Analysis and purification of ssRNA and dsRNA molecules using asymmetrical flow field flow fractionation

Eskelin , Katri

2022-11

Eskelin , K , Lampi , M , Coustau , C , Imani , J , Kogel , K-H & Poranen , M 2022 , ' Analysis and purification of ssRNA and dsRNA molecules using asymmetrical flow field flow fractionation ' , Journal of Chromatography. A , vol. 1683 . https://doi.org/10.1016/j.chroma.2022.463525

http://hdl.handle.net/10138/350653
https://doi.org/10.1016/j.chroma.2022.463525

Journal of Chromatography A 1683 (2022) 463525

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Analysis and purification of ssRNA and dsRNA molecules using asymmetrical flow field flow fractionation

Eskelin Katri[a,*], Lampi Mirka[a], Coustau Christine[b], Imani Jafargholi[c], Kogel Karl-Heinz[c], Poranen Minna M[a]

[a] Molecular and Integrative Biosciences Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland
[b] French National Centre for Scientific Research, Institut Sophia Agrobiotech, France
[c] Centre for BioSystems, Land Use and Nutrition, Institute of Phytopathology, Justus Liebig University, Germany

ARTICLE INFO

Article history:
Received 29 June 2022
Revised 30 August 2022
Accepted 18 September 2022
Available online 22 September 2022

Keywords:
Field-flow fractionation
RNA
Biophysical characterization
Macromolecule purification

ABSTRACT

Robust RNA purification and analysis methods are required to support the development of RNA vaccines and therapeutics as well as RNA interference -based crop protection solutions. Asymmetrical flow field-flow fractionation (AF4) is a gentle native purification method that applies liquid flows to separate sample components based on their hydrodynamic sizes. We recently showed that AF4 can be utilized to separate RNA molecules that are shorter than 110 nucleotides (nt), but the performance of AF4 in the analysis and purification of longer RNA molecules has not been previously evaluated. Here, we studied the performance of AF4 in separation of single-stranded (ss) and double-stranded (ds) RNA molecules in the size range of 75–6400 nt. In addition, we evaluated the power of AF4 coupling to different detectors, allowing separation to be combined with data collection on yield as well as molecular weight (*MW*) and size distribution. We show that AF4 method is applicable in RNA purification, quality control, and analytics, and results in good recoveries of ssRNA and dsRNA molecules. In addition, our results demonstrate the utility of AF4 multidetection platforms to study biophysical properties of long RNA molecules.

1. Introduction

RNA research is blooming with frequent reports on novel RNA functions. These discoveries directly translate to novel implications of RNA in applied research and beyond. Short interfering dsRNAs (siRNA) of 21–24 nucleotides (nt) are routinely applied as tools to trigger RNA interference (RNAi) pathway to downregulate specific gene functions in mammals [1–5]. This technology has resulted in the development and introduction of siRNA-based therapies for the treatment of rare genetic diseases [6]. RNAi has also been recognized as a sustainable biodegradable alternative for the use of chemical pesticides in crop protection [7–9]. In this case, RNAi pathway is activated using long dsRNA molecules sharing sequence homology with a messenger (m)RNA representing an essential gene of the target pathogen or pest. In addition to dsRNA, high quality ssRNA (mRNA) is needed to produce therapeutic proteins in *in vitro* translation systems or to be used as vaccines to program the protein production *in vivo*, as exemplified by the recently developed mRNA vaccines against Covid-19. Advantages of RNA-based drugs are their ability to act on targets that are not accessible for small molecules or proteins, their rapid and cost-effective development in comparison to small molecules or recombinant proteins, and the ability to rapidly alter the sequence of the mRNA construct for personalized treatments or to adapt to evolving pathogens [10]. RNA can also be utilized to construct nanostructures and nanoparticles with versatile functions and applications accompanied with tunable physicochemical properties including targeted drug delivery and release [11,12]. However, structural studies on RNA molecules, especially on molecules longer than 100 nt, lag behind partially due to difficulties to purify high quality RNA in native form.

RNA molecules can be produced enzymatically *in vitro* or *in vivo* using viral DNA- and RNA-dependent RNA polymerases [1,13]. These reactions may produce unwanted by-products such as RNA multimers, pre-terminated reaction products, excess of ssRNA in the case of dsRNA synthesis reaction. The by-products together with the other reaction components, such as nucleotides, DNA template and the polymerases necessitates a need for economically feasible downstream protocols to purify the RNA molecules of interest. Obviously, both the efficiency of the RNA synthesis and the number of required purification steps affect the yield and qual-

* Corresponding author.
E-mail address: katri.eskelin@helsinki.fi (E. Katri).

https://doi.org/10.1016/j.chroma.2022.463525

ity of the final product as well as the overall economic feasibility of the RNA production. Many purification methods rely on denaturation to dismantle RNA from ribonucleoprotein complexes and to melt secondary and tertiary structures of RNA molecules. The three-dimensional architecture is often important for the maximal biological functionality of the RNA molecules but refolding of molecules back to native conformation may be difficult and incomplete.

In asymmetrical flow field-flow fractionation (AF4), variations in the hydrodynamic sizes (radius of hydration, R_H) of the sample components induce differences in their separation. The principles and theory of AF4 as well as its various applications have been thoroughly described in the original papers and in many reviews [14–19]. The method applies two perpendicular flows to generate separation in a mobile phase-filled, trapezoid-shaped narrow channel without stationary phase. The cross-flow force pushes the sample towards the accumulation wall of the channel. Sample components diffuse against the applied cross-flow force due to Brownian motion.The channel flow with parabolic flow profile carries the sample components through the channel to the outlet port and ultimately to the detectors and the fraction collector. The accumulation wall consists of a porous frit that is lined with an ultrafiltration membrane that determines the minimal size of sample components that retain in the channel for separation. Small sample components migrate further from the accumulation wall, reach higher channel-flow velocities of the parabolic flow profile, and elute before the large ones. This type of normal elution mode applies to sample components that are smaller than ~0.5–1 µm in radius, and wide range of sizes can potentially be resolved in a single AF4 experiment [20]. Separation occurs in an open channel and is thus gentle as there are no strong interactions or shear forces. Accordingly, optimized conditions should provide yields that are above 70% [21]. Importantly, coupling of the AF4 separation to multiangle light scattering (MALS), dynamic light scattering (DLS) and ultraviolet (UV) or differential refractive index (dRI) concentration detectors, allows post-separation analysis of the collected data on yield, molecular weight (*MW*), and size distribution [root mean squared radius (*rms*); or radius of gyration (R_G) and (R_H)] of the separated sample components.

The potential of field flow fractionation in nucleic acid research has been recognized before. Initially, field flow fractionation was shown to separate a variety of linear and circular ssDNA and dsDNA molecules in the size range of 600–7400 bp [22–25], and was used to investigate the elution behavior of short folded and unfolded 10- to 100 nt-long ssDNA molecules [26,27]. The majority of currently published AF4 works on RNA describe experiments on RNA complexed with proteins and/or lipids [28–32]. Recently, we showed that AF4 can be utilized to analyze and purify ss- and dsRNA molecules that are shorter than 110 nt [33]. We are not aware of other published studies on free naked RNA molecules.

In this work, we further explored the applicability of AF4 for analysis and purification of RNA molecules with the focus on long ss- and dsRNA molecules. RNAs were produced *in vitro* by viral polymerases. We studied AF4 performance in the separation of differently processed ss- and dsRNA molecules in the size range of 75–6400 nt (Fig. S1). Our results show that AF4 is applicable for analysis and purification of long RNA molecules. Enzymatically produced ss- and dsRNA molecules are separated from each other and from the contaminating reaction components allowing fast purification of the dsRNA molecules of interest without a need for pre-purification. The ssRNA multimers were separated from the corresponding monomers. AF4 also allowed fractionation of segmented dsRNA genome of bacteriophage ϕ6 from total cellular RNA. Finally, we evaluated the efficacy of AF4-purified 470 nt-long dsRNA molecules sharing sequence identity with aphid salivary sheath protein (*SHP*) gene to control aphid growth and reproduction. We also discuss the limits and benefits of AF4 in analysis and purification of RNA.

2. Materials and methods

2.1. Production and pre-purification of RNA

Production and purification of RNA molecules has been described earlier [33,34] (Fig. S1). DNA templates of approximately 90–1800 bp were amplified using PCR from plasmid pLM659 encoding the S segment of bacteriophage ϕ6 genome [35]. The forward primers contained the promoter sequence for the T7 DNA-dependent RNA-polymerase followed by a 17-, 20- or 21 nt-long sequence complementary to the ϕ6 S-segment starting at the position of 80, 100, 300, 500, 700 or 1800 nt from the 3' end of the segment. Reverse primer was complementary to the 3' end of the S-segment and contained 8 nt-long ϕ6 promoter sequence at the 5' end. This 8 nt-long sequence and the last nucleotide of the T7 promoter sequence are included in the sequences of the produced RNA molecules. Full-length ϕ6 S- (2948 bp), M- (4065 bp) and L- (6374 bp) dsRNA segments and corresponding s^+, m^+ and l^+ ssRNAs were amplified using plasmids pLM659, pLM656 and pLM687 as templates, respectively [35–37] (Table S1). Plasmids pGEM-T-easy-SHP and pGEM-T-easy-eGFP [9] containing sequences of aphid (*Sitobion avenae*) salivary sheath protein (*SHP*) gene and enhanced green fluorescent protein (*eGFP*) gene, respectively, were used as templates in the production of 470 bp dsRNA molecules for RNAi experiments (see 2.5). The used PCR primers have been described earlier [9,38]. Due to the inclusion of the T7 promoter region in the primers, the DNA templates are 21 nt longer than the RNA molecules produced from the corresponding DNA templates. The names and properties of the produced RNA molecules are provided in Table S1.

Gel-purified PCR products were used as templates to transcribe ssRNA by the T7 DNA-dependent RNA-polymerase (Fig. S1). DsRNA molecules were produced in coupled reactions applying T7 DNA-dependent RNA-polymerase and ϕ6 RNA-depended RNA-polymerase [1,39]. Reaction mixtures contained 600 mM HEPES-KOH (pH 7.5), 120 mM $MgCl_2$, 100 mM DTT and 5 mM spermidine. In total, ten ssRNA and twenty dsRNA synthesis reactions, each á 50 µl were prepared and pooled.

A fraction (3/5) of the pooled reactions was pre-purified before AF4 fractionation as previously described [34] (Fig. S1). We used phenol-chloroform extraction (TRIsure, Bioline, Luckenwalde, Germany) and precipitation with 4 M LiCl (LiCl; Merck, Darmstadt, Germany) to enrich ssRNA molecules from the ssRNA synthesis reaction. For pre-purification of dsRNA, we used two consecutive LiCl precipitations at 2 and 4 M to enrich dsRNA. Precipitations were followed by washes with 70% (v/v) ethanol. For the ssRNA molecules, additional sodium acetate and ethanol precipitation was performed. Pre-purified RNA molecules were resuspended in MQ water. Reactions examined by the AF4 method were treated with DNase I (2 units per 1 µg of DNA template; RQ1 RNase-Free DNase, Promega Corporation) at 37 °C for 30 min.

Total RNA was isolated from ϕ6 -infected and noninfected *Pseudomonas syringae* HBY10 cell pellets using TRIsure (Bioline, Luckenwalde, Germany) according to manufacturer's instructions. Culturing and infection of logarithmically growing cells was done as described previously [40]. Cells were collected by centrifugation 45 min post infection before cell lysis occurred. Non-infected cells were grown and collected in the same manner.

2.2. In silico RNA analyses

Sequence information was used to determine the GC percentage and the theoretical *MW* of the studied RNA molecules (Table S1).

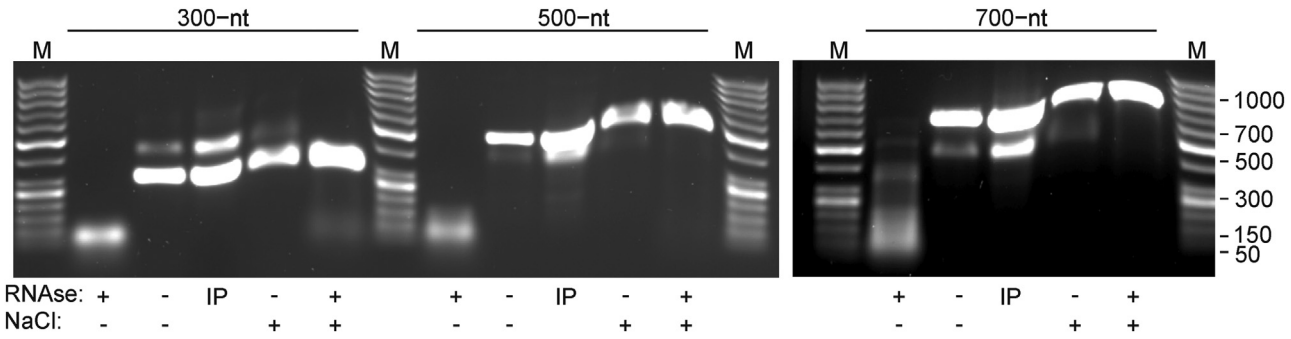

Fig. 1. RNase A sensitivity test to identify ssRNA and dsRNA molecules. Sensitivity of the 300-, 500- and 700 nt dsRNA samples to RNase A in low (-) and high (+) NaCl concentration and impact of NaCl on the electrophoretic RNA migration. Reactions were analyzed in native 1% (w/v) agarose gel. The treatments and reaction conditions are indicated below the gels. IP refers to untreated sample which contains no NaCl. M is a dsDNA size marker. The mobility of selected DNA molecules is indicated on the right.

The potential folding of ssRNA molecules at 22 °C was analyzed using M-fold web server version 2.3 (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form2.3, accessed 2.9.2020) [41]. Maximum ladder distance (MLD), the length of the longest direct path across the RNA secondary structure, was determined from the predicted structures by calculating the number of base pairs in double-stranded regions of maximal distance. R_G was estimated from the number of base paired nucleotides (N) within MLD with the following formula: $R_G = N_{MLD}^{0.33}$ nm [42]. The length of dsRNA molecules was calculated by assuming a mean rise of 0.279 nm per base pair [43]. GenBank accession number for bacterial rRNAs of *Pseudomonas syringae* are AE016853.1 and for ϕ6 L, M and S segments M17461, M17462, M12921, respectively.

2.3. AF4 setup and operation

AF4 experiments were done using an AF2000 MT instrument (Postnova Analytics, Landsberg, Germany), or Eclipse NEON (Wyatt Technology, Dernbach, Germany) at 22 °C. Prior to RNA experiments, the AF4 instruments were washed with 0.2 M NaOH supplemented with 0.2% (w/v) SDS for 30 min to remove RNAses, followed with MQ rinse until neutral pH was obtained. We initiated the study using AF2000 MT instrument and AF2000 control software (Postnova Analytics, Landsberg, Germany). The instrument was equipped with an UV detector (Shimadzu SPD-20A; Shimadzu, Kyoto, Japan) that monitors UV signal in volts (V) at 260 nm. Spacer adjusted the channel to a nominal width of 250 µm. The regenerated cellulose (RC) membrane had molecular weight cutoff of 10 g/mol (Postnova Analytics, Landsberg, Germany). The injection loop volume of manual injector was 100 µl: the volume of the injected RNA sample was adjusted to 110 µl with the used mobile phase [50 mM Tris-HCl (pH 8), 20 mM NaCl]. The injected amount of pre-purified samples corresponded to approximately 20–50 µg of RNA. For reaction mixtures, the injected sample volumes were estimated in pre-runs with varying injection volumes: up to 500 and 1000 µl of ssRNA and dsRNA synthesis reactions, respectively, were used. The AF4 experiments comprised four steps: the 5 min focusing step that included the injection step was followed by initial elution at constant cross-flow velocity of 2 ml/min for 5 min, after which cross flow decayed linearly to 0.01 ml/min in 15 min and remained constant for 15 min. However, for the 300 nt RNA molecules, the cross-flow velocity was set to 3 ml/min for 25 min. Channel flow velocity was 0.5 ml/min in all experiments carried out with the Postnova instrument (data shown in Figs. 2, 4,5, 7 and 8, and Figs. S9–10). Fractions of 1 ml were collected from the elution step.

The Eclipse NEON (Wyatt Technology, Dernbach, Germany) AF4 instrument used in this study (data shown in Figs. 3, 6, and S3–S7) was connected to DAWN NEON multi-angle light scattering (MALS, Wyatt Technology) detector that contains integrated WyattQELS for DLS, Agilent 1260 Infinity Multiple Wavelength Detector G7114A for UV, and Optilab dRI detector (Wyatt Technology). The setup also had temperature-controlled Agilent 1260 Vialsampler G7129A and Agilent 1260 G1364F fraction collector. Separation was done in a long channel that contains a dilution control module (Wyatt Technology) using a 400 µm spacer, 10 kD RC membrane (Wyatt Technology). Channel-flow velocity was 1.0 ml/min and detector-flow velocity 0.5 ml/min. The cross-flow velocity gradient initiated from 2 or 3 ml/min. Alternatively, we used a semipreparative channel with 650 µm spacer and 10 kD RC membrane (Wyatt Technology) to facilitate larger sample amounts. The separation method for semipreparative channel was adjusted: the initial cross-flow velocity was 1 ml/min and channel flow 1.5 ml/min (see Fig. 6). Detector flow was 0.5 ml/min. Total elution time (excluding the focusing) was 70 min, where the linear cross-flow gradient was 50 min (see Fig. 6). The z-average *rms* radius (R_G), R_H, average rod length and weight-average molar mass (MW) and the respective uncertainty in the obtained values was obtained using ASTRA software (version 8.1.0) using refractive index increment (dn/dc) value of 0.170 mL/g for RNA and Zimm formalism. Baseline subtraction for RI and UV signals was used. Averages and corresponding standard deviations for the major peaks were calculated from at least three independent measurements, where the % standard deviation was less than 5%. MALS-RI set-up was validated for separation, recovery and MW with BSA and using dn/dc of 0.185 ml/g and constant cross-flow velocity of 3 ml/min and channel and detector flow velocity of 1 ml/min and 0.5 ml/min, respectively).

2.4. Analysis of recovery, quality and quantity

Nucleic acid concentrations were measured using NanoDrop2000C Spectrophotometer (Thermo Scientific). Fraction contents were precipitated with sodium acetate and ethanol and solved to a small volume of MQ water prior to the concentration measurement. For qualitative analysis, RNA molecules were separated by native gel electrophoresis in 1% (w/v) agarose gel containing ethidium bromide. Electrophoresis was performed in 1 × TAE buffer (40 mM Tris-acetate (pH 7.4), 1 mM EDTA). The maximum amount of RNA loaded on the gel was 300 ng. If the RNA amount in the sample was below 300 ng, the entire content of the precipitated fraction was loaded onto the gel. After electrophoresis, gels were visualized with ChemiDoc (Bio Rad, Hercules, USA). DsDNA markers (#SM0371 and #SM0331, #SM0241, Thermo Scientific) were applied as size indicators. RNA recovery was calculated from the peak areas of the UV signal on AF4 fractograms or converted to µg amounts using ASTRA software.

Ribonuclease A (RNAse A) sensitivity of the samples was analyzed at low (20 mM) and high (500 mM) NaCl concentration to specifically digest ssRNA, respectively. RNA samples (~1 µg) were incubated in the presence or absence of RNAse A (1 µg) at 37 °C for 45 min in total volume of 30 µl. The effect of heat denaturation on ssRNA multimers was studied by heating samples at 70 °C for 10 min, followed by incubation on ice. Products were analyzed in 1% (w/v) native agarose gels.

2.5. Functional RNAi tests for AF4 purified dsRNA molecules

Two dsRNA molecules eGFP and SHP, both 470 nt, were produced *in vitro* (see 2.1 and Fig. S1). Reaction mixtures were AF4 fractionated in nine separate AF4 experiments to purify dsRNA molecules for functional tests in aphids. The buffer of the dsRNA-containing AF4 fractions was exchanged to MQ water using NAP-25 columns (GE Healthcare) and the desalted samples were concentrated with evaporation using Speed vac (Savant).

Because the efficiency of RNAi on insects may vary substantially depending on the species and the protocol used [9], two functional RNAi tests were performed independently in two laboratories and following different procedures. In the first set of experiments, dsRNA samples were diluted to 2.5 ng/µl in sterile artificial diet according to previously described procedures [44]. *Myzus persicae* aphids, approximately 60 aphids per treatment, were fed for 24 h with dsRNA-containing artificial diet and then transferred to cabbage plants. Aphids were placed in cages, approximately 20 aphids per clip cages, 3 clip cages per treatment. The survival and number of offspring were monitored for 12 days.

In a second set of experiment, dsRNA samples were mixed with artificial sucrose diet (50 mM L-serine, 50 mM L-methionine, 50 mM L-aspartic acid, sucrose (7.5% w/v), pH 7.2) at a concentration of 250 ng/µl [9]. One-day-old *Sitobion avenae* aphid nymphs were placed on artificial medium and maintained at 22 °C for 4 days. Ten synchronized nymphs with five replicates for each diet were used.

3. Results and discussion

In this study, we evaluated the applicability of AF4 to separate ss- and dsRNA molecules of 75–6400 nt size range. We initiated the study with 1800-, 700-, 500- and 300 nt RNA molecules that had identical 3' ends over the last 300 nt. These RNA preparations varied in purity as they were either unpurified reaction mixtures from ssRNA and coupled ss- and dsRNA synthesis reactions or reactions that were pre-purified by LiCl precipitation (Fig. S1). Later we amended our studies with slightly shorter (90- and 110 nt) and longer (4100- and 6400 nt) ssRNA molecules derived from the ϕ6 M and L segments, respectively (Table S1), as well as total RNA from virus-infected bacteria. The GC content of the studied RNA molecules varied from 48% to 56%, and the calculated *MW* from 25,000 to 4,100,000 g/mol (Table S1).

Agarose gel electrophoresis analysis of the studied 1800-, 700-, 500- and 300 nt pre-purified ss- and dsRNA samples showed that the coupled ss- and dsRNA synthesis reaction produced a mixture of RNA molecules suggesting that not all ssRNA molecules were converted to dsRNA by viral polymerase (Figs. S1 and 1, lanes for input (IP)). In addition, some ssRNA samples appeared to contain several RNA types with different electrophoretic mobilities (Fig. S1). The presence of ssRNA in the dsRNA samples was verified using RNase A treatment at 500 mM NaCl concentration to specifically degrade ssRNA molecules. Both the ss- and dsRNA molecules were degraded at low NaCl concentration, as expected (Fig. 1).

3.1. Separation of ss- and dsRNA molecules

In AF4, separation and thus the retention time correlates with the R_H of the sample components [14]. For ssRNAs, R_H is affected by the number of nts in the polymer, the folding of ssRNA molecules *via* secondary and tertiary interactions, and the length of double-stranded regions. The folding is also affected by the composition of the solvent: cations shield negative charges on the RNA backbone and reduce repulsive forces that promotes the formation of intramolecular interactions and more compact fold. The ssRNA molecules are typically prolate and aspherical [45]. Short dsRNA molecules behave as rigid rods, whereas long dsRNAs can bend [43]. As ssRNA can fold to more compact structures than dsRNA, we expected that ssRNA made from the same template as the dsRNA elute first during AF4 separation. However, possible differences in the three-dimensional structures of ssRNA molecules impacting R_H were expected to induce irregularity in their retention behavior and peak shape. Consequently, peak intensity should provide information on the RNA quantity and efficacy of the *in vitro* synthesis reaction or the pre-purification process, and the peak shape about the homogeneity of the RNA molecules.

We optimized AF4 flow conditions using aliquots of the pre-purified 300–1800 nt ss- and dsRNA molecules and applied those in subsequent experiments to purify RNA directly from the reaction mixtures (Figs. 2–5). The analysed RNA amounts did not induce significant overloading, as no excessive peak broadening was observed, and thus even higher amounts of RNA could be injected to the analytical AF4 channel to promote purification of more RNA from a single experiment. Pre-purified 1800, 700 and 500 nt ssRNAs and the corresponding reaction mixtures eluted as one major peak (peak 1) that was followed by a low-intensity peak (peak 2, Fig. 2A). Despite the 200 nt size difference, the 700 and 500 nt ssRNAs had similar retention times (Fig. 2A). The pre-purified 300 nt ssRNA yielded two almost equally intense peaks with no baseline separation (Fig. 2C).

All the dsRNA reaction mixtures, as well as the pre-purified 700-, 500- and 300 nt dsRNA samples produced two major peaks (Fig. 2B and D). Retention time of the first peak suggested ssRNA nature and thus incomplete conversion of ssRNA to dsRNA and inefficient pre-purification. The pre-purified 1800 nt dsRNA eluted as one major peak, but the reaction mixture had an additional broad pre-peak, indicating potential contamination with smaller RNA species (Fig. 2B).

We exposed the fractions representing the two main peaks of the AF4 separated 700 nt dsRNA sample (see Fig. 2B) to RNAse A treatment at low and high NaCl concentration followed by agarose gel electrophoresis to verify the nature of the peak contents (Fig. S2). Importantly, only the lower intensity bands that represented the faster-eluting smaller-sized molecules in AF4 were degraded in high salinity conditions, indicating that the fraction contained ssRNA molecules. The RNA species eluting at the second peak remained intact in high salinity conditions confirming ds nature.

In general, the used linear cross-flow gradient enabled separation of the 1800-, 700-, 500- and 300 nt dsRNAs from each other as well as the 1800-, 700-, and 500 nt ssRNA precursors from the corresponding dsRNA molecules. However, under these conditions the pre-purified 300 nt dsRNA sample yielded two partially overlapping peaks that eluted within 8 min (Fig. 2B). Their separation was slightly improved when we repeated the analysis using higher cross-flow velocity (Fig. 2D, see also Fig. S3). However, the peculiar elution behavior of the 300 nt RNA sample needs further studies.

The AF4 fractograms indicated that all eight RNA samples were heterogeneous mixtures of more than one RNA species and based on the AF4 fractograms the dsRNA reactions contained excessive amounts of ssRNA (Fig. 2). Templates for the production of the 300-, 500-, 700- and 1800 nt dsRNA molecules had the same 5′-

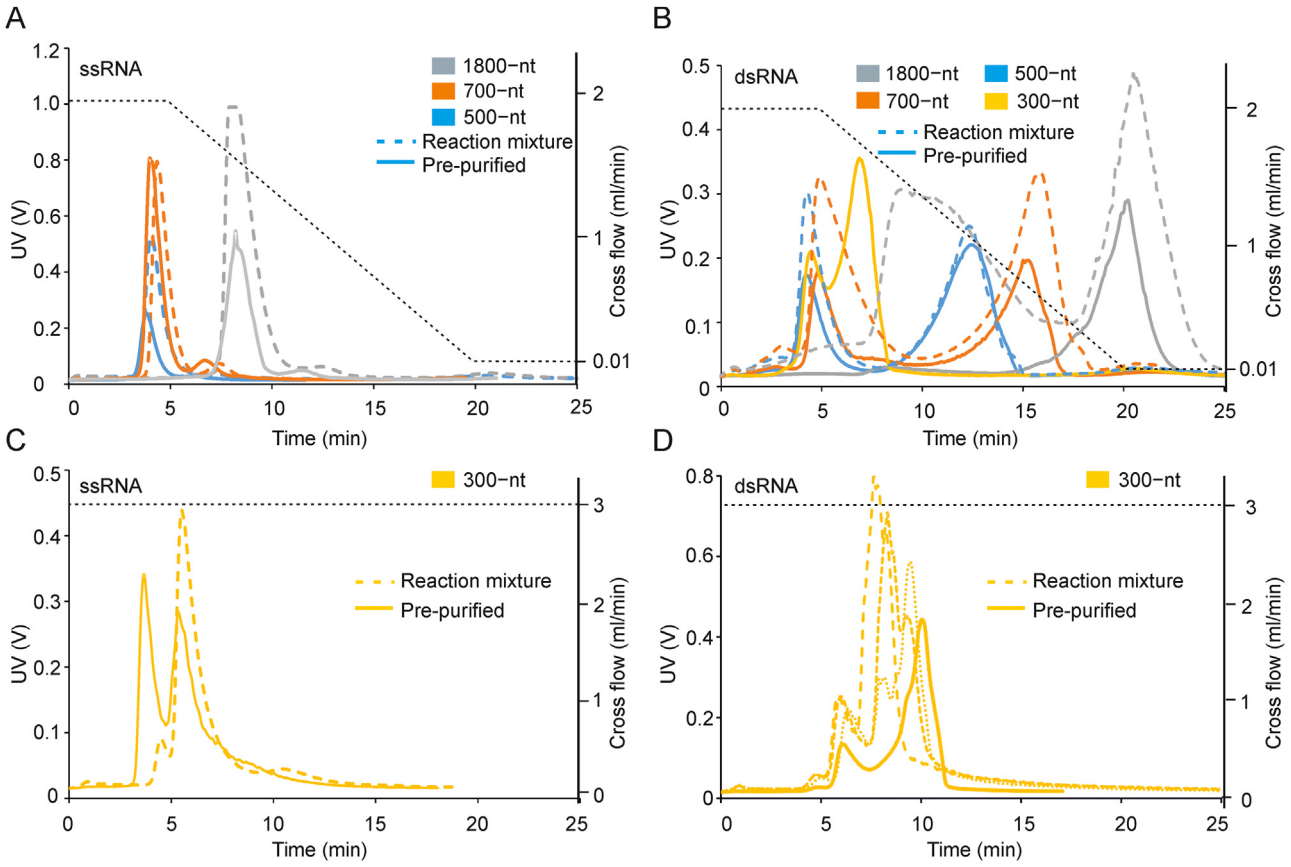

Fig. 2. AF4 fractionation of enzymatically produced ss- and dsRNA molecules. Fractograms show representative elution profiles for the pre-purified (solid lines) and non-purified (dashed lines) 300-, 500-, 700- and 1800 nt ssRNA (A, C) and dsRNA (B, D) samples. Three subsequent experiments are presented for the reaction mixture of 300 nt dsRNA with dashed lines to show the characteristic experiment-to-experiment variation for this molecule (D). Cross-flow profiles are shown with black dashed lines (right y-axis). Time axis shows the UV fractogram from the beginning of elution program excluding the focusing time. UV detector response at 260 nm is given in volts (V) (left y-axis).

UUUCC-3′ sequence at the 3' termini that should promote equal high-affinity binding of the ϕ6 RNA-dependent RNA polymerase [46]. Thus, further optimization of the reaction conditions and polymerase amounts could enhance the ratio of produced RNAs in favor of dsRNA. The pre-purification of dsRNA molecules with LiCl precipitation appeared to be inefficient for most of the studied RNAs (Fig. 2). Previous studies have suggested that the efficacy of LiCl-based RNA fractionation correlates with the length of RNA molecules [47], but we did not observe such correlation.

3.2. Molar mass distribution of studied ssRNA and dsRNA molecules

Electrophoresis in native agarose gel indicated that the pre-purified RNA samples contained at least two RNA species (Figs. 1, and S1). During AF4 separation, tailing of the peaks was observed for the ssRNA molecules, whereas dsRNA molecules showed fronting (Fig. 2) which suggests the presence of RNA molecules with varying R_H and potential distribution in *MW*. Coupling of the AF4 instrumentation to light scattering and dRI concentration detectors enabled us to analyze the *MW* (Fig. 3, Table 1), R_G (applies to molecules with R_G > 10 nm) and R_H distribution of the RNA components in samples that provided good signal intensities (see Section 3.3 for R_G and R_H data).

The molecules in the major peak of the pre-purified 1800 nt ssRNA and dsRNA had measured *MW* of 720,000 ± 60,000 g/mol and 1,450,000 ± 33,000 g/mol, respectively (Fig. 3A, Table 1). The corresponding dsRNA peak from the dsRNA reaction mixture provided slightly higher *MW* of 1,540,000 ± 100,000 g/mol. The measured *MW* values of pre-purified 1800 nt RNA molecules differed from the theoretical values by 1.2-fold (Table 1). The same trend was observed for the 700- and 500 nt ss- and dsRNA samples (Fig. 3; Table 1). For the 300 nt ssRNA and dsRNA the difference between the theoretical and observed *MW* was even more substantial (Table 1, Fig. S3A). Small differences between the measured and theoretical *MW* estimates could be explained by association of RNA molecules with some reaction mixture or mobile phase components such as polymerases, Mg^{2+}, Mn^{2+} and Na^{+}. However, this is unlikely to explain the observed difference. We recently observed a similar 20% difference for the theoretical and measured *MW* of 90- and 110 nt ssRNA and dsRNA molecules, while for smaller molecules such difference was not observed [33]. We used a dRI detector and dn/dc value of 0.17 ml/g to measure concentration for determination of *MW*. Examples of typical RI fractograms are shown in Fig. S4C and raw LS data for the 1800 nt, 700 nt, and 500 nt RNA molecules in Fig. S5. Alternatively, we could have used the UV detector to measure the concentration, the benefit being more sensitive measurement, but even then, obtaining accurate values would have depended on the used UV extinction factors that are different for ss- and dsRNA molecules [48].

The dsRNA samples contained RNA species that had *MW* values corresponding to the ssRNA and dsRNA molecules, respectively (Fig. 3B and C, Table 1). The content of the main dsRNA peak region appeared homogenous in terms of *MW* (Fig. 3). Unexpectedly, we observed a second peak in the AF4 fractograms of some ssRNA preparations (Fig. 3). The *MW* of RNA molecules present in this second peak was about twice that measured for the molecules

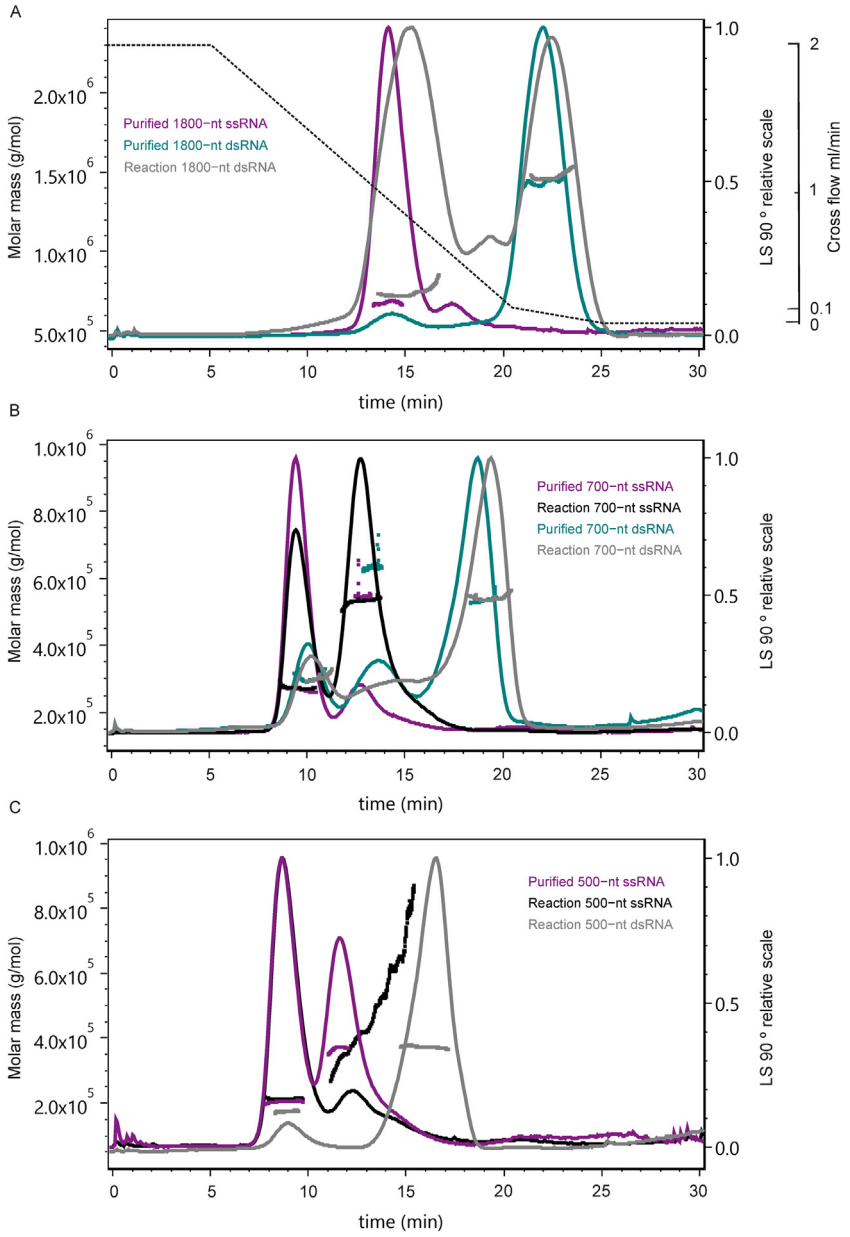

Fig. 3. AF4-MALS for molar mass distribution analysis of RNA molecules. *MW* distribution (g/mol) was compared for pre-purified RNA molecules and reaction mixtures of (A) 1800-, (B) 700-, and (C) 500 nt RNA samples. MALS signal measured at 90° is shown on a relative scale (right y-axis). Cross-flow gradient is shown with dashed line (A, right secondary y-axis). Time axis shows the fractogram from the beginning of elution excluding the focusing time.

Table 1
Theoretical and measured *MW* of selected pre-purified RNA molecules and reaction mixtures.

Name	Length (nt)	Type		Theoretical *MW* (g/mol)[1] of expected ssRNA or dsRNA	Measured *MW* (g/mol), peak 1	Measured *MW* (g/mol), peak 2
300-nt	310	ssRNA	Reaction	99,400	233,000 ± 2 000[2]	466,000 ± 9000
		dsRNA	Reaction	199,700	262,000 ± 15 000	268,000 ± 19,000
500-nt	510	ssRNA	Pre-purified	163,100	204,000 ± 2000	380,000 ± 3000
		ssRNA	Reaction	163,100	211,000 ± 2000	338,000 ± 14,000
		dsRNA	Reaction	327,000	207,000 ± 8000	373,000 ± 2000
700-nt	710	ssRNA	Reaction	228,300	273,000 ± 4000	535,000 ± 10,000
		ssRNA	Pre-purified	228,300	276,000 ± 300	533,000 ± 4000
		dsRNA	Pre-purified	456,900	310,000 ± 11,000	538,000 ± 3000
		dsRNA	Reaction	456,900	303,000 ± 1000	537,000 ± 9000
1800-nt	1810	ssRNA	Pre-purified	582,100	720,000 ± 60,000	1,170,000 ± 250,000
		dsRNA	Pre-purified	1,165,100	n.d.	1,450,000 ± 33,000
		dsRNA	Reaction	1,165,100	770,000 ± 51,000	1,540,000 ± 100,000

[1] *MW* for the ssRNA molecule representing the sense orientation and having the 5′triphosphate. For dsRNA molecules, *MW* is the sum for the sense and antisense molecules.
[2] For 300-nt ssRNA, the major first peak was preceded with a minor peak with observed *MW* of 150,000 g/mol.
n.d.: not determined.

in the first peak (Table 1) which suggests that the peak content was not monomeric ssRNA.

Agarose gel electrophoresis of the input ssRNA samples also revealed additional RNA species with slower electrophoretic mobility (Fig. S1). We used gel-purified PCR products with defined length as templates to produce RNA molecules of exact sizes. Thus, these larger RNA species were more likely non-covalent multimers formed of ssRNA molecules rather than extended monomeric ssRNA molecules. AF4 is a gentle separation method and therefore weak interactions can persist during separation [33]. Multimer formation may also be induced during the focusing step of the AF4 experiment in which sample components are concentrated that promotes sample-sample interactions, multimerization and even aggregation. If the observed multimer formation would be concentration-dependent phenomenon, higher amounts of injected RNA should increase the amount of multimers. However, we did not observe induced multimer formation or aggregation with increased RNA amounts. Instead, the intensities of peaks for monomers and multimers increased proportionally (Fig. S4).

3.3. Size distribution of the studied ssRNA and dsRNA molecules

Biophysical properties of long RNA molecules are still poorly understood. R_G value provides information on the mass-average distance of each point in the molecule from the center of gravity, whereas R_H is the radius of sphere that would have the same diffusion coefficient as the molecules under investigation in the used solvent. We estimated R_G from the number of base pairs at maximal length distance (MLD) for the MFOLD-predicted secondary structures of 1800-, 700-, 500-, and 300 nt ssRNA molecules [49]. These estimates ranged from 6.6 to 11.9 nm (Table S1). The predicted length of the corresponding dsRNA molecules ranged from 4 to 500 nm (Table S1). The estimated values were compared to those measured with MALS and DLS (Table 2).

Reliable R_G and R_H measurement requires good light scattering and dRI signal that is achievable with high injected mass and/or large molecule size that we could not achieve with all the studied RNA samples due to limits in sample volumes available for the experiments. The measured R_G value for the pre-purified 1800 nt dsRNA was 103 ± 3 nm and comparable 102 ± 2 nm value was determined for the dsRNA of the reaction mixture (Fig. S6A, Table 2). Pre-purified 1800 nt ssRNA had R_G of 21 ± 1.2 nm indicating more compact size. A similar 20 ± 0.5 nm value was measured for the first peak of the dsRNA reaction (Fig. S6A). R_H values for the ssRNA and dsRNA peaks were 15 ± 0.5 nm and 24 ± 4 nm, respectively (Fig. S6B). The same trend was observed in other measured samples: the ssRNA molecules of the first peak were significantly more compact than the dsRNA molecules of the second peak (Fig. S6C–F, Table 2). In general, the measured R_G values for ssRNA molecules were approximately 40% larger than those estimated based on formulas relying on secondary structure predictions (Table 2). There is a limited amount of information on experimentally measured R_H values for RNA molecules. Borodavka et al. have determined R_H for RNA molecules in low ionic strength buffer using fluorescence correlation spectroscopy [49]. For the 960, 1200, 1400 and 1800 nt-long RNA molecules, they measured values of 8.8–11.9 nm, but the measured R_H did not fully correlate with the number of nucleotides in the molecules [49]. Our measurements on ssRNA molecules of 500–1800 nt provided values ranging from 8.4–14.7 nm (Table 2). The shapes of ssRNA molecules are often prolate and aspherical [45] that can explain some of the variation between the measured and theoretical R_G and R_H values that assume spherical shape. In addition, these values appear to vary depending on the function of the molecules. For instance, it has been found that RNA genomes of viruses that are encapsidated to virions were significantly more compact than non-viral RNA molecules

Table 2
Predicted and measured R_G, R_H, and dsRNA length of selected pre-purified RNA molecules and reaction mixtures.

Name	Length (nt)	Type	Predicted R_G (nm) of ssRNA molecule[1]	Measured R_G (nm) of ssRNA molecule[2]	Measured R_H (nm) of ssRNA molecule[2]	Measured R_G (nm) of dsRNA molecule[2]	Measured R_H (nm) of dsRNA molecule[2]	Predicted dsRNA length (nm)[3]	Measured dsRNA length (nm)
300-nt	310	ssRNA	Reaction	6.6	11.3 ± 0.8	8.2 ± 0.2	19.9 ± 2.3	n.d.	84	n.d.
		dsRNA	Reaction	6.6	12.0 ±	n.d.	23.2 ±	n.d.	84	n.d.
500-nt	508	ssRNA	Pre-purified	7.8	11.7 ± 0.3	n.d.	16.3 ± 0.1	n.d.	140	n.d.
		ssRNA	Reaction	7.8	11.8 ± 0.4	8.4 ± 0.2	18.6 ± 1.0	n.d.	140	n.d.
		dsRNA	Reaction	7.8	n.d.	n.d.	37.6 ± 0.2	13.3 ± 0.8	140	126 ± 0.6
700-nt	710	ssRNA	Reaction	8.7	12.8 ± 1.2	n.d.	19.6 ± 4.4	n.d.	195	n.d.
		ssRNA	Pre-purified	8.7	12.1 ± 0.1	9.4 ± 0.6	17.9 ± 0.2	14.0 ± 0.6	195	n.d.
		dsRNA	Pre-purified	8.7	13.4 ± 0.9	n.d.	19.6 ± 0.1	n.d.	195	163 ± 0.5
		dsRNA	Reaction	8.7	10.8 ± 2.1	n.d.	20.1 ± 2	n.d.	195	161 ± 0.9
1800-nt	1810	ssRNA	Pre-purified	11.9	20.5 ± 1.2	n.d.	n.d.	n.d.	502	n.d.
		dsRNA	Pre-purified	11.9	n.d.	n.d.	103 ± 3	23.9 ± 2.5	502	348 ± 11
		dsRNA	Reaction	11.9	n.d.	14.7 ± 0.5	102 ± 2.4	24.7 ± 2.3	502	346 ± 4.2

[1] $R_G = N^{0.33}$ nm, where N is the number of bases in the longest predicted RNA structure (MLD) [49].
[2] Average and the corresponding standard deviation for the corresponding peak from at least three individual measurements.
[3] Length of dsRNA molecules was calculated by assuming a mean rise per base pair of 0.279 nm [43].n.d. not determined.

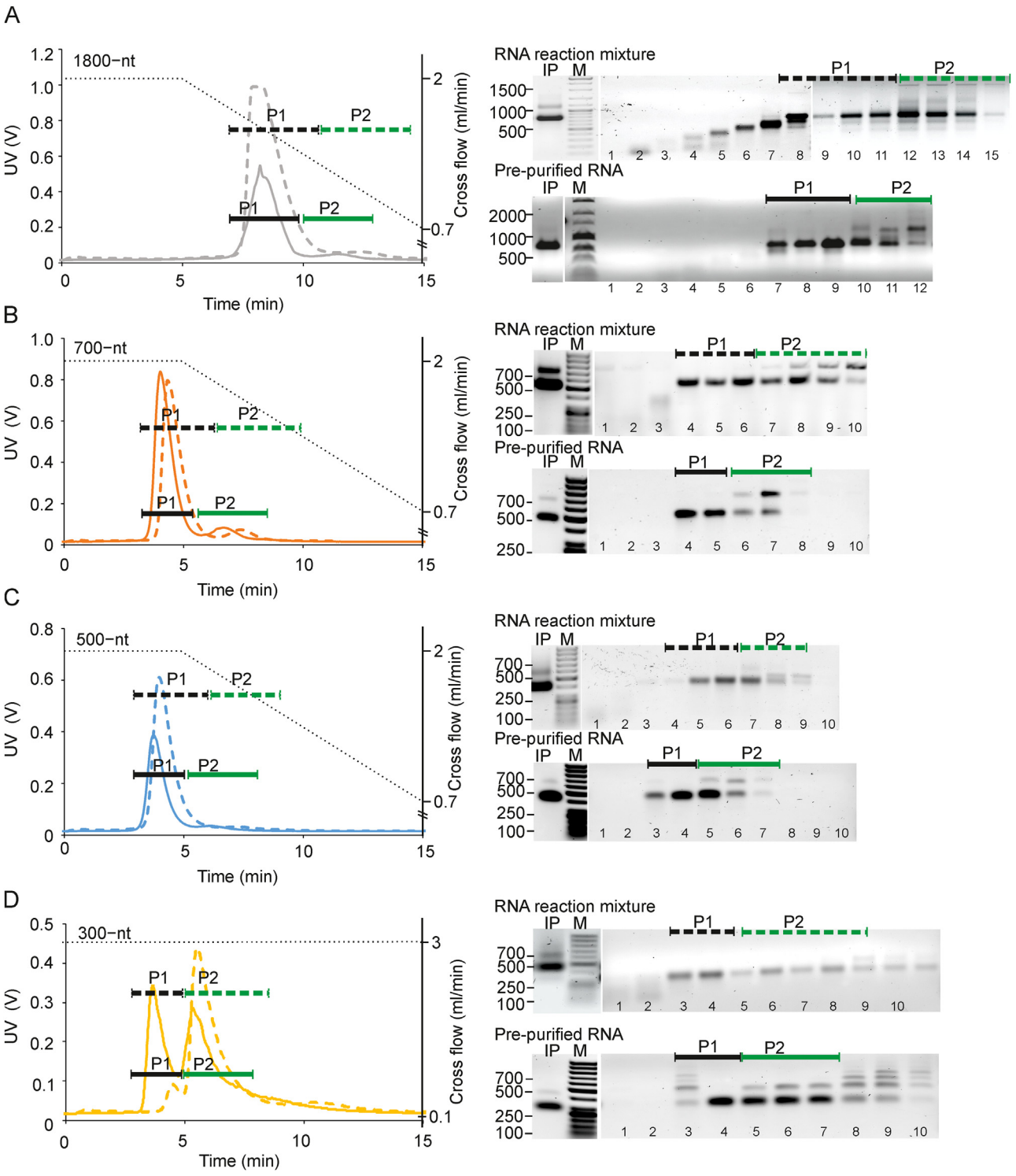

Fig. 4. Agarose gel analysis of AF4 fractionated ssRNA samples. Representative AF4 fractograms for pre-purified samples (solid lines) and non-purified reaction mixtures (dashed lines) are shown on the left and native agarose gel electrophoresis of the peak fractions as well as pre- and post-peak regions on the right for the (A) 1800-, (B) 700-, (C) 500-, and (D) 300 nt ssRNA samples. Major peak is marked as P1 (black line) and minor, later eluting low intensity peak as P2 (green line). Cross-flow profile is shown with black dashed line (right Y-axis). Time axis shows the UV fractogram from the beginning of elution excluding the focusing time. UV detector response at 260 nm is given in volts (V) (left y-axis). One ml fractions were collected starting from one min and precipitated samples analyzed by agarose gel electrophoresis with maximum load of 300 ng per well. The fraction numbers are indicated below the gel. IP refers to the input sample. M is a dsDNA size marker; mobility of selected dsDNA molecules (sizes in bp) is indicated on the left side of the agarose gel.

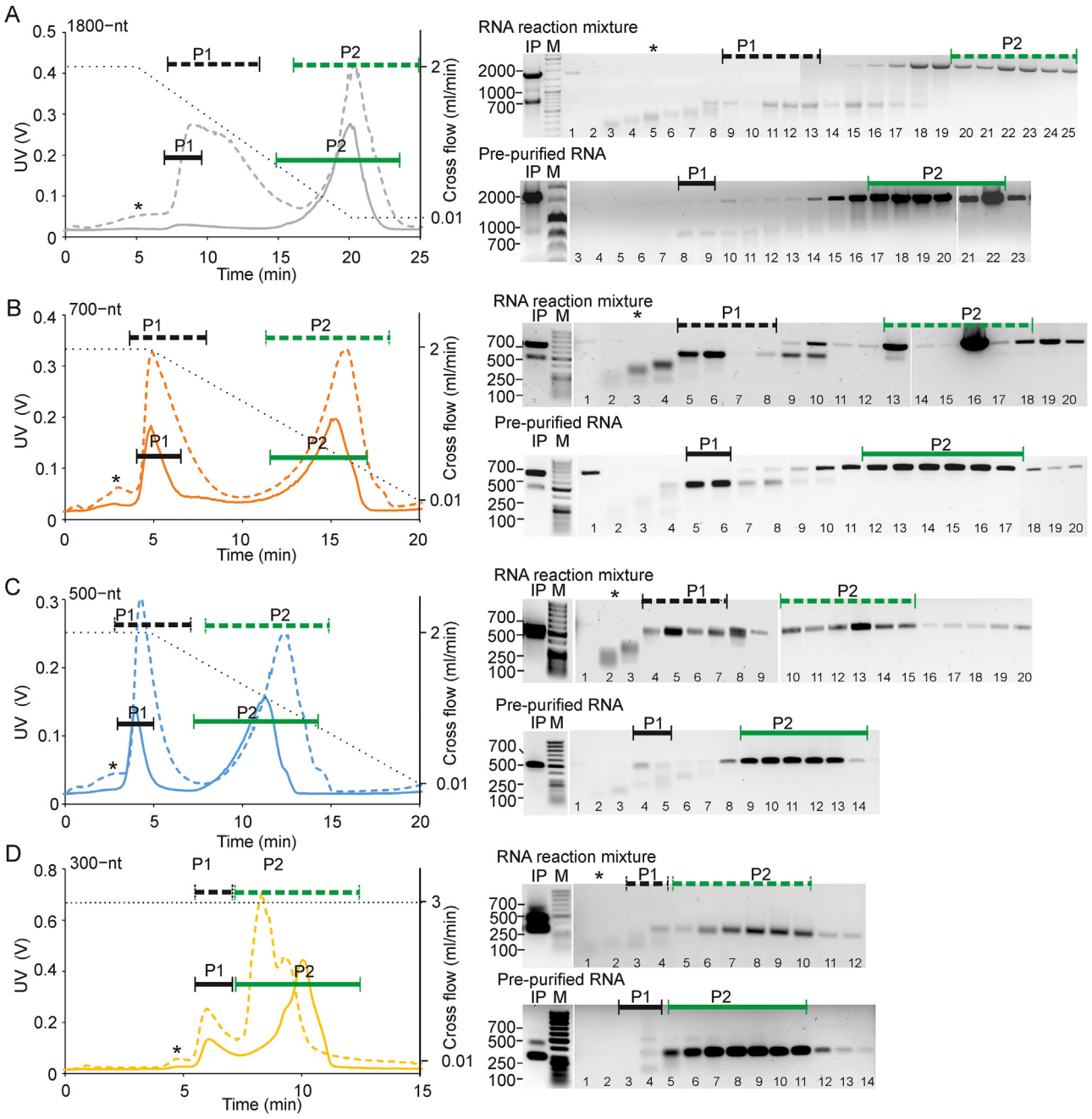

Fig. 5. Agarose gel analysis of AF4 fractionated dsRNA samples. Representative AF4 fractograms for pre-purified samples (solid lines) and non-purified (dashed lines) reaction mixtures are shown on the left, native agarose gel electrophoresis of the corresponding peak fractions as well as pre- and post-peak fractions on the right for the (A) 1800-, (B) 700-, (C) 500-, (D) 300 nt dsRNA samples. The first peak is marked as P1 (black line) and the second peak as P2 (green line). Cross-flow gradients are shown with black dashed line (right Y-axis). Time axis shows the UV fractogram from the beginning of elution excluding the focusing time. UV detector response at 260 nm is given in volts (V) (left y-axis). One ml fractions were collected starting from one min and precipitated samples analyzed by agarose gel electrophoresis with maximum load of 300 ng per well. The fraction numbers are indicated below the gel. IP refers to the input. M is a dsDNA size marker; mobility of selected dsDNA molecules (sizes in bp) is indicated on the left side of the agarose gel.

with same amount of nts [50]. Moreover, rRNA molecules appear to be more spherical than most other RNA molecules [49]. Thus, more data are needed to fully understand the folding of long RNA molecules and the impact of surrounding chemicals. Here, we have shown that coupling of AF4-mediated separation to multidetection platform offers one interesting possibility to study the biophysical characteristics of RNA molecules in various solution surroundings.

The shape factor rho, the ratio of R_G and R_H, can reveal information on the molecule conformation. We calculated it for those RNA measurements that yielded R_G and R_H information (Table 2). For ssRNA molecules the R_G/R_H ratio was close to 1.2 for the major ssRNA peak indicating deviation from a spherical shape that would have the R_G/R_H of approximately 0.78 and suggesting branched conformation [51] that can be expected for ssRNA molecules that self-fold to complex three-dimensional nanostructures. The shape factor value above 2 indicated elongated rod shape for the peak of dsRNA molecules [51], as expected.

The length of dsRNA molecules was also determined from the light scattering data (Fig. S7, Table 2). The obtained values for the 1800-, 700-, and 500 nt dsRNA molecules present in the reac-

tion mixtures were 346 ± 4, 161 ± 0.9, and 126 ± 0.6 nm, respectively. The length of the pre-purified 1800 nt dsRNA molecule was 348 ± 10 nm. In general, the measured R_H and length estimates (Table 2) were lower than the ones calculated from the base pair rise (Table S1). The sequences of the dsRNA molecules studied here are all derived from the phage ϕ6 genome which adopts a highly compact spooled structure when packed within the viral capsid [52]. The observed biophysical properties might partially reflect these functional constraints of the studied RNAs.

3.4. Agarose gel electrophoresis analysis of AF4 separated ssRNA and dsRNA

After biophysical characterization, we examined RNA content of the peaks in native agarose gels to evaluate the integrity and purity of the AF4-fractionated samples. The analysis showed that AF4 separated short RNA species, probably presented abortive ssRNA products generated by the T7 polymerase, from the full-sized ssRNAs (Fig. 4). Peak tailing and late-eluting minor peaks correlated with the larger RNA species observed in the gels (Fig. 4). As expected, reaction mixtures contained more impurities than the pre-purified samples. The 1800 nt ssRNA peak contained small RNA molecules in addition to the full-size ones (Fig. 4A). We believe that they formed after AF4 separation, because as smaller-sized molecules, their presence should have been the most prominent in fractions that eluted in front of the 1800 nt peak, but instead their amount correlated with the intensity of full-length RNA molecules (Fig. 4A). A ladder-like appearance was seen for the pre-purified 300 nt ssRNA sample in the late eluting fractions (Fig. 4D).

Two major RNA species were visible in the 1800-, 700- and 500 nt dsRNA samples in fractions that represented the two high-intensity peaks of the fractograms (Fig. 5A–C). The contaminating full-length ssRNA molecules eluted in the first peak, and dsRNA in the second. The fronting of the ssRNA peaks correlated with the observation of smaller RNA species in agarose gels (Fig. 4). Even though the 300 nt dsRNA showed inconsistent retention during AF4 fractionation (Fig. 2), the dsRNA peak contained molecules of homogenous electrophoretic mobility (Fig. 5D).

MW measurements and native agarose gel analyses suggested that the ssRNA samples contained ssRNA multimers (Figs. 3–5, Table 1). Weak intermolecular interactions can be reversed by moderate heat denaturation [53]. We compared migration of the heat-denatured and non-denatured pre-purified 300 nt ssRNA sample in agarose gel (Fig. S8A). The faint RNA species of slower electrophoretic mobility disappeared from the denatured 300 nt sample and only a single common band of faster mobility was visible. We also performed heat denaturation for the AF4-purified late-eluting fraction that migrated in the agarose gel as multiple visible bands (see Fig. 4D). Denaturation resulted in the migration of the RNA in agarose gel as a single band, indicating that the observed multiple bands represent different multimeric forms of the 300 nt ssRNA that were retained during the gentle AF4 separation (Fig. S8B).

3.5. Recoveries of AF4 purified ssRNA and dsRNA molecules

The applicability of the purification method depends on the recovered yields and on how much sample can be purified in a single experiment. We evaluated AF4 as an RNA purification method by comparing the recovered ssRNA and dsRNA amounts within the pre-purified samples and reaction mixtures using thecollected UV signals (Table 3). Depending on the efficacy of the RNA synthesis, 70–90% recoveries were achieved for the ssRNA samples when using the analytical AF4 channel. For dsRNA molecules, the yields for dsRNA molecules ranged between 40–90%. Typically, 20–30% of the injected dsRNA sample was ssRNA. The length of RNA molecules did not affect the recovered RNA yields.

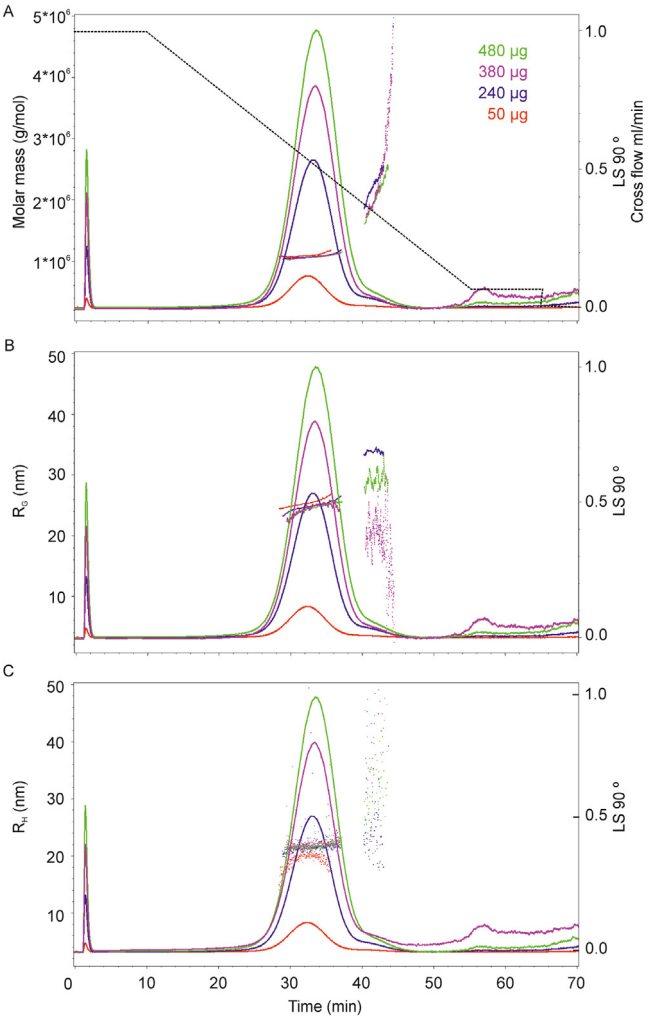

Fig. 6. Use of semipreparative channel allows for higher sample loads. A) Separation and molar mass distribution (g/mol) of pre-purified 3000 nt ssRNA molecule injected in different quantities. MALS signal measured at 90° (right y-axis) is shown on a relative scale (right y-axis). Cross-flow gradient is shown in (A) with dashed line (right secondary y-axis). Channel flow was 1.5 ml/min and detector flow 0.5 ml/min. Time axis shows the fractogram from the beginning of elution excluding the focusing time.

UV signal was also used to compare yields that were obtained from the pre-purified RNA samples and unpurified reaction mixtures (Table 3). For the ssRNA samples, direct AF4 purification of reactions yielded approximately two times higher or equal yield compared to the pre-purified samples. For dsRNA samples, direct purification of reaction mixtures with AF4 yielded comparable amounts of dsRNA that were obtained from the pre-purified samples. Thus, omitting the pre-purification process (see Fig. S1) shortens the processing time of the samples and the use of toxic chemicals such as phenol. However, the injected and recovered microgram yields were modest, if the AF4 -purified molecules would be intended for use in therapeutic or biotechnological applications. For example, milligram quantities of mRNA per patient may need to be administrated for treatment [54]. The sample load in analytical AF4 channel can be increased to some extent by further optimizing the method, but a more significant increase should be obtained by utilizing a semi-preparative channel.

We experimented the semipreparative AF4 channel using 3000 nt ssRNA as a test molecule (Table S1). Pre-purified ssRNA was injected at 50, 240, 380 and 480 µg (Fig. 6). Although in-

Table 3
Recovery of ssRNA and dsRNA molecules.

Sample		ssRNA (%)[1]	dsRNA (%)[1]	UV (V)/µl[2]	Yield, reaction versus pre-purified RNA
500-nt ssRNA	pre-purified	85.4 ± 2.2		0.026	
500-nt ssRNA	reaction	72.8 ± 3.4		0.053	2.0
500-nt dsRNA	pre-purified	24.4 ± 0.4	68.7 ± 1.2	0.023	
500-nt dsRNA	reaction	32.6 ± 0.5	54.3 ± 1.0	0.022	1.0
700-nt ssRNA	pre-purified	82.5 ± 2.4		0.067	
700-nt ssRNA	reaction	85.4 ± 3.8		0.072	1.1
700-nt dsRNA	pre-purified	25.5 ± 1.5	55.3 ± 0.7	0.040	
700-nt dsRNA	reaction	38.5 ± 1.9	50.0 ± 1.4	0.031	0.8
1800-nt ssRNA	pre-purified	85.3 ± 4.0		0.054	
1800-nt ssRNA	reaction	88.4 ± 1.1		0.052	1.0
1800-nt dsRNA	pre-purified		89.4 ± 0.7	0.032	
1800-nt dsRNA	reaction	49.2 ± 0.7	40.6 ± 0.7	0.029	0.9

[1] Recovery of injected RNA mass based on the area of the corresponding peak versus collected total UV signal.
[2] UV signal recorded in Volts (V) and divided by the volume of sample analyzed to obtain UV (V)/µl.n.d., not determined.

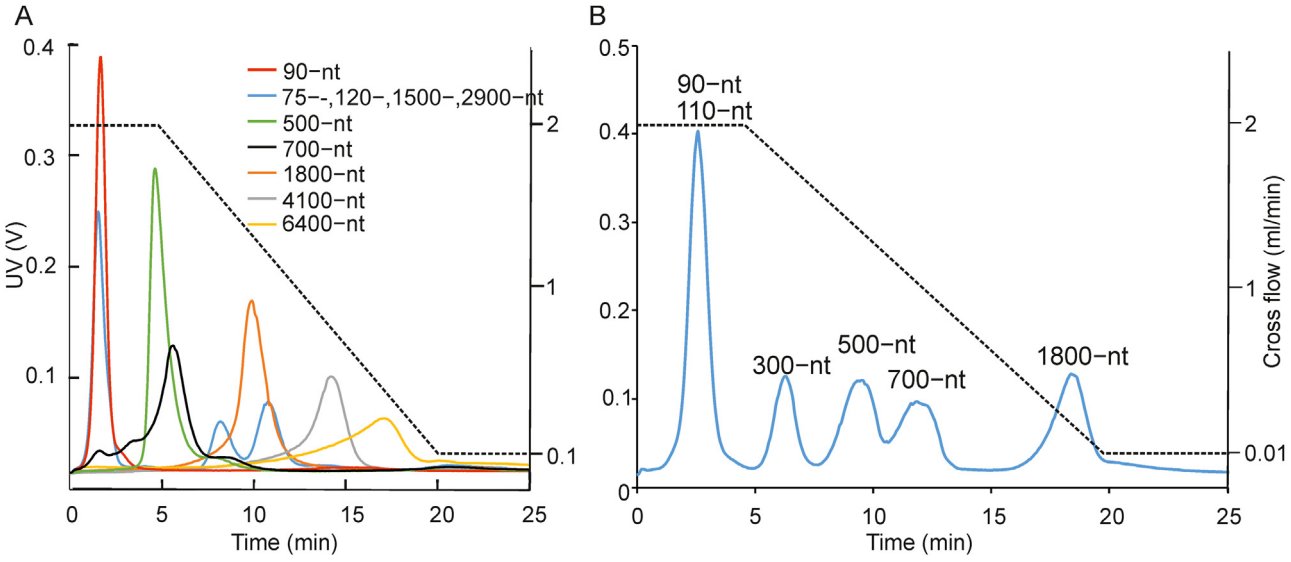

Fig. 7. Resolution power of AF4. (A) AF4 fractionation of ssRNA molecules in the size range of 75–6400 nt. (B) Fractionation of dsRNA mixture comprising molecules in a size range of 90–1800 nt. Time axis shows the UV fractogram from the beginning of elution excluding the focusing time. UV detector response at 260 nm is given in volts (V) (left y-axis). Cross-flow gradients are shown with black dashed line (right Y-axis).

creased amounts of RNA widened the major AF4 fractogram peak, the peak was separated from the following minor peak. The ssRNA recovery was 50 ± 2%. The yield can be potentially increased by further optimizing the AF4 method and the mobile phase. The measured biophysical values of *MW*, R_G and R_H were 1,070,000 ± 5000 g/mol, 26.0 ± 2.3 nm and 22.3 ± 1.0 nm (Fig. 6), whereas the theoretical *MW* estimate for the 3000 nt ssRNA was 950,000 g/mol (Table S1). Due to the limited sample quantity, we were unable to inject more than 0.5 mg RNA, which was, however, approximately ten times more than that loaded into the analytical channel. Several mg sample loads on semipreparative AF4 channel has been reported previously [55], suggesting that more RNA could be purified in a single AF4 experiment. However, the maximum amount that can be loaded to the channel obviously depends on the complexity of the sample. Monolithic anion exchange chromatography (AEX) has a high loading capacity for RNA molecules, approximately 5-8 mg/ml resin [34], and 40–50% recovery of injected RNA amount. The drawback of AEX is that RNA elutes in NaCl concentration above 0.6 M that necessitates downstream desalting steps. A clear drawback of AF4 is dilution that induces the need for concentrating steps prior to further use. The peaks from analytical and semipreparative channels eluted in 1.5–2.5 ml, respectively. Sample dilutes the most during the way from the channel end to the detectors and within their flow cells [56]. Consequently, for pure purification purposes, the extra detectors can be disconnected to diminish peak broadening.

3.6. Separation of dsRNA from the dsDNA template and proteins

RNA synthesis reaction mixtures contain dsDNA template and polymerases. Single stranded nucleic acid molecules have larger R_H than globular proteins with the same *MW* and thus they elute later in AF4 [26]. PCR products that were used as templates for RNA synthesis eluted as the corresponding dsRNA molecules in AF4 (compare Fig. 2 and Fig. S9A). This experiment indicated the importance of dsDNA removal before AF4 purification if dsRNA is the target molecule. For ssRNA molecules, the risk of co-elution with the DNA template depends on the ssRNA size. Reaction mixtures also contain ϕ6 and T7 polymerases that are 75,000 and 99,000 g/mol in *MW*, respectively. The flow conditions used did not allow separation of the monomeric and dimeric BSA (68,000 and 136,000 g/mol) or ϕ6 polymerase (75,000 g/mol) and both

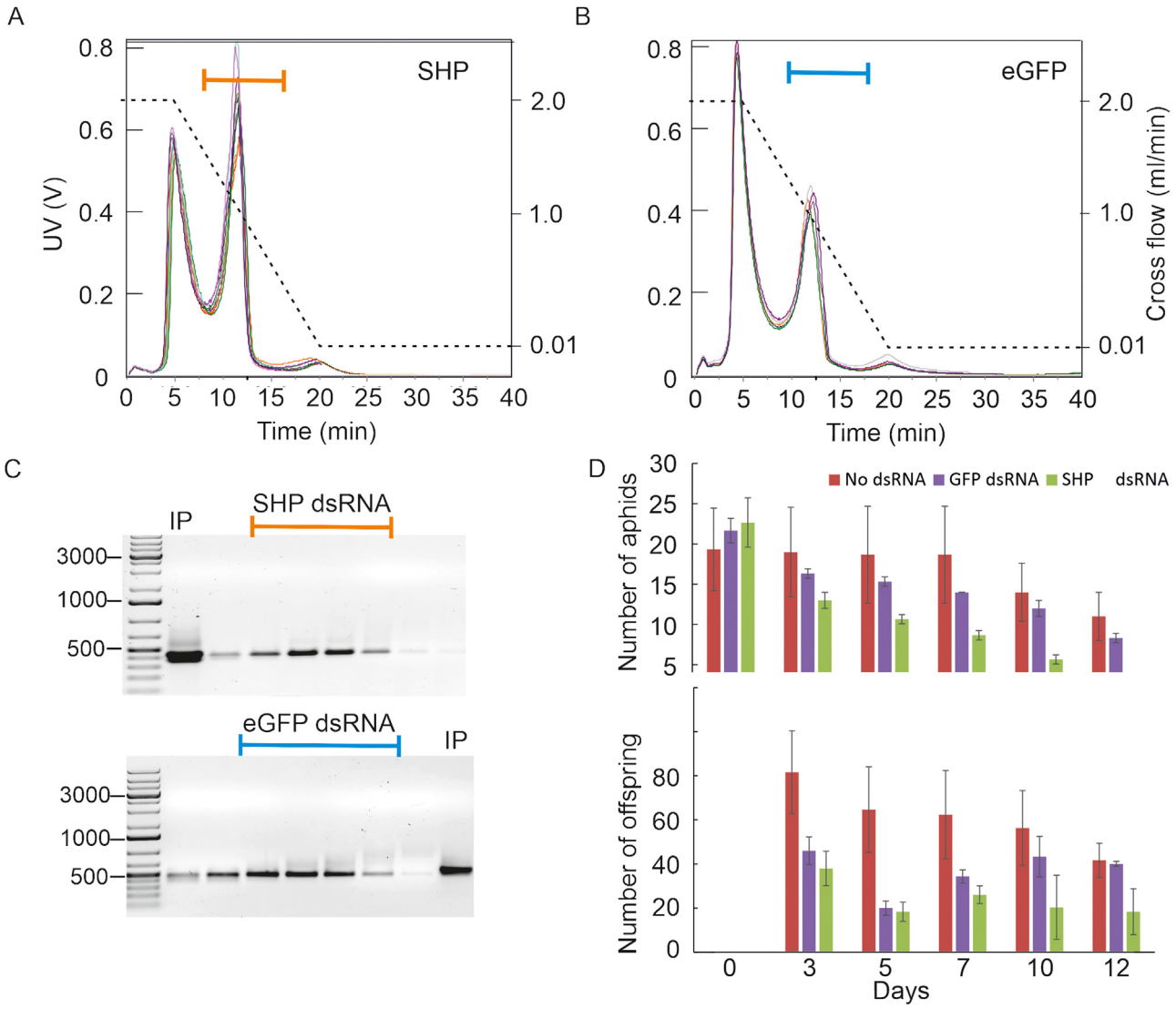

Fig. 8. AF4-purified dsRNA targeting aphid gene inhibits aphid survival. (A-B) Nine overlaid fractograms of SHP (A) and five of eGFP (B) dsRNA samples. Cross-flow gradient is shown with dashed line (right y-axis). UV detector response at 260 nm is given in volts (V) (left y-axis). Time axis shows the UV fractogram from the beginning of elution excluding the focusing time. (C) The dsRNA containing fractions [marked with orange or blue bars in (A) and (B), respectively] were analyzed in 1% (w/v) agarose gels. IP refers to input. M is a dsDNA size marker. The mobility of selected dsDNA molecules (sizes in bp) is indicated on the left side of the agarose gel. (D) The purified dsRNAs were fed to *Myzus persicae* nymphs at 2.5 ng/µl. Control feeding contained no dsRNA. The survival of adult aphids (upper panel) and production of offspring (lower panel) was monitored for 12 days. Averages and standard deviation were calculated from three replicates.

proteins eluted during the first five minutes of elution (Fig. S9B). This pinpointed the need for protein removal prior to AF4, if RNA molecules shorter than 100 nt are studied with the applied elution method.

3.7. Separation capacity of AF4

We explored the separation range of the AF4 method further by increasing the size range of the studied ssRNA molecules with 90-, 110-, 4100-, and 6400 nt RNAs as well as with bacterial total RNA, which mostly consists of ribosomal RNA (rRNA) species as well as transfer RNAs (tRNAs) (Table S1). The sizes of *Pseudomonas* rRNA species are 120 (5S), 1540 (16S) and 2910 nt (23S), whereas tRNAs are approximately 75 nt molecules. Small ssRNA species co-eluted during the 5 min constant cross-flow velocity of 2 ml/min, the 700–6400 nt ssRNAs eluted under the influence of linearly decaying cross-flow force (Fig. 7A). A difference of about twice the number of nucleotides ensured almost base-line separation for RNA molecules larger than 500 nt. For example, 16S and 23S rRNAs separated nicely, whereas the peaks for the 500- and 700 nt ssRNAs overlapped. Obviously, the R_H and shape affect the separation as well. The peaks for large, 4100- and 6400 nt, ssRNA molecules eluted under the cross-flow force, but the peaks showed strong fronting (Fig. 7A). In AF4, peak broadening and fronting is typical for large molecular species, whereas tailing is more characteristic for small molecular species [57]. A mixture of pre-purified dsRNA molecules with size range of 90–1800 nt showed good resolution for dsRNAs of 300–1800 nt size range (Fig. 7B). The difference in retention times for various dsRNA molecules was larger than for the corresponding ssRNA molecules. A similar two-fold difference in the number of nucleotides for proper separation was observed previously with AF4 separation of circular and linear dsDNA molecules [22], as well as in CIM AEX for dsRNA [34]. Compared to the AEX method that promotes baseline separation of 27–500 nt dsRNA molecules from a mixture of 27, 58, 100, 200, 500, 1800, 3000, 4100, 6400 bp dsRNAs [34], AF4 separation

with the applied flow velocities worked better for slightly larger dsRNA molecules with size range of 300–1800 bp. Size separation of dsRNA molecules beyond 1800 bp would require lower cross-flow velocity.

We evaluated further AF4 performance on large dsRNA molecules that are subject to shear forces [34]. We fractionated total RNA samples that were isolated from ϕ6 -infected *Pseudomonas* cells at late stage of infection. Such cells contain three viral genomic dsRNA segments: L (large, ~6.4 kb), M (medium, ~4.1 kb) and S (small, 2.9 kb) in addition to cellular RNAs that mostly consist of rRNA [58] (Table S1). These dsRNA molecules eluted at the end of the elution program, indicating normal elution mode despite the expected large size of 1–1.8 µm (Table S1, Fig. S10A). Agarose gel analysis indicated that these dsRNA species were of good quality and did not experience shear-induced degradation (Fig. S10B) that could potentially occur during injection of large molecules through tubing with small diameter [56]. Many pathogenic viruses have long RNA genomes and double-stranded replication intermediates. The described AF4 method provides a potential method for detection and isolation of such molecules from RNA samples. However, analysis of heterologous mixtures of ssRNA and dsRNA molecules may be challenging due to potential co-elution of ssRNA molecules of large size with dsRNA molecules of smaller size. In addition, the analysis of total RNA samples can be hampered by the large amount of rRNA.

3.8. Functionality of AF4 purified dsRNA molecules

We tested the success of AF4 purification of dsRNA in functional tests. RNAi-mediated gene expression knockdown can be achieved in aphids by feeding them on artificial diets that contain dsRNA to trigger the production of siRNAs [5]. We evaluated the effect of AF4-purified dsRNA against the aphid gene *SHP* that is involved in stylet function and phloem feeding [59]. AF4-purified eGFP dsRNA served as a non-relevant dsRNA control. We purified the 470 nt SHP dsRNAs with AF4 to separate dsRNA from ssRNA and other reaction components (Fig. 8A). The dsRNA-containing fractions were combined, desalted, concentrated and mixed with artificial diet to feed two different aphid species. Artificial medium lacking dsRNA monitored the aphid's normal mortality and reproduction and served as the control. Aphids that obtained SHP dsRNA in their diet had higher mortality than aphids that were fed with eGFP dsRNA or the control diet that contained no dsRNA (Fig. 8B for *Myzus persicae* and Fig. S11 for *Sitobion avenae*). In addition, the production of offspring was markedly decreased when the diet contained SHP dsRNA (Fig. 8B).

4. Conclusions

The use of RNA molecules in therapeutic and biotechnical applications is increasing. A recent example is Covid-19 mRNA vaccine. Such applications rely on good quality ss- and dsRNA molecules. A prerequisite for the economic viability is that the used RNA molecules can be produced and purified with high efficacy and recovery. Thus, there is demand for methods that enable purification of native RNA molecules in high quantity and quality. Also, analytical methods promoting biophysical characterization of RNA molecules in solution are in demand as biophysical properties affect their bioavailability and pharmacokinetics. We show that AF4 can be utilized to purify ss- and dsRNA molecules of wide size range. AF4 fractograms provided valuable information on the quality and quantity of the studied RNA samples. Coupling of the fractionation to concentration and light scattering detectors promoted the application of AF4 as an analytical tool to obtain data on the *MW* and size distribution as well as yield. Our experiments showed that pre-purification of reaction mixtures was not necessary prior to purification with AF4 provided that the studied RNA molecules are large enough to enable their separation from the protein components present in the reaction mixtures. Utilization of the semipreparative channel allowed upscaling of the purification. However, AF4 experiment, like any other purification method, is always a compromise between the separation capacity, dilution and time. In general, we think that AF4 coupled to light scattering detectors opens a new avenue on the study of solution properties of RNA that are essential sources of information on RNA architecture, their compactness or extendedness and dynamics. This information is still largely missing for long RNA molecules.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Eskelin Katri: Conceptualization, Investigation, Data curation, Supervision, Visualization, Writing – review & editing. **Lampi Mirka:** Investigation, Formal analysis, Visualization, Writing – original draft. **Coustau Christine:** Resources, Investigation, Writing – review & editing. **Imani Jafargholi:** Resources, Investigation, Writing – review & editing. **Kogel Karl-Heinz:** Resources, Supervision, Writing – review & editing. **Poranen Minna M:** Conceptualization, Resources, Writing – review & editing, Supervision, Funding acquisition.

Data availability

Data will be made available on request.

Acknowledgments

Tanja Westerholm, Helin Veskiväli, and Sari Korhonen are thanked for excellent technical assistance. The facilities and expertise of the HiLIFE Biocomplex unit at the University of Helsinki, a member of Instruct-ERIC Centre Finland, FINStruct, and Biocenter Finland are gratefully acknowledged. This research has been carried out under the ERA-NET Cofund SusCrop with funding from the Ministry of Agriculture and Forestry, Finland, and German Research Foundation DFG. The research has also been supported by the Academy of Finland (grant numbers 331627 and 328112) and Sigrid Juselius Foundation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2022.463525.

References

[1] A.P. Aalto, L.P. Sarin, A.A. Van Dijk, M. Saarma, M.M. Poranen, U. Arumäe, D.H. Bamford, Large-scale production of dsRNA and siRNA pools for RNA interference utilizing bacteriophage phi6 RNA-dependent RNA polymerase, RNA 13 (3) (2007) 422–429, doi:10.1261/rna.348307.

[2] A. Romanovskaya, H. Paavilainen, M. Nygårdas, D.H. Bamford, V. Hukkanen, M.M. Poranen, Enzymatically produced pools of canonical and Dicer-substrate siRNA molecules display comparable gene silencing and antiviral activities against herpes simplex virus, PloS One 7 (11) (2012) e51019, doi:10.1371/journal.pone.0051019.

[3] A. Levanova, M.M. Poranen, RNA interference as a prospective tool for the control of human viral infections, Front Microbiol. 9 (2151) (2018), doi:10.3389/fmicb.2018.02151.

[4] M.A. Valencia-Sanchez, J. Liu, G.J. Hannon, R. Parker, Control of translation and mRNA degradation by miRNAs and siRNAs, Genes Dev. 20 (5) (2006) 515–524, doi:10.1101/gad.1399806.
[5] R. Katoch, A. Sethi, N. Thakur, L.L. Murdock, RNAi for insect control: current perspective and future challenges, Appl. Biochem. Biotechnol. 171 (4) (2013) 847–873, doi:10.1007/s12010-013-0399-4.
[6] R.L. Setten, J.J. Rossi, S.P. Han, The current state and future directions of RNAi-based therapeutics, Nat. Rev. Drug. Discov. 18 (6) (2019) 421–446, doi:10.1038/s41573-019-0017-4.
[7] P.R. Das, S.M. Sherif, Application of exogenous dsRNAs-induced RNAi in agriculture: challenges and triumphs, Front. Plant. Sci. 11 (2020), doi:10.3389/fpls.2020.00946.
[8] A. Niehl, M. Soininen, M.M. Poranen, M. Heinlein, Synthetic biology approach for plant protection using dsRNA, Plant. Biotechnol. J. 16 (9) (2018) 1679–1687, doi:10.1111/pbi.12904.
[9] S. Liu, M.J. Ladera-Carmona, M.M. Poranen, A.J.E. Van Bel, K.H. Kogel, J. Imani, Evaluation of dsRNA delivery methods for targeting macrophage migration inhibitory factor MIF in RNAi-based aphid control, J. Plant. Dis. Prot. 128 (5) (2021) 1201–1212, doi:10.1007/s41348-021-00464-9.
[10] T.R. Damase, R. Sukhovershin, C. Boada, F. Taraballi, R.I. Pettigrew, J.P. Cooke, The limitless future of RNA therapeutics, Front. Biol. Eng. Biotech. 9 (2021), doi:10.3389/fbioe.2021.628137.
[11] S. Guo, C. Xu, H. Yin, J. Hill, F. Pi, P. Guo, Tuning the size, shape and structure of RNA nanoparticles for favorable cancer targeting and immunostimulation, WIREs Nanomed. Nanobiotechnol. 12 (1) (2020) e1582, doi:10.1002/wnan.1582.
[12] H. Li, T. Lee, T. Dziubla, F. Pi, S. Guo, J. Xu, C. Li, F. Haque, X.-J. Liang, P. Guo, RNA as a stable polymer to build controllable and defined nanostructures for material and biomedical applications, Nano Today 10 (5) (2015) 631–655, doi:10.1016/j.nantod.2015.09.003.
[13] R. Martins, J.A. Queiroz, F. Sousa, Ribonucleic acid purification, J. Chrom. A 1355 (2014) 1–14, doi:10.1016/j.chroma.2014.05.075.
[14] J.C. Giddings, F.J. Yang, M.N. Myers, Flow-field-flow fractionation: a versatile new separation method, Science 193 (4259) (1976) 1244, doi:10.1126/science.959835.
[15] J. Gigault, J.M. Pettibone, C. Schmitt, V.A. Hackley, Rational strategy for characterization of nanoscale particles by asymmetric-flow field flow fractionation: a tutorial, Anal. Chim. Acta 809 (2014) 9–24, doi:10.1016/j.aca.2013.11.021.
[16] K.G. Wahlund, Flow field-flow fractionation: critical overview, J. Chrom. A 1287 (2013) 97–112, doi:10.1016/j.chroma.2013.02.028.
[17] F.A. Messaud, R.D. Sanderson, J.R. Runyon, T. Otte, H. Pasch, S.K.R. Williams, An overview on field-flow fractionation techniques and their applications in the separation and characterization of polymers, Prog. Polym. Sci. 34 (4) (2009) 351–368, doi:10.1016/j.progpolymsci.2008.11.001.
[18] K. Eskelin, M.M. Poranen, H.M. Oksanen, Asymmetrical flow field-flow fractionation on virus and virus-like particle applications, Microorganisms 7 (11) (2019), doi:10.3390/microorganisms7110555.
[19] C.L. Plavchak, W.C. Smith, C.R.M. Bria, S.K.R. Williams, New advances and applications in field-flow fractionation, Ann. Rev. Anal. Chem. 14 (1) (2021) 257–279, doi:10.1146/annurev-anchem-091520-052742.
[20] K.G. Wahlund, H.S. Winegarner, K.D. Caldwell, J.C. Giddings, Improved flow field-flow fractionation system applied to water-soluble polymers: programming, outlet stream splitting, and flow optimization, Anal. Chem. 58 (3) (1986) 573–578, doi:10.1021/ac00294a018.
[21] J.C. Giddings, M.A. Benincasa, M.K. Liu, P. Li, Separation of water soluble synthetic and biological macromolecules by flow field-flow fractionation, J. Liq. Chromatogr. 15 (10) (1992) 1729–1747, doi:10.1080/10826079208018323.
[22] M.K. Liu, J.C. Giddings, Separation and measurement of diffusion coefficients of linear and circular DNAs by flow field-flow fractionation, Macromolecules 26 (14) (1993) 3576–3588.
[23] A. Litzen, K.G. Wahlund, Improved separation speed and efficiency for proteins, nucleic acids and viruses in asymmetrical flow field flow fractionation, J. Chrom. A 476 (1989) 413–421.
[24] K.G. Wahlund, A. Litzen, Application of an asymmetrical flow field-flow fractionation channel to the separation and characterization of proteins, plasmids, plasmid fragments, polysaccharides and unicellular algae, J. Chrom. A 461 (1989) 73–87.
[25] E. Haladjova, S. Rangelov, M. Geisler, S. Boye, A. Lederer, G. Mountrichas, S. Pispas, Asymmetric flow field-flow fractionation investigation of magnetopolyplexes, Macromol. Chem. Phys. 216 (18) (2015) 1862–1867, doi:10.1002/macp.201500177.
[26] J. Ashby, S. Schachermeyer, Y. Duan, L.A. Jimenez, W. Zhong, Probing and quantifying DNA–protein interactions with asymmetrical flow field-flow fractionation, J. Chrom. A 1358 (2014) 217–224, doi:10.1016/j.chroma.2014.07.002.
[27] D. Kang, M.H. Moon, Miniaturization of frit inlet asymmetrical flow field-flow fractionation, Anal. Chem. 76 (13) (2004) 3851–3855, doi:10.1021/ac0496704.
[28] M. Nilsson, S. Birnbaum, K.G. Wahlund, Determination of relative amounts of ribosome and subunits in Escherichia coli using asymmetrical flow field-flow fractionation, J. Biochem. Biophys. Methods 33 (1) (1996) 9–23.
[29] L. Pitkänen, P. Tuomainen, K.J.A. Eskelin, B. Chemistry, Analysis of plant ribosomes with asymmetric flow field-flow fractionation, Anal. Bioanal. Chem. 406 (6) (2014) 1629–1637.
[30] R. Mildner, S. Hak, J. Parot, A. Hyldbakk, S.E. Borgos, D. Some, C. Johann, F. Caputo, Improved multidetector asymmetrical-flow field-flow fractionation method for particle sizing and concentration measurements of lipid-based nanocarriers for RNA delivery, Eur. J. Pharm. Biopharm. 163 (2021) 252–265, doi:10.1016/j.ejpb.2021.03.004.
[31] J. Zhang, R.M. Haas, A.M. Leone, Polydispersity characterization of lipid nanoparticles for siRNA delivery using multiple detection size-exclusion chromatography, Anal. Chem. 84 (14) (2012) 6088–6096.
[32] J. Ashby, K. Flack, L.A. Jimenez, Y. Duan, A.K. Khatib, G. Somlo, S.E. Wang, X. Cui, W. Zhong, Distribution profiling of circulating microRNAs in serum, Anal. Chem. 86 (18) (2014) 9343–9349, doi:10.1021/ac5028929.
[33] A.L. Levanova, K. Kalke, V. Hukkanen, M.M. Poranen, K. Eskelin, Native RNA purification method for small RNA molecules based on asymmetrical flow field-flow fractionation, Pharmaceuticals 15 (261) (2022), doi:10.3390/ph15020261.
[34] A. Romanovskaya, L.P. Sarin, D.H. Bamford, M.M. Poranen, High-throughput purification of double-stranded RNA molecules using convective interaction media monolithic anion exchange columns, J. Chrom. A 1278 (2013) 54–60, doi:10.1016/j.chroma.2012.12.050.
[35] P. Gottlieb, J. Strassman, X. Qiao, M. Frilander, A. Frucht, L. Mindich, *In vitro* packaging and replication of individual genomic segments of bacteriophage phi 6 RNA, J. Virol. 66 (5) (1992) 2611–2616.
[36] L. Mindich, X. Qiao, S. Onodera, P. Gottlieb, M. Frilander, RNA structural requirements for stability and minus-strand synthesis in the dsRNA bacteriophage phi 6, Virology 202 (1) (1994) 258–263, doi:10.1006/viro.1994.1341.
[37] V.M. Olkkonen, P. Gottlieb, J. Strassman, X.Y. Qiao, D.H. Bamford, L. Mindich, *In vitro* assembly of infectious nucleocapsids of bacteriophage phi 6: formation of a recombinant double-stranded RNA virus, Proc. Natl. Acad. Sci. USA 87 (23) (1990) 9173–9177, doi:10.1073/pnas.87.23.9173.
[38] M. Jiang, P. Österlund, L.P. Sarin, M.M. Poranen, D.H. Bamford, D. Guo, I. Julkunen, Innate immune responses in human monocyte-derived dendritic cells are highly dependent on the size and the 5′ phosphorylation of RNA molecules, J. Immunol. 187 (4) (2011) 1713–1721 %J, doi:10.4049/jimmunol.1100361.
[39] E.V. Makeyev, D.H. Bamford, Replicase activity of purified recombinant protein P2 of double-stranded RNA bacteriophage phi6, EMBO J. 19 (1) (2000) 124–133, doi:10.1093/emboj/19.1.124.
[40] D.H. Bamford, P.M. Ojala, M. Frilander, L. Walin, J.K. Bamford, Isolation, purification, and function of assembly intermediates and subviral particles of bacteriophages PRD1 and σ6, in: Methods in Molecular Genetics, Elsevier, 1995, pp. 455–474.
[41] M. Zuker, P. Stiegler, Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, Nucleic Acids Res. 9 (1) (1981) 133–148, doi:10.1093/nar/9.1.133.
[42] L.T. Fang, W.M. Gelbart, A. Ben-Shaul, The size of RNA as an ideal branched polymer, J. Chem. Phys. 135 (15) (2011) 155105, doi:10.1063/1.3652763.
[43] J.A. Abels, F. Moreno-Herrero, T. Van Der Heijden, C. Dekker, N.H. Dekker, Single-molecule measurements of the persistence length of double-stranded RNA, Biophys. J. 88 (4) (2005) 2737–2744, doi:10.1529/biophysj.104.052811.
[44] E. Naessens, G. Dubreuil, P. Giordanengo, O.L. Baron, N. Minet-Kebdani, H. Keller, C. Coustau, A secreted MIF cytokine enables aphid feeding and represses plant immune responses, Curr. Biol. 25 (14) (2015) 1898–1903, doi:10.1016/j.cub.2015.05.047.
[45] C. Hyeon, R.I. Dima, D. Thirumalai, Size, shape, and flexibility of RNA structures, J. Chem. Phys. 125 (19) (2006) 194905, doi:10.1063/1.2364190.
[46] Z. Ren, R. Ghose, Slow conformational dynamics in the cystoviral RNA-directed RNA polymerase P2: influence of substrate nucleotides and template RNA, Biochemistry 50 (11) (2011) 1875–1884, doi:10.1021/bi101863g.
[47] J.R. Diaz-Ruiz, J.M. Kaper, Isolation of viral double-stranded RNAs using a LiCl fractionation procedure, Prep. Biochem. 8 (1) (1978) 1–17, doi:10.1080/00327487808068215.
[48] A.O. Nwokeoji, P.M. Kilby, D.E. Portwood, M.J. Dickman, Accurate quantification of nucleic acids using hypochromicity measurements in conjunction with UV spectrophotometry, Anal. Chem. 89 (24) (2017) 13567–13574, doi:10.1021/acs.analchem.7b04000.
[49] A. Borodavka, S.W. Singaram, P.G. Stockley, W.M. Gelbart, A. Ben-Shaul, R. Tuma, Sizes of long RNA molecules are determined by the branching patterns of their secondary structures, Biophys. J. 111 (10) (2016) 2077–2085, doi:10.1016/j.bpj.2016.10.014.
[50] A.M. Yoffe, P. Prinsen, A. Gopal, C.M. Knobler, W.M. Gelbart, A. Ben-Shaul, Predicting the sizes of large RNA molecules, Proc. Natl. Acad. Sci. USA 105 (42) (2008) 16153–16158, doi:10.1073/pnas.0808089105.
[51] L. Nilsson, Starch and other polysaccharides, in: S.K.R. Williams, K.D. Caldwell (Eds.), Field-Flow Fractionation in Biopolymer Analysis, Springer-Verlag, Wien, 2012, pp. 165–185.
[52] S.L. Ilca, X. Sun, K. El Omari, A. Kotecha, F. De Haas, F. Dimaio, J.M. Grimes, D.I. Stuart, M.M. Poranen, J.T. Huiskonen, Multiple liquid crystalline geometries of highly compacted nucleic acid in a dsRNA virus, Nature 570 (7760) (2019) 252–256, doi:10.1038/s41586-019-1229-9.
[53] A.O. Nwokeoji, S. Kumar, P.M. Kilby, D.E. Portwood, J.K. Hobbs, M.J. Dickman, Analysis of long dsRNA produced *in vitro* and *in vivo* using atomic force microscopy in conjunction with ion-pair reverse-phase HPLC, Analyst 144 (16) (2019) 4985–4994, doi:10.1039/C9AN00954J.
[54] U. Sahin, K. Karikó, Ö. Türeci, mRNA-based therapeutics–developing a new class of drugs, Nat. Rev. Drug. Dis. 13 (10) (2014) 759–780, doi:10.1038/nrd4278.
[55] C.R.M. Bria, P.W. Skelly, J.R. Morse, R.E. Schaak, S.K.R. Williams, Semi-preparative asymmetrical flow field-flow fractionation: a closer look at chan-

nel dimensions and separation performance, J. Chrom. A 1499 (2017) 149–157, doi:10.1016/j.chroma.2017.03.017.
[56] C.R. Bria, S.K. Williams, Impact of asymmetrical flow field-flow fractionation on protein aggregates stability, J. Chrom. A 1465 (2016) 155–164, doi:10.1016/j.chroma.2016.08.037.
[57] M. Marioli, W.T. Kok, Recovery, overloading, and protein interactions in asymmetrical flow field-flow fractionation, Anal. Bioanal. Chem. 411 (11) (2019) 2327–2338, doi:10.1007/s00216-019-01673-w.
[58] S. Mäntynen, L.R. Sundberg, M.M. Poranen, Recognition of six additional cystoviruses: pseudomonas virus phi6 is no longer the sole species of the family Cystoviridae, Arch. Virol. 163 (4) (2018) 1117–1124, doi:10.1007/s00705-017-3679-4.
[59] E. Abdellatef, T. Will, A. Koch, J. Imani, A. Vilcinskas, K.H. Kogel, Silencing the expression of the salivary sheath protein causes transgenerational feeding suppression in the aphid Sitobion avenae, Plant. Biotechnol. J. 13 (6) (2015) 849–857, doi:10.1111/pbi.12322.