Genetic stability of genome-scale deoptimized RNA virus vaccine candidates under selective pressure

Cyril Le Nouën[a,1], Thomas McCarty[a], Michael Brown[b], Melissa Laird Smith[b,2], Roberto Lleras[b], Michael A. Dolan[c], Masfique Mehedi[a], Lijuan Yang[a], Cindy Luongo[a], Bo Liang[a], Shirin Munir[a], Joshua M. DiNapoli[a,3], Steffen Mueller[d], Eckard Wimmer[e,1], Peter L. Collins[a], and Ursula J. Buchholz[a]

[a]RNA Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892; [b]Pacific Biosciences Inc., Menlo Park, CA 94025; [c]Bioinformatics and Computational Biosciences Branch, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892; [d]Codagenix Inc., Stony Brook, NY 11790; and [e]Department of Molecular Genetics and Microbiology, Stony Brook University, Stony Brook, NY 11794

Contributed by Eckard Wimmer, November 22, 2016 (sent for review September 12, 2016; reviewed by Mark E. Peeples and Fernando P. Polack)

Recoding viral genomes by numerous synonymous but suboptimal substitutions provides live attenuated vaccine candidates. These vaccine candidates should have a low risk of deattenuation because of the many changes involved. However, their genetic stability under selective pressure is largely unknown. We evaluated phenotypic reversion of deoptimized human respiratory syncytial virus (RSV) vaccine candidates in the context of strong selective pressure. Codon pair deoptimized (CPD) versions of RSV were attenuated and temperature-sensitive. During serial passage at progressively increasing temperature, a CPD RSV containing 2,692 synonymous mutations in 9 of 11 ORFs did not lose temperature sensitivity, remained genetically stable, and was restricted at temperatures of 34 °C/35 °C and above. However, a CPD RSV containing 1,378 synonymous mutations solely in the polymerase L ORF quickly lost substantial attenuation. Comprehensive sequence analysis of virus populations identified many different potentially deattenuating mutations in the L ORF as well as, surprisingly, many appearing in other ORFs. Phenotypic analysis revealed that either of two competing mutations in the virus transcription antitermination factor M2-1, outside of the CPD area, substantially reversed defective transcription of the CPD L gene and substantially restored virus fitness in vitro and in case of one of these two mutations, also in vivo. Paradoxically, the introduction into Min L of one mutation each in the M2-1, N, P, and L proteins resulted in a virus with increased attenuation in vivo but increased immunogenicity. Thus, in addition to providing insights on the adaptability of genome-scale deoptimized RNA viruses, stability studies can yield improved synthetic RNA virus vaccine candidates.

negative-strand RNA virus | respiratory syncytial virus | live attenuated vaccine | codon pair deoptimization | genetic stability

The availability and affordability of large-scale custom DNA synthesis opened the new field of synthetic biology (1). The combined approach of sequence design and synthetic biology allows the generation of DNA molecules with extensive targeted modifications. Synonymous genome recoding, in which one or more ORFs of a microbial pathogen are modified at the nucleotide level without affecting amino acid coding, currently is being widely evaluated to reduce pathogen fitness and create potential live attenuated vaccines, particularly for RNA viruses (reviewed in ref. 2) (3–7). The main strategies for attenuation by synonymous genome recoding are codon deoptimization (CD), codon pair deoptimization (CPD), and increasing the dinucleotide CpG and UpA content (which is usually the result of CD and CPD) (2).

The mechanisms of attenuation by these strategies are currently under intensive research. It has been suggested that the primary effect of CD and CPD is to reduce translation efficiency of pathogen mRNAs, thereby providing attenuation (8). Effects on mRNA stability also can be a factor (9). In addition, recent studies suggested that codon pair bias may be a direct consequence of dinucleotide bias (10), suggesting that the increase in the frequency of CpG and UpA dinucleotides that frequently occurs with CPD may cause attenuation by inducing stronger innate immune responses (11). Regardless of the mechanism of attenuation caused by CPD, the genetic stability of live attenuated virus vaccines is paramount to avoid deattenuation, and thus, it is important to evaluate the genetic and phenotypic stability of such vaccine candidates in detail.

Deoptimized viruses genomes contain dozens to thousands of silent nucleotide mutations in one or more ORFs. Presumably, attenuation is based on the sum of many individual mutations. This mutation multiplicity is expected to confer stability against substantial deattenuation, because the high number of mutations would present a significant barrier against reversion to virulence. In principle, on the background of thousands of attenuating mutations, any single-site reversion should yield only a minuscule selective advantage. The most likely path to reversion imaginable under this model is the progressive accumulation of many individual mutations, providing for a slow progression of deattenuation (12–14).

To date, genetic stability studies of large-scale deoptimized viruses have shown that deattenuation, indeed, seems to be low, suggesting that these viruses are genetically stable (6, 8, 15–21). However, an important limitation of these studies is that the deoptimized viruses generally have not been subjected to strong selective pressure that

Significance

Recoding viral genomes by numerous synonymous substitutions provided live attenuated vaccine candidates predicted to have a low risk of reversion. However, their stability under selective pressure was largely unknown. We evaluated the phenotypic reversion of representative genome-scale deoptimized human respiratory syncytial virus (RSV) vaccine candidates in the context of strong selective pressure. We found that a virus bearing a deoptimized L-polymerase ORF evolved to escape temperature sensitivity restriction by mutations in L and multiple other proteins. Additional analysis revealed that single mutations in the M2-1 ORF were able to substantially escape the restriction imposed by the deoptimized polymerase. Based on this information, we generated a stable deoptimized RSV vaccine candidate with improved attenuation and immunogenicity suitable for additional development.

Author contributions: C.L.N., P.L.C., and U.J.B. designed research; C.L.N., M.M., L.Y., C.L., and B.L. performed research; T.M., M.B., M.L.S., R.L., S. Munir, J.M.D., S. Mueller, and E.W. contributed new reagents/analytic tools; C.L.N., T.M., M.B., and M.A.D. analyzed data; S. Mueller and E.W. contributed to the design of the recoded respiratory syncytial virus genes; and C.L.N., P.L.C., and U.J.B. wrote the paper.

Reviewers: M.E.P., The Research Institute at Nationwide Children's Hospital; and F.P.P., Vanderbilt University.

Conflict of interest statement: C.L.N., J.M.D., S. Mueller, E.W., P.L.C., and U.J.B. are coinventors on a patent application for the development of respiratory syncytial virus vaccines by codon pair deoptimization.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. Email: lenouenc@niaid.nih.gov or Eckard.Wimmer@stonybrook.edu.

[2]Present address: Department of Genetics and Genomic Sciences, Icahn Institute for Genomics and Multiscale Biology, New York, NY 10029-5674.

[3]Present address: North America Research Department, Sanofi Pasteur Biologics LLC, Cambridge, MA 02139.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619242114/-/DCSupplemental.

would favor the outgrowth of viruses with deattenuating mutations. This study addresses this gap by evaluating phenotypic reversion of representative synthetic human respiratory syncytial virus (RSV) vaccine candidates in strong selective pressure.

RSV, a negative-strand RNA virus of the *Pneumoviridae* family, is a major worldwide human respiratory pathogen, particularly of infants and young children. A vaccine or antiviral drugs suitable for routine use are not available and are a high priority. A live attenuated vaccine is a strategy of choice, because it obviates RSV disease enhancement that is associated with inactivated and subunit RSV vaccines. The RSV genome consists of a single-stranded negative sense 15.2-kb RNA and has 10 genes in the order 3′-NS1-NS2-N-P-M-SH-G-F-M2-L-5′, preceded by a short leader region and followed by a short trailer region. The M2 mRNA encodes two separate proteins, M2-1 and M2-2, from overlapping ORFs. The genes are each flanked by short gene-start and gene-end transcription signals and transcribed as individual mRNAs by sequential transcription initiating at a single promoter in the leader region.

We recently generated four CPD RSVs that contain increasing numbers of deoptimized ORFs and exhibit a range of increasing attenuation in vitro and in vivo (22), providing live attenuated RSV vaccine candidates. Surprisingly, these viruses were temperature-sensitive. This temperature sensitivity provided a means to apply selective pressure that we used to investigate the genetic stability of these CPD RSVs. The combination of extensive replication and incrementally increasing temperatures should favor outgrowth of viruses bearing deattenuating mutations allowing for identification of genetic changes. Because a temperature gradient occurs along the human respiratory tract (23) and because respiratory virus infections progress from the upper to the lower respiratory tract, temperature stress is a relevant model to evaluate the genetic stability of live attenuated respiratory virus vaccine candidates. In this study, we applied temperature stress on two CPD RSVs that exhibited different levels of CPD and analyzed the temporal dynamics of the synthetic virus populations.

Results

CPD of Multiple RSV Genes Yielded a Very Stable Ts Phenotype Restricted to Replication at 32 °C to 34 °C. Min full-length clone (FLC) is a previously described mutant in which 9 of 11 RSV ORFs (except only M2-1 and M2-2) were CPD, resulting in a total of 2,692 silent mutations (Fig. 1*A*) (22). Min FLC is highly temperature-sensitive, with a shutoff temperature (T_{SH}) of 35 °C for plaque formation, whereas WT recombinant RSV (rRSV) readily forms plaques at 40 °C. The T_{SH} is defined as the lowest restrictive temperature at which there is a reduction in virus titer compared with the permissive temperature of 32 °C that is 100-fold or greater than that of WT RSV at this temperature.

To investigate Min FLC stability, we used a temperature stress test, representing a surrogate model for genetic stability during virus replication and spread from the cooler upper to the warmer lower respiratory tract. Ten replicate flasks of Vero cells were infected with Min FLC and subjected to serial passage at progressively increasing temperatures for a total of 18 passage stages, representing 7 mo of continuous culture. Two additional replicate flasks were infected and passaged in parallel at the permissive temperature of 32 °C as controls (Fig. 1 *B* and *C*).

At 32 °C, Min FLC replicated consistently to titers of 10^6–10^7 pfu/mL (Fig. 1*B*). Deep sequencing of the complete genomes of the two control lineages after 18 passages revealed only low-level sporadic mutations (Fig. S1), showing that Min FLC was

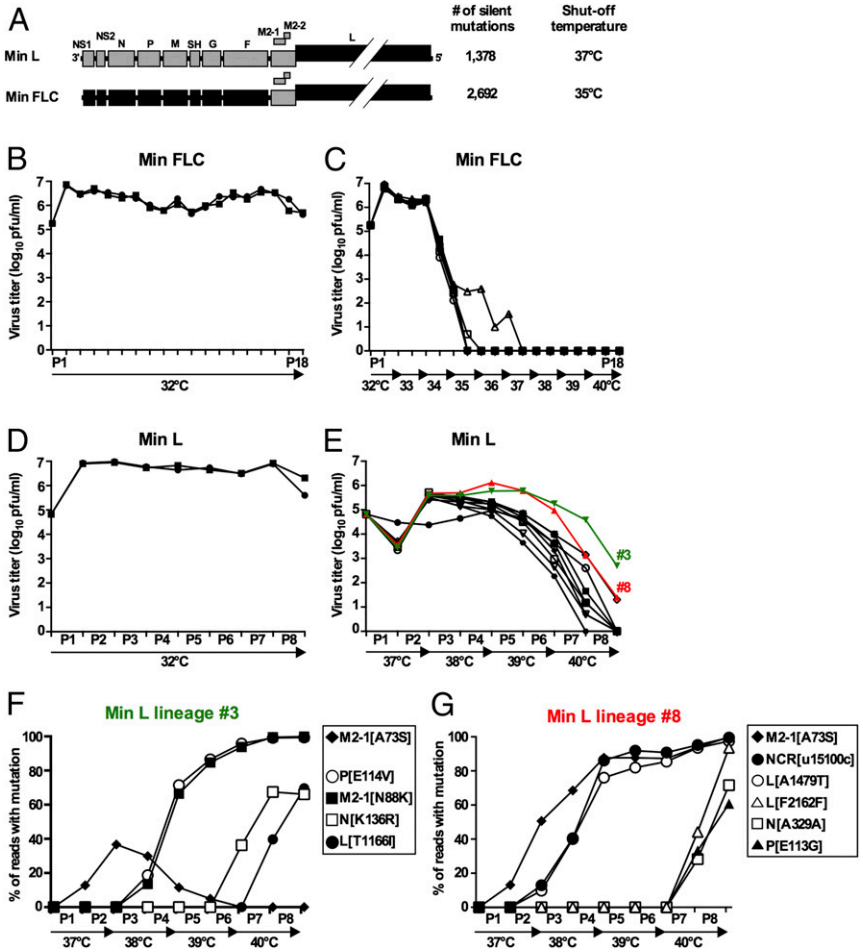

Fig. 1. Min FLC was phenotypically stable during a temperature stress test, but Min L was not. (*A*) Gene maps of Min L and Min FLC showing ORFs that are WT (gray) or CPD (black). The number of introduced mutations in each virus and the T_{SH} in Vero cells are indicated. (*B*–*E*) Incubation temperature and virus yield at each passage level during serial passage in temperature stress tests. (*C*) Replicate cultures of Vero cells in T25 flasks were infected with the indicated virus at MOI of 0.1; when the viral cytopathic effect was extensive or when cells started to detach (for passages of Min FLC at 37 °C and beyond), flasks were harvested, and clarified culture fluids were passaged 1:5 to a fresh flask. Each starting replicate flask initiated an independent serial passage (lineage). Aliquots of clarified culture fluids were frozen for titration and sequence analysis (22). (*B* and *C*) Temperature stress test of Min FLC. (*B*) Two control flasks inoculated with Min FLC were passaged 18 times at the permissive temperature of 32 °C. (*C*) Ten additional replicates were passaged from 32 °C to 40 °C with two passages at each temperature. (*D* and *E*) Temperature stress test of Min L. (*D*) Two control flasks inoculated with Min L were passaged eight times at the permissive temperature of 32 °C. (*E*) Ten additional replicates were passaged from 37 °C to 40 °C with two passages at each temperature. Lineages 3 and 8 are shown in green and red, respectively. (*F* and *G*) Accumulation of most abundant mutations (>30% of the reads in at least one passage) in lineages (*F*) 3 and (*G*) 8 during the passage series determined by deep sequencing (Tables S2 and S3 show detailed data).

genetically stable under permissive conditions. In the flasks incubated at increasing temperature, Min FLC replicated efficiently at 32 °C and 33 °C (10^6–10^7 pfu/mL) (Fig. 1*C*). However, after the first passage at 34 °C [passage 5 (P5)], virus replication was reduced 200-fold in all 10 lineages, and at the end of the second passage at 35 °C (P8), virus was undetectable in 9 lineages. In the 10th lineage, no virus was detected at the end of the first passage at 37 °C (P11). In contrast, as noted, WT rRSV exhibits no growth restriction at temperatures up to at least 40 °C.

Thus, Min FLC was highly restricted if not inactive at temperatures above 34 °C to 35 °C (the latter being its T_{SH}). Consequently, Min FLC cannot escape its Ts phenotype and is phenotypically stable under stress conditions. Sequencing was not performed on Min FLC specimens passed under increasingly restrictive temperatures because of the rapid decrease in titers. These results fulfilled the expectation of phenotypic stability for a CPD virus.

Temperature Stress on the Min L Virus Promoted the Emergence of Multiple Mutations in Multiple Genes. We also evaluated a virus in which CPD was less extensive. Specifically, the previously described virus called Min L (22), in which the L ORF alone (representing 48% of the aggregate RSV ORFs) was CPD, resulted in 1,378 silent mutations (21.2% of nucleotides of this ORF; 51% as many changes as in Min FLC). Min L has a T_{SH} of 37 °C. Ten replicate flasks were infected with Min L and passaged serially at progressively increasing temperatures for a total of eight passages, corresponding to 2 mo of continuous culture, and two additional replicate flasks were infected and passaged in parallel at 32 °C as controls (Fig. 1 *D* and *E*).

As expected, Min L replicated efficiently (10^7 pfu/mL) at each passage at 32 °C (Fig. 1*D*). Sequence analyses of RNA from the control lineages at P6 by deep sequencing (Fig. S2) and P8 by Sanger sequencing revealed only sporadic, low-level mutations. In 10 lineages passaged at increasing temperature, the titers of Min L in nine flasks were decreased by about 20-fold at the end of P1 (37 °C) (Fig. 1*E*). However, during the second passage at 37 °C, titers in the same nine lineages increased by about 200-fold, suggesting that selection and outgrowth of temperature-adapted mutants were already occurring. After P3 (38 °C), virus titers in all 10 lineages decreased steadily: at P8 (second passage at 40 °C), virus was undetectable in 7 lineages, whereas in 2 other lineages, titers were very low (20 pfu/mL each). The remaining lineage (lineage 3) (colored green in Fig. 1*E*) had a titer of 500 pfu/mL. Thus, the various Min L lineages seemed to undergo a partial loss of the temperature sensitivity phenotype but ultimately, were strongly restricted at 40 °C.

For each of 10 lineages passaged at increasing temperature, whole-genome deep sequencing was performed at the end of P6 (the second passage at 39 °C), when virus replication was still detectable in each lineage. Mutations present in ≥45% of the sequencing reads are shown in Table 1. Remarkably, many of these prominent mutations were in genes not subjected to CPD. Specifically, of these 23 prominent mutations, 21 were distributed among six ORFs (P, M, SH, G, M2-1, and L), and 2 were in extragenic regions. Of 23 mutations, 11 (48%) and 5 (22%) occurred in the M2-1 and L ORFs, respectively. Of 21 mutations present in ORFs, all but one were missense mutations, suggesting a bias for amino acid change. Some mutations were common to several lineages. Specifically, the mutation [A73S] in the antitermination transcription factor M2-1 was prominent in 8 of 10 lineages. Another M2-1 mutation (N88K) and a mutation in L (A1479T) were prominent in two lineages. M2-1 was the only gene to have one or more prominent mutations in every lineage.

Table S1 shows mutations that were present in ≥5% of the reads from the P6 specimens from the same experiment. With this lower cutoff, many more mutations were evident in every gene except NS2. Similar to the prominent mutations that were shown in Table 1, these less prominent mutations were mostly missense mutations. In the CPD L ORF, in which 21.2% of codons had been modified by CPD, 17 of the total of 31 (55%) mutations involved a nucleotide or a codon that had been modified during CPD (Table S1), suggesting that mutations frequently occurred in a codon modified by CPD. However, as shown in Table S1, only a single mutation represented a reversion to the original codon.

We also performed whole-genome deep sequencing analysis to evaluate the temporal appearance of mutations in the full-passage

Table 1. Mutations detected in individual lineages of Min L at the end of P6 (second passage at 39 °C) of the temperature stress test present at ≥45% frequency

Gene	Nucleotide mutation	Amino acid mutation	Percentage of reads with mutation in indicated lineage no.*									
			1	2	3	4	5	6	7	8	9	10
Intergenic NS2-N	g1123a	/							85			
P	a2687u	E114V			96							
M	u3798a	N179K		61								
SH	c4369a	H22Q									81	
SH	c4387g	I28M									71	
G	a5384g	E232 (silent)						47				
M2-1	g7823u	A73S	99	93		61	48	83	63	87	57	
M2-1	c7870a	N88K			94							96
M2-1	a8013g	E136G						48				
L	u10548c†	Y684H†	97									
L	u10797c†,‡	S767P†,‡										82
L	g12933a	A1479T		63						85		
L	a13783c†	Y1762S†						83				
5′ Extragenic (trailer)	u15100c	/								75		

"/" indicates that the amino acid mutation is not applicable for this particular mutation as the given mutation is localized in a nontranslated region.

*Percentage of reads with the indicated mutation; only mutations present in ≥45% of the reads are shown. Nucleotide numbering is based on RSV sequence M74568 (biological WT RSV strain A2). Mutations present in ≥5% of reads from this same experiment are shown in Table S1.

†Mutation involving a codon that had been changed as part of CPD of the L ORF.

‡Mutation involving a nucleotide position that had been changed as part of CPD of the L ORF.

series of lineages 3 and 8, which were of interest, because they maintained the highest titers during the stress test (Fig. 1*E*) and thus, have the greatest deattenuation in vitro. The appearance and frequency of the more abundant mutations are shown graphically in Fig. 1 *F*, lineage 3 and *G*, lineage 8. A more detailed listing of the mutations is shown in Tables S2 and S3.

In both lineages, a single mutation ([A73S] in M2-1) appeared at P1 (13% of each lineage) and then, increased at P2 (37 and 51% in lineages 3 and 8, respectively). From P2, the two lineages went into different evolutionary trajectories. In lineage 3, between P2 and P3, whereas the frequency of the M2-1 mutation [A73S] started to decline (30%), 10 other mutations in M2-1 appeared and constituted ~15–30% of the population (Table S2). One of these M2-1 mutations, namely [N88K], became abundant (66%) in P4, closely concurrent with an equally abundant (71%) mutation [E114V] in P4 (Fig. 1*F*). The other M2-1 mutations declined and were undetectable beyond P5, suggestive of a selective sweep (24). Two additional prominent mutations were acquired at P6 (N[K136R]) and P7 (L[T1166I]). In lineage 8, mutation [A73S] in M2-1 was fixed at P4 (88%). At P2, two additional mutations (in the 5′ trailer region and L) were acquired, and they became prominent and fixed by the end of P4. After the first passage at 40 °C (P7), some additional mutations were acquired, three of which became prominent by the end of P8: one silent in L, one silent in N, and one in P[E113G].

Two Mutations in the Antitermination Transcription Factor M2-1 Are Prominent but Incompatible. All 10 lineages at P6 had either M2-1 mutation [A73S] or [N88K] (Fig. 2*A* and Table 1). Thus, these two M2-1 mutations seemed to segregate. In addition, the disappearance of the [A73S] mutation during passage series of lineage 3 coincided with the appearance and increase of [N88K] until the latter was present in the complete population (Fig. 1*F*). We reevaluated the deep sequencing results of lineage 3, scoring only those reads that spanned both positions 73 and 88 in M2-1, thus providing a linkage analysis. At P3 and P4, only 1% of the reads had both mutations (Fig. 2*B*), suggesting that these two mutations in M2-1 are incompatible in the same genome and thus, constitute two separate virus populations.

To further characterize the dynamics of the main virus populations in lineage 3, we investigated linkage of the major mutations that appeared during the first four passages using Pacific Biosciences Inc. (PacBio) long-read, single-molecule sequencing, which provided complete reads of an entire 8.2-kb region from the 3′ genome end to the middle of the M2-2 ORF. This analysis showed that the first four passages contained four major virus subpopulations (Fig. 2*C*). One was the original Min L virus, which progressively decreased with passage. Another subpopulation that carried the M2-1 mutation [A73S] alone peaked at P2 and almost disappeared in P4. Another carried seven mutations in M2-1 (three synonymous and four nonsynonymous) that appeared together at P2, reached a maximum at P3 (about 20%), and then, disappeared. Finally, the fourth subpopulation contained the P[E114V] and M2-1[N88K] mutations that appeared together at P3 and became prominent at P4.

The Two Incompatible Mutations in the M2-1 Protein Confer Most of the Reduced Temperature Sensitivity and Increased Replication of the Min L Derivatives in Vitro. We next sought direct identification of mutation(s) responsible for the loss of temperature sensitivity of Min L. This investigation was performed by introducing into Min L, individually and in combinations, major mutations that had been identified in lineage 3, namely N[K136R], P[E114V], M2-1[N88K],

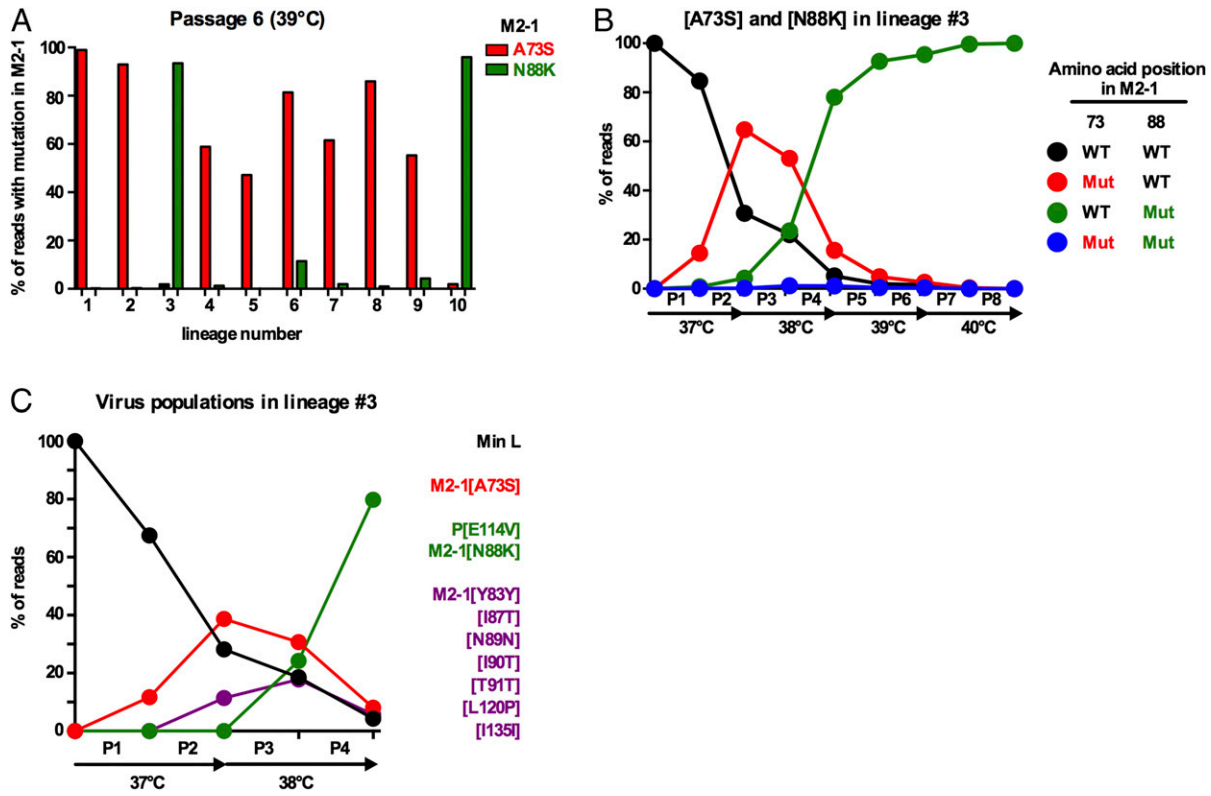

Fig. 2. M2-1 mutations [A73S] and [N88K] segregated into different viral subpopulations. (*A*) Percentage of deep sequencing reads that contained M2-1 mutation [A73S] (red bars) or [N88K] (green bars) at P6 (the second passage at 39 °C) of each of 10 lineages from the experiment in Fig. 1*E*. (*B*) Lack of linkage between M2-1 mutations [A73S] and [N88K] illustrated by the percentage of deep sequencing reads that contained the indicated combinations of assignments at codons 73 (WT vs. [A73S]) and 88 (WT vs. [N88K]) in the same read based on reads from the experiment in Fig. 1*F* that spanned both codons. (*C*) Extent of linkage between M2-1 mutations [A73S] and [N88K] and other mutations during the first four passages of lineage 3 (Fig. 1 *E* and *F*) determined by PacBio sequencing of continuous reads corresponding to an 8.2-kb region of the RSV genome from the 3′ end to the middle of the M2-2 ORF. Four major virus subpopulations were identified, and mutations that are linked on the same genomes are indicated by an identical color.

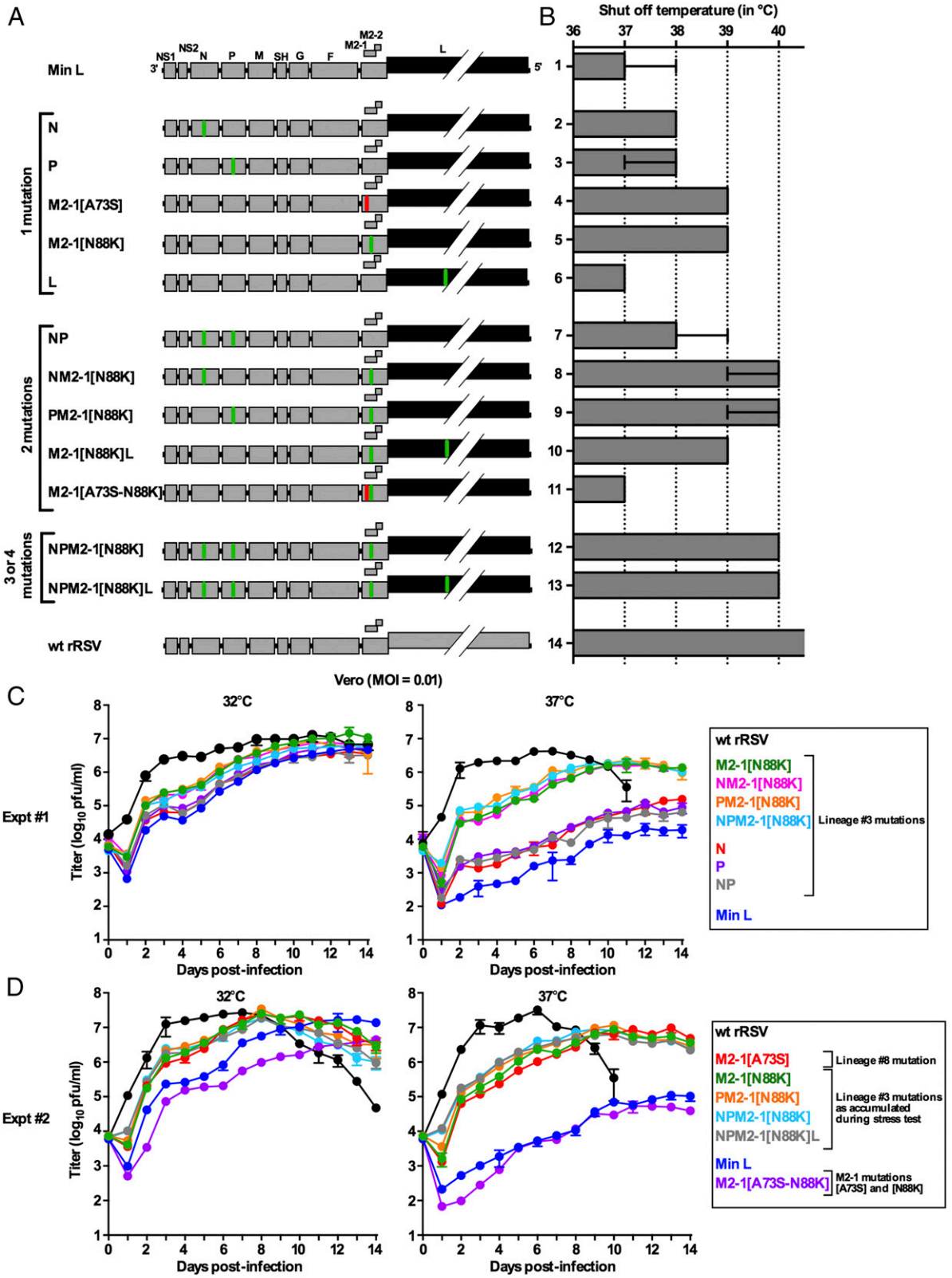

Fig. 3. Effects of specific mutations on the temperature sensitivity and in vitro replication of Min L derivatives. Five major mutations identified in lineage 3 (N [K136R], P[E114V], L[T1166I], M2-1[N88K], and M2-1[A73A]) (Fig. 1*F*) were introduced individually and in combinations by site-directed mutagenesis and reverse genetics into Min L for phenotypic analysis. The Min L-derived viruses were named based on the gene names bearing the introduced mutations, with the M2-1 mutation specified in brackets. (*A*) Mutations are indicated in the viral genome map by green bars, except that M2-1[A73S] is indicated in red. (*B*) T_{SH} determined by the efficiency of plaque formation at 32 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, and 40 °C using published methods (22). The experiment was done four times for viruses 1, 5, 12, and 14; three times for viruses 2, 3, and 7–9; two times for viruses 4 and 11, and once for viruses 6, 10, and 13 (bar graphs: medians and ranges). (*C* and *D*) Replication of Min L-derived mutants in vitro. Vero cells were infected at 32 °C and 37 °C (MOI of 0.01). Titers correspond to the mean of two replicate titrations of two replicates for each time point. The SD is indicated. Because of the large number of viruses, the analysis was divided between experiments (*C*) 1 and (*D*) 2.

and L[T1166I] (Fig. 1*F*), as well as the M2-1 mutation [A73S] that was one of the prominent mutations in replicate 8 (Fig. 1*G*). The resulting 12 viruses (Fig. 3*A*) were recovered and sequenced completely, confirming the correct sequences and absence of additional mutations.

The introduction of the N[K136R] or P[E114V] mutation alone conferred approximately a 1 °C increase in T_{SH} (Fig. 3*B*) compared with Min L, whereas L[T1166I] alone did not have an effect. Interestingly, the introduction of M2-1[A73S] or [N88K] alone induced a 2 °C increase in T_{SH}, suggesting that either of these two M2-1 mutations alone played the greatest role in the in vitro deattenuation of Min L. The combination of the N or P mutation with M2-1[N88K] conferred another small increase in T_{SH} (to an average of 2.5 °C from three independent experiments). The combination of the N, P, and M2-1[N88K] mutations induced a 3 °C increase in T_{SH} (40 °C) compared with Min L, which was not further increased by the addition of the L mutation. These data illustrated the additive role of the N, P, and M2-1[N88K] mutations in the increase in the T_{SH} of lineage 3. The combination of M2-1[A73S] and [N88K] did not confer any increase in the T_{SH} of Min L, illustrating their incompatibility, as predicted based on the deep sequencing results.

We also studied the effects of these mutations on the kinetics and efficiency of Min L replication in Vero cells (Fig. 3 *C* and *D*). The effects of the mutations were more evident at 37 °C (Fig. 3 *C*, *Right* and *D*, *Right*) than at 32 °C (Fig. 3 *C*, *Left* and *D*, *Left*), which would be expected for temperature sensitivity mutations. The N and P mutations alone and in combination had only a small effect on increasing viral replication compared with Min L. In contrast, the introduction of either the M2-1[N88K] or the [A73S] mutation alone resulted in a substantial increase in replication, and this increase in replication was not much affected by the further addition of the N, P, and L mutations. In addition, virus bearing both of the incompatible M2-1[A73S] and [N88K] mutations replicated similar to or less efficiently than Min L at 37 °C and 32 °C, respectively (Fig. 3*D*). Thus, the M2-1[N88K] or [A73S] mutation played the major role in restoring the ability of Min L to replicate in Vero cells, but they were incompatible.

Either of the Two Incompatible Mutations in the M2-1 Protein Partly Restores RNA Synthesis and Gene Expression in Vitro. We evaluated the effects of the introduced mutations on the kinetics of viral gene transcription, viral genomic RNA synthesis, protein expression, and virus particle production in a single infection cycle (Fig. 4 *A*–*E*). Vero cells were infected at a multiplicity of infection (MOI) of 3 pfu per cell with the indicated viruses, and samples were collected for analysis every 4 h for 24 h.

Analysis of the accumulation of the nine smaller RSV mRNAs (i.e., all except L) was performed by positive sense-specific quantitative RT-PCR (qRT-PCR) assays specific for each mRNA. Data for the P mRNA, which are generally representative, are shown in Fig. 4*A*, and the complete dataset for these nine mRNAs is shown in Fig. S3. In general, transcription was greatly reduced at 37 °C for Min L compared with WT rRSV. The introduction of either M2-1 mutation into Min L resulted in a substantial restoration of transcription. The further addition of the N, P, and L mutations to M2-1[N88K] provided an additional modest, but mostly consistent, increase. Western blot analysis showed that, as expected, the viral protein accumulation occurred later than that of the mRNAs; otherwise, the pattern was similar to the mRNA accumulation but reduced compared with WT (Fig. 4*B* and Fig. S4).

We also examined the accumulation of the RSV L mRNA by positive sense-specific, L-specific qRT-PCR (Fig. 4*C*). At 32 °C, a basal level of L mRNA was detected in Min L-infected cells, but there was essentially no increase with time in contrast to the progressive increase with time observed with WT L mRNA. Note that the extensive sequence differences in the WT and CPD L genes necessitated the use of different primer pairs for WT rRSV vs. Min L derivatives, precluding direct comparison of relative abundances at the different time points. At 37 °C, CPD L mRNA was undetectable, indicating a strong restriction at this temperature. The addition of the M2-1[N88K] or [A73S] mutation to Min L partly restored CPD L gene transcription at both 32 °C and 37 °C. The additional inclusion of the N, P, and L mutations further increased L gene expression.

The production of cell-associated genomic RNA (Fig. 4*D*) by Min L was almost undetectable at either 32 °C or 37 °C but was detected at 24 h postinfection (pi) at both 32 °C and 37 °C by M2-1[A73S] and M2-1[N88K] and in further increased amounts in NPM2-1[N88K]L–infected cells. Genomic RNA production by WT rRSV was detectable starting at 12 h pi at both 32 °C and 37 °C, and it was higher compared with NPM2-1[N88K]L.

The production of infectious virus particles was concurrent with the accumulation of genomic RNA (Fig. 4*E*). At 32 °C, Min L virus titers started to increase only at 24 h pi, whereas no increase was detected at 37 °C. M2-1[A73S] and M2-1[N88K] virus particles started to accumulate earlier (20 h pi at both temperatures) and at higher levels (6- and 110-fold higher at 32 °C and 37 °C, respectively) than Min L particles. NPM2-1[N88K]L virus production was first detected at 16 h pi at both temperatures and also, greater amounts (9- and 300-fold higher at 32 °C and 37 °C, respectively) than Min L virus production. Infectious WT rRSV was first observed at 12 h pi at both temperatures (Fig. 4*E*) at higher level than NPM2-1[N88K]L (10-fold higher at both 32 °C and 37 °C).

We also measured plaque sizes produced in Vero cells as an additional parameter for virus fitness (Fig. 4 *F* and *G*). WT rRSV produced plaques of significantly larger size than Min L ($P < 0.05$). Addition of the M2-1 [A73S] or [N88K] mutations to Min L increased virus fitness, resulting in plaque sizes that were not significantly different from those of WT rRSV ($P > 0.05$ compared with WT rRSV). Plaques induced by M2-1[A73S][N88K] were smaller than Min L plaques, further confirming that these two M2-1 mutations are incompatible.

Introduction of in Vitro Deattenuating Mutations from Min L into Min FLC Had Little Effect. The major mutations that were introduced into Min L as described above, namely N[K136R], P[E114V], M2-1[N88K], M2-1[A73S], and L[T1166I], were introduced in various combinations into Min FLC and assessed for virus titer after recovery (Fig. S5*A*) and T_{SH} (Fig. S5*B*). The M2-1[N88K] and [A73S] mutations individually did not increase the fitness of Min FLC as measured by viral titer or T_{SH}. The combination of the N, P, and M2-1[N88K] mutations conferred a 2 °C increase in T_{SH}, but this virus only grew to a low titer. Surprisingly, the introduction of the L[T1166I] mutation into Min FLC alone or in combinations with one or more of the other mutations seemed to inhibit recovery. Thus, none of these mutations improved the overall fitness of Min FLC, although it bears the same CPD L gene as Min L. This result suggests that multiple CPD ORFs augment phenotypic stability under selective pressure.

Evaluation of M2-1 Derivatives in Mice and Hamsters. We evaluated replication of the Min L derivatives in vivo (Fig. 5). BALB/c mice were infected intranasally (IN) with 10^6 pfu each virus. Nasal turbinates (NTs) and lungs were harvested on days 4 ($n = 8$ per virus), 5 ($n = 8$), and 10 ($n = 4$) pi. At the peak of virus replication (day 5 pi) (Fig. 5*B*), virus was detected in the NTs of only two mice infected with Min L and three mice infected with M2-1[N88K]. Replication of M2-1[A73S] was detected in the NTs of four of five mice, which was comparable with WT rRSV. NPM2-1[N88K]L replication was not detected in the NTs of any of the mice. In the lungs on day 5, replication of M2-1[N88K] and M2-1[A73S] was slightly reduced compared with Min L but was not statistically different, and replication of NPM2-1[N88K]L was strongly reduced in the lungs compared with Min L and detected in only two animals. At day 10, virus was recovered only from two animals in the M2-1[A73S] group at trace levels.

We also compared the same set of viruses in hamsters (Fig. 5*C*). On day 3, NTs and lungs were harvested from nine hamsters per virus. In the NTs, Min L replication was reduced ~100-fold compared with WT rRSV ($P \leq 0.01$). Replication of M2-1[N88K] was modestly increased compared with Min L but remained significantly attenuated compared with WT rRSV. In contrast, the titers of M2-1[A73S] were further increased compared with Min L and were not statistically different from WT rRSV. Interestingly, replication of

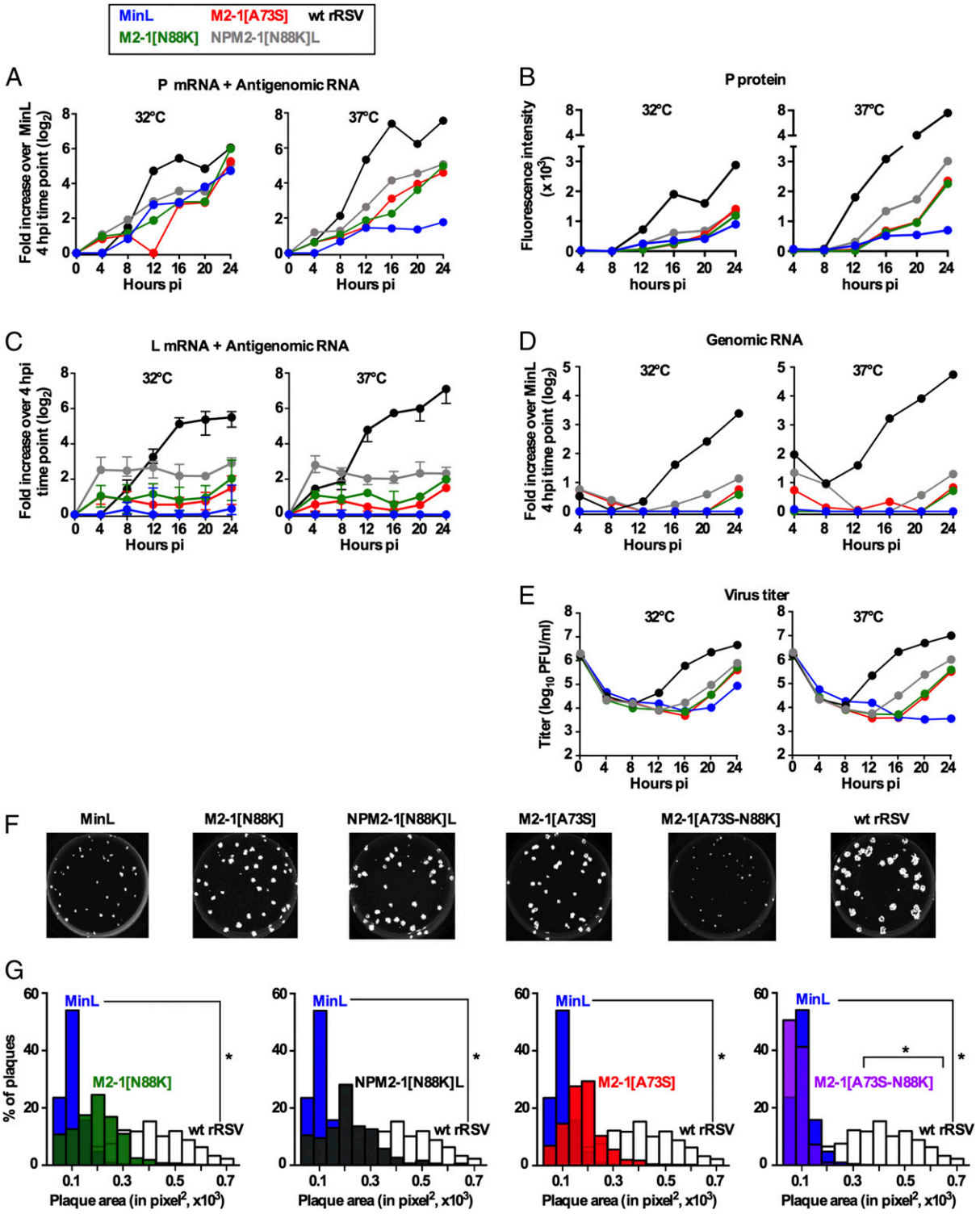

Fig. 4. Effects of specific mutations on RNA synthesis and plaque size of Min L derivatives. (*A*–*E*) Replicate cultures of Vero cells were infected (MOI of 3) with the indicated viruses. Cultures were harvested every 4 h from 4 to 24 h pi for analysis of cell-associated RNA, protein, and virus. (*A*) Positive sense viral RNA (i.e., mRNA and antigenome) was quantified in triplicate by strand-specific qRT-PCR (22). Data for P are shown. The corresponding data for NS1, NS2, N, M, SH, G, F, and M2 are shown in Fig. S3. qPCR results were analyzed using the ΔCt method, normalized to 18S rRNA, and expressed as $\log_2$ fold increase over the Min L 4 h time point. (*B*) Quantification of P protein expression by Western blotting. Data for NS1, NS2, N, G, F, and M2-1 are shown in Fig. S4. (*C*) Quantification of L mRNA and antigenome by strand-specific qRT-PCR (fold increase relative to the 4-h pi time point calculated separately for each virus, because different primer probes sets were required for WT L gene in WT rRSV vs. the CPD L gene present in Min L and its derivatives). For WT L and CPD L, data were derived from three and four different primer probe sets, respectively, designed along the L ORFs, and the median values with ranges are shown. (*D*) Quantification of cell-associated genomic RNA by strand-specific qRT-PCR (22) expressed as fold increase over the 4-h pi time point of Min L. (*E*) Virus titers from cultures incubated at 32 °C and 37 °C and assayed at 32 °C. (*F* and *G*) Virus plaque sizes. Vero cells were infected with 30 pfu per 2-cm^2 well of WT rRSV, Min L, and Min L-derived mutants and incubated under methylcellulose at 32 °C for 12 d. Plaques were visualized by immunostaining and quantified by IR imaging (Li-Cor) using ImageJ. (*F*) Representative pictures of virus plaque sizes. (*G*) Plaques size distribution of the indicated viruses. A minimum of 1,000 plaques per virus was measured. $^*P \leq 0.05$.

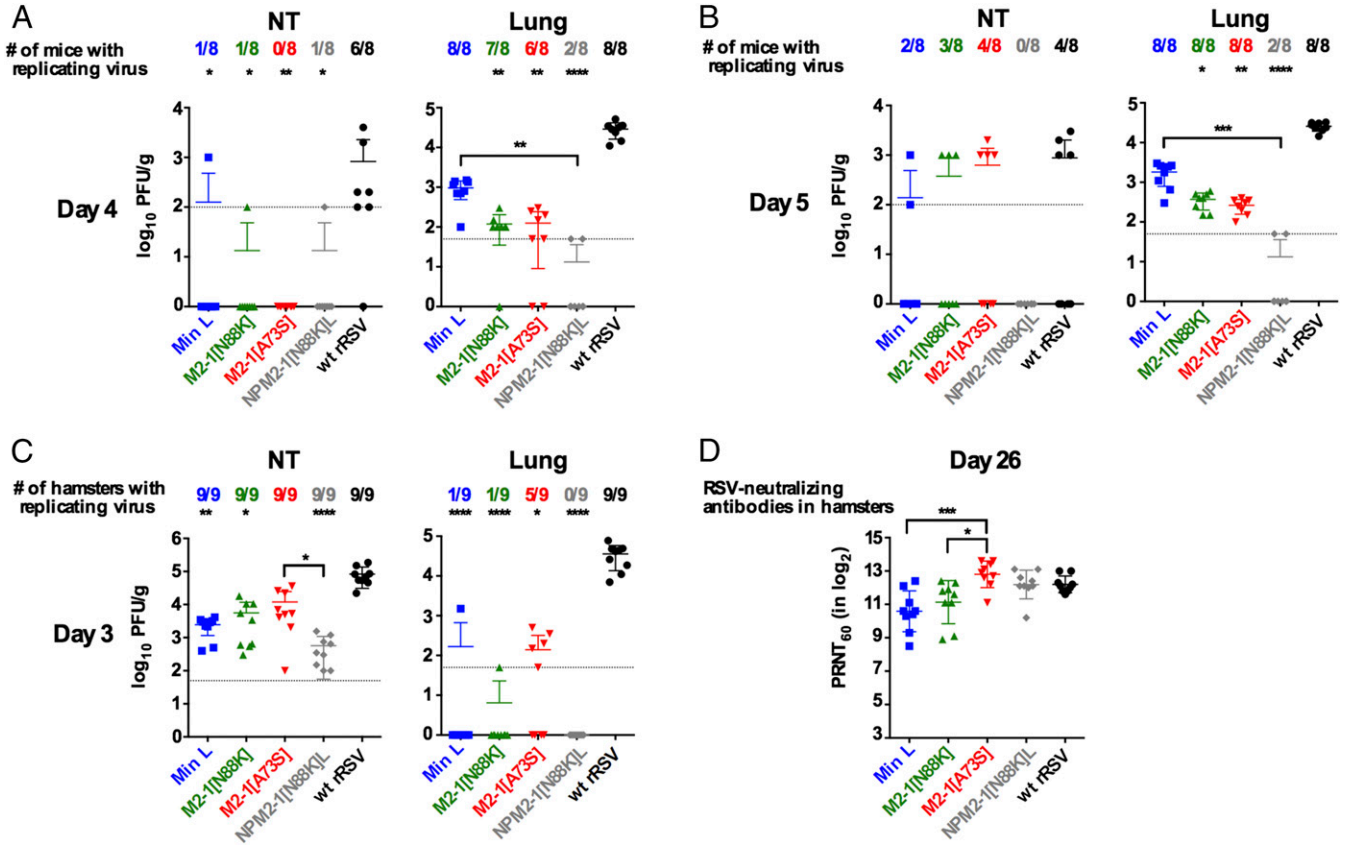

Fig. 5. Analysis of Min L derivatives in rodents shows differing effects of M2-1 mutations A73S and N88K and identifies the improved vaccine candidate NPM2-1[N88K]L. Replication of Min L, Min L mutants, and WT rRSV in mice at days (*A*) 4 and (*B*) 5 pi or (*C*) in hamsters at day 3 pi. Groups of (*A* and *B*) 20 mice or (*C*) 18 hamsters were infected IN with 10^6 pfu indicated virus per animal. At days (*A*) 4, (*B*) 5, and 10 pi for the mouse study or (*C*) at day 3 for the hamster study, RSV titers in NTs and lungs were determined as described in *Materials and Methods*. The limit of detection is indicated by a dotted line. (*D*) RSV-neutralizing antibodies at day 26 in hamsters from nine hamsters per group. The 60% plaque reduction neutralizing antibody titer ($PRNT_{60}$) values were determined as described previously (22). Statistical differences compared with WT rRSV are indicated on the top of each graph; statistical differences between Min L and the Min L-derived mutants are indicated by brackets. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$; $****P \leq 0.0001$.

NPM2-1[N88K]L in the NTs was reduced compared with Min L. In the lungs, Min L and M2-1[N88K] were detected in only one of nine hamsters for each virus, and replication of NPM2-1[N88K]L was undetectable. In contrast, replication of M2-1[A73S] was increased compared with Min L, because five of nine hamsters exhibited virus replication to about 10^2 pfu/g. Thus, in hamsters, the mutation M2-1[A73S] increased the replication of Min L, a marker of deattenuation, whereas the M2-1[N88K] mutation did not affect the replication of Min L, and the combination of the N, P, L, and M2-1[N88K] mutations decreased replication.

Despite a significant restriction of replication, Min L and the Min L-derived viruses induced titers of antibodies that were not statistically different from those induced by WT rRSV (Fig. 5*D*). The M2-1 [A73S] virus induced significantly higher levels of RSV-neutralizing serum antibodies than Min L and M21-1[N88K]. Interestingly, the NPM2-1[N88K]L virus also was comparable with WT rRSV in inducing RSV-neutralizing antibodies, despite its highly restricted replication. On day 31, hamsters were challenged IN with WT rRSV, and NTs and lungs were harvested 3 d postchallenge. No detectable challenge virus replication was detected, except for a trace of virus in one animal in the Min L group.

Genetic Stability of the NPM2-1[N88K]L Virus. The observation that the NPM2-1[N88K]L virus was more highly attenuated than Min L and yet, as immunogenic as WT rRSV identified this virus as a promising vaccine candidate. Therefore, we evaluated its stability in a temperature stress test involving four passages at 39 °C and four passages at 40 °C, corresponding to 2 mo of continuous passage (Fig. S6). Sanger sequencing of the complete genome of the final passage of 10 different stressed lineages and two control flasks did not detect any abundant mutations. These data showed that introduction of the N, P, M2-1[N88K], and L mutations into Min L to create the promising NPM2-1[N88K]L virus conferred genetic stability.

Discussion

Synonymous genome recoding of a number of RNA viruses has yielded a diverse spectrum of potential live attenuated vaccine candidates (2). However, the stability of such candidates under strong selective pressure had not been investigated. Before clinical vaccine studies are initiated, it is paramount to evaluate the stability of attenuation. In this study, we found that the RSV vaccine candidate Min FLC was genetically and phenotypically stable during 7 mo of passage in vitro at the permissive temperature of 32 °C as well as under conditions of increasing temperature during passage. These data are strong evidence for the stability of Min FLC and validates the safety of CPD of multiple genes for the development of live attenuated vaccines for RSV and related viruses, provided that extensive CPD is used.

The other vaccine candidate, Min L, was highly stable at 32 °C. Surprisingly, despite the large number of changes involved in its CPD, Min L quickly lost a substantial amount of its temperature sensitivity during the temperature stress test. Deep sequencing of the Min L lineages identified mutations in all but the NS2 ORF rather than specifically in the CPD L ORF as might have been expected. Surprisingly, many of the mutations in L occurred at nucleotides and codons that were not involved in CPD. Most of the mutations in the various ORFs involved changes in amino

acid coding. This positive selection for amino acid change suggests that at least part of the adaptation of Min L to selective stress involved changes in structure/function in various viral proteins.

The most prominent mutations acquired under stress were two missense mutations ([A73S] and [N88K]) in the M2-1 ORF, encoding the RSV transcription antitermination factor. Reintroduction of either of these mutations into Min L by reverse genetics rescued a substantial part of its replicative fitness at 37 °C, increasing viral gene transcription, protein expression, particle production, and plaque size. Notably, the increase in the expression of CPD L mRNA was evident early in infection (e.g., 4 h pi), whereas increases in the expression of the other viral mRNAs occurred later (e.g., 12 h or more pi). Furthermore, the increase in the early expression of CPD L mRNA was enhanced by the further inclusion of the N, P, and L mutations. We believe that these observations argue for a direct effect on increasing the efficiency of transcription of the CPD L gene. An alternative possibility, suggested by the normal activity of M2-1 in promoting transcriptional read through, is that the mutations in M2-1 increased read through across all of the RSV genes, resulting in the delivery of more polymerase to downstream genes, including L. However, this alternative model seems inconsistent with the kinetics of mRNA expression and the augmenting effects of the N, P, and L mutations as noted above. The partial restoration of L gene expression, which otherwise was below the level of detection at 37 °C, would be expected to increase the production of the L polymerase, although that was not directly monitored here because of its low abundance and a lack of available antibody. An increase in the production of L protein would then increase transcription of all of the RSV genes (consistent with their delayed kinetics of increase), indirectly increasing the synthesis of viral proteins, increasing RNA replication, and ultimately, indirectly increasing the production of progeny virus. Thus, we suggest that the acquisition of either of two mutations in M2-1 (and the N, P, and L mutations) adapted Min L at 37 °C by increasing transcription of the CPD L gene.

The mechanism(s) behind the rescued CPD L gene expression remain unknown. As noted, the mutations with the greatest effect were in the RSV M2-1 protein. This protein is necessary for the efficient synthesis of full-length mRNAs, which otherwise terminate prematurely (25). The M2-1 protein also increases the synthesis of polycistronic read-through mRNAs as noted. It likely binds nascent mRNA cotranscriptionally and prevents termination by the viral polymerase (26). In addition, the M2-1 protein binds directly to P (27). The binding of P and RNA to M2-1 was found to be mutually exclusive because of partially overlapping interaction surfaces (26, 28). Although A73 and N88 are away from the RNA/P binding interface, they could possibly be on the path of the exiting nascent RNA molecule. A simple model would be that the 1,378 nucleotide changes that were introduced during CPD affected the L gene template so as to reduce the efficiency of transcription elongation of the nascent L mRNA. L transcription was partly restored by the M2-1 mutations (and further enhanced by the N, P, and L mutations) through some effect on the polymerase complex, perhaps by reducing termination of the polymerase on the CPD template or nascent mRNA or by increasing processivity.

We used computer-based molecular dynamics simulations (MDSs) to investigate possible effects of the M2-1 [A73S] and [N88K] mutations on M2-1 structure (Fig. 6). The M2-1 tetramer is shown in Fig. 6*A*, with specific views in Fig. 6 *B*–*D*. In the WT M2-1 tetramer, a salt bridge is predicted to exist between K19 of one monomer and D116 of the adjoining monomer. These amino acids are shown for the red and cyan monomers (Fig. 6*B*). MDS suggests that the salt bridge helps stabilize the interaction between adjacent monomers. The A73 residue of a third monomer is predicted to be in close proximity but not involved in interactions. When A73 is changed to serine ([A73S]) (Fig. 6*C*), the salt bridge between K19 and D116 is predicted to be maintained. In addition, unlike the alanine, a serine at codon 73 is predicted to form a hydrogen bond with K19 and in some MDS timeframes, a hydrogen bond with D116. Thus, S73 could provide stabilizing links between each adjoining monomers. The predicted effect of the N88K mutation is to increase stability within rather than between monomers. Specifically,

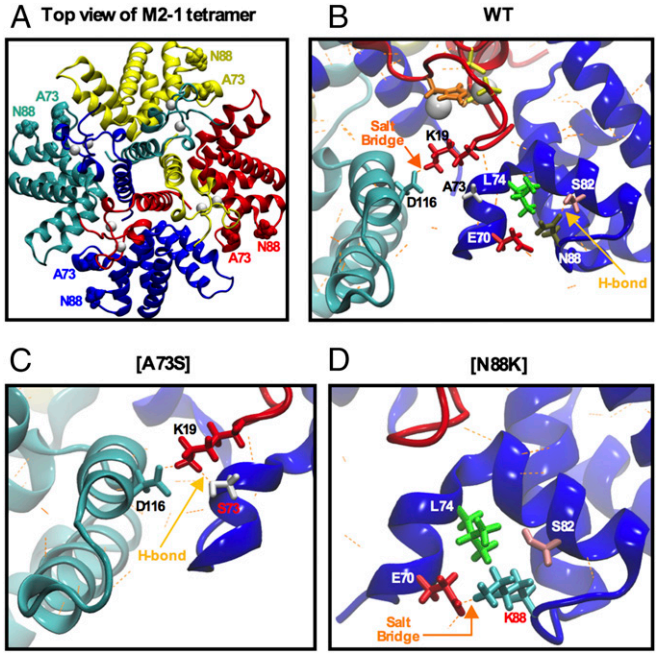

Fig. 6. Molecular modeling of the impact of deattenuating mutations on the M2-1 tetramer. (*A*) Top view of the WT M2-1 tetramer. Each monomer is represented by a different color (yellow, red, blue, and cyan). A73 and N88 are colored in each monomer and rendered as van der Waals spheres. Gray spheres represent zinc atoms. (*B*) Enlargement of one WT tetramer's region that contains amino acids A73 and N88. The residues K19 and D116 that are in close proximity to A73 are indicated in black, and their side chains are colored in red, cyan, and gray, respectively. The residues L74, S82, and E70 that are in the close proximity of N88 are indicated in white, and their side chains are colored in green, pink, red, and brown, respectively. The putative salt bridge and the hydrogen bond are indicated by orange and yellow arrows, respectively. (*C*) Molecular dynamics snapshot of the region proximal to the S73 mutation. The [A73S] mutation is shown in red, and the yellow arrow indicates the predicted new hydrogen bond. (*D*) Molecular dynamics snapshot of the K88 mutant region. The [N88K] mutation is indicated in red, with its side chain colored in cyan, and an orange arrow indicates the expected new salt bridge.

in the WT M2-1 tetramer structure, N88 is predicted to form a hydrogen bond with S82 (Fig. 6*B*). In contrast, a lysine residue at codon 88 is predicted to form an intramonomer salt bridge with E70 (Fig. 6*D*). The K88 would no longer interact with S82. In addition, the hydrophobic carbon chain of K88 is predicted to form a number of intramonomer van der Waals interactions with L74. Thus, the prominent M2-1 mutations acquired during the stress test are predicted to create interactions between (A73S) and within (N88K) M2-1 monomers. This increased stability presumably could contribute to rescue transcription of the CPD L gene. Interestingly, this increased stability is not expected to be maintained when both mutations are present together. Indeed, these two mutations could possibly form an H-bonded pair between the side chains of the S73 and K88, which would result in less flexibility of the loop on which K88 resides. This reduced flexibility could explain the incompatibility of these two mutations.

Interestingly, mutations that were found in P ([E113G] and [E114V]) are localized in the interaction domain of P with M2-1. Mutations at these two positions were shown to increase the affinity of P for M2-1 (29). This work further supports the theory that the compensatory mutations act by increasing the stability of the ribonucleoprotein complex, which we hypothesize may facilitate transcription of the CPD L gene.

As mentioned, a single mutation in the M2-1 gene (A73S) that appeared in the first passage of Min L at 37 °C and was found in 8 of 10 cultures was sufficient to rescue Min L replication at that temperature. In addition, this single mutation conferred increased replication

to Min L in hamsters. We had anticipated that deattenuation of a CPD ORF would involve multiple changes in the CPD sequence conferring incremental deattenuation. However, this study shows that a single mutation in a different gene was sufficient to yield substantial deattenuation. Therefore, deoptimization involving large numbers of nucleotide changes does not necessarily provide a stable attenuation phenotype. However, analysis of the presumptive in vitro deattenuating mutations that appeared during passage of Min L in vitro and incorporation of several of these mutations into Min L resulted in an improved vaccine candidate. Specifically, the NPM2-1[N88K]L virus (T_{SH} of 40 °C) represents a substantial improvement over Min L for a number of reasons. (*i*) It exhibited increased replication compared with Min L in Vero cells, important for vaccine manufacture. (*ii*) It did not accumulate additional mutations when passaged in stress tests at 39 °C to 40 °C. (*iii*) Of paramount importance, the virus was significantly more attenuated in vivo than Min L but as immunogenic as WT rRSV. (*iv*) Furthermore, because the M2-1[N88K] and [A73S] mutations are incompatible, this virus should be highly refractory to acquiring the M2-1[A73S] mutation that was deattenuating in the hamster model. Therefore, NPM2-1[N88K]L is an improved vaccine candidate, appropriate for evaluation in a clinical study.

Materials and Methods

Detailed protocols are in *SI Materials and Methods*.

Virus Harvest and Titration. Vero cells were scraped into the tissue culture medium, vortexed for 30 s, clarified by low-speed centrifugation, and snap-frozen. Virus titers in the clarified fluids were determined by immunoplaque assay on Vero cells at 32 °C (22).

Ion Torrent Deep Sequencing. Purified viral RNA from clarified culture fluids was copied into eight overlapping fragments spanning the viral genome. Libraries were prepared following the Ion Torrent protocol, loaded into a semiconductor sequencing chip, and sequenced on a Personal Genome Machine (Ion Torrent). A nucleotide variant was called if it occurred >50 times with an average read depth of 1,000× and a P value < 10^{-7} (quality score > 70) (30).

Deep Sequencing of Long PCR Fragments. Purified viral RNA from culture fluids was reverse-transcribed using the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific). Using RSV-specific primers and the SequalPrep Long PCR Kit (Life Technologies), the cDNAs were used to generate a PCR product of 8.2 kb spanning the genome from the 3′ end to the middle of the M2-2 ORF. DNA template libraries were prepared, sequenced, and analyzed using PacBio kits and instrumentation and CluCon software (https://github.com/mpsbpbi/clusteringConsensus).

Strand-Specific rRSV RNA Quantification. Cell-associated RNA derived from single-cycle replication experiments was used to specifically quantify viral negative sense (genomic) and positive sense (mRNA and antigenomic) RNA as described previously (22). Quantitative PCR (qPCR) results were analyzed using the comparative threshold cycle (ΔCt) method, normalized to 18S rRNA, and then, expressed as $\log_2$ fold increase over the indicated reference sample.

Evaluation of the Replication of CPD rRSVs in Mice and Hamsters. Animal studies were approved by the National Institute of Allergy and Infectious Diseases Animal Care and Use Committee and performed using previously described methods (22).

Molecular Dynamics Analysis of M2-1 Tetramer. Mutations were introduced to the crystal structure of the transcription antiterminator M2-1 protein of human RSV (Protein Data Bank ID code 4C3D) (26) using the SYBYL program (Certara). MDSs were performed using the NAMD program (v.2.9).

Statistical Analysis. Distributions of the plaques sizes were analyzed using the Kolmogorov–Smirnov test followed by Bonferroni correction. Virus replication and antibody responses in the animal experiments were analyzed using the nonparametric Kruskal–Wallis test with Dunn's posthoc analysis. A $\log_{10}$ transformation was applied to datasets when necessary to obtain equal SDs among groups. Statistics were performed using Prism 6 (GraphPad Software). Data were only considered significant at $P \leq 0.05$.

ACKNOWLEDGMENTS. This research was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, NIH (C.L.N., T.M., M.A.D., M.M., L.Y., C.L., B.L., S. Munir, J.M.D., P.L.C., and U.J.B.).

1. Abil Z, Xiong X, Zhao H (2015) Synthetic biology for therapeutic applications. *Mol Pharm* 12(2):322–331.
2. Martínez MA, Jordan-Paiz A, Franco S, Nevot M (2016) Synonymous virus genome recoding as a tool to impact viral fitness. *Trends Microbiol* 24(2):134–147.
3. Gaunt E, et al. (2016) Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. *eLife* 5:e12735.
4. Nogales A, et al. (2014) Influenza A virus attenuation by codon deoptimization of the NS gene for vaccine development. *J Virol* 88(18):10525–10540.
5. Broadbent AJ, et al. (2016) Evaluation of the attenuation, immunogenicity, and efficacy of a live virus vaccine generated by codon-pair bias de-optimization of the 2009 pandemic H1N1 influenza virus, in ferrets. *Vaccine* 34(4):563–570.
6. Cheng BY, Ortiz-Riaño E, Nogales A, de la Torre JC, Martínez-Sobrido L (2015) Development of live-attenuated arenavirus vaccines based on codon deoptimization. *J Virol* 89(7):3523–3533.
7. Diaz-San Segundo F, et al. (2015) Synonymous deoptimization of foot-and-mouth disease virus causes attenuation in vivo while inducing a strong neutralizing antibody response. *J Virol* 90(3):1298–1310.
8. Coleman JR, et al. (2008) Virus attenuation by genome-scale changes in codon pair bias. *Science* 320(5884):1784–1787.
9. Yang C, Skiena S, Futcher B, Mueller S, Wimmer E (2013) Deliberate reduction of hemagglutinin and neuraminidase expression of influenza virus leads to an ultraprotective live vaccine in mice. *Proc Natl Acad Sci USA* 110(23):9481–9486.
10. Kunec D, Osterrieder N (2016) Codon pair bias is a direct consequence of dinucleotide bias. *Cell Reports* 14(1):55–67.
11. Tulloch F, Atkinson NJ, Evans DJ, Ryan MD, Simmonds P (2014) RNA virus attenuation by codon pair deoptimisation is an artefact of increases in CpG/UpA dinucleotide frequencies. *eLife* 3:e04531.
12. Lauring AS, Jones JO, Andino R (2010) Rationalizing the development of live attenuated virus vaccines. *Nat Biotechnol* 28(6):573–579.
13. Hanley KA (2011) The double-edged sword: How evolution can make or break a live-attenuated virus vaccine. *Evolution (N Y)* 4(4):635–643.
14. Bull JJ (2015) Evolutionary reversion of live viral vaccines: Can genetic engineering subdue it? *Virus Evol* 1(1):vev005.
15. Burns CC, et al. (2006) Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. *J Virol* 80(7):3259–3272.
16. Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E (2006) Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. *J Virol* 80(19):9687–9696.
17. Bull JJ, Molineux IJ, Wilke CO (2012) Slow fitness recovery in a codon-modified viral genome. *Mol Biol Evol* 29(10):2997–3004.
18. Nougairede A, et al. (2013) Random codon re-encoding induces stable reduction of replicative fitness of Chikungunya virus in primate and mosquito cells. *PLoS Pathog* 9(2):e1003172.
19. Vabret N, et al. (2014) Large-scale nucleotide optimization of simian immunodeficiency virus reduces its capacity to stimulate type I interferon in vitro. *J Virol* 88(8):4161–4172.
20. Meng J, Lee S, Hotard AL, Moore ML (2014) Refining the balance of attenuation and immunogenicity of respiratory syncytial virus by targeted codon deoptimization of virulence genes. *MBio* 5(5):e01704–e01714.
21. Ni YY, et al. (2014) Computer-aided codon-pairs deoptimization of the major envelope GP5 gene attenuates porcine reproductive and respiratory syndrome virus. *Virology* 450-451:132–139.
22. Le Nouën C, et al. (2014) Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. *Proc Natl Acad Sci USA* 111(36):13169–13174.
23. White MD, Bosio CM, Duplantis BN, Nano FE (2011) Human body temperature and new approaches to constructing temperature-sensitive bacterial vaccines. *Cell Mol Life Sci* 68(18):3019–3031.
24. Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197–218.
25. Fearns R, Collins PL (1999) Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol* 73(7):5852–5864.
26. Tanner SJ, et al. (2014) Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proc Natl Acad Sci USA* 111(4):1580–1585.
27. Tran TL, et al. (2009) The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J Virol* 83(13):6363–6374.
28. Blondot ML, et al. (2012) Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. *PLoS Pathog* 8(5):e1002734.
29. Mason SW, et al. (2003) Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J Virol* 77(19):10670–10676.
30. Chapman MA, et al. (2011) Initial genome sequencing and analysis of multiple myeloma. *Nature* 471(7339):467–472.
31. Bukreyev A, Belyakov IM, Berzofsky JA, Murphy BR, Collins PL (2001) Granulocyte-macrophage colony-stimulating factor expressed by recombinant respiratory syncytial virus attenuates viral replication and increases the level of pulmonary antigen-presenting cells. *J Virol* 75(24):12128–12140.
32. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1):33–38.
33. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781–1802.
34. Brooks BR, et al. (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30(10):1545–1614.