Functional Analysis of the Influenza Virus H5N1 Nucleoprotein Tail Loop Reveals Amino Acids That Are Crucial for Oligomerization and Ribonucleoprotein Activities[▽]

Wai-Hon Chan,^{1,2†} Andy Ka-Leung Ng,^{1,2†} Nicole C. Robb,⁴ Mandy Ka-Han Lam,¹
Paul Kay-Sheung Chan,³ Shannon Wing-Ngor Au,^{1,2} Jia-Huai Wang,⁵
Ervin Fodor,⁴ and Pang-Chui Shaw^{1,2*}

Department of Biochemistry and Centre for Protein Science and Crystallography,¹ Molecular Biotechnology Programme,² and Department of Microbiology,³ the Chinese University of Hong Kong, Shatin, NT, Hong Kong, China; Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom⁴; and Department of Medical Oncology, Dana-Farber Cancer Institute, Department of Pediatrics, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115⁵

Received 24 November 2009/Accepted 21 April 2010

Homo-oligomerization of the nucleoprotein (NP) of influenza A virus is crucial for providing a major structural framework for the assembly of viral ribonucleoprotein (RNP) particles. The nucleoprotein is also essential for transcription and replication during the virus life cycle. In the H5N1 NP structure, the tail loop region is important for NP to form oligomers. Here, by an RNP reconstitution assay, we identified eight NP mutants that had different degrees of defects in forming functional RNPs, with the RNP activities of four mutants being totally abolished (E339A, V408S P410S, R416A, and L418S P419S mutants) and the RNP activities of the other four mutants being more than 50% decreased (R267A, I406S, R422A, and E449A mutants). Further characterization by static light scattering showed that the totally defective protein variants existed as monomers *in vitro*, deviating from the trimeric/oligomeric form of wild-type NP. The I406S, R422A, and E449A variants existed as a mixture of unstable oligomers, thus resulting in a reduction of RNP activity. Although the R267A variant existed as a monomer *in vitro*, it resumed an oligomeric form upon the addition of RNA and retained a certain degree of RNP activity. Our data suggest that there are three factors that govern the NP oligomerization event: (i) interaction between the tail loop and the insertion groove, (ii) maintenance of the tail loop conformation, and (iii) stabilization of the NP homo-oligomer. The work presented here provides information for the design of NP inhibitors for combating influenza virus infection.

Influenza has long been a major threat to public health, causing annual epidemics of respiratory illness and occasional pandemics. The death toll from influenza epidemics worldwide ranges from 250,000 to 500,000 people each year. Highly pathogenic H5N1 avian influenza virus infections in humans have resulted in a mortality rate of about 60% (http://www.who.int/csr/disease/avian_influenza/country/cases_table_2010_05_06/en/index.html). The emergence of the 2009 H1N1 swine influenza virus strain, despite having low lethality, raises the concern of the reassortment of this virus with another more virulent strain.

The genome of influenza A virus contains eight single-stranded negative-sense RNA segments (viral RNA [vRNA]). These RNA segments interact with multiple copies of nucleoprotein (NP) and three polymerase subunits (PA, PB1, and PB2) to form ribonucleoprotein (RNP) complexes (20).

NP has been implicated in the transcription-replication processes of influenza virus, as temperature-sensitive NP mutants resulted in defective transcription at nonpermissive temperatures (13, 16, 29, 32). NP may act as a cofactor that cotrans-

criptionally coats the newly synthesized cRNA (30). It may also alter the structure of the RNA template to change its mode of synthesis (11, 12), or it may interact with the polymerase subunits PB1 and PB2 to change the transcriptional activity of the polymerase (5, 19).

In the virus particle, the RNP complexes are organized into a distinct pattern, as visualized by transmission electron microscopy (EM) (23). NP not only encapsidates the viral RNA but also forms homo-oligomers to maintain the RNP structure (26). The homo-oligomerization of NP forms a major part of the RNP complex (24), as observed previously for purified intact viral RNPs by EM (27) and for mini-RNPs (2, 17). It is the RNP structure rather than the naked vRNA that acts as the template for transcription and replication in the infected cell (4). NP homo-oligomerization was previously found to be promoted by two positive elements (amino acids 189 to 358 and 371 to 465) but inhibited by a negative element (amino acids 465 to 498) (8). However, the correlation between NP homo-oligomerization and transcriptional function is ambiguous, since amino acid substitutions in both the positive (R199A and R416A) and the negative (F479A) elements showed decreased transcriptional competence (8).

Previously, we determined the atomic structure of H5N1 NP, which provides insights into the molecular mechanisms of the oligomerization of the protein (21). In the formation of NP oligomers, the tail loop of NP has to be bent toward the body

* Corresponding author. Mailing address: Department of Biochemistry, the Chinese University of Hong Kong, Shatin, NT, Hong Kong, China. Phone: (852)-26096803. Fax: (852)-26037246. E-mail: pshaw@cuhk.edu.hk.

† W.-H.C. and A.K.-L.N. have contributed equally to this work.

[▽] Published ahead of print on 12 May 2010.

domain, through the two flexible linker regions linking the tail loop (6, 21). The identification of the essential molecular contacts between neighboring NPs will shed light on the sequential event of the tail loop insertion process and provide valuable information for the design of NP inhibitors.

MATERIALS AND METHODS

Biological materials. The 293T cell line (ATCC, Manassas, VA) was cultivated in minimal essential medium (MEM) (Invitrogen, Carlsbad, CA) with 10% fetal calf serum (Invitrogen). Anti-NP serum was prepared by immunizing rabbits with purified NP. Anti-Myc antibody (Ab) (Cell Signaling, Danvers, MA) and anti-beta-actin antibody (Santa Cruz Biotechnology, Santa Cruz, CA) were purchased commercially. Plasmids pcDNA-PB1, pcDNA-PB2, pcDNA-PA, pcDNA-NP, and pPOLI-NA-RT were described previously (9, 33). NP mutants were cloned into pcDNA3 (Invitrogen) and pcDNA3.1/myc-His (Invitrogen) for the expression of untagged and Myc-tagged NP variants in 293T cells. The PA gene was also cloned into pcDNA3.1/myc-His (Invitrogen) to generate Myc-tagged PA.

RNA analysis by primer extension assay. Human kidney 293T cells were used to reconstitute RNP complexes. Plasmids pcDNA-PB1, pcDNA-PB2, pcDNA-PA, pcDNA-NP, and pPOLI-NA-RT (1.0 µg of each) was diluted to a total volume of 250 µl in Opti-MEM (Invitrogen) and subsequently added to a mix of 6 µl of Lipofectamine 2000 (Invitrogen) in 250 µl of Opti-MEM. The transfection mixture was incubated for 30 min before being added to 1.5 ml (about 10⁶ cells) 293T cells in a suspension in MEM containing 10% fetal calf serum in 35-mm dishes. Cells were harvested at 48 h posttransfection, and total RNA was extracted with TRIzol reagent (Invitrogen).

Primer extension assays were performed as described previously (9, 33). Briefly, an excess of DNA primer (about 10⁵ cpm), labeled at its 5' end with ³²P, was mixed with 5 µg of total RNA in 5 µl of water and denatured at 95°C for 3 min. The mixture was cooled on ice and subsequently incubated at 45°C for 1 h with the addition of 50 U SuperScript II RNase H⁻ reverse transcriptase (Invitrogen) in first-strand buffer (Invitrogen). Two NA gene-specific primers and one 5S rRNA primer (used as an internal control) were used in the same reverse transcription reaction: 5'-GGACTAGTGGGAGCATCAT-3' (to detect vRNA), 5'-TCCAGTATGGTTTTGATTTCGG-3' (to detect mRNA and cRNA), and 5'-TCCCAGGCGGTCTCCCATCC-3' (to detect 5S rRNA). Reactions were stopped by the addition of 8 µl of 90% formamide and heating at 95°C for 3 min. Transcription products were analyzed on 6% polyacrylamide gels containing 7 M urea in Tris-borate-EDTA (TBE) buffer and detected by autoradiography. Phosphorimaging analysis with ImageQuant TL (GE Healthcare, Waukesha, WI) was used for quantification. An unpaired Student's *t* test was used for analyses of significance.

Co-IP. For the NP homo-oligomerization experiment, 1 µg each of untagged and Myc-tagged NP plasmids was transfected into 10⁶ human kidney 293T cells in suspension. Coimmunoprecipitation (co-IP) was performed at 48 h posttransfection. Cells were resuspended in a lysis buffer containing 50 mM Tris-HCl (pH 7.6), 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100 (co-IP buffer) and lysed by sonication. The lysate was centrifuged at 16,000 × *g* for 10 min at 4°C. The supernatant was incubated at 4°C overnight with or without anti-Myc Ab. The mixture was then incubated with protein A beads for 1.5 h at 4°C with shaking. The beads were centrifuged and washed with co-IP buffer three times before being boiled in SDS loading dye and analyzed by Western blotting. For the NP-polymerase interaction experiment, Myc-tagged PA and untagged PB1 and PB2 were transfected with the various untagged NP mutants, and a protocol similar to that described above was followed.

Protein expression and purification. Maltose binding protein (MBP)-tagged NP was expressed in *Escherichia coli* BL21(DE3)pLysS cells. The cells were lysed by sonication, and the lysate was passed through an amylose column (New England Biolabs, Ipswich, MA). The bound protein was eluted with a 0 to 20 mM maltose gradient in 20 mM sodium phosphate (pH 6.5) and 150 mM NaCl. The eluate was incubated with thrombin (100 U) (Sigma, St. Louis, MO) and RNase A (100 U) (Sigma) at 4°C overnight to remove MBP from NP and then passed through a heparin HP column (GE Healthcare). NP was eluted with a 0 to 1.5 M NaCl gradient in the same buffer. Gel filtration was performed with Superdex 200 (GE Healthcare). RNase A was removed after passage through a heparin high-performance (HP) column and a gel filtration column. NP mutants were generated by site-directed mutagenesis of wild-type (WT) plasmid pRSETMBP-NP (21) according to a standard protocol and were purified as described above for the wild-type protein.

Gel shift assay. A 24-nucleotide (nt) 2'-O-methylated RNA oligonucleotide with the sequence 5'-UUU GUU ACA CAC ACA CAC GCU GUG-3' was

purchased (RiboBioscience, Guangzhou, China). A fixed amount of RNA (10 µM) was incubated with an increasing amount of purified wild-type NP or protein variants (0, 5, 10, and 20 µM) for 30 min at room temperature. The mixture was subjected to agarose gel electrophoresis and visualized by ethidium bromide staining.

SPR. A biotinylated 2'-O-methylated RNA oligonucleotide with the sequence 5'-UUU GUU ACA CAC ACA CAC GCU GUG-3' was immobilized on an SA sensor chip (GE Healthcare) until the surface density reached 30 to 35 response units (RU), according to manufacturer's instructions (GE Healthcare). Surface plasmon resonance (SPR) measurements were carried out with a BIAcore 3000 system at 25°C. Data were analyzed with BIAevaluation v. 4.1 software.

Static light scattering. Wild-type or mutant NP proteins were subjected to static light scattering analysis by using a miniDAWN triangle (45°, 90°, and 135°) light scattering detector (Wyatt Technology Corporation, Santa Barbara, CA) connected to an Optilab DSP interferometric refractometer (Wyatt Technology Corporation). This system was connected to a Superdex 200 column (GE Healthcare) controlled by an AKTAexplorer chromatography system (GE Healthcare). Before sample injection, the miniDAWN detector system was equilibrated with 20 mM sodium phosphate (pH 7.0) and 150 mM NaCl for at least 2 h to ensure a stable baseline signal. The flow rate was set to 0.7 ml/min, and the sample volume was 100 µl. The laser scattering (687 nm) and the refractive index (690 nm) of the respective protein solutions were recorded during the measurement processes. Wyatt ASTRA software was used to evaluate all data obtained. For the R267A mutant, an 1,850-nt-long RNA was transcribed from pTRI-Xef *in vitro* (Ambion, Austin, TX) and added to the protein solution.

RESULTS

The tail loop insertion is maintained by intra- and intermolecular interactions. As observed for the trimeric structure, NP homo-oligomerizes by inserting the tail loop (amino acids 402 to 428) into the groove of the body domain of its neighboring NP (Fig. 1A). It was hypothesized that NP uses the same mechanism in forming oligomeric NP in the RNP structure (21). Two forces appear to govern the insertion, the maintenance of the tail loop structure and interaction between the tail loop and the groove, and these forces are summarized by schematic diagrams in Fig. 1B and C, respectively. The tail loop structure is apparently maintained by three clusters of residues, namely, T1, T2, and T3, at the tip, stem, and base. T1 involves a salt bridge centered at R416, whose side chain interacts with the main chain of F412 and V414. T2 is situated at the stem and forms a zigzag pattern of hydrophobic interactions between V408, P410, L418, and P419. T3 is another hydrophobic cluster centered at I406, which interacts with E421 and I425. T3 seems to play a role in stabilizing the base of the tail loop. The tail loop insertion event was observed to involve five clusters of residues, I1 to I5. I1 and I2 involve salt bridges at the tip region, featuring interactions among R416, E339, T411, S335, and T390. I3 is located at the stem and is centered at R267, which interacts with S407, V408, L418, and P419. I4 is a hydrophobic cluster of F420, I265, and P453 at the stem. I5, situated at the base, involves a salt bridge between R422 and E449.

NP mutants display defective transcription-replication activity. Since NP is a major component of the RNP complex (25), it is conceivable that mutations in NP may alter the transcription-replication activity. Based on structural information, we first constructed 13 single-point mutants and 1 double-point mutant targeting the residues in the eight above-mentioned clusters (Fig. 1B and C) in the hope of identifying the crucial residues for the tail loop insertion event. Charged residues were mutated to alanine to remove the hydrogen bond interaction, while other residues were mutated to serine to

FIG. 1. Crystal structure and tail loop interaction of NP. (A) NP consists of head and body domains and a long tail loop, which inserts into the neighboring NP to form homo-oligomers. (B) Intrachain interaction of the tail loop. Three regions of molecular contacts were identified from the NP crystal structure. T1 (dotted lines) involves hydrophilic interactions, while T2 and T3 (double solid lines) involve hydrophobic interactions. (C) Interchain interaction of the tail loop. Most of the interactions are hydrophilic in nature (I1, I2, I3, and I5), except for I4, which is hydrophobic. The thicknesses of the dotted lines indicate the strength of the hydrogen bond between the two residues. Amino acid residues of the tail loop are indicated by circles, while amino acid residues of the insertion groove are indicated by rectangles.

remove the hydrophobic interaction. The mutated NP plasmids were then cotransfected with plasmids expressing polymerase proteins and a reporter plasmid. The amounts of plasmid were adjusted for individual NP mutants to obtain similar expression levels of the protein (Fig. 2A). A plasmid expressing wild-type NP was used as the positive control, while an empty plasmid was the negative control. At 48 h posttransfection, RNA was extracted, and the vRNA, cRNA, and mRNA levels of the reporter gene were quantified by a primer extension assay, followed by polyacrylamide gel electrophoresis and autoradiography (Fig. 2B). The various RNA levels were normalized to the internal 5S rRNA control and compared with those of wild-type NP.

Among those mutants that were constructed to assess the importance of intersubunit tail loop insertion events for transcription-replication activity, the E339A and R416A mutants showed nearly undetectable activity, and the R267A, R422A, and E449A mutants showed a more-than-50% decrease, while the T390S, S407A, T411S, and F420S mutants did not show a significant change (Fig. 2C). On the other hand, among mutants that could possibly destroy the tail loop conformation, the L418S P419S and R416A mutants showed undetectable activ-

ity, and the I406S mutant displayed a more-than-50% decrease in RNP activity, while the V408S, P410S, and I425S mutants did not. When considering the hydrophobic cluster of T2, we wondered if a double mutant of V408S and P410S would give a result similar to that for the L418S P419S mutant and whether the L418S and P419S single-point mutants would resemble the V408S or P410S mutant. To address this question, we constructed three more mutants (V408S P410S, L418S, and P419S mutants) and performed further experiments as described above. We found that the V408S P410S mutant resulted in nearly nondetectable RNP activity, while the L418S and P419S mutants had no significant change (Fig. 2B and C).

Defective NP variants interact with RNA and the polymerase complex. The above-described work showed that eight NP single- or double-point mutants at the tail loop and the groove of insertion exerted a negative effect on the transcription-replication activity of RNP. We hypothesized that these mutations may disrupt the tail loop insertion event and, thus, NP homo-oligomerization. However, for a functional RNP, NP has to interact with both RNA and the trimeric polymerase complex (3, 5, 18). Therefore, we first tested whether the eight defective NP variants still possessed these essential contacts.

The eight NP variants were expressed and purified from *E. coli*. Increasing amounts of purified proteins were incubated with a 24-nucleotide 2'-O-methylated RNA oligonucleotide. 2'-O-methylated RNA was used to confer resistance to RNase. Samples were then analyzed by agarose gel electrophoresis and ethidium bromide staining. As shown in Fig. 3A, the NP variants retained their affinities for the RNA.

Since a gel shift assay is not sensitive and the amount of protein used might have already been in saturation, surface plasmon resonance (SPR) was employed to provide a more quantitative estimation of the protein-RNA binding affinity (Table 1 and see www.bch.cuhk.hk/pcshaw/jvi.pdf). As shown in Table 1, the affinity between the variant NP and RNA remained at a submicromolar level. Interestingly, the K_D (dissociation constant) values may be categorized into two groups: one similar to that of wild-type NP RNA and another about 20 to 60 times higher. The discrepancy is probably attributed by the different *in vitro* oligomeric states during the SPR experiment. The more oligomeric variants (the WT and the I406S, R422A, and E449A variants) have more RNA binding sites and thus contribute to a higher association rate and a higher affinity (Table 1). On the other hand, these differences in affinities for RNA do not correlate with their RNP activities. For example, the affinities of some fully active NP variants were significantly lower than those of wild-type NP but similar to those of partially active and inactive variants. We therefore conclude that the variation in RNA affinities detected by SPR is probably not crucial for the RNP activity. As illustrated below, we found that RNP activity is instead related to the oligomerization state of the protein.

To investigate their interaction with the viral polymerase, NP mutants were coexpressed with PB1, PB2, and Myc-tagged PA in 293T cells. Coimmunoprecipitation was performed by using anti-Myc antibodies at 48 h posttransfection. Samples were analyzed by SDS-PAGE, followed by Western blotting (Fig. 3B). The absence of NP in the negative control lacking Myc-tagged PA indicated that NP alone was not coimmunoprecipitated by the anti-Myc anti-

FIG. 2. Identification of important amino acid residues involved in the tail loop insertion by an RNP reconstitution assay. (A) Western blotting of cell lysates expressing NP mutants. The amount of NP-expressing plasmids was adjusted to give similar expression levels. Beta-actin was used as an internal control. (B) *In vivo* RNP reconstitution followed by primer extension assay of NP mutants. Different pcDNA-NP plasmids were mixed with pcDNA-PB1, pcDNA-PB2, pcDNA-PA, and pPOLI-NA-RT and transfected into 293T cells. pcDNA3 instead of pcDNA-NP was used for the negative control. RNA was isolated from cells after 48 h of incubation and analyzed by a primer extension assay as described in Materials and Methods. Positions of vRNA, mRNA, cRNA, and 5S rRNA are indicated on the right. A representative result of three independent experiments is shown. (C) Quantitation of mRNA, cRNA, and vRNA. RNA levels for each NP mutant were compared to those of wild-type NP, which was set to 100%. 5S rRNA was used as an internal control to standardize RNA levels. The results represent the mean percentages $\pm$ standard deviations from three independent experiments (*, $P < 0.02$; **, $P < 0.0005$).

bodies. NP was also not detected if anti-Myc antibodies were not present during coimmunoprecipitation, demonstrating that NP and the NP-polymerase complex did not associate with the protein A beads themselves. In contrast, wild-type NP was coimmunoprecipitated in the presence of Myc-tagged PA and anti-Myc antibodies. In addition, all NP variants were shown to be coimmunoprecipitated with the Myc-tagged polymerase complex, demonstrating that they were able to associate with the RNA polymerase.

Defective NP variants possess abnormal oligomeric states *in vitro*. We then further tested the oligomerization states of the variants by static light scattering. The purified NP variants were passed through a gel filtration column, a light scattering detector, and a refractive interferometer. The oligomeric state of the proteins was revealed by the calculated native molecular weight. The proteins were treated with RNase A during the

purification process. Wild-type NP appeared as a mixture of trimers and tetramers, which is consistent with data from our previous study (21). All eight defective NP variants showed aberrant oligomeric states compared with wild-type NP (Fig. 4), suggesting that NP-NP homo-oligomerization plays a significant role in the transcription-replication process of the virus.

The NP E339A, R416A, V408S P410S, and L418S P419S mutants showed almost undetectable levels of transcription-replication activity (Fig. 2B). Static light scattering of these four variants revealed that they were monomeric (Fig. 4A), with a single symmetric peak at around 50 kDa. These four variants completely lost their oligomerization ability *in vitro*, which is consistent with their total loss of activity in the RNP assay.

The remaining four NP variants, R267A, I406S, R422A, and E449A, showed detectable but reduced transcription-replica-

FIG. 3. NP mutants defective in RNP activity possess the ability to interact with RNA and polymerase proteins. (A) Gel shift assay of the RNA binding activities of the defective NP variants. Purified NP was incubated with a 24-nt RNA, and the complexes were analyzed by agarose gel electrophoresis. Bovine serum albumin (BSA) instead of NP was used as a negative control. (B) Coimmunoprecipitation of NP using a Myc-tagged polymerase complex. Wild-type NP was used as a positive control. The Myc-tagged PA plasmid was left out in the negative control. Loading controls are shown at the top. For each variant, “+” refers to the presence of anti-Myc antibodies, while “–” refers to their absence.

tion activity (Fig. 2B). The latter three were eluted in a skewed and broad peak during static light scattering (Fig. 4B). The peaks were of higher molecular weights than wild-type NP. The broad peak also suggests that they are a mixture of different types of oligomers. It is conceivable that some of the oligomeric forms are nonproductive for RNP activity.

Although the R267A mutant retained 45% of the RNP ac-

tivity (Fig. 2C), surprisingly, it appeared as a single monomeric peak by static light scattering (Fig. 4C). This discrepancy is most likely due to differences between the *in vitro* and *in vivo* conditions. To mimic the *in vivo* situation, we synthesized an 1,850-nt-long RNA by *in vitro* transcription. The RNA was mixed with the R267A NP variant and subjected to static light scattering again. An extra peak of higher molecular weight

TABLE 1. Kinetic and affinity constants for NP-RNA interactions^a

Protein	RNP activity	Mean k_a (1/ms) $\pm$ SE	Mean K_d (1/s) $\pm$ SE	Mean K_D (nM) $\pm$ max error
WT	Fully active	$6.00 \times 10^5 \pm 0.05 \times 10^5$	$9.48 \times 10^{-3} \pm 0.15 \times 10^{-3}$	15.8 ± 0.4
I406S	Partially active	$6.21 \times 10^5 \pm 0.07 \times 10^5$	$8.23 \times 10^{-3} \pm 0.16 \times 10^{-3}$	13.3 ± 0.4
R422A	Partially active	$5.40 \times 10^5 \pm 0.07 \times 10^5$	$1.11 \times 10^{-2} \pm 0.02 \times 10^{-2}$	20.6 ± 0.6
E449A	Partially active	$6.39 \times 10^5 \pm 0.08 \times 10^5$	$8.74 \times 10^{-3} \pm 0.19 \times 10^{-3}$	13.7 ± 0.5
P410S	Fully active	$6.59 \times 10^3 \pm 0.09 \times 10^3$	$6.21 \times 10^{-3} \pm 0.09 \times 10^{-3}$	942 ± 27
P419S	Fully active	$7.49 \times 10^3 \pm 0.08 \times 10^3$	$2.83 \times 10^{-3} \pm 0.07 \times 10^{-3}$	378 ± 13
R267A	Partially active	$1.09 \times 10^4 \pm 0.02 \times 10^4$	$8.13 \times 10^{-3} \pm 0.10 \times 10^{-3}$	745 ± 23
E339A	Inactive	$3.40 \times 10^3 \pm 0.06 \times 10^3$	$2.92 \times 10^{-3} \pm 0.09 \times 10^{-3}$	858 ± 42
R416A	Inactive	$3.63 \times 10^3 \pm 0.04 \times 10^3$	$3.54 \times 10^{-3} \pm 0.06 \times 10^{-3}$	975 ± 27
V408S P410S	Inactive	$2.03 \times 10^3 \pm 0.04 \times 10^3$	$1.20 \times 10^{-3} \pm 0.09 \times 10^{-3}$	593 ± 56
L418S P419S	Inactive	$8.25 \times 10^3 \pm 0.07 \times 10^3$	$3.31 \times 10^{-3} \pm 0.05 \times 10^{-3}$	401 ± 9

^a $K_D = k_d/k_a$, where K_D is affinity, k_d is the dissociation rate, and k_a is the association rate. Errors of k_a and k_d represent standard errors; errors of K_D represent the maximum errors from each experiment.

appeared, which resembled the profile of wild-type NP. This was not observed for those variants with undetectable RNP activity (data not shown). This finding suggests that the R267A variant was able to form NP homo-oligomers upon RNA binding, which could explain the RNP activity observed *in vivo*.

NP variants with impaired RNP activity cannot form homo-oligomers *in vivo*. We found that the four NP mutants with undetectable RNP activity (the E339A, R416A, V408S P410S, and L418S P419S mutants) were unable to form oligomers *in vitro* (Fig. 2B and 4A). We were also interested in determining whether they behave similarly *in vivo*. Therefore, C-terminally Myc-tagged WT and E339A, R416A, V408S P410S, and L418S P419S mutant NPs were constructed. Each of these plasmids was cotransfected with their untagged counterparts into 293T cells. Coimmunoprecipitation was performed by using anti-Myc antibodies. NP was detected with anti-NP antibodies. Since the tagged and untagged proteins have distinct molecular weights and can be separated by SDS-PAGE, the presence of untagged NP indicates the homo-oligomerization of NP.

The amount of plasmid used was optimized so that the expression levels of individual proteins were similar (Fig. 5A). The cotransfection of Myc-WT NP and WT NP was used as a positive control. The presence of untagged NP indicates that coimmunoprecipitation was successful. The transfection of WT NP without Myc-WT NP was used as the negative control. No untagged NP was observed if Myc-WT NP was omitted. All four untagged NP variants were detected by Western blotting (Fig. 5B), showing that they were coimmunoprecipitated by their Myc-tagged counterparts. It appears that these four NP variants have the ability to form oligomers *in vivo*. However, NP is an RNA binding protein (7, 28), and the observed NP-NP interaction could therefore have been facilitated by the presence of RNA rather than by the tail loop insertion.

A previous report concluded that the removal of RNA did not alter the oligomeric state of NP (27). Therefore, RNase A was added to the coimmunoprecipitation experiment. If the observed interaction between Myc-tagged and untagged NP variants was due to the tail loop insertion, then RNase A treatment would not disrupt the binding. On the other hand, if the interaction was mediated by RNA, the band of untagged NP should disappear. Wild-type NP was still coimmunoprecipitated with Myc-tagged wild-type NP after RNase A treatment, with the band intensity being similar to that of the

untreated one (Fig. 5B). This finding demonstrated that tail loop-mediated NP homo-oligomerization is not affected by the removal of RNA. The untagged NP bands of the E339A and R416A mutants had reduced intensity but were still detectable, suggesting that their tail loops were largely impaired and could not form homo-oligomers efficiently. The number of oligomers formed may be below the level of detection of the static light scattering experiment (Fig. 4A). The lower bands of the V408S P410S and L418S P419S mutants were barely detectable or totally absent (Fig. 5B, bottom), suggesting that the previously observed interaction (Fig. 5B, top) was due to the presence of RNA and further implying that their tail loops are defective in forming oligomers, which also correlates with their monomeric nature *in vitro* (Fig. 4A).

DISCUSSION

The recent pandemic outbreak of A/H1N1 swine influenza virus has highlighted the fact that the number of inhibitors of influenza virus that are available is limited and that further antivirals are required. Among the influenza viral proteins, NP is crucial for encapsidating RNA and the transcription-replication process (22, 31). In addition, it is highly conserved, and therefore, it is an attractive candidate for inhibitor design (14).

The determination of the H1N1 and H5N1 NP crystal structures has revealed the importance of the tail loop insertion for NP homo-oligomerization (21, 34). In contrast to the NP structures of rabies virus (1) and vesicular stomatitis virus (10), the tail loop of each influenza A virus NP is deeply inserted into the groove of its neighboring NP, generating extensive interactions between the two proteins.

The intersubunit interaction between the tail loop and the insertion groove is mediated solely by salt bridges. Hydrogen bonds between E339 and R416 at the tip region (cluster I1) play a critical role in this event. The mutation of R416A totally abolished RNP activity (Fig. 2B), which is consistent with previous findings by other groups (6, 8, 15). The mutation of E339A also leads to the complete loss of RNP activity (Fig. 2B). Both the E339A and R416A variants were found to be monomeric *in vitro* (Fig. 4A), similar to previous findings by gel filtration analysis (34). These two variants also formed mainly monomers *in vivo* in the absence of RNA (Fig. 5B, bottom). Since the tail loop conformation of the E339A and R416A

FIG. 4. Static light scattering analysis of purified NP variants. Wild-type or mutant NP proteins were subjected to static light scattering analysis using a miniDAWN triangle light scattering detector. The light scattering signal and refractive index were measured during the experiment. Curves are the light scattering signals. Horizontal lines are the calculated native molecular weights (MW) (in thousands) of proteins, which are proportional to the ratio of the light scattering signal to the refractive index. (A) NP variants totally defective in RNP activity are monomeric, while wild-type NP exists as a mixture of trimers and tetramers. (B) Partially defective NP variants (I406S, R422A, and E449A) gave a skewed and broad peak and higher native molecular weights than wild-type NP, indicating that they exist as a mixture of different types of oligomers. (C) R267A exists as monomer in the absence of RNA. The addition of an 1,850-nt RNA to both the wild type and the R267A variant led to a high-molecular-weight peak on the left, suggesting the formation of NP-RNA complexes. Errors of native molecular weight were calculated from the percent error of the experiments.

FIG. 5. Inability of NP mutants defective in RNP activity to form homo-oligomers *in vivo*. (A) Plasmids expressing untagged NP and C-terminally Myc-tagged NP of the respective mutants were transfected into 293T cells. Cell lysates were analyzed by Western blotting (WB) using polyclonal anti-NP antibodies after 48 h of incubation. The negative control contained untagged wild-type NP alone. (B) Coimmunoprecipitation of untagged NP variants with the corresponding Myc-tagged NP in the presence or absence of RNA. RNase A was used to remove RNA during coimmunoprecipitation. For each variant, “+” refers to the presence of anti-Myc antibodies, while “-” refers to its absence.

variants was not disrupted, their tail loops may still be able to gain access to the groove, although this interaction is not maintained, hence giving rise to the low degree of interaction observed (Fig. 5B, bottom). The similar behaviors of these two protein mutants strengthen the conclusion about the importance of this ion pair. Analysis of 2,500 influenza A virus NP sequences also revealed that both residues are strictly conserved (Table 2). Sequence alignment shows that both residues in influenza A virus NP are identical to the corresponding residues in influenza B and C virus NPs (Table 2). This finding further supports the functional importance of such an interaction and implies that the interaction is conserved throughout all influenza virus NPs. In addition to cluster I1, cluster I3, which is located at the stem region, also exerts a supportive role in the intersubunit interaction. The mutation of R267A, the central residue of cluster I3, resulted in a reduction in RNP activity (Fig. 2B). The R267A variant also lost its ability to form oligomers *in vitro*, which was restored in the presence of RNA (Fig. 4C). Mutations at clusters I2 and I4 did not show significant reductions in RNP activity and therefore are unlikely to be important for NP homo-oligomerization.

In contrast to intersubunit interactions, the tail loop conformation of NP is maintained mainly through hydrophobic interactions at the stem region. Four nonpolar residues, V408, P410, L418, and P419 (cluster T2), form a zigzag pattern of interaction. The V408S P410S and L418S P419S double-point mutants showed no detectable RNP activity (Fig. 2B). Both protein variants were monomeric *in vitro* (Fig. 4A) and *in vivo* (Fig. 5B, bottom) in the absence of RNA. Interestingly, single-point mutants of these four nonpolar residues all showed wild-type RNP activity. This finding suggests a “backup” mechanism of tail loop maintenance. Normal RNP activities of the V408S

and L418S mutants (Fig. 2B) indicate that either the P410-L418 or the V408-P419 interaction is sufficient for the conformational maintenance of the tail loop. The presence of two pairs of nonpolar residues could maintain the tail loop conformation and possibly sustain virus survival even if one of the four residues is mutated. This can also be observed naturally, as V408 is polymorphic (can be either I or T) in influenza A virus NP, although the other three residues remain conserved (Table 2). Influenza B virus NP may also use hydrophobic interactions to maintain the stem of the tail loop but without the backup mechanism. P410 is identical, while L418 is replaced with an I in influenza B virus NP. The P410-I418 interaction is conserved in influenza B virus NP, and one hydrophobic interaction might be adequate for tail loop maintenance. However, the corresponding hydrophobic residues are not present in influenza C virus NP, suggesting that it may use another mechanism for such tail loop maintenance. Cluster T1 appears to play a role in tail loop maintenance, as the mutation of R416A abolished RNP activity (Fig. 2B). However, R416 also takes part in cluster I1 for intersubunit interactions. Since the V408S P410S and L418S P419S double-point mutants showed no detectable RNP activity in the presence of R416, R416 by itself is not sufficient to maintain the tail loop conformation.

In addition to the intersubunit interaction and tail loop maintenance, the stabilization of the NP homo-oligomer is necessary for a fully functional RNP. Stabilizing power is contributed by the base region. The tail loop itself is stabilized by cluster T3, centered at I406, while the intersubunit interaction is stabilized by cluster I5, with the R422-E449 ion pair. Protein variants of I406S, R422A, and E449A all displayed abnormal oligomeric states by static light scattering (Fig. 4B). Abnormal oligomerization is likely to result in the reduction of RNP activity (Fig. 2B). Although I406 is polymorphic in influenza A virus NP (can be either V or T), the corresponding residues in influenza B and C virus NPs are identical, further suggesting the functional importance of this residue. The R422-E449 interaction is also partially conserved in influenza B and C virus NPs (Table 2). This ion pair was previously reported to be the molecular contact for trimerization (21), hence implying that the modes of interaction of trimeric NP in the crystal structure and the oligomeric NP inside cells are highly similar.

In conclusion, we have identified three factors that govern the NP homo-oligomerization event: (i) interaction between

TABLE 2. Variability of residues important for NP homo-oligomerization^a

Influenza A virus NP		Influenza B virus NP		Influenza C virus NP	
Position	Residue	Position	Residue	Position	Residue
267	R	325	R	272	K
339	E	395	E	354	E
406	I	462	I	416	I
408	I/V	464	C	418	N
410	P	466	P	420	A
416	R	472	R	426	R
418	L	474	I	428	I
419	P	475	A	429	T
422	R/K	478	K	432	V
449	E	505	N	459	E

^a Corresponding residues of each type of NP are listed. Conserved residues are shown in boldface type.

the tail loop and the insertion groove, (ii) maintenance of the tail loop conformation, and (iii) stabilization of the NP homo-oligomer. The first two factors are crucial for NP-NP interactions as well as the transcription-replication activity of the viral RNP. The third factor plays a minor but supportive role in the process, since destabilizing the homo-oligomer results in reduced but not totally abolished RNP activity. In general, the tip and the stem regions of the tail loop are more important than the base region.

The data presented in this report also give insights into the sequential events of the tail loop insertion process. First, it is essential for the tail loop to maintain a proper conformation. The disruption of the conformation by the V408S P410S or L418S P419S mutant led to a greatly reduced interaction in the coimmunoprecipitation study (Fig. 5B, bottom), possibly because the disrupted tail loop could not access the insertion groove. Second, the intersubunit interactions between the tail loop and the insertion groove are crucial. Mutations at R416A and E339A did not affect the tail loop conformation and therefore could access the insertion groove; nevertheless, they resulted in minimal NP-NP interactions (Fig. 5B) and possessed no detectable RNP activities due to the inability of the intersubunit interaction. Third, the NP homo-oligomer is further stabilized by intra- and intersubunit interactions of clusters T3 and I5 at the base of the tail loop.

We have identified the E339-R416 ion pair and the hydrophobic interactions between V408, P410, L418, and P419 as the essential molecular contacts for the NP homo-oligomerization process. Based on these findings, inhibitors that can disrupt either the ion pair or the hydrophobic interactions, and therefore disrupt NP homo-oligomerization, should restrict the transcription-replication processes of the virus. Such antiviral drugs would inhibit influenza virus and could relieve the symptoms of influenza in patients. This strategy could be used in the future to design novel influenza virus inhibitors.

ACKNOWLEDGMENTS

This work was supported by a General Research Fund grant (grant CUHK 472808) from the Research Grants Council of Hong Kong SAR and by the Medical Research Council (grant G0700848), United Kingdom.

REFERENCES

- Albertini, A. A., A. K. Wernimont, T. Muziol, R. B. Ravelli, C. R. Clapier, G. Schoehn, W. Weissenhorn, and R. W. Ruigrok. 2006. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **313**:360–363.
- Area, E., J. Martin-Benito, P. Gastaminza, E. Torreira, J. M. Valpuesta, J. L. Carrascosa, and J. Ortin. 2004. 3D structure of the influenza virus polymerase complex: localization of subunit domains. *Proc. Natl. Acad. Sci. U. S. A.* **101**:308–313.
- Baudin, F., C. Bach, S. Cusack, and R. W. Ruigrok. 1994. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.* **13**:3158–3165.
- Bishop, D. H., J. F. Objieski, and R. W. Simpson. 1971. Transcription of the influenza ribonucleic acid genome by a virion polymerase. I. Optimal conditions for *in vitro* activity of the ribonucleic acid-dependent ribonucleic acid polymerase. *J. Virol.* **8**:66–73.
- Biswas, S. K., P. L. Boutz, and D. P. Nayak. 1998. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J. Virol.* **72**:5493–5501.
- Coloma, R., J. M. Valpuesta, R. Arranz, J. L. Carrascosa, J. Ortin, and J. Martin-Benito. 2009. The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog.* **5**:e1000491.
- Compans, R. W., J. Content, and P. H. Duesberg. 1972. Structure of the ribonucleoprotein of influenza virus. *J. Virol.* **10**:795–800.
- Elton, D., E. Medcalf, K. Bishop, and P. Digard. 1999. Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology* **260**:190–200.
- Fodor, E., M. Crow, L. J. Mingay, T. Deng, J. Sharps, P. Fichter, and G. G. Brownlee. 2002. A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *J. Virol.* **76**:8989–9001.
- Green, T. J., X. Zhang, G. W. Wertz, and M. Luo. 2006. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* **313**:357–360.
- Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal, and P. Palese. 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.
- Klump, K., R. W. Ruigrok, and F. Baudin. 1997. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* **16**:1248–1257.
- Krug, R. M., M. Ueda, and P. Palese. 1975. Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. *J. Virol.* **16**:790–796.
- Krug, R. M., and J. M. Aramini. 2009. Emerging antiviral targets for influenza A virus. *Trends Pharmacol. Sci.* **30**:269–277.
- Li, Z., T. Watanabe, M. Hatta, S. Watanabe, A. Nanbo, M. Ozawa, S. Kakugawa, M. Shimojima, S. Yamada, G. Neumann, and Y. Kawaoka. 2009. Mutational analysis of conserved amino acids in the influenza A virus nucleoprotein. *J. Virol.* **83**:4153–4162.
- Markushin, S. G., and Y. Z. Ghendon. 1984. Studies of fowl plague virus temperature-sensitive mutants with defects in synthesis of virion RNA. *J. Gen. Virol.* **65**:559–575.
- Martin-Benito, J., E. Area, J. Ortega, O. Llorca, J. M. Valpuesta, J. L. Carrascosa, and J. Ortin. 2001. Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep.* **2**:313–317.
- Medcalf, L., E. Poole, D. Elton, and P. Digard. 1999. Temperature-sensitive lesions in two influenza A viruses defective for replicative transcription disrupt RNA binding by the nucleoprotein. *J. Virol.* **73**:7349–7356.
- Mena, I., E. Jambrina, C. Albo, B. Perales, J. Ortin, M. Arrese, D. Vallejo, and A. Portela. 1999. Mutational analysis of influenza A virus nucleoprotein: identification of mutations that affect RNA replication. *J. Virol.* **73**:1186–1194.
- Murti, K. G., R. G. Webster, and I. M. Jones. 1988. Localization of RNA polymerases on influenza viral ribonucleoproteins by immunogold labeling. *Virology* **164**:562–566.
- Ng, A. K., H. Zhang, K. Tan, Z. Li, J. H. Liu, P. K. Chan, S. M. Li, W. Y. Chan, S. W. Au, A. Joachimiak, T. Walz, J. H. Wang, and P. C. Shaw. 2008. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* **22**:3638–3647.
- Ng, A. K., J. H. Wang, and P. C. Shaw. 2009. Structure and sequence analysis of influenza A virus nucleoprotein. *Sci. China C Life Sci.* **52**:439–449.
- Noda, T., H. Sagara, A. Yen, A. Takada, H. Kida, R. H. Cheng, and Y. Kawaoka. 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**:490–492.
- Pons, M. W., I. T. Schulze, G. K. Hirst, and R. Hauser. 1969. Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* **39**:250–259.
- Portela, A., and P. Digard. 2002. The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J. Gen. Virol.* **83**:723–734.
- Prokudina-Kantorovich, E. N., and N. P. Semenova. 1996. Intracellular oligomerization of influenza virus nucleoprotein. *Virology* **223**:51–56.
- Ruigrok, R. W., and F. Baudin. 1995. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J. Gen. Virol.* **76**:1009–1014.
- Scholtissek, C., and H. Becht. 1971. Binding of ribonucleic acids to the RNP-antigen protein of influenza viruses. *J. Gen. Virol.* **10**:11–16.
- Scholtissek, C., and B. R. Murphy. 1978. Host range mutants of an influenza A virus. *Arch. Virol.* **58**:323–333.
- Shapiro, G. I., and R. M. Krug. 1988. Influenza virus RNA replication *in vitro*: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. *J. Virol.* **62**:2285–2290.
- Shu, L. L., W. J. Bean, and R. G. Webster. 1993. Analysis of the evolution and variation of the human influenza A virus nucleoprotein gene from 1933 to 1990. *J. Virol.* **67**:2723–2729.
- Thierry, F., and O. Danos. 1982. Use of specific single stranded DNA probes cloned in M13 to study the RNA synthesis of four temperature-sensitive mutants of HK/68 influenza virus. *Nucleic Acids Res.* **10**:2925–2938.
- Vreede, F. T., T. E. Jung, and G. G. Brownlee. 2004. Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *J. Virol.* **78**:9568–9572.
- Ye, Q., R. M. Krug, and Y. J. Tao. 2006. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**:1078–1082.