

Cite this: DOI: 10.1039/c0cs00092b

www.rsc.org/csr

TUTORIAL REVIEW

Targeting the protein–protein interactions of the HIV lifecycle

Ali Tavassoli

Received 27th August 2010

DOI: 10.1039/c0cs00092b

The human immunodeficiency virus (HIV), the causative agent of acquired immunodeficiency syndrome (AIDS), relies heavily on protein–protein interactions in almost every step of its lifecycle. Targeting these interactions, especially those between virus and host proteins, is increasingly viewed as an ideal avenue for the design and development of new therapeutics. In this *tutorial review*, we outline the lifecycle of HIV and describe some of the protein–protein interactions that control and regulate each step of this process, also detailing efforts to develop therapies that target these interactions.

Introduction

There are multiple compounds currently on the market that target HIV, with the majority down-regulating the activity of a viral enzyme associated with the ailment. As a result, these therapies are very susceptible to emergence of drug resistance strains, of particular concern with HIV, whose genome-replication process has a high error rate that encourages mutations. This is reflected in drug-resistance rates of up to 19% to some HIV inhibitors,¹ and a transmitted drug resistance of over 9%.² The World Health Organisation

estimates 33 million people to be infected by HIV worldwide, resulting in 2 million deaths a year; given the number of

University of Southampton, Southampton, SO17 1BJ, UK.
E-mail: a.tavassoli@soton.ac.uk;
Web: <http://www.personal.soton.ac.uk/ali1>



Ali Tavassoli

Ali is a lecturer in chemical biology at The University of Southampton. He is the recipient of the European Association for Chemical and Molecular Sciences' silver medal for European Young Chemist, and a Career Establishment Award from Cancer Research UK. He has a multidisciplinary background in synthetic chemistry and molecular biology. His lab's research interests are at the interface of chemistry and biology; developing biological

and molecular tools, high throughput screening methodologies and uncovering compounds that tackle diseases through inhibition of specific protein–protein interactions. His research group are also active in the field of synthetic biology. For more information see the Tavassoli group website at www.personal.soton.ac.uk/ali1.

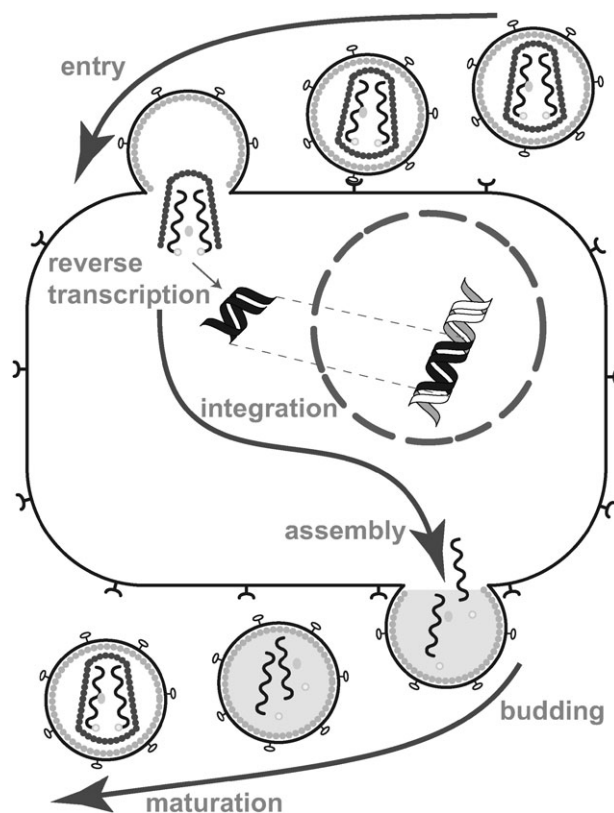


Fig. 1 The HIV lifecycle. HIV entry involves binding and internalisation of the viral capsid into the host, where the HIV genome is reverse transcribed from RNA to DNA and integrated into the host's genome. The proteins necessary for the production of viral progeny are produced by the host (through transcription and translation of the integrated viral genome), assembled and released. The immature particles undergo maturation and go on to infect other cells in the host.

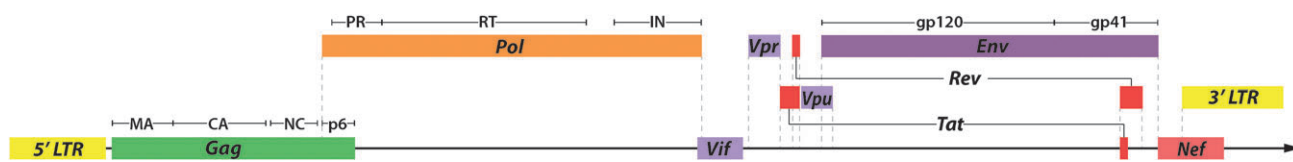


Fig. 2 The HIV genome. A single RNA strand codes for the 15 proteins of the HIV lifecycle, utilising all three reading frames. There is considerable overlap in the genetic information, which allows the controlled production of specific proteins (through modulation of mRNA splicing), as they are needed during the lifecycle. The Pol proteins are produced from a Gag-Pol mRNA by a frame-shift during translation.

infected individuals harbouring drug-resistant strains, there is now a pressing need for the development of novel therapies that enable new approaches to targeting HIV.

As with other viruses, HIV heavily relies on usurping the host's cellular machinery at almost every stage of its lifecycle (Fig. 1). These processes typically occur through the interaction of viral and host proteins, or homodimeric/multimeric viral protein interactions. The reliance on protein-protein interactions allows the possibility of developing novel therapeutic agents that inhibit HIV by targeting these interactions. There are however, major obstacles to this approach; the interacting surfaces involved are typically large and featureless, making logical design of inhibitors extremely challenging.³ Even with crystal structures of the interacting protein pair at hand, uncovering the region to target with a small molecule is not trivial. The alternative to logical design is high-throughput screening, which enables all potential binding sites on the target proteins to be assessed and assayed, with the most potent being identified by the system (rather than predetermined).⁴ This approach is increasingly used in drug discovery, with many modern drug candidates initially uncovered by screening of molecular libraries, with the lead compounds being optimised through conventional medicinal chemistry. Both approaches have been successfully applied to development of HIV therapies currently in the market.

Here we detail the steps in the HIV lifecycle that have been (or are currently being) targeted for the development of novel therapeutics that function by inhibiting specific protein-protein interactions. As the majority of current HIV therapies target enzymatic activity, there is much potential for inhibitors with this alternative mode of action.

The virus

HIV is part of the *Lentivirus* genus and a member of the retrovirus family. The genetic information of the virus is therefore stored on two copies of positive-sense RNA strands. The HIV genome (Fig. 2) contains nine reading frames, three of which (Gag, Pol and Env) encode polypeptides that are further proteolyzed to give a total of 15 proteins that are required for the lifecycle of HIV. The four Gag (matrix, capsid, nucleocapsid, p6) and two Env (gp120, gp41) proteins form structural components that make up the virus core and outer membrane. Pol encodes three enzymes (reverse transcriptase, protease, integrase) that provide essential functions in the virus lifecycle (not available *via* the host machinery), which are also encapsulated within the virus particle. The six other HIV proteins (Vif, Vpr, Nef, Tat, Rev, Vpu) are accessory proteins that play key roles at various

points during the virus lifecycle. The virus can therefore be considered as a molecular entity that consists of 15 proteins and RNA.⁵

The virus itself (Fig. 3) is roughly spherical, with a diameter of 120 nm. The two RNA strands at the core of the virus particle are tightly bound to nucleocapsid proteins (NC, also known as p7) and are surrounded by a conical capsid, composed of the p24 capsid protein (CA). Within the capsid, the enzymes and proteins needed by the virus (and not available in the host) are present; these include reverse transcriptase, integrase, protease, Vif, Vpr, Nef and p6. The capsid is surrounded by the matrix protein (MA), which ensures the integrity of the virus particle. The viral envelope surrounding the matrix is made up of phospholipids taken from the host's membrane during budding, and is embedded with gp41 proteins. The viral envelope is completed by gp120, which non-covalently associates to gp41. The envelope proteins arrange into homotrimeric "spikes" that are essential for entry into the host cell.

Viral entry

Infection begins with virus entry into the host. This is an intricate, complex, multistep process that involves the viral envelope proteins gp120 and gp41, and corresponding receptors on the host cell. The envelope proteins are synthesized as a single polypeptide (gp160) that is cleaved during its transit to the cellular membrane. The gp41 protein is anchored to, and spans the viral membrane, with its extracellular domain binding (non-covalently) to gp120. These viral proteins

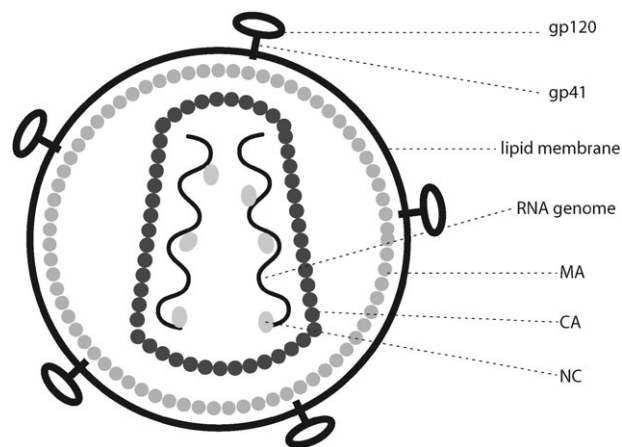


Fig. 3 The HIV virus. The virus is roughly spherical with the RNA genome enclosed within a conical capsid that also contains the enzymes and proteins needed for infection. These include reverse transcriptase, integrase, protease, as well as p6, Vif, Vpr and Nef.

associate into homotrimers⁶ that mediate binding and entry to target cells. The primary receptor for HIV is the glycoprotein CD4, which is expressed on the surface of T helper cells, regulatory T cells, monocytes, macrophages and dendritic cells. CD4 normally functions as a co-receptor (with the T cell receptor) to activate T-cells following interaction with antigen-presenting cells. The binding of gp120 (through a highly conserved, unglycosylated region)⁷ to CD4 initiates viral entry, and causes gp120 to undergo a dramatic change in conformation (whilst still maintaining its association with gp41)⁸ to expose a second (co-receptor) binding site. The primary co-receptors of HIV are CCR5 and CXCR4,⁹ both chemokine receptors and members of the G protein-coupled receptor family. Different strains of HIV can display selectivity for a particular co-receptor; those using CCR5 are termed R5-tropic (or M-tropic as CCR5 is found mostly in macrophages), those using CXCR4 are termed X4-tropic (or T-tropic, as CXCR4 receptors are typically found in T lymphocytes). Those that bind either CCR5 or CXCR4 are termed mixed-tropic (or X4R5). Following binding to the co-receptor, gp120 (now anchored to the host by two protein–protein interactions) undergoes a further conformational shift that brings a hydrophobic region in gp41 close to the host cell, resulting in its insertion into the host cell's membrane. This insertion causes an energetically favoured conformation rearrangement of the heptad repeat regions (HR1 and HR2) of gp41,¹⁰ which brings the transmembrane region of gp41 (in the virus membrane) in contact with the gp41 hydrophobic fusion peptide (inserted into the host's membrane). Thus the fusion pore is formed, enabling the virus capsid to enter the cell.

There are several protein–protein interactions that are central to HIV entry: the binding of gp120 to CD4, the binding of gp120 to the co-receptor and the internal interactions of gp41 during membrane fusion. Targeting these interactions holds considerable potential for the treatment of HIV, especially for individuals harbouring strains of HIV that are resistant to drugs targeting reverse transcriptase or protease (the majority of therapeutics currently used). There are currently two entry-inhibiting drugs (Maraviroc and Fuzeon) on the market, with several undergoing clinical trials.¹¹ The HIV envelope proteins contain regions that are highly diverse and variable in their primary sequence, as well as highly conserved regions (where any mutations result in non-functional viruses). This variability is a considerable challenge facing entry inhibitors, which results in widely differing patient baseline sensitivity and response to inhibitors targeting envelope proteins.¹²

The first point of contact between HIV and the host is the interaction of CD4 with gp120, which has been extensively targeted using a variety of strategies, without yielding any compounds in the clinic.¹¹ The next point of intervention in HIV entry is the interaction of gp120 with its chemokine co-receptor (CCR5 or CXCR4). The importance of CCR5 for (R5-tropic) HIV entry is demonstrated by a well-established association between a 32 base pair deletion variant of CCR5 (CCR5-Δ32) in 5–14% of individuals of European Caucasians, and protection from HIV infection.¹³ The presence of one copy of this allele has been shown to delay

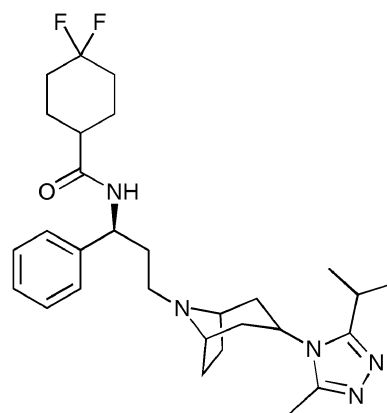


Fig. 4 Maraviroc. A viral entry inhibitor that prevents the interaction of the viral gp120 protein with its chemokine co-receptor CCR5.

the onset of AIDS by around 2 years, while those with two copies of the CCR5-Δ32 allele have strong protection against R5-tropic HIV strains and may not become infected at all. As these individuals (with the CCR5-Δ32 allele) are healthy, it is suggested that CCR5 is largely dispensable. As HIV has not existed long enough in the population to exert selective pressure for the CCR5-Δ32 allele, this mutation is thought to have arisen as a result of either bubonic plague or smallpox, which may explain the intense selectivity observed in Europeans.¹⁴ Unlike CCR5 however, CXCR4 is essential for the viability of multiple physiological processes, and deletion of the gene causes embryonic death in mice.¹⁵

The only inhibitor of the interaction of gp120 with a chemokine co-receptor is Maraviroc (Fig. 4, UK-427,857, Pfizer, Inc.),¹⁶ a potent (IC₅₀ of 2 nM), orally bioavailable antagonist of CCR5 that inhibits viral entry by R5- and mixed-tropic strains of HIV. The compound is the product of medicinal chemistry development and optimisation of a lead identified by high-throughput screening. It is typically prescribed to patients with R5 tropic HIV strains that are resistant to multiple antiretroviral agents. As Maraviroc binds a human chemokine receptor (rather than a viral protein), it is expected to be less susceptible to acquired resistance. However, Maraviroc-resistant strains of HIV (with mutations at residues 316 and 323 in V3 loop of gp120) have been identified in primary isolates¹⁷ that are seemingly able to use Maraviroc-bound CCR5 for entry. Another source of resistance observed in patients with virological failure is changes in viral tropism from R5 to X4 or mixed-tropic.¹⁷ An additional complication is that as the disease progresses, there is a natural shift from R5 to X4-tropic viruses;¹⁸ as Maraviroc is approved for use in patients with multiple-drug resistant strains of HIV (typically in the latter stages of the disease), the possibility of treatment failure as a result of X4-tropic viral outgrowth is increased. There have therefore been suggestions that Maraviroc may be better utilized in patients at earlier stages of HIV. There are several other entry inhibitors currently being developed that target the interaction of gp120 with CCR5.¹¹ Antibodies that block the CCR5 receptor and prevent binding of gp120 are another promising class of entry inhibitors that are currently undergoing clinical trials.¹⁹

Although there are no current clinical trials of compounds that inhibit the interaction of gp120 with CXCR4,¹¹ there are several in development.²⁰ These compounds, when administered in combination with a CCR5 inhibitor, may provide an effective method for inhibiting HIV entry. Such a combination therapy would be effective regardless of viral tropism and would eliminate the selective pressure for X4-tropic HIV strains in patients taking inhibitors targeting R5 strains.

The other HIV entry inhibitor currently on the market is Fuzeon (Enfuvirtide, Roche), a linear 36 amino acid peptide, derived from the HR2 region of gp41.²¹ Fuzeon is a membrane fusion inhibitor that was approved for use in 2003. Due to its cost and the need for intravenous administration, it is typically used as salvage therapy in patients with multi-drug resistant HIV. As outlined above, gp41 mediated membrane fusion involves the formation of a six-helix bundle made up of the HR1 and HR2 domains of the three gp41 proteins of the functional trimeric spikes. Fuzeon works by competitively binding to HR1, thus preventing the formation of the six-helix bundle that is necessary for membrane fusion. Acquired resistance requires mutations in the 10 amino acid motif between residues 36 and 45 of gp41. As this 10 amino acid motif is also critical for viral fusion, fuzeon-resistant mutants show poor replicative capacity compared with wild type. Clinical isolates have therefore shown variations in sensitivity to Fuzeon, but primary resistance has not been observed.²² There are several other fusion inhibitors currently in development.¹¹ These include several “next generation” peptides that have improved efficacy and are active against some fuzeon-resistant strains.²³ As these compounds target a different part of gp41 (to that targeted by fuzeon) there is potential for the development of synergistic therapies based on entry inhibitors.

Inhibition of the host's antiviral defences

There exists in each host cell sophisticated defence mechanisms that offer protection from viral infections. Viruses in turn have evolved infectivity factors that inactivate the host's innate defences, allowing infection to proceed unchecked. The HIV viral infectivity factor (ViF)²⁴ is present in the capsid, and neutralises a potent antiviral pathway present in nonpermissive host cells (T lymphocytes, macrophages and several leukemic T-cell lines). In the absence of ViF, this antiviral pathway is sufficient to effectively inactivate HIV-1,²⁵ and thus needs to be deactivated (by the virus). The key factor in this antiviral response is APOBEC3G,²⁶ a host protein that functions by deaminating cytidines from the negative strand of HIV-1 DNA during reverse transcription. Deamination of cytidine converts it to a uridine, thus rapidly introducing a very large number of mutations into the viral genome (GC base-pairs are effectively changed to AT base pairs) rendering the viral genomic information useless, and coding for nonsense. ViF binds to APOBEC3G, inducing the rapid degradation of APOBEC3G *via* a proteasome-dependant pathway, thus silencing the host's antiviral defences.²⁷ Inhibiting this protein–protein interaction is expected to restore the host's defences and inhibit HIV infection. This promise has been

demonstrated by a compound recently uncovered through screening a library of 30 000 small molecules for inhibitors of the ViF/APOBEC3G interaction.²⁸ The compound was shown to antagonise ViF function and reduce viral infectivity by increasing APOBEC3G incorporation into virions, enhancing cytidine deamination of the viral genome and inhibiting HIV-1 replication. Although the presence of permissive cells (those that do not contain APOBEC3G and its antiviral activity) limits the benefit of this approach as a single therapy, such agents will be of potential benefit in combination therapies.

Reverse transcription

HIV stores its genomic information as RNA, whereas the host uses DNA. In order to transplant viral genomic information into the host cells, it must first be copied from single-strand RNA onto duplex DNA, in a process called reverse transcription. As the host lacks the cellular machinery to carry out this conversion (genomic information typically flows from DNA to RNA to protein; the central dogma of molecular biology), the virus needs to supply its own reverse transcriptase enzyme for this purpose. HIV reverse transcriptase (Fig. 5) is a heterodimer, made up of a 560-residue subunit (p66) and a 440-residue subunit (p51); the dimeric form is essential for all enzymatic activity.²⁹ The functional complex is generated from a p66 homodimer by HIV protease, which cleaves one of the C-terminal RNase H domains during maturation of the virus particle. The resulting p66/p51 heterodimer is a multi-functional enzyme, with polymerase and nuclease activity. It is interesting to note that the p66 subunit performs both the polymerase and nuclease functions, while the p51 subunit is inactive, but acts as structural support.³⁰

All reverse transcription reactions require a primer with a free 3'-hydroxyl group to initiate cDNA synthesis; HIV reverse transcriptase uses a lysine tRNA as the replication



Fig. 5 HIV reverse transcriptase bound to DNA. The enzyme is a p66/p51 heterodimer. The RNase H domain (orange) of p66 (yellow) is highlighted; p51 is coloured red. This figure was made using PyMol 1.2 and structure 2HMI from The Protein Data Bank.

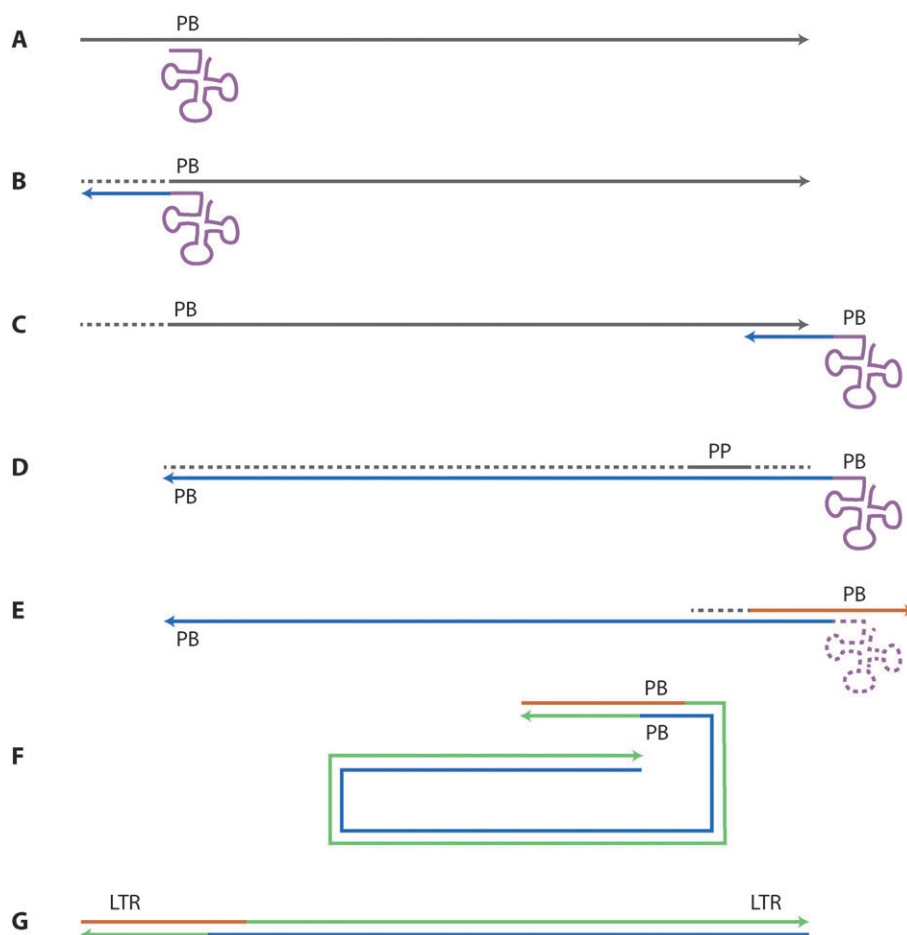


Fig. 6 Mechanism of HIV reverse transcription. (A) A lysine tRNA binds to the primer binding (PB) region on the RNA genome. (B) As DNA is synthesized (reverse transcription only occurs in the 3' to 5' direction), the complementary RNA strand is degraded. (C) The DNA-tRNA hybrid is transferred to the 3' end of the template, binding to a complementary site. (D) The viral single strand RNA template is degraded except for the polypurine (PP) section. (E) PP serves as a primer for second strand synthesis. (F) The tRNA is degraded. Note the matching PB sites on the first and second strand. (G) The first and second DNA strands hybridise at their PB site allowing double strand DNA synthesis to be completed (each serving as template for the other) by the DNA polymerase function of reverse transcriptase. (G) Reverse transcriptase will add LTR regions at each terminus of the DNA.

primer (Fig. 6A).³¹ The 18 nucleotides of the tRNA^{Lys,3} anneal to the complementary primer binding (PB) region in the viral genome (in the untranslated leader region), initiating reverse transcription (Fig. 6B). As the single strand DNA is synthesized, the complementary single strand RNA (viral genome) is degraded by the RNase H function of the enzyme. The DNA polymerase activity of HIV reverse transcriptase completes the synthesis of the double strand DNA copy of the viral genome³² (Fig. 6C–G), adding long terminal repeats, used by HIV integrase to incorporate the viral genome into that of the host.

The HIV reverse transcriptase is essential to the virus, and therefore an attractive target for antiviral therapies. A large group of current HIV therapies on the market target reverse transcriptase, with the majority being nucleoside or nucleotide analogues, such as azidothymidine (AZT)³³ or Didanosine (ddI).³⁴ These compounds³⁵ all lack the 3' hydroxyl group of deoxyribose, thus their incorporation into the growing oligonucleotide causes chain termination (there is no 3' hydroxyl group on the inhibitor to bind the 5' phosphate group of the

next incoming nucleotide). Although an integral component of almost all current HIV treatment programs, these inhibitors are very susceptible to the evolution of drug-resistant strains of HIV. The rapid emergence of mutations conferring drug-resistance to HIV observed in the clinic is due to the poor fidelity of the reverse transcription process, estimated to be around 1 mistake per 1700 bases, and as high as 1 in 70 when copying certain regions of the genome.³⁶ The high error rate is the result of the absence of an intrinsic exonucleolytic proofreading mechanism in HIV reverse transcriptase. This is actually advantageous to the virus, as it allows the rapid production of mutant strains that undergo growth advantage selection in the host. Each time the reverse transcriptase copies the HIV genome it has the potential to make mistakes that introduce mutations into the proteins of the progeny that may confer a growth advantage on the resulting virus (through resistance to a drug or better evasion of the host immune system). The resulting strain will have a growth advantage in the host and become dominant. The low fidelity of HIV reverse transcriptase is (ironically) essential for

the inhibitory activity of nucleoside and nucleotide analogues, as it enables their incorporation into the reverse transcribed DNA chain without subsequent excision and repair.

A promising alternative strategy for inhibition of HIV reverse transcriptase is to disrupt the protein–protein interaction between the p66 and p51 subunits, as the heterodimeric form is required for enzyme function. There have been two examples of this approach in the literature,³⁷ one a small molecule derived *via* structure-based ligand design,^{37c} and the other a linear peptide derived from a structural motif located between residues 284 and 300 of p51.^{37b,c} Both of these compounds are far from the clinic, but they represent a start for the development of compounds that target the dimerisation of p66 and p51. Another interesting reverse transcriptase inhibitor is the thymine derivative TSAO,³⁸ which acts by destabilising the heterodimeric form of reverse transcriptase, leading to the loss of DNA binding activity. It is worth mentioning that regardless of structure or mechanism of action, all compounds targeting reverse transcriptase will be susceptible to acquired drug resistance, due to the high mutation rates of HIV genome replication.

Integration

Integration finalises infection, with the viral (reverse transcribed) cDNA being incorporated into the host's genome. This occurs *via* a series of carefully coordinated reactions, mediated by the HIV integrase enzyme. As with reverse transcription, there are no functional equivalents to integrase in the host, the enzyme is therefore essential to the HIV lifecycle. Integrase is made up of three domains: the zinc-binding N-terminal domain, the catalytic core domain and the DNA-binding C-terminal domain.³⁹ The catalytic core domain contains a conserved D64, D116, E152 “catalytic triad” that bind up to two divalent metal ions (Mg^{2+} or Mn^{2+}). These metal ions are essential for integrase activity and are present in other DNA processing enzymes in nature. Structural studies suggest that the active integrase complex is a multimer (possibly a homodimer or tetramer).³⁹

The reverse transcribed viral DNA is linear, blunt-ended and contained within a pre-integration complex including integrase, matrix and Viral Protein R (Vpr). The first step of integration is 3'-end processing, in which integrase recognises specific sequences in the long terminal repeats (LTRs) of the viral cDNA, and cleaves two nucleotides from the 3' end of each strand to expose terminal hydroxyl groups on the overhanging termini. The viral DNA remains bound to the multimeric integrase complex, and is transported into the nucleus *via* the multiple nuclear localisation signals in integrase, matrix and Vpr.⁴⁰ This initiates DNA strand transfer, which involves the insertion of processed viral cDNA ends into the host chromosomal DNA. Integrase catalyses the attack by the 3'-hydroxyl groups (at the termini of the viral cDNA strands) on a pair of phosphodiester bonds in the host's chromosomal DNA, with the site of attack on each strand being separated by 5 base pairs.⁴¹ Strand transfer⁴² produces an integration intermediate in which the 3'-end of each viral cDNA strand is joined to the host's chromosome, with two overhanging bases on the 5'-ends and a gap to the host

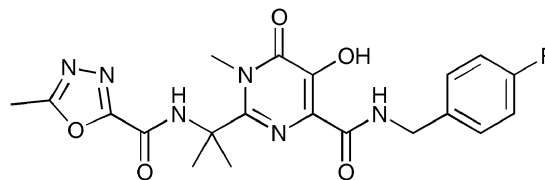


Fig. 7 Raltegravir. The only integrase inhibitor currently on the market, Raltegravir inhibits the integration of integrase with its DNA substrate.

chromosomal DNA. Integration is completed by the host's DNA repair enzymes trimming the overhanging bases, filling in the single strand gaps to the host chromosome, and ligating the ends together. There is no specific site of attack on the host's chromosome and integration can occur into any location. The integrated viral genome will lay dormant, especially in the latent stages of the disease, until cellular transcription factors enhance transcription of the viral genome and trigger the production of viral proteins (discussed in detail in the next section).

The only integrase inhibitor currently on the market is Raltegravir (Isentress, Merck & Co),⁴³ an orally bioavailable hydroxypyrimidinone carboxamide (Fig. 7). Raltegravir acts by binding to the divalent metal ions in the integrase catalytic core domain, preventing their interaction with DNA and inhibiting the strand transfer step. Inhibition of integrase by Raltegravir is followed by a very strong reduction in viral loads; however, resistance has evolved readily in the clinic,⁴⁴ necessitating the development of second-generation inhibitors or alternative strategies for targeting integrase.⁴⁵ There have been attempts to inhibit integrase activity by disrupting the dimerization of the enzyme with peptides derived from the interface region of the catalytic domain of the enzyme.⁴⁶ These peptides were inhibitors of the 3'-endonuclease activity of the enzyme (IC_{50} values in the low μM range), and were found to inhibit the cross-linking of the dimeric form of integrase. Another peptide inhibitor of integrase dimerisation is Indolicidin, a naturally occurring, 13-mer antimicrobial peptide that inhibits both the 3'-endonuclease activity and strand transfer steps of integrase. Further derivatisation and optimisation of this peptide have substantially improved its activity (IC_{50}) from 60 μM to 600 nM.⁴⁷ But the relative ease with which Raltegravir-resistance has evolved in the clinic⁴⁴ strongly suggests that direct targeting of the HIV integrase enzyme may not be the best approach for inhibition of the integration step. As integration is a highly organised, multistep process requiring several host factors,⁴⁸ a better approach may be to indirectly target integrase by inhibiting its interaction with an essential host protein. One such co-factor is LEDGF/p75 (lens epithelium-derived growth factor), whose interaction with integrase modulates the tethering of the pre-integration complex to chromatin; LEDGF/p75 binds to chromatin *via* a PWWP domain in its N-terminus, and binds to integrase through a domain in its C-terminus. This interaction has been shown to be crucial for HIV integrase activity *via* RNAi and knockout studies. An attractive alternative approach to inhibition of integrase would be to uncover small molecules that are capable of inhibiting this protein–protein

interaction. Such a compound was recently identified by analysis of the co-crystal structure of HIV integrase catalytic core domain with the C-terminus of LEDGF/p75; this was followed by virtual screening of a library of 200 000 compounds, and the lead compounds were further optimised by rational design.⁴⁹ The most potent inhibitor blocked the replication of primary isolates of HIV with an IC₅₀ of 580 nM, by binding to integrase and inhibiting its essential interaction with LEDGF/p75. Although the lead compound was found to be active against HIV strains resistant to Raltegravir, a double point mutation in integrase (A128T, E170G) is sufficient for the virus to become fully resistant to the inhibitor. This demonstrates the pitfalls of targeting binding pockets on viral proteins, which have an exceptionally high mutation rate, and therefore enable the virus to escape the effects of a single compound. A better approach may have been to develop compounds that bind to LEDGF/p75 and inhibit/block its interaction with integrase, as the host protein will have a far lower mutation rate, thus its binding pockets will be more robust and durable (than those on viral proteins).

Transcription and translation

Transcription of HIV genes from the integrated viral DNA (known as the provirus) requires several host and viral proteins. The 5' LTR of the viral genome contains the HIV promoter, as well as sites for several host transcription factors. Among these, the binding site(s) for NF- κ B (a host protein complex that controls DNA transcription) play a central role in mediating and inducing HIV gene expression. Interestingly, HIV strains have co-evolved with variations in their LTR promoter-proximal region. The HIV-1 subtype E that predominates in Asia contains one NF- κ B binding site, the LTR from subtype C that predominates in Africa contains three NF- κ B binding sites, and the North American subtype B contains two NF- κ B binding sites. NF- κ B and other cellular factors determine the rate of transcription from the integrated viral genome, and variations in their intracellular levels (in various cell types, or points in their lifecycle) is thought to be a key factor in determining whether the provirus is laying dormant or actively replicating. Despite the effective use of host signalling pathways to facilitate and enhance transcription of the integrated viral genome, the production of elongated transcripts is quite inefficient. The viral protein Tat (Trans-Activator of Transcription) plays a key role in enhancing and rapidly up-regulating transcription. Initially, only a small number of transcripts are produced, resulting in the build up of Tat, which binds cellular kinases that trigger the phosphorylation of the C-terminal domain of the largest subunit of RNA polymerase II. This results in a dramatic increase (around 100-fold) in transcription of the ~9 kb viral genome.

The complexity of the HIV genome (Fig. 2) results in the possibility of more than 30 different viral mRNA species in the host cell during transcription and translation. As previously mentioned, the expression of each of the multiple overlapping genes within a single region of HIV proviral DNA is regulated by mRNA splicing. These diverse, subgenomic species are generated *via* four different 5'-splice sites and eight different 3'-splice sites. This leads to various fully spliced (~2 kb

mRNAs that encode Rev, Tat or Nef; several partially spliced (~4 kb) mRNAs that encode Env, Vif, Vpr, or Vpu; as well as the full-length transcript (~9 kb) that encodes the Gag and Gag-Pol genes. The viral mRNA is initially fully spliced by the host's splicing factors that retain and splice (in the nucleus) all intron containing pre-RNA. To produce the partially spliced, and full length viral mRNA (that encode the rest of the viral proteins), the virus suppresses the host's nuclear retention mechanism to enable the translocation of intron-containing unspliced and partially spliced mRNA to the cytoplasm. This is the role of the Rev (Regulator of Virion) protein⁵⁰ that functions by binding to the Rev Response Element (REE) in the env-coding region of the HIV mRNA and enhancing the export of unspliced mRNAs from the nucleus to the cytoplasm. In the early phase of provirus transcription, all the HIV mRNA is in fully spliced resulting in an increase in the cellular levels of Rev, Tat and Nef. As already mentioned, the up-regulation of Tat increases levels of full length mRNA, while the increase in Rev activates the transport of unspliced or partially-spliced mRNA to the cytoplasm, resulting in the production of the rest of the HIV proteins.

Tat and Rev play essential roles in the transcription and translation of the provirus genome (and therefore the lifecycle of HIV), and there are examples of compounds that directly or indirectly inhibit Tat and Rev in the literature.⁵¹ Such inhibitors are unlikely to effectively combat HIV infection on their own, and will be redundant during the latent phase of infection. They may however prove beneficial in combination with other therapeutic agents, especially to those in the latter stages of infection.

Assembly and budding

The assembly and release (known as budding) of HIV from the host occurs in a series of organised steps that are driven by the viral Gag protein.⁵² Gag is necessary and sufficient for virus assembly; in the absence of other components of HIV, Gag drives the spontaneous assembly and release of spherical virus-like particles. The Gag polyprotein is made up of four domains: matrix (MA), capsid (CA), nucleocapsid (NC) and p6, as well as two smaller spacer peptides SP1 and SP2. The N-terminal MA domain of Gag binds to the host's plasma membrane where Gag-Gag interactions (and their intrinsic sphere-forming properties) form the structural shell of budding virus particles.⁵³ The MA domain also recruits Env glycoproteins to the host's plasma membrane, ensuring the presence of the HIV envelope proteins in the progeny virus particles. The NC region of Gag recognises and binds the dimeric full-length viral RNA transcript *via* a four stem-loop element located at the end of the gag gene. Gag also binds to, and actively directs the packaging of, other viral proteins (such as Gag-Pol) into the assembling particle. Although Gag enables the formation and assembly of viral spheres on membrane surfaces, components of the host's endosomal sorting complex are essential for budding, which involves separation of the nascent virion envelope from the cell membrane, releasing the virus particle. TSG101 is a central component of the endocytic machinery (ESCRT-I complex), whose activity is normally mediated by the endosomal protein

HRS (hepatocyte growth factor-regulated tyrosine kinase substrate) binding *via* a PSAP motif. HIV mimics the binding of HRS through a competing PTAP motif contained within the p6 region of Gag, which enables the recruitment of TSG101 and associated ESCRT-I complex proteins from their normal site of action on the late endosome to the plasma membrane to mediate budding.⁵⁴ Studies have shown that siRNA mediated ablation of TSG101 severely impairs viral production by arresting the release of viral particles from the plasma membrane of host cells, and disruption of the TSG101/Gag interaction by mutation of the Gag PTAP motif also blocks viral budding. There are currently no drugs on the market that target this step of the HIV lifecycle. Our own efforts have focused on targeting the interaction between the p6 region of Gag and TSG101. We have screened a library of over 10^8 cyclic peptides,⁵⁵ and have identified an inhibitor (8-mer cyclic peptide) of the above interaction that functions by binding to TSG101 and preventing its interaction with p6.⁵⁶ An alternative approach is to develop analogues of the PTAP motif that will competitively bind to TSG101 in the place of p6.⁵⁷ Although in the early stages of development TSG101-targeting inhibitors represent a very promising approach to inhibiting HIV in infected individuals. As such inhibitors act by binding to a host protein, it is hoped that they may not be as readily susceptible to viral mutations as compounds that directly bind to viral proteins.

Maturation

Upon release of the immature and non-infectious virus particles from the cell, the Gag polyprotein is cleaved by HIV protease to generate the mature Gag proteins MA, CA, NC, and p6. The proteolytic processing of Gag occurs in a series of ordered-sequential events, controlled by the differential rate of processing at each of the five cleavage sites. Gag is initially cleaved in two to give MA-CA-SP1 and NC-SP2-p6. Subsequent cleavages liberate MA and p6, followed by proteolysis of the remaining CA-SP1 and NC-SP2 fragments. Gag-Pol is also cleaved by HIV protease to release protease, reverse transcriptase and integrase. Accurate processing of Gag is essential to the infectivity of the virus particle, making HIV protease a prime target for inhibition, with multiple compounds having been developed



Fig. 8 HIV protease. The functional enzyme is a homodimer, with a DTG catalytic triad (highlighted) from each subunit making up the active site at the dimer interface. This figure was made using PyMol 1.2 and structure 3KF1 from The Protein Data Bank.

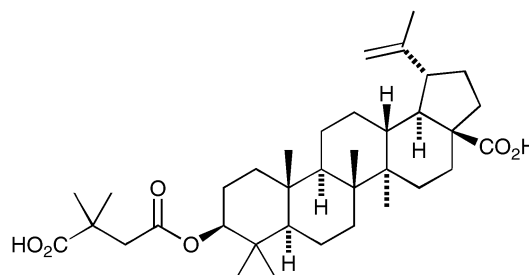


Fig. 9 Bevirimat. A maturation inhibitor that prevents the interaction of the CA-SP1 fragment with protease, therefore inhibiting its cleavage.

and available for use in the clinic. As with other compounds targeting viral proteins however, protease inhibitors are very susceptible to drug resistance rapidly acquired through mutation. They are therefore typically administered in combination with drugs targeting other steps in the HIV lifecycle (typically reverse transcription). An alternative approach is to target the homodimerisation of HIV protease, which is essential for its function. Each subunit contributes an Asp-Thr-Gly catalytic triad to form the active site, which is located at the dimer interface (Fig. 8). Several peptides have been identified (by logical design or high-throughput screening) that inhibit dimerisation of HIV protease, as well as a group of triterpenes identified by computational screening.⁵⁸ These compounds are still in the development stage and quite far from being deployed in the clinic. Despite their novel mode of action, any HIV protease dimerisation inhibitor will be as susceptible to acquired drug resistance as those currently on the market.

Upon cleavage by protease, the processed Gag proteins co-ordinate the rearrangement of the immature virion to infectious, mature virus particles. Maturation changes the morphology of the virus from spherical particles to those containing a conical capsid core composed of around 1500 CA molecules. The capsid self-assembles through CA/CA protein-protein interactions and contains the viral RNA genome in complex with NC, as well as other HIV proteins required for infection of a new host cell. CA/CA interactions are central to virus maturation and capsid formation and potentially an attractive target for the development of novel antiviral agents that inhibit virus maturation.⁵⁹ A first step in this process is the discovery of a 12-mer linear peptide (by phage display) that binds to a conserved hydrophobic groove in CA, altering the dimer interface and inhibiting assembly of immature and mature capsid particles *in vitro*.⁶⁰ The peptide was not active in various *in vivo* viral release assays, probably due to its inability to cross the cell membrane. The study nonetheless identified an allosteric pocket on CA, which could serve as the starting point for the design and development of a small molecule maturation inhibitor.

An alternative approach to preventing virus maturation is to inhibit Gag processing by targeting its interaction with HIV protease. One of the first compounds in this category is Bevirimat (Fig. 9, also known as PA-457),⁶¹ which has been shown to inhibit the replication of various HIV strains by blocking the cleavage of the CA-SP1 Gag fragment, preventing

the formation of CA and virus maturation. The mechanism of action of this compound is yet to be fully determined, but there is strong evidence suggesting that Bevirimat binds to Gag near or at the CA-SP1 junction, preventing its interaction with protease (either directly, or by altering the conformation of Gag). The compound has been undergoing clinical trials, performing well in early studies, but *in vitro* studies have shown that Bevirimat resistance may be acquired by a single point mutation (A1V) in SP1. This mutant was observed in the virus from 2 patients (out of 46) participating in Phase IIb clinical trials, alongside several other mutations in protease that render reduced susceptibility to treatment.⁵⁹ As a result, further development of Bevirimat was halted by the parent company earlier this year.

Concluding remarks

It has become increasingly clear that we are unlikely to see a single, 'magic-bullet' treatment for HIV, and that the best approach is likely the combination of compounds that attack the virus at multiple stages in its lifecycle. The ever-evolving nature of the virus challenges even this approach, resulting in a pressing need for new, innovative therapies. Targeting protein–protein interactions with small molecules holds much promise, but is a relatively new approach and still in its infancy. There are multiple challenges associated with uncovering and developing compounds that inhibit protein–protein interactions (such as lack of easily identifiable substrate pockets on protein interfaces), causing progress to be slower than traditional drug discovery approaches (targeting enzymes and their active sites). The urgent need for new HIV therapeutics that function *via* novel and robust mechanisms however make overcoming these challenges a matter of priority and urgency.

References

- P. Cane, I. Chrystie, D. Dunn, B. Evans, A. M. Geretti, H. Green, A. Phillips, D. Pillay, K. Porter, A. Pozniak, C. Sabin, E. Smit, J. Weber and M. Zuckerman, *Br. Med. J.*, 2005, **331**, 1368.
- A. M. Wensing, D. A. van de Vijver, G. Angarano, B. Asjo, C. Balotta, E. Boeri, R. Camacho, M. L. Chaix, D. Costagliola, A. De Luca, I. Derdelinckx, Z. Grossman, O. Hamouda, A. Hatzakis, R. Hemmer, A. Hoepelman, A. Horban, K. Korn, C. Kucherer, T. Leitner, C. Loveday, E. MacRae, I. Maljkovic, C. de Mendoza, L. Meyer, C. Nielsen, E. L. Op de Coul, V. Ormaasen, D. Paraskevis, L. Perrin, E. Puchhammer-Stockl, L. Ruiz, M. Salminen, J. C. Schmit, F. Schneider, R. Schuurman, V. Soriano, G. Stanczak, M. Stanojevic, A. M. Vandamme, K. Van Laethem, M. Violin, K. Wilbe, S. Yerly, M. Zazzi and C. A. Boucher, *J. Infect. Dis.*, 2005, **192**, 958.
- L. Lo Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177.
- (a) *Nat. Chem. Biol.*, 2007, **3**, 433; (b) C. P. Austin, L. S. Brady, T. R. Insel and F. S. Collins, *Science*, 2004, **306**, 1138; (c) R. L. Stein, *J. Biomol. Screening*, 2003, **8**, 615.
- A. D. Frankel and J. A. Young, *Annu. Rev. Biochem.*, 1998, **67**, 1.
- P. Zhu, J. Liu, J. Bess Jr., E. Chertova, J. D. Lifson, H. Grise, G. A. Ofek, K. A. Taylor and K. H. Roux, *Nature*, 2006, **441**, 847.
- P. D. Kwong, R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski and W. A. Hendrickson, *Nature*, 1998, **393**, 648.
- M. Pancera, S. Majeed, Y. E. Ban, L. Chen, C. C. Huang, L. Kong, Y. D. Kwon, J. Stuckey, T. Zhou, J. E. Robinson, W. R. Schief, J. Sodroski, R. Wyatt and P. D. Kwong, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 1166.
- (a) H. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhardt, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman and N. R. Landau, *Nature*, 1996, **381**, 661; (b) T. Dragic, V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore and W. A. Paxton, *Nature*, 1996, **381**, 667; (c) E. Oberlin, A. Amara, F. Bachelierie, C. Bessia, J. L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J. M. Heard, I. Clark-Lewis, D. F. Legler, M. Loetscher, M. Baggiolini and B. Moser, *Nature*, 1996, **382**, 833.
- D. C. Chan, D. Fass, J. M. Berger and P. S. Kim, *Cell*, 1997, **89**, 263.
- J. C. Tilton and R. W. Doms, *Antiviral Res.*, 2010, **85**, 91.
- P. Pugach, N. Ray, P. J. Klasse, T. J. Ketas, E. Michael, R. W. Doms, B. Lee and J. P. Moore, *Virology*, 2009, **387**, 296.
- P. C. Sabeti, E. Walsh, S. F. Schaffner, P. Varilly, B. Fry, H. B. Hutcheson, M. Cullen, T. S. Mikkelsen, J. Roy, N. Patterson, R. Cooper, D. Reich, D. Altshuler, S. O'Brien and E. S. Lander, *PLoS Biol.*, 2005, **3**, e378.
- (a) J. C. Stephens, D. E. Reich, D. B. Goldstein, H. D. Shin, M. W. Smith, M. Carrington, C. Winkler, G. A. Huttley, R. Allikmets, L. Schriml, B. Gerrard, M. Malasky, M. D. Ramos, S. Morlot, M. Tzetzis, C. Oddoux, F. S. di Giovine, G. Nasiouas, D. Chandler, M. Aseev, M. Hanson, L. Kalaydjieva, D. Glavac, P. Gasparini, E. Kanavakis, M. Claustres, M. Kambouris, H. Ostrer, G. Duff, V. Baranov, H. Sibul, A. Metspalu, D. Goldman, N. Martin, D. Duffy, J. Schmidtke, X. Estivill, S. J. O'Brien and M. Dean, *Am. J. Hum. Genet.*, 1998, **62**, 1507; (b) A. P. Galvani and M. Slatkin, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 15276.
- Y. R. Zou, A. H. Kottmann, M. Kuroda, I. Taniuchi and D. R. Littman, *Nature*, 1998, **393**, 595.
- P. Dorr, M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammen, A. Wood and M. Perros, *Antimicrob. Agents Chemother.*, 2005, **49**, 4721.
- M. Westby, C. Smith-Burchnell, J. Mori, M. Lewis, M. Mosley, M. Stockdale, P. Dorr, G. Ciaramella and M. Perros, *J. Virol.*, 2007, **81**, 2359.
- H. Schuitemaker, M. Koot, N. A. Kootstra, M. W. Dercksen, R. E. de Goede, R. P. van Steenwijk, J. M. Lange, J. K. Schattenkerk, F. Miedema and M. Tersmette, *J. Virol.*, 1992, **66**, 1354.
- J. M. Jacobson, M. S. Saag, M. A. Thompson, M. A. Fischl, R. Liporace, R. C. Reichman, R. R. Redfield, C. J. Fichtenbaum, B. S. Zingman, M. C. Patel, J. D. Murga, S. M. Pemrick, P. D'Ambrosio, M. Michael, H. Kroger, H. Ly, Y. Rotshteyn, R. Buice, S. A. Morris, J. J. Stavola, P. J. Maddon, A. B. Kremer and W. C. Olson, *J. Infect. Dis.*, 2008, **198**, 1345.
- (a) T. Murakami, S. Kumakura, T. Yamazaki, R. Tanaka, M. Hamatake, K. Okuma, W. Huang, J. Toma, J. Komano, M. Yanaka, Y. Tanaka and N. Yamamoto, *Antimicrob. Agents Chemother.*, 2009, **53**, 2940; (b) Y. Iwasaki, H. Akari, T. Murakami, S. Kumakura, M. Z. Dewan, M. Yanaka and N. Yamamoto, *Cancer Sci.*, 2009, **100**, 778.
- C. Wild, T. Greenwell and T. Matthews, *AIDS Res. Hum. Retroviruses*, 1993, **9**, 1051.
- M. L. Greenberg and N. Cammack, *J. Antimicrob. Chemother.*, 2004, **54**, 333.
- J. J. Dwyer, K. L. Wilson, D. K. Davison, S. A. Freel, J. E. Seedorff, S. A. Wring, N. A. Tvermoes, T. J. Matthews, M. L. Greenberg and M. K. Delmedico, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12772.
- D. H. Gabuzda, K. Lawrence, E. Langhoff, E. Terwilliger, T. Dorfman, W. A. Haseltine and J. Sodroski, *J. Virol.*, 1992, **66**, 6489.
- U. von Schwedler, J. Song, C. Aiken and D. Trono, *J. Virol.*, 1993, **67**, 4945.
- A. M. Sheehy, N. C. Gaddis, J. D. Choi and M. H. Malim, *Nature*, 2002, **418**, 646.
- M. Marin, K. M. Rose, S. L. Kozak and D. Kabat, *Nat. Med.*, 2003, **9**, 1398.
- R. Nathans, H. Cao, N. Sharova, A. Ali, M. Sharkey, R. Stranska, M. Stevenson and T. M. Rana, *Nat. Biotechnol.*, 2008, **26**, 1187.

- 29 T. Restle, B. Muller and R. S. Goody, *FEBS Lett.*, 1992, **300**, 97.
- 30 L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice and T. A. Steitz, *Science*, 1992, **256**, 1783.
- 31 (a) S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole and M. Alizon, *Cell*, 1985, **40**, 9; (b) M. Guyader, M. Emerman, P. Sonigo, F. Clavel, L. Montagnier and M. Alizon, *Nature*, 1987, **326**, 662.
- 32 T. E. Abbink and B. Berkhout, *Virus Res.*, 2008, **134**, 4.
- 33 H. Mitsuya, K. J. Weinhold, P. A. Furman, M. H. St Clair, S. N. Lehrman, R. C. Gallo, D. Bolognesi, D. W. Barry and S. Broder, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 7096.
- 34 R. Yarchoan, H. Mitsuya, R. V. Thomas, J. M. Pluda, N. R. Hartman, C. F. Perno, K. S. Marczyk, J. P. Allain, D. G. Johns and S. Broder, *Science*, 1989, **245**, 412.
- 35 Y. El Safadi, V. Vivet-Boudou and R. Marquet, *Appl. Microbiol. Biotechnol.*, 2007, **75**, 723.
- 36 (a) B. D. Preston, B. J. Poiesz and L. A. Loeb, *Science*, 1988, **242**, 1168; (b) J. D. Roberts, K. Bebenek and T. A. Kunkel, *Science*, 1988, **242**, 1171.
- 37 (a) A. Agopian, E. Gros, G. Aldrian-Herrada, N. Bosquet, P. Clayette and G. Divita, *J. Biol. Chem.*, 2009, **284**, 254; (b) M. C. Morris, V. Robert-Hebmann, L. Chaloin, J. Mery, F. Heitz, C. Devaux, R. S. Goody and G. Divita, *J. Biol. Chem.*, 1999, **274**, 24941; (c) D. Grohmann, V. Corradi, M. Elbasyouny, A. Baude, F. Horenkamp, S. D. Laufer, F. Manetti, M. Botta and T. Restle, *ChemBioChem*, 2008, **9**, 916.
- 38 J. Balzarini, M. J. Perez-Perez, A. San-Felix, D. Schols, C. F. Perno, A. M. Vandamme, M. J. Camarasa and E. De Clercq, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 4392.
- 39 A. Engelman, F. D. Bushman and R. Craigie, *EMBO J.*, 1993, **12**, 3269.
- 40 M. D. Miller, C. M. Farnet and F. D. Bushman, *J. Virol.*, 1997, **71**, 5382.
- 41 A. Engelman, K. Mizuuchi and R. Craigie, *Cell*, 1991, **67**, 1211.
- 42 K. Mizuuchi and K. Adzuma, *Cell*, 1991, **66**, 129.
- 43 (a) V. Summa, A. Petrocchi, F. Bonelli, B. Crescenzi, M. Donghi, M. Ferrara, F. Fiore, C. Gardelli, O. Gonzalez Paz, D. J. Hazuda, P. Jones, O. Kinzel, R. Laufer, E. Monteagudo, E. Muraglia, E. Nizi, F. Orvieto, P. Pace, G. Pescatore, R. Scarpelli, K. Stillmock, M. V. Witmer and M. Rowley, *J. Med. Chem.*, 2008, **51**, 5843; (b) V. Summa, A. Petrocchi, V. G. Matassa, C. Gardelli, E. Muraglia, M. Rowley, O. G. Paz, R. Laufer, E. Monteagudo and P. Pace, *J. Med. Chem.*, 2006, **49**, 6646.
- 44 I. Malet, O. Delelis, M. A. Valantin, B. Montes, C. Soulie, M. Wiriden, L. Tchertanov, G. Peytavin, J. Reynes, J. F. Mouscadet, C. Katlama, V. Calvez and A. G. Marcelin, *Antimicrob. Agents Chemother.*, 2008, **52**, 1351.
- 45 V. Nair and G. Chi, *Rev. Med. Virol.*, 2007, **17**, 277.
- 46 L. Zhao, M. K. O'Reilly, M. D. Shultz and J. Chmielewski, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 1175.
- 47 K. Krajewski, C. Marchand, Y. Q. Long, Y. Pommier and P. P. Roller, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5595.
- 48 B. Van Maele, K. Busschots, L. Vandekerckhove, F. Christ and Z. Debyser, *Trends Biochem. Sci.*, 2006, **31**, 98.
- 49 F. Christ, A. Voet, A. Marchand, S. Nicolet, B. A. Desimmie, D. Marchand, D. Bardiot, N. J. Van der Veken, B. Van Remoortel, S. V. Strelkov, M. De Maeyer, P. Chaltin and Z. Debyser, *Nat. Chem. Biol.*, 2010, **6**, 442.
- 50 V. W. Pollard and M. H. Malim, *Annu. Rev. Microbiol.*, 1998, **52**, 491.
- 51 (a) I. D'Orso, J. R. Grunwell, R. L. Nakamura, C. Das and A. D. Frankel, *J. Virol.*, 2008, **82**, 9492; (b) D. Daelemans, E. Afonina, J. Nilsson, G. Werner, J. Kjems, E. De Clercq, G. N. Pavlakis and A. M. Vandamme, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 14440; (c) E. S. DeJong, C. E. Chang, M. K. Gilson and J. P. Marino, *Biochemistry*, 2003, **42**, 8035.
- 52 E. O. Freed, *Virology*, 1998, **251**, 1.
- 53 (a) B. K. Ganser-Pornillos, M. Yeager and W. I. Sundquist, *Curr. Opin. Struct. Biol.*, 2008, **18**, 203; (b) P. D. Bieniasz, *Cell Host Microbe*, 2009, **5**, 550.
- 54 D. G. Demirov, J. M. Orenstein and E. O. Freed, *J. Virol.*, 2002, **76**, 105.
- 55 A. Tavassoli and S. J. Benkovic, *Nat. Protoc.*, 2007, **2**, 1126.
- 56 A. Tavassoli, Q. Lu, J. Gam, H. Pan, S. J. Benkovic and S. N. Cohen, *ACS Chem. Biol.*, 2008, **3**, 757.
- 57 F. Liu, A. G. Stephen, C. S. Adamson, K. Gousset, M. J. Aman, E. O. Freed, R. J. Fisher and T. R. Burke Jr., *Org. Lett.*, 2006, **8**, 5165.
- 58 M. J. Camarasa, S. Velazquez, A. San-Felix, M. J. Perez-Perez and F. Gago, *Antiviral Res.*, 2006, **71**, 260.
- 59 C. S. Adamson, K. Salzwedel and E. O. Freed, *Expert Opin. Ther. Targets*, 2009, **13**, 895.
- 60 J. Sticht, M. Humbert, S. Findlow, J. Bodem, B. Muller, U. Dietrich, J. Werner and H. G. Krausslich, *Nat. Struct. Mol. Biol.*, 2005, **12**, 671.
- 61 F. Li, R. Goila-Gaur, K. Salzwedel, N. R. Kilgore, M. Reddick, C. Matallana, A. Castillo, D. Zoumplis, D. E. Martin, J. M. Orenstein, G. P. Allaway, E. O. Freed and C. T. Wild, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13555.