Published online 26 March 2011

Nucleic Acids Research, 2011, Vol. 39, No. 13 ***5611–5621***
doi:10.1093/nar/gkr171

Paired mutations abolish and restore the balanced annealing and melting activities of ORF1p that are required for LINE-1 retrotransposition

James D. Evans[1], Suresh Peddigari[1], Kathy R. Chaurasiya[2], Mark C. Williams[2,3,*] and and Sandra L. Martin[1,4,*]

[1]Department of Cell and Developmental Biology, University of Colorado School of Medicine, Aurora, CO 80045, [2]Department of Physics, [3]Center for Interdisciplinary Research on Complex Systems, Northeastern University, 111 Dana Research Center, Boston, MA 02115 and [4]Program in Molecular Biology, University of Colorado School of Medicine, Aurora, CO 80045, USA

Received January 12, 2011; Revised March 9, 2011; Accepted March 10, 2011

ABSTRACT

Retrotransposition amplifies LINE-1 (L1) to high copy number in mammalian genomes. The L1 protein encoded by ORF1 (ORF1p) is required for retrotransposition. This dependence on ORF1p was investigated by mutating three highly conserved residues, R238, R284 and Y318 to alanine, thereby inactivating retrotransposition. R284A and Y318A were rescued by further substituting the alanine with the appropriate conservative amino acid, e.g. lysine or phenylalanine, respectively, whereas R238K remained inactive. Quantification of the steady-state levels of L1 RNA and ORF1p failed to discriminate active from inactive variants, indicating loss of L1 retrotransposition resulted from loss of function rather than reduced expression. The two biochemical properties known for ORF1p are high-affinity RNA binding and nucleic acid chaperone activity. Only R238A/K exhibited significantly reduced RNA affinities. The nucleic acid chaperone activities of the remaining paired mutants were assessed by single-molecule DNA stretching and found to mirror retrotransposition activity. To further examine ORF1p chaperone function, their energetic barriers to DNA annealing and melting were derived from kinetic work. When plotted against each other, the ratio of these two activities distinguished functional from non-functional ORF1p variants. These findings enhance our understanding of the requirements for ORF1p in LINE-1 retrotransposition and, more generally, nucleic acid chaperone function.

INTRODUCTION

Long interspersed element-1 (LINE-1, or L1) is a successful and active retrotransposon in mammals. L1 sequence directly accounts for 17% of the human and 19% of the mouse genomes. L1 also mobilizes non-autonomous elements including short interspersed elements (SINEs) and processed pseudogenes, making it responsible for at least 30% of mammalian DNA (1). The structure and function of the genome is impacted by L1 at multiple levels, which range from insertional mutagenesis to altering gene expression and promoting recombination [reviewed in ref. (2)].

A majority of the 660 000 L1 elements in the mouse genome are truncated and therefore cannot retrotranspose, but approximately 3000 are estimated to be competent for retrotransposition (3). Active mouse L1 is characterized by a series of monomer promoter motifs followed by a 5′-UTR, two open reading frames (ORF1 and ORF2), a 3′-UTR, poly-A signal and an A-rich region, all embedded within a short target site duplication. Retrotransposition begins with transcription of one of the active elements, followed by translation of the two L1-encoded proteins, ORF1p and ORF2p. Both L1 proteins are essential for retrotransposition (4), and both are required in *cis* (5). The L1 RNA ultimately serves as a template for target-primed reverse transcription [TPRT, (6)] resulting in a new L1 insertion in the genome.

The role of ORF1p in L1 retrotransposition is incompletely understood. The protein forms a stable trimer via an N-terminal coiled-coil domain (7), followed by an RRM domain (8) and a CTD [(9), Figure 1A]. Two biochemical properties of ORF1p are known: RNA binding and nucleic acid chaperone activity. ORF1p from mouse L1 is a high-affinity, non-sequence-specific RNA binding protein (10), which is likely critical for packaging the L1

*To whom correspondence should be addressed. Tel: +303 724 3467; Fax: +303 724 3420; Email: Sandy.Martin@UCDenver.edu
*Correspondence may also be addressed to Mark C. Williams. Tel: +617 373 7323; Fax: +617 373 2943; Email: mark@neu.edu

RNA during retrotransposition (11–14). The RNA binding region of the protein is C-terminal of the coiled-coil, encompassing the RRM and CTD domains, which are both required for high-affinity RNA binding (8,15). ORF1p is also a nucleic acid chaperone protein, a property which is required for retrotransposition (16), and may facilitate TPRT (17,18). Mutations exist in ORF1p that do not disrupt the high-affinity interaction of the protein with RNA but nevertheless abolish retrotransposition (16,17,19), consistent with a critical role for the nucleic acid chaperone activity of ORF1p in L1 replication.

Nucleic acid chaperones paradoxically promote both annealing of complementary single stranded (ss) and melting of double-stranded (ds) nucleic acids. A consequence of these two fundamental activities is rearrangement of nucleic acid secondary structure, a property that is essential in a variety of biological processes including retroviral replication (20,21), functional splicing of the T4 phage thymidylate synthase gene (22) and proper folding of a yeast tRNA anticodon (23).

Single-molecule stretching is a sensitive assay of nucleic acid chaperone activity. When single dsDNA molecules are stretched, the applied force facilitates melting of the DNA. When the DNA molecule is relaxed, DNA reannealing occurs. By measuring DNA stretching and relaxation in the presence of nucleic acid chaperone proteins, the extent to which the proteins facilitate DNA melting and annealing, important components of nucleic acid chaperone activity, can be determined [reviewed in refs (24,25)].

Here, we investigate the effect of six amino acid substitutions in ORF1p on L1 retrotransposition, expression, protein stability, RNA binding, the cooperativity of DNA melting and the kinetics of ORF1p-mediated annealing and strand exchange. We present a novel analysis of the latter data, and use it to quantify the balance between these two opposing components of nucleic acid chaperone activity by comparing wild-type (wt) to mutant proteins. Taken together, these data demonstrate that L1 retrotransposition is exquisitely sensitive to changes in the biophysical properties of ORF1p that determine its nucleic acid chaperone activity.

MATERIALS AND METHODS

Autonomous retrotransposition assay

Mutations were introduced into L1 ORF1 by site-directed mutagenesis and assayed for effects on retrotransposition in 143 B cells using eGFP expression as a marker of retrotransposition as described (5,16).

DNA, RNA and protein analysis

Whole-cell lysates (WCL) were prepared from L1 transfected 143 B cells every 24 h post-transfection as described (17). Protein was quantified (Bradford assay, Bio-Rad) and ORF1p was detected by western blotting in 20 μg of WCL as described (17). hnRNP-Q was detected on the same blots using a mouse monoclonal hnRNP-Q antibody (Abcam). Blot images were captured on a GE Typhoon 9400. L1 RNA was quantified by quantitative real time PCR using primers with a probe specific to the ORF2 region. Total RNA was isolated from WCL using TRIzol-LS (Invitrogen) and quantified (Nanodrop, Thermo Scientific). RNA (2 μg) was treated with RQ1 DNase (Promega) and cDNA was synthesized from 1 μg using a High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). Two microliter cDNA were used for quantitative PCR using TaqMan probe assay with 5′-CTCAGAATGAAAGGCTGGAAAAC as the forward primer, 5′-AGGATGGCTACTCCTGCTTGTT as the reverse primer and 5′-FAM-CCAAGCAAATGG TATGAAG-NFQMGB as the probe. A standard curve was generated using 10-fold dilutions of L1-containing plasmid. 18S rRNA was quantified using the TaqMan Gene Expression assay for 18S rRNA (Applied Biosystems). The quantities of L1 and 18S rRNA in the cDNAs were calculated from their respective standard curves using the absolute standard curve method (Applied Biosystems) and then L1 RNA values were normalized to 18S rRNA values.

Recombinant ORF1p purification

ORF1p coding sequences with single residue mutations were cloned into a modified pFastBac bacmid and used for recombinant expression in SF9 cells as before (18), according to the Bac-to-Bac baculovirus expression system protocol (Invitrogen). We replaced the 6His to ORF1 AUG sequence of pFastBac with the corresponding sequence from pBluBac to facilitate ORF1p purification as described previously (7), except that the ammonium sulfate precipitation was 27% $NH_4(SO_4)_2$ followed by 73% $NH_4(SO_4)_2$.

Circular dichroism and thermal melt

Circular dichroism (CD) and melt scans were performed on a Jasco 815 spectropolarimeter with purified ORF1p between 2.1 and 4.2 μM trimer concentration in 50 mM phosphate, pH 7.6; 250 mM NaCl and 0.1 mM EDTA, and data were collected as the observed ellipticity in millidegree. Thermal stability was determined by observing the differential absorption at 222 nm while the temperature was raised at a rate of 2°C per minute, from 4 to 70°C. Data were normalized to the 4°C data point.

RNA filter binding

Assays were performed in 250 mM NaCl and analyzed as described previously (10). Differences in K_{Dapp} of 2- to 3-fold were not significant because of experimental variation (16).

Single-molecule DNA stretching

A dual beam optical tweezers instrument was used to stretch 5′ biotin-end-labeled bacteriophage λ DNA as described elsewhere in detail (26–28). Briefly, the DNA was captured on streptavidin-coated polystyrene beads (Bangs Labs), then stretched and relaxed after rinsing out other DNA molecules. All stretching experiments were performed at a pulling rate of 100 nm/s in 10 mM HEPES; 50 mM Na^+; pH 7.5 buffer. Specific protein

solutions were exchanged with the buffer around the DNA molecule to probe protein–DNA interactions. To quantify the effects of protein on the DNA stretching curves, the transition width was determined from the transition slope at its midpoint as described previously (29). Significant differences among protein groups were assessed by one way ANOVA followed by Tukey's pairwise comparisons using KaleidaGraph 4.03 (Synergy Software).

Annealing and displacement assays

For annealing, ORF1p and two complementary, 48-nt DNA oligonucleotides, one end-labeled with ^{32}P, were incubated at equimolar concentrations in 30 µl reaction buffer (20 mM HEPES pH 7.6; 25 mM NaCl; 1 mM EDTA pH 8.0; 1 mM $MgCl_2$; 1 mM dithiothreitol and 0.1% Triton-X 100 w/v). Reactions with ORF1p were incubated at 1, 15, 22 and 37°C; those without protein were incubated at 15, 22, 37 and 42°C. Aliquots (3 µl) were taken at the noted times, mixed with 3 µl stop buffer (40 mM HEPES pH 7.6; 0.4 mg/ml tRNA; 0.2% sodium dodecyl sulfate w/v; 10 mM EDTA pH 8.0; 3% Ficoll-400 w/v; 0.25% bromophenol blue w/v and 0.25% xylene cyanol w/v) and placed on ice. Samples were fractionated through a 15%, 19:1 bis:acrylamide gel at 4°C in 0.5× TBE, the gel dried and exposed to a storage phosphor screen. Images were collected on a Typhoon 9400 and bands containing ss and ds oligonucleotides were quantified with ImageQuaNT 5.2 (GE). KaleidaGraph was used for curve fitting and graphical representation of data.

Full duplex exchange experiments to quantify melting catalysis facilitating strand exchange were performed by mixing 1 nM 38-bp dsDNA, 5′ end-labeled with ^{32}P on one of the strands, with 10 nM of the identical unlabeled ss oligonucleotide competitor and 10 nM ORF1p trimer in reaction buffer. This strand exchange assay differs from a strand-exchange assay used previously to study ORF1p because the large thermodynamic advantage that drove the previous reactions [replacement of a short (18) or imperfectly matched (16) strand in a preformed duplex with the perfect complement, provided at 50× molar excess] has been removed. Reactions were incubated with ORF1p at 22, 27, 32 and 37°C, or 37, 40, 42, 45, 47 and 50°C for reactions without protein, and then treated as described for the annealing reactions.

RESULTS

Previous studies demonstrated that retrotransposition of mouse L1 was reduced or abolished by the substitutions D159H and R297K in ORF1p (16,17). These residues lie in the coiled-coil and CTD, respectively, far apart in the primary sequence of ORF1p (Figure 1A). To further probe the functional requirements for ORF1p in L1 retrotransposition we substituted alanine for three conserved residues in the RRM and CTD domains of the protein in the context of an otherwise active L1: R238A, R284A or Y318A (Figure 1B). All three of these mutations rendered L1 inactive for retrotransposition in a cultured cell assay. Replacing the alanine to make the

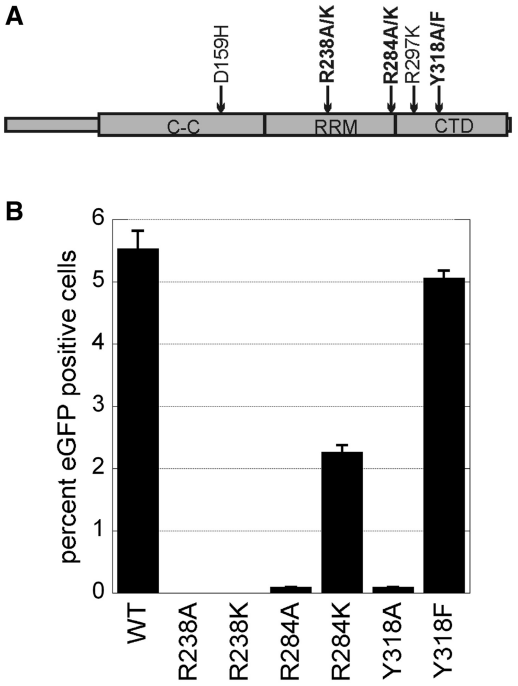

Figure 1. Mutations to conserved ORF1p residues in the murine ORF1p sequence perturb retrotransposition. (**A**) Schematic of ORF1p with single amino acid substitutions studied here (bold) or shown previously to affect nucleic acid chaperone activity (16,17). (**B**) Average percent (±SD, $n = 3$) eGFP-expressing cells 6 days after transfection with antisense intron-containing eGFP reporter in wt or ORF1p mutant L1.

variants R284K and Y318F nearly or completely restored retrotransposition, respectively, but R238K remained inactive (Figure 1B). Retrotransposition detected by eGFP fluorescence using flow cytometry was confirmed by PCR amplification of the spliced eGFP from genomic DNA (data not shown).

Single residue mutations may decrease protein stability (30,31). Mutations in ORF1p would likely interfere with L1 retrotransposition if the protein became unstable or unable to interact with L1 RNA, which might in turn cause instability of L1 RNA. Therefore the steady-state levels of L1 RNA and ORF1p were measured daily from cells transfected with wt and mutant L1 constructs. Based on qRT-PCR, L1 RNA was not reduced below wt levels in any of the inactive mutants. Moreover, differences in abundance of L1 RNA did not distinguish among active and inactive elements (Figure 2A). The highest steady-state levels of ORF1p detected in the lysates by western blotting were consistently observed 2 days post-transfection. Small fold changes of ORF1p abundance among mutants again failed to distinguish between active and inactive ORF1p variants (Figure 2B). These data indicate the defects in retrotransposition must be due to something other than instability or loss of L1 RNA or ORF1p.

Mutant proteins were purified (Supplementary Figure S1A) in order to study their biochemical properties. The final purification step is size exclusion

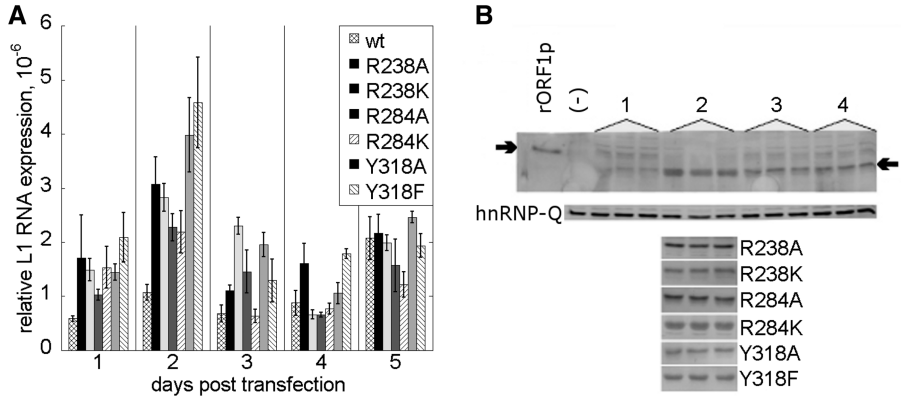

Figure 2. Effect of ORF1p mutations on steady-state levels of L1 RNA and ORF1p in cells transfected with retrotransposition assay vectors. (**A**) L1 RNA normalized to 18 S rRNA measured by qRT-PCR on Days 1–5 following transfection of 143 B cells, average (±SD, $n = 3$). (**B**) Western blots of ORF1p. Top panel, triplicate wt extracts Days 1–4 post-transfection as indicated; ORF1p is not detectable on Days 5 and 6 (data not shown). Right arrow points to ORF1p, left arrow to signal obtained from 2.5 ng his-tagged recombinant ORF1p (rORF1p); (−) untransfected control cell extract. Middle panel shows hnRNP-Q detected on the same blot. Below are triplicate lanes containing protein extracts on Day 2 post-transfection with the indicated ORF1 mutants.

chromatography. All of the mutant proteins eluted in the same fraction as wt ORF1p, indicating that they share its homotrimeric asymmetric dumbbell-shaped structure (7).

Mutant proteins were characterized by their differential absorption of circularly-polarized light. All mutant and wt ORF1p produced absorption patterns consistent with a generally helical molecule (data not shown). If local secondary structure were perturbed by some of these mutations, it was insufficient to affect the overall absorption characteristics of the molecule. Additionally, the effect of increasing temperature on elliptical absorption at 222 nm revealed that the proteins denatured between 38 and 46°C (Supplementary Figure S1B). Although R238K and Y318F were more temperature sensitive than wt ORF1p, this property was not correlated with retrotransposition competence.

A nitrocellulose filter binding assay was used to assess whether any of the ORF1p mutants were compromised for RNA binding. As shown in Figure 3, the relative affinities were: Y318F ≥ Y318A ≥ wt ≥ R284K ≥ R284A >> R238K ≥ R238A. Because of experimental variation, however, only the two substitutions at R238 were significantly different than wt, each exhibiting an ~10-fold reduction in their apparent affinity for RNA. It is likely that the retrotransposition defects in R238A and R238K are due to their reduced affinity for RNA (19), however, altered RNA binding does not explain the loss of retrotransposition in R284A or Y318A. The inability of these two mutants to retrotranspose cannot be explained by the loss of L1 RNA or ORF1 protein, a failure to trimerize, gross perturbation of secondary structure, decreased thermal stability or altered K_{Dapp} for RNA. Thus, some other property of ORF1p must be affected by these mutations.

Single-molecule stretching experiments were used to assess the nucleic acid chaperone activities of the paired mutants at R284 and Y318. Figure 4A shows a stretching and relaxation cycle for bacteriophage λDNA in the absence of protein and in the presence of wt ORF1p

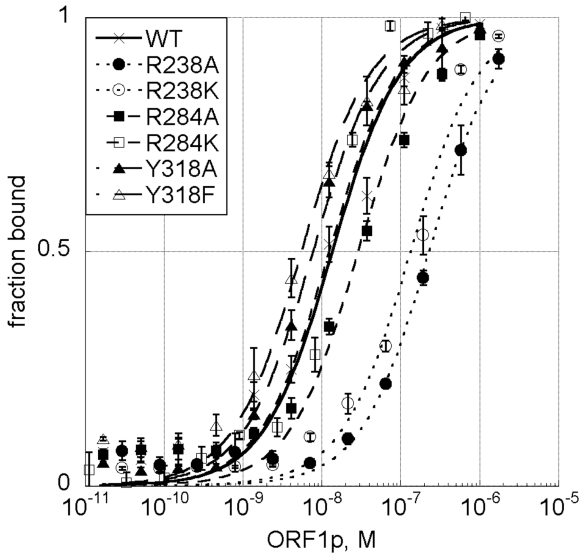

Figure 3. Binding of ORF1p variants to RNA. Nitrocellulose filter binding assay with 25–50 pM 111 nt RNA and increasing amounts of wt and mutant ORF1p. Points are averages of (±SD, $n = 3$). Lines are fits as described previously (10). Most K_{Dapp} values fell in the low nanomolar range (5.7–28 nM), with the exception of R238K and R238A (130 and 230 nM, respectively).

using optical tweezers. In the absence of protein, very little force is required to stretch the DNA to its full contour length. Near the contour length the force increases dramatically, reflecting the elasticity of the double helix. The approximately constant force plateau of ~60 pN represents a cooperative transition from dsDNA to ssDNA, or a force-induced melting transition (26,27,32,33). In addition, the force-extension curve in the absence of protein was almost completely reversible, showing little hysteresis, i.e. disagreement between stretching and relaxation. The amount of hysteresis is characterized by the area between the stretching and

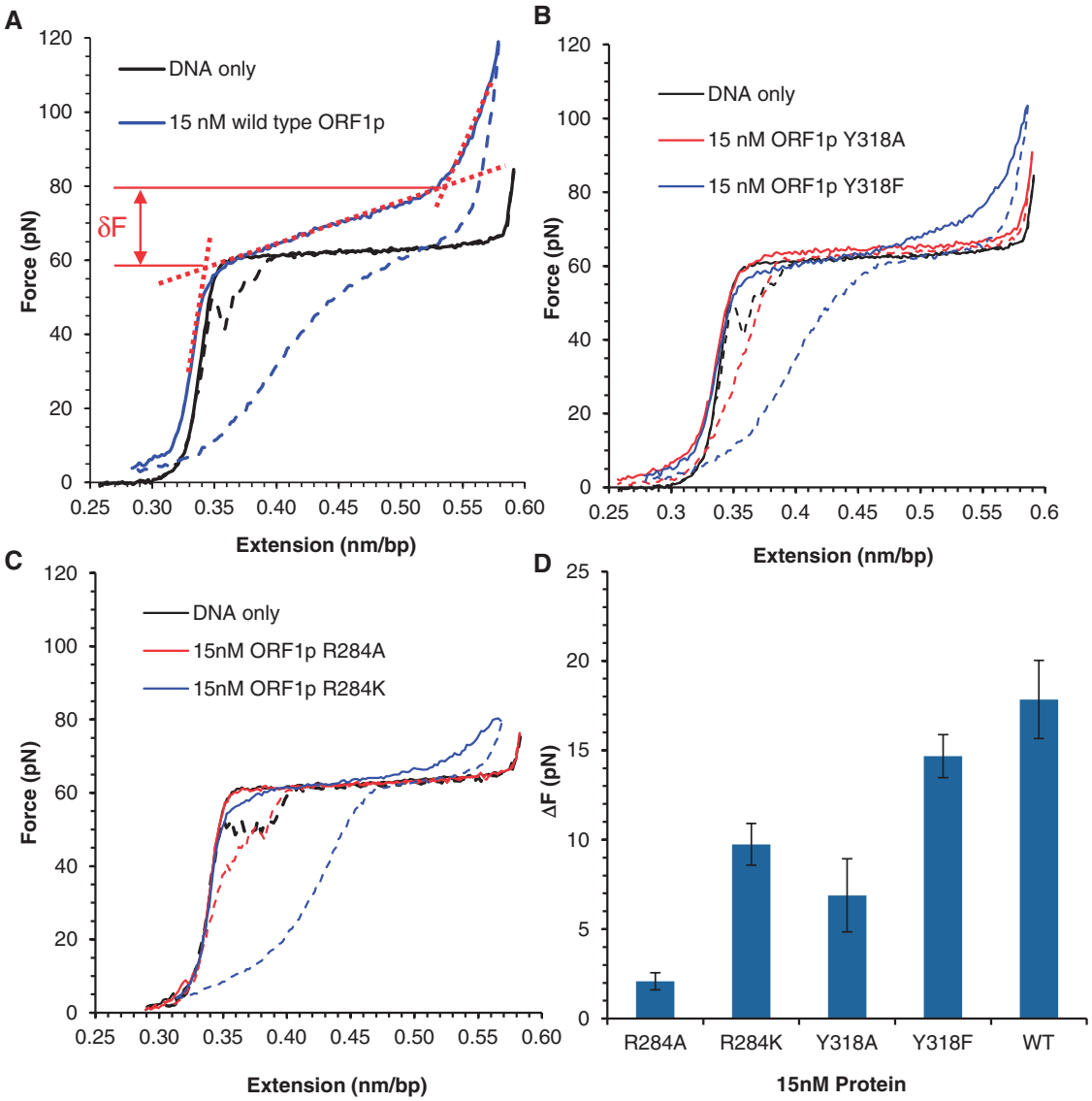

Figure 4. Single-molecule stretching. Typical force-extension (solid) and relaxation (dashed) curves of DNA only (black) and in the presence of the indicated ORF1p variants; blue lines are used for retrotransposition active variants and red lines for retrotransposition inactive variants. (**A**) wt, $\Delta F = 17.8 \pm 2.2$; red lines illustrate the calculation of the transition width, δF. (**B**) R284A, $\Delta F = 2.1 \pm 0.5$; and R284K, $\Delta F = 9.7 \pm 1.2$. (**C**) Y318A, $\Delta F = 6.9 \pm 2.0$ and Y318F, $\Delta F = 14.7 \pm 1.2$. (**D**) Change in the transition width of DNA force-induced melting in the presence of wt and variant ORF1 proteins, where we have also subtracted the value of the transition width in the absence of protein ($\Delta F = \delta F - \delta F_0$), with $\delta F_0 = 3.66 \pm 0.16$ pN (29). Error bars are standard error for at least three measurements. Significant differences among the five ORF1p variants were found by ANOVA ($P < 0.001$); the active and inactive pairs at each site are different, Y318F $\neq$ Y318A ($P = 0.036$), R284K $\neq$ R284A ($P = 0.017$), but the active R284K is not distinguishable from the inactive Y318A ($P = 0.658$).

relaxation curves. In the presence of wt ORF1p, the transition was altered significantly, as also shown in Figure 4A. The DNA force-induced melting plateau was sloped, indicating that the transition cooperativity was significantly reduced. A reduction in melting cooperativity indicates that DNA melting can be more easily initiated and smaller regions of DNA can be melted at once to facilitate DNA rearrangements. This change in shape of the transition was quantified by measuring the transition width, as described in Figure 4A. A greater transition width indicates a less cooperative transition. Therefore, when the transition width is greater, DNA will more easily undergo conformational rearrangements, as expected for a nucleic acid chaperone protein (28,34). The amount of hysteresis observed in the presence of wt ORF1p is greater than that observed for the nucleic acid chaperone HIV-1 NC (25), but significantly less than that observed for the slower nucleic acid chaperone protein HTLV-1 NC (35) and ssDNA binding proteins like T4 gene 32 protein (36). Thus, wt ORF1p exhibits moderately rapid kinetics, intermediate between other nucleic acid chaperones.

While Y318F was similar to wt ORF1p, Y318A had markedly less effect on DNA stretching (average $\Delta F = 14.7$, 17.8 and 6.9, respectively, Figure 4B and D), consistent with the relative retrotransposition results for

this mutant pair. Likewise, R284A, which was also defective for retrotransposition, had the least effect on DNA stretching while R284K strongly altered the DNA stretching curve (average $\Delta F = 2.1$ and 9.7, respectively, Figure 4C and D). Thus, as observed with other mutations in ORF1p (16,17), inactivation of L1 retrotransposition in Y318A and R284A was correlated with loss of nucleic acid chaperone activity as measured by single-molecule stretching. Similarly, mutations that restored retrotransposition also restored chaperone activity measured by DNA stretching.

Although the results from single-molecule DNA stretching for the paired mutations at each site correlated well with retrotransposition, the single-molecule data were not able to distinguish within error between two mutants that clearly differed in retrotransposition, R284K and Y318F (Figure 4D). This result may be partially explained by the fact that R284K is less active than wt and Y318F (Figure 1B); however, our understanding of the relationship between the nucleic acid chaperone activity of ORF1p and L1 retrotransposition activity was clearly incomplete.

Two aspects of nucleic acid chaperone function (37), the kinetics of DNA annealing and melting, were then examined in detail. To accomplish this, we developed assays to measure the magnitude of the energetic barriers to annealing and melting of complementary DNA oligonucleotides in the presence of ORF1p. First, complementary DNA oligonucleotides were incubated for 2.5 min at 21°C in the presence of increasing concentrations of ORF1p trimer to determine the mid-point and saturating nt:ORF1p ratio for annealing under these conditions. The annealing rate saturated when the concentration of protein to nucleic acid was ~9:1, and the half-maximal point was 1.15 nM (Supplementary Figure S2). Thus, all further experiments were performed with equimolar protein and oligonucleotides.

The effect of wt and mutant ORF1p on DNA annealing was determined at varying temperatures. The reaction can be expressed as:

$$S_1 + S_2 \overset{k_{annel}}{\rightleftharpoons} S_1S_2. \quad (1)$$

Here S_1 and S_2 are the ssDNA strands that anneal to form duplex S_1S_2 with total rate k_{anneal}. The fraction of labeled oligonucleotide incorporated into duplex was plotted against time and these points were fitted to a capped, inverse exponential decay (Figure 5A):

$$fr(t) = fr_\infty \times \left(1 - e^{-k_{anneal} \times t}\right), \quad (2)$$

where fr is the fraction of labeled oligonucleotide incorporated into duplex as a function of time, fr_∞ is the equilibrium fraction of annealed oligonucleotide, t is the reaction time in seconds and k_{anneal} is the reaction rate constant in s^{-1}.

To find the temperature dependence of the reaction rate, the natural logarithm of the resulting rate coefficient was then plotted against the inverse temperature according to transition state theory (Figure 5B):

$$\ln(k) = \ln\left(\frac{k_B T}{h}\right) + \frac{\Delta S_a}{R} - \frac{\Delta H_a}{R} \times \frac{1}{T}, \quad (3)$$

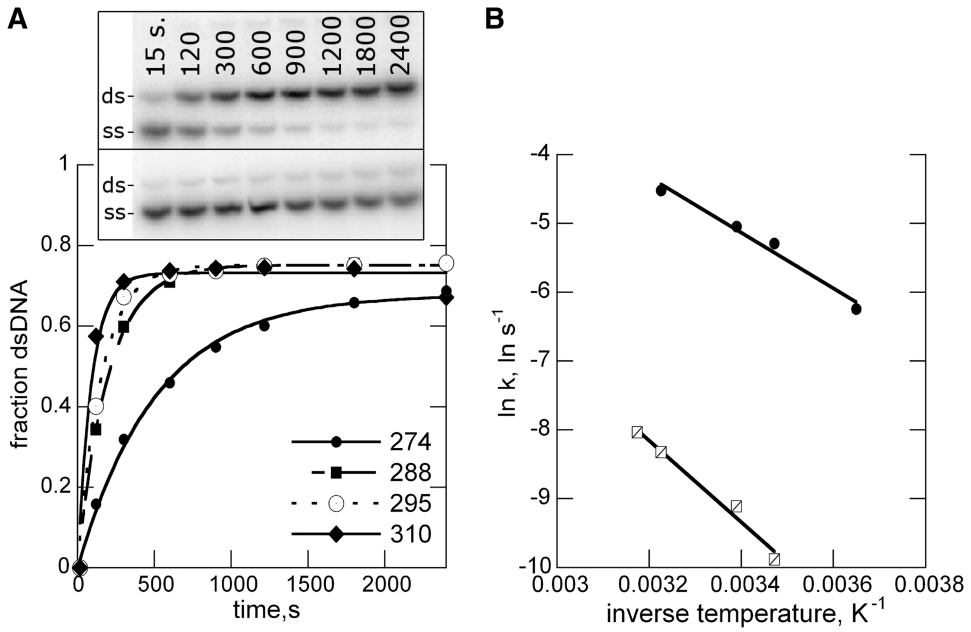

Figure 5. Temperature dependence of ORF1p-mediated annealing catalysis. (**A**) Phosphorimage of complementary oligonucleotides, one labeled with ^{32}P, annealing in the presence (inset, top), or absence (inset, bottom), of wt ORF1p, respectively. The fraction of duplex DNA formed in the presence of wt ORF1p was plotted against time for each temperature and fit to an inverse exponential decay, Equation (2). (**B**) Natural logarithms for the reaction rates were plotted against the inverse temperature, and points were fitted to a line for presence (circles) and absence (squares) of wt ORF1p.

where R is the ideal gas constant, 8.31 J mol^{-1} K^{-1}, T is the reaction temperature in Kelvin, k_B is Boltzmann's constant, h is Planck's constant, ΔS_a is the entropic contribution to the activation energy for annealing in J mol^{-1} K^{-1}, and ΔH_a is the enthalpic contribution to the activation energy for annealing in J mol^{-1}. A line was fitted to these points, allowing us to determine the enthalpic and entropic activation energies for this process. Strand exchange assays were used to determine the energetic barrier to melting and subsequent strand exchange in order to probe the other fundamental property of nucleic acid chaperones. A complementary DNA duplex, one strand labeled with ^{32}P, was incubated with a 10-fold excess of competing, unlabeled oligonucleotide and equimolar ORF1p, again varying only temperature. The excess of unlabeled complementary oligonucleotide assured that the displaced ^{32}P -labeled strand remained displaced for quantification. The time dependence of the reaction could only be fit to two exponentials, and was therefore biphasic:

$$S_1+S_2S_3 \overset{k_1}{\rightleftharpoons} (S_1S_2S_3)^* \overset{k_2}{\rightleftharpoons} S_1S_3+S_2. \quad (4)$$

here S_1 is the unlabeled ssDNA strand and S_2S_3 is the duplex that contains the labeled ssDNA strand as well as one unlabeled strand. The intermediate complex $(S_1S_2S_3)^*$ will be formed when an unlabeled strand encounters a partially melted duplex. This represents the first step in the reaction, which occurs with the faster rate k_1. When the final duplex S_1S_3 is formed, S_2 must be dissociated with rate k_2. Based on this model, the data points were fit to a two-factor exponential decay (Figure 6A):

$$\mathrm{fr}(t) = \mathrm{fr}_\infty + (1-\mathrm{fr}_\infty)\left(P \times \left(e^{-k_1 \times t}\right) + (1-P) \times \left(e^{-k_2 \times t}\right)\right), \quad (5)$$

where P is the fractional contribution to the overall rate from the first rate, k_1 and k_2 are the first and second rates, respectively, and the remaining factors are the same as Equation (2).

Based on previous observations of a biphasic reaction for DNA and RNA hairpin annealing (21,37), the fast rate was assumed to be the primary DNA melting step required for initial annealing of the complex. This hypothesis is also supported by the observation that the fast rate is similar to that observed for other reported chaperone-mediated DNA-duplex melting reactions (21), whereas the rate constant of the slow phase was two to three orders of magnitude slower than that observed for other nucleic acid chaperones. We will show later that the analysis of the fast rates correlate with mutations in ORF1p, further supporting our assumption that the fast rate involves DNA melting needed to initiate strand transfer, a process facilitated by nucleic acid chaperones. In contrast, the slow rates did not depend significantly on the presence of ORF1p. The temperature dependence of the fast rate was plotted and fit to the transition state model as before (Figure 6B).

The enthalpic and entropic contributions to the activation energy for product formation were quantified from Equation (3) (Table 1) for both annealing and exchange and these terms were combined into a total free energy of activation according to:

$$\delta\Delta G_a = \delta\Delta H_a - T\delta\Delta S_a, \quad (6)$$

where $\delta\Delta H_a$ is the difference in enthalpic contributions between a particular mutant and the no protein condition, $\delta\Delta S_a$ is the difference between entropic contributions between a mutant and no protein, T is the temperature

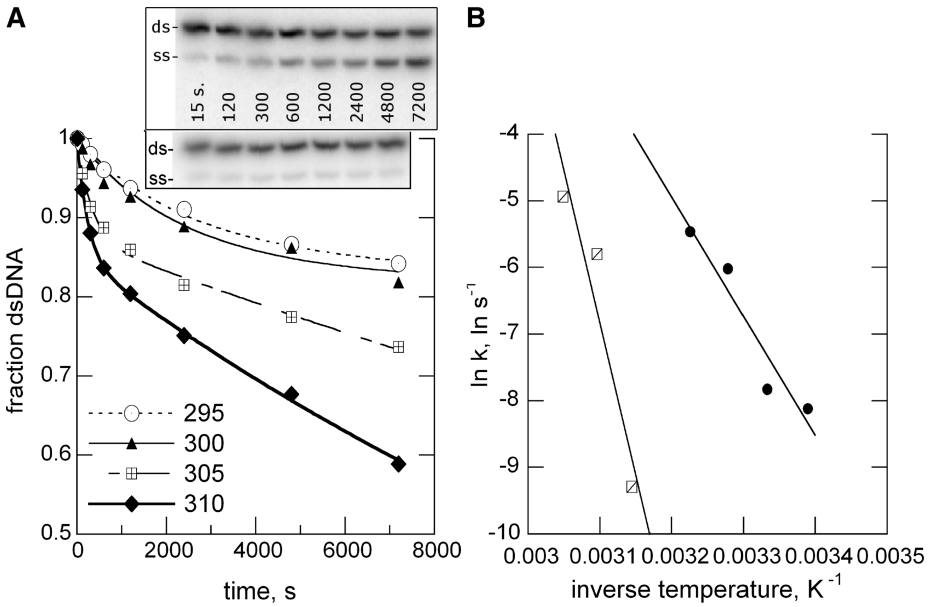

Figure 6. Temperature dependence of ORF1p-mediated strand displacement catalysis. (**A**) Phosphorimages of a ^{32}P labeled oligonucleotide melting from duplex in the presence of wt ORF1p (top inset) or without protein (bottom inset) at 37°C for the indicated times. Graph plots duplex melted at the four temperatures indicated in the presence of wt ORF1p, lines show fit to a two-factor exponential decay, Equation (5). (**B**) Plot of temperature dependence of overall reaction rate. The natural logarithm of the found overall rate plotted against the inverse temperature, with fitted line for presence (circles) and absence (squares) of wt ORF1p.

Table 1. Slopes and *y*-intercepts of lines fitted to points plotted according to transition state theory

ORF1p	Annealing (K•ln s^{-1})	Annealing $1/T = 0$	Exchange (K•ln s^{-1})	Exchange $1/T = 0$
Wt	−4060 ± 340 (200)	8 ± 1 (1)	−16 800 ± 900 (500)	48 ± 3 (2)
R284A	−3070 ± 740 (430)	5 ± 2 (1)	−20 300 ± 2200 (1300)	60 ± 7 (4)
R284K	−3670 ± 210 (120)	7 ± 1 (1)	−17 800 ± 3000 (1700)	52 ± 10 (6)
Y318A	−3180 ± 500 (290)	6 ± 2 (1)	−8200 ± 3900 (2300)	20 ± 13 (7)
Y318F	−3540 ± 300 (170)	7 ± 1 (1)	−14 800 ± 3700 (1800)	42 ± 12 (6)
No protein	−5970 ± 350 (200)	11 ± 1 (1)	−42 700 ± 4100 (2900)	125 ± 13 (9)

Slopes of lines fitted to points as graphed from Equation (3) express the temperature dependence of the reaction rate; more negative values denote a more temperature-dependent reaction and thus a higher enthalpic barrier to product formation. Intercepts are used according to Equation (3) to determine the relative magnitude of entropic contributions to energetic barriers. Values are expressed as ±standard deviation and (standard error). The points comprising these lines were derived for each mutant from experimental data collected as in Figures 5 and 6.

in K (here set to 310, or 37°C), and $\delta\Delta G_a$ is the total difference in the free energy of activation to product formation in the presence of ORF1p.

All mutant and wt ORF1 proteins catalyzed both annealing and strand displacement, although to differing magnitudes. The activation energy for annealing was reduced by 400–900 J mol^{-1} bp^{-1} compared to reactions without protein, and the energetic barrier to strand displacement was reduced by 9–15 kJ mol^{-1} bp^{-1}.

The change in free energy of melting was plotted as a function of that of annealing (Figure 7). This plot revealed an optimum magnitude for ORF1p-mediated annealing and strand-displacement catalytic activities in the retrotransposition-competent elements. All of the tested ORF1p variants exhibited an enhanced reduction in the energetic barrier to annealing, i.e. improved catalysis of annealing, compared to wt independent of retrotransposition activity. However, the retrotransposition-active mutants have both very similar reductions in the barrier to melting and annealing. In contrast, R284A shows strongly enhanced annealing and reduced melting. This strongly enhanced annealing is unable to overcome the reduction in the protein's ability to melt nucleic acids, which is clearly required for chaperone activity. Perhaps most interesting is the result for Y318A, which shows strongly enhanced melting and annealing, yet does not facilitate retrotransposition. This demonstrates that optimum melting and annealing is necessary for retrotransposition. This optimum melting and annealing is restored by substituting to form Y318F, which shows melting and annealing kinetics in these studies that are similar to wt.

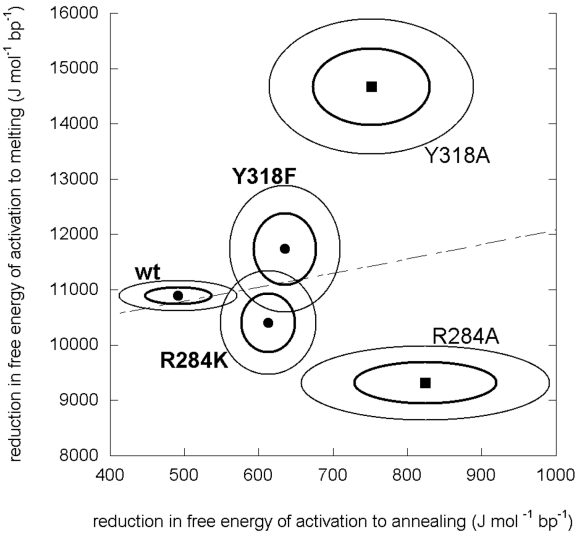

Figure 7. Retrotransposition requires a precise ratio between the reductions in the energetic barriers to strand displacement and annealing. Average energetic reductions for displacement and annealing (values from Table 1, multiplied by $-R$ (gas constant) and normalized to oligonucleotide length); ovals represent standard deviations (thin arcs) and standard errors (thick arcs) in the x and y dimensions. The three retrotransposition-competent ORF1p variants (bold, circles) group closely together, with a generally observed upward trend given by the regression line, slope = 2.57, $R = 0.29$, distinguishing them from the inactivating mutants (squares) by at least one standard error.

DISCUSSION

ORF1p from mouse L1 is a robust nucleic acid chaperone protein that is required for retrotransposition. Previously, substitutions that compromise L1 retrotransposition were mapped to single amino acids in either the coiled-coil (17) or the CTD (16) domains (D159H and R297K in Figure 1A, respectively). Both of these substitutions affected nucleic acid chaperone activity and not RNA binding affinity.

Here, an additional three highly conserved amino acids, two (R238, R284) in the recently described RRM domain (8) and the third, Y318, in the CTD (9), were replaced with alanine, and also found to inactivate L1 retrotransposition. Two of these three new alanine mutations were rescued by substitution of the appropriate amino acid to restore the basic (R284K) or hydrophobic (Y318F) character of the original arginine or tyrosine, whereas the third (R238K) was not (Figure 1). Retrotransposition deficits were independent of changes observed in steady-state levels of L1 RNA or ORF1p, or aberrant ORF1p structure (Figure 2 and Supplementary Figure S1B). Both RNA binding and nucleic acid chaperone activities of ORF1p appear to be critical for L1 retrotransposition because R238A and R238K compromised high-affinity RNA binding (Figure 3), and R284A and Y318A compromised nucleic acid chaperone activity based on reduced cooperativity of DNA melting (Figure 4), and poorly optimized annealing and strand displacement catalysis (Figure 7).

Residues R238 and R284 lie in the RRM domain; the structure of the corresponding fragment from human L1 ORF1p has been solved by X-ray crystallography, revealing that R202 (homologous to R238 in mouse L1) forms a salt bridge together with E169. Disruption of the salt bridge is expected to destabilize ORF1p (8). In the case of mouse L1, this salt bridge does not appear to be important for ORF1p stability because eliminating it with R238A did not alter the thermal melting behavior compared to wt (Supplementary Figure S1B) or reduce its abundance in cells (Figure 2B). Moreover, the salt bridge partner of R202, E169 (D204 in mouse), is rotated to the outside of ORF1p in the modeled structure (Supplementary Figure S3); thus it is unlikely that a salt bridge can form between these two residues in the mouse protein. Interestingly, R238K fails to restore the RNA affinity of ORF1p. Taken together, the available data indicate that R238 is important for interactions between ORF1p and nucleic acids, although the interactions involve more than simple electrostatics.

The second residue of the RRM domain studied here, R284, has not been examined previously. Based upon the crystal structure of the human homolog, R284 is on the surface of the RRM domain in the vicinity of several other arginine residues that form electrostatic interactions with RNA (8). The interpretation that this residue is important for electrostatic interactions is supported by the observed inactivation of retrotransposition by R284A and restoration of activity by R284K. We were unable to measure a significant reduction in RNA affinity in R284A, although a reproducible trend in this direction was observed. The nucleic acid chaperone activity of R284A is, however, significantly altered by an imbalance of annealing and displacement activities because of a markedly decreased ability to catalyze the displacement reaction. In contrast, R284K restores the wt balance between annealing and strand displacement, suggesting that the electrostatic interaction between R284 and nucleic acids is important for the melting component of the nucleic acid chaperone activity.

The final residue examined, Y318 in the CTD, is conserved in placental mammals, marsupials and frogs. The homologous site, however, is substituted with phenylalanine in swimmer, a LINE element in Japanese medaka (8), consistent with the observed retrotransposition with Y318F here. Y318F also restored the balance between the melting and annealing activities of ORF1p that was lost in Y318A. Because annealing catalysis was not significantly different between the alanine and phenylalanine variants at residue 318, the biochemical problem with Y318A appears to be its excessive basepair melting catalysis. This residue is likely involved in hydrophobic or steric interactions because the terminal hydroxyl group on tyrosine is dispensable.

Based upon the results of earlier studies of wt and mutant ORF1 proteins, we proposed that at least some of the interactions involved in high-affinity binding of ORF1p to RNA are distinct from those involved in the nucleic acid chaperone activity (16,17), i.e. that these reactions involve distinct sites on the protein. The present results provide further support for this hypothesis. If the melting activity in the strand displacement assay is functionally equivalent to the RNA binding affinity, then the paired variants at residues 284 and 318 should have been indistinguishable in the strand displacement assay as they were in the RNA binding assay, but they were not.

A 3D structure of the intact trimer with nucleic acids will be necessary to fully understand the relationship between the structure of ORF1p and its RNA binding and nucleic acid chaperone activities. Nevertheless, it seems unlikely that all of the residues that affect the chaperone activity of ORF1p are closely juxtaposed in the tertiary structure of the protein, because they are so widely distributed along its primary structure. Instead we predict that residues responsible for chaperone activity are distributed in the folded structure as well as the primary sequence of this protein.

When taken together and compared with retrotransposition experiments, the *in vitro* single molecule and solution experiments demonstrate the importance of the finely balanced nucleic acid chaperone capabilities of ORF1p for biological activity. They also show that the nucleic acid chaperone activity of ORF1p requires optimal interactions with ssDNA and dsDNA that result in a reduction in melting cooperativity, as well as an optimal enhancement of strand annealing and strand displacement capabilities. This enhancement of annealing and displacement is likely a result of the combination of the ability of ORF1p to facilitate strand attraction as well as nucleic acid unwinding. Specific single amino acid substitutions in ORF1p inhibit this chaperone activity by interfering with at least one of these capabilities, as demonstrated by the properties of the mutations at R284 and Y318. A mutation that strongly enhanced both melting and annealing (Y318A) inhibited retrotransposition, but a compensatory mutation that reduced the melting and annealing capabilities of ORF1p and restored these properties to wt levels also restored retrotransposition. Similarly, a mutation that enhanced annealing but significantly reduced the ability of ORF1p to facilitate melting also inhibited retrotransposition. A compensatory mutation that restored the kinetics to wt values also restored retrotransposition activity (Figure 7).

Other groups have published kinetic data on DNA or RNA annealing (21,35,37–39), DNA melting as an intermediate step to annealing (21), steady-state duplex destabilization, and chaperone effects on force-induced DNA melting (29,40,41). Although it has been suggested that the melting and annealing activities of nucleic acid chaperones are in a tightly constrained balance (42) this is, to our knowledge, the first work to quantitatively determine an optimum kinetics of strand displacement and annealing that can be destroyed and restored with specific point mutations.

Finally, it is noteworthy that all of the ORF1p mutations that retained wt RNA affinity were nevertheless nucleic acid chaperone proteins—all ORF1p variants tested accelerated both base pair annealing and melting, and presumably allow for the rearrangement of secondary structure as a consequence, regardless of their retrotransposition competence. The property that discriminates between biologically active and inactive variants of

ORF1p, however, is the balanced ratio of the energetic barriers to annealing and melting. This finding implies that L1 retrotransposition depends upon an ORF1p that provides the two optimized energetic reductions in specific kinetic properties, consistent with observations that dispersed mutations in ORF1p could disrupt this finely tuned balance and inactivate retrotransposition. These quantities must have a basis in the specific nucleic acid rearrangements that are required for L1 retrotransposition. It may be a general feature of nucleic acid chaperone proteins to maintain specifically tuned energetic reductions to annealing and melting that satisfy their biological function, as determined here for L1 ORF1p.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank D. Branciforte, D. Bushman and J. Cummiskey for constructs, and Dr I. Rouzina, Dr D. Bain and Dr K. Maluff for valuable discussion.

FUNDING

National Institutes of Health Grants (R01GM40367 to S.L.M.); (P30 CA046934 for the Protein Production Core of the Univ Colorado Cancer Center); (R01GM75965 to M.C.W.); National Science Foundation (MCB-0744456 to M.C.W.). Funding for open access charge: National Institutes of Health Grant (R01GM40367 to S.L.M.).

Conflict of interest statement. None declared.

REFERENCES

1. Cordaux,R. and Batzer,M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.*, **10**, 691–703.
2. Han,J.S. and Boeke,J.D. (2005) LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? *Bioessays*, **27**, 775–784.
3. Goodier,J.L., Ostertag,E.M., Du,K. and Kazazian,H.H.J. (2001) A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.*, **11**, 1677–1685.
4. Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H. Jr. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917–927.
5. Wei,W., Gilbert,N., Ooi,S.L., Lawler,J.F., Ostertag,E.M., Kazazian,H.H., Boeke,J.D. and Moran,J.V. (2001) Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.*, **21**, 1429–1439.
6. Luan,D.D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chomosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, **72**, 595–605.
7. Martin,S.L., Branciforte,D., Keller,D. and Bain,D.L. (2003) Trimeric structure for an essential protein in L1 retrotransposition. *Proc. Natl Acad. Sci. USA*, **100**, 13815–13820.
8. Khazina,E. and Weichenrieder,O. (2009) Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc. Natl Acad. Sci. USA*, **106**, 731–736.
9. Januszyk,K., Li,P.W., Villareal,V., Branciforte,D., Wu,H., Xie,Y., Feigon,J., Loo,J.A., Martin,S.L. and Clubb,R.T. (2007) Identification and solution structure of a highly conserved C-terminal domain within ORF1p required for retrotransposition of long interspersed nuclear element-1. *J. Biol. Chem.*, **282**, 24893–24904.
10. Kolosha,V.O. and Martin,S.L. (2003) High affinity, non-sequence-specific RNA binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). *J. Biol. Chem.*, **278**, 8112–8117.
11. Basame,S., Li,P.W.-L., Howard,G., Branciforte,D., Keller,D. and Martin,S.L. (2006) Spatial assembly and RNA binding stoichiometry of a protein essential for LINE-1 retrotransposition. *J. Mol. Biol.*, **357**, 351–357.
12. Hohjoh,H. and Singer,M.F. (1996) Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *EMBO J.*, **15**, 630–639.
13. Kolosha,V.O. and Martin,S.L. (1997) In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl Acad. Sci. USA*, **94**, 10155–10160.
14. Martin,S.L. (1991) Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol. Cell. Biol.*, **11**, 4804–4807.
15. Martin,S.L., Li,J. and Weisz,J.A. (2000) Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *J. Mol. Biol.*, **304**, 11–20.
16. Martin,S.L., Cruceanu,M., Branciforte,D., Wai-Lun Li,P., Kwok,S.C., Hodges,R.S. and Williams,M.C. (2005) LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J. Mol. Biol.*, **348**, 549–561.
17. Martin,S.L., Bushman,D., Wang,F., Li,P.W.-L., Walker,A., Cummiskey,J., Branciforte,D. and Williams,M.C. (2008) A single amino acid substitution in ORF1 dramatically decreases L1 retrotransposition and provides insight into nucleic acid chaperone activity. *Nucleic Acids Res.*, **18**, 5845–5854.
18. Martin,S.L. and Bushman,F.D. (2001) Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell. Biol.*, **21**, 467–475.
19. Kulpa,D.A. and Moran,J.V. (2005) Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum. Mol. Gen.*, **14**, 3237–3248.
20. Liu,H.W., Cosa,G., Landes,C.F., Zeng,Y., Kovaleski,B.J., Mullen,D.G., Barany,G., Musier-Forsyth,K. and Barbara,P.F. (2005) Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription. *Biophys. J.*, **89**, 3470–3479.
21. Stewart-Maynard,K.M., Cruceanu,M., Wang,F., Vo,M.N., Gorelick,R.J., Williams,M.C., Rouzina,I. and Musier-Forsyth,K. (2008) Retroviral nucleocapsid proteins display nonequivalent levels of nucleic acid chaperone activity. *J. Virol.*, **82**, 10129–10142.
22. Zhang,A., Derbyshire,V., Salvo,J.L. and Belfort,M. (1995) Escherichia coli protein StpA stimulates self-splicing by promoting RNA assembly in vitro. *RNA*, **1**, 783–793.
23. Chakshusmathi,G., Kim,S.D., Rubinson,D.A. and Wolin,S.L. (2003) A La protein requirement for efficient pre-tRNA folding. *EMBO J.*, **22**, 6562–6572.
24. Chaurasiya,K.R., Paramanathan,T., McCauley,M.J. and Williams,M.C. (2010) Biophysical characterization of DNA binding from single molecule force measurements. *Phys. Life Rev.*, **7**, 299–341.
25. Wu,H., Rouzina,I. and Williams,M.C. (2010) Single-molecule stretching studies of RNA chaperones. *RNA Biol.*, **7**, 73–84.
26. McCauley,M.J. and Williams,M.C. (2007) Mechanisms of DNA binding determined in optical tweezers experiments. *Biopolymers*, **85**, 154–168.
27. Williams,M.C., Rouzina,I. and Bloomfield,V.A. (2002) Thermodynamics of DNA interactions from single molecule stretching experiments. *Acc. Chem. Res.*, **35**, 159–166.
28. Williams,M.C., Rouzina,I., Wenner,J.R., Gorelick,R.J., Musier-Forsyth,K. and Bloomfield,V.A. (2001) Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein

revealed by single molecule stretching. *Proc. Natl Acad. Sci. USA*, **98**, 6121–6126.
29. Cruceanu,M., Urbaneja,M.A., Hixson,C.V., Johnson,D.G., Datta,S.A., Fivash,M.J., Stephen,A.G., Fisher,R.J., Gorelick,R.J., Casas-Finet,J.R. *et al.* (2006) Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.*, **34**, 593–605.
30. Spector,S., Wang,M., Carp,S.A., Robblee,J., Hendsch,Z.S., Fairman,R., Tidor,B. and Raleigh,D.P. (2000) Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry*, **39**, 872–879.
31. Goldberg,A.L. (2003) Protein degradation and protection against misfolded or damaged proteins. *Nature*, **426**, 895–899.
32. Williams,M.C. and Rouzina,I. (2002) Force spectroscopy of single DNA and RNA molecules. *Curr. Opin. Struct. Biol.*, **12**, 330–336.
33. Shokri,L., McCauley,M.J., Rouzina,I. and Williams,M.C. (2008) DNA overstretching in the presence of glyoxal: structural evidence of force-induced DNA melting. *Biophys. J.*, **95**, 1248–1255.
34. Williams,M.C., Gorelick,R.J. and Musier-Forsyth,K. (2002) Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc. Natl Acad. Sci. USA*, **99**, 8614–8619.
35. Qualley,D.F., Stewart-Maynard,K.M., Wang,F., Mitra,M., Gorelick,R.J., Rouzina,I., Williams,M.C. and Musier-Forsyth,K. (2010) C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 nucleocapsid protein via an electrostatic mechanism. *J. Biol. Chem.*, **285**, 295–307.
36. Shokri,L., Rouzina,I. and Williams,M.C. (2009) Interaction of bacteriophage T4 and T7 single-stranded DNA-binding proteins with DNA. *Phys. Biol.*, **6**, 025002.
37. Vo,M.N., Barany,G., Rouzina,I. and Musier-Forsyth,K. (2006) Mechanistic studies of mini-TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. *J. Mol. Biol.*, **363**, 244–261.
38. Sharma,K., Didier,P., Darlix,J.L., de Rocquigny,H., Bensikaddour,H., Lavergne,J.P., Penin,F., Lessinger,J.M. and Mely,Y. (2010) Kinetic analysis of the nucleic acid chaperone activity of the hepatitis C virus core protein. *Nucleic Acids Res.*, **38**, 3632–3642.
39. Urbaneja,M.A., Wu,M., Casas-Finet,J.R. and Karpel,R.L. (2002) HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J. Mol. Biol.*, **318**, 749–764.
40. Cruceanu,M., Gorelick,R.J., Musier-Forsyth,K., Rouzina,I. and Williams,M.C. (2006) Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. *J. Mol. Biol.*, **363**, 867–877.
41. McCauley,M.J. and Williams,M.C. (2009) Optical tweezers experiments resolve distinct modes of DNA-protein binding. *Biopolymers*, **91**, 265–282.
42. Levin,J.G., Guo,J., Rouzina,I. and Musier-Forsyth,K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.*, **80**, 217–286.