Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 612

Escherichia coli LysU is a potential surrogate for human lysyl tRNA synthetase in interactions with the C-terminal domain of HIV-1 capsid protein

Nonlawat Boonyalai,[a,b] James R. Pullen,[a] Mohd Firdaus Abdul Wahab,[a,c] Michael Wright[a,d] and Andrew D. Miller*[a,d]

Human lysyl-tRNA synthetase (hLysRS) is known to interact directly with human immunodeficiency virus type-1 (HIV-1) GagPol polyproteins, and both hLysRS with $tRNA^{Lys3}$ are selectively packaged into emerging HIV-1 viral particles. This packaging process appears to be mediated by contact between the motif 1 helix h7 of hLysRS and the C-terminal dimerization domain of the HIV-1 capsid protein (CA) segment of Gag or GagPol. Given similarities between hLysRS and *Escherichia coli* (*E. coli*) heat shock protein LysU, we investigate if LysU might be an hLysRS surrogate for interactions with Gag or GagPol proteins. We report on a series of studies involving three CA *C*-domains: CA_{146} (intact domain), CA_{151} (truncated domain), and CA_{146}-M185A (M185A, CA dimer interface mutant). After confirming that LysU and CA_{146} are dimeric whilst CA_{151} and M185A remain monomeric, we use glutathione *S*-transferase (GST) pull-down assays to demonstrate the existence of specific interactions between LysU and all three CA-*C* domains. By means of ^{1}H-NMR titration experiments, we estimate K_d values of 50 μM for the interaction between LysU and CA_{146} or >500 μM for interactions between LysU and CA_{151} or LysU and M185A. The reason for these binding affinity differences may be that interactions between LysU and CA_{146} take place through dimer–dimer interactions resulting in a $\alpha_2\beta_2$ heterotetramer. LysU/CA-*C* protein interactions are weaker than those reported between hLysRS and the Gag, CA or CA_{146} proteins, and hLysRS/Gag binding interactions have also been suggested to involve only αβ heterodimer formation. Nevertheless, we propose that LysU could act as a surrogate for hLysRS with respect to Gag and GagPol polyprotein interactions although arguably not sufficiently for LysU to act as an inhibitor of the HIV-1 life cycle without further adaptation or mutation. Potentially, LysU and/or LysU mutants could represent a new class of anti-HIV-1 therapeutic agent.

Received 31st July 2012,
Accepted 16th November 2012

DOI: 10.1039/c2ob26499d

www.rsc.org/obc

Introduction

Escherichia coli (*E. coli*) lysyl-tRNA synthetase is a class II synthetase and is unusual in existing in two distinct synthetase isoforms: LysS and LysU.[1] These share a high degree of sequence identity (88%) and similar aminoacylation activities but are regulated differently. LysS is constitutively expressed under normal growth conditions and appears to be responsible for the tRNA charging activity, while LysU is the product of a normally silent gene that is induced to high-expression during heat-shock.[2,3] Current research indicates that LysU is a truly multifunctional protein with multiple catalytic activities. Other than tRNA charging, LysU is an unusually effective sequential catalyst of diadenosine 5′-5‴-P^1,P^4-tetraphosphate (Ap_4A) and diadenosine 5′-5‴-P^1,P^3-triphosphate (Ap_3A) formation from ATP.[4–6] Moreover a new glycerol kinase capability was very recently reported.[7] Otherwise, LysU is known to be phosphorylated and may be involved in the phosphorylation of heat shock proteins such as heat shock protein 70 (Hsp70, DnaK).[6] There appears a growing likelihood that other new catalytic functions can emerge as well (Chen *et al.*, unpublished data).

Current research into the human homologue of LysU, namely human lysyl-tRNA synthetase (hLysRS), also suggests that this enzyme has a tendency towards multifunctionality. Other than tRNA charging, the homologue hLysRS is known to be involved in the human immunodeficiency virus (HIV) life

[a]*Imperial College Genetic Therapies Centre, Department of Chemistry, Flowers Building, Armstrong Road, Imperial College London, London SW7 2AZ, UK. E-mail: a.miller07@btinternet.com*
[b]*Department of Biochemistry, Kasetsart University, Pahonyothin Road, Chatuchak, Bangkok 10900, Thailand*
[c]*Department of Biological Sciences, Faculty of Biosciences and Bioengineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia*
[d]*Institute of Pharmaceutical Science, King's College London, Franklin-Wilkins Building, Waterloo Campus, 150 Stamford Street, London, SE1 9NH, UK*

cycle by acting as a carrier for the $tRNA^{Lys}$ isoacceptor, known as $tRNA^{Lys3}$, that itself acts as a primer to initiate viral mRNA reverse transcription, viral amplification and infection post human immunodeficiency virus type 1 (HIV-1) viral particle entry into host white blood cells.[8–11] Protein hLysRS forms a hLysRS–$tRNA^{Lys3}$ carrier complex that then interacts with HIV-1 Gag polyprotein to enable packaging of the complex within HIV-1 viral particles.[12–14] Post HIV-1 viral particle entry into host white blood cells, $tRNA^{Lys3}$ must dissociate from the hLysRS carrier in order to prime copying of the viral genome from RNA to double-stranded complementary DNA (cDNA) by a viral enzyme reverse transcriptase (RT). Newly reverse-transcribed cDNA then translocates to the nucleus for chromosomal integration. Subsequent transcription of this integrated viral DNA results in at least two polyproteins: Gag and GagPol. In the final step of the viral life cycle, Gag, GagPol, and genomic RNA, plus fresh hLysRS–$tRNA^{Lys3}$ carrier complex and other host cell-derived factors assemble at plasma membranes for final maturation into new HIV-1 viral particles. Here, Gag is processed into component viral proteins known as the matrix (MA), capsid (CA) and nucleocapsid proteins (NC). CA is a 231 amino acid residue protein composed of two domains: CA-*N* (residues 1–145) an N-terminal "core" domain, and CA-*C* or CA_{146} (residues 146–231) a C-terminal "dimerization" domain. The structure of CA_{146} has been solved to a resolution of 1.7 Å, as has the truncated protein CA_{151} (residues 151–231).[15] The structures are similar but the CA_{146} readily forms a dimer characterized by a dissociation constant of 10 ± 3 μM which is nearly the same as the full-length protein CA (18 ± 1 μM), whilst CA_{151} appears to remain a monomer.[15,16] Each CA monomer contains a 3_{10} helix followed by an extended strand and four α-helices.[17] The major homology region forms a compact strand-turn helix motif.

The structural relationship between LysU and hLysRS is close in that both are considered to be class II lysyl tRNA synthetases (class II LysRSs) with a close structural relationship to the other known class II aspartyl- and asparaginyl-tRNA synthetases.[18] All have a modular arrangement of functional domains with a C-terminal aminoacylation catalytic domain and an N-terminal anticodon-binding site. In addition, class II LysRSs are some of the most conserved aminoacyl-tRNA synthetase enzymes known. Indeed the sequence identity between human hLysRS and the two *Escherichia coli* LysRS isozymes LysS and LysU is such that the determinants for novel interactions and functions are likely to involve only modest structural differences. Excellent X-ray crystal structure data exist for all three,[18–20] with hLysRS found as an α_4 homotetramer, in the crystalline state, that readily dissociates to α_2 homodimers in solution, and both LysS and LysU are found as robust α_2 homodimers (Fig. 1). Therefore, considering the structural and family relationship between LysU and hLysRS, we became curious about the likelihood of LysU acting as a surrogate for hLysRS in its interaction with human CA-*C* domains. This seemed all the more possible since hLysRS/Gag interactions involve the CA portion (helix h4 in CA_{146}) of the Gag polyprotein and amino acid residues 208–259 of the hLysRS motif 1 helix h7.[14,21,22] The level of sequence homology between the LysU and hLysRS motif 1 h7 helices is approximately 80% identical. Furthermore, if LysU were able to act as a surrogate of hLysRS in interactions with the Gag or CA proteins, then the protein might also act potentially as an inhibitor of the HIV-1 life cycle by interfering adversely in HIV-1 capsid assembly at the stage of hLysRS/Gag interactions. Here we describe our first efforts to study LysU and the CA-*C* domains in this context, leading to first stage proof of concept for this suggestion.

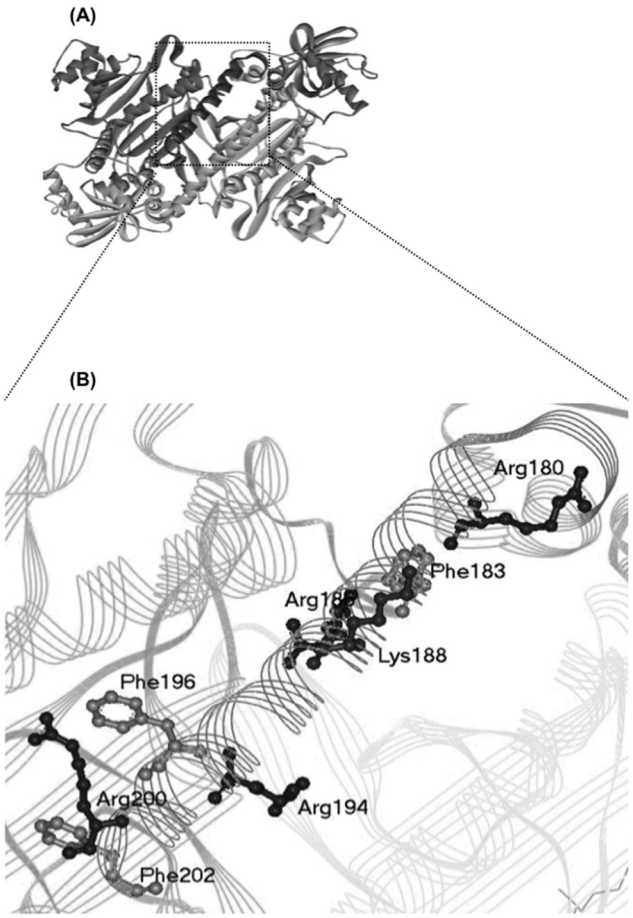

Fig. 1 Ribbon representation of the homodimeric LysU.[19] One complete monomer is shown in light grey and the other in grey, whilst helix 7 (h7) in the motif 1 of LysU is represented in dark grey. We show a close-up of h7. Some charged and aromatic amino acid residues are shown in ball-stick representation. These residues are putative interaction centres with the CA-*C* domain.

Results and discussion

Characterization of LysU

The DNA sequence encoding LysU was sub-cloned into pET-14b and pGEX-4T1 vectors to create plasmids expressing LysU protein with N-terminal His_6-tag and glutathione-*S*-transferase (GST)/LysU fusion protein (GST-LysU) respectively. Both proteins were expressed and purified in good yield (~90 mg L^{-1}), then characterized by circular dichroism (CD)

spectroscopy and found to possess similar unfolding characteristics and melting transition temperatures (T_m) to wild-type LysU (results not shown).[23–25] An enzymatic Ap_4A biosynthesis assay was then used to determine their catalytic integrity.[6,26–29] The measured k_{cat} values for Ap_4A formation catalysed by LysU and GST-LysU were found to be 128 ± 20 and 224 ± 24 min^{-1} respectively (under standard assay conditions). These values are equivalent to those seen for functional LysU WT.[6] Consequently, both His_6-tag and GST-fusion do not appear to unduly affect the overall structure or catalytic integrity of LysU.

Characterization of the CA-*C* domains

For our purposes, three different CA-*C* domains: CA_{146}, CA_{151} and mutant CA_{146}-M185A (M185A) were created for functional studies with LysU. Full-length CA_{146} is expected to be a dimer at high concentration, whilst truncated CA_{151} and dimer interface mutant M185A thought to be exclusively monomeric.[15] The reason that all three different CA-*C* domains were selected for functional studies was to allow the opportunity to explore how dimerisation differences and potentially other differences in CA-*C* domain physical properties might impact upon interactions with LysU. Accordingly for our purposes, all three were produced as N-terminal His_6-tag proteins and purified as described. The far-UV CD spectra demonstrated that each retained α-helical structure (typified by negative maxima at ΔA_{222} and ΔA_{208}) (Fig. 2A) in agreement with the crystallographic structures.[30] CD melting experiments were also performed (Fig. 2B) demonstrating that the melting profile of CA_{146} is in accord with previously published data with a T_m value of 64.7 ± 0.1 °C.[30] The profiles of CA_{151} and M185A differed somewhat and T_m values were found to be 64.4 ± 0.1 °C and 74.8 ± 0.2 °C respectively. M185A seems to be significantly more robust with respect to temperature suggesting an additional element of conformational stability. An extensive thermodynamic characterization of full-length CA-*C* domains has recently been reported using a variety of different techniques.[31] A comparison with our structural data suggests that the N-terminal His_6-tag was not significantly perturbing the structural characteristics of any of our three CA-*C* domains.

Oligomeric states

Size-exclusion chromatography and isothermal titration calorimetry (ITC) were used to confirm the oligomeric states of the three CA-*C* domains. Size-exclusion chromatography profiles on Superdex 200 medium each yielded a single peak with CA_{146} eluting in advance of CA_{151} and M185A that co-eluted (Fig. 3A). Average distribution constant values, K_{av}, were

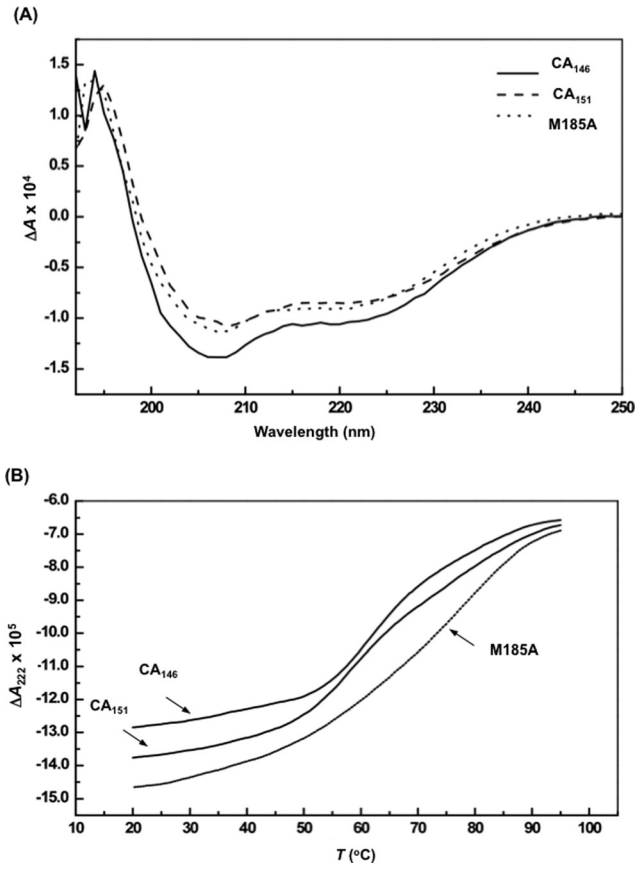

Fig. 2 (A) CD spectra of the three CA-*C* domains. CD spectral data were collected from 250 to 190 nm with protein samples at a concentration of 100 μM in 50 mM Tris-HCl buffer pH 8.0, 0.1 M NaCl. (B) Heat-induced unfolding of the three CA-*C* domains followed by CD at ΔA_{222}.

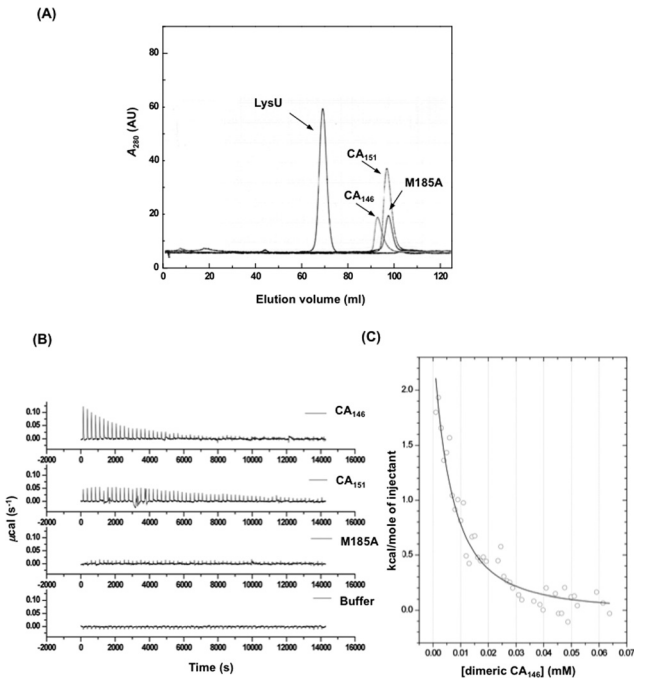

Fig. 3 (A) Gel filtration analysis of LysU, and the three CA-*C* domains. Chromatography was performed on Superdex 200 (Hi-load 16/60) FPLC column as described in Materials and Methods, with a corresponding protein monomer concentration of 115, 258 and 135 μM for CA_{146}, CA_{151}, and M185A respectively. (B) Raw ITC calorimetric data from the dilutions of CA_{146}, CA_{151} and M185A at 20 °C. Starting protein monomer concentrations were 280, 302 and 269 μM respectively, whilst the cell contained 10 mM PBS buffer. Each protein was administered to the cell in 2.5 × 25 × 10 μL aliquots (C) Fitted dimer dissociation curve for CA_{146} generated during step-wise dilution of CA_{146} into buffer.

measured and converted to approximate hydrodynamic molecular weight data. According to this exercise, CA_{146} possessed a molecular weight of approx. 25 kDa under the given elution conditions while CA_{151} and M185A each demonstrated molecular weights of approx. 14 kDa. These results are in agreement with previous experiments that suggested CA_{146} to be homodimeric in oligomeric state at concentrations >100 μM but exclusively monomeric in state <1 μM.[30] Further, our results also suggested that CA_{151} and M185A were eluted in an exclusively monomeric state under the given column conditions. Finally LysU was confirmed as a robust dimer (Fig. 3) with K_{av} values suggesting a molecular weight of around 114 kDa, very close to known molecular weight.[25]

When ITC dilution experiments were conducted, some interesting discrepancies were observed. When high concentration aliquots of dimeric CA_{146} were titrated progressively into phosphate-buffered saline (PBS) buffer, a standard endothermic data set was collected (Fig. 3B). This can be explained in terms of the thermodynamic effects of dimer dissociation induced by dilution. Initial endothermic heat uptake is high owing to extensive dissociation but as the CA_{146} concentration increases which each aliquot (towards a final dimer concentration of 65 μM), then the corresponding magnitude of heat uptake diminishes as the extent of dissociation also decreases. From these data, a dimer dissociation constant (K_{diss}) was extracted for CA_{146} of 13 ± 0.2 μM (Fig. 3), which agrees closely with the value reported previously.[15] By contrast, experiments performed with aliquots of mutant M185A resulted in an enthalpically neutral data set (Fig. 3) consistent with monomeric protein being diluted into buffer without state change. The raw ITC data set obtained from aliquots of CA_{151} appeared to exhibit elements of both the CA_{146} and M185A data sets (Fig. 3B). On the basis of this data we would suggest that CA_{151} has at least some propensity for dimer formation in solution at higher concentrations (above 250 μM), in contrast to the alternative suggestion made elsewhere,[15,16] but is otherwise monomeric in character.

Binding of LysU to CA-*C* domains

Initial binding interaction studies between LysU and the CA-*C* domains were carried out using a GST-pull down assay. GST-LysU protein was immobilized on batches of glutathione-sepharose beads, using GST-protein to block excess binding sites. Each of the CA-*C* domains was then incubated with the immobilized GST-LysU protein, and excess unbound protein was removed by washing. The sepharose beads were resuspended directly in sodium dodecyl sulphate (SDS) sample buffer, boiled and subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were analyzed by Western blotting with either anti-GST or anti-CA specific antibodies (Fig. 4). In the first instance, anti-GST probing was able to demonstrate that GST-LysU was interacting with the glutathione-sepharose beads in all experiments in which it was used (Fig. 4A). In the second instance, anti-CA probing was able to demonstrate that GST-LysU was binding with reasonable affinity to all three CA-*C* domains (Fig. 4B). Control experiments replacing GST-LysU with GST alone established that GST was able to bind to glutathione-sepharose beads but not to CA-*C* domains (data not shown). The data (Fig. 4B) also indicated that anti-CA antibodies exhibit some background non-specific interactions on the gel (Fig. 4B) with GST and with GST-LysU as well.

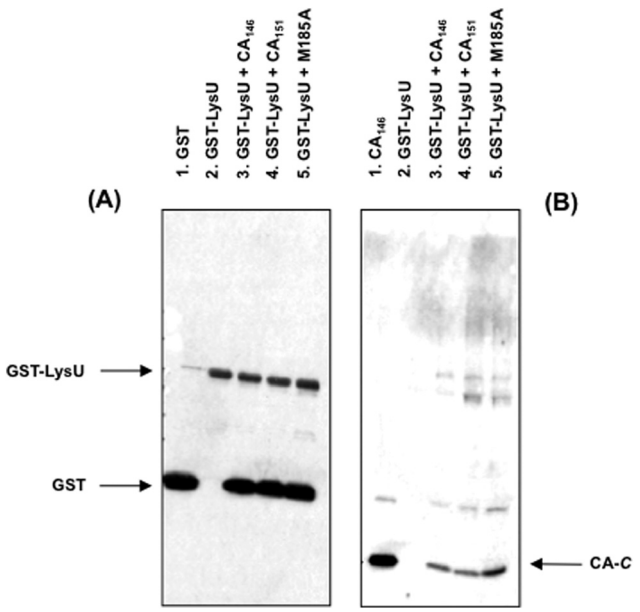

Fig. 4 *In vitro* interaction between CA_{146} and LysU. GST-LysU was adsorbed to glutathione-agarose beads, and after washing and blocking, the beads were incubated in binding buffer containing 6 μg of purified recombinant CA-*C* domains. After further washes, beads were resuspended directly in SDS sample buffer, boiled, and subject to SDS-PAGE. Western blots of the eluted material from the GST pull down assay were probed with anti-GST (panel A) and anti-CA (panel B) antibodies.

^{1}H-NMR titration binding experiments were used in an attempt to quantify these interactions. A surprisingly clear 1D-^{1}H-NMR (400 MHz) spectrum of each of the CA-*C* domains could be obtained post Fourier transformation of free induction decays due to their relatively small size (the dimer MWt is only about 25 kDa). By contrast, binding of CA-*C* to LysU was expected to result in considerable line-broadening of CA-*C* ^{1}H-resonance signals owing to the formation in solution of a much larger LysU/CA-*C* hetero-complex with a reduced spin-spin relaxation time constant (T_2) compared to the lone CA-*C* domains. On close analysis, the main difference observed between the spectra of LysU and the CA-*C* domains was a series of very structured methyl ^{1}H-resonance signals (around 0.4 to −0.4 ppm) found in CA-*C* domain but not LysU (Fig. 5). ^{1}H-NMR titration binding experiments were thus performed with fixed concentrations of CA-*C* domains titrated with increasing quantities of LysU. Averaged increases in line broadening of the methyl ^{1}H-resonance signals ($\Delta\nu$) were plotted as a function of LysU concentration to give binding isotherms from which dissociation constant (K_d) data could be derived, assuming a single site, single affinity binding model. The interaction between LysU and CA_{146} was found strongest with a K_d value estimated at 50 μM (Fig. 5), whereas others K_d values were estimated at >500 μM. These data do suggest that

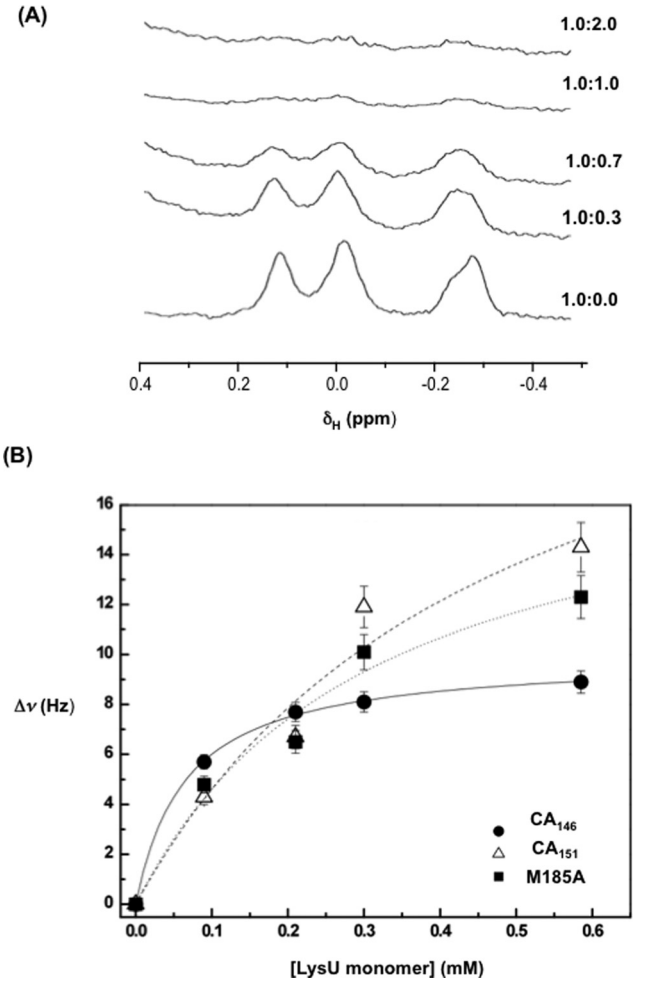

Fig. 5 ^{1}H-NMR titration binding experiments involving LysU and CA_{146} protein. (A) Titration of a fixed concentration of CA_{146} protein (0.3 mM [monomer] in 10 mM PBS buffer, pH 7.4, 10% D_2O) with an increasing concentration of LysU (as indicated by CA_{146} : LysU monomer ratios). ^{1}H-NMR spectrum at the upfield methyl region of CA_{146} and effects on line width and signal intensity with added LysU, added until these signals disappeared into baseline through line broadening caused by LysU/CA_{146} binding interactions. (B) Binding isotherms for interactions between LysU/CA_{146}, LysU/CA_{151}, and LysU/M185A were realized by plotting change in line broadening ($\Delta\nu$) as a function of LysU concentrations in order to extract approximate K_d values.

CA-*C* domains are able to bind to LysU, in agreement with the GST pull down assay. However, interactions are much weaker than those reported by fluorescence anisotropy for the interaction between hLysRS and Gag protein, or Δ1-219 hLysRS with the CA_{146} protein.[21,22]

These higher K_d values and the more robust natures of the LysU dimer and CA_{146} homodimers, suggested to us that LysU–CA_{146} interactions should logically give rise to an $\alpha_2\beta_2$ heterotetrameric complex in solution as a result of LysU dimer–CA_{146} dimer interactions. In the case of the other CA-*C* domains (CA_{151} and M185A), both would seem to be even more likely to interact as monomers with homodimeric LysU. This potential difference in association mechanism may explain why the binding affinities of CA_{151} and M185A for LysU were at least an order of magnitude weaker than the binding affinity of CA_{146} for LysU.

Implications for CA_{146} interactions

Our primary interest here was to test the hypothesis of LysU functional equivalence with hLysRS by studying for putative interactions between LysU and three CA-*C* domains derived from the CA capsid protein region of Gag. Of these three, only homodimeric CA_{146} appears to interact soundly with homodimeric LysU (Fig. 4 and 5) potentially forming a putative $\alpha_2\beta_2$ heterotetramer structure. Interactions between LysU and monomeric CA_{151} or even M185A are much weaker. However, even interactions between LysU–CA_{146} measured here are weaker than those reported interactions between hLysRS and Gag or CA protein, or between Δ1-219 hLysRS and CA_{146} (K_d values approx 300 to 700 nM respectively).[21,22] Also hLysRS has only been detected forming αβ heterodimeric complexes,[22] and the possibility that $\alpha_2\beta_2$ heterotetramers may also be formed has not been discussed elsewhere. Having said this hLysRS is believed to bind to more than one CA-*C* domain in order to assist Gag multimerisation during viral assembly,[32] with or without the influence of the CA-*N* domain.[33] Therefore, hLysRS and Gag may in fact associate together, at least initially, as a $\alpha_2\beta_2$ heterotetramer but this remains to be shown categorically.

In general, viral assembly is a particularly attractive candidate for anti-viral intervention because viral structures are formed by multiple, relatively weak, non-covalent interactions.[34,35] So far attempts to develop HIV-1 capsid assembly inhibitors have involved Gag protein derived peptides,[36] and a phage-display derived peptide (capsid assembly inhibitor, CAI) that binds reasonably strongly (K_d value of 15 μM) to helices h1 and h2 of CA-*C*.[37,38] Kovaleski *et al*. have reported alternatively on the design of a CA-*C* helix h4 mimic that binds hLysRS much more strongly (K_d value of 270 nM).[22] This peptide has yet to be investigated as an inhibitor of hLysRS/Gag protein interactions and further more as an HIV-1 capsid assembly inhibitor. Other new agents including cyclic peptide inhibitors of HIV-1 capsid-human lysyl-tRNA synthetase interaction have recently been disclosed.[39,40] Still, there is some way to go before putative HIV-1 capsid assembly inhibitors may be properly evaluated as anti-HIV therapeutic agents. Our data here suggest that LysU is a potential surrogate for hLysRS in interactions with CA-*C* domains. Therefore, LysU and/or appropriate mutants could be a potential new class of inhibitor of HIV-1 capsid assembly.

Overall, the multifunctional promiscuity of LysU suggests potentially important implications in biology and chemistry. The presentiment that one gene should code for one protein with one function appears increasingly outdated. It is well established that a given protein fold (such as the β-barrel) can catalyse a variety of different chemistries involving different enzymes (*e.g.*, isomerase, mutase, kinase, racemase, reductase, deaminase *etc.*). By contrast, LysU represents an example of an alternative scenario, where single enzymes constructed from

single folds have the capacity to promote a variety of significantly different chemistries depending upon surrounding conditions. Moreover, on the basis of data reported here, the LysU fold does not only carry out catalysis but has the capacity for viral capsid protein binding and putative inhibition of capsid assembly. Therefore, LysU can be said to have multiple physical functions to arraign with multiple catalytic capabilities. As the understanding of viral capsid assembly processes grows,[41] and as flexibilities inherent in active-site based protein biocatalysis become increasingly understood,[42] we would expect that the functional promiscuity of LysU should be encountered with many other proteins both known and yet to be discovered.

Experimental

Plasmid construction

LysU and GST-LysU expressing plasmids were constructed as follows. A 1.5 kb fragment corresponding to the entire *LysU* coding sequence was amplified using the plasmid pADH2 as a template.[6,23–25] Polymerase chain reaction (PCR) oligonucleotide primers were designed with *Nde*I and *Xho*I sites as well as initiation and termination codons immediately flanking the *LysU* coding sequence. Amplified DNA was sub-cloned into the *Nde*I-*Xho*I sites of expression vector pET14b(+) (Novagen, Merck Biosciences, Nottingham, UK) for LysU to be expressed as an N-terminal His_6-tagged protein. The following primers were used to construct the LysU expressing plasmid (forward primer, 5′-GGAATTCCATATGTCTGAACAAGAAAC-3′, and reverse primer, 5′-CAGCCCTCGAGTTATTTCTGTG-3′). The GST-LysU expressing plasmid was similarly constructed from pADH2 using oligonucleotide primers designed with *BamH*I and *Xho*I sites. Amplified DNA was then sub-cloned into the *BamH*I-*Xho*I sites of pGEX-4T1 (GE Healthcare Life Sciences, Bucks, UK). The following primers were used to construct the GST-LysU expressing plasmid (forward primer, 5′-GGAATTCGGATCC ATGTCTGAACAAGAA-3′, and reverse primer, 5′-ACAGCCCTCG AGTTATTTCTGTGGG-3′).

The DNA segments coding for CA_{146} and CA_{151} were amplified from pGag by PCR.[25] Again, *Nde*I and *Xho*I sites were introduced as above. Amplified fragments were then sub-cloned into the *Nde*I-*Xho*I sites of pET14b(+), creating an expression system for N-terminal His_6-tagged CA-*C* domains. The following primers were used to construct the CA_{146} expressing plasmid (forward primer, 5′-TAAGAACATATGAG CCCTACCAGCATTCTG-3′, and reverse primer, 5′-CTCTCTC GAGTTAAGCCAAAACTCTTGC-3′). The same reverse primer and the following forward primer (5′-CCCTACCCATATGCT GGACATAAGACAAGGA-3′) were used to construct the CA_{151}. Plasmid DNA expressing the M185A mutant was generated using the QuickChange™ site directed mutagenesis kit (Stratagene, La Jolla, CA) acting on the CA_{146} expressing plasmid as a template using the following primers to introduce the desired mutation (forward primer, 5′-GAGGTAAAAAATTGGGCGACAG AAACCTTGTTG-3′, reverse primer, 5′-CAACAAGGTTTCTGT CGCCCAATTTTTTACCTC-3′). Once all plasmid assembly and mutations were completed, the DNA sequences of all CA-*C* constructs were confirmed by DNA sequencing.

Protein expression and purification

All CA-*C* domains were expressed from the relevant plasmids in 1 L cultures of *E. coli* BL21 (DE3), pLysS (Novagen) grown at 37 °C until the A_{600} reached 0.6 to 0.7. Cultures were induced by addition of isopropyl D-1-thiogalactopyranose (IPTG) (1 mM) and further incubated at 37 °C for 3 h. Cells were harvested by centrifugation and cell pellets were kept at −20 °C until required. Pellets were washed and resuspended (30 mL) with extraction buffer (50 mM tris(hydroxymethyl)amino methane-hydrochloride [Tris-HCl] buffer, pH 8.0, 0.5 M NaCl, 5 mM imidazole, 1% Triton X-100, 5 μg mL^{-1} of DNase, 1 tablet of ethylenediamine tetraacetic acid [EDTA]-free protease inhibitor cocktail, 2 mM β-mercaptoethanol [β-ME], and 0.6 mM benzamidine) then disrupted with an ultrasound probe. After centrifugation, supernatants were applied to a 5 mL Ni^{2+}-HisTrap HP column (GE Healthcare, Piscataway, NJ, USA) attached to a fast protein liquid chromatography (FPLC) purification system to resolve proteins *via* their His_6-tag. In each case, the column was rinsed with washing buffer (50 mM Tris-HCl buffer, pH 8.0, 0.5 M NaCl, and 30 mM imidazole) to remove non-specifically bound proteins, before elution with an increasing gradient of imidazole to 0.5 M. Fractions were analyzed by SDS-PAGE and those with purity >97% were pooled and extensively dialysed against PBS buffer (10 mM PBS, 138 mM NaCl, 27 mM KCl, pH 7.4). Aliquots were flash-frozen and stored at −70 °C. The CA-*C* monomer concentration was determined by ultra-violet (UV) absorbance using an extinction coefficient of ε_{280} = 8530 M^{-1} cm^{-1} for CA_{146} and other CA-*C* monomer proteins.[15]

His_6-LysU proteins were also expressed from relevant plasmids in *E. coli* BL21 (DE3), pLysS, whilst GST-LysU and GST were expressed from *E. coli* BL21 (Novagen). Cells containing pET14b-LysU or pGEX4T1-LysU were induced with 1 mM or 0.1 mM IPTG at 37 °C for 3 h, then harvested, washed and resuspended (30 mL) in PBS buffer supplemented with EDTA-free protease inhibitor cocktail, 100 mM $MgCl_2$, 5 μg mL^{-1} of DNase, 1% Triton X-100, 2 mM β-ME, 0.6 mM benzamidine. The cell suspension was lysed and centrifuged as before. His_6 proteins were purified as described above, GST proteins using a 5 mL GSTRap column (GE Healthcare) equilibrated with 4× column volume of PBS buffer before application of the lysate. The column was then washed with 10 volumes of PBS and eluted with 5 volumes of 50 mM Tris-HCl, pH 8.0, 10 mM reduced glutathione. Analysis by SDS-PAGE indicated the isolation of highly pure GST-LysU protein. The LysU (monomer) concentration was determined by UV spectrophotometer using an extinction coefficient of ε_{280} = 30 580 M^{-1} cm^{-1}.[25] The bicinchoninic acid (BCA) assay (Pierce Biotechnology, Rockford, IL, USA) was used to determine the concentration of GST-LysU.

CD studies

These were performed on a Jasco J-715 spectropolarimeter using a quartz cuvette with a path length of 0.1 cm. The recorded spectra were the average of three scan between 190 and 250 nm obtained at a rate of 20 nm min^{-1}, a response time of 4 s and a bandwidth of 1 nm. Protein solutions were in 50 mM Tris-HCl buffer, pH 8.0 and the temperature was kept constant at 25 °C. Temperature-induced unfolding was followed by observing the change in ΔA_{222} ($\Delta\Delta A_{222}$) of solutions of protein (20 μM) in the same buffer on heating using a scan rate of 60 °C h^{-1}.

Analytical size-exclusion chromatography

Size-exclusion chromatography was performed at room temperature on a Superdex 200 Hi-load 16/60 FPLC column (GE Healthcare) equilibrated in 50 mM Tris-HCl, 0.1 M NaCl buffer, pH 8.0 at a flow rate of 1 mL min^{-1}. The column was calibrated using a low molecular weight (LMW) gel filtration kit (GE Healthcare) as molecular mass standards. Protein elution was monitored by A_{280} absorbance. Average distribution constant values were determined using expression (1):

$$K_{av} = \frac{(V_e - V_o)}{(V_t - V_o)} \quad (1)$$

where V_e, V_o and V_t correspond to the elution volume of the solute, the void volume and the total volume of the bed, respectively. These values were converted into hydrodynamic protein molecular weight values.

ITC studies

Dilution ITC was employed to analyze the monomer–dimer equilibrium of our CA-*C* domains. These studies used a Microcal VP-ITC calorimeter following standard instrumental procedure at 20 °C with a 250 μL injection syringe and 400 rpm stirring. In a typical experiment, small aliquots of concentrated CA-*C* domains in 10 mM PBS buffer, pH 7.4 were slowly injected into the calorimeter reaction vessel (1.4 mL volume) containing an identical buffer. The thermal reference cell was filled with water. Titrations typically consisted of a preliminary 1 μL injection followed by 2.5 × 25 × 10 μL injections with duration of 20 s each and 3 min spacing. Integrated heat pulse data (after correction for blank controls) were analyzed by nonlinear regression in terms of a simple monomer–dimer equilibrium model to give dimer dissociation constant values (K_{diss}).

In vitro binding assay

Glutathione-sepharose beads (Life Technologies, Paisley, UK) (30 μL) were washed with PBS plus added Tween 20 and ethylenediamine tetraacetic acid (PBST-EDTA) (10 mM PBS, pH 7.4, 0.1% Tween 20, 0.1 mM EDTA) before incubation with purified GST-LysU fusion protein (about 16 μg) at 4 °C for 1 h. After 3 additional washes to remove unbound material, the beads were resuspended in 30 μL PBST-EDTA and incubated with GST-control protein at 4 °C for another hour to block any unoccupied sites. The mixture was then washed three times and resuspended in an equal volume PBST-EDTA. Purified recombinant CA-*C* domains (6 μg) were then incubated with the washed beads at 4 °C for 1 h. The beads were washed a further 3 times with PBST-EDTA-200N (10 mM PBS, pH 7.4, 0.1% Tween 20, 0.1 mM EDTA, 200 mM NaCl) before suspension in SDS sample buffer. The samples were then heated at 95 °C and separated by SDS-PAGE.

Immunodetection of LysU and CA-*C* domains

To detect GST-LysU and CA-*C* domains by Western blot analysis, the SDS-PAGE bands were transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore, MA, USA) and non-specific sites blocked by soaking in tris-buffered saline with added tween 20 (TBS-T) buffer (20 mM Tris-HCl, 137 mM NaCl, 0.1% Tween 20, pH 7.5) containing 5% skimmed milk, at 4 °C overnight on an orbital shaker. After blocking the membrane was washed three times for 5 min with TBS-T at ambient temperature and probed with either 1 : 30 000 of a rabbit polyclonal antibody to HIV-1 Capsid protein (Advanced Biotechnologies Inc., Columbia, MA, USA) or 1 : 20 000 of an anti-GST antibody (Novagen), diluted in 1% milk TBS-T buffer for 1 hour at ambient temperature. Thereafter the membrane was washed 3 times and then incubated with 1 : 3000 of a rabbit IgG antibody horse radish peroxidase (HRP) conjugate, anti-R-HRP (Abcam, Cambridge, UK) for 1 h at ambient temperature. After washing a further three times with TBS-T buffer, immunodetection patterns were analysed by enhanced chemiluminescence (ECL) Western blotting reagents (GE Healthcare) and revealed by exposure of Hyperfilm ECL autoradiography film as described in the manufacturer's protocol.

^{1}H-NMR titration binding experiment

These experiments were performed at 25 °C (unless otherwise stated) on a Bruker Avance 400 MHz NMR spectrometer equipped with dual resonance probe with z-gradient. Titration of LysU and its mutants into samples of CA-*C* domains [0.3 mM, monomer] was carried out using ratios of 1 : 0, 1 : 0.3, 1 : 0.7, 1 : 1 and 1 : 2, in 10 mM PBS pH 7.4, 10% D_2O, with water suppression achieved with excitation sculpting. ^{1}H NMR spectra were taken with 128 scans and $\Delta\nu$ line broadening values were calculated from the half maximum height of the appropriate resonance. The dissociation constant (K_d) was calculated by fitting the data with the equation:

$$\Delta\nu = \frac{\Delta\nu_o[B]}{[B] + K_d} \quad (2)$$

where $\Delta\nu$ and $\Delta\nu_o$ are line broadening values at observation and maximum respectively, and $[B]$ is the concentration of LysU monomer.

Ion-exchange HPLC assays

The biosynthesis of Ap_4A was followed by ion-exchange HPLC as previously described.[25–27] The reaction mixture contained 5–10 μM of protein catalyst, 2 mM of lysine, 75 μg of pyrophosphatase, 10 mM ATP, 160 μM $ZnCl_2$ and 10 mM $MgCl_2$ in

25 mM Tris-HCl, pH 8.0. Fractions were identified by their retention time compared to the following standards: AMP, ADP, ATP, and Ap_4A.

Conclusion

In conclusion, we propose that LysU and/or LysU mutants could be a potential new class of inhibitor of hLysRS/Gag protein interactions and inhibitor of HIV-1 capsid assembly. Efforts will now be made going forward to discover modified LysU derivatives or LysU mutants with improved affinities for CA-*C* domains and arguably improved potential as anti-HIV-1 therapeutic agents. The potential advantages of using LysU and or LysU mutants as anti-HIV-1 therapeutic agents are multifold. LysU is typically simple to over-express and purify from bacterial cell culture, using our system reported here. The protein is also facile to mutate.[7] Furthermore, therapeutic screening of LysU mutants in HIV-1 capsid assembly assays should also be routine to implement. Accordingly, the combination of these three factors together suggest that therapeutic LysU mutants once identified should have real potential as biopharmaceutical protein agents (or as gene therapy agents) that could also be adapted with ease to retain therapeutic potency in the face of the usual tendency of all viruses to undergo viral escape through selective mutation when faced with selection pressures from chronic exposure to anti-viral therapeutic agents. Therefore, LysU mutants might actually have the potential to open the door to real, personalized anti-viral therapeutics.

Abbreviations

hLysRS	human-lysyl-tRNA synthetase
HIV-1	human immunodeficiency virus type-1
MA	matrix of HIV-1
NC	nucleocapsid protein of HIV-1
CA	capsid protein of HIV-1
CA-*C*	C-terminal domain of CA
CA-*N*	N-terminal domain of CA
CAI	capsid assembly inhibitor
LysU	*E. coli* lysyl-tRNA synthetase (heat inducible isozyme)
LysS	*E. coli* lysyl-tRNA synthetase (constitutively expressed isozyme)
GST	glutathione-*S*-transferase
Ap_4A	diadenosine 5′,5‴-P^1,P^4-tetraphosphate
Ap_3A	diadenosine 5′,5‴-P^1,P^3-triphosphate
hsp70	heat shock protein 70
RT	reverse transcriptase
CD	circular dichroism
T_m	melting transition temperature
ITC	isothermal titration calorimetry
PBS	phosphate-buffered saline
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
IPTG	isopropyl D-1-thiogalactopyranose
Tris-HCl	tris(hydroxymethyl)-aminomethane-hydrochloride
EDTA	ethylenediamine tetraacetic acid
β-ME	β-mercaptoethanol
FPLC	fast protein liquid chromatography
UV	ultraviolet
BCA assay	bicinchoninic acid assay
LMW	low molecular weight
PBST-EDTA	PBS with included Tween-20 and EDTA
PVDF	polyvinylidene fluoride
TBS-T	tris-buffered saline with added Tween 20
HRP	horseradish peroxidase
ECL	enhanced chemiluminescence

Acknowledgements

We thank Dr Steve Pettit at University of North Carolina-Chapel Hill for providing pGag plasmid as a gift. We thank Mitsubishi Chemical Corporation, IC-Vec Ltd and the EPSRC for supporting the Imperial College Genetic Therapies Centre.

References

1 F. Leveque, P. Plateau, P. Dessen and S. Blanquet, *Nucleic Acids Res.*, 1990, **18**, 305.
2 R. L. Clark and F. C. Neidhardt, *J. Bacteriol.*, 1990, **172**, 3237.
3 B. Rudolph, K. M. Gebendorfer, J. Buchner and J. Winter, *J. Biol. Chem.*, 2010, **285**, 19029.
4 Y. Nakamura and K. Ito, *Mol. Microbiol.*, 1993, **10**, 225.
5 A. Brevet, J. Chen, F. Leveque, S. Blanquet and P. Plateau, *J. Biol. Chem.*, 1995, **270**, 14439.
6 M. Wright, N. Boonyalai, J. A. Tanner, A. D. Hindley and A. D. Miller, *FEBS J.*, 2006, **273**, 3534.
7 X. Chen, N. Boonyalai, C. Lau, S. Thipayang, Y. Xu, M. Wright and A. D. Miller, *FEBS J.*, 2012, DOI: 10.1111/febs.12053.
8 L. Kleiman and S. Cen, *Int. J. Biochem. Cell Biol.*, 2004, **36**, 1776.
9 L. Kleiman, R. Halwani and H. Javanbakht, *Curr. HIV Res.*, 2004, **2**, 163.
10 J. Mak and L. Kleiman, *J. Virol.*, 1997, **71**, 8087.
11 J. Gabor, S. Cen, H. Javanbakht, M. Niu and L. Kleiman, *J. Virol.*, 2002, **76**, 9096.
12 S. Cen, A. Khorchid, H. Javanbakht, J. Gabor, T. Stello, K. Shiba, K. Musier-Forsyth and L. Kleiman, *J. Virol.*, 2001, **75**, 5043.
13 S. Cen, H. Javanbakht, S. Kim, K. Shiba, R. Craven, A. Rein, K. Ewalt, P. Schimmel, K. Musier-Forsyth and L. Kleiman, *J. Virol.*, 2002, **76**, 13111.

14 H. Javanbakht, R. Halwani, S. Cen, J. Saadatmand, K. Musier-Forsyth, H. Gottlinger and L. Kleiman, *J. Biol. Chem.*, 2003, **278**, 27644.
15 T. R. Gamble, S. Yoo, F. F. Vajdos, U. K. Von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist and C. P. Hill, *Science*, 1997, **278**, 849.
16 S. Yoo, D. G. Myszka, C.-Y. Yeh, M. McMurray, C. P. Hill and W. I. Sundquist, *J. Mol. Biol.*, 1997, **269**, 780.
17 D. K. Worthylake, H. Wang, S. Yoo, W. I. Sundquist and C. P. Hill, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 85.
18 M. Guo, M. Ignatov, K. Musier-Forysth, P. Schimmel and X.-L. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 2331.
19 S. Onesti, A. D. Miller and P. Brick, *Structure*, 1995, **3**, 163.
20 G. Desogus, F. Todone, P. Brick and S. Onesti, *Biochemistry*, 2000, **39**, 8418.
21 B. J. Kovaleski, R. Kennedy, M. K. Hong, S. A. Datta, L. Kleiman, A. Rein and K. Musier-Forsyth, *J. Biol. Chem.*, 2006, **281**, 19449.
22 B. J. Kovaleski, R. Kennedy, A. Khorchid, L. Kleiman, H. Matsuo and K. Musier-Forsyth, *J. Biol. Chem.*, 2007, **282**, 32274.
23 M.-E. Theoclitou, E. P. L. Wittung, A. D. Hindley, T. S. H. El-Thaher and A. D. Miller, *J. Chem. Soc., Perkin Trans. 1*, 1996, 2009.
24 M. Wright and A. D. Miller, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 943.
25 S. Hughes, J. A. Tanner, A. D. Hindley, A. D. Miller and I. R. Gould, *BMC Struct. Biol.*, 2003, **3**, 5.
26 S. Blanquet, P. Plateau and A. Brevet, *Mol. Cell. Biochem.*, 1983, **52**, 3.
27 J. Charlier and R. Sanchez, *Biochem. J.*, 1987, **248**, 43.
28 M. Wright, J. A. Tanner and A. D. Miller, *Anal. Biochem.*, 2003, **316**, 135.
29 M. Wright and A. D. Miller, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2813.
30 M. G. Mateu, *J. Mol. Biol.*, 2002, **318**, 519.
31 M. C. Lidón-moya, F. N. Barrera, M. Bueno, R. Pérez-jiménez, J. Sancho, M. G. Mateu and J. L. Neira, *Protein Sci.*, 2005, **14**, 2387.
32 R. Swanstrom and J. W. Wills, in *Retroviruses*, ed. J. Coffin, S. Hughes and H. Varmus, Cold Spring Harbor Laboratory Press, New York, 1997, pp. 33–47.
33 J. Lanman, J. Sexton, M. Sakalian and P. E. Pervelige Jr., *J. Virol.*, 2002, **76**, 6900.
34 Y. Morikawa, *Curr. HIV Res.*, 2003, **1**, 1.
35 B. K. Ganser-Pornillos, U. K. von Schwedler, K. M. Stray, C. Aiken and W. I. Sundquist, *J. Virol.*, 2004, **78**, 2545.
36 M. T. Garzón, M. C. Lidon-Moya, F. N. Barrera, A. Prieto, J. Gomez, M. G. Mateu and J. L. Neira, *Protein Sci.*, 2004, **13**, 1512.
37 J. Sticht, M. Humbert, S. Findlow, J. Bodem, B. Müller, U. Dietrich, J. Werner and H.-G. Kräusslich, *Nat. Struct. Mol. Biol.*, 2005, **12**, 671.
38 F. Ternois, J. Sticht, S. Duquerroy, H.-G. Kräusslich and F. A. Rey, *Nat. Struct. Mol. Biol.*, 2005, **12**, 678.
39 L. Kleiman, C. P. Jones and K. Musier-Forsyth, *FEBS Lett.*, 2010, **584**, 359.
40 V. Dewan, T. Liu, K. M. Chen, Z. Qian, Y. Xiao, L. Kleiman, K. V. Mahasenan, C. Li, H. Matsuo, D. Pei and K. Musier-Forsyth, *ACS Chem. Biol.*, 2012, **7**, 761.
41 D. H. Chen, M. L. Baker, C. F. Hryc, F. DiMaio, J. Jakana, W. Wu, M. Dougherty, C. Haase-Pettingell, M. F. Schmid, W. Jiang, D. Baker, J. A. King and W. Chiu, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 1355.
42 L. Wang, E. A. Althoff, J. Bolduc, L. Jiang, J. Moody, J. K. Lassila, L. Giger, D. Hilvert, B. Stoddard and D. Baker, *J. Mol. Biol.*, 2012, **415**, 615.

