Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer

Bahram Arezi* and Holly Hogrefe

Agilent Technologies, Stratagene Products Division, La Jolla, CA 92037, USA

ABSTRACT

In an effort to increase the thermostability of Moloney Murine Leukemia Virus reverse transcriptase (MMLV RT), we screened random and site-saturation libraries for variants that show increased resistance to thermal inactivation. We discovered five mutations E69K, E302R, W313F, L435G and N454K that collectively increase the half-life of MMLV RT at 55°C from less than 5 min to ~30 min in the presence of template-primer. In addition, these mutations alter the thermal profile by increasing specific activity of the pentuple mutant (M5) over a broad range of cDNA synthesis temperatures (25–70°C). We further show that M5 generates higher cDNA yields and exhibits better RT–PCR performance compared to wild-type RT when used at high temperature to amplify RNA targets containing secondary structure. Finally, we demonstrate that M5 exhibits tighter binding (lower K_m) to template-primer, which likely protects against heat inactivation.

INTRODUCTION

Reverse transcriptases are derived from RNA-containing retroviruses and are either monomeric (e.g. Moloney Murine Leukemia Virus, MMLV RT) or heterodimeric (e.g. Avian Myeloblastosis Virus, AMV RT) in structure. Reverse transcriptases are multi-functional enzymes with three enzymatic activities including RNA- and DNA-dependent DNA polymerization activity, and an RNase H activity that catalyses the cleavage of RNA in RNA–DNA hybrids (1).

MMLV RT is widely used in a number of research applications, including cDNA cloning, end-point and real-time RT–PCR, microarray analysis and RACE. The associated RNase H activity of MMLV RT, which is instrumental to biological activity *in vivo*, limits the efficiency of synthesizing long cDNA *in vitro* (2,3). This problem was overcome with the development of mutants that exhibit diminished RNase H activity (H⁻ MMLV RT), while retaining full RNA-directed DNA polymerase activity (4). Besides providing improved synthesis of full-length cDNA, H⁻ MMLV RT exhibits higher thermo-stability (4-fold longer half-life at 50°C) compared to wild-type MMLV RT (5).

Another challenge to copying RNA is the presence of RNA secondary structure which can stall cDNA synthesis (6,7), resulting in truncated cDNA molecules. Interference from RNA secondary structure can be overcome by performing cDNA synthesis at higher temperature. For example, MMLV RT is generally recommended for use at 37°C (or 42°C for H⁻ MMLV RT), and increased temperatures (e.g. by up to 8°C) are advocated for relaxing RNA secondary structure and improving transcription of problematic sequences. Increasing cDNA synthesis temperature has also been reported to reduce non-specific priming and increase MMLV RT fidelity by lowering the stability of mismatches (8). Ultimately, however, cDNA length and yield are compromised by lower enzyme activity and reduced RNA integrity at elevated temperatures (9).

Although several strategies have been used to facilitate cDNA synthesis above 50°C, each has disadvantages that limit utility. For example, cDNA synthesis can be performed at PCR temperatures using thermophilic DNA-dependant DNA polymerases with low intrinsic reverse transcriptase activity (10–12), that can be further enhanced by substituting Mn^{2+} for Mg^{2+} (13). Under these conditions, pausing at secondary structures is minimized; however, cDNA synthesis itself is relatively inefficient and highly error-prone [Mn^{2+} is mutagenic (14)]. To increase efficiency and fidelity, thermophilic DNA polymerases have been further engineered to improve utilization of RNA templates in the presence of Mg^{2+} (15–18); those that are commercially available allow single-enzyme RT–PCR but target-length capability is limited. Methods that provide efficient reverse

*To whom correspondence should be addressed. Tel: +1 858 373 6389; Fax: +1 858 373 5300; Email: bahram.arezi@agilent.com

transcription at higher non-PCR temperatures (up to 55–65°C) include the use of avian RTs (e.g. AMV RT) and enhancing the thermoactivity of MMLV RT with stabilizers. While AMV RT retains more DNA synthetic activity at elevated temperatures than MMLV RT (5,19) (T_{opt} 45–50°C; recommended up to 60°C), avian RTs are less commonly used perhaps because the reverse transcription yield (cDNA conversion efficiency) for AMV RT is significantly lower than MMLV RT (2% for AMV compared to 44% for MMLV) (20).

In the presence of trehalose, H⁻ MMLV RT was shown to retain full activity up to 60°C; however, effective concentrations (0.6 M) are inhibitory to PCR, precluding its use in one-tube RT–PCR formats (21). Efforts to engineer MMLV RT mutants with increased thermostability have also been described (22). As mentioned above, MMLV RT mutants lacking RNase H activity also show greater resistance to thermal inactivation in the presence of template-primer. In this case, differences in thermostability are attributed to differences in the degree of protection provided by intact (H⁻ MMLV RT) versus degraded RNA:DNA templates (RNA portion degraded by wild-type RNase H activity) (5). Additional mutations have been reported to increase the intrinsic stability of H⁻ MMLV RT without affecting template binding (23); however, the identity of such mutations has not been disclosed (e.g. SuperScript III RT, Invitrogen).

In this study, we expand upon earlier efforts to increase the thermal activity of MMLV RT. Using random and site-directed mutagenesis in combination with high-temperature activity screens, we successfully identify multiple point mutations that collectively increase the half-life of MMLV RT at elevated temperature. As we will show, the resulting pentuple mutant (MMLV RT M5) generates full-length cDNAs up to 55°C and shows great utility in the amplification of RNA targets with difficult secondary structures.

MATERIALS AND METHODS

PicoMaxx® DNA polymerase (blend of Taq and Pfu), human total RNA (skeletal muscle, HeLa), Human Universal Reference total RNA, GeneMorph® II EZClone Domain Mutagenesis Kit, StrataPrep® Plasmid Miniprep kit, QuikChange® Site-Directed Mutagenesis Kit, QuikChange® Multi Site-Directed Mutagenesis Kit, and BL21-CodonPlus® (DE3)-RIL competent cells were from Stratagene. Oligo(dG)₁₈ was synthesized by TriLink BioTechnologies. Poly(rC) was purchased from Amersham. RNA Millenium size marker was purchased from Ambion (#7150). Costar 96-well plates (29444-102) and DE81 filters (21426-088) were from VWR. Airpore tape sheets (19571) were from Qiagen. EasyTides deoxyguanosine [α³³P] (NEG614H001MC) was purchased from PerkinElmer. His-tagged proteins were purified in native form as described in QIAexpressionist (June 2003). Untagged reverse transcriptases were purified by sequential purification over anion exchange (Q), phophocellulose, poly-rU and SP cation exchange columns.

Generating random mutant RT library

A 2 kb sequence encoding an RNaseH-deficient MMLV RT mutant (MMLV RT D524N) was randomly mutagenized using the GeneMorph® Random Mutagenesis Kit and primers pSTRAT-F and pSTRAT-R (Supplementary Table 1) according to manufacturer's recommendations (100 ng plasmid, 20 PCR cycles). Mutated PCR products served as 'mega primers' to replace the parental DNA sequence using the 'domain swapping' protocol from the GeneMorph® II EZClone Domain Mutagenesis Kit. Resulting plasmids were cloned into XL-10 Gold competent cells. Library size was 5×10^4 (total) and clones contained roughly 1–6 mutations per kilo basepair based on sequencing of 20 random clones. Colonies were scraped off plates and plasmid DNA was extracted from the entire library using the StrataPrep® Plasmid Miniprep Kit. To minimize clone redundancy, a portion of the plasmid DNA was then transformed into BL21-CodonPlus® (DE3)-RIL competent cells to generate a library size of 5×10^4.

RT thermostability screening

Mutant BL21-DE3-RIL colonies were inoculated into 120 μl LB media (containing 100 μg/ml ampicillin and 35 μg/ml chloramphenicol) in Costar 96-well plates sealed with Airpore tape sheets, and grown over night at 37°C. Ten microlitres of each culture was inoculated into 110 μl LB media (with antibiotics and 1 mM IPTG) and grown overnight. Cells were lysed using 30 μl lysis buffer (125 mM Tris pH 8, 4.5% glucose, 50 mM EDTA, 2.5% Triton, 5 mg/ml lysozyme, 50 mM DTT). Reverse transcriptase activity was measured by adding 10 μl of lysate to 50 μl (total volume) reaction cocktails containing 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 8 mM MgCl₂, 2 μg poly(rC), 0.5 μg oligo(dG)₁₈, 10 mM DTT, 50 μM dGTP and 0.5 μCi α³³pdGTP. Reactions were incubated at 42°C or 55°C for 60 min. Four microlitres of each reaction was spotted on DE-81 filters, and dried. The filters were washed five times with 2× SSC and dried. The filters were then exposed to Kodak BioMax MR-1 films (VWR IB8941114) overnight.

Saturation and multi-site mutagenesis

Saturation mutagenesis was performed using the QuikChange® Site-Directed Mutagenesis Kit and primer pairs containing one degenerate (sense: 5′-NNG/T; antisense: 5′-A/CNN) codon at E69, E302, F303, G305, W313, L435 or N454 (Supplementary Table 1). Ten nanograms of plasmid DNA and 125 ng of each primer were used according to the QuikChange® kit. Five colonies from each library were sequenced, and resulting libraries exhibited 75–100% mutant frequencies.

Independent mutations were combined by generating a mutant library with the QuikChange® Multi Site-Directed Mutagenesis Kit, pMMLV RT E302R/D524N plasmid DNA (for H⁺ M5: pMMLV RT E302R plasmid DNA), and primers pE69K, pW313F, pL435G and pN454K (Supplementary Table 1). Ten clones from each library were sequenced.

RT activity assays

RNA-dependent DNA polymerase activity was assayed as described above (RT thermostability screening) in $50 \mu l$ reactions containing $poly(rC)/oligo(dG)_{18}$ and $0.5 pmol$ RT. Incorporated radioactivity was measured by scintillation counting. Reactions that lacked enzyme were set up along with sample incubations to determine 'total cpms' (omitted filter wash steps) and 'minimum cpms'(washed filters as above). Sample cpms were subtracted by minimum cpms to determine 'corrected cpms'.

Full-length cDNA profiling was performed using a poly(A)-tailed RNA ladder (Ambion). Reactions $(20 \mu l)$ contained $1 \mu g$ RNA ladder, $0.5 \mu g$ $oligo(dT)_{18}$, $4 mM$ total dNTPs, $192 ng$ RT and $1 \times$ RT buffer [consisting of $50 mM$ Tris–HCl (pH 8.3), $75 mM$ KCl and $3 mM$ $MgCl_2$]. Reactions were incubated at varying temperatures ($42–60°C$) for 60 min, and then electrophoresed on 1% alkaline agarose gels that were stained with SYBR Gold.

Circular dichroism (CD) spectra measurements

CD measurements were performed at UMDNJ (University of Medicine and Dentistry of New Jersey)—Circular Dichroism facility. Measurements were obtained on an Aviv model 400 spectropolarimeter with a 5-cuvette holder (Lakewood, NJ), using reverse transcriptase preparations dialysed against $20 mM$ potassium phosphate buffer (pH 8.0) containing $100 mM$ KCl. Protein concentrations were at $0.3 mg/ml$ for CD spectra measurements at different temperatures, and at $0.2 mg/ml$ for melting CD spectra.

Two-step RT–PCR assays

cDNA synthesis was carried out in reactions $(20 \mu l)$ containing $0.5 \mu g$ $oligo(dT)_{18}$, $4 mM$ total dNTPs, $10 mM$ DTT, $1 \times$ RT buffer, and the indicated amounts of RNA and reverse transcriptase. cDNA incubations were carried out at indicated temperatures. Information on the RT–PCR systems employed is provided in Supplementary Data (Tables 1 and 2).

Steady-state kinetic measurements for template-primer

K_m and V_{max} values for template-primer were measured using limiting amounts of enzyme (0.03 pmol) in the presence of $100 \mu M$ dGTP, α^{33}p-dGTP (3 μCi/pmol), $1 \times$ RT buffer, and varying concentrations of template-primer (1–10 000 nM). Samples were incubated at $37°C$ for 5 min, and reactions were quenched with ice-cold 0.2 M EDTA (final concentration). Five microlitres of each reaction was spotted on DE-81 filters, and incorporated radioactivity was measured by scintillation counting as described above.

RESULTS

MMLV RT mutagenesis

Round 1—screening random mutant library. A sequence encoding H⁻ MMLV RT (MMLV RT D524N) was randomly mutagenized, and mutant colonies were tested for increased activity at $55°C$ as described (Materials and Methods section). We used an RNase H⁻ backbone to prevent isolating clones that were more thermostable by virtue of diminished RNase H activity (5). Typical results obtained from cultures grown, lysed and assayed in a 96-well format are shown in Supplementary Data (Figure 1). We screened 3400 clones from the primary library and selected 116 clones that showed higher activity at $55°C$ compared to H⁻ MMLV RT. To account for differences in protein expression, positives were re-assayed three times and seven promising candidates were selected, expressed and affinity purified. Increased thermostability was assessed by measuring reverse transcriptase activity with synthetic $poly(rC):oligo(dG)_{18}$ substrate at $42°C$ and $52°C$, and calculating percent activity at $52°C$ compared to $42°C$ (T_{opt} for H⁻ MMLV RT). As shown in Table 1, mutant clones 70, 90, 98, 113 and 115 exhibited improved thermostability ($52°C/42°C$ activity ratios: $\geq 32\%$) compared to H⁻ MMLV RT ($52°C/42°C$ activity ratio: 20%).

Sequencing results (Table 1) identified only one point mutation in Mut70 (E69K), Mut90 (T306R), Mut98 (N454K) and Mut113 (L435M), while Mut115 contained four mutations including S23F, E176D, L351I and G608A. Mutations at T306 (T306K) and F309 (F309N) were previously shown to increase resistance against thermal inactivation (22). As this region may contribute to MMLV RT thermostability, we further investigated residues E302, F303, G305 and W313, whose side chains lie in close proximity to those of T306 and F309 in the MMLV RT crystal structure (24). We hypothesized
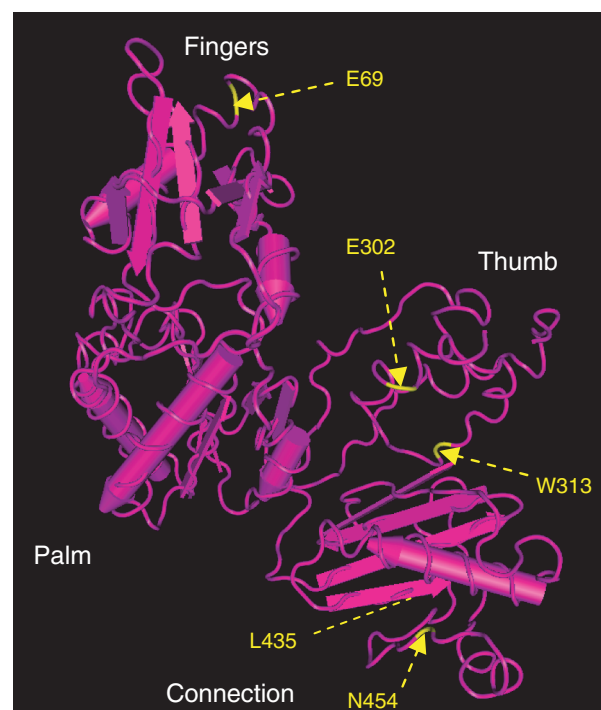
Figure 1. Structure of MMLV RT highlighting the locations of thermostable mutations. The crystal structure of MMLV RT (consisting of fingers, palm, thumb and connection domains) lacking the RNase H domain was adopted from Das and Georgiadis (24). E69 resides in the fingers, E302 and W313 in the thumb, and L435 and N454 in the connection domain.

that neighboring residues might show similar effects when substituted with a more highly charged side chain (e.g. akin to T→R/K and F→N replacements at 306 and 309, respectively).

Round 2—saturation mutagenesis of residues that increase heat resistance. Residues E69, E302, F303, G305, W313, L435 and N454 were independently subjected to saturation mutagenesis. For NNG/T codons, the frequency of the least-represented mutants can be calculated as 1/32 (1/4 × 1/4 × 1/2). Assuming that the mutation efficiency is 100%, there is more than a 99% chance of detecting all possible mutants in a sample of 200 clones ($0.99 = 1 - (1 - f)^n$, where f is the frequency of the least-represented mutants and n the number of clones screened) (25). Therefore, we screened 200 clones from each of the seven saturation libraries to find the side chain replacement that provides the highest activity at 55°C. As Table 2 shows, the most effective substitutions at positions E69, W313 and L435 are lysine (K), phenylalanine (F) and glycine (G), respectively. The predominant mutations at residues E302 and N454 are arginine (R; 17 out of 20 clones), and lysine (K; three out of four clones), respectively. We failed to recover thermal resistant mutants from F303 and G305 saturation libraries.

Round 3—combinations of thermal-resistance mutations. To this point, five sites that independently increase RT activity at 55°C were identified by random (E69, N454, L435) or rational design (E302, W313) approaches, and the most effective substitution at each position was selected using saturation mutagenesis. We then examined various combinations of the most effective mutations by creating a library of multi-site mutants using pMMLV RT

Table 1. Polymerase activity of random mutants at elevated temperature

RT	Mutation(s)	Percentage activity at 52°C/42°C
Wild-type	–	20% ± 1%
Mut70	E69K	37% ± 0.2%
Mut84	No data	22% ± 0.7%
Mut86	No data	21% ± 0.5%
Mut90	T306R	73% ± 2.3%
Mut98	N454K	32% ± 1%
Mut113	L435M	36% ± 0.7%
Mut115	S23F, E176D, L351I, G608A	34% ± 2%

Table 2. Side-chain replacements providing highest thermal resistance

Randomized amino acid	Clones sequenced	Preferred mutation(s) (number of clones)
E69	2	K
E302	20	K (3), R (17)
F303	–	–
G305	–	–
W313	2	F
L435	4	G
N454	4	K (3), R (1)

E302R/D524N plasmid DNA and a combination of pE69K, pW313F, pL435G and pN454K primers. Clones (containing one to five mutations each) were assayed for activity at the more restrictive temperature of 57°C. In general, when the mutations are combined, the effects on thermostability appear to be additive, as shown by an increase in activity with increasing number of mutations (Supplementary Figure 2): (lowest) wt < R, K < RF < RFG < RFGK, RKFGK (highest).

Figure 1 shows the location of each mutation in the crystal structure of MMLV RT [lacking the RNase H domain; adopted from Das and Georgiadis (24)]. E69 is located in the fingers domain, E302 and W313 are in the thumb, and L435 and N454 are in the connection domain. Therefore, the mutations are distributed in three distinct functional domains of MMLV RT. Fingers domain has been proposed to provide an intermediate binding site for template-primer in between phosphonucleotidyl transfer reactions (26). The thumb domain is also suggested to play a role in substrate binding and processivity (27), while the connection domain offers conformational flexibility between the polymerase and the RNase H domain (24).

Temperature optimum

The H⁻ MMLV RT pentuple mutant (E69K/E302R/W313F/L435G/N454K), hereafter referred to as H⁻ M5, was subcloned to remove the affinity tag and purified (Materials and Methods section). Comparisons indicate that H⁻ M5 exhibits the same reaction condition optima (pH, salt and Mg^{2+}) and error rate as H⁻ MMLV RT (data not shown). We compared the thermal profile of H⁻ M5 to H⁻ MMLV RT using poly(rC):oligo(dG)$_{18}$. As shown in Figure 2, H⁻ M5 exhibits a broad thermal profile with nearly full (>85%) activity between 40°C and 55°C. In contrast, the range of peak activity for H⁻ MMLV RT is narrower, with a maximum at 40–50°C. H⁻ M5 also shows significantly higher activity at both extremes (≤40°C and ≥55°C) compared to H⁻ MMLV RT. Similar experiments were carried out with an RNase H-proficient version of the pentuple mutant, hereafter referred to as M5 (reverted to D524). M5 exhibits
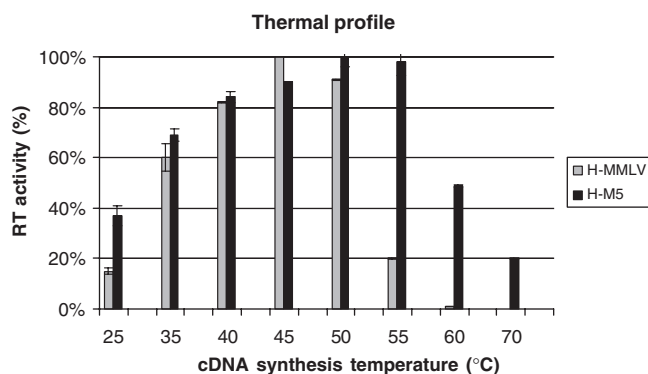

Figure 2. Thermal profiles. Relative RT activities were measured (Materials and Methods section) after 5 min incubations in the presence of template-primer at the temperatures indicated. Percent activity was normalized relative to the maximum activity exhibited by each RT (100%).

virtually the same thermal profile as H⁻ M5, showing >90% activity at 55°C (data not shown).

Full-length cDNA synthesis at elevated temperature

To assess full-length cDNA synthesis as a function of temperature, we used a mixture of cRNAs ranging in length from 0.5 to 9 kb as template [poly(A)-tailed RNA ladder]. In reverse transcription reactions employing H⁻ MMLV RT, cDNA length decreases with increasing temperature, and at 55°C, no full-length product of any size is visible (Figure 3). In contrast, both H⁺ and H⁻ versions of M5 synthesize full-length cDNA (up to 9 kb) over a broad range of temperatures (42–55°C). Since these assays were
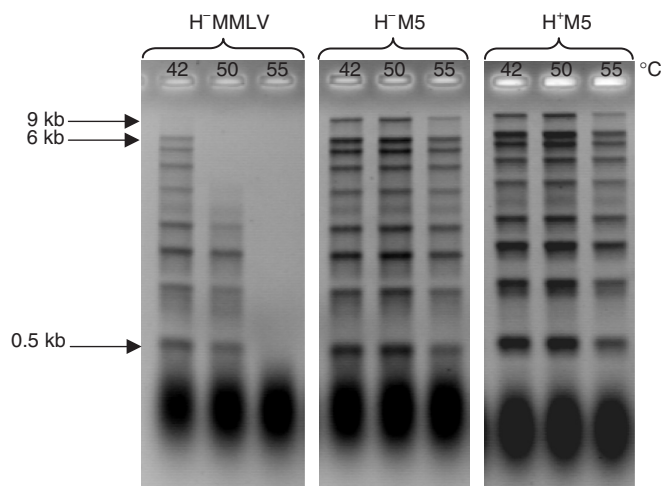
Figure 3. Full-length cDNA profiling using poly(A)-tailed RNA ladder. Reactions contained 192 ng (Materials and Methods section) of each enzyme and were incubated at the indicated temperatures.

performed with identical amounts of protein, we can conclude that M5 exhibits higher specific activity on natural (mixed base) RNA templates compared to H⁻ MMLV RT at its optimal temperature of 42°C. Moreover, the RNase H activity of M5 does not appear to appreciably affect yield of long cDNAs.

Contribution of template-primer to thermal resistance

Enzyme half-life ($t_{1/2}$) was measured in the presence and absence of template-primer at 55°C, the maximum temperature where M5 shows near full activity (Figure 2). In the presence of template-primer, $t_{1/2}$ values were approximately 30 min and less than 5 min for H⁻ M5 and H⁻ MMLV RT, respectively (Supplementary Figure 3). In the absence of template-primer, however, activities of both enzymes are at background levels after 5 min at 55°C. These results suggest that the M5 mutations improve thermostability through tighter binding to template-primer rather than increased intrinsic stability.

CD was then employed to further investigate the contribution of template-primer binding to heat resistance of M5 (Figure 4). CD spectra of M5 and MMLV RT were obtained in the absence and presence of poly(rC):oligo(dG)₁₈. As shown in Figure 4A, the spectra are virtually identical for wild-type and mutant enzymes in the absence of template-primer, and consistent with high alpha-helix content (24). Differences in ellipticity between M5 and MMLV RT can be attributed to differences in protein concentration (as confirmed by SDS–PAGE analysis; data not shown).

Melting CD spectra were measured at 212.5 nm, where ellipticity changes were most pronounced for enzyme samples and minimal for template-primer (alone). As temperature increased over the range of 25–80°C, all protein samples showed signs of aggregation immediately after unfolding, including the formation of visible precipitate,
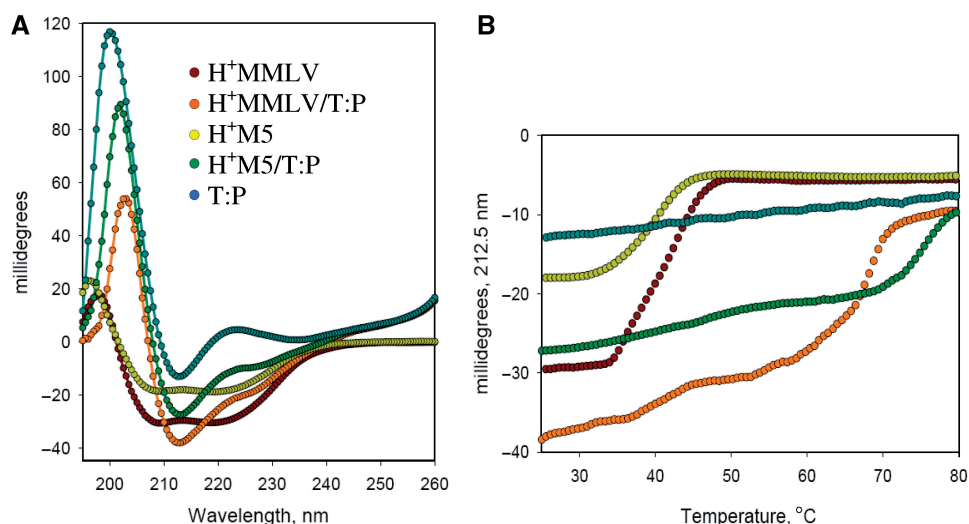
Figure 4. CD spectra. The CD spectra of MMLV RT and M5 were measured in the absence and presence of template-primer (T:P) at 25°C (**A**). Melting CD spectra were measured at 212.5 nm across a range of temperatures, from 25°C to 80°C (**B**). Protein spectra (reported in millidegrees) were not presented in terms of molar ellipticites due to contributions from template-primer. Enzyme/template-primer spectra were not the sum of individual enzyme and template-primer spectra (indicative of conformation changes accompanying complex formation; data not shown), and thus, could not be corrected by subtracting the primer-template spectra and computing mean residue ellipticity from calculated mean amino-acid weight.

a reduction in ellipticity values (consistent with reduced protein concentration), and the absence of spectral changes that accompany conversion of ordered structure to random coil (data not shown). As shown in Figure 4B, MMLV RT and M5 start to unfold (and aggregate) at $\sim 34°C$ in the absence of template-primer. In the presence of template-primer, both enzymes are significantly more resistant to heat-denaturation, and calculated transition temperatures (first derivative of melting spectra) increase dramatically to $68°C$ (MMLV RT) and $75.5°C$ (M5). Since MMLV RT aggregates as it is unfolded, these temperatures should not be considered true melting temperatures, as they reflect kinetics of aggregation and solubility of the unfolded form, in addition to intrinsic conformational stability of MMLV RT.

Template-primer binding affinity

Steady-state kinetic parameters were then measured to determine whether increased thermal resistance in the presence of template-primer could be explained by higher binding affinity. K_m and V_{max} values were measured under conditions (5 min reaction time; 0.03 pmol RT) where incorporation is linear over the range of poly(rC): oligo(dG)$_{18}$ concentrations used (data not shown). As shown in Table 3, K_m (RNA–DNA) values for H$^-$ M5 and H$^-$ MMLV RT are 100 nM and 1000 nM, respectively, indicating that M5 binds template-primer with significantly higher (10-fold) affinity compared to wild-type MMLV RT. These results indicate that mutations isolated in our high-temperature activity screens provide increased resistance to heat-inactivation through tighter binding to primed RNA, which is not too surprising given that thermostability was monitored in the presence of poly(rC)–oligo(dG)$_{18}$ and assessed in terms of relative incorporation at elevated temperature.

Two-step RT–PCR

In Figure 5, we illustrate the broad temperature capability of M5 and the benefit of high cDNA synthesis temperatures with certain RT–PCR targets. In Figure 5A and B, M5 and H$^-$ M5 generate amplifiable cDNA from human skeletal muscle RNA at 37°C, 42°C and 55°C, while MMLV RT and H$^-$ MMLV RT can only generate cDNA at lower temperatures of 37°C and 42°C, respectively. In Figure 5C, amplification of a 0.6 kb fragment containing some secondary structure (28) is dramatically improved by carrying out cDNA synthesis with H$^-$ M5 at 55°C instead of 42°C. Additional studies have shown that M5 exhibits the same multi-temperature capability as H$^-$ M5, and supports the use of cDNA synthesis temperatures up to 60°C to relax intrinsic mRNA secondary structures

Table 3. Steady-state kinetic parameters

RT	K_m (RNA–DNA, nM)	V_{max} (pmol/h)
H$^-$ MMLV	1000	0.04
H$^-$ M5	100	0.07

All values have $<\pm30\%$ error and were obtained from three independent experiments.

and improve yields of certain messages (data not shown). In Figure 5D and E, we illustrate that sensitivity-of-detection is comparable for H$^-$ and H$^+$ versions of M5. In Figure 5D, identical yields of a 2 kb target (pol ε) are amplified from varying amounts (down to 1 ng) of reverse-transcribed total RNA. In panel E, a 0.6-kb fragment located at the 5′-end of 20 kb human nebulin transcript is amplified from as low as 10 ng reverse-transcribed RNA, indicating that both H$^+$ and H$^-$ M5 exhibit similar sensitivity when synthesizing very long cDNA targets.

DISCUSSION

In this report, we identified novel mutations residing in the fingers, thumb and connection domains of MMLV RT (Figure 1) that independently improve polymerase activity at elevated temperature (55°C; Tables 1 and 2). In combination, five of these mutations (E69K/E302R/W313F/L435G/N454K) broaden the thermal profile of MMLV RT (>85% activity between 40°C and 55°C; Figure 2), and allow synthesis of long (9 kb) cDNAs up to 55°C (Figure 3). Increased thermal resistance, however, is dependant on the presence of template-primer, as shown in both enzyme half-life and CD spectral analyses (Figure 4 and Supplementary Figure 3). When compared to wild-type in the presence of template-primer, the pentuple mutant (M5) exhibits a longer half-life at 55°C (30 min versus less than 5 min) and a higher transition temperature (75.5°C compared to 68°C). However in the absence of substrate, wild-type and mutant enzymes are equally sensitive to temperature. Both enzymes start to unfold at the same temperature (transition temperature, 34°C) and are completely inactive within 5 min of heating at 55°C.

A discrepancy was observed between the activity and spectral profiles of MMLV RT, obtained in the presence of template-primer. Polymerase activity of MMLV RT decreases dramatically between 50°C and 60°C (H$^-$, Figures 2 and 3; H$^+$, data not shown), even though spectral data indicate that the enzyme/template-primer complex remains intact up to 60°C (Figure 4; MMLV/T:P). These results suggest that between 50°C and 60°C, small structural changes occur in the active site that interfere with nucleotide binding and ternary complex formation, but do not affect template-primer binding or global MMLV RT structure.

As template-primer is required, increased thermal resistance of M5 can be attributed to changes in enzyme–substrate interactions, rather than to increased intrinsic stability. This conclusion is supported by steady-state kinetic measurements showing that the K_m [RNA–DNA] value of M5 is 10-fold lower than that of wild-type enzyme (Table 3). Our results are consistent with previous studies showing that interactions with template-primer protect both MMLV RT and AMV reverse transcriptases against thermal inactivation (5,19). In one study (5), the $t_{1/2}$ (50°C) of H$^-$ AMV RT was found to be 15 times higher than that of H$^-$ MMLV RT in the presence of template-primer (but similar in the absence of template-primer), which the authors attributed to greater binding affinity
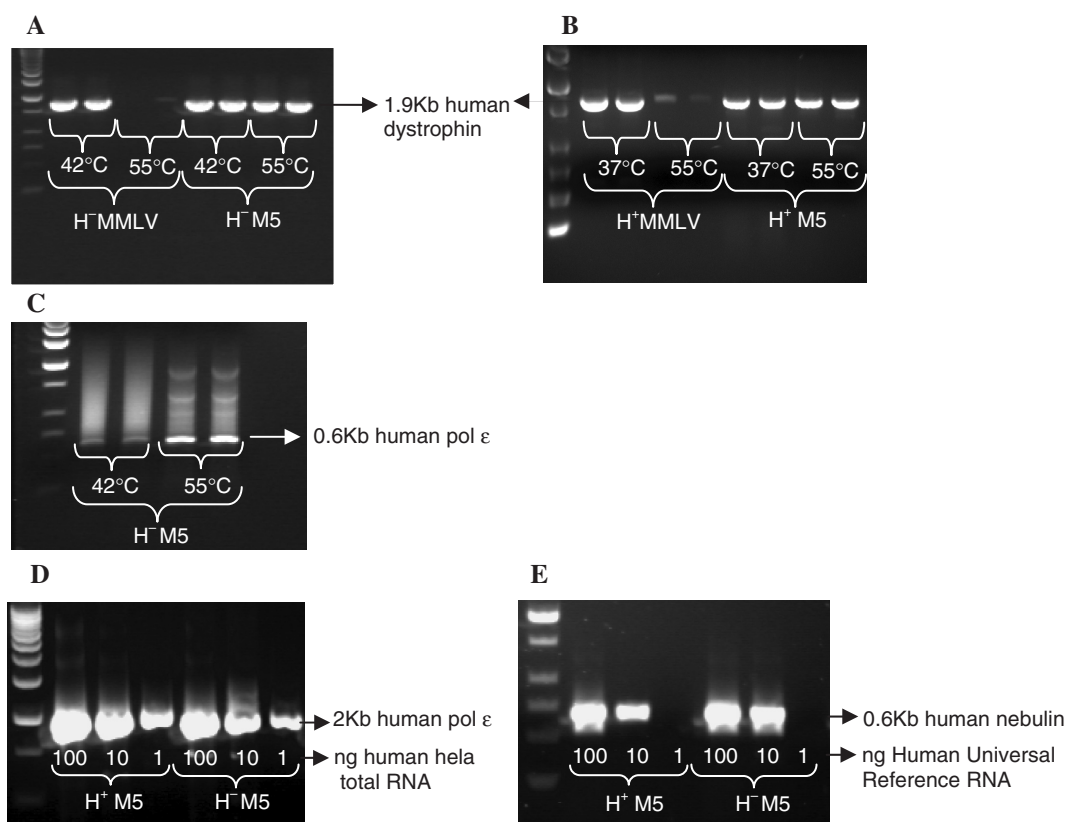
Figure 5. High-temperature cDNA synthesis and two-step RT–PCR. All reactions contained 384 ng RT and were incubated at indicated temperatures for 1 h, followed by 15 min at 70°C. Two microlitres of cDNA and 2.5 U of PicoMaxx were used in all amplifications. In (**A**) and (**B**), reactions also contained 500 ng human skeletal muscle total RNA. A 1.9 kb fragment of dystrophin was amplified using primers Dys8F and Dys2R. In (**C**), reactions contained 100 ng human hela RNA and 4 pmol Eps0.6-R. A 0.6 kb fragment of polymerase ε was amplified using primers Eps0.6-F and Eps0.6-R. In (**D**), indicated amounts of human hela total RNA were used in each reaction. RT incubation temperature was at 50°C. A 2 kb fragment of polymerase ε was amplified using primers Eps2-F and Eps2-R. In (**E**), indicated amounts of Human Universal Reference RNA were used in each reaction. RT incubation temperature was at 42°C. A 0.6-kb fragment from the 5′-end of 20-kb human nebulin transcript was amplified using primers NebF3 and NebR3.

for template-primer (e.g. K_D of H⁻ AMV RT is 50-fold lower than that of H⁻ MMLV RT).

As discussed above, mutations that increase thermal resistance of M5 through tighter binding to template-primer reside in three distal locations within the fingers (E69K), thumb (E302**R**/K, where bold is preferred mutation; W313F), and connection (L435**G**/M, N454**K**/R) domains. Three of the five mutations are located at positions that have been previously implicated in RT interactions with template-primer or in MMLV RT solubility.

The fingers domain of MMLV RT is thought to play a role in processive DNA synthesis by providing an intermediate binding site for template-primer between phosphonucleotidyl transfer reactions (26). In crystal structure of an N-terminal MMLV RT fragment (amino acids 24–278; fingers-palm domains) bound to template-primer, side chains S67 (O$^\gamma$) and E69 (O$^{\varepsilon 1}$) formed hydrogen bonds with the single-stranded template overhang (26). However, when compared to the more recent structure of full-length MMLV RT (24), the fingers domain binding site was not accessible for DNA binding due to packing constraints. Increased thermal resistance exhibited by MMLV RT E69K is consistent with the proposal that

E69 [which is conserved between MMLV and HIV RT; (24)] contributes to template-primer binding by providing a secondary binding site in the fingers domain. Moreover, saturation mutagenesis indicates that lysine is the only side chain substitution that provides increased thermal activity, suggesting that M5/template-primer interactions are stabilized through additional bond formation between the ε-amino group of K69 and negatively charged phosphate groups on the RNA template strand.

The thumb of MMLV RT (amino acids 276–361) is thought to play a role in substrate binding and processivity. Four of the mutations in MMLV RT that provide increased thermal activity reside in an alpha helical region that is considered analogous to α_H of HIV RT (24) and encompasses α_J (E302R/K; identified in this study) and the α_J-α_K interconnecting loop [T306R/K, F309N, W313F; identified here and (22)]. In a structure-based sequence alignment, MMLV RT E302 and T306 were identified as counterparts to HIV RT G262 and W266 (29), which reside in helix α_H facing the minor groove of the template-primer duplex (30) and interact with the primer strand sugar-phosphate backbone (31). Compared to wild-type, HIV RT G262A and W266A mutants exhibit lower

processivity (32) and reduced template-primer binding affinity [14- and 4-fold higher K_m (RNA–DNA) values, respectively; (30)], which is consistent with an assigned role in facilitating translocation of template-primer during polymerization (27).

Increased thermal resistance exhibited by MMLV RT E302R/K, T306R/K, F309N and W313F mutants is consistent with the proposal that E302 and T306 are analogous to HIV RT G262 and W266 [despite differences in the structures of HIV RT α_H and MMLV RT α_J, (24)], and that interactions with template-primer extend to F309 and W313. Failure to recover thermally resistant mutants by saturation mutagenesis of F303 and G305 (Table 2) suggests that these side chains face away from the surface (E302-T306-F309-W313) that interacts by analogy with HIV RT (29), with the minor groove of the template-primer duplex. Introducing positively charged (E302R/K, T306R/K) or polar (F309N) side chains may improve template-primer binding and thermal resistance by increasing the number of hydrogen bonds formed with minor groove base or phosphate backbone atoms. It is difficult to speculate as to how a relatively conservative hydrophobic mutation (W313F) enhances thermal activity. In a study of 129 protein–DNA complexes, tryptophan was rarely found in DNA backbone or base interactions; in contrast, phenylalanine was frequently involved in van der Waals contacts with the sugar–phosphate backbone or with DNA bases, positioned such that the plane of the ring faced DNA (no exposed bases) or was involved in ring-stacking interactions (exposed bases) (33).

The connection domain (amino acids 362–474) provides conformational flexibility between the RNA-/DNA-dependant polymerase and the C-terminal RNase H domain. The mutations that provide increased thermal activity in MMLV RT reside at the C-terminal end of β_{19} (L435G/M) and in a large unstructured loop (N454K/R) that connects to the RNase H domain. L435 is a solvent-exposed residue (24), residing in a stretch of hydrophobic amino acids (432–436: LVI<u>L</u>A) that were independently replaced by lysine in an effort to increase solubility (34). L435K mutants exhibited increased solubility in the absence of detergents and near wild-type ($\sim$78%) levels of polymerase activity (34), indicating that this position is 'mutatable' but not essential for function. In our screen for increased thermal resistance, we recovered L435M (random mutant library screen; Table 1) and L435G (saturation mutagenesis; 4/4 thermally active clones contained G; Table 2) mutants, but not L435K, suggesting that local perturbations in β_{19} (e.g. side chain loss, L435G; conservative replacement, L435M) can lead to changes in template-primer interactions. L435G/M mutations are not expected to confer thermostability through increased solubility or stability [as is the case for L435K; (34)], since M5 (contains L435G) shows the same heat-sensitivity and denaturation/aggregation behaviour as wild-type MMLV RT in the absence of template-primer (Figure 4 and Supplementary Figure 3). Residue N454, which is located distal to L435 in the connection domain, has not been previously implicated in template-primer interactions in MMLV RT.

In summary, we have engineered a pentuple mutant of MMLV RT (M5) that exhibits higher affinity for primed RNA and increased thermal activity in the presence template-primer. Compared to wild-type MMLV RT, M5 exhibits higher specific activity over a broad range of temperatures and improved synthesis of long cDNAs at elevated temperature ($>50^\circ$C). As a result, M5 can serve as a single reagent for all cDNA synthesis systems, whether priming at 25°C with random hexamers, or at higher temperatures (55°C –60°C) to enhance priming specificity or resolve secondary structures. A commercial preparation of MMLV RT, containing the five thermal resistance mutations identified here, is available for cDNA synthesis and RT–PCR applications (AffinityScript RT).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

REFERENCES

1. Molling,K., Bolognesi,D.P., Bauer,H., Busen,W., Plassmann,H.W. and Hausen,P. (1971) Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nat. New Biol.*, **234**, 240–243.
2. Berger,S.L., Wallace,D.M., Puskas,R.S. and Eschenfeldt,W.H. (1983) Reverse transcriptase and its associated ribonuclease H: interplay of two enzyme activities controls the yield of single-stranded complementary deoxyribonucleic acid. *Biochemistry*, **22**, 2365–2372.
3. Krug,M.S. and Berger,S.L. (1987) First-strand cDNA synthesis primed with oligo(dT). *Methods Enzymol.*, **152**, 316–325.
4. Kotewicz,M.L., Sampson,C.M., D'Alessio,J.M. and Gerard,G.F. (1988) Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.*, **16**, 265–277.
5. Gerard,G.F., Potter,R.J., Smith,M.D., Rosenthal,K., Dhariwal,G., Lee,J. and Chatterjee,D.K. (2002) The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.*, **30**, 3118–3129.
6. Harrison,G.P., Mayo,M.S., Hunter,E. and Lever,A.M. (1998) Pausing of reverse transcriptase on retroviral RNA templates is influenced by secondary structures both 5' and 3' of the catalytic site. *Nucleic Acids Res.*, **26**, 3433–3442.
7. DeStefano,J.J., Buiser,R.G., Mallaber,L.M., Myers,T.W., Bambara,R.A. and Fay,P.J. (1991) Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *J. Biol. Chem.*, **266**, 7423–7431.
8. Malboeuf, C.M., Isaacs, S.J., Tran, N.H. and Kim, B. (2001) Thermal effects on reverse transcription: improvement of accuracy and processivity in cDNA synthesis. *Biotechniques*, 30, 1074–1078, 1080, 1082.
9. Gerard, G.F., Collins, S. and Smith, M.D. (2002) Excess dNTPs minimize RNA hydrolysis during reverse transcription. *Biotechniques*, 33, 984, 986, 988.
10. Jones,M.D. and Foulkes,N.S. (1989) Reverse transcription of mRNA by Thermus aquaticus DNA polymerase. *Nucleic Acids Res.*, **17**, 8387–8388.
11. Shandilya,H., Griffiths,K., Flynn,E.K., Astatke,M., Shih,P.J., Lee,J.E., Gerard,G.F., Gibbs,M.D. and Bergquist,P.L. (2004)

Thermophilic bacterial DNA polymerases with reverse-transcriptase activity. *Extremophiles*, **8**, 243–251.

12. Markau, U., Ebenbichler, C., Achhammer, G. and Ankenbauer, W. (1998) US patent, 6,399,320.

13. Myers,T.W. and Gelfand,D.H. (1991) Reverse transcription and DNA amplification by a Thermus thermophilus DNA polymerase. *Biochemistry*, **30**, 7661–7666.

14. Cadwell,R.C. and Joyce,G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl.*, **2**, 28–33.

15. Ong,J.L., Loakes,D., Jaroslawski,S., Too,K. and Holliger,P. (2006) Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide. *J. Mol. Biol.*, **361**, 537–550.

16. Vichier-Guerre,S., Ferris,S., Auberger,N., Mahiddine,K. and Jestin,J.L. (2006) A population of thermostable reverse transcriptases evolved from Thermus aquaticus DNA polymerase I by phage display. *Angew. Chem. Int. Ed. Engl.*, **45**, 6133–6137.

17. Smith, E.S., Elfstrom, C.M., Gelfand, D.H., Higuchi, R.G., Myers, T.W., Schonbrunner, N.J. and Wang, A.M. (2002) High temperature reverse transcription using mutant DNA polymerases. US Patent Application 20020012970.

18. Sauter,K.B. and Marx,A. (2006) Evolving thermostable reverse transcriptase activity in a DNA polymerase scaffold. *Angew. Chem. Int. Ed. Engl.*, **45**, 7633–7635.

19. Yasukawa,K., Nemoto,D. and Inouye,K. (2008) Comparison of the thermal stabilities of reverse transcriptases from avian myeloblastosis virus and Moloney murine leukaemia virus. *J. Biochem.*, **143**, 261–268.

20. Stahlberg,A., Kubista,M. and Pfaffl,M. (2004) Comparison of reverse transcriptases in gene expression analysis. *Clin. Chem.*, **50**, 1678–1680.

21. Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. (1998) Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl Acad. Sci. USA*, **95**, 520–524.

22. Smith, M.D., Potter, R.J., Djariwal, G., Gerard, G.F. and Rosenthal, K. (2006) US patent 7,078,208.

23. Potter,J., Zheng,W. and Lee,J. (2003) Thermal stability and cDNA synthesis capability of SuperScript III reverse transcriptase. *Focus*, **25.1**, 19–24.

24. Das,D. and Georgiadis,M.M. (2004) The crystal structure of the monomeric reverse transcriptase from Moloney murine leukemia virus. *Structure*, **12**, 819–829.

25. Hogrefe, H.H., Cline, J., Youngblood, G.L. and Allen, R.M. (2002) Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit. *Biotechniques*, **33**, 1158–1160, 1162, 1164–1155.

26. Najmudin,S., Cote,M.L., Sun,D., Yohannan,S., Montano,S.P., Gu,J. and Georgiadis,M.M. (2000) Crystal structures of an N-terminal fragment from Moloney murine leukemia virus reverse transcriptase complexed with nucleic acid: functional implications for template-primer binding to the fingers domain. *J. Mol. Biol.*, **296**, 613–632.

27. Bebenek,K., Beard,W.A., Darden,T.A., Li,L., Prasad,R., Luton,B.A., Gorenstein,D.G., Wilson,S.H. and Kunkel,T.A. (1997) A minor groove binding track in reverse transcriptase. *Nat. Struct. Biol.*, **4**, 194–197.

28. Xu,R.H., Schuster,D.M., Lee,J.E., Smith,M., Potter,J., Dhariwal,G., Rosenthal,K., Nathan,M., Gerard,G.F. and Rashtchian,A. (2000) One-step analysis and quantification of RNA by RT-PCR: Using high-temperature reverse transcription. *Focus*, **22**, 3–6.

29. Shindyalov,I.N. and Bourne,P.E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.*, **11**, 739–747.

30. Beard,W.A., Stahl,S.J., Kim,H.R., Bebenek,K., Kumar,A., Strub,M.P., Becerra,S.P., Kunkel,T.A. and Wilson,S.H. (1994) Structure/function studies of human immunodeficiency virus type 1 reverse transcriptase. Alanine scanning mutagenesis of an alpha-helix in the thumb subdomain. *J. Biol. Chem.*, **269**, 28091–28097.

31. Jacobo-Molina,A., Ding,J., Nanni,R.G., Clark,A.D. Jr, Lu,X., Tantillo,C., Williams,R.L., Kamer,G., Ferris,A.L., Clark,P. *et al.* (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc. Natl Acad. Sci. USA*, **90**, 6320–6324.

32. Bebenek,K., Beard,W.A., Casas-Finet,J.R., Kim,H.R., Darden,T.A., Wilson,S.H. and Kunkel,T.A. (1995) Reduced frameshift fidelity and processivity of HIV-1 reverse transcriptase mutants containing alanine substitutions in helix H of the thumb subdomain. *J. Biol. Chem.*, **270**, 19516–19523.

33. Luscombe,N.M., Laskowski,R.A. and Thornton,J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.*, **29**, 2860–2874.

34. Das,D. and Georgiadis,M.M. (2001) A directed approach to improving the solubility of Moloney murine leukemia virus reverse transcriptase. *Protein Sci.*, **10**, 1936–1941.