

B-CePs as cross-linking probes for the investigation of RNA higher-order structure

Alice Sosic^{1,*}, Richard Göttlich², Dan Fabris³ and Barbara Gatto¹

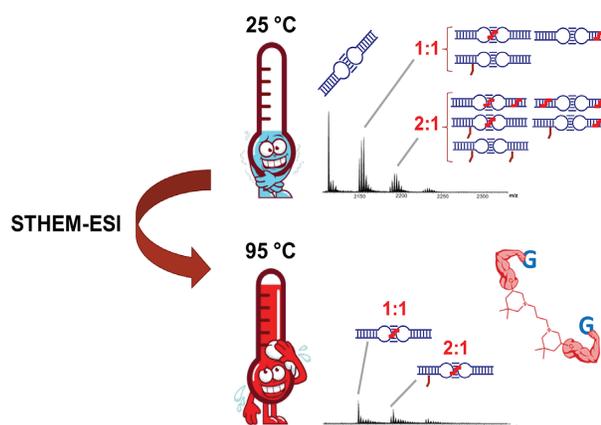
¹Department of Pharmaceutical and Pharmacological Sciences, University of Padova, 35131 Padova, Italy, ²Institute of Organic Chemistry, Justus Liebig University Giessen, 35392 Giessen, Germany and ³Departments of Chemistry and Biological Sciences, University at Albany-SUNY, Albany, NY, 12222, USA

Received January 04, 2021; Revised May 13, 2021; Editorial Decision May 14, 2021; Accepted May 19, 2021

ABSTRACT

Elucidating the structure of RNA and RNA ensembles is essential to understand biological functions. In this work, we explored the previously uncharted reactivity of bis-chloropiperidines (B-CePs) towards RNA. We characterized at the molecular level the different adducts induced by the fast reacting compound B-CeP 1 with RNA. Following an approach based on solution thermal melting coupled with ESI mass spectrometry (STHEM-ESI), we proved the ability of B-CePs to induce inter-molecular cross-links between guanines in double stranded RNA. These results open the possibility of using B-CePs as structural probes for investigating higher-order structures, such as the kissing loop complex established by the dimerization initiation site (DIS) of the HIV-1 genome. We confirmed the potential of B-CePs to reveal the identity of RNA structures involved in long-range interactions, expecting to benefit the characterization of samples that are not readily amenable to traditional high-resolution techniques, and thus promoting the elucidation of pertinent RNA systems associated with old and new diseases.

GRAPHICAL ABSTRACT



INTRODUCTION

Substantial progress in RNA research is profoundly changing our understanding of the biological significance of this biopolymer, while driving technological applications and scientific innovation (1,2). The RNA revolution has highlighted the importance of a myriad of non-coding RNAs (ncRNAs) that play pivotal roles in essential biological processes in all living organisms (3). Besides, the increasing threat to human health due to old and new RNA viruses, coupled with their highly disrupting economic and social consequences, make the challenge of probing and blocking key viral RNAs even more urgent (4). The structures adopted by RNA sequences are essential to carry out their intended functions and are critical for binding partners such as other nucleic acids, proteins and small molecule ligands. However, methods and tools employed to characterize the interactions of RNA at the molecular level are still limited as compared to those available for proteins. An established approach involves the utilization of small molecules with well-defined activity as chemical probes to accurately and reliably investigate specific structures (1). In the context of RNA structural studies, widely used methods for structure-

*To whom correspondence should be addressed. Tel: +39 049 8275691; Fax: +39 049 8275366; Email: alice.sosic@unipd.it
Present address: Dan Fabris, Department of Chemistry, University of Connecticut, Storrs, CT 06269, USA.

probing experiments at single nucleotide-resolution employ chemicals reacting with folded RNA in a quantitative and structurally informative way, helping to elucidate mechanistic and phenotypic questions (5).

Bis-3-chloropiperidines (B-CePs) are an emerging class of alkylating agents with excellent reactivity toward nucleic acids (6–9). B-CePs react with DNA leading to depurination and strand cleavage at guanines (dG) (6–9). The analysis of a large library of B-CePs showed that a high reactivity toward DNA is detrimental to their development as cytotoxic agents (10,11), but might be instead ideal to support possible applications as chemical probes for structural studies. To test this hypothesis, we selected compound **1** (1,3-bis(5-chloro-3,3-dimethylpiperidin-1-yl)propane, Figure 1A) as a possible prototype of B-CePs structural probe by virtue of its high reactivity with nucleophiles (6,8). In this report, we specifically explored its activity toward structured RNA by using both single- and double-stranded substrates (Figure 1B). Electrophoretic and mass spectrometric analysis were employed in concert to elucidate molecular mechanism of the reaction of B-CeP **1** with this type of biopolymer. We implemented an approach that coupled solution thermal melting with electrospray ionization (STEM-ESI) (12,13) mass spectrometry to probe rapidly and unambiguously the presence of inter-strand conjugates in folded RNA structures.

This approach offered us the ability to clarify the identity of the stable bi-functional adducts between adjacent guanines in paired RNA strands. We finally evaluated the potential of B-CeP **1** as a possible cross-linking probe by investigating the spatial arrangement of the kissing complex formed by the dimerization initiation site (DIS) domain of the HIV-1 genome (Figure 1C), which represents an excellent example of a highly structured RNA of great biological and pharmacological relevance.

MATERIALS AND METHODS

RNA substrates

Stock solutions of individual dinucleotides (i.e. ApA, CpC, GpG, UpU) were 2'-ACE deprotected as suggested by the manufacturer (GE Healthcare Dharmacon, CO, USA), dried and then dissolved in nuclease-free water. Synthetic oligoribonucleotides were purchased in lyophilized form from commercial sources. The RNA oligonucleotides used for gel electrophoresis analysis were chemically synthesized by Metabion International AG (Martinsried, Germany), whereas those used for mass spectrometric analysis were purchased from Integrated DNA Technologies (IDT, Coralville, IA, USA). All oligonucleotides were used as received with no further purification. All names, sequences and secondary structures are reported in Figure 1. Each substrate possessed specific features explained in the text. RNA duplex such as dsRNA was obtained by annealing equimolar amounts of the complementary strands RNA1 and RNA2 in 1× BPE buffer (BPE 10×: NaH₂PO₄·2H₂O 2 mM, Na₂HPO₄·12H₂O 6 mM, Na₂EDTA·2 H₂O 1 mM, pH 7.4). The mixture was denatured at 95°C for 5 min, and then left to cool to room temperature to obtain the desired RNA duplex (Figure 1B). The DIS kissing-loop (KL) dimer

was prepared by heating to 95°C for 5 min the RNA solution and then quickly cooling on ice to allow for proper folding of the RNA hairpins and obtain the kissing-loop conformer (KL, Figure 1C) (14,15).

Chemical reagents

Bis-3-chloropiperidine **1** (Figure 1A) was synthesized in house as previously described (8). Chlorambucil (Chl) was purchased from Sigma-Aldrich (Milan, Italy). Aliquots of chemical probes were freshly prepared by diluting an 8 mM DMSO stock in MilliQ water and were instantly reacted with the RNA substrate to avoid the typical quenching effects of the aqueous environment. All the other chemical reagents, including salts and solvents, were purchased from Sigma-Aldrich (Milan, Italy).

B-CeP **1** hydroxylation kinetic

The formation of B-CeP **1** reactive and hydrolyzed species upon incubation at 37°C in 1× BPE buffer (pH 7.4) was followed in time by ESI-MS. A fresh solution of test compound was prepared at the concentration of 80 μM and incubated at 37°C. Samples were taken after 5 min, 15 min, 30 min, 1 h, 2 h and 24 h, diluted 1:10 with methanol and analyzed on a Thermo Fisher Scientific (West Palm Beach, CA, USA) LTQ-Orbitrap Velos mass spectrometer. The analyses were performed in nanoflow mode by using quartz emitters produced in house by using a Sutter Instruments Co. (Novato, CA, USA) P2000 laser pipette puller. Up to 5 μl samples were typically loaded onto each emitter by using a gel-loader pipette tip. A stainless steel wire was inserted in the back-end of the emitter to supply an ionizing voltage around 1 kV. Typical source temperature was 200°C, whereas desolvation voltage was in the range of 40–50 V. Data were processed by using Xcalibur 2.1 software (Thermo Scientific). We calculated the percentage of each species over total compound in each sample to obtain relative percentages of B-CeP **1** reacted species, whose chemical structures are schematized in panel B of Supplementary Figure S1.

Probing reactions

Based on the final aim of the experiment, different reaction conditions were explored, including different substrates, probe to substrate ratios and incubation times. Case by case reaction conditions are detailed within the text. Typical reaction mixtures consisted of 20 μl of a final 2 μM solution substrate, properly folded when necessary, in 1× BPE buffer. RNA solutions were added with compound **1** or Chlorambucil as reference substance at final concentrations ranging from 5 to 50 μM and incubated at 37°C for different times, depending on the purpose of the experiment.

Gel electrophoretic analysis

An initial assessment of probing reactions results was achieved by submitting reaction mixtures to sequencing polyacrylamide gel electrophoresis (PAGE) analysis. Appropriate RNA strand was 5'-radioactively labeled by using

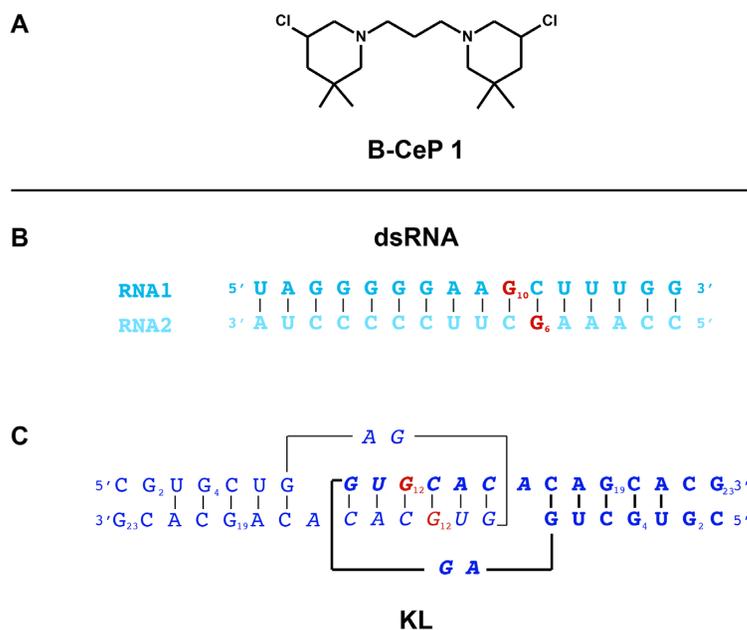


Figure 1. Representative chemical probe and RNA model substrates employed in the study. (A) Chemical structure of bis-3-chloropiperidine derivative **1**. (B) Sequence and secondary structure of the RNA duplex (dsRNA) formed by annealing RNA1 and RNA2. (C) Sequence and secondary structure of the kissing loop (KL) complex formed by two copies of the HIV-1 DIS. The two monomers of DIS are shown in blue, normal and bold respectively. Loop sequences are shown in italic letters. Guanines with specific features mentioned in the text are numbered and/or highlighted in red.

T4 polynucleotide kinase (Ambion, Applied Biosystems, TX, USA) in the presence of [γ - 32 P] ATP (3000 Ci/mmol, PerkinElmer, MA, USA) at 37°C for 1 h, following manufacturer's instructions. The kinase activity was inactivated by heating the reaction mixture at 95°C for 5 min, followed by two phenol extractions to remove the enzyme. Reaction mixtures were prepared by mixing a final 2 μ M solution of 5'-radiolabeled RNA substrate in 1 \times BPE buffer with compound **1** at final concentrations of either 5 or 50 μ M. Reaction mixtures were incubated at 37°C for 1, 2, 4, 7, 15 and 24 h, as indicated in the text. Samples were dried in a vacuum centrifuge, re-suspended in 5 μ l of denaturing gel loading buffer (10 mM Tris-HCl, 80% formamide, 0.025% bromophenol blue), and then loaded onto a 20% denaturing polyacrylamide gel (7 M urea) in 1 \times TBE buffer (Tris-HCl 89 mM, borate 89 mM, EDTA 2 mM). Acrylamide/bis-acrylamide (19:1) 40% solution was purchased from VWR International PBI Srl (Milan, Italy). 5'-radiolabeled oligonucleotide products were detected by using a STORM B40 Phosphorimager (GE Healthcare, Italy).

Mass spectrometric analysis

Samples prepared in 1 \times BPE were buffer-exchanged by performing ethanol precipitation in the presence of 1 M ammonium acetate. In case of reaction mixtures, this treatment served also to achieve reaction quenching. Samples were re-dissolved and diluted in 150 mM ammonium acetate (pH adjusted to 7.0) to achieve a final 2 μ M concentration of total RNA. All samples were analyzed by direct infusion electrospray ionization (ESI) on a Thermo Fisher Scientific (West Palm Beach, CA, USA) LTQ-Orbitrap Velos mass spectrometer. The analyses were performed in nanoflow mode as above described for the 'B-CeP **1** hydroxylation ki-

netic'. Source temperature and desolvation conditions were adjusted by closely monitoring the incidence of ammonium adducts and water clusters. Typical source temperature was 200°C, whereas desolvation voltage was in the range of 40–50 V, enabling the detection of intact RNA and RNA adducts. All experiments were performed in negative ion mode. The instrument was calibrated by using a 0.5 mg/ml solution of CsI in 50/50 water/methanol, which provided a typical 2 ppm mass accuracy. Data were processed by using Xcalibur 2.1 software (Thermo Scientific). Melting experiments were performed by using a home-built temperature-controlled nanospray source (described in the next paragraph) that allowed us to perform thermal melting of any sample solution contained in the nanospray emitter. All experiments were performed in negative ion mode.

Home-built temperature-controlled nanospray source

The temperature-controlled heater for nanospray analysis was built from a brass block, which was cut in two halves and carved to obtain a channel that fit snugly around a typical quartz emitter (see Supplementary Figure S4). The halves have through holes to accommodate a thermocouple probe (Omega, PR-21-1) and two 25W heater cartridges (Tempco, HDC19100). The channel allows for a stainless-steel wire to be inserted from the back end of the emitter to provide the necessary spray voltage. The wire is insulated from the block by a PTFE sleeve to avoid voltage discharge. When the two halves are reassembled, the entire block can be encased in a water jacket made of silicone material. The temperature is set by a programmable controller (Laird, TC-XX-PR-59) unit with a precision of $\pm 0.1^\circ\text{C}$ at 1 Hz. The water jacket is connected to a temperature-controlled circulator, which can be used to maintain temperatures below

ambient and to expedite cooling. The apparatus is capable of raising the temperature from 4°C to 25°C in <1 min. The fact that the large mass of the block requires some time to reach the set temperature is more than compensated by long-term temperature stability afforded by the apparatus. This feature is a particularly beneficial for studying slow nucleic acid interactions, which can take minutes to hours.

RESULTS AND DISCUSSION

B-CeP 1 reacts fast with nucleophiles

B-CePs reactivity lies in the fast formation of the reactive aziridinium ion (N⁺) that is readily attacked by nucleophiles (6,8). In aqueous solution, aziridinium ions react with water to produce several species comprising mono-(OH) and di-hydroxylated (2 OH) products (6,10). It is therefore possible to assess the kinetics of B-CePs reactivity by analyzing the species obtained with water in the absence of competing nucleic acids (10). To do so, a solution of **1** was incubated at 37°C in BPE buffer from 5 min to 24 h. Small aliquots were taken at different incubation times and directly analyzed by high resolution electrospray ionization mass spectrometry (ESI-MS). Representative spectra and the structure of all identified species are thoroughly reported in Supplementary Figure S1 (see Supporting Information), while graphs in Figure 2 show the relative percentage of B-CeP **1** species at different incubation time.

The results at 5 min displayed the immediate formation of the reactive electrophilic aziridinium ion (species **1**_{N⁺}, turquoise in Figure 2). Subsequent attack by the nucleophilic water was revealed by the appearance of the 3-hydroxy substituted species (species **1**_{OH}, mint green in Figure 2). After 15 min incubation, the intact compound **1** was no longer observed, whereas the formation of aziridinium ion on the second 3-chloropiperidine moiety (species **1**_{N⁺OH}, orange in Figure 2) could be readily noted. The fact that all reactive intermediates were completely converted into water-reacted **1**_{2OH} species (peach-colored in Figure 2) after 2 h incubation provided a measure of the rather high rate of reaction of compound **1** in aqueous environment. No further changes could be noted in the sample incubated up to 24 h. The kinetics of reactant consumption observed here were consistent with earlier observations (6,8), which supported its possible utilization as a fast-acting chemical probe.

B-CeP 1 reacts with guanines in RNA

With the aim to gain information on the base-specific reactivity with RNA, we treated compound **1** with an equimolar combination of homo-dinucleotides, namely ApA (green), CpC (blue), GpG (red), UpU (black). A representative ESI-MS spectrum obtained after 2 h reaction is provided in Figure 3A, whereas that of an untreated control mixture is in Supporting Information, Supplementary Figure S2. Characteristic covalent adducts were readily detected, thus demonstrating the ability of B-CePs to alkylate RNA substrates. More specifically, the products corresponded to the monofunctional addition of compound **1** to GpG, ApA, and CpC (species **a**, **b** and **c** in Figure 3A and Table 1).

In contrast, the mass of species **d** was consistent with a putative bifunctional adduct of GpG with compound **1** bridging across adjacent guanines. At the same time, no adduct whatsoever could be observed for UpU, probably due to the significantly lower nucleophilic character of the corresponding type of nucleobase. These data allowed us to quantify the yields of the various alkylating reactions by using the unreacted UpU as a putative internal standard. A safe estimate of probe reactivity should also take in account the expected variation of ionization efficiency between reactants and products. In ESI-MS experiments, the intensity of signals afforded by nucleic acid samples is susceptible to their overall hydrophobicity and gas-phase proton affinity, which are established by their base composition (16,17). This effect can be readily appreciated by examining the spectrum of the untreated homo-dinucleotide mixture (Supporting Information, Supplementary Figure S2), in which different signal intensities were recorded for the various species in spite of their equimolar concentration in solution. The effects of chemical modifications may affect the ionization characteristics of the species in an unpredictable way, but they are likely to change substantially for adducts of small and singly-charged substrates such as the dinucleotides employed in the experiment. The intensities observed for each dinucleotide immediately before addition of compound **1** and after 2 h incubation were normalized against that of UpU, thus providing the means for obtaining an accurate measure of the percentage of the initial material consumed during the reaction. As shown in Figure 3B, the results afforded a G >> A ≈ C scale of relative nucleobase-reactivity. Considering that the formation of bifunctional alkylation products must necessarily involve the intramolecular closing of an initial mono-functional precursor, the lower overall rate of formation of such product exhibited by adenine and cytosine could explain the absence of cross-linked products noted for the corresponding dinucleotides under the selected experimental conditions. These results reveal the preferential attack of B-CeP **1** to guanines, leading to both mono- and bi-functional adducts.

Fast reaction of B-CeP 1 with single-stranded RNA yields stable adducts

In light of the encouraging results obtained from the dimeric oligoribonucleotides, we evaluated the reactivity of bis-3-chloropiperidine **1** towards single-stranded RNA by using a combination of gel electrophoresis and mass spectrometry analyses. As a model system, we utilized a 16-mer RNA construct (named RNA1, Figure 1B), which replicated a DNA oligonucleotide described in previous studies (6). Its sequence contained a stretch of five adjacent guanine ribonucleotides (G), which were expected to constitute multiple nucleophilic centers susceptible to alkylation. Compound **1** was reacted with RNA1 according to the experimental conditions described previously for the DNA analog (6), i.e. by mixing 2 μM of nucleic acid substrate with either 5 or 50 μM compound in BPE buffer at 37°C for incubation times up to 24 h. A radioactively labeled version of RNA1 was also prepared to enable product analysis by high-resolution polyacrylamide gel electrophoresis (PAGE). The detection of numerous discrete bands with

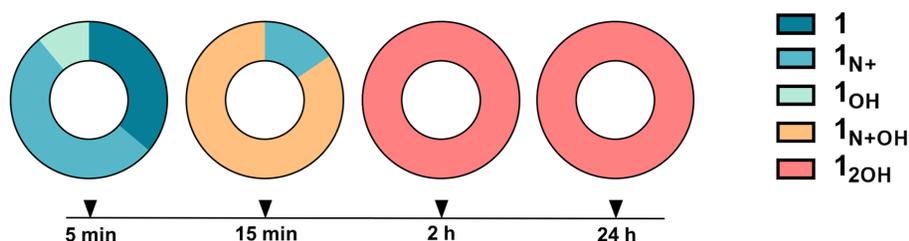


Figure 2. Time-dependent hydroxylation of B-CeP **1** in BPE buffer pH 7.4 at 37°C. The graphs report the percentage of each hydrolysis products listed in the legend, which was obtained by ESI-MS analysis (see Supporting Information, Supplementary Figure S1, for the corresponding spectra). Each incubation interval is reported on the time axis.

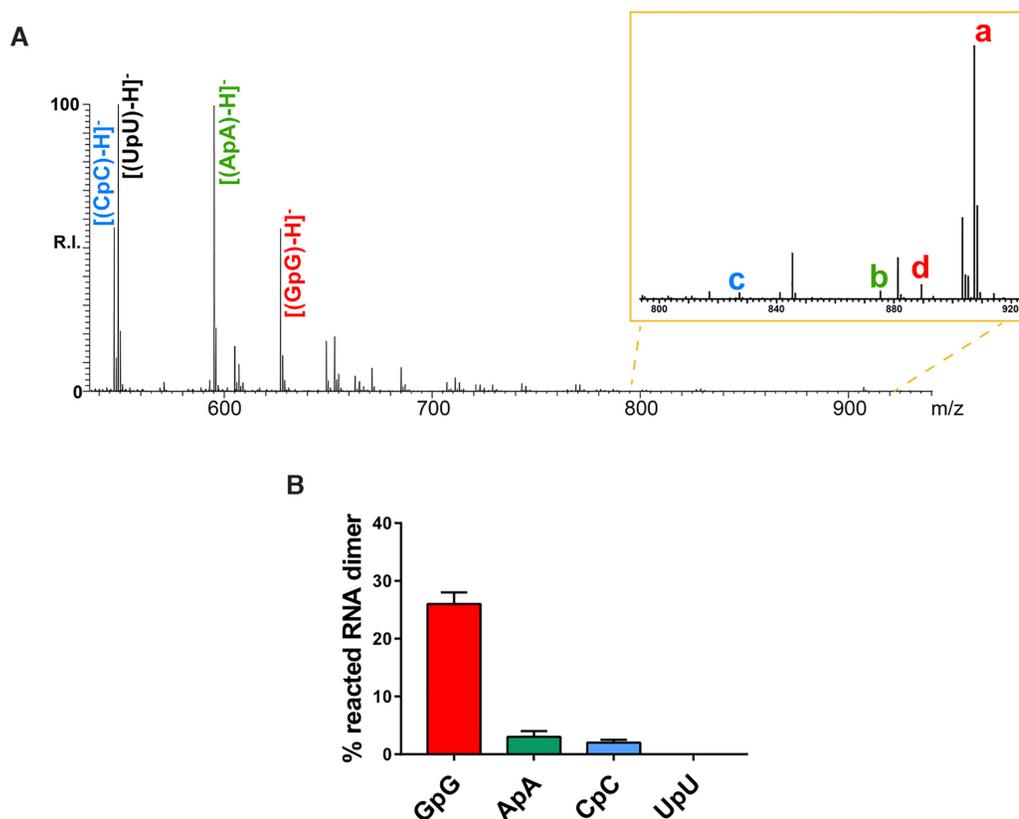


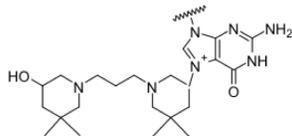
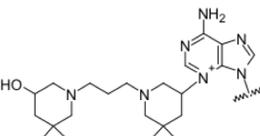
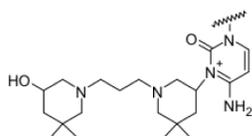
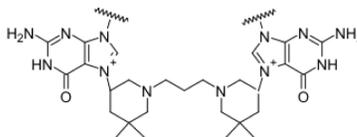
Figure 3. (A) Representative ESI-MS spectrum of a sample containing 50 μM B-CeP **1** and equimolar amounts of RNA homo-dinucleotides (5 μM each) in BPE buffer (pH 7.4) after 2 h incubation at 37°C. Spectrum was obtained by direct infusion ESI-MS in negative ion mode. In the inset, **a**, **b** and **c** correspond to the monofunctional alkylation products of GpG, ApA and CpC, respectively; **d** is the bifunctional product of GpG; additional signals correspond to typical sodium and ammonium adducts. (B) Histograms displaying the percentages of reacted RNA species observed in panel A, which provide a measure of the relative reactivities manifested by B-CeP **1** towards each type of dinucleotide.

lower mobilities than that of the untreated control was attributed to covalent B-CeP **1** adducts (Figure 4A). Increasing the amount of probe in solution increased both abundance and stoichiometry of adducts in the reaction mixture, to the point in which the initial substrate could be no longer detected at the higher concentration (i.e. 50 μM). The reactivity of B-CeP **1** was compared with that of the nitrogen mustard chlorambucil, a potent and effective DNA alkylating agent (18). In clear contrast with the former, the latter was unable to induce RNA adducts even at the highest concentration employed in the experiment (i.e. 100 μM) after 24 h incubation (lane Chl in Figure 4A). Interestingly, the fact that no band migrating faster than the control was ever de-

tected indicated that the B-CeP **1** products were very stable and unable of producing cleavage of the biopolymer chain. These exquisite features provided a striking contrast with the reactivity exhibited by **1** toward DNA, which produced extensive backbone cleavage after only 1 h incubation under otherwise identical conditions (6).

Prompted by this observation, we performed ESI-MS analysis of reaction mixtures to achieve the characterization of RNA alkylation products at the molecular level. The representative data in Figure 4B were obtained from a sample containing 2 μM of RNA1 and 5 μM of compound **1**, which was incubated for 2 h at 37°C. Only the region containing the 4- charge state is shown for the sake of clar-

Table 1. Nomenclature, nucleobases involved, symbols, representative chemical structure, experimental and calculated masses of species detected in the spectrum reported in Figure 3A

Name	RNA base	Symbol	Chemical structure	Experimental mass (u)	Calculated mass (u)
a	G			908.39	908.39
b	A			876.40	876.40
c	C			828.38	828.38
d	G - G			890.38	890.38

ity. The base peak labeled $[\text{RNA1-4H}]^{4-}$ was readily recognized as the unreacted substrate, whereas additional signals were assigned to covalent RNA adducts by matching experimental masses with values calculated from the RNA1 sequence and the structure of compound **1** (see Supporting Information, Supplementary Table S1). More specifically, the $\text{RNA1}+\mathbf{1}_B$ species displaying an incremental mass of 262.24 u over that of unmodified RNA1 was readily assigned to putative bi-functional adducts formed by reaction of both 3-chloropiperidine moieties with the RNA strand (indicated by  in Figure 4B, inset). In contrast, the $\text{RNA1}+\mathbf{1}_M$ species with a 280.24 u increment was diagnostic for mono-functional adducts in which only one of the 3-chloropiperidines had reacted with the substrate, whereas the other had been hydrolyzed to 3-hydroxyl by the aqueous environment (indicated by  in Figure 4B, inset).

Analyses performed as a function of concentration and time corroborated the observations afforded by the gel electrophoresis experiments by revealing the identity of the species with lower mobility. The increasing amounts of products with masses matching different combinations of mono and bifunctional adducts (labelled , , ) in Supplementary Figure S3, Supporting Information) con-

firmed the dose-dependence of the reaction between B-CeP **1** and RNA1. The coexistence of free RNA1 and the alkylation products with two equivalent of probe indicates that the second alkylation events occurred before saturation of a first reaction site was complete. The absence of a predominant product, which might be clearly favoured over the others, thus suggested that the various guanines in the construct possessed similar susceptibility to alkylation. This finding was consistent with the absence of any stable secondary structure that may prevent access by the reagent. The fact that no variations of the alkylation pattern were observed for incubation times greater than 2 h confirmed the relatively fast rate of reaction observed with GpG in the dinucleotide mixture. In this direction, the time-dependence of the B-CeP **1** reaction with RNA1 (Figure 4A) matched very closely that of hydrolysis in the aqueous buffer (Figure 2). The majority of alkylating probes have an unfavourable balance between hydrolysis and reaction with the intended target, which lead to anemic yields and call for greater reagent concentrations (19). The comparable reaction rates afforded by B-CeP **1** would be expected to enable a fairer competition, which should afford better probing reaction yields.

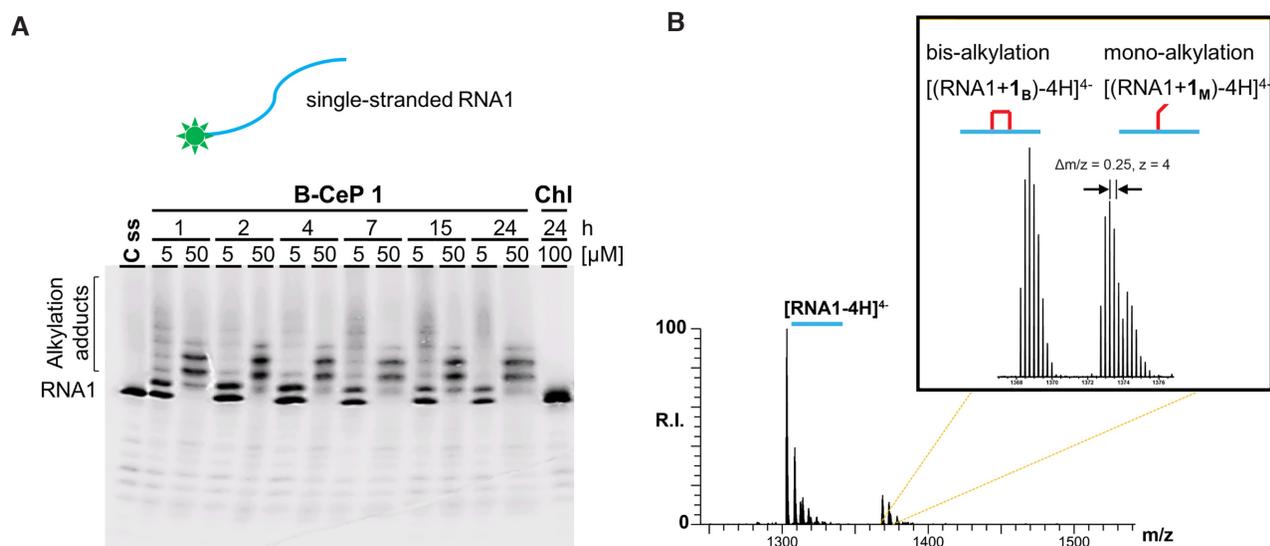


Figure 4. (A) Concentration- and time-dependence of the reaction between B-CeP **1** and single-stranded RNA (RNA1). 2 μM aliquots of radiolabeled RNA were treated with compound **1** at either 5 or 50 μM final concentrations at 37°C in BPE buffer, pH 7.4, and incubated for the indicated time. The reaction mixtures were analyzed by denaturing polyacrylamide gel electrophoresis (PAA 20%, 7M urea, TBE 1X). ‘C-ss’ (single-stranded control) indicates an untreated sample of RNA1. Chlorambucil (Chl) was used as reference substance. (B) Representative ESI-MS spectrum of reaction mixture obtained by incubating unlabeled RNA1 (2 μM) with 5 μM bis-3-chloropiperidine **1** for 2 h at 37°C. Spectra were recorded in 150 mM ammonium acetate in negative ion mode. Lower intensity signals near free/bound species are typical sodium and ammonium adducts. The inset shows an enlarged view of the 4- charge state region, containing the signals of bis-alkylated RNA1+**1**_B (—) and mono-alkylated RNA1+**1**_M (—) adducts.

B-CePs can effectively cross-link double-stranded RNA

We next evaluated the potential of B-CeP **1** as a putative structural probe by reacting it with different structured RNAs. The first RNA assembly consisted of a double-stranded construct (dsRNA, Figure 1B), which was obtained by annealing RNA1 with its complementary strand RNA2. The reaction conditions employed in these experiments replicated those identified earlier for the single-stranded RNA1. Analysis by denaturing PAGE revealed the formation of a series of alkylation products of radiolabeled RNA1, which displayed the usual dose- and time-dependent trends (Figure 5A). These bands matched very closely those produced by the reaction with individual single-stranded RNA1 (Figure 4A). The absence of radioactive tracer on the complementary RNA2 prevented its detection in these experiments. However, alkylation adducts were observed when the dsRNA construct included the radioactively labeled RNA2 (Figure 5B). In both cases, a new product with much lower mobility could be clearly detected at the higher reagent concentrations and longer reaction times (indicated by red arrows in Figure 5), which was absent in the reaction mixture of individual single-stranded RNA1 (Figure 4A). The fact that no band was detected for intact dsRNA in the control lane containing untreated samples (i.e. C ds lanes in Figure 5) confirmed that the denaturing conditions employed in the analysis were definitively capable of inducing complete dissociation of the annealed strands. Therefore, the slow-moving band could be plausibly assigned to bifunctional conjugated products of dsRNA, in which inter-strand cross-linking by B-CeP **1** prevented the expected dissociation.

The rate of reaction with the double-stranded construct was comparable to that observed with the single-stranded

one. Complete consumption of the initial material and formation of putative inter-strand conjugates were already observed after only 1 h incubation with 50 μM of B-CeP **1**. By comparison, treatment with the reference probe chlorambucil induced no significant amounts of adducts upon more extended incubation with higher reagent concentration, as exemplified by the lane containing the products of 24 h incubation with 100 μM probe (Figure 5). Under these conditions, the inefficient reactivity of chlorambucil could be explained by a combination of possible kinetics and structural factors. In particular, the high yield of hydrolysis products exhibited by typical nitrogen mustards in aqueous environments would be expected to favor hydroxylated over alkylated products. Upon formation of an initial mono-functional product, the competing pathways involving the second 2-chloroethyl function could thus lead almost exclusively to ‘dangling’ hydroxylated products rather than ‘closed’ bifunctional conjugates across opposite strands (16,19). Moreover, it should be noted that the structure of the dsRNA construct is characterized by the presence of a unique GC step at the 10–11 position of the RNA1 sequence (Figure 1B). These nucleotides are base-paired with a complementary CG sequence in RNA2. The arrangement dictated by the Watson-Crick rules places G10 of the former in close spatial proximity with G6 of the latter. Owing to the presence of only one G in the entire RNA2 sequence, a similar arrangement would not be possible anywhere else in the double-stranded structure. Therefore, the two guanines on opposite strands were considered as the most likely sites to be bridged by B-CeP **1** in the inter-strand conjugate. This hypothesis was further supported by the predominant reactivity of B-CeP **1** with guanine, which minimized any possible involvement of other types

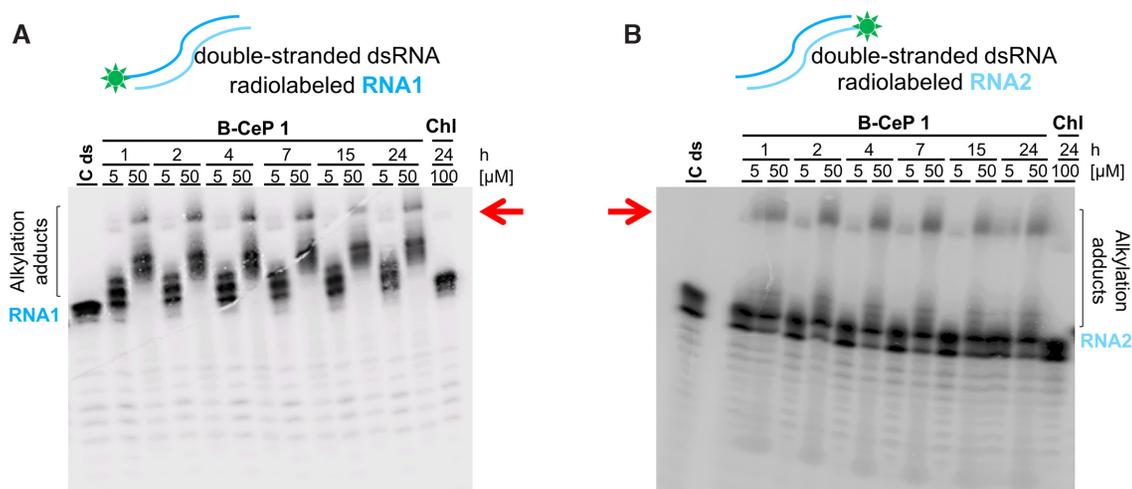


Figure 5. Denaturing PAGE analysis of cross-linking reaction mixtures between B-CeP **1** and double-stranded RNA (dsRNA) substrate. Either radiolabeled RNA1 (A) or RNA2 (B) strand was employed to enable visualization. 2 μM aliquots of RNA were treated with compound **1** at either 5 or 50 μM final concentration at 37°C in BPE buffer, pH 7.4, and incubated for the indicated time. The reaction mixtures were analyzed by denaturing polyacrylamide gel electrophoresis (PAA 20%, 7M urea, TBE 1 \times). ‘C ds’ (double-stranded control) indicates an untreated sample of RNA duplex. Chlorambucil (Chl) was used as reference probe. The red arrow indicates slow-migrating bands consistent with a putative inter-strand cross-link.

of nucleotides in the cross-linking reactions. Furthermore, the steric situation of guanine in a typical helical structure leaves the N7 position (the most nucleophilic function in the nucleotide) exposed to the solvent and readily accessible to any alkylating agent. Differently from chlorambucil, which was unable to produce significant bifunctional conjugates across opposite strands (Figure 5), the detection of stable inter-strand conjugates for B-CeP **1** indicated that the spacing between 3-chloropiperidine functions was clearly appropriate to cover the distance between susceptible N7 positions.

STHEM-ESI analysis of inter-strand RNA cross-links

The identity of the products obtained from the double-stranded substrate was further investigated by ESI-MS analysis. Figure 6A provides the spectrum obtained from a representative reaction mixture containing 2 μM dsRNA and 5 μM B-CeP **1**, which was incubated for 2 h at 37°C. These conditions enabled the detection of abundant products with different stoichiometries including up to three equivalents of B-CeP **1** adduct. Based on the respective masses, the various signals could be immediately assigned to the initial unreacted substrate (indicated with — in Figure 6A) and ‘dangling’ monofunctional products in which only one of the 3-chloropiperidine functions had reacted with dsRNA (labelled — , see inset in Figure 6A). In contrast, putative intra- and inter-strand bifunctional products, labelled respectively — and — , shared the same elemental composition and thus could not be unambiguously distinguished according to mass (see inset in Figure 6A). The fact that the selected experimental conditions were unequivocally non-denaturing, as demonstrated by the detection of intact duplex construct, precluded the possibility of differentiating the bifunctional conjugates by inducing strand dissociation. This usually desirable feature prevented

us from verifying the information obtained from the same type of sample by denaturing PAGE.

Over the years, different approaches based on gas-phase activation have been employed to induce strand dissociation in cross-linked samples, which have met variable levels of success as a function of system complexity and overall stability of binding between complementary strands (19–21). In this report, we explored an approach that coupled solution thermal melting with electrospray ionization mass spectrometry (STHEM-ESI) to achieve strand dissociation directly in the sample solution rather than in the gas phase. The home-built apparatus employed in these experiments consisted of a copper block surrounding the nanospray emitter, which could be heated under computer control in 0.2°C increments (see Supporting Information, Supplementary Figure S4, for additional details) (12,13). This arrangement allowed us to spray any melting product directly into the ion source for analysis.

Preliminary experiments were conducted on untreated dsRNA control by raising the emitter temperature from 25 to 95°C (Supporting Information, Supplementary Figure S5). The results indicated that the duplex was fully dissociated in individual strands at the higher end of the temperature range. When these conditions were applied to the reaction mixture of dsRNA with B-CeP **1**, we observed the complete dissociation of any unreacted dsRNA remaining in the sample, as well as a variety of mono-functional and intra-strand bifunctional products (Figure 6B). In contrast, a series of inter-strand conjugated products were observed in a still intact duplex form, owing to the presence of a bridging B-CeP **1** adduct that had effectively prevented strand dissociation. The mono-functional products of the RNA1 strand displayed multiple stoichiometries, consistent with the susceptibility of any of the guanines in the sequence to B-CeP **1** alkylation. In the context of the dsRNA, no equivalent mono-functional product could be detected for RNA2, thus

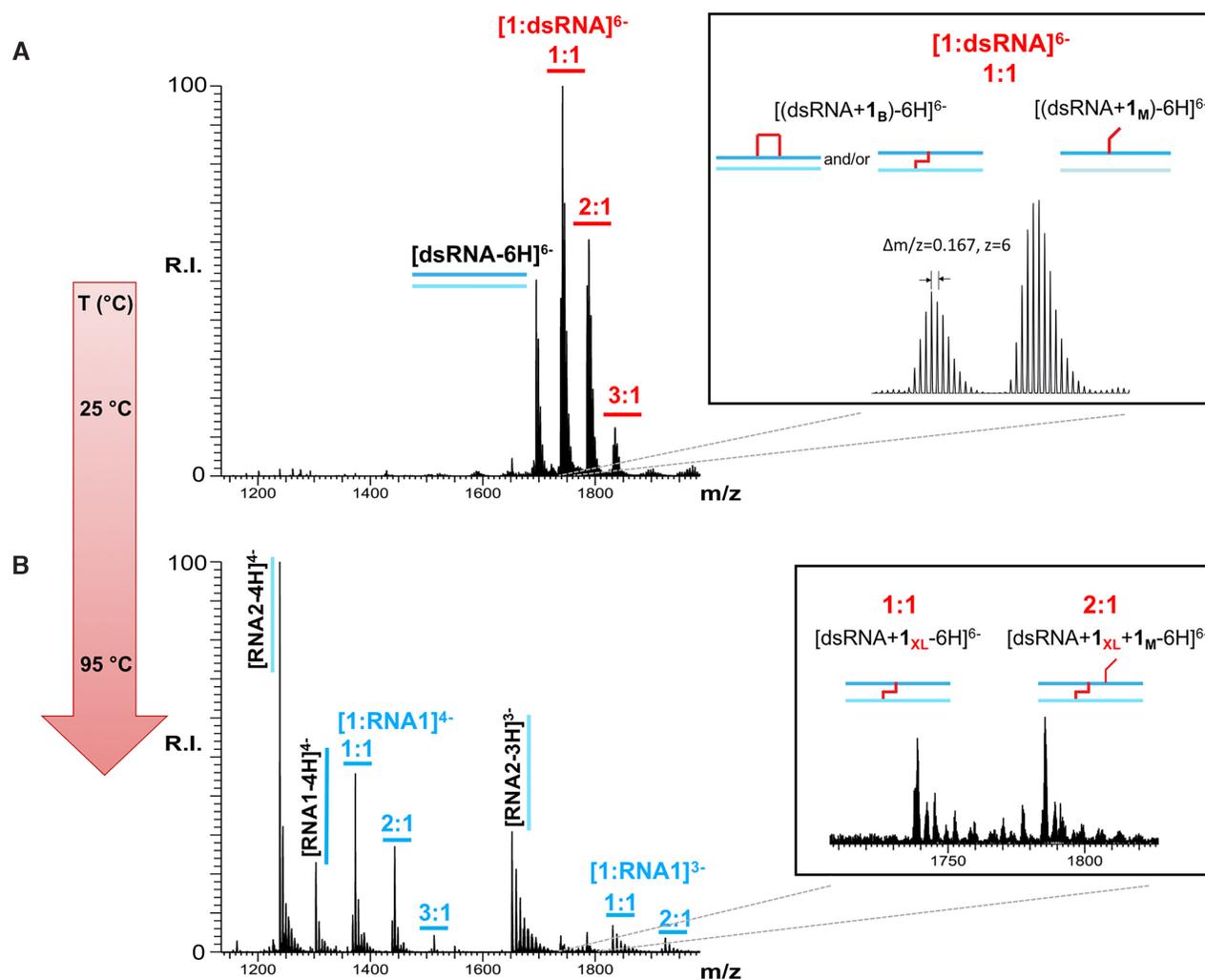


Figure 6. STHEM-ESI analysis of a sample obtained by mixture 2 μ M of double-stranded oligoribonucleotide dsRNA with 5 μ M bis-3-chloropiperidine **1** and incubating for 2 h at 37°C. (A) Representative ESI-MS spectrum obtained at 25°C. The inset shows an enlarged view of the 6- charge state of bis-alkylation dsRNA+**1**_B (intra- and/or inter-strand cross-links) and mono-alkylated dsRNA+**1**_M adducts of B-CeP **1**. Multiple stoichiometries with masses that match combinations of mono- and bifunctional adducts with up to three adduct equivalents were identified. (B) Representative ESI-MS spectrum obtained at 95°C. The inset shows an enlarged view of the 6- charge state of inter-strand cross-links (dsRNA+**1**_{XL} and dsRNA+**1**_{XL}+**1**_M), which remained unaffected by the higher temperature. Spectra were recorded in 150 mM ammonium acetate. Lower intensity signals near free/bound species are the typical sodium and ammonium adducts observed in the ESI-MS analysis of nucleic acids. A specific color is associated to each RNA substrate: red for the dsRNA, light blue for RNA2 and blue for RNA1. The ratios refer to the different stoichiometries of compound **1** to each substrate (1:1, 2:1 and 3:1 ratios of compound **1**:substrate).

indicating that the only G present in this sequence was fully involved in the inter-strand conjugate with RNA1. In fact, classic tandem mass spectrometry experiments (MS/MS) (16,22,23) performed at 25°C confirmed that the bifunctional adducts, including both intra- and inter-strand products, occurred in the dsRNA only between two Gs (see Supporting Information, Supplementary Figure S6). Unlike the dsRNA construct, the single-stranded RNA2 afforded the direct identification of both mono- and bi-functional alkylation adducts produced by compound **1** (see Supplementary Figure S7, Supporting Information). Based on the presence of a single G in position 6 of this construct, the latter could be readily attributed to bridging reactions involving the contiguous A5 and C7 nucleotides placed on either side, which would be consistent with the spacing between the reactive functions of B-CeP **1**. However, when Gs are

advantageously located in a dsRNA context to form inter-strand cross-links, they prevail over the other nucleotides to react with the probe. Moreover, it should be noted that the stability of the bond between B-CeP **1** and RNA was verified by analyzing a sample of the RNA1 reaction mixture held at 95°C, which reported no variations of the alkylation pattern. Even at the higher temperatures tested in the study, no evidence of possible cleavage of the covalent backbone was ever observed.

B-CeP compounds as probes for RNA higher-order structure

The experiments performed on both single- and double-stranded substrates afforded valuable insights into the structural determinants of B-CeP **1** activity. The feature that stood out the most was the ability of this compound to effi-

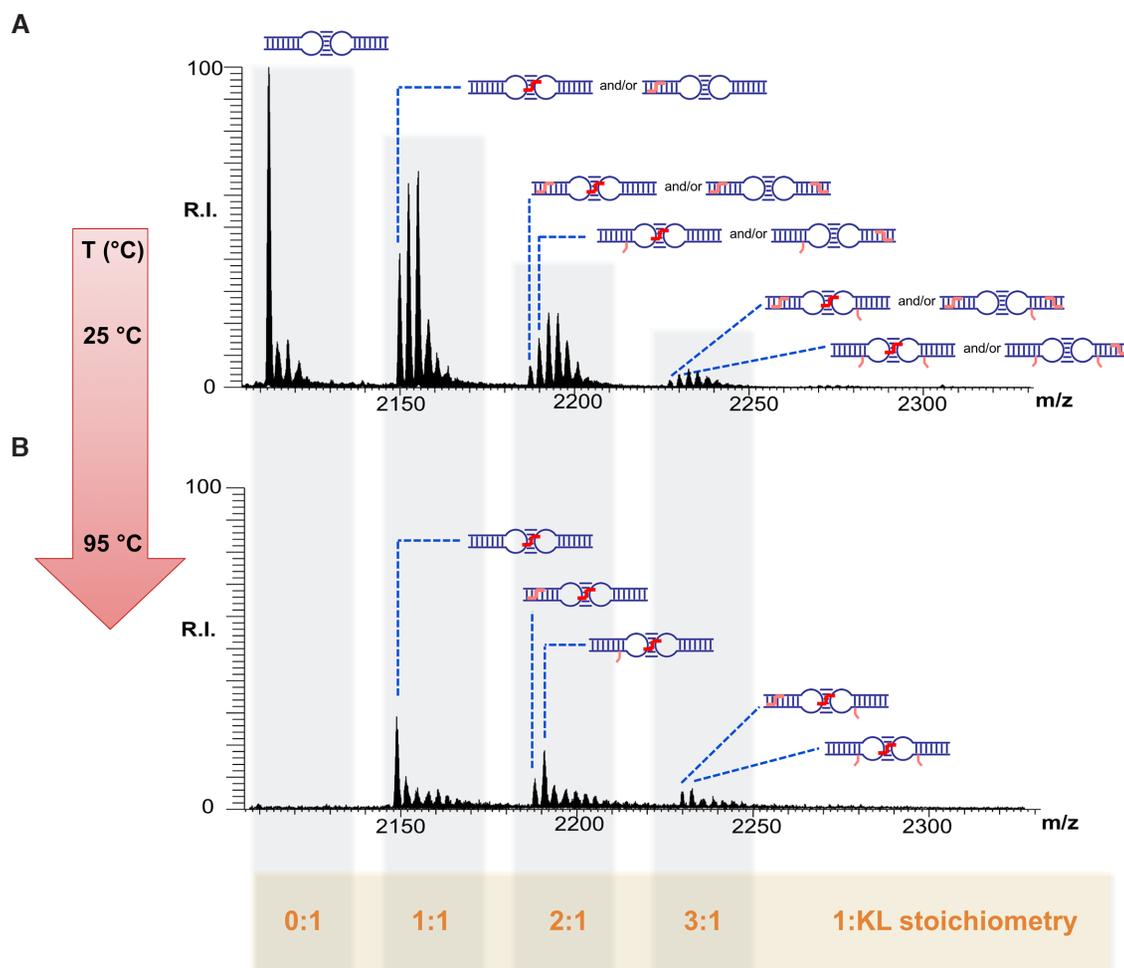


Figure 7. Thermal denaturation experiment monitored using mass spectrometry of reaction mixture obtained by incubating DIS KL (2 μM) with 5 μM bis-3-chloropiperidine **1** incubated for 2 h at 37°C. (A) Representative ESI-MS spectrum obtained at 25°C. Only the 7- charge state of unmodified and modified DIS KL complex is shown for the sake of clarity. Multiple binding stoichiometries with masses that match combinations of mono and bifunctional adducts with up to three equivalents of compound were identified and are highlighted as grey box with the indicated 1:KL stoichiometry. In the figure, a schematic representation of unmodified DIS KL and the putative B-CePs adducts are indicated. Monoalkylation and bifunctional adducts within each hairpin are in pink. Inter-strand cross-links connecting the two DIS subunits are highlighted in red. (B) Representative ESI-MS spectrum obtained at 95°C. Only the 7- charge state of unmodified and modified DIS KL complex is shown for the sake of clarity. The inter-strand cross-links induced by B-CeP **1** on the DIS KL complex (KL+1_{XL}, KL+1_{XL}+1_B, KL+1_{XL}+1_M, KL+1_{XL}+1_B+1_M, KL+1_{XL}+2_M), which remain unaffected at high temperature, are indicated. Spectra were recorded in 150 mM ammonium acetate. Lower intensity signals near free/bound species are typical sodium and ammonium adducts.

ciently alkylate not only guanines that are solvent-exposed in single-stranded substrates, but also those involved in base-pairing interactions in helical structures. Such feature could be particularly appealing in the identification of regions involved in tertiary interactions, such as those established between secondary structures folded by distal sequences of the same strand, or quaternary contacts between discrete domains present in separate strands. To test these capabilities, we selected a system mimicking the kissing loop (KL) complex formed by two copies of the dimerization initiation site (DIS) of the HIV-1 genome (24). The quaternary interaction that characterizes this system is stabilized by the annealing of complementary palindromic sequences located on the single-stranded loop of each individual hairpin (Figure 1C). Any mutation in this highly conserved region has been shown to reduce viral infectivity by disrupting

both genome dimerization and packaging *in vivo*, thus making this domain a very desirable target for the development of new antiviral strategies (24–28).

The desired KL complex employed in these experiments was prepared from a synthetic oligonucleotide by following an established heat-refolding procedure (29). Complex formation was verified by ESI-MS analysis under non-denaturing conditions (Supporting Information, Supplementary Figure S8). The probing reaction was then accomplished by adding 5 μM of B-CeP **1** to 2 μM KL RNA. STEM-ESI analysis was performed at both 25 and 95°C to enable the discrimination between intra- and intermolecular crosslinks, as described above. At 25°C, the representative spectrum in Figure 7A revealed the presence of both unreacted and alkylated species, which were assigned according to the observed masses. Among them, a series of

mono- and bi-functional alkylated products, which include both intra- and inter-strand, could be immediately recognized with a range of stoichiometries up to three equivalents of B-CeP **1**. In contrast, the higher temperature induced the melting of the quaternary interaction between hairpins. All KL species present in the spectrum (Figure 7B) contained inter-molecular cross-links that bridged across the annealed palindromic sequences and prevented complex dissociation.

A close examination of the structure of the KL construct showed that the duplex formed by the base-paired palindromes contained only one site in which two guanines (i.e., the G₁₂ of each hairpin, Figure 1C) were placed in proper position to enable inter-molecular cross-linking. In contrast, at least two different sites comprising G₂–G₂₃ and G₄–G₁₉ of each hairpin (Figure 1C) were available to support the formation of intra-molecular cross-links that bridged across the opposite strands of the stem structure. The actual presence of these intra-molecular conjugates was confirmed by the detection of corresponding signals in the low m/z range of the spectrum obtained at 95°C (Supporting Information, Supplementary Figure S9). The formation of numerous mono-functional adducts could be explained by the initial alkylation of any of the guanines in the construct, followed by failure to produce a second alkylation event because of the absence of another guanine with spatial reach, or the lost competition with the hydrolysis reaction in the aqueous solution. Indeed, each alkylation product observed in these experiments could be clearly explained on the basis of a well-defined structural context, thus indicating that B-CeP **1** could serve as a valid structural probe.

CONCLUSIONS

Bis-chloropiperidine **1** is an alkylating agent capable of reacting at relatively high rate with guanine nucleotides present in any type of nucleic acid substrate. In the case of RNA, B-CeP **1** can lead to stable mono- and bi-functional adducts without inducing the depurination and backbone cleavage observed in DNA (6–9). The results obtained from dinucleotide and oligonucleotide models provided a clear understanding of the reactivity toward RNA and enabled the characterization of typical alkylation products. The utilization of a double-stranded substrate afforded valuable insights into the structural determinants of B-CeP **1** activity, which revealed the uncommon ability to form bifunctional adducts bridging across guanines in opposite strands. This feat is typically hindered by the rather constrained steric context of helical structures and the need to match the well-defined distance between susceptible N7 positions in the opposite strands. The experiments involving the KL complex demonstrated that, aided by STEM-ESI analysis, this capability could be leveraged to identify quaternary interactions established by complementary sequences located in different structural domains. In a series of earlier reports, the merits of combining footprinting and cross-linking reagents with MS detection to probe the structure of RNA molecules have been explored (30,31). The information on solvent accessibility and spatial contiguity afforded by such reagents could be translated into valid spatial constraints to support full-fledged molecular modeling operations (32), and allowed to generate all-atom structures

for the feline immunodeficiency virus (FIV) frameshifting pseudoknot and the HIV-1 genome packaging signal (Ψ -RNA), which had not been previously solved (33,34). In this context, B-CeP **1** will become an excellent addition to the available toolkit of bifunctional crosslinkers employed for the characterization of long-range tertiary and quaternary interactions in complex RNA systems. Their observed preference for purine nucleotides could readily complement the selectivity of psoralens and related crosslinkers for pyrimidines (35,36), which has recently brought the spotlight on these types of reagents for their ability to support transcriptome-wide structural investigations (37–39). However, while the rather low yields afforded by the latter necessitate detection strategies based on strand amplification, such as Next-Generation Sequencing (40), the crosslinking efficiency afforded by the former has been shown sufficient to support the greater demands of direct MS-analysis, which must be carried out exclusively on genuine samples, not their amplification copies. Furthermore, while B-CePs conjugates were shown to be stable during the entire analytical workflow, those generated by psoralens and other photoactivated reagents tend to be prone to unwanted instability and, in some cases, can be readily reversed by UV light exposure (40). These favorable characteristics are shared with mechlorethamine and related nitrogen mustards, which are also prominent members of the crosslinking toolkit (16,19). Moreover, B-CePs produce inter-strand crosslinks only on RNA–RNA duplexes, thus making these reagents an excellent choice for applications aimed at elucidating the structure–function relationships in biologically relevant RNA systems and assessing the drug-gability of RNAs associated with diseases (1,41–43). The recent SARS-CoV-2 pandemic has keenly highlighted the relevance of RNA viruses to human health and underscored the importance of structural characterization to expedite the functional elucidation of viral genomes (4,44,45). Therefore, alternative experimental approaches based on structural probes such as B-CeP **1** will be expected to play an ever-expanding role in complementing traditional high resolution techniques.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr B. Toro for building the temperature-controlled nanospray apparatus.

FUNDING

European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie Actions [751931 to A.S.]; National Institutes of Health [R01 GM121844 to D.F.]. Funding for open access charge: A.S. personal research funds.

Conflict of interest statement. None declared.

REFERENCES

- Morgan, B.S., Forte, J.E. and Hargrove, A.E. (2018) Insights into the development of chemical probes for RNA. *Nucleic Acids Res.*, **46**, 8025–8037.
- Breaker, R.R., Conklin, D.S., Gold, L., Soll, D., Montimurro, J.S. and Agris, P.F. (2012) RNA science and its applications—a look toward the future: Albany, NY USA, November 3–4, 2011. *RNA Biol.*, **9**, 1050–1053.
- Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution—trashing old rules to forge new ones. *Cell*, **157**, 77–94.
- Rangan, R., Zheludev, I.N., Hagey, R.J., Pham, E.A., Wayment-Steele, H.K., Glenn, J.S. and Das, R. (2020) RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *RNA*, **26**, 937–959.
- Kwok, C.K., Tang, Y., Assmann, S.M. and Bevilacqua, P.C. (2015) The RNA transcriptome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem. Sci.*, **40**, 221–232.
- Sosic, A., Zuravka, I., Schmitt, N.K., Miola, A., Gottlich, R., Fabris, D. and Gatto, B. (2017) Direct and topoisomerase II mediated DNA damage by bis-3-chloropiperidines: the importance of being an earnest G. *ChemMedChem*, **12**, 1471–1479.
- Zuravka, I., Roesmann, R., Sosic, A., Gottlich, R. and Gatto, B. (2015) Bis-3-chloropiperidines containing bridging lysine linkers: influence of side chain structure on DNA alkylating activity. *Bioorg. Med. Chem.*, **23**, 1241–1250.
- Zuravka, I., Roesmann, R., Sosic, A., Wende, W., Pingoud, A., Gatto, B. and Gottlich, R. (2014) Synthesis and DNA cleavage activity of Bis-3-chloropiperidines as alkylating agents. *ChemMedChem*, **9**, 2178–2185.
- Zuravka, I., Sosic, A., Gatto, B. and Gottlich, R. (2015) Synthesis and evaluation of a bis-3-chloropiperidine derivative incorporating an anthraquinone pharmacophore. *Bioorg. Med. Chem. Lett.*, **25**, 4606–4609.
- Helbing, T., Carraro, C., Francke, A., Sosic, A., De Franco, M., Gandin, V., Gottlich, R. and Gatto, B. (2020) Aromatic linkers unleash the antiproliferative potential of 3-chloropiperidines against pancreatic cancer cells. *ChemMedChem*, **15**, 2040–2051.
- Carraro, C., Helbing, T., Francke, A., Zuravka, I., Sosic, A., De Franco, M., Gandin, V., Gatto, B. and Gottlich, D.R. (2021) Appended aromatic moieties in flexible bis-3-chloropiperidines confer tropism against pancreatic cancer cells. *ChemMedChem*, **16**, 860–868.
- Marchand, A., Rosu, F., Zenobi, R. and Gabelica, V. (2018) Thermal denaturation of DNA G-quadruplexes and their complexes with ligands: thermodynamic analysis of the multiple states revealed by mass spectrometry. *J. Am. Chem. Soc.*, **140**, 12553–12565.
- Hommersom, B., Porta, T. and Heeren, R.M.A. (2017) Ion mobility spectrometry reveals intermediate states in temperature-resolved DNA unfolding. *Int. J. Mass Spectrom.*, **419**, 52–55.
- Hagan, N.A. and Fabris, D. (2007) Dissecting the protein-RNA and RNA-RNA interactions in the nucleocapsid-mediated dimerization and isomerization of HIV-1 stemloop 1. *J. Mol. Biol.*, **365**, 396–410.
- Turner, K.B., Hagan, N.A., Kohlway, A.S. and Fabris, D. (2006) Mapping noncovalent ligand binding to stemloop domains of the HIV-1 packaging signal by tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **17**, 1402–1411.
- Zhang, Q., Yu, E.T., Kellersberger, K.A., Crosland, E. and Fabris, D. (2006) Toward building a database of bifunctional probes for the MS3D investigation of nucleic acids structures. *J. Am. Soc. Mass Spectrom.*, **17**, 1570–1581.
- Null, A.P., Nepomuceno, A.I. and Muddiman, D.C. (2003) Implications of hydrophobicity and free energy of solvation for characterization of nucleic acids by electrospray ionization mass spectrometry. *Anal. Chem.*, **75**, 1331–1339.
- Rajski, S.R. and Williams, R.M. (1998) DNA cross-linking agents as antitumor drugs. *Chem. Rev.*, **98**, 2723–2796.
- Scalabrin, M., Dixit, S.M., Makshood, M.M., Krzemien, C.E. and Fabris, D. (2018) Bifunctional cross-linking approaches for mass spectrometry-based investigation of nucleic acids and protein-nucleic acid assemblies. *Methods*, **144**, 64–78.
- Scalabrin, M., Siu, Y., Asare-Okai, P.N. and Fabris, D. (2014) Structure-specific ribonucleases for MS-based elucidation of higher-order RNA structure. *J. Am. Soc. Mass Spectrom.*, **25**, 1136–1145.
- Smith, S.I. and Brodbelt, J.S. (2010) Rapid characterization of cross-links, mono-adducts, and non-covalent binding of psoralens to deoxyoligonucleotides by LC-UV/ESI-MS and IRMPD mass spectrometry. *Analyst*, **135**, 943–952.
- Nordhoff, E., Kirpekar, F. and Roepstorff, P. (1996) Mass spectrometry of nucleic acids. *Mass Spectrom. Rev.*, **15**, 67–138.
- Quinn, R., Basanta-Sanchez, M., Rose, R.E. and Fabris, D. (2013) Direct infusion analysis of nucleotide mixtures of very similar or identical elemental composition. *J. Mass Spectrom.*, **48**, 703–712.
- Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. and Marquet, R. (1996) A loop-loop “kissing” complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 5572–5577.
- Berkhout, B. and van Wamel, J.L. (1996) Role of the DIS hairpin in replication of human immunodeficiency virus type 1. *J. Virol.*, **70**, 6723–6732.
- Clever, J.L. and Parslow, T.G. (1997) Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J. Virol.*, **71**, 3407–3414.
- Paillart, J.C., Berthou, L., Ottmann, M., Darlix, J.L., Marquet, R., Ehresmann, B. and Ehresmann, C. (1996) A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *J. Virol.*, **70**, 8348–8354.
- Shen, N., Jette, L., Liang, C., Wainberg, M.A. and Laughrea, M. (2000) Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. *J. Virol.*, **74**, 5729–5735.
- Turner, K.B., Kohlway, A.S., Hagan, N.A. and Fabris, D. (2009) Noncovalent probes for the investigation of structure and dynamics of protein-nucleic acid assemblies: the case of NC-mediated dimerization of genomic RNA in HIV-1. *Biopolymers*, **91**, 283–296.
- Yu, E. and Fabris, D. (2003) Direct probing of RNA structures and RNA-protein interactions in the HIV-1 packaging signal by chemical modification and electrospray ionization fourier transform mass spectrometry. *J. Mol. Biol.*, **330**, 211–223.
- Yu, E. and Fabris, D. (2004) Toward multiplexing the application of solvent accessibility probes for the investigation of RNA three-dimensional structures by electrospray ionization-Fourier transform mass spectrometry. *Anal. Biochem.*, **334**, 356–366.
- Yu, E.T., Hawkins, A., Eaton, J. and Fabris, D. (2008) MS3D structural elucidation of the HIV-1 packaging signal. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 12248–12253.
- Yu, E.T., Hawkins, A., Kuntz, I.D., Rahn, L.A., Rothfuss, A., Sale, K., Young, M.M., Yang, C.L., Pancerella, C.M. and Fabris, D. (2008) The collaboratory for MS3D: a new cyberinfrastructure for the structural elucidation of biological macromolecules and their assemblies using mass spectrometry-based approaches. *J. Proteome Res.*, **7**, 4848–4857.
- Yu, E.T., Zhang, Q. and Fabris, D. (2005) Untying the FIV frameshifting pseudoknot structure by MS3D. *J. Mol. Biol.*, **345**, 69–80.
- Calvet, J.P. and Pederson, T. (1979) Photochemical cross-linking of secondary structure in HeLa cell heterogeneous nuclear RNA in situ 1. *Nucleic Acids Res.*, **6**, 1993–2001.
- Garrett-Wheeler, E., Lockard, R.E. and Kumar, A. (1984) Mapping of psoralen cross-linked nucleotides in RNA. *Nucleic Acids Res.*, **12**, 3405–3423.
- Lu, Z., Gong, J. and Zhang, Q.C. (2018) PARIS: psoralen analysis of RNA interactions and structures with high Throughput and Resolution. *Methods Mol. Biol.*, **1649**, 59–84.
- Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T., Davidovich, C., Gooding, A.R., Goodrich, K.J., Mattick, J.S. et al. (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell*, **165**, 1267–1279.
- Sharma, E., Sterne-Weiler, T., O’Hanlon, D. and Blencowe, B.J. (2016) Global mapping of human RNA-RNA interactions. *Mol. Cell*, **62**, 618–626.
- Kudla, G., Wan, Y. and Helwak, A. (2020) RNA conformation capture by proximity ligation. *Annu. Rev. Genomics Hum. Genet.*, **21**, 81–100.
- Kubota, M., Tran, C. and Spitale, R.C. (2015) Progress and challenges for chemical probing of RNA structure inside living cells. *Nat. Chem. Biol.*, **11**, 933–941.

42. Disney, M.D. (2019) Targeting RNA with small molecules to capture opportunities at the intersection of chemistry, biology, and medicine. *J. Am. Chem. Soc.*, **141**, 6776–6790.
43. Velema, W.A., Park, H.S., Kadina, A., Orbai, L. and Kool, E.T. (2020) Trapping transient RNA complexes by chemically reversible acylation. *Angew. Chem. Int. Ed. Engl.*, **59**, 22017–22022.
44. Rangan, R., Zheludev, I.N. and Das, R. (2020) RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses. *RNA*, **26**, 937–959.
45. Wacker, A., Weigand, J.E., Akabayov, S.R., Altincekic, N., Bains, J.K., Banijamali, E., Binas, O., Castillo-Martinez, J., Cetiner, E., Ceylan, B. *et al.* (2020) Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Res.*, **48**, 12415–12435.