# Universal and strain specific structure features of segment 8 genomic RNA of influenza A virus—application of 4-thiouridine photocrosslinking

Marta Soszynska-Jozwiak[1], Maciej Pszczola[1], Julita Piasecka[1], Jake M. Peterson[2], Walter N. Moss[2], Katarzyna Taras-Goslinska[3], Ryszard Kierzek[1],*, and Elzbieta Kierzek[1],*

From the [1]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland; [2]Roy J. Carver Department of Biophysics, Biochemistry and Molecular Biology, Iowa State University, Ames, Iowa, USA; [3]Department of Chemistry, Adam Mickiewicz University, Poznan, Poland

Edited by Karin Musier-Forsyth

RNA structure in the influenza A virus (IAV) has been the focus of several studies that have shown connections between conserved secondary structure motifs and their biological function in the virus replication cycle. Questions have arisen on how to best recognize and understand the pandemic properties of IAV strains from an RNA perspective, but determination of the RNA secondary structure has been challenging. Herein, we used chemical mapping to determine the secondary structure of segment 8 viral RNA (vRNA) of the pandemic A/California/04/2009 (H1N1) strain of IAV. Additionally, this long, naturally occurring RNA served as a model to evaluate RNA mapping with 4-thiouridine (4sU) crosslinking. We explored 4-thiouridine as a probe of nucleotides in close proximity, through its incorporation into newly transcribed RNA and subsequent photoactivation. RNA secondary structural features both universal to type A strains and unique to the A/California/04/2009 (H1N1) strain were recognized. 4sU mapping confirmed and facilitated RNA structure prediction, according to several rules: 4sU photocross-linking forms efficiently in the double-stranded region of RNA with some flexibility, in the ends of helices, and across bulges and loops when their structural mobility is permitted. This method highlighted three-dimensional properties of segment 8 vRNA secondary structure motifs and allowed to propose several long-range three-dimensional interactions. 4sU mapping combined with chemical mapping and bioinformatic analysis could be used to enhance the RNA structure determination as well as recognition of target regions for antisense strategies or viral RNA detection.

Influenza A virus (IAV) is a fast-evolving RNA virus that causes annual epidemics and frequent pandemics (1). Five documented influenza pandemics occurred in the 20th and early 21st centuries (1889, 1918, 1957, 1968, and 2009), where the virus circulated in nonseasonal patterns from human to human, causing excess morbidity and high mortality levels (2–5). The emergence of pandemic strains has been connected to antigenic shifts, which occur when influenza strains exchange entire gene segments during coinfection of a host (6). A/California/04/2009 (H1N1) influenza virus (CalvRNA8) is one such strain and is still circulating worldwide by way of human hosts.

Great work has already been performed in the analysis of influenza RNA secondary structure. Many of influenza's viral processes have been tied to RNA structures, and key motifs are highly conserved across strains (7–11). Previously, the secondary structure of three entire viral genomic segments (8, 7, and 5) of A/Vietnam/1203/2004 (H5N1) was determined (12–15). Connections between viral RNA structure variances and their pandemic potential have been shown in several cases (16–18).

RNA structure is closely related to RNA function. Several experimental methods exist for the analysis of secondary and tertiary structure of long RNAs (19–22). Most chemical reagents show a preference for reacting with single-stranded or flexible nucleotides: dimethyl sulfide (DMS), 1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT), kethoxal, and selective 2′-hydroxyl acylation and primer extension (SHAPE) reagents (23–26). Lead ion cleavage and in-line probing can also provide information about flexible, unpaired nucleotides (27, 28). Several approaches have been developed for identification of RNA-RNA helices. SPLASH and PARIS are two relatively recent methods for identifying RNA-RNA intramolecular interactions (29). Both of these methods rely on the use of reversible UV crosslinking and psoralen, an intercalating small molecule. While these approaches provide for a very powerful analysis of RNA structure and its interactions, they do have their limitations. Psoralen has vaguely defined sequence and structural preferences that affect its affinity for intercalation and thus its crosslinking efficiency. Additionally, psoralen intercalation distorts the native helical conformation and can therefore affect RNA structure and stability. As an alternative to intercalating reagents, new methods for base pair detection (including tertiary interactions) would be beneficial. Historically, only enzymatic mapping could provide information about both paired and unpaired regions of RNA; however, enzymatic approaches (due

\* For correspondence: Elzbieta Kierzek, elzbieta.kierzek@ibch.poznan.pl; Ryszard Kierzek, rkierzek@ibch.poznan.pl.

to their large size) cannot provide single-nucleotide resolution, and the need for specific enzyme sequences and conditions limits their general applicability. Additionally, few reagents (DMS, some SHAPE reagents, and lead ion cleavage) are able to be used in cells.

Applying several methods with different reaction mechanisms allows one to better determine RNA structure, as each distinct approach can offer complementary, nonoverlapping data. These datasets could be crucially important in the design of therapeutics targeting pathogenic RNA. Therefore, it is of great importance to provide additional methods to broaden our structure determination capabilities.

4-thiouridine (4sU), in comparison to uridine, has a sulfur atom at C-4 rather than an oxygen. The base pairing properties of 4sU are the same as unmodified uridine, preserving RNA secondary structure (17). Additionally, 4sU can be incorporated into RNA during transcription (17, 30) and is not cytotoxic (31). This modification occurs naturally in some bacterial and archaeal tRNAs, as natural 4sU is proposed to be a physiological sensor for near UV radiation exposure (32, 33). 4sU can be efficiently photoactivated by wavelengths in the 330 to 366 nm range and induces photocross-linking with substrates (nucleic acids or amino acids) in close proximity (34). 4sU cross-linking is used to detect RNA–protein interactions *in vivo* by forming adducts. That feature is utilized in such methods as Photo-Activatable Ribonucleoside enhanced Cross-Linking and ImmunoPrecipitation (PAR-CLIP) (35) and viral Photo-Activatable Ribonucleoside Cross-Linking (vPAR-CL) (36). 4sU was also used as a label to track RNA from synthesis to degradation and identify their associated RNA-binding proteins (37) and RNA indicators (38, 39). Additionally, site-specific 4sU photocrosslinking in RNA was used to solve several structural problems (17, 40) such as monitoring the formation of ribozyme–substrate complexes as ribozymes proceeded along their folding pathways (41), interactions in RNA of yeast spliceosomes (17, 42), and the arrangement of the central pseudoknot region of 16S rRNA in the 30S ribosomal subunit (17, 43). Many 3D and RNA-RNA intramolecular interactions can be distinguished by site-directed or random incorporation of 4sU followed by photocrosslinking, including 16S rRNA (43, 44) and RNA components in the spliceosome (17, 42).

Although 4sU was previously used as a tool for solving very specific RNA structural problems or RNA–RNA, RNA–

protein interactions it had not been utilized for the global structural mapping of RNA. 4sU photocrosslinking of RNA possesses several useful features. 4sU is able to form photo-adducts with both pyrimidines and purines, reducing sequence dependency in data (Fig. 1). However, the reactivity with each base is not equal (in decreasing order: U, A, G, and C) (45). Previous studies have shown that 4sU crosslinks formed efficiently at double-stranded regions of RNA with some flexibility, in the ends of helices, and across bulges and loops when their structural mobility is allowed (45, 46). Additionally, long-range 3D interactions have been recognized (17).

This paper presents the secondary structure of CalvRNA8, which is proposed based on classical chemical mapping, bioinformatics analysis, and the new application of 4sU photocrosslinking. Comparison of standard mapping methods with the 4sU method allows us to validate the new 4sU mapping method and highlight several tertiary features of CalvRNA8. Further, we define features universal for type A strains, as well as those unique to CalvRNA8.

Herein, 4sU crosslinking data verified and greatly enhanced the secondary structure prediction of CalvRNA8 based on chemical mapping data. It is important to note that the 4sU method can be used both *in vitro* and *in cellulo*, and that high crosslinking efficiency allows for usage of low-cost detection methods such as capillary electrophoresis. In this work, we introduce quantification and clear application of obtained data in the facilitation of RNA structure determination.

## Results and discussion

### Secondary structure of segment 8 genomic RNA of A/California/04/2009

The secondary structure of genomic segment 8 RNA of A/California/04/2009 (CalvRNA8) was determined using chemical mapping. CalvRNA8 was obtained from *in vitro* transcription, and the RNA folding was optimized to obtain single-molecule folding to avoid homodimer formation (Fig. S1 in Supporting Information 3). Also, refolding of RNA prior to secondary structure determination is a confirmed way to obtain native folding and removes unnatural base pairings formed during preparation and storage. Different folding buffers with range of natural concentration of ions and native pH (50 mM HEPES, pH 7.5, 100–300 mM NaCl, and
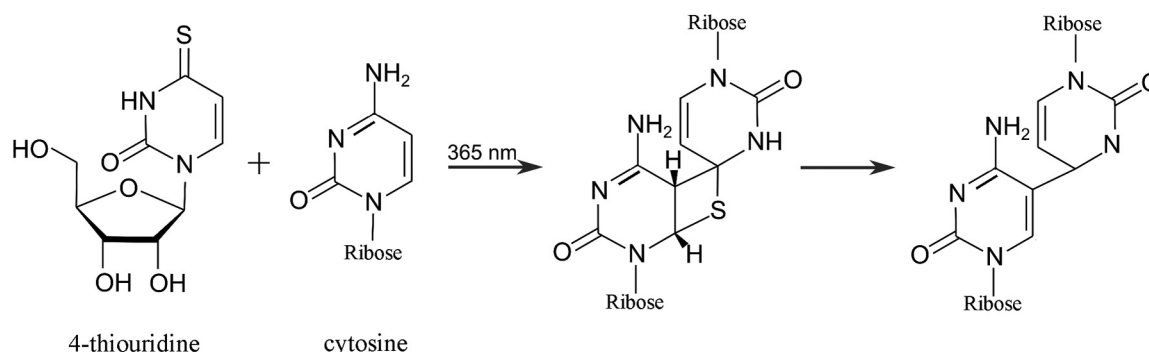


**Figure 1. Reaction between 4-thiouridine (4sU) and cytosine during irradiation, forming a photoadduct.**

5–20 mM MgCl$_2$) were tested. Finally, buffer containing 300 mM NaCl, 5 mM MgCl$_2$, 50 mM HEPES, pH 7.5 was selected in which CalvRNA8 forms a single-molecule structure (which was true for other conditions, but with no dimer occurrence). Folded CalvRNA8 was mapped with canonical chemical reagents—DMS (methylates N1 of A and N3 of C when unpaired), CMCT (modifies N3 of U and N1 of G when unpaired), and NMIA (SHAPE method) (23, 47, 48). Mapping was carried out at 37 °C. The mapping data were analyzed by reverse transcription followed by capillary electrophoresis.

DMS strongly modified 76 nucleotides (nts) and moderately modified 38 nts, which is 27% of all A and C within CalvRNA8. CMCT modified 30% of all U, especially the highly reactive regions 202 to 210, 270 to 279, 333 to 342, and 733 to 746. NMIA strongly modified 91 nts and moderately modified 93 nts, representing 21% of the CalvRNA8 segment. Reactivities from SHAPE correlate with classical chemical mapping results. Several CalvRNA8 regions (220–249, 270–279, 300–321, 384–401, 509–539, 732–763) characterized higher than average reactivities, whereas regions 550 to 600 and 358 to 383 lower than average.

Using these chemical mapping results, RNAstructure 6.0.1 (49) was used to determine the secondary structure of CalvRNA8 (Fig. 2, Supporting Information 1). Also, selected conserved base pairs for influenza virus type A were included in prediction according to previous analysis of all available sequences of segment 8 vRNA (vRNA8) (12) and secondary structure of strain A/Vietnam/1203/2004 (H5N1) (VietvRNA8) (Experimental procedures). The chosen conserved base pairs correlate with chemical mapping results, with conservation greater than 95% (Fig. 2).

CalvRNA8 is strongly structured *in vitro*, with 58% of nucleotides forming base pairs across four domains (Fig. 2). Domain I (1–28/890–850) (the designation means that region 1–20 nt is base-paired with region 850–890 nt) contains the panhandle motif, a well-known conserved influenza motif that takes part in viral propagation (18, 50). Domain II (29–168) is relatively unremarkable, while domains III (169–468) and IV (469–849) are the most reactive, containing hairpins and internal loops. SHAPE and chemical mapping show ten relatively large flexible regions across these domains (171–176, 244–249, 272–277, 333–342, 384–389, 509–515, 425–434, 530–535, 655–660, 732–741), each with at least six reactive nucleotides.

Additionally, the ProbKnot algorithm (51) of RNAstructure was used to predict possible pseudoknots within CalvRNA8. The partition function mode was used, in which all experimental constraints were implemented. This resulted in one predicted pseudoknot region (500–547) (Fig. S2 in Supporting Information 3), which is in agreement with chemical mapping data and 4sU crosslinking. In the secondary structure model of CalvRNA8, a hairpin is formed in this region. It is possible that either the hairpin or pseudoknot is formed depending on the
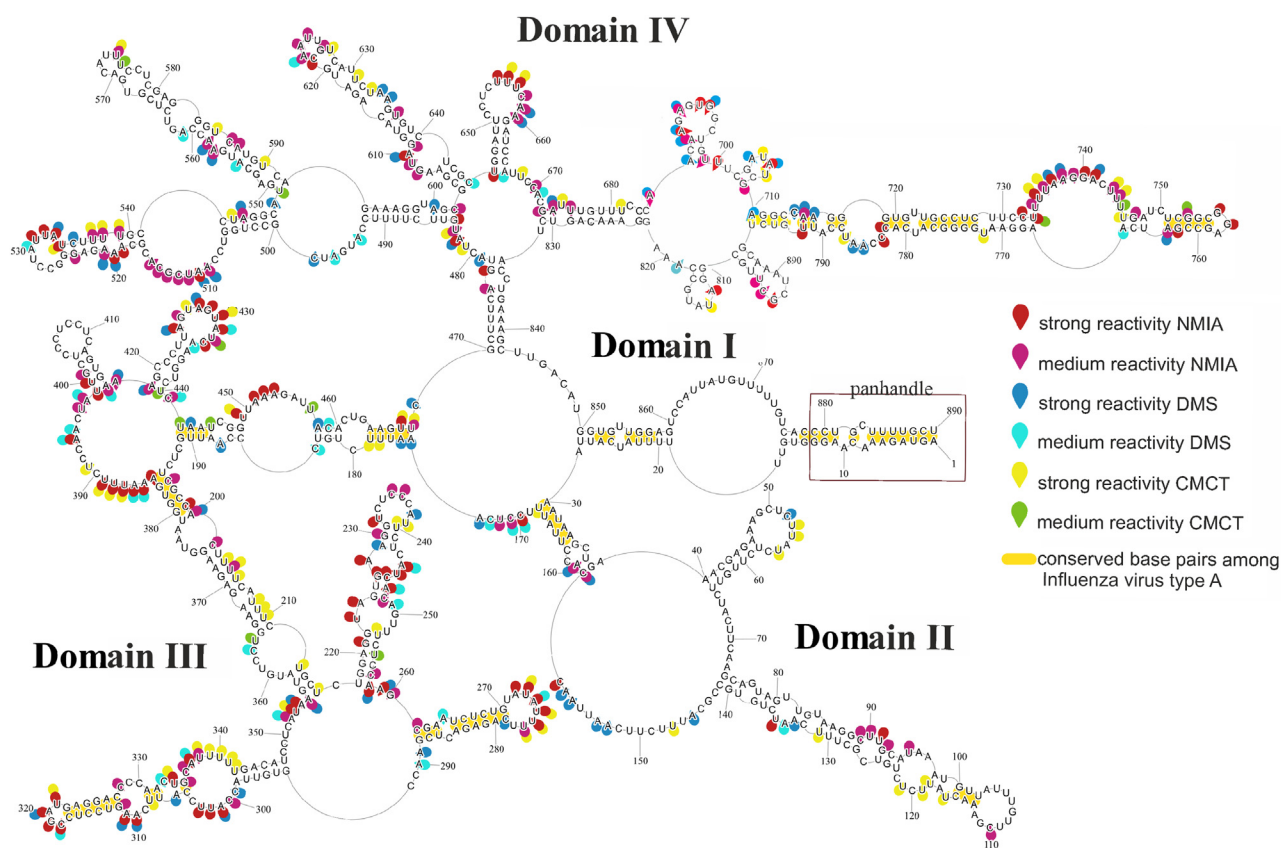


**Figure 2. The secondary structure of influenza virus segment 8 vRNA for A/California/04/2009 (H1N1) (CalvRNA8) as predicted by RNAstructure 6.0.1 using experimental chemical mapping data as constraints.** Strong DMS, CMCT, as well as SHAPE reactivities converted to pseudo-free energies were used. Additionally, medium reactivity DMS and CMCT are annotated. *Yellow lines* indicated conserved base pairs among influenza virus type A used also as restrains. The numbering of CalvRNA8 is from its 5′ end.

specific stage of replication. This situation is common in influenza virus in order to regulation of viral replication (12). However, this pseudoknot region does appear to be strain-specific; it cannot be created in A/Vietnam/1203/2004 and A/WSN/1933 (H1N1) due to sequence differences that prevent the formation of crucial base pairs.

### Base pairing probabilities for segment 8 genomic RNA of A/California/04/2009

RNAstructure 6.0.1 (49) was used to calculate base pairing probabilities for the secondary structure of CalvRNA8 (Fig. 3). This calculation indicates the accuracy of a proposed model. For the partition function calculation, experimental data were included. Results indicate that the motifs with highest probable accuracy are predicted in domains I, II, and IV. There are several regions with secondary structures (certain base pairs and unpaired nucleotides) with a greater than 90% probability: 1–76/876–890, 74–141, 646–665, 710–796. Domain III characterized fewer base pair probabilities than the other domains.

### Base pair conservation across type A influenza

Base pair conservation was calculated comparing our predicted CalvRNA8 secondary structure to a database of influenza sequences (Supporting Information 2). This conservation estimates the universality of the model across type A influenza. In calculations, all available (34,248) IAV sequences were

downloaded from the NCBI Influenza Virus Database. On average, canonical base pairing in CalvRNA8 is 89% conserved. The highest levels of conservation are seen in the panhandle region (1–16/890–876, 98.2%) and in several hairpins (82–133, 89.6%; 217–258, 94.2%; 261–288, 97.2%; 293–347, 97.2%) (Fig. 4).

### Covariance analysis

The cm-builder script (52) with the RNAFramework (53) toolkit (which automates steps in covariance analysis with the Infernal (54) suite and R-Scape (55)) was used to analyze covariance within the whole segment secondary structure, as well as the predicted tertiary interactions. Structural motifs were assembled into dot bracket notation and analyzed against the same NCBI Influenza Virus Database used for base pair conservation. Resulting data showed no statistically significant covariation within our model. This result is unsurprising, however, as prior research (56) has shown a lack of covariance across influenza subtypes due to the sheer genetic diversity within the species. Covariance data is available as part of the supporting information (Supporting Information 4).

### Mapping of segment 8 genomic RNA of A/California/04/2009 using photocrosslinking of 4-thiouridine

Application of 4sU for structure mapping requires random incorporation of 4sU in transcribed RNA. CalvRNA8 with 4sU
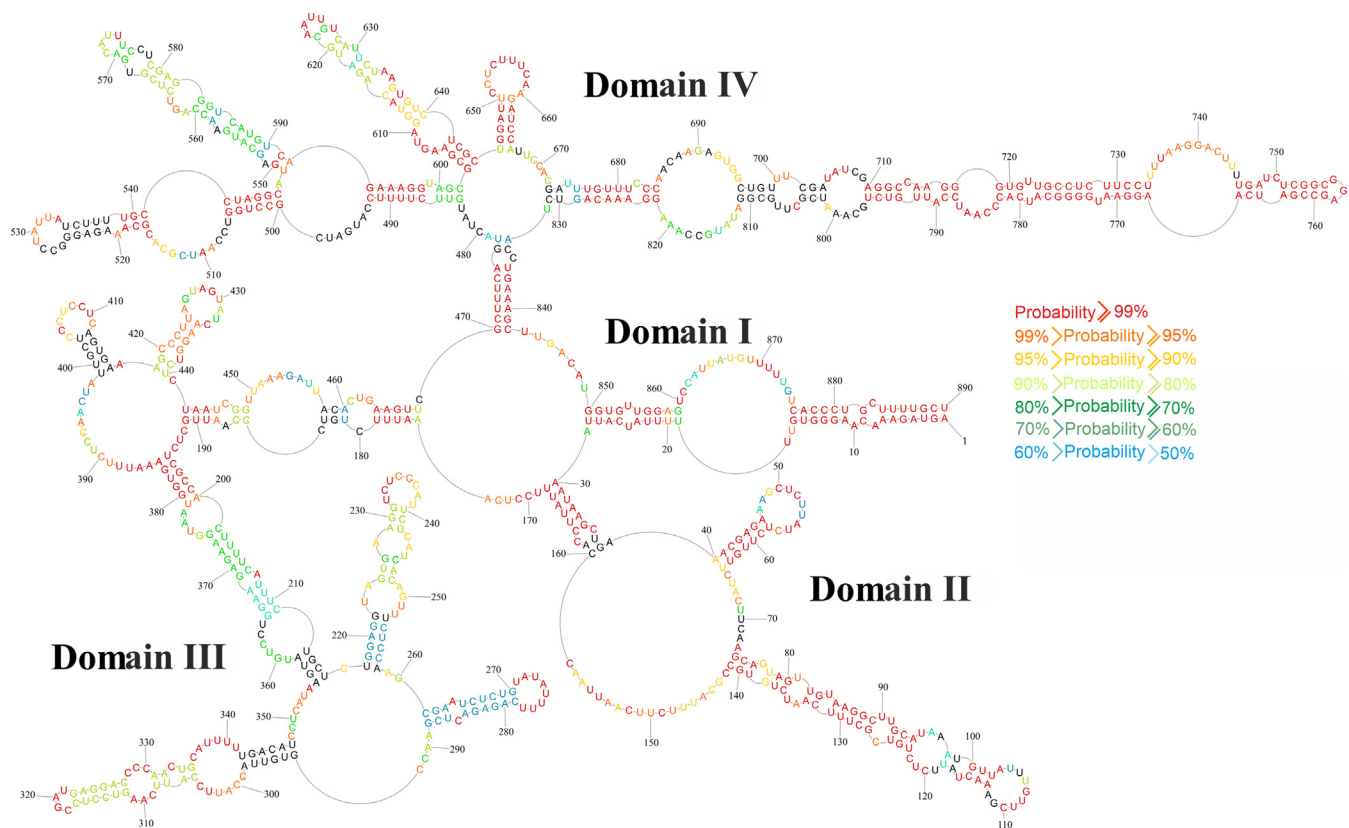


**Figure 3. Predicted probability of nucleotides being paired or single stranded in segment 8 vRNA A/California/04/2009 as calculated by RNAstructure.** Probabilities lower than 50% are in *black text*. The partition function calculation incorporated constraints from strong reactivity of DMS and CMCT and converted pseudo-energy reactivities from SHAPE.
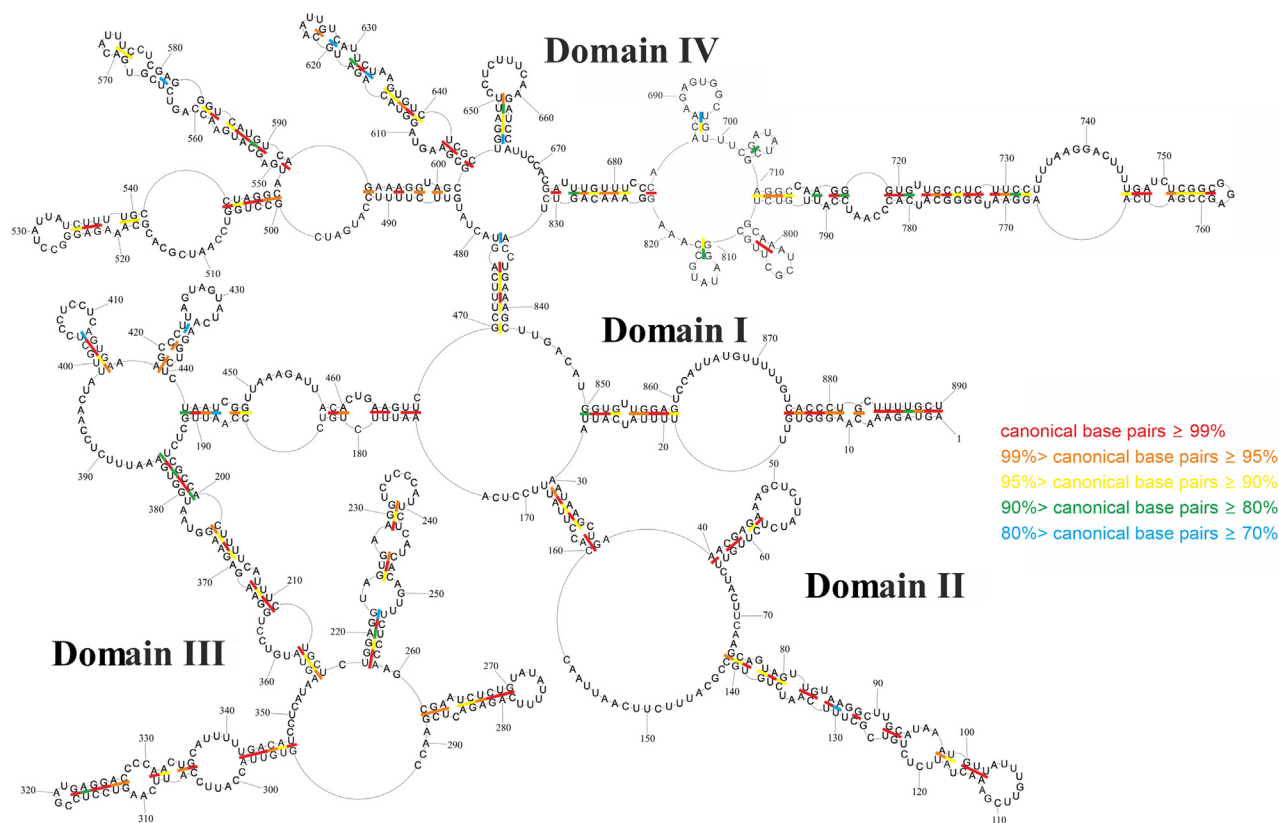
**Figure 4. Conservation of segment 8 vRNA A/California/04/2009 secondary structure in type A viruses.** *Colors* indicate the percentage of canonical base pairings preserved across type A influenza strains for segment 8 vRNA.

was obtained from transcription *in vitro* using 4-thiouridine-5′-triphosphate. 4sU was incorporated with a stoichiometry of one modification per one RNA. Next, CalvRNA8 with 4sU was folded as described in Experimental procedures and irradiated at 365 nm. Different times of irradiation were tested: 10 s, 15 s, 30 s, 15 min, 30 min, and 45 min 10 s was selected as the optimal time for all experiments. Sites of photocrosslinking of CalvRNA8 were analyzed by reverse transcription followed by capillary electrophoresis. Rules to recognize the strength of hits (strong, medium, and weak) were established (Experimental procedures). Folded CalvRNA8 with 4sU that was not irradiated served as control. Interpreting reverse transcriptase termination events from 4sU-X conjugates allows for recognition of crosslinking sites. 80 U, 6 A, 6 G, and 3 C were identified as photoadducts in the modified CalvRNA8. This is consistent with the known preference of 4-thiouridine to form photoadducts (45).

Identified photocrosslinks correlate with chemical mapping data and the proposed secondary structure model of CalvRNA8. Photoadducts were observed at or adjacent to GU pairs and at the ends of helices. In terms of the helical occurrences, most hits were detected at the first or second Watson–Crick base pair of a proposed helix. There were only two exceptions (U32, U165 in a 6-bp helix with five A-U pairs) that are thermodynamically weak and probably unstable. All these hits could be related to the flexibility of local secondary structure motifs, such as hairpins with internal loops and bulges. There are also observed 4sU hits opposite of hairpin

regions (marking local interactions such as hairpins: 549–593, 603–644). There are hairpins with crosslinking sites in hairpin loops and one side of stems at U and C (*e.g.*, 517–541, 40–64) that could result from the preference of 4sU cross-links to U and C. Additionally, 4sU at certain sites could (and usually does) react with several nucleotides in close proximity, which in turn gives higher stop signals from U in comparisons to other nucleotides (A, C, G). Crosslinking confirmed the modeled large hairpin loops if the distance in the primary sequence allowed for conjugation reactions. There was also crosslinking in single-stranded junction sequences between local motifs whose occurrence could be explained by the 3D folding of the local domains (*e.g.*, domain I), as often these hits are in "blocks" of several nucleotides. Interestingly, there are secondary structure motifs without any 4sU crosslinking signals: helices 710–733/782–767 and 261–270/288–279, and hairpin 305–335. These regions could be highly structured and rigid, which would make 4sU crosslinking impossible. We established the following criteria to verify the RNA structure model *via* 4sU mapping and deduce if the crosslinking hits come from a local structural motif or a long-distance interaction (*e.g.*, between two motifs placed in sequence distance from each other): (1) For a local motif, the hit should be in the terminal or penultimate base pair of a helix, at or adjacent to a GU pair, in a hairpin loop, in a bulge, or in an internal loop. (2) Verifying long-distance crosslinking in single-stranded regions of predicted 2D RNA structure requires further analysis. First, find accessible complementary regions using bioinformatics

and manual validation; the proposed interaction should be possible between regions. Second, 4sU crosslinking needs to be detected at or near predicted base pairings to confirm that both single-stranded regions are located in near proximity (45) and can create interactions. It is assumed that proposed long-range interactions are flexible and rather weak, so the regions may be partially accessible to chemicals. It needs to be noted that one 4sU can crosslink with several partners, all possibilities of which we see in our data (45).

One of the above two requirements, either criteria 1 and 2, must be met for all 4sU hits to positively verify the RNA secondary structure model (Fig. S5 in Supporting Information 3). In the above long-distance interaction criteria, we narrowed the possibilities to thermodynamically stable classic Watson–Crick and G-U base pairs. In this way, the most possible long-distance base pairing interactions were proposed and confirmed by both complementary of several nucleotides and presence of 4sU crosslinking. In the long distance interactions, following the general rules of 4sU crosslinking, the close proximity and flexible structure of proposed interacting regions are assumed for efficient crosslinking.

Following the criteria described above, several long-distance interactions were considered. We present the most probable and stable long-distance interactions according to our data, which sometimes involve the unfolding of other interactions. Additionally, these interactions were analyzed with bioinformatic analysis *via* base pair conservation. Potential long-range interactions longer than 3 nts were considered, with putative long-range interactions shown in Figures 5 and 6. These interactions are proposed within specific domains, such as several hairpin loop–internal loop interactions (272–275/361–358; 360–365/432–427; 605–612/658–651), as well as a proposed interaction between two internal loops (384–395/457–449 or 383–395/745–733). A hairpin loop (46–50) and single-stranded junction region (148–143) are also proposed to interact.

Interactions are also possible between domains. For example, a hairpin loop (426–432) interacts with a junction region (483–477). Two internal loops interactions are proposed (384–395/745–734; 383–395/745–733), as well as two examples of internal loop–hairpin loop interactions (384–395/457–449; 374–381/528–535).

The probability of the proposed base pairings is supported by conservation across type A influenza, with ranges from 87.5% (427–432/365–360) to 69.1% (384–395/745–734) with some pairs exceeding 90% conservation (Supporting Information 2).

Multiple loops and internal loops take part in tertiary interactions, making CalvRNA8 highly folded. 4sU mapping analysis also shows that for two cases one region can interrelate with two other regions. Likely, these tertiary interactions are flexible and rather weak, highlighting the universality of type A interactions. However, it cannot be excluded that less conserved or additional interactions occur that rule the tertiary folding of CalvRNA8 and are strain-specific.
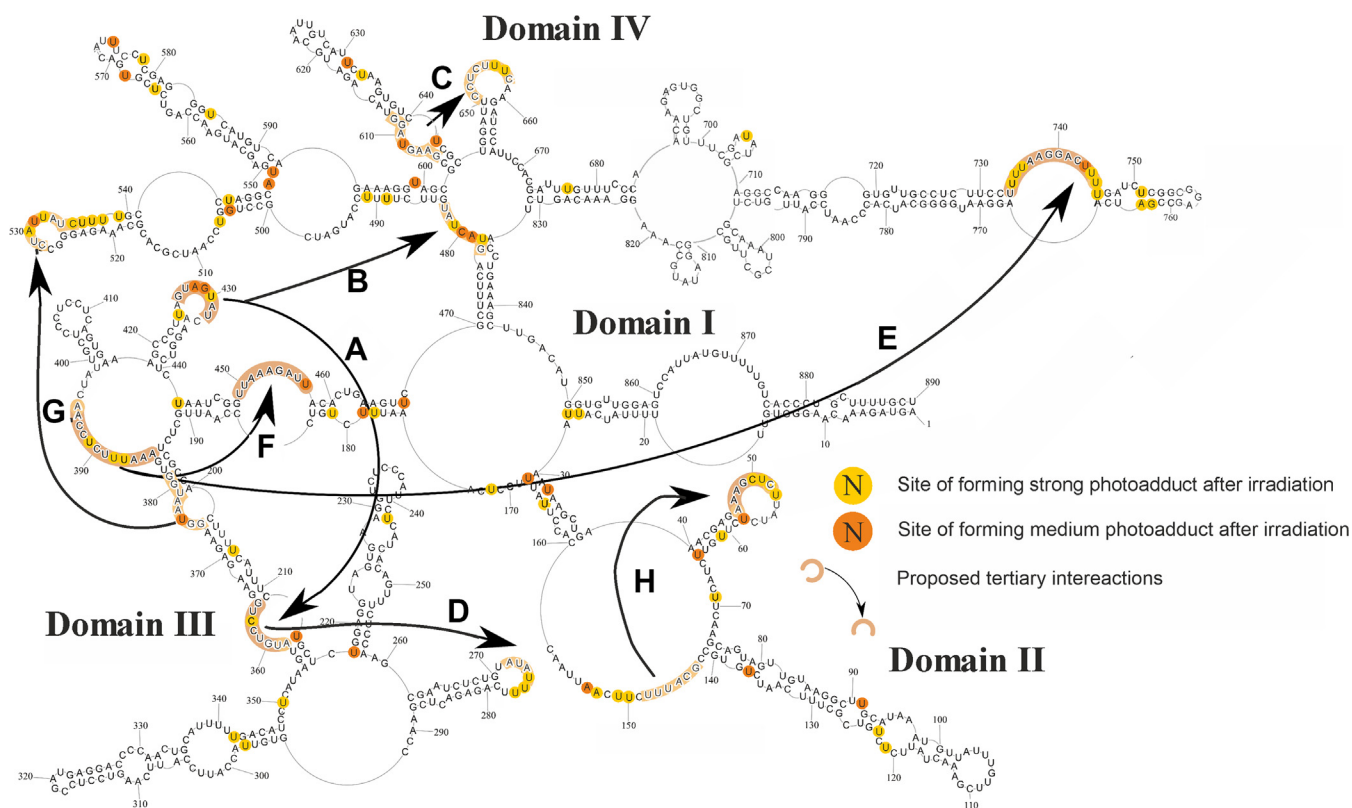


**Figure 5. Proposed long-range interactions in segment 8 genomic RNA of A/California/04/2009.** All likely tertiary interactions longer than three base pairs were considered. *Arrows* marked with *letters* indicated interaction between regions highlighted in *dark orange*.
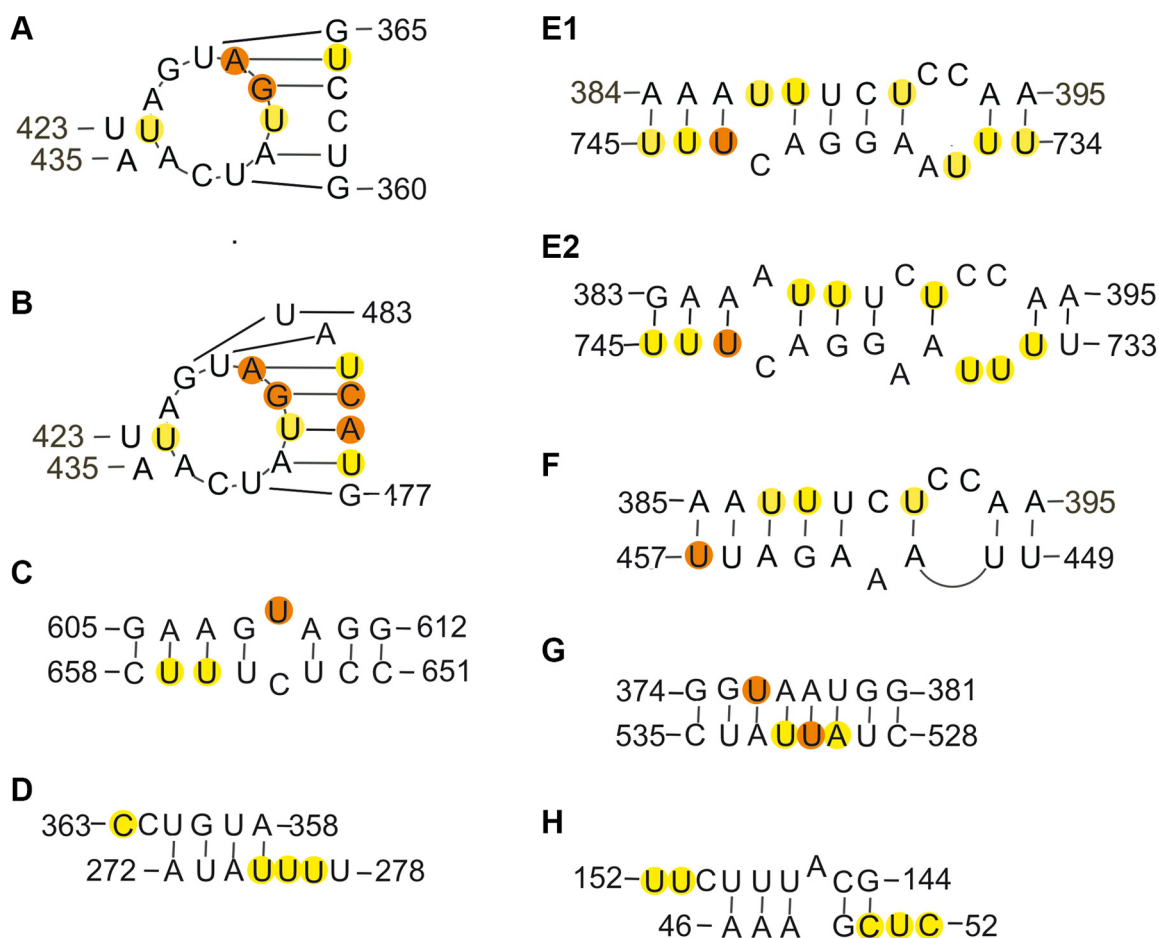
**Figure 6. Proposed long-range interactions in segment 8 vRNA of A/California/04/2009.** Panels *A-H* show base pairing in each proposed long-range interaction. *Numbers* identify the RNA nucleotide sequence from the 5′ site. *Yellow* and *orange circles* indicated detected sites of strong or medium 4sU photocrosslinking after irradiation, respectively.

Interestingly, some regions were modified strongly by both CMCT and 4sU, such as between 384–395 and 734–745. It is possible because the 3D interaction between these regions contains internal bulge, GU pair as well as base pairs adjacent to GU that are easily mapped by chemical reagents (Fig. 6). Also, 4sU crosslinking is known to occur in structurally flexible regions like these (45).

Comparison of our model to chemical mapping data shows full agreement in the long-distance interactions 360–365/432–427 and 383–395/745–733. This indicates that they are preferred to possible interactions 426–432/477–483 and 384–395/734–745 or 385–395/449–457, respectively (Fig. 6). Moreover, the moderate NMIA hits are evidence of structural dynamics that support our assumption that the long-range interactions are flexible (Fig. S4 in Supporting Information 3).

We also investigated if any 4sU long-range interactions were preserved across different strains. Published data from *in virio* psoralen studies of vRNA showed that the pattern of intermolecular interactions is flexible and different in virion of different strains (57). Also, "hot spots" of interactions (regions involved in multiple bindings) are recognized, from which few are conserved for different strains (58). An important region in vRNA8 of A/WSN/1933 (H1N1) is 605–780 nt, which is

involved in the binding of vRNA1, vRNA3, and vRNA7. Further involvement of the 500–550 and 800–850 nt regions of vRNA8 in binding vRNA5 has also been reported (58) for the same strain. Our CalvRNA8 model forms a hairpin region at 500–548 containing a helix of high BP probability (Fig. 3) and high conservation among type A influenza (87.3%) (Fig. 4). Interestingly, regions 605–780 and 800–850 belong to domain IV, which is present in both VietvRNA8 and CalvRNA8. The CalvRNA8 domain IV is characterized by high BP probability of helices. Additionally, two hairpins (603–644 and 719–782) are highly structurally conserved for IAV (86.5% and 93.6% base pairs conservation, respectively). Based on these results, it appears that domain IV and its putative motifs are responsible for intersegmental interactions, potentially during viral packaging.

Dadonaite *et al.* (57) detected a general pattern of long-range interactions in the vRNA of A/WSN/1933 (H1N1). The data showed interactions between three distinct vRNA8 regions 220–400, 460–620, and 700–800. While the pattern of long-distance base pairing for CalvRNA8 (Fig. 5) is unlikely to perfectly match A/WSN/1933, several interactions were revealed by s4U mapping between the abovementioned regions: 384–395/745–734 (Fig. 6E), 384–395/457–449 (Fig. 6F), and 374–381/535–528 (Fig. 6G). One interaction falls within

domain III and two fall between domains III and IV. The highest base pair conservations calculated for these interactions are 73.0% for 374– 381/535–528 and 76.7% for 384–395/457–449. These interactions are not among those highlighted for A/WSN/1933; although they are sequentially possible, they would be weaker due to two sequence mismatches. Despite sharing some general similarity, the long-range interactions are predicted to be different between these two strains. Intersegmental interactions could be overshadowing the vRNA8 structure in A/WSN/1933, resulting in the observed structural differences. Also, there may be flexibility in the interaction between domain III and IV during virion assembly, allowing domain IV to engage in intersegmental bindings.

We have shown that 4sU mapping is a valuable method that, when coupled to chemical mapping and structure conservation bioinformatics analysis, gives information about the flexibility of local motifs and 3D folding properties with possible long-distance interactions.

### Antiviral effects of antisense oligonucleotides support the RNA structure model

A previous study involving antisense oligonucleotides (ASO) complementary to CalvRNA8 showed significant inhibition of A/California/04/2009 replication in Madin–Darby canine kidney (MDCK) cells (59). The most effective ASOs targeted single-stranded (63–73, 861–874) and partially single-stranded (160–176, 181–194, 404–412) regions of determined herein secondary structure that supports our new model. These five ASOs inhibited viral propagation 5- to 25-fold, with the highest inhibition caused by ASOs 187-14L and 404-14L targeting the last listed regions. In terms of their accessibility, the structure of the five targets did not significantly change when compared with a previous *in silico* vRNA8 model. It is possible that these motifs exist *in vivo* and perform important functions. Interestingly, the most effective ASOs target domains correlated to s4U mapped long-range interactions. The disruption of the domain's structural organization seems to be crucial for significant inhibitory effects.

Recently, a study has been published investigating RNA secondary structures in influenza using SHAPE-MaP coupled to NGS (57). The structure analysis was carried out on A/WSN/1933 (H1N1) vRNAs associated with proteins in vRNP (*in virio*), purified from deproteinated virus particles (*ex virio*), and transcribed (*in vitro*). The analysis was focused on local secondary structure motifs with a maximum pairing distance limited to 150 nucleotides. Results from the *in virio* experiments show that, within vRNP complexes, vRNAs are able to form a number of stable secondary structure motifs. Structural analyses of vRNA8 revealed nine secondary structure motifs with high base-pairing probability exceeding 80%. Four of them are also present in the CalvRNA8 secondary structure model presented herein. Common motifs are formed in regions: 261–263/288–286, 265–284, 293–298/37–342, 312–317/327–322, 751–755/763–759, and 797–799/809–807.

These strains possess 85% sequence identity. Interestingly, first two motifs were also present in the previously published *in vitro* secondary structure model of VietvRNA8, an avian-origin strain (12). This similarity is shown in Figure 7. In the same paper, there was also *in silico* analysis of vRNA8 folding carried out on over 14,000 sequences obtained from the NCBI Influenza Virus Database. This study revealed that the abovementioned regions—forming common secondary structure motifs—encompass stems at 261– 270/277–288 and 312–317/322–327, with base pair conservation of 97.2% and 94.9%, respectively (12). Further analyses on A/WSN/1933 (H1N1) vRNA8 showed that in the absence of proteins (*ex virio* and *in vitro* conditions), RNA is more structured (57). These predictions revealed other structural similarities between strains. A/WSN/1933 (H1N1) vRNA8 is characterized by a high probability of base pairing (>80%) in the regions 611–615/640–636, 647–650/664–661, and 710–733/796–767, which are also present in the CalvRNA8 secondary structure model determined by us (Fig. 7). These findings are meaningful, as our studies were performed in a protein-free environment where the structural constraints of RNA are strongly manifested. Overall, the presence of conserved secondary structure motifs indicates their important role in the influenza replication cycle and suggests that RNA conformation is preserved for certain functions. These RNA regions may also serve as universal targets for anti-influenza strategies.

### Conservation analysis of model structure

The determined secondary structure of CalvRNA8 was compared with a model of segment 8 genomic RNA A/Vietnam/1203/2004 (H5N1) (VietvRNA8) that was published in 2016 (12). The sequence identity between CalvRNA8 and VietvRNA8 is 83% (Fig. S6 in Supporting Information 3), allowing for analysis of both inclusive and strain-specific folding. Sites of CMCT, DMS, or NMIA modifications of CalvRNA8 mainly correlate with those found to be reactive in homologous regions of VietvRNA8.

Many RNA structural motifs are the same between VietvRNA8 and CalvRNA8. Several examples include: 717–789, 1–28/850–890, 261–288, 293–347, 421–424/437–434, 175– 179/468–464, 696–704/811–802, 710–713/796–793, and 98–102/122–113 (Fig. 7). There are also regions that fold differently such as base pairs shifts and different folding patterns. These examples include: 484–602, 217–258, 201–215/374–365, and 40–64. CalvRNA8 shows an overall 89% base pair conservation, which is higher than that of VietvRNA8 secondary structure (82.6%). Several helical motifs are conserved across the models. For example, hairpin 217–258 (94.2%), helix 201–215/374–355 (88.6%), and motif 500–547 (87.3%) (Fig. 3). Additionally, numerous base pairings and motifs (261–262/288–286, 265–284, 293–298/347–342, 312–327, 710–713/796–793, 717–733/789–767, 751–763) are conserved across CalvRNA8, VietvRNA8, and A/WSN/1933 (H1N1) vRNA8 (Fig. 7). Data and modeling also show 5′ and 3′ interactions forming a *panhandle* structure in all compared vRNA8 strains (11, 57).
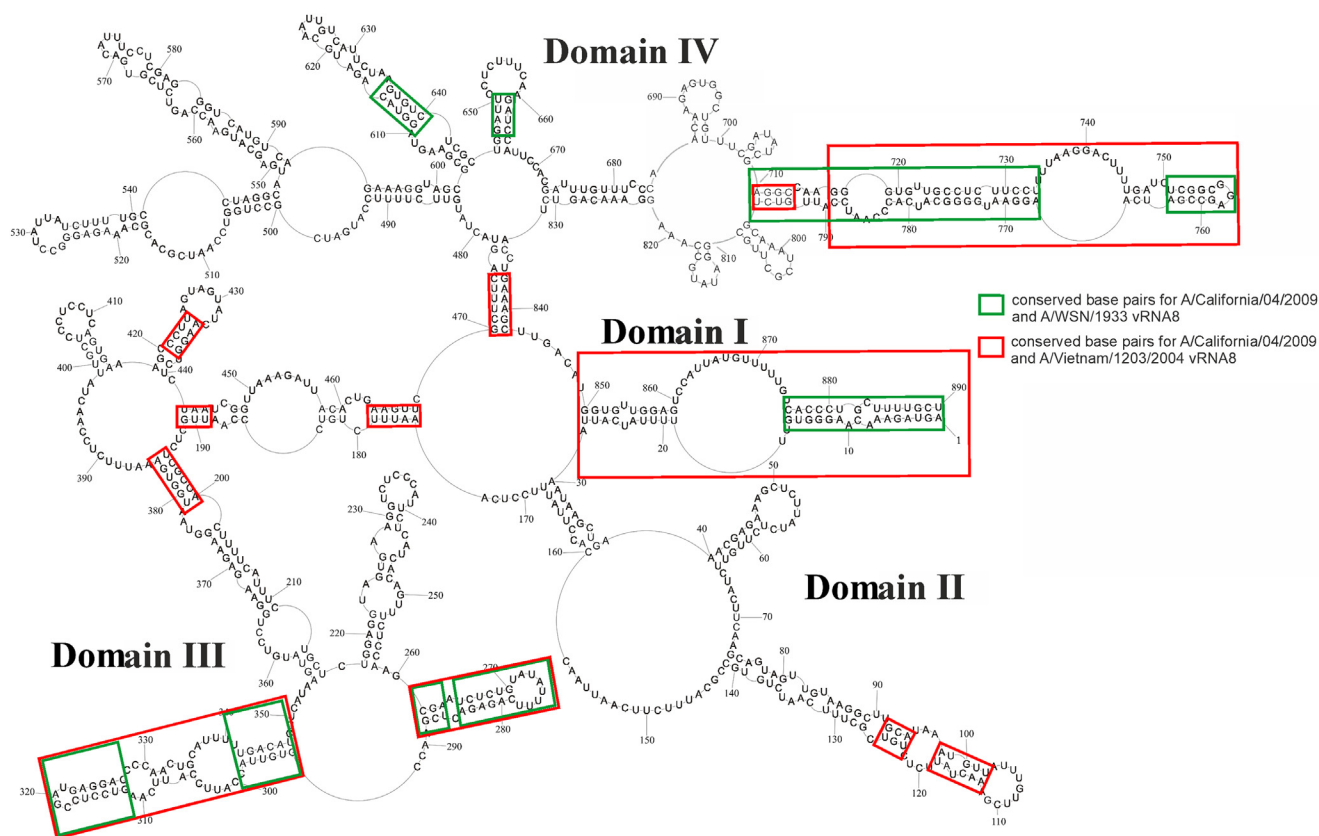
**Figure 7. Comparison of segment 8 vRNA secondary structures of A/California/04/2009, A/Vietnam/1203/2004 (H5N1) and A/WSN/1933 (H1N1).** *Red rectangles* indicate the conserved base pairs for A/California/04/2009 and A/Vietnam/1203/2004 (H5N1) vRNA8. *Green rectangles* indicate conserved base pairs for A/California/04/2009 and A/WSN/1933 (H1N1) vRNA8.

Analysis of VietvRNA8 secondary structure led to the determination of two possible long-range interactions (400–404/431–428, and 426–430/852–848). It should be noted that VietvRNA8 and CalvRNA8 are nearly identical in sequence, with CalvRNA8 having a 15 nucleotide insert between VietvRNA8's nucleotide 614 and 615. Both interactions are possible in CalvRNA8 when taking into account the sequence conservation between these two strains, chemical mapping profile, and crosslinking data on only one side (for that reason they are not marked in Figs. 5 and 6). A third interaction (227–231/409–405) through changes in sequence would be weaker in CalvRNA8, as it would contain a mismatch. However, there is no confirmation of 4sU crosslinking data and mapping data to prove this interaction.

### Conclusion

Secondary structure and tertiary interactions were determined for CalvRNA8 based on the experimental data from several mapping methods and bioinformatics analysis. The secondary structure of CalvRNA8 allows for the validation of structural conservation for type A and strain-specific motifs. The new data show that vRNA8 secondary structure for type A is largely universal, but there are a few essential differences that are believed to be strain-specific.

The incorporation of 4sU into natural RNA could be used in a new method of RNA structure mapping where photocross-

linking with proximal nucleotides is generated. In this method, 4sU is randomly incorporated into RNA during transcription with stoichiometrically one modification per RNA, which can be easily optimized for studied RNA. Proposed quantification and interpretation of detected crosslinking sites make the method straightforward and readily applied to other RNA studies. Results on the model CalvRNA8 showed that photocrosslinking of 4sU occurred in flexible regions near or in double-stranded RNA, allowing for secondary and tertiary interaction recognition. In-cell culture studies showed that 4sU could photocrosslink with other RNA, DNA, or proteins, and any analysis or interpretation *via* this method should take this into account. Further, naked RNA needs to be used as a control for cellular experiments. In general, the 4sU mapping method is a valuable technique for getting more detailed information on RNA folding.

Studies of RNA structure using 4sU photocrosslinking can be performed *in vitro* and *in vivo*. Exposure of cells to 4sU results in rapid uptake, phosphorylation to 4sU-triphosphate, and incorporation into newly transcribed RNA (60, 61). Next, UV exposure will generate cross-linking, which is already a well-understood and routine procedure where 4-thiouridine is apply for assorted applications (38, 61, 62). The ability to use this mapping method in cells is a big advantage. Additionally, high 4sU-crosslinking efficiency allows for a low-cost, fast interpretation method for hits detections. Coupling 4sU mapping with chemical mapping and structure/sequence

analysis can support discovery of regions prone to antisense strategies or detections of viral RNA.

## Experimental procedures

### Experimental constructs

DNA template for the synthesis of CalvRNA was obtained in several steps. First, 0.1 MOI (multiplicity of infection) of A/California/04/2009 (H1N1) (a gift from Prof. Luis Martinez-Sobrido, Texas Biomedical Research Institute) was used to infect MDCK cells. Infected MDCK cells were incubated at 33 °C in the air with 5% $CO_2$ for 24 h. Next, total RNA was isolated from infected MDCK cells, and reverse transcription was carried out using SuperScript III (Invitrogen) with a specific primer (ATGAGTCTTCTAACCGAGGTCG), which is complementary to the 22 nucleotides at the 3′ end of CalvRNA8. Next, PCR reactions using cDNA as template and TTCTGCAGTTACTCTAGCTCTATGTTGACAAAATGACCATC and TTGAATTCATGAGTCTTCTAACCGAGGTCGAAAC primers were done to add an EcoRI site on the 5′ end and a Pst I site on the 3′ end of the template of segment 8 A/California/04/2009. DNA was purified using a PCR/DNA Clean-up Purification Kit (Eurx). The DNA template was cloned into pUC19 and sequenced using TGTAAAACGACGGCCAGT, TCACACAGGAAACAGCTATGAC, CAAGAAGGTCATCTTTCAGACCAG primers for confirmation of proper sequence.

### Oligonucleotide synthesis and labeling

Primers for PCR and reverse transcription (Experimental Constructs, Tables 1 and 2) were synthesized by the phosphoramidite approach on a MerMade12 synthesizer. Primers were deprotected and purified according to published protocols (63–65). Concentrations of all oligonucleotides were measured using a UV Spectrophotometer (NanoDrop2000 Thermo Scientific). Primers for reverse transcription were synthesized with aminolinker on the 5′ end and, after deprotection and desalting, were labeled from the 5′ end with fluorophores: 5-FAM, 5-ROX, 6-TAMRA, and 6-JOE (dyes from ANASPEC). For labeling, a reaction mixture containing 300 μg primer (11 μl in $H_2O$), 220 μg of diluted in DMSO fluorophores (14 μl), 75 μl 0.1 M sodium tetraborate, at pH 8.5 was incubated overnight at room temperature in a shaker oscillating at low speed. Labeled primers were precipitated with ethanol and purified by electrophoresis on a 12% denaturing PAGE.

### RNA synthesis

*In vitro* transcription was performed using an Ampliscribe T7 Flash kit (Epicenter) according to manufacture protocol.

**Table 1**
Primers for polymerase chain reaction using to obtain DNA template for CalvRNA8

| Name | Length (nt) | Sequence 5′→3′ |
|---|---|---|
| FC8 | 48 | GCGTAATACGACTCACTATAGGGAGTAGAAACAAGGGTGTTTTTTATC |
| RC8 | 29 | AGCAAAAGCAGGGTGACAAAACATAATG |

The underlined nucleotides residues are the polymerase T7 promoter.

**Table 2**
Primers for reverse transcription

| Name | Length (nt) | Sequence 5′→3′ |
|---|---|---|
| 1_8 | 23 | AGCAAAAGCAGGGTGACAAAAAC |
| 2_8 | 22 | ACATGACCCTCGAGGAAATGTC |
| 3_8 | 25 | CTTGAATGGAATGGTAACACGGTTC |

Each primer was labeled with 6-TAMRA, 5-FAM, 6-JOE, 5-ROX at 5′ end.

DNA template for transcriptions was obtained from PCR reaction with FC8 and RC8 primers (Table 1) and pUC19 with cloned segment 8 A/California/04/2009 sequence. DNA and RNA were purified using PCR/DNA Clean-up Purification Kit (Eurx) and RNeasy MiniElute Cleanup Kit (Qiagen), respectively. RNA quality was checked on an agarose gel.

### RNA synthesis with 4-thiouridine-5′-triphosphate

*In vitro* transcription with 4-thiouridine-5′-triphosphate (Tebu-Bio) was performed using modified Ampliscribe T7 Flash (Epicenter) manufacturer's protocol. Briefly, 2.5 mM 4-thiouridine-5′-triphosphate, 5.63 mM UTP, 7 mM of each ATP, CTP, GTP, 2 μl T7 RNA polymerase, and 1000 ng DNA temple were used. The reaction was incubated for 3 h at 37 °C. Next, 1.5 μl DNase I (Epicenter) was added and sample was incubated for 15 min. In the end, RNA was purified using RNeasy MiniElute Cleanup Kit (Qiagen). RNA quality was checked on an agarose gel.

### RNA folding

Before each experiment, RNA was folded using the same protocol. RNA was heated to 80 °C in water for 5 min and slowly cooled to 50 °C. Once at this temperature, the sample volume was doubled using 2× folding buffer (the final folding buffer contained 300 mM NaCl, 5 mM $MgCl_2$, 50 mM HEPES, pH 7.5), and samples were slowly cooled to 37 °C (1 °C/min). After this folding procedure, the RNA was folded into a single-molecule structure (Fig. S1 in Supporting Information 3).

### Photocrosslinking

Before each experiment, RNA was folded as described above. Steady-state photolysis experiments were performed in a 0.1 × 1 cm rectangular cell on an optical bench irradiation system using a Genesis CX355STM OPSL laser (Coherent), with 355 nm emission wavelength (the output power used was set at 60 mW). Samples were irradiated for 10 s, 30 s, 15 min, 30 min, and 45 min. Finally, 10 s was selected as the optimal time for experiments. For each experiment, a nonirradiated sample containing 4-thiouridine was used as control. Photocrosslinking was performed in the Faculty of Chemistry of Adam Mickiewicz University. Irradiation products were analyzed by primer extension with primers shown in Table 2 and read out with capillary electrophoresis as described below. The experiments were performed in at least technical triplicate with the average results presented. For the obtained reactivity of each nucleotide, the standard deviation (SD) was calculated (Supporting Information 1).

### Chemical mapping using NMIA, DMS, and CMCT

Folding of RNA was carried out as described above. Chemical mapping was conducted according to published procedures with appropriate optimizations (23, 47, 48). Briefly, 28.2 mM of NMIA (N-methylisatoic anhydride) or 30 mM of CMCT (1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate) or 0.18% of DMS (dimethyl sulfate) was used to study secondary structure of CalvRNA8. Chemical mapping was performed at 37 °C with DMS, CMCT, or NMIA for 15, 20, or 40 min, respectively. Parallel control reactions were done in the same condition but without mapping reagents. Modified nucleotides were read out by primer extension using a stoichiometry of 2 pmol primer/1.5 pmol RNA. Primer extension was performed at 55 °C with reverse transcriptase SuperScript III (Invitrogen) using the manufacturer's protocol. Primers labeled with 6-JOE were used for detection of modification by DMS, CMCT and NMIA. Primers labeled by 5-FAM were used for detection of reverse transcription products that serve as a control of the quality of RNA. For each primer labeled with 5-ROX and 6-TAMRA, there were prepared ddNTP ladders (most often ddGTP and ddATP). cDNA fragments were separated by capillary electrophoresis (Laboratory of Molecular Biology Techniques at Adam Mickiewicz University in Poznan). All experiments were performed at least in technical triplicate with the average results presented. For the obtained reactivity of each nucleotide, the SD was calculated (Supporting Information 1).

### Processing of chemical mapping and 4sU mapping data

ShapeFinder program was used to analyze mapping data according to the published method (66). Quantitative NMIA reactivities for individual datasets were normalized to a scale in which 0 indicates an unreactive site and the average intensity at highly reactive sites is set to 1.0. The normalization factor for each dataset was determined by first excluding the most reactive 2% of peak intensities and then calculating the average for the next 8% of peak intensities. All reactivities were then divided by this average. Reactivities ≥0.7 are treated as strong, 0.5 to 0.7 as medium and <0.5 as weak. All calculated reactivities were used in the prediction of RNA secondary structure. Nucleotides with no data were designated as −999. Normalized SHAPE reactivities from extension reactions of each primer were processed independently. DMS and CMCT modifications analysis was conducted similar to NMIA reactivity calculations, except that only strong modifications were used in RNAstructure prediction.

4sU photoadducts were detected using reverse transcription. The crosslinked nucleotides cause RT stops and were detected and analyzed in the same manner as classical chemical mapping with assignment of strong and medium crosslinking sites. For each mapping method (SHAPE, DMS, CMCT mapping, and using 4sU photocrosslinking), at least three datasets were obtained from each primer (three technical repeats) and the average of results was used.

Classical chemical mapping methods results were used in RNAstructure 6.0.1 (49) for the prediction of secondary structure of CalvRNA8. Normalized SHAPE reactivity (as described above) was used in "Read SHAPE reactivity—pseudo free energy" mode with a slope of 1.8 and an intercept of −0.6 kcal/mol (24). DMS and CMCT strong reactivities were introduced in the same prediction using "chemical modification" mode (67, 68). Additionally, conserved type A base pairs (conservation >95%, in agreement with chemical mapping results) were used as constraints.

### Data availability

All relevant data are included in the main text and supporting information.

*Abbreviations*—The abbreviations used are: ASO, antisense oligonucleotide; CMCT, 1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate; DMS, dimethyl sulfide; IAV, influenza A virus; MDCK, Madin–Darby canine kidney; SD, standard deviation; SHAPE, selective 2′-hydroxyl acylation and primer extension.

### References

1. Bouvier, N. M., and Palese, P. (2008) The biology of influenza viruses. *Vaccine* **26**, D49–D53
2. *Epidemic and Pandemic Alert Response Unit WHO Guidelines for Investigation of Human Cases of Avian Influenza A (H5N1).* (2007). World Health Organization, Geneva, Switzerland
3. Adalja, A. A., Watson, M., Toner, E. S., Cicero, A., and Inglesby, T. V. (2019) Characteristics of microbes most likely to cause pandemics and global catastrophes. *Curr. Top. Microbiol. Immunol.* **424**, 1–20
4. Li, K. S., Guan, Y., Wang, J., Smith, G. J., Xu, K. M., Duan, L., Rahardjo, A. P., Puthavathana, P., Buranathai, C., Nguyen, T. D., Estoepangestie, A. T., Chaisingh, A., Auewarakul, P., Long, H. T., Hanh, N. T., *et al.* (2004) Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* **430**, 209–213
5. Neumann, G., Noda, T., and Kawaoka, Y. (2009) Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* **459**, 931–939
6. Kim, H., Webster, R. G., and Webby, R. J. (2018) Influenza virus: Dealing with a drifting and shifting pathogen. *Viral Immunol.* **31**, 174–183

7. Wayne, K.,D., Lazniewski, M., and Plewczynski, D. (2017) RNA structure interactions and ribonucleoprotein processes of the influenza A virus. *Brief Funct. Genomics* **17**, 402–414

8. Ferhadian, D., Contrant, M., Printz-Schweigert, A., Smyth, R. P., Jean-Paillart, C., and Marquet, R. (2018) Structural and functional motifs in influenza virus RNAs. *Front. Microbiol.* **9**, 559

9. Piasecka, J., Lenartowicz, E., Soszynska-Jozwiak, M., Szutkowska, B., Kierzek, R., and Kierzek, E. (2020) RNA secondary structure motifs of the influenza A virus as targets for siRNA-mediated RNA interference. *Mol. Ther. Nucleic Acids* **19**, 627–642

10. [preprint] Takizawa, N., Higashi, K., Kawaguchi, R. K., Gotoh, Y., Suzuki, Y., Hayashi, T., and Kurokawa, K. (2020) A functional RNA structure in the influenza A virus ribonucleoprotein complex for segment bundling. *bioRxiv.* https://doi.org/10.1101/2020.03.05.975870

11. Priore, S. F., Kierzek, E., Kierzek, R., Baman, J., Moss, W. N., Dela-Moss, L. I., and Turner, D. H. (2013) Secondary structure of a conserved domain in the intron of influenza A NS1 mRNA. *PLoS One* **8**, e70615

12. Lenartowicz, E., Kesy, J., Ruszkowska, A., Soszynska-Jozwiak, M., Michalak, P., Moss, W. N., Turner, D. H., Kierzek, R., and Kierzek, E. (2016) Self-folding of naked segment 8 genomic RNA of influenza A virus. *PLoS One* **11**, e0148281

13. Ruszkowska, A., Lenartowicz, E., Moss, W. N., Kierzek, R., and Kierzek, E. (2016) Secondary structure model of the naked segment 7 influenza A virus genomic RNA. *Biochem. J.* **473**, 4327–4348

14. Michalak, P., Soszynska-Jozwiak, M., Biala, E., Moss, W. N., Kesy, J., Szutkowska, B., Lenartowicz, E., Kierzek, R., and Kierzek, E. (2019) Secondary structure of the segment 5 genomic RNA of influenza A virus and its application for designing antisense oligonucleotides. *Sci. Rep.* **9**, 3801

15. Michalak, P., Piasecka, J., Szutkowska, B., Kierzek, R., Biala, E., Moss, W. N., and Kierzek, E. (2021) Conserved structural motifs of two distant IAV subtypes in genomic segment 5 RNA. *Viruses* **13**, 525

16. Gultyaev, A. P., Heus, H. A., and Olsthoorn, R. C. L. (2007) An RNA conformational shift in recent H5N1 influenza A viruses. *Bioinformatics* **23**, 272–276

17. Chevalier, C., Goffic, R. L., Jamme, F., Leymarie, O., Réfrégiers, M., and Delmas, B. (2016) Synchrotron infrared and deep UV fluorescent microspectroscopy study of PB1-F2 β-aggregated structures in influenza A virus-infected cells. *J. Biol. Chem.* **291**, 9060–9072

18. Gultyaev, A., Fouchier, R., and Olsthoorn, R. (2010) Influenza virus RNA structure: Unique and common features. *Rev. Immunol.* **29**, 533–556

19. Mathews, D. H., Moss, W. N., and Turner, D. H. (2010) Folding and finding RNA secondary structure. *Cold Spring Harb. Perspect. Biol.* **2**, a003665

20. Mailler, E., Paillart, J., Marquet, R., Smyth, R. P., and Vivet-Boudou, V. (2019) The evolution of RNA structural probing methods: From gels to next-generation sequencing. *Wiley Interdiscip. Rev. RNA* **10**, e1518

21. Kierzek, E. (2009) Binding of short oligonucleotides to RNA: Studies of the binding of common RNA structural motifs to isoenergetic microarrays. *Biochemistry* **48**, 11344–11356

22. Kierzek, R., Turner, D. H., and Kierzek, E. (2015) Microarrays for identifying binding sites and probing structure of RNAs. *Nucleic Acids Res.* **43**, 1–12

23. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P., and Ehresmann, B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128

24. Deigan, K. E., Mathews, D. H., and Weeks, K. M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 97–102

25. Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2005) RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA(Asp) transcripts. *J. Am. Chem. Soc.* **127**, 4659–4667

26. Busan, S., Weidmann, C. A., Sengupta, A., and Weeks, K. M. (2019) Guidelines for SHAPE reagent choice and detection strategy for RNA structure probing studies. *Biochemistry* **58**, 2655–2664

27. Regulski, E. E., and Breaker, R. R. (2008) In-line probing analysis of riboswitches. *Methods Mol. Biol.* **419**, 53–67

28. Palou-Mir, J., Barceló-Oliver, M., and Sigel, R. K. O. (2017) The role of lead(II) in nucleic acids. *Met. Ions Life Sci.* **17**, 403–434

29. Aw, J. G. A., Shen, Y., Wilm, A., Sun, M., Lim, X. N., Boon, K. L., Tapsin, S., Chan, Y. S., Tan, C. P., Sim, A. Y. L., Zhang, T., Susanto, T. T., Fu, Z., Nagarajan, N., and Wan, Y. (2016) Mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* **62**, 603–617

30. Testa, S. M., Disney, M. D., Turner, D. H., and Kierzek, R. (1999) Thermodynamics of RNA-RNA duplexes with 2-or 4-thiouridines: Implications for antisense design and targeting a group I intron. *Biochemistry* **38**, 16655–16662

31. Sidaway-Lee, K., Costa, M. J., Rand, D. A., Finkenstadt, B., and Penfield, S. (2014) Direct measurement of transcription rates reveals multiple mechanisms for configuration of the Arabidopsis ambient temperature response. *Genome Biol.* **15**, R45

32. Sontheimer, E. J. (1994) Site-specific RNA crosslinking with 4-thiouridine. *Mol. Biol. Rep.* **20**, 35–44

33. Lauhon, C. T., Erwin, W. M., and Ton, G. N. (2004) Substrate specificity for 4-thiouridine modification in Escherichia coli. *J. Biol. Chem.* **279**, 23022–23029

34. Peak, M. J., Midden, W. M., Babasick, D. M., and Haugen, D. A. (1988) Photochemistry of 4-thiouridine and thymine. *Photochem. Photobiol.* **48**, 229–232

35. Williams, G. D., Townsend, D., Wylie, K. M., Kim, P. J., Amarasinghe, G. K., Kutluay, S. B., and Boon, A. C. M. (2018) Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat. Commun.* **9**, 465

36. Zhou, Y., and Routh, A. (2019) Mapping RNA-capsid interactions and RNA secondary structure within authentic virus particles using next-generation sequencing. *Nucleic Acids Res.* **48**, e12

37. Garibaldi, A., Carranza, F., and Hertel, K. J. (2017) Isolation of newly transcribed RNA using the metabolic label 4-thiouridine. *Nucleic Acids Res.* **19**, 3653–3660

38. Duffy, E. E., Schofield, J. A., and Simon, M. D. (2019) Gaining insight into transcriptome-wide RNA population dynamics through the chemistry of 4-thiouridine. *Wiley Interdiscip. Rev. RNA* **10**, e1513

39. Rutkowski, A. J., and Dölken, L. (2017) High-resolution gene expression profiling of RNA synthesis, processing, and decay by metabolic labeling of newly transcribed RNA using 4-thiouridine. *Methods Mol. Biol.* **1507**, 129–140

40. Wollenzien, P., Expert-Bezançon, A., and Favre, A. (1991) Sites of contact of mRNA with 16S rRNA and 23S rRNA in the Escherichia coli ribosome. *Biochemistry* **30**, 1788–1795

41. Reymond, C., Ouellet, J., Bisaillon, M., and Perreault, J. P. (2007) Examination of the folding pathway of the antigenomic hepatitis delta virus ribozyme reveals key interactions of the L3 loop. *RNA* **13**, 44–54

42. Ryan, D. E., Kim, C. H., Murray, J. B., Adams, C. J., Stockley, P. G., and Abelson, J. (2004) New tertiary constraints between the RNA components of active yeast spliceosomes: A photo-crosslinking study. *RNA* **10**, 1251–1265

43. Juzumiene, D. I., and Wollenzien, P. (2001) Arrangement of the central pseudoknot region of 16S rRNA in the 30S ribosomal subunit determined by site-directed 4-thiouridine crosslinking. *RNA* **7**, 71–84

44. Baranov, P. V., Dokudovskaya, S. S., Oretskaya, T. S., Dontsova, O. A., Bogdanov, A. A., and Brimacombe, R. (1997) A new technique for the characterization of long-range tertiary contacts in large RNA molecules: Insertion of a photolabel at a selected position in 16S rRNA within the Escherichia coli ribosome. *Nucleic Acids Res.* **25**, 2266–2273

45. Dubreuil, Y. L., Expert-Bezancon, A., and Favre, A. (1991) Conformation and structural fluctuations of a 218 nucleotides long rRNA fragment: 4-thiouridine as an intrinsic photolabelling probe. *Nucleic Acids Res.* **19**, 3653–3660

46. Favre, A., Saintomé, C., Fourrey, J. L., Clivio, P., and Laugâa, P. (1998) Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *J. Photochem. Photobiol. B Biol.* **42**, 109–124

47. Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231

48. Ziehler, W. A., and Engelke, D. R. (2001) Probing RNA structure with chemical reagents and enzymes. *Curr. Protoc. Nucleic Acid Chem.* Chapter 6:Unit 6.1

49. Reuter, J. S., and Mathews, D. H. (2010) RNAstructure: Software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**, 129

50. Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M., and Palese, P. (1987) Genomic RNAs of influenza-viruses are held in a circular conformation in virions and in infected-cells by a terminal panhandle. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8140–8144

51. Bellaousov, S., and Mathews, D. H. (2010) ProbKnot: Fast prediction of RNA secondary structure including pseudoknots. *RNA* **16**, 1870–1880

52. Manfredonia, I., Nithin, C., Ponce-Salvatierra, A., Ghosh, P., Wirecki, T. K., Marinus, T., Ogando, N. S., Snijder, E. J., van Hemert, M. J., Bujnicki, J. M., and Incarnato, D. (2020) Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.* **48**, 12436–12452

53. Incarnato, D., Morandi, E., Simon, L. M., and Oliviero, S. (2018) RNA framework: An all-in-one toolkit for the analysis of RNA structures and post-transcriptional modifications. *Nucleic Acids Res.* **46**, e97

54. Nawrocki, E. P., and Eddy, S. R. (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935

55. Rivas, E., Clements, J., and Eddy, S. R. (2020) Estimating the power of sequence covariation for detecting conserved RNA structure. *Bioinformatics* **36**, 3072–3076

56. Gultyaev, A. P., Spronken, M. I., Richard, M., Schrauwen, E. J., Olsthoorn, R. C., and Fouchier, R. A. (2016) Subtype-specific structural constraints in the evolution of influenza A virus hemagglutinin genes. *Sci. Rep.* **6**, 38892

57. Dadonaite, B., Gilbertson, B., Knight, M. L., Trifkovic, S., Rockman, S., Laederach, A., Brown, L. E., Fodor, E., and Bauer, D. L. V. (2019) The structure of the influenza A virus genome. *Nat. Microbiol.* **4**, 1781–1789

58. Le Sage, V., Kanarek, J. P., Snyder, D. J., Cooper, V. S., Lakdawala, S. S., and Lee, N. (2020) Mapping of influenza virus RNA-RNA interactions reveals a flexible network. *Cell Rep.* **31**, 107823

59. Lenartowicz, E., Nogales, A., Kierzek, E., Kierzek, R., Martínez-Sobrido, L., and Turner, D. H. (2016) Antisense oligonucleotides targeting influenza A segment 8 genomic RNA inhibit viral replication. *Nucleic Acid Ther.* **26**, 277–285

60. Rädle, B., Rutkowski, A. J., Ruzsics, Z., Friedel, C. C., Koszinowski, U. H., and Dölken, L. (2013) Newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. *J. Vis. Exp.* **8**, 50195

61. Spitzer, J., Hafner, M., Landthaler, M., Ascano, M., Farazi, T., Wardle, G., Nusbaum, J., Khorshid, M., Burger, L., Zavolan, M., and Tuschl, T. (2014) PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation): A step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods Enzymol.* **539**, 113–161

62. Schueler, M., Munschauer, M., Gregersen, L. H., Finzel, A., Loewer, A., Chen, W., Landthaler, M., and Dieterich, Ch. (2014) Differential protein occupancy profiling of the mRNA transcriptome. *Genome Biol.* **15**, R15

63. Xia, T. B., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X. Q., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **37**, 14719–14735

64. Kierzek, E., and Kierzek, R. (2003) The thermodynamic stability of RNA duplexes and hairpins containing N-6-alkyladenosines and 2-methylthio-N-6-alkyladenosines. *Nucleic Acids Res.* **31**, 4472–4480

65. Fratczak, A., Kierzek, R., and Kierzek, E. (2009) LNA-modified primers drastically improve hybridization to target RNA and reverse transcription. *Biochemistry* **48**, 514–516

66. Vasa, S. M., Guex, N., Wilkinson, K. A., Weeks, K. M., and Giddings, M. G. (2008) ShapeFinder: A software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA* **14**, 1979–1990

67. Schroeder, S. J., Stone, J. W., Bleckley, S., Gibbons, T., and Mathews, D. M. (2011) Ensemble of secondary structures for encapsidated satellite tobacco mosaic virus RNA consistent with chemical probing and crystallography constraints. *Biophys. J.* **101**, 167–175

68. Schroeder, S. J. (2014) Probing viral genomic structure: Alternative viewpoints and alternative structures for satellite tobacco mosaic virus RNA. *Biochemistry* **53**, 6728–6737

ASBMB

*J. Biol. Chem.* (2021) 297(6) 101245 **13**