3, 128.3, 127.8, 126.8, 121.2, 119.1, 86.9, 86.2, 84.2, 71.5, 70.3, 70.2, 69.8, 69.5, 67.7, 64.3, 54.9, 39.9, 39.5, 39.3, 36.0, 35.8, 25.8, 25.6, 20.2, 18.9, 18.8, 18.4; HRMS (ESI) *m/z* calcd for $[M+H]^+$ $C_{50}H_{64}N_7O_{11}$ 938.4664, found 938.4662, and calcd for $[M+Na]^+$ $C_{50}H_{63}N_7NaO_{11}$ 960.4483, found 960.4475.

Compound 6c: The compound **6c** was synthesized similarly as **6a** using **5** (375 mg, 0.61 mmol, 1.0 equiv) and N^4-acetyl-2'-deoxycytidine (164.2 mg, 0.61 mmol, 1.0 equiv). The product was purified with flash column chromatography (SiO_2, EtOAc/MeOH/Et_3N, 9:1:0.5) giving **6c** as a white foam (402 mg, 0.46 mmol, 76%): R_f = 0.4 (EtOAc/MeOH/Et_3N, 9:1:0.5); ^{1}H NMR (400 MHz, acetone-d_6): δ 8.26 (d, J = 7.6 Hz, 1H), 7.46 (d, J = 7.2 Hz, 1H), 7.40-7.19 (m, 9H), 7.00-6.71 (m, 4H), 6.20 (t, J = 6.0 Hz, 1H), 5.70 (s, 1H), 5.29 (s, 1H), 4.61-4.50 (m, 1H), 4.15-3.31 (m, 17H), 3.77 (s, 3H), 2.80-2.52 (m, 4H), 2.21 (s, 3H), 1.89 (s, 3H), 1.74 (pentet, J = 7.6 Hz, 2H), 1.64 (pentet, J = 7.6 Hz, 2H), 1.54-1.40 (m, 2H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.7, 170.8, 168.1, 162.9, 159.0, 158.5, 155.1, 144.8, 144.7, 140.7, 136.0, 136.2, 130.3 (4C), 128.3 (2C), 128.0 (2C), 127.0, 118.7, 113.9 (2C), 113.4 (2C), 95.7, 87.0, 86.8, 86.7, 70.3, 70.2, 69.9, 69.5, 67.8, 63.3, 63.2, 59.9, 42.0, 39.4, 39.2, 35.9, 25.8, 25.6, 24.2, 18.4, 18.3; HRMS (ESI) *m/z* calcd for $[M+H]^+$ $C_{47}H_{60}N_5O_{11}$ 870.4289, found 870.4277, and calcd for $[M+Na]^+$ $C_{47}H_{59}N_5NaO_{11}$ 892.4109, found 892.4098.

Compound 6d: The compound **6d** was similarly as **6a** using **5** (363 mg, 0.59 mmol, 1.0 equiv) and N^6-benzoyl-2'-deoxyadenosine (209.7 mg, 0.59 mmol, 1.0 equiv). The product was purified with flash column chromatography (SiO_2, EtOAc/MeOH/Et_3N, 8:2:0.5) giving **6d** as a white foam (422 mg, 0.44 mmol, 75%): R_f = 0.4 (EtOAc/MeOH/Et_3N, 8:2:0.5); ^{1}H NMR (400 MHz, acetone-d_6): δ 8.56 (s, 1H), 8.41 (s, 1H), 8.11 (d, J = 7.2 Hz, 2H), 7.61 (tt, J = 7.2, 2.4 Hz, 1H), 7.52 (dt, J = 7.2, 1.2 Hz, 2H), 7.42 (dd, J = 7.2, 1.6 Hz, 2H), 7.32-7.11 (m, 7H), 6.80-6.72 (m, 4H), 6.56 (t, J = 6.4 Hz, 1H), 5.68 (s, 1H), 5.27 (t, J = 1.6 Hz, 1H), 4.19 (q, J = 3.6 Hz, 1H), 3.96-3.86 (m, 2H), 3.73 (d, J = 4.0 Hz, 2H), 3.59 (t, J = 6.0 Hz, 1H), 3.51 (s, 3H), 3.50 (t, J = 5.6

Hz, 2H), 3.46 (t, *J* = 5.2 Hz, 2H), 3.36 (pentet, *J* = 5.6 Hz, 2H), 3.31 (pentet, *J* = 6.0 Hz, 2H), 3.04 (pentet, *J* = 6.8 Hz, 1H), 3.92-3.80 (m, 1H), 2.63 (t, *J* = 5.2 Hz, 2H), 2.56-2.42 (m, 2H), 2.15 (dt, *J* = 7.2, 2.0 Hz, 2H), 1.88 (s, 3H), 1.74-1.66 (m, 2H), 1.61 (t, *J* = 7.6 Hz, 2H), 1.48-1.37 (m, 2H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.3, 168.1, 165.8, 158.9, 158.4, 152.1, 151.8, 150.5, 145.0, 142.9, 140.7, 136.3, 136.1, 135.9, 134.4, 132.6, 130.4, 130.3, 130.2, 128.7, 128.6, 128.3, 127.9, 126.8, 125.4, 118.1, 113.7, 113.2, 86.9, 86.3, 84.9, 71.9, 70.3, 70.2, 69.8, 69.5, 67.7, 64.4, 54.9, 39.6, 39.5, 39.2, 36.0, 28.7, 25.8, 25.5, 18.3; HRMS (ESI) *m/z* calcd for $[M+H]^+$ $C_{53}H_{62}N_7O_{10}$ 956.4558, found 956.4553, and calcd for $[M+Na]^+$ $C_{53}H_{61}N_7NaO_{10}$ 978.4378, found 978.4370.

Phosphoramidite 1a: An oven-dried round-bottomed flask containing **6a** (0.7 g, 0.83 mmol, 1.0 equiv) was dried by four cycles of evacuation and refilling nitrogen. Freshly distilled CH_2Cl_2 (3 mL) was added via a syringe. 2-Cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphoramidite (289 µL, 0.91 mmol, 1.1 equiv) and diisopropyl ammonium tetrazolide (142.2 mg, 0.83 mmol, 1.0 equiv) were added sequentially. After stirring at rt for 5 h, the mixture was concentrated to dryness with a nitrogen flow. The residue was purified with flash column chromatography (SiO_2, EtOAc/acetone/Et_3N, 9:1:0.5) to give **1a** as a light yellow oil (0.75 g, 0.72 mmol, 87%): R_f = 0.5 (EtOAc/acetone/Et_3N, 7:3:0.5); ^{1}H NMR (400 MHz, acetone-d_6): δ 7.62 (d, *J* = 11.6 Hz, 1H), 7.50-7.47 (m, 2H), 7.38-7.30 (m, 4H), 7.29-7.21 (m, 2H), 7.20-7.10 (m, 1H), 6.92-6.80 (m, 4H), 6.36 (dt, *J* = 6.4, 2.4 Hz, 1H), 5.69 (s, 1H), 5.29 (t, *J* = 1.2 Hz, 1H), 4.76 (br s,1H), 4.22-4.12 (m, 1H), 3.98-3.92 (m, 2H), 3.54 (s, 3H), 3.90-3.30 (m, 17H), 2.75 (t, *J* = 6.0 Hz 2H), 2.62 (t, *J* = 6.4 Hz, 2H), 2.53-2.40 (m, 2H), 2.18 (t, *J* = 7.2 Hz, 2H), 1.90 (s, 3H), 1.76 (pentet, *J* = 7.6 Hz, 2H), 1.64 (pentet, *J* = 7.2 Hz, 2H), 1.51-1.43 (m, 5H), 1.23-1.07 (m, 12H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.4, 167.9, 163.7, 159.1, 158.6, 150.6, 145.2, 140.8, 135.8, 130.4, 128.4, 128.1, 127.1, 118.5, 118.3, 113.9, 113.4, 110.5, 86.8, 85.5, 84.6, 73.8, 70.4, 70.2, 69.9, 69.6, 67.8, 63.6, 58.9, 58.7, 54.9, 43.3, 39.4, 39.1, 35.9, 25.8, 25.5, 24.3, 24.2, 20.1, 18.3, 11.6; ^{31}P NMR (160 MHz, acetone-d_6): δ 149.3, 149.2; MS (ESI) *m/z* calcd for $[M+Na]^+$ $C_{55}H_{75}N_6NaO_{12}P$ 1065.5, found 1065.4.

Phosphoramidite 1b: Compound **1b** was synthesized similarly as **1a** using **6b** (240 mg, 0.26 mmol, 1.0 equiv), 2-cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphoramidite (92.1 µL, 0.29 mmol, 1.1 equiv) and diisopropyl ammonium tetrazolide (44.5 mg, 0.26 mmol, 1.0 equiv). The product was purified by flash column chromatography (SiO_2, EtOAc/acetone/Et_3N, 3:1:0.2) giving **1b** as a light yellow oil (250 mg, 0.22 mmol, 86%): R_f = 0.5 (EtOAc/MeOH/Et_3N, 9:1:0.5); ^{1}H NMR (400 MHz, acetone-d_6): δ 8.05 (s, 1H), 8.01 (d, *J* = 4.8 Hz, 1H), 7.45-7.11 (m, 9H), 6.85-6.41 (m, 4H), 6.30 (t, *J* = 6.4 Hz, 1H), 5.72 (s, 1H), 5.29 (s, 1H), 5.78 (br s, 2H), 4.40-4.18 (m, 3H), 4.08-3.29 (m, 20H), 3.20-2.52 (m, 7H), 2.40-2.18 (m, 2H), 1.91 (s, 3H), 1.80-1.61 (m, 4H), 1.58- 1.42 (m, 2H), 1.32-1.12 (m, 18H); ^{13}C NMR (100 MHz, acetone-d_6): δ 180.3, 172.9, 170.3, 158.9, 158.4, 155.5, 148.8, 148.6, 137.5, 130.3, 129.0, 128.9, 128.3, 127.9, 126.9, 121.5, 118.8, 118.5, 113.7, 113.2, 86.4, 84.2, 84.0, 74.4 (d, *J* = 16.7 Hz), 73.8 (d, *J* = 18.2 Hz), 70.4, 70.2, 69.9, 69.6, 67.7, 63.8 (d, *J* = 10.7 Hz), 58.9, 58.8 (d, *J* = 8.2 Hz), 57.6, 54.9, 43.3, 43.2, 39.5, 39.2, 35.8, 25.8, 25.6, 24.3, 20.2, 18.7; ^{31}P NMR (160 MHz, acetone-d_6): δ 149.7, 149.5, 149.4, 149.1; MS (ESI) *m/z* calcd for $[M+H]^+$ $C_{59}H_{81}N_9O_{12}P$ 1138.6, found 1138.3.

Phosphoramidite 1c: The compound **1c** was synthesized similaryly as **1a** using **6c** (225 mg, 0.26 mmol, 1.0 equiv), 2-cyanoethyl-*N*,*N*,*N'*,*N'*-tetraisopropylphosphoramidite (92.1 µL, 0.29 mmol, 1.1 equiv) and diisopropyl ammonium tetrazolide (44.5 mg, 0.26 mmol, 1.0 equiv). The product was purified by flash column chromatography (SiO_2, EtOAc/acetone/Et_3N, 3:1:0.2) giving **1c** as a light yellow oil (223 mg, 0.21 mmol, 81%): R_f = 0.4 (EtOAc/acetone/Et_3N, 1:1:0.1); ^{1}H

NMR (400 MHz, acetone-d_6): δ 8.36-8.22 (m, 1H), 7.47 (t, J = 6.8 Hz, 1H), 7.38-7.16 (m, 9H), 6.98-6.84 (m, 4H), 6.22 (pentet, J = 5.2 Hz, 1H), 5.70 (s, 1H), 5.29 (t, J = 1.2 Hz, 2H), 4.78-4.61 (m, 1H), 4.30-4.16 (m, 1H), 4.08-3.92 (m, 2H), 3.90-3.10 (m, 18H), 2.82-2.52 (m, 6H), 2.42-2.32 (m, 1H), 2.32-2.12 (m, 3H), 2.04 (s, 3H), 1.90 (s, 3H), 1.80-1.70 (m, 2H), 1.70-1.59 (m, 2H), 1.52-1.40 (m, 2H), 1.40-1.00 (m, 12H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.6, 169.2, 168.0, 163.0, 159.0, 158.6, 154.9, 145.1, 144.8, 140.8, 135.8, 135.5, 130.4, 128.4, 128.1, 127.1, 118.6, 118.3, 113.9, 113.4, 95.8, 86.9, 70.4, 70.3, 69.9, 69.6, 67.8, 62.6, 59.0, 54.9, 43.3, 39.4, 39.2, 35.9, 33.9, 25.8, 25.6, 24.4, 24.5, 24.6, 22.4, 18.3, 14.6; ^{31}P NMR (160 MHz, acetone-d_6): δ 149.7, 149.6, 149.4, 149.1; MS (ESI) *m/z* calcd for $[M+H]^+$ $C_{56}H_{77}N_7O_{12}P$ 1070.5, found 1070.1.

Phosphoramidite 1d: Compound **1d** was synthesized similarly as **1a** using **6d** (300 mg, 0.31 mmol, 1.0 equiv), 2-cyanoethyl-*N*,*N*,*N*′,*N*′-tetraisopropylphosphoramidite (108 µL, 0.34 mmol, 1.1 equiv) and diisopropyl ammonium tetrazolide (53.1 mg, 0.31 mmol, 1.0 equiv). The residue was purified by flash column chromatography (SiO_2, EtOAc/acetone/Et_3N, 3:1:0.2) giving **1d** as a light yellow oil (280 mg, 0.24 mmol, 77%): R_f = 0.5 (EtOAc/acetone/Et_3N, 1:1:0.1); ^{1}H NMR (400 MHz, acetone-d_6): δ 8.57 (d, J = 2.8 Hz, 1H), 8.44 (d, J = 3.6 Hz, 1H), 8.11 (d, J = 8.0 Hz, 2H), 7.58 (t, J = 7.2 Hz, 1H), 7.49 (t, J = 7.6 Hz, 2H), 7.47-7.40 (m, 2H), 7.40-7.03 (m, 7H), 6.90-6.72 (m, 4H), 6.66-6.52 (m, 1H), 5.64 (s, 1H), 5.27 (s, 1H), 5.06-4.91 (m, 1H), 4.39-4.26 (m, 1H), 3.99-3.11 (m, 21H), 2.89-2.55 (m, 5H), 2.25-2.11 (m, 2H), 1.89 (s, 3H), 1.79-1.55 (m, 4H), 1.52-1.39 (m, 2H), 1.40-1.11 (m, 12H); ^{13}C NMR (100 MHz, acetone-d_6): δ 172.7, 169.5, 168.0, 158.9, 158.4, 152.1, 151.8, 150.7, 145.5, 143.2, 140.7, 136.2, 135.9, 134.4, 132.6, 130.3, 130.2, 128.7, 128.6, 128.3, 127.9, 126.9, 125.6, 118.5, 118.7, 113.7, 113.1, 86.4, 85.9, 85.8, 85.0, 74.2 (d, J = 18.2 Hz), 73.6 (d, J = 18.3 Hz), 70.4, 70.3, 69.9, 69.6, 67.7, 64.0, 63.9, 59.9, 59.2, 59.1, 59.0, 58.9, 54.9, 43.4, 43.3, 39.5, 39.2, 36.0, 34.0, 25.8, 25.6, 24.4, 24.3, 20.2 (d, J = 12.2 Hz), 20.1 (d, J = 18.3 Hz), 18.4; ^{31}P NMR (160 MHz, acetone-d_6): δ 149.4, 149.2; MS (ESI) *m/z* calcd for $[M+Na]^+$ $C_{62}H_{78}N_9NaO_{11}P$ 1178.5, found 1178.4.

Small scale ODN purification

ODN synthesis: The 20-mer ODN **7** was used as example. Its synthesis was carried out on a 1 µmol scale under standard conditions using phosphoramidite chemistry. The synthesizer manufacturer suggested 1 µmol synthetic cycle was copied to a new file. After the coupling steps and before the capping steps, a 180 sec waiting step was inserted, and was set to be active only for base 5. The 0.1 M solution of **1a** in dry acetonitrile was attached to bottle position 5. The volume should be around 500 µl, which was needed for running the bottle exchange cycle, the begin cycle (optional), as well as the coupling steps in one cycle. When editing sequence, the dT at the 5′-end was input as 5. CPG with a long chain alkylamine succinyl ester linker and a glass pore size of 500 Å was used as solid support. The 5′-DMTr 2-cyanoethyl benzoyl-dA, isobutyryl-dG, acetyl-dC, and dT phosphoramidites (0.1 M acetonitrile solutions) were used as coupling monomers. The 0.25 M DCI solution in acetonitrile was used as activator. Failure sequences were capped with Ac_2O in each synthetic cycle. At the end of synthesis, detritylation was not carried out as this would remove the methacrylamide tag. Trityl assay indicated an overall yield of 86% for the 19-mer. Cleavage was carried out on the synthesizer with conc. NH_4OH. Deprotection was achieved by heating the solution to 55 °C for 15 h. After cooling to rt, 200 µl DIEA was added. The solution was divided into 5 equal portions, and evaporated to dryness under vacuum, separately. One portion (~0.2 µmol) was dissolved in 300 µl water, 20 µl was injected into RP HPLC to generate trace a (Figure 2).

Polymerization: To the second portion (~0.2 μmol) of ODN in a centrifuge tube was added 50 μl water. The tube was vortexed and spun shortly for efficient dissolution. Polymerization solution [12 μl, 44.4 μmol, 3.70 M *N*,*N*-dimethylacrylamide, and 4.4 μmol, 0.18 M *N*,*N*′-methylenebis(acrylamide)], and initiators $(NH_4)_2S_2O_8$ (5%, 5 μl, 1.1 μmol) and TMEDA (0.66 M, 5 μl, 3.3 μmol) were added sequentially. The mixture was vortexed and spun briefly, and then transferred immediately into the top compartment of a centrifugal filter unit. The bottom of the compartment could be stopped with a cap or parafilm, but the solution will not pass through the filter disc even without stopping. Care needs to be taken to transfer all solution to the center of the filter disc, which can be achieved with steady releasing and pushing the pipette. Washing the residue solution to the compartment is not recommended as this would dilute the polymerization solution. The solution was allowed to stand at rt with the cap of the unit closed. A gel was formed in ~15 min, but the polymerization was allowed to continue for 1 h.

Washing: The gel was loosened with a clean spatula, and water containing 5% Et_3N (~250 μl) was added. After 3 min, the liquid was removed with a spin. The washing was repeated six times with the last wash using pure water.

Cleavage and extraction: A minimum amount of AcOH solution (80%) that could cover the gel (~100 μl) was added. After 5 min, the liquid was collected with a spin in a centrifuge. The procedure was repeated two more times. A minimum amount of water that could cover the gel (~100 μl) was added, and the extract was collected with a spin. The extraction was repeated four times. The extracts including those of AcOH solution and water were combined and evaporated to dryness. To the residue, concentrated NH_4OH (100 μl) was added. After a short vortex, the closed tube was heated at 80 ^oC for 15 min (later we found that heating is not needed). The tube was cooled to rt, and *n*-BuOH (900 μl) was added. The mixture was vortexed for 30 sec and centrifuged for 3 min at 14.1K. The supernatant was carefully removed with a pipette, and the residue, which was the pure ODN, was dried under vacuum to evaporate residue *n*-BuOH. The ODN was dissolved in 300 μL water, vortexed, and spun, and 20 μl was injected into RP HPLC to generate trace b (Figure 2). The recovery yield of the ODN was estimated to be 68% by dividing the area of the peak in trace b at 19 minutes by that in trace a at 45 minutes. ESI MS are in ESI.

The 31-mer **8** and 43-mer **10** were synthesized under the same conditions for **7** except that cleavage and deprotection were achieved using a mixture of conc. NH_4OH and $MeNH_2$ solutions (1:1 v/v) at rt for 3 h. The ODNs were purified at 0.2 μmol scale using the same procedure and the same amount of reagents for purifying **7**. The RP HPLC traces for crude and pure ODNs are in ESI. The recovery yields of purification were estimated to be 71% (**8**) and 93% (**10**) by comparing areas of peaks in the HPLC profiles of pure and crude ODN. MALDI-TOF MS are in ESI. ODN **10** was also purified using the acidic buffers 0.1 mM Et_3NHCl and 10 mM $Et_3NHO_2CCF_3$ at pH 3.0 to cleave full-length sequences from gel under otherwise identical conditions. The profiles of RP HPLC analysis of pure ODN are in ESI.

Large scale ODN purification

The 37-mer ODN **9** was used as example. Its synthesis was achieved in five 10 μmol columns with 1000 Å CPG. The same procedure and reagents for synthesizing **7** were used except that a different synthetic cycle was used. The standard 10 μmol synthesis cycle was copied into a new file. After the coupling steps and before the capping steps, a 180-second waiting step was inserted, and was set to be active only for base 5. Approximately, a total of 5.9 ml of 0.08 M (0.1 M would be better) solution of **1a** was used for tagging full-length sequences. According to trityl

essay before the cycle for attaching **1**, the average yield of 36-mer for the five synthesis was 50%. Cleavage and deprotection were achieved under the same conditions for **8**. DIEA (~500 µl) was added to minimize the possibility of premature tag cleavage during evaporation. RP HPLC analysis of the crude indicated that only about 50% of the 36-mer intermediate was successfully methacrylated with **1a** (ESI). As a result, the synthesis yield of the 37-mer was 25%. For purification, the crude ODN was dissolved in 1.25 ml water in a 25 ml round-bottomed flask, which was flushed with nitrogen (Figure 3). Polymerization solution (1.25 ml) and initiator solutions (10 µl each) were added sequentially. The flask was hand-shaken briefly, and polymerization was allowed to proceed at rt for 1 h. The gel was transferred into a 15 ml Büchner funnel with a sintered glass disc and broken into several pieces with a spatula. Water (~4 ml) containing 5% Et_3N was added to the gel. After 3 min, vacuum was applied and the liquid was removed. The washing was repeated 4 times. The gel was further washed with water (~4 ml) containing 1% Et_3N for 4 times without waiting between the washes, and finally, the gel was washed with pure water (~ 4 ml). For cleaving full-length sequences, 80% AcOH (~4 ml) was added into the funnel. After 5 min, vacuum was applied and the extract was collected in a centrifuge tube placed in the filter flask. The procedure was repeated two times. The gel was further washed with water (4 ml × 5) without waiting between washes. The extracts including AcOH solution and water were combined and evaporated to dryness under vacuum. The residue was dissolved in 1 ml concentrated NH_4OH and heated at 80 ^oC for 15 min. After cooling to rt, *n*-BuOH (9 ml) was added, and the mixture was vortexed for 30 sec and centrifuged at 14K for 3 min. The supernatant was removed, and the residue was dried under vacuum. Pure ODN was obtained as white cotton-like fluffy fibers (Figure 3). The weight was 98 mg, which corresponds to a 70% recovery yield of the purification procedure. A portion of the ODN was dissolved in water, and analyzed with RP HPLC (ESI). The images of MALDI-TOF MS and ^1H and ^{31}P NMR of the pure ODN are in ESI.

Long sequence purification

The sequences **11** (61-mer), **12** (81-mer), **13** (151-mer), **14** (196-mer) and **15** (197-mer) were used to demonstrate the power of the technique for purification of long ODN sequences. ODN **15** is used for the description. The ODN was synthesized at 0.2 µmol scale under similar conditions for **7** using **1c** as polymerizable tagging agent with slight modifications. The standard 0.2 µmol (instead of 1 µmol) synthetic cycle was used to create the new cycle. CPG with a pore size of 2000 Å was used as solid support. Before synthesis, the capping solutions were manually delivered to fill the column, and the mixture was allowed to stand at rt for 20 min. The CPG was then washed with acetonitrile. The synthesis yield before the last cycle was 9% according to trityl essay. The ODN was cleaved and deprotected using the mixture of conc. NH_4OH and $MeNH_2$ (1:1 v/v) at rt for 3 h. About 200 µl DIEA was added, and the solution was divided into four equal portions, which were evaporated to dryness under vacuum. One portion was dissolved in 50 µl water, and 10 µl was injected into RP HPC to generate trace c (Figure 2). A second portion was used for purification. The polymerization step including reagent volumes was exactly the same as for **7** except that the water used for dissolving ODN and the polymerization solution contained 7 M urea. In the washing step, 3.0 M NaOAc (200 µl × 6 with 3 min waiting between the washes), water containing 5% Et_3N (200 µl × 1) and pure water (200 µl × 1) were used, sequentially. Cleavage, extraction, and precipitation were carried out under the same conditions for **7**. The pure ODN was dissolved in 20 µl water, and all was injected into RP HPLC to give trace d. The recovery yield of the purification process was estimated to be 78%.

The long sequences **11**-**14** was synthesized using **1a**-**b** or **1d** as polymerizable tagging agent and purified under the same conditions described for **15**. According to trityl assay before the last cycle for tagging full-length sequence with **1a**-**b** or **1d**, the synthesis yields were 83% (**11**), 48% (**12**), 22% (**13**), and 35% (**14**). For purification, the recovery yields were estimated to be 83% (**11**), 80% (**12**), 50% (**13**) and 82% (**14**) by comparing peak areas in crude and pure HPLC profiles (theoretically more concentrated pure than crude ODN was injected). RP HPLC profiles are in ESI. In addition to use 5% Et_3N and 3.0 M NaOAc to wash away failure sequences from gel, 10% piperidine in DMF, 10% piperidine in water and 1.0 M NaOH were also tested. While all gave acceptable results, 3.0 M NaOAc gave higher recovery yields, and narrower HPLC peak. ESI MS of **12** is in ESI.

Acknowledgements: Financial support from US NSF (CHE-0647129 and CHE-1111192), Michigan Translational Research & Commercialization (M-TRAC) of Michigan Economic Development Corporation, Michigan Tech Research Excellence Fund for Technology Commercialization (REF-TC), and Michigan Universities Commercialization Initiatives (MUCI); assistance from Mr. Dean W. Seppala (electronics) and Mr. Jerry L. Lutz (NMR); and NSF equipment grants (CHE-1048655 and CHE-9512455) are gratefully acknowledged.

References

1 Y. Singh, P. Murat and E. Defrancq, *Chem. Soc. Rev.*, 2010, **39**, 2054.
2 R. L. Juliano, X. Ming and O. Nakagawa, *Accounts Chem. Res.*, 2012, **45**, 1067.
3 A. Fiszer and W. J. Krzyzosiak, *Nucleic Acids Res.*, 2014, **42**, 6787.
4 S. Grijalvo, A. Avino and R. Eritja, *Expert Opin. Ther. Pat.*, 2014, **24**, 801.
5 D. Castanotto and C. A. Stein, *Curr. Opin. Oncol.*, 2014, **26**, 584.
6 A. Y. Borovkov, A. V. Loskutov, M. D. Robida, K. M. Day, J. A. Cano, T. Le Olson, H. Patel, K. Brown, P. D. Hunter and K. F. Sykes, *Nucleic Acids Res.*, 2010, **38**, 1.
7 S. Y. Ma, N. Tang and J. D. Tian, *Curr. Opin. Chem. Biol.*, 2012, **16**, 260.
8 S. Mueller, J. R. Coleman and E. Wimmer, *Chem. Biol.*, 2009, **16**, 337.
9 S. Kosuri and G. M. Church, *Nat. Methods*, 2014, **11**, 499.
10 J. D. Tian, K. S. Ma and I. Saaem, *Mol. Biosyst.*, 2009, **5**, 714.
11 R. A. Hughes, A. E. Miklos and A. D. Ellington, *Synth. Biol., Part B*, 2011, **498**, 277.
12 T. Abramova, *Molecules*, 2013, **18**, 1063.
13 M. Schulte, N. Luhring, A. Keil and Y. S. Sanghvi, *Org. Process Res. Dev.*, 2005, **9**, 212.
14 Y. S. Sanghvi and M. Schulte, *Curr. Opin. Drug Disc.*, 2004, **7**, 765.
15 R. R. Deshmukh, W. E. Leitch and D. L. Cole, *J. Chromatogr. A*, 1998, **806**, 77.
16 R. R. Deshmukh, T. N. Warner, F. Hutchison, M. Murphy, W. E. Leitch, P. De Leon, G. S. Srivatsa, D. L. Cole and Y. S. Sanghvi, *J. Chromatogr. A*, 2000, **890**, 179.
17 S. Y. Fang, S. Fueangfung, X. Lin, X. Zhang, W. Mai, L. Bi and S. A. Green, *Chem. Commun.*, 2011, **47**, 1345.
18 S. Y. Fang and S. Fueangfung, *Org. Lett.*, 2010, **12**, 3720.
19 D. Pokharel, Y. N. Yuan, S. Fueangfung and S. Y. Fang, *RSC Adv.*, 2014, **4**, 8746.
20 D. Pokharel and S. Y. Fang, *Open Org. Chem. J.*, 2014, **8**, 15.
21 Y. N. Yuan, S. Fueangfung, X. Lin, D. Pokharel and S. Y. Fang, *RSC Adv.*, 2012, **2**, 2803.

22 D. J. Dellinger, B. E. Boyes and G. R. Nicol, *U.S. Pat. Appl. Publ., 20050186580, 25 Aug 2005*, 2005,
23 W. Tang and S. Y. Fang, *Tetrahedron Lett.*, 2008, **49**, 6003.
24 K. Pappas, X. Zhang, W. Tang and S. Y. Fang, *Tetrahedron Lett.*, 2009, **50**, 5741.
25 K. Pappas, X. A. Zhang and S. Y. Fang, *Lett. Org. Chem.*, 2010, **7**, 539.
26 W. H. Pearson, D. A. Berry, P. Stoy, K. Y. Jung and A. D. Sercel, *J. Org. Chem.*, 2005, **70**, 7114.
27 D. C. Capaldi, H. Gans, A. H. Krotz, J. Arnold, R. L. Carty, M. N. Moore, A. N. Scozzari, K. Lowery, D. L. Cole and V. T. Ravikumar, *Org. Process Res. Dev.*, 2003, **7**, 832.