

## Purification and Sequencing of DNA Guides from Prokaryotic Argonaute

Daan C. Swarts\*, Edze R. Westra, Stan J. J. Brouns and John van der Oost

Department of Agrotechnology and Food Sciences, Wageningen University, Wageningen, the Netherlands

\*For correspondence: [daan.swarts@wur.nl](mailto:daan.swarts@wur.nl)

**[Abstract]** Some proteins utilize nucleic acids to guide them to complementary nucleic acid targets. One example is prokaryotic Argonaute protein, which, binds small single stranded DNA molecules as guides (Swarts et al., 2014). This protocol describes a method to purify DNA guides from these proteins. It also describes a PCR-based method to enrich the guides by PCR amplification. This methods relies on addition of a poly-A tail at the 3'-end of the ssDNA molecules by Terminal Deoxynucleotidyl Transferase (TdT), followed by ligation of a oligonucleotide to the 5'-end of the ssDNA molecule using T4 RNA ligase, and amplification by PCR. The generated dsDNA products are suitable for traditional cloning and sequencing and high-throughput sequencing. Importantly, the information which strand matches the ssDNA molecule is not lost during this process.

### **Materials and Reagents**

#### A. Purifying nucleic acids from proteins

1. Purified protein and co-purified nucleic acids in purification buffer (e.g. *TtAgo* with siDNA guides)

*Note: Please refer to the protocol “[Expression and Purification of the Thermus thermophilus Argonaute Protein](#)” (Swarts et al., 2014b).*

2. Proteinase K solution (20 mg/ml) (Life Technologies, Ambion®, catalog number: AM2548)

3. CaCl<sub>2</sub> solution (50 μM)

4. Roti phenol/chloroform/isoamyl alcohol (pH 7.5-8.0) (Carl Roth, catalog number: A156)

5. 99% ethanol

6. 70% ethanol (pre-cooled to -20 °C)

7. Linear acrylamide (5 mg/ml) (Life Technologies, Ambion®, catalog number: AM9520)

*Note: Alternatively, a house-made Linear Acrylamide can be used (<http://www.uvm.edu/~tpdelane/lab/protocols/LinearPolyAcryl.htm>).*

8. RNase free MQ water

9. Additional materials and reagents required for analysis of purified nucleic acids (optional)

- a. T4 polynucleotide kinase (PNK) (New England Biolabs, catalog number M0201)

- b. T4 polynucleotide kinase reaction buffer (New England Biolabs, catalog number M0201)
  
- B. Enriching and preparing ssDNA molecules for sequencing
  1. Purified single stranded DNA nucleotides in MQ water (e.g. see Procedure part A "purifying nucleic acids from proteins")
  2. RNase A (DNase and protease-free) (10 mg/ml) (Thermo Fisher Scientific, catalog number: EN0531)
  3. Terminal Deoxynucleotidyl transferase (TdT) (recombinant) (Life Technologies, Invitrogen™, catalog number: 10533-073)
  4. Terminal Deoxynucleotidyl transferase (TdT) buffer (Life Technologies, Invitrogen™, catalog number: 16314-015)
  5. dATP (100 mM) (Thermo Fisher Scientific, catalog number: R0141)
  6. QIAquick nucleotide removal kit (QIAGEN, catalog number: 28304)
  7. Primer A (5'-GAGAGAGGATCCGAATTGTGCAGCTGTCAATCAACC-3') (5 µM)
  8. T4 RNA ligase 1 (New England BioLabs, catalog number: M0204)
  9. T4 RNA ligase reaction buffer (10x, supplied with T4 RNA ligase 1)
  10. ATP (10 mM) (supplied with T4 RNA ligase 1)
  11. PEG8000 (supplied with T4 RNA ligase 1)
  12. Primer B (5'-GAGAGAGGATCCTTTTTTTTTTTTTTTTTTTVN-3') (5 µM)
  13. Pfu DNA polymerase (or equivalent)
  14. 10x Pfu DNA polymerase buffer (or equivalent, supplied with MgCl<sub>2</sub> or MgSO<sub>4</sub>)
  15. dNTPs (5 µM)
  16. 2% agarose gel
  17. Generuler low range DNA ladder (Thermo Fisher Scientific)
  18. GeneJET gel extraction kit (Thermo Fisher Scientific, Fermentas, catalog number: K0691)

*Note: Additional materials and reagents required only if ssDNA is not 5'-end phosphorylated: T4 polynucleotide kinase (PNK, New England Biolabs, catalog number: M0201) and T4 polynucleotide kinase reaction buffer (New England Biolabs, catalog number: M0201).*

## **Equipment**

- A. Purifying nucleic acids from proteins
  1. Heat block
  2. PCR machine
  3. Cooled table top Eppendorf centrifuge
  4. Freezer (-20 °C)

B. Enriching and preparing ssDNA molecules for sequencing

1. Heat block
2. PCR machine
3. Table top centrifuge
4. Agarose gel electrophoresis equipment

**Procedure**

A. Purifying nucleic acids from proteins

1. Take purified protein sample (e.g. 50-500 µl, protein concentration up to 10 µM).
2. Add CaCl<sub>2</sub> to a final concentration of 5 µM.
3. Add 1/10 volume of Proteinase K solution (e.g. to 450 µl protein/CaCl<sub>2</sub> solution, add 50 µl Proteinase K solution).
4. Incubate 4 h at 65 °C in a heat block. Alternatively, small volumes (up to 200 µl) can be incubated 4 h at 65 °C in a PCR machine.
5. Add Roti phenol/chloroform/isoamyl alcohol (pH 7.5-8.0) in a 1:1 ratio (e.g. to 500 µl, add 500 µl phenol/chloroform/isoamyl alcohol).
6. Vortex at max speed for 5 sec.
7. Centrifuge the sample for 15 min in a table top centrifuge at maximum speed.
8. Carefully pipet the upper fraction (aqueous phase containing the nucleic acids) to a clean centrifuge tube. Take less than the actual upper fraction (e.g. from a 500 µl sample, take only 450 µl). This avoids mixing with the lower fraction (organic phase containing proteins) which contains the phenol which might interfere with downstream steps. Discard the lower fraction.
9. If the total volume of the upper fraction is larger than 500 µl, distribute the sample equally over multiple tubes.
10. Add 99% ethanol in a 2:1 ratio (ethanol: sample).
11. Add 10 µl linear polyacrylamide (5 µg/µl) as carrier.
12. Incubate the samples at -20 °C for at least 16 h. Small nucleic acids take longer to precipitate.
13. Centrifuge in a table top centrifuge at 4 °C for 30 min at maximum speed (13,000 rpm).
14. Remove all ethanol with a pipette, a small white/transparent pellet should be visible- try not to disturb the pellet.
15. Add 500 µl pre-cooled -20 °C, 70% ethanol to the pellet.
16. Centrifuge in a table top centrifuge at 4 °C for 15 min at maximum speed.
17. Remove all ethanol with a pipette-a small white/transparent pellet should be visible- try not to disturb the pellet.
18. Incubate the reaction tube at 50 °C until all ethanol is evaporated (5-10 min).
19. Resuspend the nucleic acids in 50 µl RNase-free MQ water.

20. Analyze nucleic acids by  $^{32}\text{P}$ -labeling (see Figure 1) or use for enriching and preparing ssDNA molecules for sequencing (procedure B below). For  $^{32}\text{P}$ -labeling, use T4 polynucleotide kinase (T4 PNK) according to the protocol provided by the manufacturer.

B. Enriching and preparing ssDNA molecules for sequencing

1. Add 1  $\mu\text{l}$  of DNase free RNase A to 50  $\mu\text{l}$  purified ssDNA.

2. Incubate at RT for 10 min.

3. To polyadenylate the ssDNAs, use Terminal Deoxynucleotidyl transferase according to the protocol provided by the manufacturer.

10  $\mu\text{l}$  5x TdT reaction buffer

36.5  $\mu\text{l}$  ssDNA (30-100 ng, if necessary, add MQ water)

1.5  $\mu\text{l}$  dATP (100 mM)

1.5  $\mu\text{l}$  TdT

4. Mix and incubate for 60 min at 37 °C.

5. Incubate for 10 min at 70 °C to inactivate TdT.

6. Purify the samples using the QIAquick nucleotide removal kit according to the manual provided by the manufacturer. Elute in 30  $\mu\text{l}$  MQ water.

Notes:

a. *Other nucleic acid purification kits might also be suitable for purifying the nucleic acids. However, most nucleic acid purification kits do not purify nucleic acids <100 nt. Ensure the kit you use is suitable for the size of nucleic acid you are purifying.*

b. *The next step is dependent on the presence of a phosphorylated 5'-end on your ssDNA molecules. If you are not sure if your small DNA molecule has a phosphorylated 5'-end, it is possible to phosphorylate it with T4 polynucleotide kinase according to the manual provided by the manufacturer. After 5'-end phosphorylation, purify the samples using the QIAquick nucleotide removal kit according to the manual provided by the manufacturer. Elute in 30  $\mu\text{l}$  MQ water.*

7. Ligate primer A (see Notes a and e) to the phosphorylated 5'-ends of the ssDNA molecules with 3'-end poly A tails by mixing 16  $\mu\text{l}$  ssDNA, 1  $\mu\text{l}$  primer A, 6  $\mu\text{l}$  ATP (10 mM), 6  $\mu\text{l}$  10x Reaction buffer, 30  $\mu\text{l}$  50% PEG8000 (w/v) and 1  $\mu\text{l}$  T4 RNA ligase. Incubate for 16 h at 25 °C. T4 RNA ligase is used as it can ligate a 3'-end of a ssDNA strand (the primer) to a 5'-end phosphorylated ssDNA strand (the ssDNA molecules with 3'-end poly A tails).

8. Purify the samples using the QIAquick nucleotide removal kit according to the manual provided by the manufacturer. Elute in 30  $\mu\text{l}$  MQ water.

9. To enrich the ssDNA molecules, PCR amplify the ssDNA molecule using primers A and B.

2  $\mu\text{l}$  10x Pfu polymerase buffer (containing MgCl<sub>2</sub> or MgSO<sub>4</sub>)

14 µl ssDNA with 5' oligonucleotide and 3' poly A tail

1 µl primer A (5 µM)

1 µl primer B (5 µM)

1 µl dNTPs (5 µM)

1 µl Pfu polymerase

10. Run a PCR

1 min at 95 °C

-----

20 cycles:

1 min at 95 °C (denaturation)

30 sec at 40 °C (annealing)

1 min at 72 °C (elongation)

-----

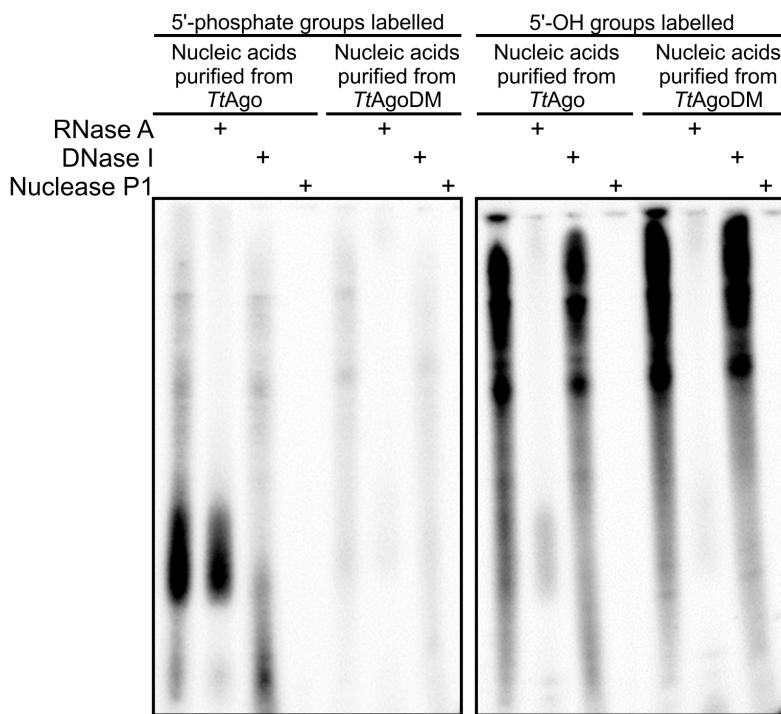
1 min at 72 °C

Notes:

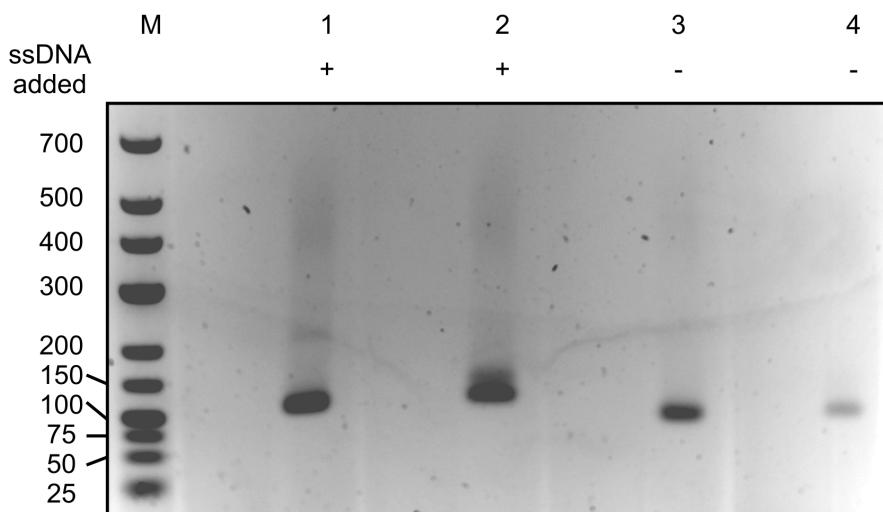
- a. *The annealing temperature is dependent on your primers. The elongation time is dependent on the DNA that you want to amplify. The amount of cycles required is dependent on the amount of template and the amount of product required. Generally, less cycles is better for dsDNA quality, while more cycles generate more dsDNA fragments.*
- b. *Run the PCR product on a 2% agarose gel.*
- c. *Sometimes two (or more) bands appear after the PCR. The lowest (smallest) band often corresponds with a PCR product containing only the primers but no DNA of interest. Based on the size of your co-purified ssDNA, you should be able to predict the expected size of a successful generated and amplified dsDNA product (see Figure 2). Cut out the band of the correct size and purify the dsDNA fragments from this band using the GeneJET gel extraction kit according to the protocol provided by the manufacturer.*
- d. *If the yield of the dsDNA molecules is too low for cloning/sequencing, the PCR can be repeated on the dsDNA purified from the gel. Ensure to repeat the gel purification, as also products of other lengths might be amplified.*
- e. *The purified dsDNA PCR products can be sequenced directly by a modified Illumina sequencing protocol. In order to do this, ensure that primer A contains a sequence suitable for the sequencing platform used (discuss this with your sequencing company). Alternatively, the purified dsDNA PCR products can be cloned in a vector using the introduced KpnI restriction sites on either end of the PCR product (or other restriction sites, if you designed your own primers). Alternatively, the products can be cloned by TOPO cloning (Life Technologies) after addition of 3'-end adenine groups with Taq polymerase (see protocol provided by the manufacturer). After transformation of these vectors, plasmids*

can be purified and sequenced using standard protocols. After sequencing, the strand originating from the single stranded molecule can be identified by the presence of the 5'-end Primer A sequence and 3'-end poly-A sequence.

### **Representative data**



**Figure 1. 20% denaturing polyacrylamide gels with resolved nucleic acids that co-purified with *TtAgo* or *TtAgoDM* (a catalytic mutant of *TtAgo*, unable to acquire guides *in vivo*) (Swarts et al., 2014a).** Nucleic acids were  $^{32}\text{P}$  labelled by replacing 5'-phosphate (left gel) or 5'-OH groups (right gel) with  $^{32}\text{P}$  using T4 polynucleotide kinase. Labelled nucleic acids were cleaved with RNase A, DNase I or nuclease P1 (cleaves ssRNA and ssDNA). The left gel shows that *TtAgo* co-purified with small ssDNA molecules with a 5'-end phosphate. The right gel shows that both *TtAgo* and *TtAgoDM* co-purify with long ssRNA molecules of undefined length. Figure adapted from Swarts et al. (2014a).



**Figure 2. 2% agarose gel with resolved enriched dsDNA molecules.** M: Generuler low range DNA ladder. Sizes corresponding with the DNA in the marker lane are indicated in bp. Lanes 1 and 2 contain generated dsDNA samples to which 21 nt ssDNA was added in step B1 of the protocol. Lanes 3 and 4 contain negative controls with generated dsDNA molecules consisting of only the oligonucleotide and 3'-end poly A tail without an ssDNA insert. Note that sometimes, bands representing both molecules with and without an insert will appear in the sample of interest. In this case, make sure you select the right band (usually the upper band, as it contains an insert and thus is larger) for further processing and sequencing.

### Acknowledgments

This work was financially supported by a grant from the Netherlands Organization of Scientific Research (NWO) to J.O. (NWO-TOP, 854.10.003).

### References

1. Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., Snijders, A. P., Wang, Y., Patel, D. J., Berenguer, J., Brouns, S. J. and van der Oost, J. (2014a). [DNA-guided DNA interference by a prokaryotic Argonaute](#). *Nature* 507(7491): 258-261.
2. Swarts, D. C., Jore, M. M. and van der Oost, J. (2014b). [Expression and purification of the \*Thermus thermophilus\* argonaute protein](#). *Bio-protocol* 4(19): e1253.