View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Online

University of Wollongong

Research Online

Faculty of Science, Medicine and Health - Papers: part A

Faculty of Science, Medicine and Health

1-1-2016

The discovery of allyltyrosine based tripeptides as selective inhibitors of the HIV-1 integrase strand-transfer reaction

Neal Dalton
University of Wollongong, nd15@uow.edu.au

Christopher P. Gordon
University of Wollongong, cpg02@uow.edu.au

Timothy Boyle
University of Wollongong

Nicholas Vandegraaff
Avexa Ltd

John Deadman
Avexa Ltd

See next page for additional authors

Follow this and additional works at: https://ro.uow.edu.au/smhpapers

Part of the Medicine and Health Sciences Commons, and the Social and Behavioral Sciences Commons

Recommended Citation

Dalton, Neal; Gordon, Christopher P.; Boyle, Timothy; Vandegraaff, Nicholas; Deadman, John; Rhodes, David I.; Coates, Jonathon A.; Pyne, Stephen G.; Keller, Paul A.; and Bremner, John B., "The discovery of allyltyrosine based tripeptides as selective inhibitors of the HIV-1 integrase strand-transfer reaction" (2016). *Faculty of Science, Medicine and Health - Papers: part A*. 4511.
https://ro.uow.edu.au/smhpapers/4511

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

The discovery of allyltyrosine based tripeptides as selective inhibitors of the HIV-1 integrase strand-transfer reaction

Abstract

From library screening of synthetic antimicrobial peptides, an O-allyltyrosine-based tripeptide was identified to possess inhibitory activity against HIV-1 integrase (IN) exhibiting an IC_{50} value of 17.5 µM in a combination 3'-processing and strand transfer microtitre plate assay. The tripeptide was subjected to structure-activity relationship (SAR) studies with 28 peptides, incorporating an array of natural and non-natural amino acids. Resulting SAR analysis revealed the allyltyrosine residue was a key feature for IN inhibitory activity whilst incorporation of a lysine residue and extended hydrophilic chains bearing a terminal methyl ester was advantageous. Addition of hydrophobic aromatic moieties to the N-terminal of the scaffold afforded compounds with improved inhibitory activity. Consolidation of these functionalities lead to the development of the tripeptide 96 which specifically inhibited the IN strand-transfer reaction with an IC_{50} value of 2.5 µM.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Dalton, N., Gordon, C. P., Boyle, T. P., Vandegraaff, N., Deadman, J., Rhodes, D. I., Coates, J. A., Pyne, S. G., Keller, P. A. & Bremner, J. B. (2016). The discovery of allyltyrosine based tripeptides as selective inhibitors of the HIV-1 integrase strand-transfer reaction. Organic and Biomolecular Chemistry, 14 (25), 6010-6023.

Authors

Neal Dalton, Christopher P. Gordon, Timothy Boyle, Nicholas Vandegraaff, John Deadman, David I. Rhodes, Jonathon A. Coates, Stephen G. Pyne, Paul A. Keller, and John B. Bremner

This journal article is available at Research Online: https://ro.uow.edu.au/smhpapers/4511

The Discovery of Allyltyrosine Based Tripeptides as Selective Inhibitors of the HIV-1 Integrase Strand-Transfer Reaction

Neal Dalton,[a] Christopher P. Gordon,*[a,b‡] Timothy P. Boyle,[a] Nicholas Vandegraaf,[c,d] John Deadman,[c] David I. Rhodes,[c,e] Jonathan A. Coates,[c] Stephen G. Pyne,[a] Paul A. Keller,*[a] and John B. Bremner[a]

[a] School of Chemistry, University of Wollongong, Northfields Avenue, Wollongong NSW 2522 Australia. E-mail: keller@uow.edu.au; Tel: +61 (0)242214692

[b] ‡Present address: School of Science and Health, Western Sydney University, Locked Bag 1797, Penrith South DC, Australia. E-mail: c.gordon@westernsydney.edu.au; Tel: +61 (02) 4620 3201

[c] Avexa Ltd, 576 Swan St, Richmond, Vic 3121, Australia.

[d] Present Address: Genera Biosystems, Scoresby, Vic 3179, Australia

[e] Present Address: Monash Institute of Materials Engineering, Monash University, Clayton, Vic 3800, Australia

Abstract: From library screening of synthetic antimicrobial peptides, an *O*-allyltyrosine-based tripeptide was identified to possess inhibitory activity against HIV-1 integrase (IN) exhibiting an IC_{50} value of 17.5 μM in a combination 3'-processing and strand transfer microtitre plate assay. The tripeptide was subjected to structure-activity relationship (SAR) studies with 28 peptides, incorporating an array of natural and non-natural amino acids. Resulting SAR analysis revealed the allyltyrosine residue was a key feature for IN inhibitory activity whilst incorporation of a lysine residue and extended hydrophilic chains bearing a terminal methyl ester was advantageous. Addition of hydrophobic aromatic moieties to the scaffold *N*-terminal afforded compounds with improved inhibitory activity. Consolidation of these observations led to the development of tripeptide **96** which specifically inhibited the IN strand-transfer reaction with an IC_{50} value of 2.5 μM.

Introduction

For those living in the developed world HIV infection has increasingly been considered as a chronic disease.[1] This remarkable turnaround is due primarily to the advent of highly active antiretroviral therapy (HAART) in which a combination of drugs, typically three to four, which target different steps in the viral lifecycle are taken. At present clinicians have a palette of drugs to formulate HAART schedules with twenty-seven FDA approved drugs for HIV therapy.[2] Clinically available agents include eight nucleoside and five non-nucleoside reverse transcriptase inhibitors, nine protease inhibitors, three integrase inhibitors, in addition to the fusion inhibitor Enfuvirtide, and the CCR5-blocker Maraviroc.[2] As a result of this array of agents, a 20-year-old HIV-positive patient in the U.S. or Canada today who is diagnosed at an early stage of infection and prescribed HAART is expected to live into their early 70's, a life expectancy approaching that of the general population.[3] Further, the roll-out of cheaper generic drugs across resource-poor settings has resulted in dramatic improvements in life expectancy. For example in Zimbabwe over the last decade the average life expectancy for HIV sufferers has increased 5.5 years to around 53 years.[2]

Despite these immense gains it is important to note that HIV remains an incurable disease with about 35.3 million people currently living with the condition across the globe.[4] Further, infection rates have not abated with about 2.1 million new infections reported in 2013 which equates to about 6300 new infections per day.[4] Against this backdrop HIV-1 strains displaying resistance against one or more of the aforementioned twenty seven currently FDA approved agents have been characterised.[5, 6] Additionally the rate of resistance evolution remains extremely rapid. For example since the latest integrase inhibitor Dolutegravir received FDA approval on August 13, 2013, four point mutations conferring resistance have been characterised.[5] Consequently until a cure is found it is essential that next generation anti-HIV agents are continually being progressed through the drug development pipeline.

Of the current set of utilised drug targets, HIV integrase (IN) remains relatively underexploited with only three inhibitors currently approved by the FDA, although interest in such inhibitors is strong.[7] This, in addition to IN having no counterparts in mammalian cells, continues to frame the enzyme as an attractive drug target. The IN enzyme is indispensable to the HIV life cycle and catalyses two distinct reactions, these being 3′-processing and strand-transfer. During 3′-processing, which occurs within the cytoplasm of an infected cell, integrase catalyses the excision of a 5′-GT dinucleotide from each end of the viral genome thereby generating the nucleophilic 3′-hydroxyl ends required for strand transfer.[8-10] This water-mediated endonucleolytic cleavage of the 5′-GT dinucleotides occurs immediately on the 3′ side to a highly conserved CA dinucleotide.[11-14] Following 3′-processing, integrase undergoes a structural change in preparation for the binding of the acceptor (chromosomal) DNA.[15, 16] Integrase, still bound to the 3′-processed viral DNA, translocates to the nucleus of the infected cell as part of a pre-integration complex (PIC), wherein the terminal 3′-OH of the viral DNA attacks the host DNA.[12, 17-19] This integration event is a point of no return for the host cell which then becomes a permanent carrier of the virus.[20]

The three IN inhibitors that have received FDA approval for HIV therapy are Raltegravir (RAL, **2**), Elvitegravir (EVG, **3**), and Dolutegravir (DTG, **4**) (Figure 1). Each of these agents selectively inhibits stand-transfer and as outlined in figure 1, each binding to the active-site *via* a similar mechanism. Having evolved from the first generation diketoacid inhibitors such as L-731988 (**1**), these analogues possess a diketoacid bioisostere which chelates the two catalytic magnesium ions within the HIV integrase active-site.[20, 21] Thus, whilst this paradigm provides a conduit to potent inhibition, single point mutations can endow cross resistance; for example, the clinically observed mutants F121Y and Q148H display cross resistance to RAL, EVG, and DTG.[5] Consequently the development of competitive inhibitors which bind to the active-site through alternative interactions, or elicit inhibition *via* allosteric mechanisms, would provide significant additions to the current HAART arsenal.

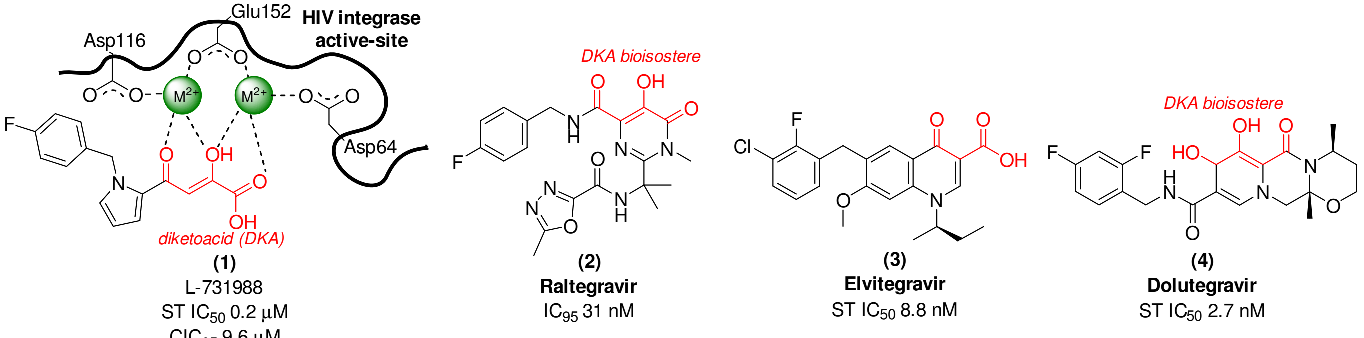


Fig. 1: Structure of a 1st generation IN inhibitor L-7931988 with schematic outlining the diketoacid-Mg^{2+} interactions along with the chemical structures of RAL, EVG, and DTG.

Indeed, small molecule IN allosteric inhibitors (ALLINIs) have been reported recently along with a number of co-crystallised structures.[22-24] These molecules disrupt the protein-protein interaction between transcriptional co-activator lens epithelium derived growth factor (LEDGF) and the IN catalytic core.[22-26] LEDGF has been shown to be a dominant factor to promote localisation of the PIC to the host chromatin as well as enhancing strand transfer in isolated protein assays. Full-length LEDGF was also shown to promote tetramerisation of full-length HIV-IN, which is essential for the integration of both viral DNA ends into the chromosomal DNA.[22-26] Of these analogues the most recently reported GSK1264 (5) has been co-crystallised within the LEDGF binding pocket of IN catalytic core (Figure 2). This compound inhibited HIV-1 replication with an EC50 value of ~38 nM.[22]

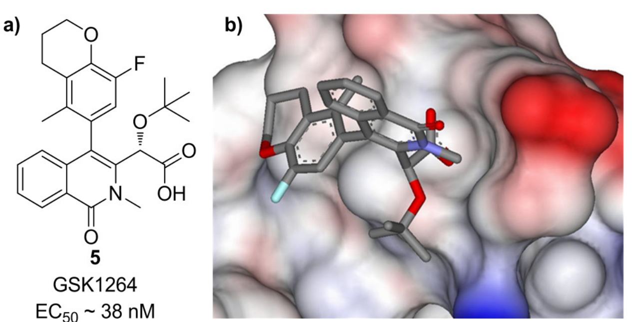
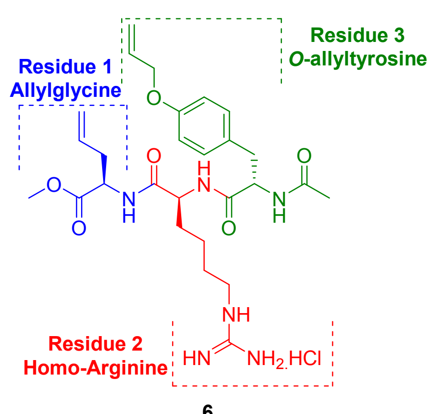


Fig. 2: a) The structure of GSK1264; b) Structure of co-crystallised, of GSK126 within the LEDGF binding pocket of the IN catalytic core (PDB accession code 4OJR).



3'-Processing and Strand-Transfer IC_{50} 17.5 μM

Fig. 3: *O*-Allyl-L-tyrosine-based tripeptide HIV-1 integrase inhibitor discovered from a screening program of a number of 'in-house' compound libraries.

Given our ongoing interest in the development of HIV inhibitors[23, 24, 27-29] and the renewed vigour for the development of next generation IN inhibitors, we conducted a screening program utilising a number of 'in-house' compound libraries from which an O-allyltyrosine-based tripeptide (Compound 6, Figure 3) was identified to inhibit IN with an IC50 value of 17.5 μM. This tripeptide, which emerged from our ongoing antibacterial drug design program,30-32 presented as an appealing scaffold for drug development endeavours since: a) analogues could be rapidly accessed via standard peptide coupling approaches, b) the scaffold is amenable to diverse structural and functional group alterations and c) the tripeptide bears no significant structural similarity to any currently reported peptide-based integrase inhibitors.[24, 25, 30, 33-37] Consequently we embarked on an extended structure-activity-relationship investigation of the *O*-allyltyrosine tripeptide scaffold in a bid to generate a pharmacophore for HIV-1 integrase inhibition.

Results and Discussion

To initiate structure-activity-relationship studies, the lead compound **6** was segmented into three regions; residue 1 (allylglycine), residue 2 (homo-arginine), and residue 3 (*O*-allytyrosine), respectively (Figure 3). It was envisaged that these libraries could be efficiency accessed through relatively standard peptide coupling approaches with minor alterations of the procedure utilised to synthesise **6**. Briefly, in the initial synthesis of **6**, the *N*-acetyl-*O*-allyltyrosine residue (**7**, Scheme 1) was prepared *via* nucleophilic *O*-allylation of commercially available (*S*)-*N*-acetyltyrosine ethyl ester (**8**) with allyl bromide (**9**). Subsequent ester hydrolysis afforded **10** which was coupled to **11** using typical EDCI-HOBt-mediated amide formation conditions with the resulting dipeptide hydrolysed to furnish the acid **13**. The methyl ester protected allylglycine residue **14** was obtained *via* thionyl chloride mediated esterification of commercially available allylglycine, and was coupled to dipeptide **13** again *via* EDCI-HOBt-mediated coupling. *N*-Boc-deprotection of the resulting tripeptide **16** used trifluoroacetic acid and the crude material was subsequently reacted with $(BocNH)_2C{=}NSO_2CF_3$ affording the protected arginine analogue **17** with final Boc-protection giving **6**.

Scheme 1: Synthetic procedure to access the lead allyltyrosine based tripeptide 6. Reagents and Conditions: i) K_2CO_3(aq) (2 eq.), ii) $LiOH.H_2O$ (2 eq.), THF/H_2O (3:1), iii) $NaHSO_4$ (2 M); iv) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, v) $SOCl_2$, CH_3OH vi) TFA/CH_2Cl_2 (1:1), vii) $(BocNH)_2C{=}NSO_2CF_3$ (1 eq.), viii) 1 M HCl/diethyl ether.

With an effective synthetic procedure in hand, initial investigations focused on the allylglycine region and as outlined in Scheme 2, eight analogues (compounds **18** – **25**) were synthesised. Here the specific aim was to probe for potential H-bond donating/accepting interactions whilst **22** was prepared to investigate pi-stacking interactions. Compound **18** was synthesised utilising **12** which was initially *N*-Boc-deprotected and subsequently treated with $(BocNH)_2C{=}NSO_2CF_3$ to afford the protected arginine analogue **26** with final TFA mediated de-protection affording **18** (Scheme 2). The remainder of the first series compounds were also prepared using **12** which was initially hydrolysed and the resulting free carboxylic acid was coupled to the desired amines using typical EDCI-HOBt-mediated conditions. The resulting *N*-Boc-protected analogous **29** – **35** were de-protected and subsequently treated with $(BocNH)_2C{=}NSO_2CF_3$ to afford the protected arginine analogues **36** – **42** and a final TFA mediated *N*-Boc-deprotection and then treatment with HCl in ether afforded the final desired analogues **19** – **25**.

This initial series of analogues were subjected to a previously reported combination 3'-processing and strand transfer microtitre plate assay[38, 39] and as outlined in Table 1 the inhibitory activities afforded by the assay indicated that the incorporation of nitrogen rich functionalities within the residue-1 region of the scaffold (e.g. **19** and **20**) was detrimental to IN inhibitory activity as was simplification to a methyl ester or glycine moiety (**18** and **21**, respectively) and inclusion of R-phenylalanine was also detrimental to inhibitory activity (e.g. **22**). However restoration of inhibitory activity was observed with inclusion of extended methyl ester moieties with the β-alanine analogue 23 displaying an IC_{50} value of 33 μM whilst the γ-aminobutyric analogue **24** and β-glutamic analogue **25** displayed superior activity to the lead with IC_{50} values of 10 μM, respectively.

Scheme 2: Synthetic procedures to access analogues the allyltyrosine modified analogues **18** through **25**. Reagents and conditions: i) $TFA{:}CH_2Cl_2$ 1:1; ii) $(BocNH)_2C{=}NSO_2CF_3$ (1 eq.); iii) 1 M HCl/diethyl ether; iv) $LiOH.H_2O$ (2 eq.), THF/H_2O (3:1); v) $NaHSO_4$ (2 M); vi) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF.

Table 1: HIV-IN inhibitory activity of the ally glycine modified analogues **18** – **25**

Compound	R_1	IC_{50} (μM) [a]
18		> 100
19	HN, N, NH2.HCl	> 100
20	NH, HCl.H2N, N, H	> 100
21	O, O, N, H	> 100
22	O, O, N, H	80
23	O, O, N, H	33
24	O, O, N, H	10
25	O, NH, O, O, O	10

[a]IC50 determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

Upon identifying a γ-aminobutyric methyl ester and β-glutamic dimethyl ester as superior alternatives to allylglycine, attention turned to the homoarginine residue with a specific aim of simplifying the guanidino moiety to a primary amine. Accordingly three analogues (**43** – **45**) were prepared. Compounds **43** and **44** contained a lysine moiety and were accessed through *N*-Boc-deprotection of the previously prepared analogues **33** and **34** (Scheme 2). The ornithine derivative **45** was synthesised in a four step procedure (Scheme 3) whereby the *N*-Boc-protected ornithine analogue **46** was coupled with the previously synthesised analogue **10** (e.g. Scheme 1, step 2) under typical EDCI-HOBt conditions and the resulting analogue **47** was hydrolysed to afford the acid **48**. Subsequent coupling with the ester **27** furnished **49** and final TFA mediated *N*-Boc-deprotection followed by treatment with HCl in ether yielded the desired analogue **45**.

The IN inhibitory activities of the homo-arginine modified analogues (**43** and **44**, Table 2) were similar to those of the corresponding homo-arginine analogues (**23** and **24**, Table 1), whilst the ornithine analogue **45** displayed a minor decrease in potency relative to the parent **44**. Thus this data indicated that the guanidino group was not essential for activity.

SAR analysis of the initial compound series (Table 1) indicated that extension of the carbon linker between the terminal methyl ester and amide moiety (i.e. **23** and **24**, Table 1) was advantageous for activity whilst incorporation of rigid or amine rich functionalities (i.e. compounds **19**, **20**, and **21**, Table 1) was detrimental. Thus to further investigate flexibility and polarity, an additional series of eight compounds was prepared (Scheme 4), each prepared in a two-step protocol from the previously synthesised **13,** which was coupled to the required amine under HOBt-EDCI mediated conditions with subsequent *N*-Boc-deprotection and HCl treatment affording the final analogues **50** – **58**.

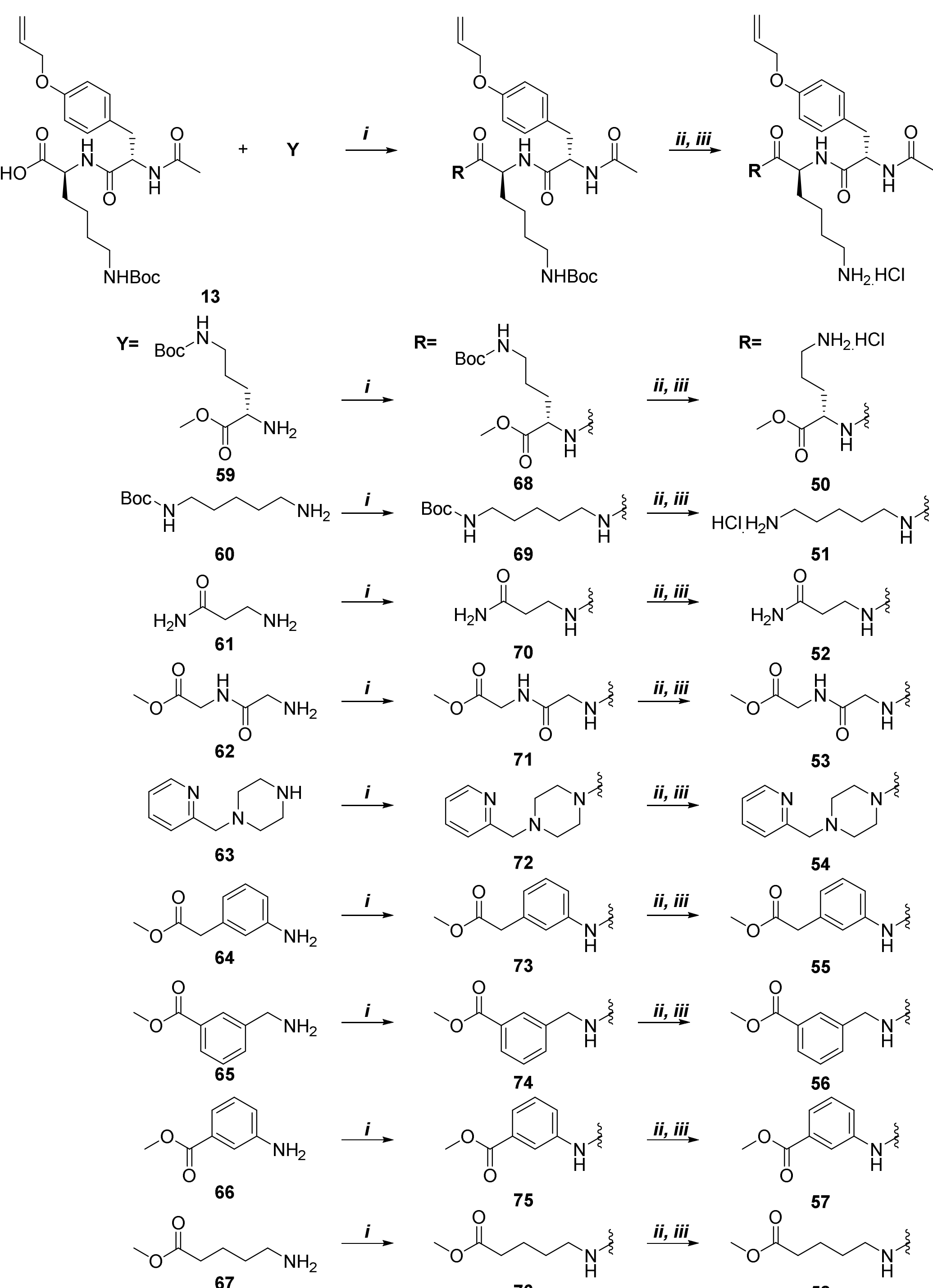
Scheme 3: Synthesis of the ornithine based analogue **45**. ***Reagents and conditions*: *i*)** EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, ***ii)*** LiOH.H_2O (2 eq.), THF/H_2O (3:1); ***iii***) $NaHSO_4$ (2 M); ***iv***) TFA:CH_2Cl_2 1:1; ***v***) 1 M HCl/diethyl ether.

Table 2: HIV IN inhibitory activities of the homo-arginine modified analogues **43** – **44**

Compound	R^1	R^2	IC_{50} (μM) [a]
43		NH_2·HCl	23
44		NH_2·HCl	10
45		NH_2·HCl	17

[a]IC_{50} determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

This third series of analogues was subjected to the combined 3'-processing and ST assay with the flexible primary amine analogues **50** and **51**, and the amide analogue **52** displaying similar inhibitory activity to the lead compound whereas once again rigid amine rich moieties (e.g. **53** and **54**) displaying reduced IN inhibitory activity (Table 3). However, as demonstrated by **55** and **57**, introduction of inflexible moieties did not automatically bestow reduced activity whilst the premise that extension of the carbon linker between the terminal methyl ester and amide afforded increased activity was further supported by compound **58**.



Scheme 4: Synthesis of the allylglycine modified analogues **50 – 58**. ***Reagents and conditions*: *i***) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, ***ii)*** TFA:CH_2Cl_2 1:1; ***iii***) 1 M HCl/diethyl ether.

Table 3: HIV-IN inhibitory activity of the ally glycine modified analogues **50** – **58**.

Compound	R^1	IC_{50} (μM)[a]
50		22
51		23
52		19
53		58
54		32
55		5
56		25
57		15
58		7

[a]IC_{50} determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

Upon establishing substantial SAR for the allylglycine region of the scaffold subsequent attention turned to the allyltyrosine region with three analogues prepared, **77** – **79**. As outlined is Scheme 5, **77** was prepared in a four-step procedure in which the previously prepared *N*-Boc-protected lysine derivative **11** was coupled to *N*-acetylphenylalanine to afford **81** which was successively hydrolysed, coupled with **27**, and *N*-Boc-deprotected to give **77** after HCl treatment.

Scheme 5: The four-step synthesis of the phenylalanine analogue **77**. ***Reagents and conditions*: *i*)** EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF; ***ii)*** LiOH.H_2O (2 eq.), THF/H_2O (3:1); ***iii)*** TFA:CH_2Cl_2 1:1; ***iv***) 1 M HCl/diethyl ether.

Compound **78** was prepared utilising the *N*-Boc-protected tripeptide **34** (Scheme 2), which was converted to the phenol **84** by treatment with catalytic tetrakis(triphenylphosphine)palladium(0) and

10 equivalents of morpholine in THF under nitrogen (Scheme 6). Subsequent *N*-Boc-deprotection of **84** and hydrochloride formation afforded the tyrosine analogue **78**.

Scheme 6: The two-step synthesis of the tyrosine analogue 78. ***Reagents and conditions:*** *i)* $Pd(PPh_3)_4$ (10 mol %), THF, rt, 10 min, then morpholine (10 eq.), 3 h; *ii)* TFA:CH_2Cl_2 1:1; *iii)* 1 M HCl/diethyl ether.

The final analogue in this series **79** was accessed *via* a five step procedure, and in contrast to the previously employed strategies, **79** was produced by sequential coupling to the C-terminal residue (Scheme 7). Initially the commercially available (*S*)-*N*-acetyltyrosine ethyl ester **85** was converted to the acetate ester **87** and subsequently hydrolysed to give **88**. Coupling of **88** with **89** furnished **90** and piperidine mediated Fmoc-deprotection afforded **91**. Unexpectedly, final Boc-deprotection of **91** and hydrochloride formation afforded **79** which arose from acid catalysed ester interchange in MeOH/HCl. Nevertheless this analogue was subjected to the integrase assay.

Scheme 7: The five-step synthesis of the methyl ester analogue **79** with the unexpected conversion to **79** from **91**. ***Reagents and conditions***: ***i***) K_2CO_3 (2 eq.), DMF; ***ii***) $LiOH.H_2O$ (2 eq.), THF/H_2O (3:1); ***iii***) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF; iv) 1% piperidine/acetonitrile; v) TFA:CH_2Cl_2 1:1; vi) 1 M HCl/diethyl ether MeOH.

As outlined in Table 4, each of the allyltyrosine modified derivatives displayed reduced inhibitory activities and the significantly reduced activities displayed by **78** and **79** indicates that the allyl moiety plays a crucial binding role, potentially participating in hydrophobic/π-stacking interactions with the enzyme active-site.

Having collated SAR data for the residues 1, 2 and 3 of the lead compound **6**, attention turned to *N*-terminal amide functionalised analogues. Each of these comprised the previously identified active functionalities of an extended carbon chain possessing a terminal methyl ester in the allylglycine region of the scaffold, a lysine at residue 2, and the allyltyrosine moiety at residue 3. As illustrated in reaction Schemes 8, 9, and 10, five analogues, **92** – **96** were prepared in this series. The synthesis of **92** was achieved with the initial coupling of Boc-Tyr(All)-OH (**97**) with the lysine methyl ester analogue **11** to afford the dipeptide **98** which was subsequently hydrolysed to give **99**. EDCI-HOBt

mediated amide coupling of **99** with **27** furnished the di-Boc-protected analogue **100** and final deprotection ultimately afforded the desired analogue **92** (Scheme 8).

Table 4: The HIV IN inhibitory activities of the allyltyrosine modified analogues **77** – **79**.

Compound	R^4	IC_{50} (μM) [a]
77		19
78	HO	50
79		60

[a]IC_{50} determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

Scheme 8: Synthesis of the lysine-based analogue **92**. ***Reagents and conditions***: ***i***) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, ***ii***) LiOH.H_2O (2 eq.), THF/H_2O (3:1); ***iii***) $NaHSO_4$ (2 M); iv) TFA:CH_2Cl_2 1:1; v) 1 M HCl/diethyl ether.

Compounds **93** to **95** were synthesised using a semi-convergent protocol from **101** (Scheme 9). Compound **101** was prepared in a four step procedure with the initial amide coupling of **102** to **27** furnishing **103**. Sequential Boc-deprotection and coupling of **89** with *N*-Boc-Tyr(All)-OH (**104**) provided **105** and final TFA mediated deprotection afforded the key Fmoc-protected intermediate **101**. Utilising **101**, the desired analogue **93** was prepared in two steps by the amide coupling with 2-pyridinecarboxylic acid followed by piperidine mediated Fmoc-deprotection. The synthesis of both **94** and **95** required the initial preparation of **107** which was obtained in a two-step procedure whereby methyl salicylate was converted to the benzyl ether derivative **109** using typical ether formation conditions and benzyl bromide. Prior to coupling, the ester was hydrolysed to the desired acid **107** under basic conditions. Standard EDCI-HOBt mediated amide formation utilising **107** and the key intermediate **101** afforded **110** and a final Fmoc-deprotection furnished the desired analogue **94**. Analogue **95** was prepared through the debenzylation of **110** and final Fmoc-deprotection afforded the desired analogue.

Scheme 9: The semi-convergent synthesis of the *N*-terminal modified analogues **93** – **95.** ***Reagents and conditions*****:** ***i***) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, ***ii)*** TFA:CH_2Cl_2 1:1; ***iii***) 1 M HCl/diethyl ether; ***iv***) 2-pyridinecarboxylic acid (1 eq.), EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF; ***v***) 1% piperidine/acetonitrile; ***vi***) $PhCH_2Br$ (2 eq.), K_2CO_3 (2 eq.), DMF; ***vii***) KOH (4 eq.), MeOH:H_2O (3:1); ***viii***) thioanisole (50 eq), TFA (2 mL).

Scheme 10: The four-step synthesis of the cbz-based analogue **96**. ***Reagents and conditions*****:** ***i***) K_2CO_3 (3 eq.), DMF; ***ii)*** LiOH.H_2O (2 eq.); ***iii***) **89** (1.0 eq) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF; ***iv***) 1% piperidine/acetonitrile; ***v***) 1 M HCl/diethyl ether.

The final compound **96** in this *N*-terminal series was furnished through a five step procedure whereby the phenolic moiety Cbz-Tyr-OMe (**111**) was initially *O*-allylated with allyl bromide to afford **112** and subsequent ester hydroylysis provided **113**. This was then coupled with the previously prepared **89** using EDCI-HOBt mediated amide formation conditions and final piperidine Fmoc-deprotection afforded **96** (Scheme 10).

As outlined in Table 4, modifications within this region had mixed effects on the HIV-IN inhibition, *e.g.* removal of the acetyl group or addition of a benzyl ester (e.g. **92** and **96** respectively), imparted minimal effects on inhibitory activity, while introduction of pyridine (**93**) or phenol (**95**) moieties resulted in significant activity reductions. However, the benzyl protected phenol derivative **94** displayed higher potency (IC_{50} 4 μM) than the previous most active analogue **55**.

Table 5: HIV-IN inhibitory activities of the *N*-terminal modified analogues **92 – 96**.

Compound	R^3	IC_{50} (μM) [a]
92	NH_2.HCl	14
93	N, H, O, N	31
94	N, H, O, O	4
95	N, H, O, OH	60
96	N, H, O, O	10

[a] IC_{50} determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

Thus, whilst the development of the five targeted compound libraries provided significant SAR data and culminated in the development of the most potent derivative **94**, little information relating to precise inhibitory activity was established as the IC_{50} values for each of the analogues were determined using a combination 3'-processing and strand transfer assay.[38, 39] Whilst the assay provides an expedient means of identifying general IN inhibitors, it affords limited insights to potential mechanisms of action. Consequently subsequent investigations focused on subjecting a number of the most active analogues to individual 3′-processing and strand-transfer inhibition assays. Initially 3′-processing inhibitory activity was examined in the presence of magnesium, as it is generally accepted that Mg^{2+} is the co-factor for integration in cells.[40] However, the assay was also performed using manganese as a co-factor as Mn^{2+} appears to be required *in vitro* for the DKAs to produce potent inhibition.[40-43]

As summarised in Table 6, the *O*-allyltyrosine analogues are specific inhibitors of the strand-transfer reaction as no inhibitory activity was observed in the 3′-processing assay up to compound concentrations of 30 μM with the most potent analogue **96** inhibiting the ST reaction with an IC_{50} value of 2.5 μM. These results suggest that the *O*-allyltyrosine analogues function *via* an alternative mechanism to previously reported peptide based inhibitors and the LEDGF/p75 allosteric inhibitors. For example the previously reported Vpr- and Env-derived peptides inhibit both 3'-processing and ST,[36] similarly both series of cell-permeable stapled Vpr-derived[35], IN-derived[34] peptides and combinatorial-derived hexapeptides[45] are also inhibitors of both 3'-processing and ST. To date, the only other reported peptide analogue to specifically inhibit ST was a heptapeptide which also displays cationic character.[46]

Moreover the previously reported series of small molecule inhibitors of the LEDGF/p75 interaction were equipotent against 3'-processing and ST[26] whilst the most recently reported LEDGF/p75 inhibitor GSK1264 is a potent inhibitor of 3'-processing.[22] Together this information suggests that the *O*-allyltyrosine analogues may function *via* a competitive mechanism similar to diketoacid-based analogues which are also specific inhibitors of the stand-transfer reaction.[41]

Table 6: The 3'-processing (3'-P) inhibitory activities of the most potent analogues in the presence of Mg^{2+}, of Mn^{2+}, along with the strand-transfer inhibitory activities of the most potent analogues (IC_{50} values in μM).

Compound	R^1	3′-P (Mg^{2+})[a]	3′-P (Mn^{2+})[a]	ST[a]
6		> 30	NT	5.2
24		> 30	> 30	5.5
25		> 30	NT	35
	R^1	3′-P (Mg^{2+})[a]	3′-P (Mn^{2+})[a]	ST[a]
55		> 30	NT	9
58		> 30	> 30	9
	R^3	3′-P (Mg^{2+})[a]	3′-P (Mn^{2+})[a]	ST[a]
94		> 30	NT	5
96		> 30	> 30	2.5

[a]IC_{50} determinations are the mean ±95% confidence interval (CI) of one experiment performed in triplicate. NT = not tested

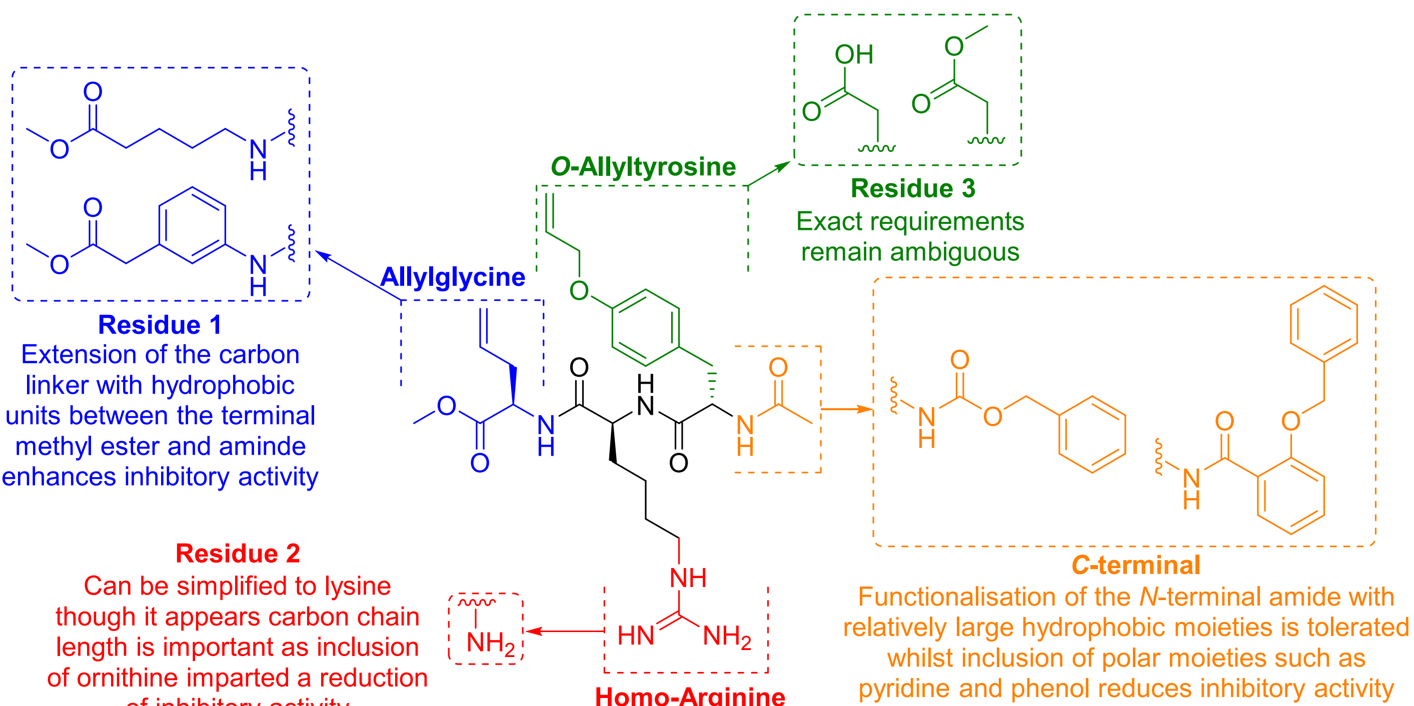


Fig 4: Summary of the key requirements within the allyl-tyrosine scaffold required to elicit HIV IN ST inhibitory activity.

Conclusions

From an in-house screening program an *O*-allyltyrosine-based tripeptide (**6**) was identified as an inhibitor with an IC_{50} value of 17.5 μM. As outlined in figure 4 subsequent SAR analysis identified a number of crucial features required for IN inhibitory activity. Specifically in relation to residue-1, extension of the carbon linker was advantageous for activity whilst the incorporation of amine rich

functionalities were detrimental. Functionalisation of the *C*-terminal amide moiety with relatively large hydrophobic moieties is tolerated whilst inclusion of polar pyridyl or phenolic within this region was detrimental. The homo-arginine residue could be simplified to lysine though it appears that the length of the carbon chain is important as inclusion of ornithine imparted a minimal reduction of inhibitory activity. Finally the role of the *O*-allyltyrosine residue is indeed significant as removal of the allyl moiety reduced inhibitory activity whilst the introduction of heteroatoms significantly reduced activity.

Consolidation of these optimum binding requirements afforded **96** which specifically inhibited ST with an IC_{50} value of 2.5 μM. Additionally it is proposed that this compound functions *via* an alternative mechanism to previously reported peptide based inhibitors and LEDGF/p75 allosteric inhibitors. Thus **96** provides a unique scaffold for further elaboration and current investigations include resolving a co-crystallised structure of **96** with the IN catalytic core in addition to molecular docking studies. The resulting advances in this class of HIV IN inhibitors and studies into the molecular mechanisms of activity will be reported in due course.

Experimental section

General Chemistry Procedures

Reagents and solvents were purchased reagent grade and used without further purification unless stated. CH_2Cl_2 was distilled from $CaCO_3$. Melting points (mp) were determined using a Gallenkamp (Griffin) melting point apparatus. Temperatures are uncorrected and expressed in degrees Celsius (°C). Optical rotations were measured using a Jasco polarimeter with a 10 mm path length.

Nuclear magnetic resonance (NMR) spectra were measured using a Varian Unity 300 MHz spectrometer. ^1H NMR spectra were acquired at 300.0 MHz whereas ^{13}C NMR spectra were acquired at 75.4 MHz. Spectra were recorded in deuterated chloroform ($CDCl_3$) containing 0.5% trimethylsilane TMS (δ 0.00 ppm), used as the internal standard, unless otherwise stated. Chemical shifts (δ) are expressed in ppm and coupling constants (*J*) are expressed in Hertz (Hz), both relative to the internal standard. Multiplicities are denoted generically as singlet (s), broad singlet (bs), doublet (d), doublet of doublets (dd), broad doublet (bd), doubleted triplet (dt), triplet (t), triplet of doublet of doublets (tdd), triplet of doublets (td), quartet (q) and multiplet (m). Each peak is listed according to the following convention: chemical shift, multiplicity, coupling constant, integration, assignment. Interchangeable peaks are denoted by letters in superscript.

Chemical ionization (CI) mass spectra (MS) were obtained on a Shimadzu QP-5000 MAT-44 quadrupole spectrometer. Electrospray (ESI) mass spectra were obtained on a VG Quattro-triple quadrupole. CI and ES were both performed via direct insertion with an electron beam of 70 eV at source temperatures < 200°C. The principal ion peaks *m/z* values are reported with their relative intensities in parentheses. ESI high resolution mass spectra (HRMS) were obtained using a Q-Tof mass spectrometer.

Thin layer chromatography (TLC) was performed using Merck Silica Gel F_{254} aluminium sheets. Column chromatography was performed using Merck silica gel 60 (70-230 mesh), under gravity, unless otherwise stated. All chromatographic solvent proportions are volume to volume. Solvents were evaporated by rotary evaporation *in vacuo*.

Procedure A: Allyl Ether Formation

The phenol derivative (1 eq.) and anhydrous potassium carbonate (K_2CO_3) (2 eq.) were combined and dried under vacuum for 1 hr. The vessel was then sealed and flushed with N_2, before anhydrous DMF (5 mL) was added. The mixture was allowed to stir at rt for 30 min before allyl bromide (2 eq.) was added, and the reaction stirred for 12 h. at rt. The reaction was then quenched with water (30 mL) and the solution was extracted with EtOAc (3 x 30 mL). The combined organic fractions were washed with water (5 x 50 mL), dried ($MgSO_4$) and the solvent evaporated to dryness under reduced pressure to yield the allyl ether product.

Procedure B: Methyl Ester Formation

To a stirred solution of the appropriate amino acid (1 eq.) in MeOH (10 mL) at 0 °C $SOCl_2$ (3 eq.) was slowly added. The solution was then removed from the ice bath and stirred at rt for 3 h. The reaction was then concentrated to dryness leaving the methyl ester amino acid as the hydrochloride salt.

Procedure C: Methyl /Ethyl Ester Hydrolysis

To a solution of the ester (1 eq.) in THF/H_2O (3:1, 60 mL) was added LiOH.H_2O (2 eq.) and the resulting suspension was allowed to stir for 12 h at rt, before being quenched with water (30 mL), and evaporated *in vacuo* to remove the THF. The resulting aqueous solution was extracted with CH_2Cl_2 (20 mL) to remove any unreacted materials. The aqueous phase was then acidified to pH 1 with a 2 M $NaHSO_4$ solution. The mixture was extracted with CH_2Cl_2 (3 x 50 mL) and the combined CH_2Cl_2 extracts were dried ($MgSO_4$) and the solvent removed under reduced pressure to yield the desired acid.

Procedure D: Amide Coupling

The acid (1 eq.), HOBt (1.1 eq), EDCI (1.1 eq.) and the amine hydrochloride (1.2 eq.) were placed in a flask, and then placed under high vacuum to dry. The vessel was then sealed and flushed with N_2. Anhydrous DMF (2 mL) and DIPEA (1 eq.) were added at rt and the solution was allowed to stir at rt for 12 h (in cases where the amine was present as the free base, DIPEA was not necessary and therefore excluded). The reaction was quenched with water until precipitation occurred (30 mL). The aqueous mixture was extracted with CH_2Cl_2 (3 x 50 mL) and the combined CH_2Cl_2 extracts were thoroughly washed with water (3 x 30 mL) dried ($MgSO_4$) and evaporated to dryness under reduced pressure to yield the desired amide.

Procedure E: *N*-Boc Deprotection

A solution of the N-Boc protected amine in CH_2Cl_2/TFA (1:1, 2 mL) was stirred at rt for 3 h. The solvent was removed under reduced pressure to yield the crude amine as the trifluoroacetate salt, which was either used as is or converted to the hydrochloride salt.

Procedure F: *N*-Fmoc Deprotection

A solution of the *N*-Fmoc protected amine in 1% piperidine/acetonitrile was stirred at rt for 3 h. The solvent was then removed under reduced pressure. The resulting crude product was purified by silica gel column chromatography using 15:1 CH_2Cl_2/MeOH as the eluting solvent. A 1% ninhydrin/ethanol detection solution was used to monitor the progress of elution of the desired amine by TLC analysis. The resulting product was either used as the free base or converted to the hydrochloride salt.

Procedure G: Guanidino Group Formation

The amine (1 eq.), as either the free base or the trifluoroacetate salt, was placed in a flask with $(BocNH)_2C{=}NSO_2CF_3$ (1 eq.) and dried under high vacuum. The flask was then sealed and flushed with N_2. Dry CH_2Cl_2 (2 mL) and triethylamine (NEt_3) (1.1 eq.) were added to the flask and the solution was allowed to stir at rt for 3 h. The solvent was then evaporated *in vacuo* and the resulting crude product was purified by silica gel column chromatography using 15:1 CH_2Cl_2/MeOH as the eluting solvent. The fractions were monitored by TLC analysis using uv light absorption (254 nm) for the detection of components, and those containing the desired compound were pooled and evaporated *in vacuo* to give the product which was used without further purification

Procedure H: Hydrochloride Salt Formation

The amine, as either the free base or trifluoroacetate salt was suspended in a minimum volume of MeOH. The solution was then treated with excess 1 M HCl/diethyl ether solution and concentrated *in vacuo*. The product was purified by precipitation from a MeOH solution by the addition of anhydrous diethyl ether.

General Assay Procedures

Assays were performed at Avexa Ltd. Initial Anti-HIV integrase inhibitory activity was determined using a combination 3'-processing and strand transfer *via* a microtitre plate assay, based on a

reported procedure[38] with some modifications. The oligonucleotide labelled with DIG had an additional GT on the 3' end (which is processed off in the 3'-processing portion of the assay) and the reaction buffer differed using 25 mM Tris-Cl at pH 7.5, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 25 mM NaCl, 50 μg/mL BSA, 5 mM β-mercaptoethanol, 30 nM substrate, and 10% DMSO. Assays were performed for 2 h at 37 °C. Reaction products bound to plates were detected using anti-DIG alkaline phosphatase Fab fragments (Roche) and 4-nitrophenol substrate. Colour was measured at 405 nm after 2 h. Positive control reactions typically absorbed at 405 nm of 1.2 to 1.8 with negatives values of 0.05 to 0.1.

Individual 3' processing assays used a gel based method as described in Ovenden *et al.*[39] using individually either Mg^{2+} or Mn^{2+}. 3'-Processing assays utilised the Chow *et al.*[44] procedure without modification.

Compound Characterisation

Nomenclature - New compounds were named according to the following order of precedence acid > ester > amide; due to the frequent use of several carbamate protecting groups in the synthesis, for simplicity, this functionality was excluded from the naming hierarchy. The aza/oxo substitution method was then used, where the longest chain of the highest priority was found and the remaining functional groups named as substituents of that chain.

Methyl (2*S*,5*S*,8*S*)-2-allyl-8-(4-allyloxybenzyl)-3,6,9-triaza-5-(4-guanidinobutyl)-4,7,10-trioxoundecanoate hydrochloride (6)

Compound **17** (50 mg, 0.07 mmol) was converted to the uncharacterised *N*-Boc deprotected trifluoroacetate salt *via* procedure E and the resulting solid was then converted immediately, *via* procedure H to the hydrochloride salt **6** (35 mg, 0.06 mmol, 86%) as a hygroscopic light brown amorphous solid. MS (ESI^+), *m/z* 545 (100%) [MH^+], 446 (30), 273 (20). HRMS (ESI^+) calcd for $C_{27}H_{40}N_6O_6$ + H: 545.3088; found 545.3085. $[\alpha]_D^{25}$ +62.9 (*c*. 0.12, EtOH). ^{1}H NMR (300 MHz, CD_3OD): δ 8.25 (d, *J* = 7.8 Hz, 1H, NH); 8.14 (d, *J* = 6.9 Hz, 1H, NH); 7.15 (d, *J* = 8.7 Hz, 2H, 2′-CH and 6′-CH); 7.08 (d, *J* = 7.8 Hz, 1H, NH); 6.96 (d, *J* = 7.7 Hz, 1H, NH); 6.83 (d, *J* = 8.7 Hz, 2H, 3′-CH and 5′-CH); 6.04 (tdd, *J* = 17.3, 10.4, 5.2 Hz, 1H, $OCH_2C\underline{H}=CH_2$); 5.78 (tdd, *J* = 17.1, 10.1, 6.9 Hz, 1H, C-2$CH_2C\underline{H}=CH_2$); 5.38 (dd, *J* = 17.3, 1.7 Hz, 1H, $OCH_2CH=C\underline{H}H$ *trans*); 5.23 (dd, *J* = 10.6, 1.6 Hz, 1H, $OCH_2CH=CH\underline{H}$ *cis*); 5.14 (dd, *J* = 17.3, 1.5 Hz, 1H, $CHCH_2CH=C\underline{H}H$ *trans*); 5.09 (dd, *J* = 10.5, 1.5 Hz, 1H, $CHCH_2CH=CH\underline{H}$ *cis*); 4.55-4.45 (m, 3H, $OC\underline{H}_2CH=CH_2$ and 8-CH); 4.42-4.38 (m, 2H, 2-CH and 5-CH); 3.70 (s, 3H, OCH_3); 3.16 (t, *J* = 6.9 Hz, 2H, 4″-CH_2); 3.03 (dd, *J* = 13.9, 5.8 Hz, 1H, 8-$CHC\underline{H}_aH_b$); 2.82 (dd, *J* = 13.9, 9.1 Hz, 1H, 8-$CHCH_a\underline{H}_b$); 2.55-2.47 (m, 2H, $CHC\underline{H}_2CH=CH_2$); 1.92 (s, 3H, 11-$CH_3$); 1.84-1.77 (m, 1H, 1″-$C\underline{H}_aH_b$); 1.70-1.64 (m, 1H, 1″-$CH_a\underline{H}_b$); 1.62-1.54 (m, 2H, 3″-$CH_2$); 1.45-1.42 (m, 2H, 2″-$CH_2$). ^{13}C NMR (75 MHz, CD_3OD): δ 172.7 (C-7); 172.6 (C-1); 172.3 (C-4); 172.2 (C-10); 157.8 (C-4′); 157.4 (C=N); 133.8 ($CHCH_2\underline{C}H=CH_2$); 133.0 ($OCH_2\underline{C}H=CH_2$); 130.0 (C-1′); 129.2 (C-2′ and C-6′); 117.7 ($OCH_2CH=\underline{C}H_2$); 116.2 ($CHCH_2CH=\underline{C}H_2$); 114.6 (C-3′ and C-5′); 68.6 ($O\underline{C}H_2CH=CH_2$); 55.5 (C-8); 52.7 (C-5); 52.5 (C-2); 51.5 (OCH_3); 41.1 (C-4″); 36.6 (8-$CH\underline{C}H_2$); 35.5 ($CH\underline{C}H_2CH=CH_2$); 31.5 (C-1″); 28.0 (C-3″); 22.4 (C-11); 21.2 (C-2″).

Methyl (7*S*,10*S*)-10-(4-allyloxybenzyl)-5,8,11-triaza-7-(4-guanidinobutyl)-6,9,12-trioxotridecanoate hydrochloride (24)

Compound **41** (99 mg, 0.14 mmol) was converted to the *N*-Boc deprotected trifluoroacetate salt *via* procedure E, the resulting solid was then converted, *via* procedure H to give hydrochloride salt **24** (64 mg, 0.11 mmol, 83%) as a hygroscopic light brown amorphous solid. MS (ESI^+), *m/z* 533 (100%) [MH^+], 534 (35). HRMS (ESI^+) calcd for $C_{26}H_{40}N_6O_6$ + H: 533.3088; found 533.3072. $[\alpha]_D^{25}$ +89.7 (*c*. 0.13, EtOH). ^{1}H NMR (300 MHz, CD_3OD): δ 7.70 (bs, 1H, NH); 7.62 (bs, 1H, NH); 7.17 (d, *J* = 8.4 Hz, 2H, 2′-CH and 6′-CH); 7.09 (d, *J* = 7.8, 1H, NH); 7.07 (bs, 1H, NH); 6.85 (d, *J* = 8.5 Hz, 2H, 3′-CH and 5′-CH); 6.12-5.94 (m, 1H, $OCH_2C\underline{H}=CH_2$); 5.38 (dd, *J* = 17.3, 1.6 Hz, 1H, $OCH_2CH=C\underline{H}H$ *trans*); 5.22 (dd, *J* = 10.6, 1.5 Hz, 1H, $OCH_2CH=CH\underline{H}$ *cis*); 4.55-4.45 (m, 3H, $OC\underline{H}_2CH=CH_2$ and 10-CH); 4.24 (dd, *J* = 9.4, 4.6 Hz, 1H, 7-CH); 3.64 (s, 3H, OCH_3); 3.20-3.11 (m, 4H, 4″-CH_2 and 4-CH_2); 3.03 (dd, *J* = 13.9, 6.4 Hz, 1H, 10-$CHC\underline{H}_aH_b$); 2.86 (dd, *J* = 13.7, 8.5 Hz, 1H, 10-$CHCH_a\underline{H}_b$); 2.33 (t, *J* = 7.4 Hz, 2H, 2-CH_2); 1.93 (s, 3H, 13-CH_3); 1.86-1.69 (m, 3H, 1″-

$C\underline{H}_aH_b$ and C-3H_2); 1.66-1.51 (m, 3H, 1″-$CH_a\underline{H}_b$ and 3″-CH_2); 1.47-1.25 (m, 2H, 2″-CH_2). ^{13}C NMR (75 MHz, CD_3OD): δ 174.1 (C-1); 172.9 (C-6); 172.8 (C-9 and C-12); 157.8 (C-4′); 157.4 (C=NH); 133.8 ($OCH_2\underline{C}H$=CH_2); 130.2 (C-2′ and C-6′); 129.2 (C-1′); 116.3 (OCH_2CH=$\underline{C}H_2$); 114.6 (C-3′ and C-5′); 68.7 ($O\underline{C}H_2CH$=CH_2), 55.8 (C-10); 53.5 (C-7); 51.1 (OCH_3); 41.2 (C-4″); 38.6 (C-4); 36.5 (10-$CH\underline{C}H_2$); 31.3 (C-1″); 30.9 (C-13); 28.1 (C-3″); 24.5 (C-3); 22.8 (C-13); 21.5 (C-2″).

_N_-{(1_S_,4_S_)-1-(4-Allyloxybenzyl)-8-amino-3-aza-4-[3-(methoxycarbonylmethyl)phenylcarbamoyl]-2-oxooctyl}acetamide hydrochloride (55)

Using procedure E, **73** (79 mg, 0.12 mmol) was deprotected to the *N*-Boc deprotected trifluoroacetate salt, and the resulting solid reacted *via* procedure H giving the hydrochloride salt **55** (46 mg, 0.08 mmol, 64%) as a hygroscopic brown amorphous solid. MS (ESI^+), *m/z* 539 (100%) [MH^+], 540 (33), 406 (70). HRMS (ESI^+) calcd for $C_{29}H_{38}N_4O_6$ + H: 539.2870; found 539.2876. $[\alpha]_D^{25}$ -25.8 (*c*. 0.14, EtOH). ^{1}H NMR (300 MHz, CD_3OD): δ 7.59-7.50 (m, 2H, 2″-CH and 6″-CH); 7.30-7.15 (m, 2H, 2′-CH and 6′-CH); 7.07-7.01 (m, 1H, 5″-CH); 6.92-6.82 (m, 2H, 3′-CH and 5′-CH); 6.80-6.71 (m, 1H, 4″-CH); 6.19-5.86 (m, 1H, $OCH_2C\underline{H}$=CH_2); 5.39 (bd, *J* = 17.5 Hz, 1H, OCH_2CH=$C\underline{H}H$ *trans*); 5.20 (bd, *J* = 10.0 Hz, 1H, OCH_2CH=$CH\underline{H}$ *cis*); 4.61-4.46 (m, 3H, $OC\underline{H}_2CH$=CH_2 and 1-CH); 4.38-4.27 (m, 1H, 4-CH); 3.67 (s, 3H, OCH_3); 3.37 (s, 2H, 3″-CCH_2); 3.12-2.81 (m, 4H, 8-CH_2 and 1-$CHC\underline{H}_2$); 2.03-1.89 (m, 4H, 5-$C\underline{H}_aH_b$ and $COCH_3$); 1.79-1.61 (m, 3H, 5-$CH_a\underline{H}_b$ and 7-CH_2); 1.56-1.40 (m, 2H, 6-CH_2). ^{13}C NMR (75 MHz, CD_3OD): δ 173.6 (4-$CH\underline{C}$=O); 173.3 (C-5); 172.9 (3″-$CCH_2\underline{C}$=O); 171.7 ($\underline{C}OCH_3$); 158.8 (C-4′); 139.4 (C-1″); 136.2 (C-3″); 134.9 ($OCH_2\underline{C}H$=CH_2); 131.4 (C-1′); 131.2 (C-2′ and C-6′); 129.9 (C-5″); 126.3 (C-4″); 122.1 (C-2″); 120.0 (C-6″); 117.3 (OCH_2CH=$\underline{C}H_2$); 115.6 (C-3′ and C-5′); 69.6 ($O\underline{C}H_2CH$=CH_2); 56.9 (C-1); 55.0 (C-4); 52.6 (OCH_3); 41.7 (3″-$C\underline{C}H_2$); 40.6 (C-8); 37.7 (1-$CH\underline{C}H_2$); 32.4 (C-5); 28.0 (C-7); 23.7 ($CO\underline{C}H_3$); 22.7 (C-6).

Methyl (8_S_,11_S_)-11-(4-allyloxybenzyl)-8-(4-aminobutyl)-6,9,12-triaza-7,10,13-trioxotetradecanoate hydrochloride (58)

Compound **76** (70 mg, 0.12 mmol) was converted to the *N*-Boc deprotected trifluoroacetate salt *via* procedure E, and the resulting solid was then converted, *via* procedure H, to give the hydrochloride salt **58** (53 mg, 0.10 mmol, 84%) as a hygroscopic brown amorphous solid. MS (ESI^+), *m/z* 505 (100%) [MH^+], 508 (80), 509 (23). HRMS (ESI^+) calcd for $C_{26}H_{40}N_4O_6$ + H: 505.3026; found 505.3035. $[\alpha]_D^{25}$ +53.5 (*c*. 0.22, EtOH). ^{1}H NMR (300 MHz, CD_3OD): δ 8.22 (bs, 1H, NH); 8.10 (bs, 1H, NH); 7.50 (bs, 1H, NH); 7.16 (d, *J* = 8.0 Hz, 2H, 2′-CH and 6′-CH); 6.86 (d, *J* = 7.9 Hz, 2H, 3′-CH and 5′-CH); 6.11-5.97 (m, 1H, $OCH_2C\underline{H}$=CH_2); 5.38 (bd, *J* = 17.3 Hz, 1H, OCH_2CH=$C\underline{H}H$ *trans*); 5.23 (bd, *J* = 10.5 Hz, 1H, OCH_2CH=$CH\underline{H}$ *cis*); 4.56-4.41 (m, 3H, $OC\underline{H}_2CH$=CH_2 and 11-CH); 4.30-4.21 (m, 1H, 8-CH); 3.63 (s, 3H, OCH_3); 3.15-3.06 (m, 2H, 5-CH_2); 3.04-2.81 (m, 4H, 4″-CH_2 and 11-$CHC\underline{H}_2$); 2.33 (t, *J* = 7.1 Hz, 2H, 2-CH_2); 1.93 (s, 3H, 14-CH_3); 1.78-1.25 (m, 10H, 1″-$C\underline{H}_aH_b$, 1″-$CH_a\underline{H}_b$, 3-CH_2, 4-CH_2, 2″-CH_2 and 3″-CH_2). ^{13}C NMR (126 MHz, CD_3OD): δ 175.2 (C-1); 173.7 (C-10); 173.4 (C-7); 173.3 (C-13); 158.5 (C-4′); 134.7 ($OCH_2\underline{C}H$=CH_2); 131.1 (C-2′ and C-6′); 129.9 (C-1′); 117.3 (OCH_2CH=$\underline{C}H_2$); 115.5 (C-3′ and C-5′); 69.8 ($O\underline{C}H_2CH$=CH_2); 56.6 (C-11); 54.2 (C-8); 52.3 (OCH_3); 40.9 (C-4″); 39.8 (C-5); 37.4 (11-$CH\underline{C}H_2$); 34.2 (C-2); 32.3 (C-2″); 29.5 (C-3″); 27.8 (C-4); 23.6 (C-14); 23.0 (C-2″); 22.8 (C-3).

Methyl (7_S_,10_S_)-10-(4-allyloxybenzyl)-7-(4-aminobutyl)-5,8,11-triaza-12-benzyloxy-6,9,12-trioxododecanoate hydrochloride (96)

Compound **114** (264 mg, 0.33 mmol) was converted to the uncharacterised *N*-Fmoc deprotected amine *via* procedure F, using 1% piperidine in 9:1 acetonitrile/DMF (10 mL). This was then converted, *via* procedure H to the hydrochloride salt **96** (142 mg, 0.23 mmol, 70%) as a hygroscopic brown amorphous solid. MS (ESI^+), *m/z* 583 (100%) [MH^+], 584 (35) [MD^+]. HRMS (ESI^+) calcd for $C_{31}H_{42}N_4O_7$ + H: 583.3132; found 583.3135. $[\alpha]_D^{25}$ +176.9 (*c*. 0.1, EtOH). ^{1}H NMR (300 MHz, CD_3OD): δ 8.08 (bs, 1H, NH); 7.56 (bs, 1H, NH); 7.34-7.25 (m, 5H, 2‴-CH, 3‴-CH, 4‴-CH, 5‴-CH, 6‴-CH,); 7.15 (d, *J* = 7.0 Hz, 2H, 2′-CH and 6′-CH); 6.84 (d, *J* = 7.0 Hz, 2H, 3′-CH and 5′-CH); 6.14-5.96 (m, 1H, $OCH_2C\underline{H}$=CH_2); 5.38 (bd, *J* = 17.2 Hz, 1H, OCH_2CH=$C\underline{H}H$ *trans*); 5.23 (bd, *J* = 10.2 Hz, 1H, OCH_2CH=$CH\underline{H}$ *cis*); 5.09-4.97 (m, 2H, 12-$COOCH_2$); 4.50 (d, *J* = 4.0 Hz, 2H, $OC\underline{H}_2CH$=CH_2); 4.34-4.24 (m, 2H, 7-CH, 10-CH); 3.64 (s, 3H, OCH_3); 3.23-3.11 (m, 2H, 4-CH_2); 3.04-2.95 (m, 2H, 4″-CH_2); 2.91-2.82 (m, 2H, 10-$CHCH_2$); 2.36-2.29 (m, 2H, 2-CH_2); 1.84-

1.60 (m, 6H, 1″-CH_2, 3-CH_2 and 3″-CH_2); 1.44-1.30 (m, 2H, 2″-CH_2). ^{13}C NMR (75 MHz, CD_3OD): δ 174.2 (C-1); 173.5 (C-9); 173.0 (C-6); 157.9 (C-4′); 157.3 (C-12); 137.1 (C-1‴); 134.3 ($OCH_2\underline{C}H=CH_2$); 130.8 (C-2′ and C-6′); 129.4 (C-1′); 129.0 (C-3‴ and C-5‴); 128.4 (C-4‴); 128.0 (C-2‴ and C-6‴); 117.3 ($OCH_2CH=\underline{C}H_2$); 115.2 (C-3′ and C-5′); 69.7 ($O\underline{C}H_2CH=CH$); 67.4 (12-$COO\underline{C}H_2$); 57.4 (C-10); 53.8 (C-7); 52.9 (OCH_3); 41.5 (C-4″); 39.5 (C-4); 37.4 (10-$CH\underline{C}H_2$); 32.1 (C-1″); 31.9 (C-2); 27.8 (C-3″); 25.3 (C-3); 23.6 (C-2″).

Acknowledgements

CPG is the recipient of an ARC DECRA fellowship and acknowledges the postgraduate scholarship provided the University of Wollongong and the ARC Linkage program. ND acknowledges the University of Wollongong and Avexa for a matching scholarship.

Notes and references

1. S. Srinivasa and K. Grinspoon Steven, *European journal of endocrinology / European Federation of Endocrine Societies*, 2014, **170**, R185-202.
2. J. M. A. Lange and J. Ananworanich, *Antiviral Therapy*, 2014, **19**, 5-14.
3. H. Samji, A. Cescon, R. S. Hogg, S. P. Modur, K. N. Althoff, K. Buchacz, A. N. Burchell, M. Cohen, K. A. Gebo, M. J. Gill, A. Justice, G. Kirk, M. B. Klein, P. T. Korthuis, J. Martin, S. Napravnik, S. B. Rourke, T. R. Sterling, M. J. Silverberg, S. Deeks, L. P. Jacobson, R. J. Bosch, M. M. Kitahata, J. J. Goedert, R. Moore and S. J. Gange, *PLoS One*, 2013, **8**, e81355/81351-e81355/81358, 81358 pp.
4. M. Cohen Stacy, M. Gray Kristen, M. C. B. Ocfemia, S. Johnson Anna and H. I. Hall, *Public health reports (Washington, D.C. : 1974)*, 2014, **129**, 335-341.
5. M. Wensing Annemarie, V. Calvez, F. Gunthard Huldrych, A. Johnson Victoria, R. Paredes, D. Pillay, W. Shafer Robert and D. Richman Douglas, *Topics in antiviral medicine*, 2014, **22**, 642-650.
6. M. Roche, H. Salimi, R. Duncan, B. L. Wilkinson, K. Chikere, M. S. Moore, N. E. Webb, H. Zappi, J. Sterjovski, J. K. Flynn, A. Ellett, L. R. Gray, B. Lee, B. Jubb, M. Westby, P. A. Ramsland, S. R. Lewin, R. J. Payne, M. J. Churchill and P. R. Gorry, *Retrovirology*, 2013, **10**, 43.
7. P. N. Dube, *Journal of Chemical, Biological and Physical Sciences*, 2014, **4**, 1152-1170, 1119 pp.
8. S. P. Lee, H. G. Kim, M. L. Censullo and M. K. Han, *Biochemistry*, 1995, **34**, 10205-10214.
9. E. Tramontano, P. La Colla and Y.-C. Cheng, *Biochemistry*, 1998, **37**, 7237-7243.
10. Y. Wang, H. Klock, H. Yin, K. Wolff, K. Bieza, K. Niswonger, J. Matzen, D. Gunderson, J. Hale, S. Lesley, K. Kuhen, J. Caldwell and A. Brinker, *Journal of Biomolecular Screening*, 2005, **10**, 456-462.
11. A. Mazumder, A. Engelman, R. Craigie, M. Fesen and Y. Pommier, *Nucleic Acids Research*, 1994, **22**, 1037-1043.
12. A. Mazumder, M. Gupta and Y. Pommier, *Nucleic Acids Research*, 1994, **22**, 4441-4448.
13. J. Snasel, D. Rejman, R. Liboska, Z. Tocik, T. Ruml, I. Rosenberg and I. Pichova, *European Journal of Biochemistry*, 2001, **268**, 980-986.
14. *Application: RU RU Pat.,* 2002-126506 2234535, 2004.
15. N. Neamati, C. Marchand and Y. Pommier, *Advances in Pharmacology (San Diego, CA, United States)*, 2000, **49**, 147-165, 141 plate.
16. C. Marchand, X. Zhang, G. C. G. Pais, K. Cowansage, N. Neamati, T. R. Burke, Jr. and Y. Pommier, *Journal of Biological Chemistry*, 2002, **277**, 12596-12603.
17. A. S. Espeseth, P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J. Y. Melamed, S. Young, T. Hamill, J. L. Cole and D. J. Hazuda, *Proceedings of the National Academy of Sciences of the United States of America*, 2000, **97**, 11244-11249.
18. G. Chi, N. Neamati and V. Nair, *Bioorganic & Medicinal Chemistry Letters*, 2004, **14**, 4815-4817.
19. M. L. Barreca, L. De Luca, N. Iraci and A. Chimirri, *Journal of Medicinal Chemistry*, 2006, **49**, 3994-3997.
20. C. P. Gordon, R. Griffith and P. A. Keller, *Medicinal Chemistry*, 2007, **3**, 199-220.
21. R. Di Santo, *Journal of Medicinal Chemistry*, 2014, **57**, 539-566.

22. K. Gupta, T. Brady, M. Dyer Benjamin, N. Malani, Y. Hwang, F. Male, T. Nolte Robert, L. Wang, E. Velthuisen, J. Jeffrey, D. Van Duyne Gregory and D. Bushman Frederic, *The Journal of biological chemistry*, 2014, **289**, 20477-20488.

23. T. S. Peat, D. I. Rhodes, N. Vandegraaff, G. Le, J. A. Smith, L. J. Clark, E. D. Jones, J. A. V. Coates, N. Thienthong, J. Newman, O. Dolezal, R. Mulder, J. H. Ryan, G. P. Savage, C. L. Francis and J. J. Deadman, *PLoS One*, 2012, **7**, e40147.

24. D. I. Rhodes, T. S. Peat, N. Vandegraaff, D. Jeevarajah, J. Newman, J. Martyn, J. A. V. Coates, N. J. Ede, P. Rea and J. J. Deadman, *ChemBioChem*, 2011, **12**, 2311-2315.

25. Z. Hayouka, M. Hurevich, A. Levin, H. Benyamini, A. Iosub, M. Maes, D. E. Shalev, A. Loyter, C. Gilon and A. Friedler, *Bioorganic & Medicinal Chemistry*, 2010, **18**, 8388-8395.

26. T. W. Sanchez, B. Debnath, F. Christ, H. Otake, Z. Debyser and N. Neamati, *Bioorganic & Medicinal Chemistry*, 2013, **21**, 957-963.

27. T. B. Dupree, P. A. Keller and R. Griffith, *Australian Journal of Chemistry*, 2011, **64**, 916-918.

28. P. A. Keller, C. Birch, S. P. Leach, D. Tyssen and R. Griffith, *Journal of Molecular Graphics & Modelling*, 2003, **21**, 365-373.

29. S. J. Titmuss, P. A. Keller and R. Griffith, *Bioorganic & Medicinal Chemistry*, 1999, **7**, 1163-1170.

30. T. P. Boyle, J. B. Bremner, J. Coates, J. Deadman, P. A. Keller, S. G. Pyne and D. I. Rhodes, *Tetrahedron*, 2008, **64**, 11270-11290.

31. T. P. Boyle, J. B. Bremner, J. A. Coates, J. Deadman, P. A. Keller, S. G. Pyne and K. Somphol, *European Journal of Medicinal Chemistry*, 2009, **44**, 1001-1009.

32. J. B. Bremner, P. A. Keller, S. G. Pyne, T. P. Boyle, Z. Brkic, J. Morgan, K. Somphol, J. A. Coates, J. Deadman and D. I. Rhodes, *Bioorganic & Medicinal Chemistry*, 2010, **18**, 4793-4800.

33. Z. Hayouka, A. Levin, M. Hurevich, D. E. Shalev, A. Loyter, C. Gilon and A. Friedler, *Bioorganic & Medicinal Chemistry*, 2012, **20**, 3317-3322.

34. Y.-Q. Long, S.-X. Huang, Z. Zawahir, Z.-L. Xu, H. Li, T. W. Sanchez, Y. Zhi, S. De Houwer, F. Christ, Z. Debyser and N. Neamati, *Journal of Medicinal Chemistry*, 2013, **56**, 5601-5612.

35. W. Nomura, H. Aikawa, N. Ohashi, E. Urano, M. Metifiot, M. Fujino, K. Maddali, T. Ozaki, A. Nozue, T. Narumi, C. Hashimoto, T. Tanaka, Y. Pommier, N. Yamamoto, J. A. Komano, T. Murakami and H. Tamamura, *ACS Chemical Biology*, 2013, **8**, 2235-2244.

36. S. Suzuki, E. Urano, C. Hashimoto, H. Tsutsumi, T. Nakahara, T. Tanaka, Y. Nakanishi, K. Maddali, Y. Han, M. Hamatake, K. Miyauchi, Y. Pommier, J. A. Beutler, W. Sugiura, H. Fuji, T. Hoshino, K. Itotani, W. Nomura, T. Narumi, N. Yamamoto, J. A. Komano and H. Tamamura, *Journal of Medicinal Chemistry*, 2010, **53**, 5356-5360.

37. M. Maes, A. Levin, Z. Hayouka, D. E. Shalev, A. Loyter and A. Friedler, *Bioorganic & Medicinal Chemistry*, 2009, **17**, 7635-7642.

38. Y. Hwang, D. Rhodes and F. Bushman, *Nucleic Acids Research*, 2000, **28**, 4884-4892.

39. S. P. B. Ovenden, J. Yu, S. San Wan, G. Sberna, R. Murray Tait, D. Rhodes, S. Cox, J. Coates, N. G. Walsh and B. M. Meurer-Grimes, *Phytochemistry (Elsevier)*, 2004, **65**, 3255-3259.

40. I. K. Pemberton, M. Buckle and H. Buc, *Journal of Biological Chemistry*, 1996, **271**, 1498-1506.

41. J. A. Grobler, K. Stillmock, B. Hu, M. Witmer, P. Felock, A. S. Espeseth, A. Wolfe, M. Egbertson, M. Bourgeois, J. Melamed, J. S. Wai, S. Young, J. Vacca and D. J. Hazuda, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 6661-6666.

42. C. Marchand, A. A. Johnson, R. G. Karki, G. C. G. Pais, X. Zhang, K. Cowansage, T. A. Patel, M. C. Nicklaus, T. R. Burke, Jr. and Y. Pommier, *Molecular Pharmacology*, 2003, **64**, 600-609.

43. I. J. Chen, N. Neamati and A. D. MacKerell, Jr., *Current Drug Targets: Infectious Disorders*, 2002, **2**, 217-234.

44. Y. Shibagaki, M. L. Holmes, R. S. Appa and S. A. Chow, *Virology*, 1997, **230**, 1-10.