

Accepted Manuscript



A structure-based design approach to advance the allyltyrosine-based series of HIV integrase inhibitors

Christopher P. Gordon, Neal Dalton, Nicholas Vandegraaff, John Deadman, David I. Rhodes, Jonathan A. Coates, Stephen G. Pyne, Renate Griffith, John B. Bremner, Paul A. Keller

PII: S0040-4020(17)31211-5

DOI: [10.1016/j.tet.2017.11.052](https://doi.org/10.1016/j.tet.2017.11.052)

Reference: TET 29132

To appear in: *Tetrahedron*

Received Date: 18 September 2017

Revised Date: 16 November 2017

Accepted Date: 17 November 2017

Please cite this article as: Gordon CP, Dalton N, Vandegraaff N, Deadman J, Rhodes DI, Coates JA, Pyne SG, Griffith R, Bremner JB, Keller PA, A structure-based design approach to advance the allyltyrosine-based series of HIV integrase inhibitors, *Tetrahedron* (2017), doi: 10.1016/j.tet.2017.11.052.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Graphical Abstract

To create your abstract, type over the instructions in the template box below.
 Fonts or abstract dimensions should not be changed or altered.

A structure-based design approach to advance the allyltyrosine-based series of HIV integrase inhibitors

Leave this area blank for abstract info.

Christopher P. Gordon,*^{a,b,†} Neal Dalton,^a Nicholas Vandegraaff,^{c,d} John Deadman,^{c,e} David I. Rhodes,^{c,f} Jonathan A. Coates,^c Stephen G. Pyne,^a Renate Griffith,^f John B. Bremner,^a and Paul A. Keller,^a

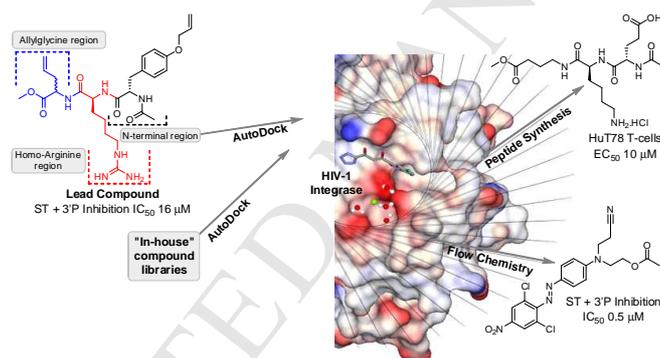
^aSchool of Chemistry, University of Wollongong, Northfields Avenue, Wollongong NSW 2522 Australia.

^{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

^cAvexa Ltd., 576 Swan St, Richmond, Vic 3121, Australia.

^dPresent Address: Genera Biosystems Ltd., Scoresby, Vic 3179, Australia.

^eChemocopeia Pty Ltd, 157 Arnold Street, Melbourne, Vic. 3054, Australia





A structure-based design approach to advance the allyltyrosine-based series of HIV integrase inhibitors

Christopher P. Gordon,^{*a,b†} Neal Dalton,^a Nicholas Vandegraaff,^{c,d} John Deadman,^{c,e} David I. Rhodes,^{c,f} Jonathan A. Coates,^c Stephen G. Pyne,^a Renate Griffith,^f John B. Bremner,^a and Paul A. Keller,^a

^aSchool of Chemistry, University of Wollongong, Northfields Avenue, Wollongong NSW 2522 Australia.

^{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

^cAvexa Ltd., 576 Swan St, Richmond, Vic 3121, Australia.

^dPresent Address: Genera Biosystems Ltd., Scoresby, Vic 3179, Australia.

^eChemocopeia Pty Ltd, 157 Arnold Street, Melbourne, Vic. 3054, Australia

^fPresent Address: Monash Institute of Materials Engineering, Monash University, Clayton, Vic 3800, Australia

[†]School of Medical Sciences, UNSW Australia, Sydney, NSW 2052, Australia

ARTICLE INFO

Article history:

Received

Received in revised form

Accepted

Available online

Keywords:

HIV-1 integrase

Strand-transfer inhibitors

Molecular docking

Flow chemistry

ABSTRACT

As of mid-2017, only one structure of the human immunodeficiency virus (HIV) integrase core domain co-crystallised with an active site inhibitor was reported. In this structure (1QS4), integrase is complexed with a diketo-acid based strand-transfer inhibitor (INSTI). This structure has been a preferred platform for the structure-based design of INSTIs despite concerns relating to structural irregularities arising from crystallographic packing effects. A survey of the current pool of 297 reported integrase catalytic core structures indicated that the anatomy of the active site in the complex structure 1QS4 exhibits subtle variations relative to all other structures examined. Consequently, the 1QS4 structure was employed for docking studies. From the docking of twenty-seven allyltyrosine analogues, a 3-point inhibitor binding motif required for activity was established and successfully utilised in the development of a tripeptide displaying an EC₅₀ value of 10 ± 5 μM in HIV infected human T-cells. Additional docking of “in-house” compound libraries unearthed a methyl ester based nitrile derivative displaying an IC₅₀ value of 0.5 μM in a combined 3'-processing and strand-transfer assay.

2009 Elsevier Ltd. All rights reserved.

1. Introduction

Since the advent of highly active antiretroviral therapy (HAART) in 1996, the prognosis and life expectancy for those infected with HIV has improved dramatically. Currently, a 20-year-old HIV-positive patient in the U.S. or Canada diagnosed at an early stage of infection and prescribed a current HAART regime is expected to live into their early 70's.¹ Traditionally, HAART regimes comprise two nucleoside reverse transcriptase inhibitors in addition to either a non-nucleoside inhibitor such as efavirenz, or a protease inhibitor.^{2,3} However, development of cross-resistance and poor tolerability necessitated the development of new chemotherapeutics targeting alternative components of the viral machinery including the HIV integrase (IN) enzyme.⁴

The IN enzyme mediates integration of the viral genome into the chromatin of the host's T-helper cells, an event representing a point of no return with the host cell becoming a permanent carrier of the viral cDNA.^{5,6} Integration occurs across a multi-step sequence which is initiated in the host cell cytosol with

3'-processing (3'P) whereby IN cleaves a dinucleotide from each viral DNA terminus at a conserved CA sequence, yielding two reactive 3' hydroxyl groups.^{6,7} Following this processing step, IN associates with a number of viral and cellular proteins, including Lens Epithelium Derived Growth Factor (LEDGF/p75), to form a pre-integration complex (PIC) which subsequently migrates to the nucleus.⁷ Within the nucleus, IN catalyses nucleophilic attack of the reactive hydroxyl groups upon the host chromosomal DNA in a process known as strand transfer (ST).⁷ In addition to IN having an indispensable role in the life cycle of HIV, there is no known counterpart in mammalian cells which renders the enzyme a highly attractive target for chemotherapeutic development.

Initial studies with IN inhibitors focused on the development of compounds capable of coordinating the divalent metal ions within the active site.⁶ Predominately, these analogues were constructed around a diketo-acid (DKA) moiety, or a bioisostere thereof, and specifically inhibited the ST reaction (INSTIs). Optimisation of this class culminated in the U.S. Food and Drug Administration (FDA) approved INSTIs raltegravir (RAL), elvitegravir (EVG, Fig. 1, compound 1), and more recently

dolutegravir (DTG, Fig. 1 compound 2). To complement this arsenal of INSTIs is an emerging class of allosteric inhibitors often referred to as LEDGINs were developed. These compounds bind within the LEDGF/p75 binding pocket and exhibit a multimodal mechanism of action.⁸⁻¹² In addition to impeding LEDGF/p57-IN binding interactions the LEDGINs induce aberrant IN multimerisation, and while only modestly interfering with early steps of HIV replication, they potently disrupt late steps including particle assembly and maturation.⁸⁻¹² A member of this class (BI-224436, Fig. 1, compound 3) was progressed to Phase I clinical development. However, this trial was ceased presumably due to enterohepatic recirculation-related PK issues.¹⁰

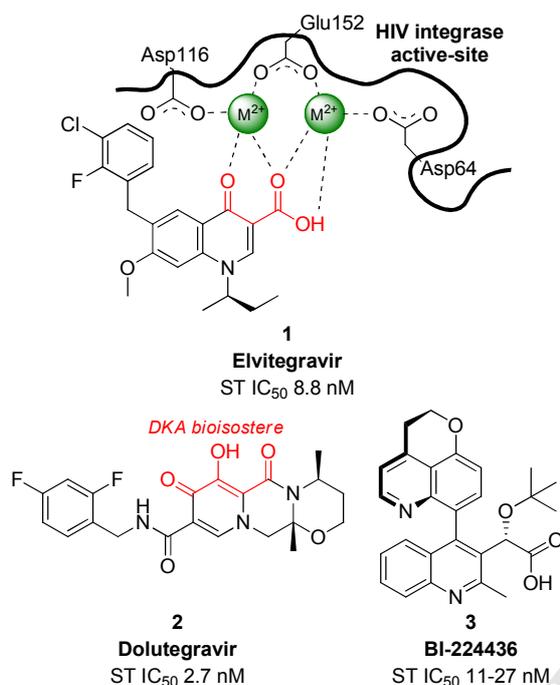


Fig. 1: Structure of elvitegravir (EVG, 1) with schematic outlining the diketoacid-Mg²⁺ interactions along with the chemical structures of dolutegravir (DTG, 2) and the BI-224436 (3) which progressed to phase I clinical trials.

While IN inhibitors are a relatively new addition to the clinician's palette of drugs to formulate HAART regimes, resistance to both INSTIs and LEDGINs has been observed. For example, even within Australia where it is estimated that 25,313 people are currently living with HIV¹³ and the infection rate has remained at a relatively stable rate of \approx 1,000 cases per year, resistance to IN inhibitors has been observed in a HAART naïve patient.¹⁴ Four point mutations conferring resistance to DTG have been characterised since the drug received FDA approval in mid-2013,¹⁵ and various mutations conferring resistance to LEDGINs have also been mapped. For example, the inhibitory activity of BI-224436 is reduced by 2.6- and 64-fold against the A128T and A128N mutants respectively relative to the wild-type virus.¹⁶ Consequently, if IN inhibitors are to remain a vital component of HAART regimes, it is essential that new generations of inhibitors are continually entering the drug development pipeline.

To this end, we recently identified a series of allyltyrosine based tri-peptides which displayed specific ST inhibitory activity.¹⁷ In an initial attempt to enhance potency, the scaffold was subjected to a comprehensive structure-activity-relationship (SAR) campaign, and while analogues with incrementally enhanced IC₅₀ values were developed (i.e. compound 5, Fig. 2), significant activity enhancements were not forthcoming. Hence,

as a consequence of the recent plethora of HIV IN crystal structures deposited in the protein databank, we were moved to explore the possibility of enhancing the inhibitory activity of the scaffold through a structure-based approach.

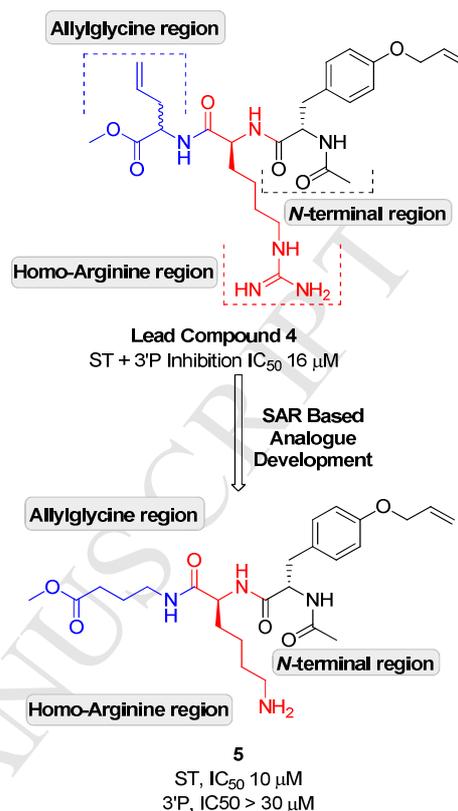


Fig. 2: *O*-Allyl-L-tyrosine-based tripeptide HIV-1 integrase inhibitor (4) discovered from a screening program of a number of 'in-house' compound libraries and the structure of the INSTI 5 which evolved from SAR-based design.¹⁷

IN is comprised of three independently folded domains.^{12,18-22} The *N*-terminal domain (NTD; residues 1–50) binds Zn²⁺ via a conserved His-His-Cys-Cys motif. The catalytic core domain (CCD) (residues 50–212) adopts an RNase H superfamily fold and contains a DDE motif that binds Mg²⁺ or Mn²⁺ ions and mediates DNA cleaving and joining. The *C*-terminal domain (CTD; residues 223–268) features a Src homology domain 3 (SH3)-like fold that contributes to DNA binding and is connected to the catalytic core domain by a α -helical linker (residues 213–222). All three domains are required for activity, and each contributes to the correct assembly of a catalytically active IN tetramer.^{12,18-22} However as the CCD contains the crucial determinants required for catalytic activity including the DDE catalytic triad, which mediates both 3'P and ST, we elected to focus on the active site of the CCD for the structure-based design development of the allyltyrosine analogues. The results of this work and some biological evaluations are now reported in this paper.

2. Results and Discussion

2.1. Development of an integrase catalytic core domain model

As of mid-2017 a total of 297 individual structures of the IN CCD in monomeric, dimeric, or trimeric complexes, were deposited in the protein databank. Of these, 56^{8,19-21,23-30} were crystallised in the absence of an inhibitor while 240^{8-10,12,18,20,21,25-29,31-50} were co-crystallized with either a LEDGIN or a fragment,

bound within in LEDGF/p57 binding pocket. However, with the exception of recently reported structures of the orthologous retroviral IN from the prototype foamy virus (PFV) co-crystallised with a series 1,8-naphthyridine-3-carboxamides INSTIs,⁵¹ to date only one structure of HIV IN (1QS4)³⁰ has been reported with an inhibitor co-crystallized in the active site. The inhibitor, known as 5CITEP is a member of the diketo-acid family and is a specific inhibitor of strand-transfer (IC_{50} 2.1 ± 0.1 μ M). Consequently, 1QS4³⁰ has been favoured for docking of INSTIs.⁵²⁻⁵⁷ Other structures have also been utilised with 2B4J⁸ typically favoured for the docking of LEDGINS⁵⁸⁻⁶² as this structure is co-crystallized with the IN binding domain of LEDGF, whilst high-resolution structures such as 3L3U²⁶ and 1BL3²⁷ have also been employed.^{34,63,64}

In this study, we utilised 1QS4 as a platform for model development. The A chain of 1QS4, in which the inhibitor 5CITEP is bound, exists as a dimer and is thus representative of a catalytically active complex. The structure also possesses a divalent Mg^{2+} ion which is essential for catalytic activity. Moreover, the structure conforms with critical structure evaluation parameters. For example 90.9 % of backbone Φ and Ψ angles fall into the most favoured regions of the Ramachandran plot while the R factor and R_{free} values of 0.209 and 0.255, respectively, are within the ideal 25% threshold (supplementary S.2).⁶⁵

It should be noted that 1QS4 is a less than optimal structure for docking as four amino acid residues within the region known as the “140s-loop”, which is critical to catalytic function,¹⁹ are not resolved. Furthermore, it has been proposed that the A-chain binding site environment is unlikely to resemble a physiologically relevant situation.⁶⁶ Structural irregularities may arise from crystallographic packing effects as the active site of the A chain is in close contact with the active site of an adjacent IN molecule in a neighbouring unit cell.⁶⁷ This arrangement of the core domains creates a larger binding cavity allowing the ligand to interact not only with its primary “receptor” but also with the symmetry-related protein and its bound ligand. Hence these additional contacts may perturb the resolved orientation of the ligand.^{66,67}

Thus in light of the above, we surveyed the current pool of 297 published IN CCD structures to ascertain whether the active site of the 1QS4 A-chain contained significant discrepancies relative to other structures (supplementary S.1). We employed a manual approach measuring inter-residue distances between key active site-residues. As illustrated in Fig. 3a five distances (defined as D1 – D5) (Fig. 3a) were measured within each crystal structure. D1 was proposed to provide a qualitative measure of active site depth, D2 a measure of active site width, D3 and D4 would provide an indication of the shape of the “phosphate-binding pocket” (predominantly formed by His-67, Lys-156, Lys 159), whilst D5 would provide a measure of active site length. Further, the angle formed by the terminal side-chain carbons of Glu-152, Lys-156, and Lys-159 was also measured to provide further insight to the shape of the “phosphate-binding pocket”. As illustrated in figure 3b the D1 – D5 measures and the phosphate-binding pocket angles were clustered and analysed in eight groups. Group 1 incorporated non-inhibited structures crystallised as a monomeric unit, group 2 combined CCDs crystallised within dimeric or trimeric complexes whilst group 3 contained monomeric units crystallised with an inhibitor. The 4th and 5th groups combined respective A and B chains from dimeric structures which were co-crystallised with LEDGINS whereas group 6 incorporated CCDs from dimeric complexes in which an inhibitor was resolved only in the opposing CCD (i.e. the B

chains of 3VQ6⁴⁴ or 4NYF¹⁰). Finally, group 7 was comprised of the B and C chain of 1QS4 whilst group 8 was the A chain of 1QS4.

From the analysis of the groups it was found that the D5 measure remained remarkably constant across all 297 structures (≈ 14 Å) whilst D1 and D4 displayed negligible variations with average distances of ≈ 5 Å and ≈ 10 Å, respectively. The D2 (width of active site) was notably larger in both the A and B chains of the LEDGIN bound complexes, and the A-chain of 1QS4, suggesting a possible inhibited conformation were Glu-152 is separated from the other residues of the catalytic triad (Fig. 3c – d). However, the most significant variations within the 1QS4 A-chain active site were the distance between Glu-152 and Lys156 side-chains (D3) and between Lys156 and Lys159 side-chains (D4), which were approximately 4 Å and 2 Å less than average, respectively. Also, the phosphate-binding pocket angle within the 1QS4 A-chain was approximately 30° higher than average. Regarding structural effects, these variations rendered the active site of the 1QS4 A-chain and the active sites of the allosterically inhibited CCDs considerably wider than observed in non-inhibited structures (Fig. 3d – e). A further unique feature of the 1QS4 A-chain is that the “phosphate binding pocket” is considerably more enclosed indicating that binding of 5CITEP within the active site (Fig. 3f) does induce a unique “inhibited conformation”, albeit subtly varied relative to other structures. Consequently, the A-chain of 1QS4 was considered to be a suitable platform for docking runs.

2.2. Validation studies of the 1QS4 A-chain model

Prior to subjecting the A-chain of 1QS4 to docking runs, the missing residues of the “141s-loop” (e.g. Ile-141, Pro-142, Tyr-143 and Asn-144; *see protocol 4.3.2*; supplementary, S.3) were ‘spliced in’ to investigate the proposed biologically relevant two Mg^{2+} model (DDE + 2Mg).^{57,68} Five separate receptor models were created (I-V, Table 1) with IV and V constructed with two Mg^{2+} ions to replicate models utilised by Perryman *et al.*⁵⁷ and Diamond *et al.*⁶⁸ To evaluate the validity of these receptor models each was subjected to docking runs (AutoDock-3⁶⁹) using 5CITEP as a ligand (*protocol 4.3.3*). As outlined in Table 1, models III and IV were the best performed best, affording final root means squared differences (RMSD) of less than 1 Å relative to the crystallographically resolved conformation of 5CITEP. However, of these two models, the single hydrated receptor was deemed the most relevant as a Mg^{2+} ion would not exist in a non-solvated form in a physiological environment. Thus, the single hydrated Mg^{2+} ion construct was employed for all subsequent docking runs.

2.3. Docking studies of the first generation allyltyrosine derivatives

With a validated model established, attention turned to the allyltyrosine scaffold¹⁷ with an initial aim of establishing an archetypal binding motif required for activity. Therefore, 27 analogues (supplementary S.5), encompassing variations within the three altered quadrants of the scaffold (i.e. Fig. 2), were docked (*protocol 4.3.3*). Of these, 14 contained variations within the allylglycine region (IC_{50} values 4.5 - > 100 μ M), three contained variations in the homo-arginine region (IC_{50} 10 – 60 μ M) whilst the remaining ten derivatives contained *N*-terminal appendages. From these docking runs (supplementary S.6), a consistent 3-point binding motif was established where the terminal methyl ester moiety resided within the “phosphate-binding pocket” (i.e. His-67, Lys-156 and Lys-159), and the terminal amine of the homo-arginine region formed electrostatic interactions with Asp-64. Additionally, the allyltyrosine

functionality occupied the hydrophobic zone of the active site region of the scaffold was required to adopt a strained and compacted conformation to establish interactions within the (Figure 4a).

From a design perspective, the docking runs indicated that the chain bearing the methylester moiety within the allylglycine

ACCEPTED MANUSCRIPT

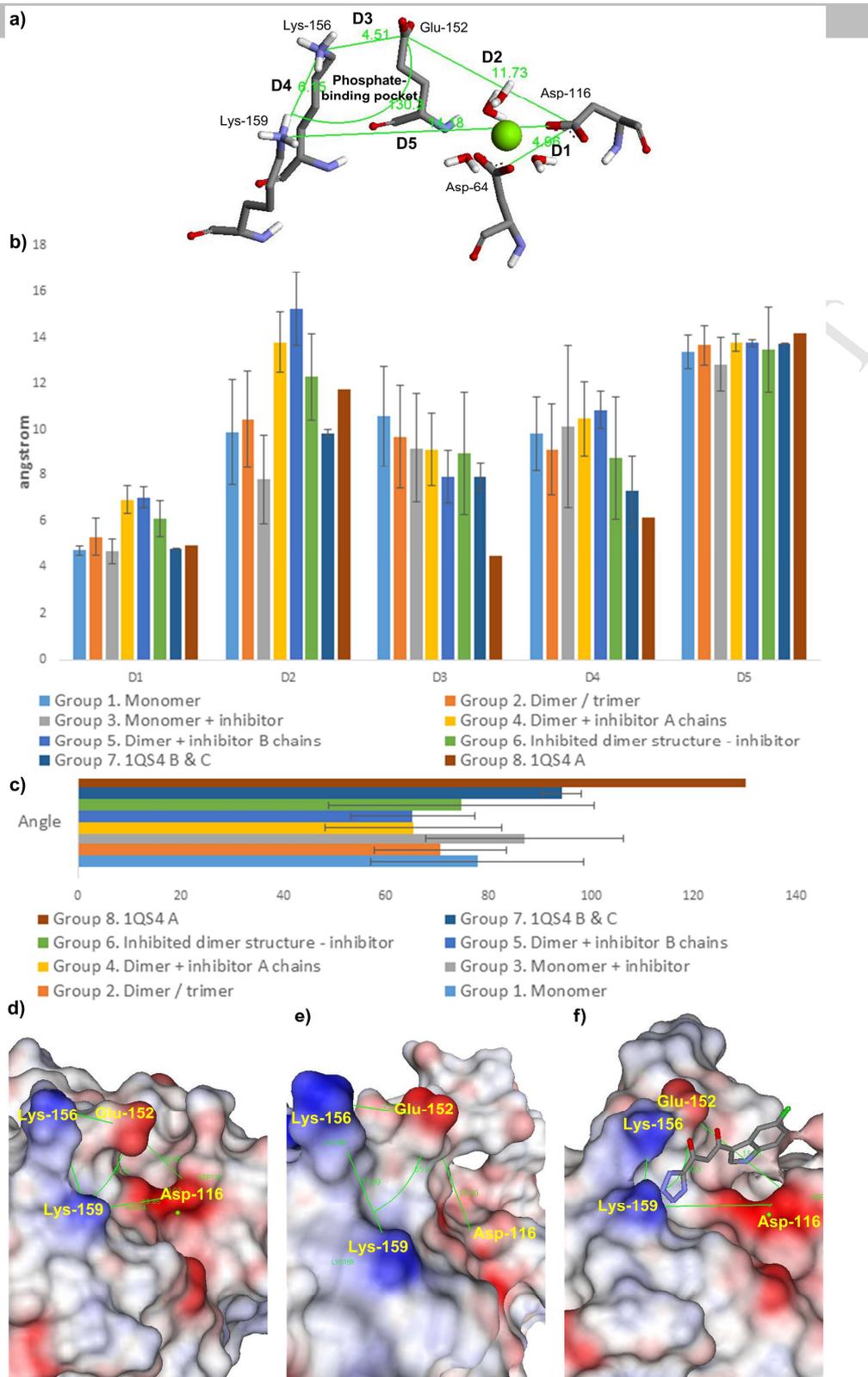


Figure 3: **a)** Representation of the five inter-residue distances measured in Å to investigate specific structural differences within the active sites of all 297 reported CCDs. The measurements displayed are from the A-chain of 1QS4. For simplicity, side-chain hydrogens and 5CITEP have been removed. The angle formed by the terminal side-chain carbons of Glu-152, Lys-156, and Lys-159 (the phosphate binding-pocket angle) is represented as a green arc. D1 = distance between C χ of Asp-64 and C χ of Asp-116. D2 = distance between C χ of Asp-116 and C δ of Glu-152. D3 = distance between C δ of Glu-152 and C ϵ of Lys-156. D4 = distance between C ϵ of Lys-156 and C ϵ of Lys-159. D5 = distance between C ϵ of Lys-159 and C χ of Asp-116. **b)** Graphical representation of the inter-residue distances, D1 through D5 (error bars represent standard deviation). **c)** Measurement of the angle formed by the terminal side-chain carbons of Glu-152, Lys-156, and Lys-159 across all 297 CCDs (error bars represent standard deviation); **d)** The non-inhibited crystal structure 1BL3 which is representative of the majority of reported HIV-1 IN crystal structures showing solvent accessible surface determined with a probe radius of 1.4 Å.; **e)** The non-inhibited crystal structure 3ZT2 which is representative of the majority of reported HIV-1 IN structures co-crystallised with LEDGINS; **f)** 1QS4 A-chain co-crystallised with 5CITEP.

Table 1: Receptor models generated for validation docking runs using 5CITEP with final docked energies and RMSD relative to the crystallographically resolved conformation of 5CITEP.

Model	Receptor Setup	Docked Energy (Kcal/mol)	RMSD (Å)
I	No Mg ²⁺ ion present	-6.96	3.61
II	Single non-hydrated Mg ²⁺ ion	-7.05	4.32
III	Single hydrated Mg ²⁺ ion	-6.35	0.97
IV	One hydrated Mg ²⁺ and one non-hydrated Mg ²⁺ ion	-6.32	0.99
V	Two hydrated crystallographically resolved and modelled Mg ²⁺ ions	-6.96	8.02

phosphate binding pocket (supplementary, S.6). As the initial crystal structure survey indicated that the length of the active site remained at a relatively consistent 14 Å across all structures, it was proposed that truncation along with ester hydrolysis could prove advantageous. Consequently, compound **7** (Figure 5) was highlighted for synthesis. Additionally, the docking studies suggested that the phosphate-binding pocket may potentially accommodate an additional ester/coxylate moiety and, as this inference was supported by docking runs of analogues **8** – **10** (supplementary S.6), they were also marked for synthesis (Figure 5).

Whilst the primary object of this study was the advancement of the allyltyrosine series, given our ongoing interest in the discovery of IN inhibitors we concurrently performed docking of “in-house” compound libraries. From these studies, a series of methyl ester based nitrile analogues which formed interactions with the three-point binding motif (i.e. Fig. 4a) were identified. As outlined in figure 4b the nitrile group established electrostatic interactions with Asp-64 and Asp-116, the methyl ester moiety occupied the phosphate-binding pocket while the nitrobenzene head resided within the hydrophobic zone of the active site. In light of these docking results, four analogues (compounds **6**, **11** – **13**, Fig. 5) were subjected to biological evaluation.

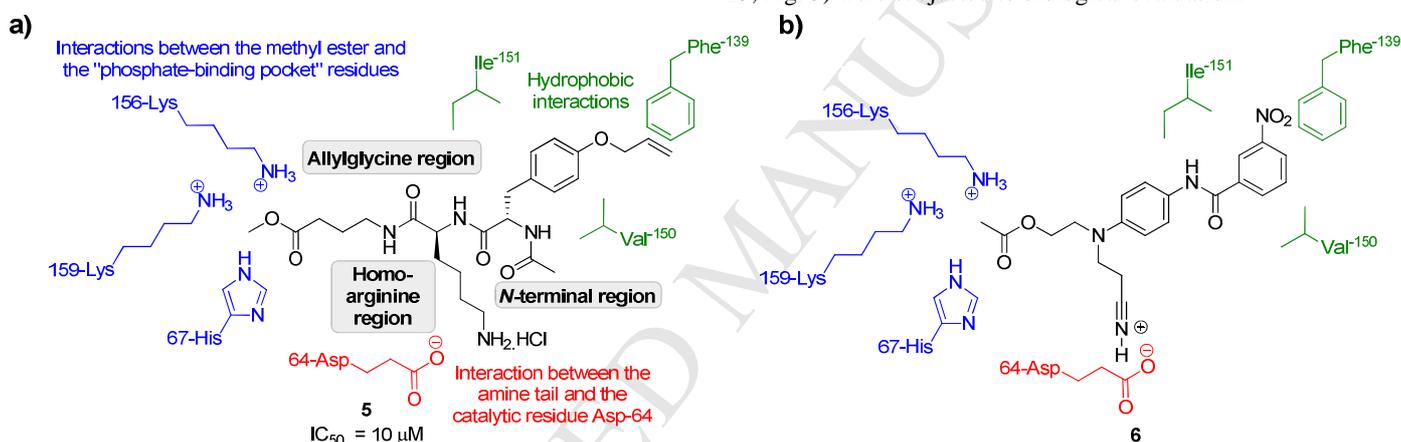


Fig. 4: a) Proposed 3-point binding motif by which the allyltyrosine lead compound **5** binds within the IN active site; b) schematic representation of the docked conformation obtained for the nitrile-based compound **6**.

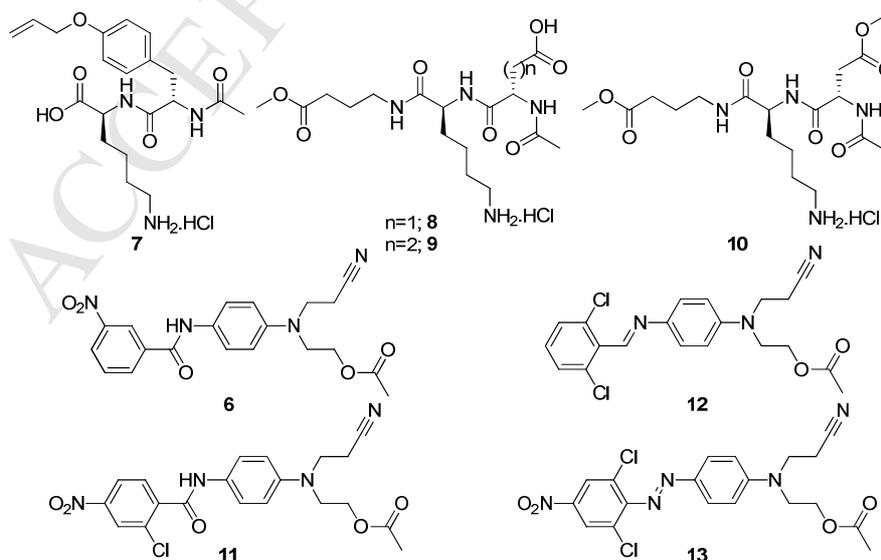
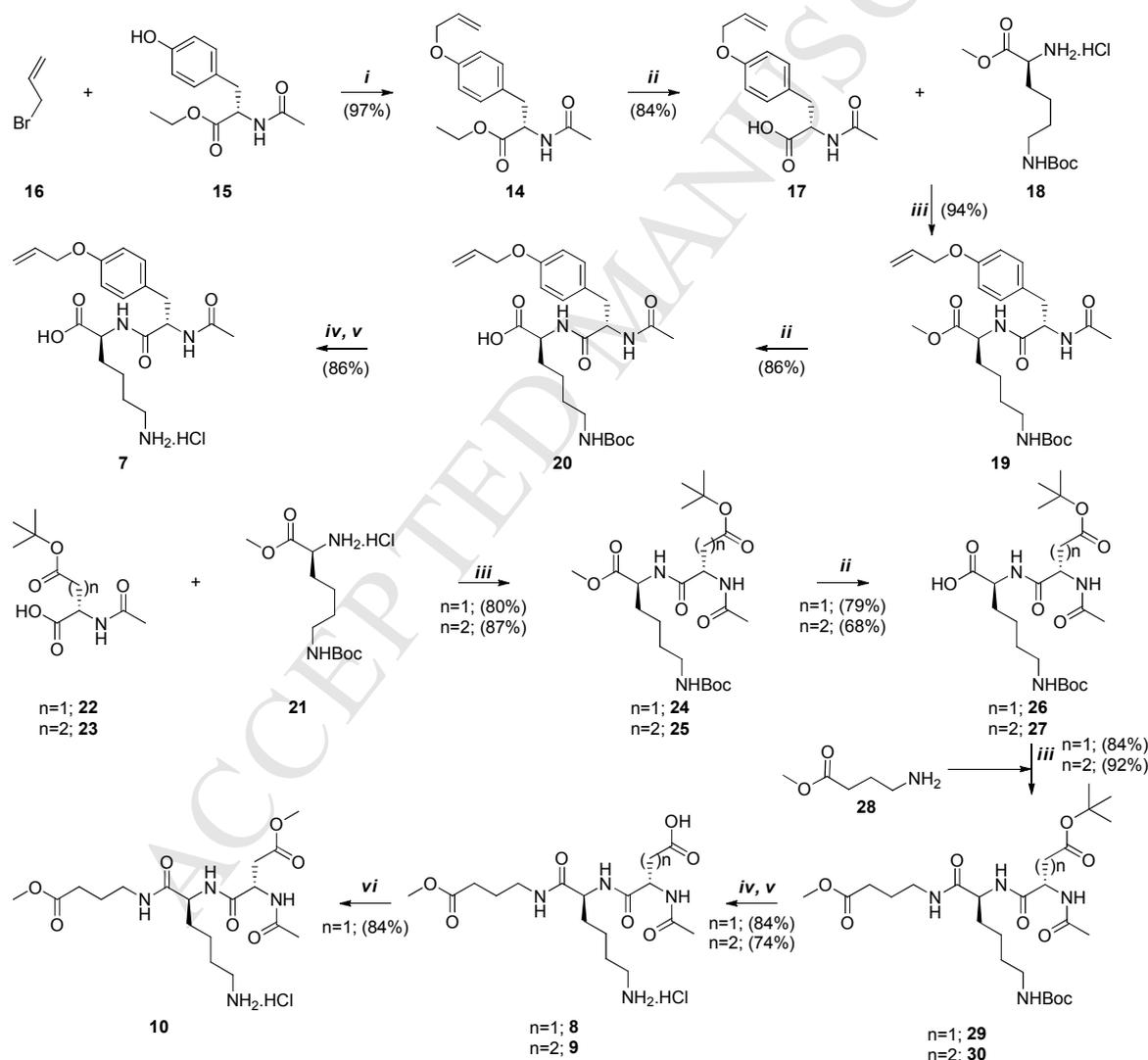


Figure 5: Structure of the dipeptide (**7**) and tripeptide derivatives (**8** – **10**) which were designed and marked for synthesis from structure-based design along with the nitrile-based analogues **6**, **11** – **13** which were predicted to possess IN inhibitory activity from docking runs. Note; final docked energies were not analysed given the structural dissimilarities, particularly in relation to number of rotatable bonds.

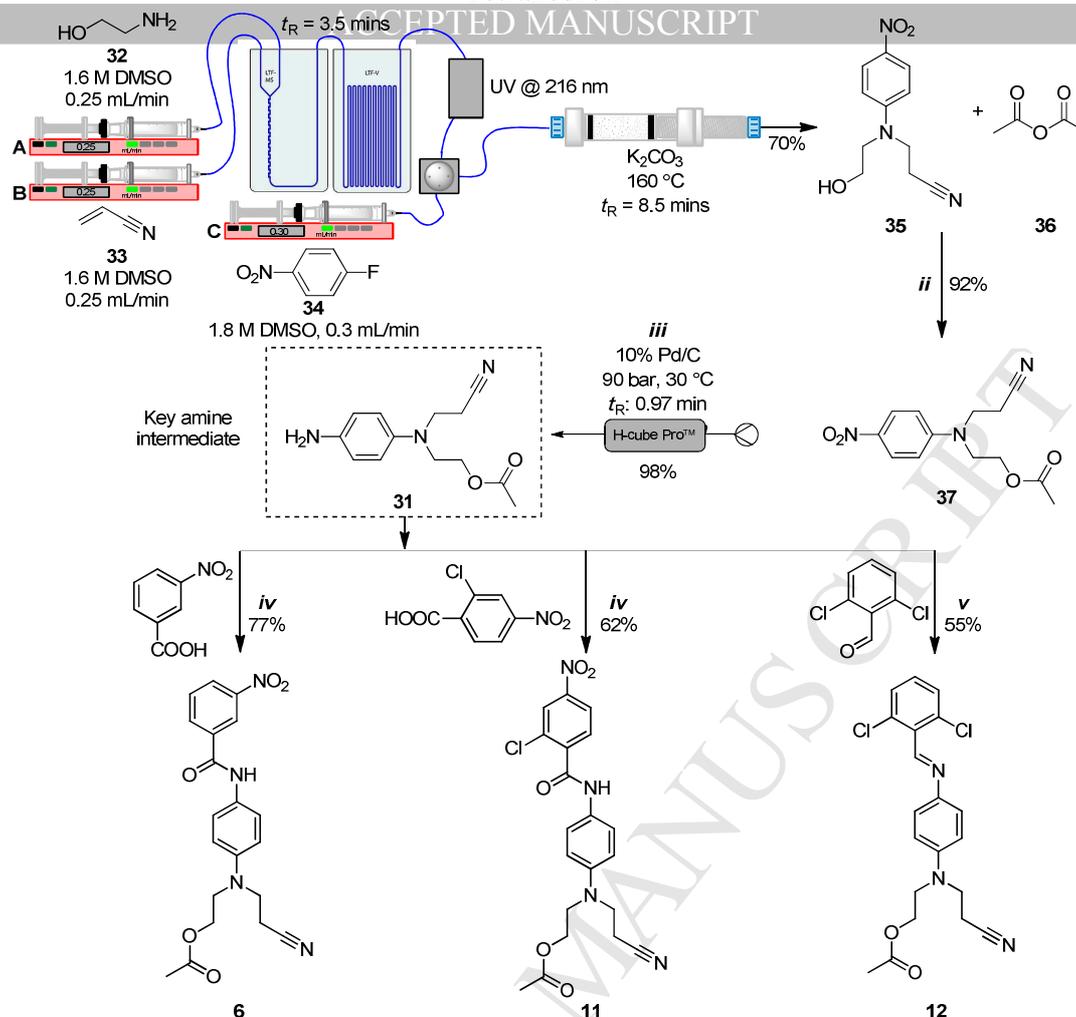
As illustrated in Scheme 1, the dipeptide analogue **7** was accessed in a five-step procedure from **14**, which was initially prepared *via* nucleophilic *O*-allylation of commercially available (*S*)-*N*-acetyltyrosine ethyl ester (**15**) with allyl bromide (**16**). Subsequent ester hydrolysis afforded **17** which was coupled to **18** under typical EDCI-HOBt-mediated amide formation conditions and the resulting dipeptide **19** was hydrolyzed to furnish **20**. Finally, TFA mediated *N*-Boc-de-protection followed by treatment with HCl in ether yielded the desired analogue **7**. The carboxylate tripeptide analogues **8** and **9** were synthesised in a four-step procedure which initially entailed EDCI-HOBt mediated coupling of **21** with **22** or **23**, respectively. The resultant dipeptides **24** and **25** were hydrolyzed and the exposed carboxylate moieties of **26** and **27** were then coupled to **28** utilising EDCI-HOBt amide coupling conditions. Finally, concurrent Boc and *t*-Bu de-protection of **29** and **30** was effected utilising a 1:1 solution of CH₂Cl₂/TFA with the resulting intermediates converted to the hydrochloride salts **8** and **9**. Lastly the third desired tripeptide, **10** was accessed from **8** utilising SO₂Cl₂/MeOH esterification conditions.

was commercially available while **6**, **11**, and **12** were accessed *via* a mixed batch-flow synthetic approach. The required key amine intermediate **31** was acquired from a four-step procedure (Scheme 2). The first phase of the procedure entailed *in situ* formation of 3-((2-hydroxyethyl)amino)propanenitrile *via* an aza-Michael reaction. Upon the reaction mixture eluting from the microreactor solution of 1-fluoro-4-nitrobenzene (**34**) was introduced into the reaction stream. The resulting mixture then flowed through an Omnifit® column packed with K₂CO₃ and upon eluting from the column, the reaction stream was collected in H₂O with the resulting precipitate collected to afford **35** in a 70% overall yield. Compound **35** was acetylated to afford **37** and subsequent hydrogenation afforded the key amine **31** in a quantitative yield. From the key amine **31**, compounds **6** and **11** were obtained from DCC/DMAP mediated amide bond formation condition, whilst **12** was furnished under typical imine formation conditions.



Scheme 1: Synthetic procedure to access the desired allyltyrosine based dipeptide **7** and the tripeptide derivative **8** – **10**.

Reagents and Conditions: *i*) K₂CO₃ (aq.) (2 eq.), CH₂=CH-CH₂-Br *ii*) LiOH.H₂O (2 eq.), THF/H₂O (3:1), *iii*) EDCI (1.1 eq.), HOBt (1.1 eq.), DIPEA (1 eq.), DMF, *iv*) TFA/CH₂Cl₂ (1:1), *v*) 1 M HCl/diethyl ether, *vi*) SOCl₂, CH₃OH.



Scheme 2: Synthesis of the nitrile-based analogues **6**, **11**, and **12**. *Reagents and conditions:* *i*) syringe pump A 0.25 mL/min 1.6 M ethanolamine (DMSO, 10 mL), syringe pump B 0.25 mL/min 1.6 M 3-butenenitrile (DMSO, 10 mL), syringe pump C 0.25 mL/min 1.8 M 1-fluoro-4-nitrobenzene (DMSO, 50 mL); *ii*) Et₃N, rt, 16 h; *iii*) H-cube pro, **37** in EtOH (0.05 M), 50 °C, 50 bar, 100 % H₂, 0.5 mL/min, 10 % Pd/C (70 mm, t_R = 0.97 min); *iv*) DCC (1.1 eq.), DMAP (1.1 eq.), DMF, rt, 16 h; *v*) 3 Å molecular sieves, anhydrous MeOH, rt, 18 h.

2.5. HIV IN inhibitory assays

The inhibitory activity of **6** – **13** was assessed using a previously reported combination 3'-processing and strand transfer microtitre plate assay.^{70,71} As outlined in Table 2, whilst each of the peptide analogues maintained activity comparable to the lead

allyltyrosine **5**, (e.g. IC₅₀ = 10 μM), significant potency enhancements were not forthcoming. Nonetheless, in support of the model and docking, each of the nitrile-based derivatives displayed inhibitory activity with **6**, **11**, and **12** returning modest IC₅₀ values of 85 μM, 35 μM, and 40 μM, respectively, whilst **13** returned an IC₅₀ value of 0.5 μM.

Table 2: Structure and activity of the dipeptide and tripeptide derivatives **7**, and **8** – **10** along with the nitrile-based analogues **6**, **11** – **13**. The activities were ascertained from a combined 3'-processing (3'P) and strand-transfer (ST) assay.

Compound	R ¹	R ²	IC ₅₀ μM ^a	Compound	R ³	IC ₅₀ μM ^a
7	HO-CH ₂ -CH ₂ -NH ₂	CH ₂ -CH ₂ -O-CH ₂ -CH ₂ -CH ₂ -NH ₂	12	6	NO ₂ -C ₆ H ₄ -NH-CO-	85
8	CH ₃ -CO-(CH ₂) ₃ -NH-	HO-CH ₂ -CH ₂ -NH ₂	17	11	Cl-C ₆ H ₃ (NO ₂)-NH-CO-	35
9	CH ₃ -CO-(CH ₂) ₃ -NH-	HO-CH ₂ -CH ₂ -CH ₂ -NH ₂	15	12	Cl-C ₆ H ₃ (Cl)-NH-CO-	40
10	CH ₃ -CO-(CH ₂) ₃ -NH-	CH ₃ -CO-CH ₂ -NH ₂	19	13	Cl-C ₆ H ₃ (Cl)(NO ₂)-NH-CO-	0.5

^aIC₅₀ values are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

Table 3: The 3'-processing (3'-P) inhibitory activities of **5**, **7** – **10** and **13** in the presence of Mg²⁺ or Mn²⁺, along with the strand-transfer inhibitory activities of the most potent analogues (IC₅₀ values in μM).

Compound	R ¹	R ²	IC ₅₀ μM	3'-P (Mg ²⁺) IC ₅₀ (μM) ^a	3'-P (Mn ²⁺) IC ₅₀ (μM) ^a	ST IC ₅₀ (μM) ^a
5			10	> 30	>30	9
7			10	> 30	>30	9
8			17	>30	>30	13
9			15	>30	>30	12
10			19	>30	>30	15
13			0.5	28	> 30	1

^aIC₅₀ values are the mean ±95% confidence interval (CI) of one experiment performed in triplicate.

In accordance with the previous study,¹⁷ to ascertain whether the peptide analogues (e.g. **7** – **10**) maintained strand-transfer specific inhibitory activity, and to gain insight into the potential inhibitory mechanism of **13**, these derivatives were subjected to individual 3'-processing and strand-transfer inhibition assays. Initially, 3'-processing inhibitory activity was examined in the presence of Mg²⁺, as it is generally accepted that Mg²⁺ is the co-factor for integration in cells.^{72–74} However, the assay was also performed using Mn²⁺ as a co-factor as Mn²⁺ appears to be required *in vitro* for the DKAs to produce potent inhibition.^{22,25,72–74} As outlined in Table 3, while specific stand-transfer specific inhibition was retained by **5** and **7** – **10**, compound **13** proved to be less specific displaying an IC₅₀ value of 28 μM in the Mg²⁺ mediated 3'-processing assay.

Despite maintaining comparable activity and strand transfer inhibition selectivity to the lead allyltyrosine **5**, given the anionic nature of **7** – **10**, the ability of these derivatives to elicit whole cell activity was questioned. Hence **7** and **9**, in addition to **13**, were subjected to whole cell assays previously reported and utilised by Ovenden *et al.*⁷¹ to assess both potency and toxicity against human T78 T-cells (Table 4). Further, as the whole cell activity of **5** had not been evaluated in the previous study¹⁷ it was also examined. As charted in Table 3, consistent with the lead **5**, **7** did not display significant cytotoxicity or whole cell activity. Likewise, **13** was devoid of whole cell activity, however, gratifyingly, **9** displayed an EC₅₀ value of 10 ± 5 μM while remaining devoid of cytotoxic activity up to 100 μM. Thus while significant improvements of IN inhibitory activity were not forthcoming, the current structure-based approach did afford an analogue with whole cell activity. As the lead compound displayed no quantifiable inhibitory activity at the cellular level, this indeed marks a significant advancement for the scaffold.

3. Conclusions

From the analysis of the current 297 crystal structures of the IN CCD deposited in the protein databank, it was established that the catalytic-site of the IQS4 A-chain adopts an “inhibited-

conformation”, albeit only subtly varied relative to non-inhibited structures. This structure appears to be a valid platform for docking studies of INSTIs. Further, it seems that the allosteric LEDGIn induce a similar conformation in which the Glu-152 side-chain becomes significantly separated from the two other residues (e.g. Asp-116 and Asp-64) of the catalytic triad.

Table 4: Inhibitory activity and cytotoxicity of **7**, **9**, **13**, and **5** at the cellular level against HuT78 T-cells.

Compound	EC ₅₀ (μM) ^{a,b}	CC ₅₀ (μM) ^a
7	> 100	> 100
9	10 ± 5.0	> 100
13	> 100	> 100
5	> 100	> 100

^aThe whole cell antiviral assay was conducted by infecting 50,000 HuT78 T-cells in RPMI with 10% heat-inactivated foetal calf serum with 65 TCID₅₀ units of HIV-1_{NL4-3} in the presence of different concentrations of the drug.

^bIC₅₀ values are the mean ±95% confidence interval (CI) of one experiment performed in triplicate

In relation to the allyltyrosine scaffold, from docking runs of the first generation series,¹⁷ a 3-point binding motif required for inhibitory activity was proposed and subsequently exploited to develop a new set of four peptide analogues which displayed comparable activity to the lead compound **5**. Additionally from docking studies focusing on “in-house” compound libraries a nitrile based analogue **13** displaying an IC₅₀ value of 0.5 μM in the combined 3'-P and ST was identified. While seven of the eight investigated derivatives were devoid of whole cell activity, the most potent tripeptide **9** displayed promising whole cell activity (EC₅₀ 10 ± 5 μM) and as the initial lead compound **5** was devoid of whole cell activity this represents a significant advancement of the scaffold. Moreover, whilst the *in vitro* activity of **9** is significantly higher than the majority of most current INSTIs inhibitors the whole-cell activity is comparable to a number of first generation analogues. For example the activity of **9** is equivalent to L-708906, (EC₅₀ 10.1 μM, CC₅₀ 88.3 μM),¹¹ and L-731988 (EC₅₀ 1 μM, CC₅₀ 520 μM),⁷⁵ and only one order of magnitude lower than S-1360^{76,77} (EC₅₀ 0.2 μM, CC₅₀ 110 μM) which was the first integrase inhibitor to reach clinical studies.⁷⁸

Thus from this preliminary study, a working model of HIV-1 integrase was developed from which a selective inhibitor of the strand-transfer reaction displaying whole-cell activity and negligible cytotoxicity was produced. Compound **9** provides a promising scaffold for further elaboration, and current investigations include resolving a co-crystallised structure of **9** with the IN catalytic core in addition to further structure-based design iterations.

4. Materials and Methods

4.1. General Experimental Procedures

Reagents and solvents were purchased reagent grade and used without further purification unless stated. All reactions were performed in standard glassware. CH₂Cl₂ was distilled from CaCO₃. 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 determined with a Varian Unity 300 MHz spectrometer. Proton NMR (¹H NMR) spectra were acquired at 300.0 MHz. Carbon NMR (¹³C NMR) spectra were acquired at 75.4 MHz. Spectra were recorded in deuterated chloroform (CDCl₃) with 0.5% trimethylsilane (TMS), obtained from Cambridge Isotope Laboratories Inc., using TMS (δ 0.00 ppm) 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), doublet of triplet (dt), triplet (t), triplet of doublet of doublets (tdd), triplet of doublets (td), quartet (q) and multiplet (m). Each resonance is listed according to the following convention: chemical shift, multiplicity, coupling constant, integration, assignment. Interchangeable resonances 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 stated with their relative intensities in parentheses. ES high resolution mass spectra (HRMS) were obtained using a Q-ToF mass spectrometer.

Thin layer chromatography (TLC) was performed using Merck Silica Gel F₂₅₄ 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 (Büchi rotary evaporator).

4.2. Flow chemistry system

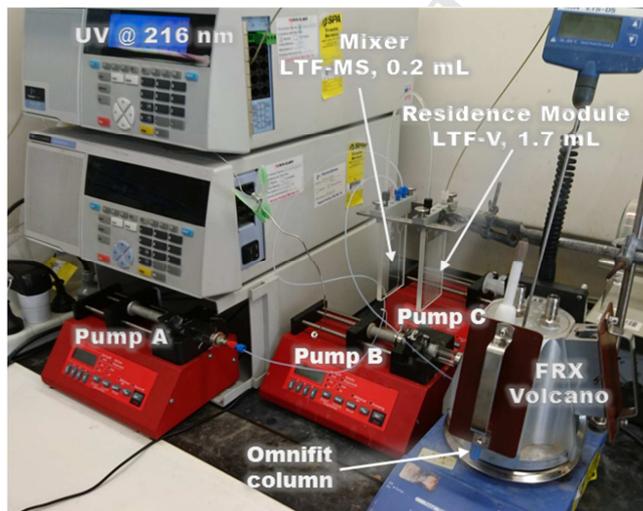


Figure 6: The flow chemistry system utilised to synthesise the key amine intermediate **31**. *Syringe pumps*, World Precision Instruments Ltd AL300-220 Aladdin Infusion Pump; *UV detector*, PerkinElmer series 200 UV/Vis detector, λ 290 nm, range 3.000 AUFS, response time 0.5; *Little Things Factory Microreactors*, mixer MR-Lab MS (volume 0.2 mL), residence stretch MR-Lab V (volume 1.7 mL); *Omnifit*[®] column, L × I.D. 100 mm × 10 mm, bed volume 5.6 mL; *Column heating block*; syrris FRX Volcano.

4.3. Modelling Protocols

4.3.1. Software and Hardware

Construction and manipulation of all ligand and enzyme structures was performed using the BUILDER module of InsightII. All ligand and enzyme minimisations were performed within the consistent valence force field (CVFF) with charges using a combination of the steepest descents and conjugate algorithms as specified. Minimisation calculations were conducted using the DISCOVER module of InsightII. InsightII was installed on a Silicon Graphics O2, MIPS R12000, 270 MHz processor with 384 Mb RAM.

All docking runs were performed using the AutoDock suite.⁶⁹ AutoDockTools (ADT) was used to prepare ligand and enzyme structures, analyse docking runs and to set AutoGrid and AutoDock parameters. AutoGrid and AutoDock calculations were performed on a Dell PowerEdge 2800 Server, Intel Xeon processor at 2.8GHz/1MB Cache, 800MHz FSB with 36 GB RAM.

Visualisation and analysis of docking runs results were performed using DSViewerPro (standard PC platform software). With the exception of hydrogen bonding, which was automatically detected with a distance between the two heteroatoms of up to 3.0 Å, ligand enzyme interactions were determined *via* manual inspection. Enzyme solid surfaces were represented as solvent assessable surfaces with a probe radius of 1.4 Å.

4.3.2. Enzyme Preparation and Insertion of the Unresolved Flexible Loop

The atomic coordinates of 1QS4 were obtained from the Protein Data Bank. Hydrogen atoms were added to the structure by setting the pH to 7.0 and the C-chain, the magnesium ion Mg1003, all water molecules, with the exception of those coordinated to the A and B-chain Mg²⁺ ions (HOH 44, 258, 266, 443, 529, 286, 337, and 520), were removed. The inhibitor, 5CITEP, was assigned as its enol form with the enolic proton placed on the 3-oxygen (oxygen closest to the tetrazole) while the tetrazole was assigned as the 1''H tautomer. The unresolved A-chain residues, IleA141, ProA142, TyrA143, and AsnA144, were sequentially added to GlyA140 and a *trans* amide bond was created between AsnA144 and ProA145. The two unresolved B-chain residues, ProB142 and TyrB143, were sequentially added from IleB141 and a *trans* amide bond was generated between ProB142 and TyrB143. Potentials, partial charges and formal charges were set in accordance with the CVFF. The two magnesium ions, Mg1001 and Mg1002, were assigned a formal and partial charge of +2. An "active site" set, which included all residues within 10 Å of 5CITEP, was generated and a pseudo atom was placed at the centre of the set. The pseudo atom was transformed to 0, 0, 0, in absolute coordinates thus placing the A-chain active site at the centre of the atomic grid.

Minimisation of the inserted flexible was performed over six steps. During the first two steps, all heavy atoms were constrained to their current positions ("fixed") and the steepest descents algorithm was employed for 10000 iterations or until the derivative converged to 0.01 cal. followed by further minimisation using the conjugate gradient algorithm over 100000 iterations or until the derivative converged to 0.001. During the third and fourth steps constraints were removed from backbone atoms and minimisation was performed as described in steps 1 and 2, while in the remaining two steps all constraints were removed and minimisations was again performed as described in steps 1 and 2.

4.3.3. Docking Run Parameters

Docked conformations were generated using the Lamarckian genetic algorithm. Global optimisation started with a population of 50 individuals and a maximum of 1.0×10^6 energy evaluations for small ligands, or 5.0×10^6 evaluations for peptide analogues. A maximum of 27000 generations were produced with a total of 250 docking runs (ga runs) performed. Remaining parameters were retained at AutoDock 3.0 defaults with the elitism value (number of top individuals that automatically survived) set to 1. The average of the worst energy was calculated over a window of 10 previous generations and the rate of crossover was 0.8. The rate at which random gene mutation occurred was 0.02, the mean of Cauchy distribution for gene mutation was zero, and the variance of Cauchy distribution for gene mutation was 1.0. For the local search, the pseudo-Soils and Wets local search method was employed with a maximum of 300 iterations and 50 local search (LS) runs. The maximum number of consecutive successes or failures before halving or doubling the local step size (ρ) was four. The Soils and Wets parameter defining the initial variance and size of the local space to be sampled was 1.0. The lower bound on ρ , or the termination criterion for the local search, was 0.01. The probability of any particular phenotype being subjected to a local search was 0.06. For the dock options, the platform-independent library was utilised as the random number generator with the two random number generator seeds being PID and time. The energy parameter for the external grid energy was $1000.0 \text{ kcal mol}^{-1}$, the maximum allowable initial energy was $0.0 \text{ kcal mol}^{-1}$, the maximum number of retries was 1000, whilst the internal electrostatic energy was not calculated.

The step size parameter for translation was 2.0 Å per step and the quaternion and torsion were both set at 50° per step. The root-mean-square positional deviation was set at 2.0 Å for larger flexible molecules or 1.0 Å for smaller rigid molecules.

4.3.4. Analysis of Docking Runs

Analysis of docking runs was conducted with ADT using cluster analysis. The RMSD tolerance was set to 1.0 Å for small molecules and 2.0 Å for the peptide derivatives. The conformation with the lowest energy within the most populated cluster was used as the final docked conformation.

4.3.5. Reproduction of the 5CITEP Crystallographic Resolved Orientation

The validation docking runs were conducted on five receptor models: 1) the completed dimer (denoted as **model III**), 2) the completed dimer with the Mg²⁺ coordinated water molecules removed (denoted as **model II**), 3) The completed dimer with the Mg²⁺ ions and coordinated water molecules removed (denoted as **model I**). In the remaining two receptor models, the crystallographically resolved Mg²⁺ ions and coordinated water molecules along with a second Mg²⁺ ion which was hydrated (denoted as **model V**), 5) the crystallographically resolved Mg²⁺ ions and coordinated water molecules along with a second Mg²⁺ ion which was not hydrated (denoted as **model IV**). In each model, 5CITEP was manually translated 10 Å directly above the active site and saved as a separate entity before being removed from the model's atomic coordinates. The inhibitor, 5CITEP, was assigned as its enol form with the enolic proton placed on the 3-oxygen (oxygen closest to the tetrazole) while the tetrazole was assigned as the 1''H tautomer. The ligand non-polar hydrogens and lone pairs were merged, and charges were assigned using the Gasteiger-Marsilli formalism. The ligand rigid root, to which the rotatable groups are connected, was assigned automatically and all rotatable bonds were fixed. Utilising the utilities implemented by ADT, enzyme non-polar hydrogens were merged, Kollman united-atom partial charges were assigned, lone pairs were merged and solvation parameters were added to each of the five models (**model I-V**). A formal charge of +2 was assigned to the Ca²⁺ ions and the charges of the coordinated water molecules were assigned using the Gasteiger-Marsilli formalism. Each model (**model I-V**) was embedded in a grid centred at atomic coordinates 0, 0, 0, with dimensions of $26.25 \times 22.5 \times 22.5 \text{ Å}$ ($70 \times 60 \times 60$ points) and a spacing of 0.375 Å between grid points. Electrostatic grid maps were subsequently generated for each atom type in the ligand using the auxiliary program AutoGrid3. Docking runs were conducted as described in section 4.2.3.

4.4. General Synthetic Procedures

4.4.1. Procedure A: Allyl Ether Formation¹⁷

The phenol derivative and anhydrous potassium carbonate (K₂CO₃) were combined and dried under vacuum for 1 h. The vessel was then sealed and flushed with N₂, before anhydrous DMF was added. The mixture was allowed to stir at rt for 30 min before allyl bromide 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 (MgSO₄) and concentrated *in vacuo* to yield the allyl ether product.

4.4.2. Procedure B: Methyl /Ethyl Ester Hydrolysis¹⁷

To a solution of the ester in THF/H₂O (3:1) was added LiOH.H₂O and the resulting suspension was allowed to stir for 12 h at rt, before being diluted 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 concentrated *in vacuo* to yield the desired acid.

4.4.3. Procedure C: Amide Coupling¹⁷

The acid, HOBt, EDCI, and the amine hydrochloride 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 and DIPEA 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 diluted 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 concentrated *in vacuo* yield the desired amide.

4.4.4. Procedure D: N-Boc De-protection¹⁷

A solution of the N-Boc protected amine in $\text{CH}_2\text{Cl}_2/\text{TFA}$ (1:1) was stirred at rt for 3 h. The crude was concentrated *in vacuo* to yield the crude amine as the trifluoroacetate salt, which was converted to the hydrochloride salt.

4.4.5. Procedure E: Hydrochloride Salt Formation¹⁷

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.

4.4.6. Procedure F: Methyl Ester Formation¹⁷

To a stirred solution of the appropriate amino acid in MeOH at 0°C SOCl_2 was slowly added. The solution was then removed from the ice bath and stirred at rt for 3 h. The reaction was then concentrated *in vacuo* leaving the methyl ester amino acid as the hydrochloride salt.

4.5. HIV-integrase assays

4.5.1. In vitro assays^{71,79}

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 procedure⁷⁰ 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 $\mu\text{g}/\text{mL}$ 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.*⁷¹ using individually either Mg^{2+} or Mn^{2+} . 3'-Processing assays utilised the Chow *et al.*^{71,80} procedure without modification.

4.5.2. Whole Cell Assays⁷¹

The whole cell antiviral assay was conducted by infecting 50,000 HuT78 T-cells in RPMI + 10% heat inactivated foetal calf serum, in each well of a 96 well plate, with 65 TCID₅₀ units of HIV-1_{NL4-3} in the presence of different concentrations of drug. After an overnight infection in a total volume of 100 μL and

equal volume of fresh media containing drug was added. After an additional 48 h, 100 μL of media was removed and replaced with 100 μL fresh media containing drug. A sample of culture supernatant was diluted 1 in 2000 and p24 measured using a p24 ELISA assay (Vironos-tika HIV-1 Antigen kit, Organon Teknika). For cytotoxicity assays, cells were passaged in the presence of compound as described for the infection assay: however, no virus was added. The viability of cells was assessed using the Cytolux assay (Perkin-Elmer Life Sciences) following the manufacturer's instructions.

4.6. Compound Characterisation

4.6.1. Ethyl (S)-2-(4-allyloxybenzyl)-3-aza-4-oxopentanoate (14)⁸¹

Using **procedure A**, and the commercially available (S)-N-acetyltyrosine ethyl ester monohydrate (2.60 g, 10.35 mmol), anhydrous potassium carbonate (3.20 g, 23.15 mmol), allyl bromide (3.0 g, 24.80 mmol) and DMF (15 mL) as solvent, the ester (**14**) (2.74 g, 9.38 mmol, 97%) was obtained as a white solid, mp. 69-71°C. MS (CI + ve) *m/z* 292 (100%) [MH^+]. HRMS (ESI^+) calcd for $\text{C}_{16}\text{H}_{21}\text{NO}_4 + \text{H}$: 292.1543, found 292.1531. ¹H NMR (300 MHz, CDCl_3): δ 7.00 (d, $J = 8.7$ Hz, 2H, 2'-CH and 6'-CH); 6.83 (d, $J = 8.7$ Hz, 2H, 3'-CH and 5'-CH); 6.04 (tdd, $J = 17.3, 10.5, 5.3$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CH}_2$); 5.91 (bs, 1H, 3-NH); 5.40 (dd, $J = 17.3, 1.6$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CHH trans}$); 5.28 (dd, $J = 10.5, 1.4$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CHH cis}$); 4.82 (td, $J = 7.6, 5.6$ Hz, 1H, 2-CH); 4.51 (d, $J = 5.3$ Hz, 2H, $\text{OCH}_2\text{CH}=\text{CH}_2$); 4.17 (q, $J = 7.1$ Hz, 2H, OCH_2CH_3); 3.06 (d, $J = 5.4$ Hz, 2H, 2- CHCH_2); 1.99 (s, 3H, 5- CH_3); 1.25 (t, $J = 7.1$ Hz, 3H, OCH_2CH_3). ¹³C NMR (75 MHz, CDCl_3): δ 172.0 (C-1); 169.9 (C-4); 157.9 (C-4'); 133.5 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 130.5 (C-2' and C-6'); 128.3 (C-1'); 117.9 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 115.7 (C-3' and C-6'); 69.0 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 61.7 (OCH_2CH_3); 53.5 (C-2); 37.3 (2- CHCH_2); 23.3 (C-5); 14.4 (OCH_2CH_3).

4.6.2. (S)-2-(4-Allyloxybenzyl)-3-aza-4-oxopentanoic acid (17)⁸¹

Compound **17** was synthesised using **procedure B**, from the ester **14** (2.71 g, 9.31 mmol) in THF/water 3:1 (80 mL) with $\text{LiOH}\cdot\text{H}_2\text{O}$ (837 mg, 19.35 mmol) to yield **17** (2.07 g, 7.86 mmol, 84%) as white granular crystals, mp. 175-178°C. MS (ESI^-), *m/z* 261 (20%) [$\text{M} - \text{H}^-$], 114 (25), 112 (100). HRMS (ESI^-) calcd for $\text{C}_{14}\text{H}_{17}\text{NO}_4 - \text{H}$: 262.1085; found 262.1078. ¹H NMR (300 MHz, $(\text{CD}_3)_2\text{CO}$): δ 7.25 (bs, 1H, 3-NH); 7.03 (d, $J = 8.7$ Hz, 2H, 2'-CH and 6'-CH); 6.81 (d, $J = 8.7$ Hz, 2H, 3'-CH and 5'-CH); 6.01 (tdd, $J = 17.3, 10.6, 5.3$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CH}_2$); 5.37 (dd, $J = 17.3, 1.6$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CHH trans}$); 5.25 (dd, $J = 10.5, 1.4$ Hz, 1H, $\text{OCH}_2\text{CH}=\text{CHH cis}$); 4.75-4.70 (m, 1H, 2-CH); 4.47 (ddd, $J = 5.3, 1.5, 1.5$ Hz, 2H, $\text{OCH}_2\text{CH}=\text{CH}_2$); 3.11 (dd, $J = 14.1, 5.6$ Hz, 1H, 2- CHCH_2); 3.01 (dd, $J = 14.1, 5.9$ Hz, 1H, 2- CHCH_2); 1.95 (s, 3H, 5- CH_3). ¹³C NMR (75 MHz, $(\text{CD}_3)_2\text{CO}$): δ 172.5 (C-1); 169.7 (C-4); 157.8 (C-4'); 134.2 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 130.5 (C-2' and C-6'); 129.6 (C-1'); 116.6 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 114.6 (C-3' and C-6'); 68.5 ($\text{OCH}_2\text{CH}=\text{CH}_2$); 53.9 (C-2); 36.7 (2- CHCH_2); 22.0 (C-5).

4.6.3. Methyl (2S,5S)-5-(4-allyloxybenzyl)-3,6-diaza-2-[4-(tert-butoxycarbonyl)aminobutyl]-4,7-dioxooctanoate (19)

The ester was synthesised using procedure C from the acid **17** (330 mg, 1.25 mmol), the commercially available methyl (S)-2-amino-6-(tert-butoxycarbonylamino)hexanoate hydrochloride (**18**) (330 mg, 1.12 mmol), EDCI (240 mg, 1.26 mmol), HOBt (170 mg, 1.26 mmol), DIPEA (0.20 mL, 1.15 mmol) and DMF (3 mL) as the solvent. The reaction mixture was allowed to stir for 14 h before being quenched to yield **19** (530 mg, 1.05 mmol,

94%) as an off white powder, mp. 102-105°C. MS (ESI⁺), *m/z* 506 (100%) [MH⁺], 528 (40) [M + Na⁺], 450 (20), 406 (55). HRMS (ESI⁺) calcd for C₂₆H₃₉N₃O₇ + Na: 528.2686; found: 528.2690. ¹H NMR (300 MHz, CDCl₃): δ 7.11 (d, *J* = 8.5 Hz, 2H, 2'-CH and 6'-CH); 6.82 (d, *J* = 8.5 Hz, 2H, 3'-CH and 5'-CH); 6.59 (d, *J* = 6.7 Hz, 1H, 3-NH); 6.04 (tdd, *J* = 16.9, 10.5, 5.3 Hz, 1H, OCH₂CH=CH₂); 5.40 (dd, *J* = 17.3, 1.5 Hz, 1H, OCH₂CH=CH₂ *trans*); 5.27 (dd, *J* = 10.5, 1.3 Hz, 1H, OCH₂CH=CH₂ *cis*); 4.95 (bs, 1H, 5-CH); 4.76-4.68 (m, 1H, 2-CH); 4.50 (bd, *J* = 5.3 Hz, 2H, OCH₂CH=CH₂); 3.70 (s, 3H, OCH₃); 3.10-2.90 (m, 6H, 5-CHCH₂, 4''-CH₂, 6-NH and 4''-CH₂NH); 1.96 (s, 3H, 8-CH₃); 1.86-1.71 (m, 1H, 1''-CH₂H_b); 1.67-1.58 (m, 1H, 1''-CH₂H_b); 1.49-1.37 (m, 11H, OC(CH₃)₃ and 3''-CH₂); 1.27 (m, 2H, 2''-CH₂). ¹³C NMR (75 MHz, CDCl₃): δ 172.4 (C-4); 171.5 (C-1); 170.5 (C-7); 157.8 (C-4'); 156.3 (NHCOOC(CH₃)₃); 133.5 (OCH₂CH=CH₂); 130.5 (C-1'); 128.9 (C-2' and C-6'); 117.8 (OCH₂CH=CH₂); 115.0 (C-3' and C-5'); 79.3 (COOC(CH₃)₃); 69.0 (OCH₂CH=CH₂); 54.7 (C-5); 52.5 (C-2); 52.4 (OCH₃); 40.3 (C-4''); 37.6 (5-CHCH₂); 32.0 (C-1''); 29.6 (C-3''); 28.7 (OC(CH₃)₃); 23.2 (C-8); 22.7 (C-2'').

4.6.4. (2*S*,5*S*)-5-(4-Allyloxybenzyl)-3,6-diaza-2-[4-(*tert*-butoxycarbonyl)aminobutyl]-4,7-dioxooctanoic acid (**20**)

The acid **20** was synthesised using **procedure B** from the ester **19** (1.54 g, 3.05 mmol), LiOH.H₂O (256 mg, 6.1 mmol) and THF/water 3:1 (80 mL). The reaction was stirred for 14 h before being diluted with water (25 mL). Unreacted starting materials were extracted using CH₂Cl₂ (30 mL) then EtOAc (30 mL). After the aqueous solution was acidified the product was extracted with CH₂Cl₂ (2 x 30 mL) and EtOAc (2 x 30 mL), the combined organic fractions were dried and evaporated to yield the acid **20** (1.29 g, 2.63 mmol, 86%) as white crystals, mp. 62-66°C. MS (ESI⁺), *m/z* 492 (100%) [MH⁺], 436 [MH⁺ - OCH₂CH=CH₂] (25), 392 [MH⁺ - Boc] (70). HRMS (ESI⁺) calcd for C₂₅H₃₇N₃O₇ + H: 492.2704; found 492.2715. ¹H NMR (300 MHz, (CD₃)₂CO): δ 7.15 (d, *J* = 8.7 Hz, 2H, 2'-CH and 6'-CH); 6.81 (d, *J* = 8.7 Hz, 2H, 3'-CH and 5'-CH); 6.01 (tdd, *J* = 17.2, 10.5, 5.2 Hz, 1H, OCH₂CH=CH₂); 5.36 (dd, *J* = 17.3, 1.7 Hz, 1H, OCH₂CH=CH₂ *trans*); 5.19 (dd, *J* = 10.6, 1.5 Hz, 1H, OCH₂CH=CH₂ *cis*); 4.61 (dd, *J* = 8.9, 5.2 Hz, 1H, 5-CH); 4.48 (d, *J* = 5.2, Hz, 2H, OCH₂CH=CH₂); 4.41-4.33 (m, 1H, 2-CH); 3.07 (dd, *J* = 14.1, 5.1 Hz, 1H, 5-CHCH₂H_b); 3.01 (t, *J* = 6.8 Hz, 2H, 4''-CH₂); 2.78 (dd, *J* = 14.0, 8.9 Hz, 1H, 5-CHCH₂H_b); 1.86 (s, 3H, 8-CH₃); 1.83-1.79 (m, 1H, 1''-CH₂H_b); 1.72-1.68 (m, 1H, 1''-CH₂H_b); 1.43-1.31 (m, 13H, OC(CH₃)₃, 3''-CH₂ and 2''-CH₂). ¹³C NMR (75 MHz, (CD₃)₂CO): δ 175.3 (C-1); 174.1 (C-4); 174.0 (C-7); 159.6 (C-4'); 157.8 (NHCOOC(CH₃)₃); 135.9 (OCH₂CH=CH₂); 132.2 (C-1'), 131.4 (C-2' and C-6'), 118.2 (OCH₂CH=CH₂); 116.4 (C-3' and C-5'); 80.2 (COOC(CH₃)₃); 70.4 (OCH₂CH=CH₂); 56.7 (C-5); 54.2 (C-2); 41.9 (C-4''); 38.8 (5-CHCH₂); 33.3 (C-1''); 31.3 (C-3''); 29.6 (OC(CH₃)₃); 24.8 (C-8); 23.4 (C-2'').

4.6.5. (2*S*,5*S*)-5-(4-Allyloxybenzyl)-2-(4-aminobutyl)-3,6-diaza-4,7-dioxooctanoic acid hydrochloride (**7**)

Compound **20** (101 mg, 0.21 mmol) was converted to the uncharacterised *N*-Boc deprotected trifluoroacetate salt via **procedure D**, and the resulting solid was then converted, via **procedure E**, using acetonitrile as the solvent, to give the hydrochloride salt **7** (77 mg, 0.18 mmol, 86%) as a hygroscopic light brown amorphous solid. MS (ESI⁺), *m/z* 392 (100%) [MH⁺], 393 (22) [MD⁺]. HRMS (ESI⁺) calcd for C₂₀H₂₉N₃O₅ + H: 392.2185; found 392.2187. [α]_D²⁵ +64.9 (c. 0.12, EtOH). ¹H NMR (300 MHz, CD₃OD): δ 7.28-7.18 (m, 2H, 2'-CH and 6'-CH); 6.94-6.84 (m, 2H, 3'-CH and 5'-CH); 6.18-6.00 (m, 1H,

OCH₂CH=CH₂); 5.42 (bd, *J* = 17.7 Hz, 1H, OCH₂CH=CH₂ *trans*); 5.27 (bd, *J* = 8.4 Hz, 1H, OCH₂CH=CH₂ *cis*); 4.89-4.76 (1H, 5-CH); 4.62-4.43 (m, 3H, OCH₂CH=CH₂ and 2-CH); 3.17-2.77 (m, 4H, 4''-CH₂ and 5-CHCH₂); 1.95 (bs, 3H, 8-CH₃); 1.80-1.65 (m, 4H, 1''-CH₂ and 3''-CH₂); 1.56-1.45 (m, 2H, 2''-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 175.6 (C-1); 174.7 (C-4); 173.9 (C-7); 158.9 (C-4'); 134.9 (OCH₂CH=CH₂); 131.2 (C-1'); 130.4 (C-2' and C-6'); 117.3 (OCH₂CH=CH₂); 115.6 (C-3' and C-5'); 69.7 (OCH₂CH=CH₂); 56.4 (C-2); 51.0 (C-5); 40.5 (C-4''); 37.8 (5-CHCH₂); 32.0 (C-1''); 27.8 (C-3''); 23.6 (C-8); 22.4 (C-2'').

4.6.6. Methyl (2*S*,5*S*)-3,6-diaza-2-[4-(*tert*-butoxycarbonyl)aminobutyl]-5-(*tert*-butoxycarbonyl)methyl-4,7-dioxooctanoate (**24**)

The ester was synthesised using **procedure C** from (*S*) *N*-acetyl-4-(1,1-dimethylethyl)ester aspartic acid **22** (395 mg, 1.71 mmol), methyl (*S*)-2-amino-6-(*tert*-butoxycarbonylamino)hexanoate hydrochloride (**21**) (501 mg, 1.69 mmol), EDCI (444 mg, 2.32 mmol), HOBT (340 mg, 2.52 mmol) and DIPEA (0.30 mL, 1.72 mmol), to yield **24** (640 mg, 1.35 mmol, 80%) as a white powder, mp. 130-133°C. MS (ESI⁺), *m/z* 474 (100%) [MH⁺], 419 (20), 374 (45). HRMS (ESI⁺) calcd for C₂₂H₃₉N₃O₈ + H: 474.2815; found 474.2801. ¹H NMR (300 MHz, CDCl₃): δ 7.15 (d, *J* = 8.0 Hz, 1H, 3-NH); 6.97 (d, *J* = 7.1 Hz, 1H, 6-NH); 4.78-4.74 (m, 1H, 5-CH); 4.50-4.46 (m, 1H, 2-CH); 3.68 (s, 3H, OCH₃); 3.04 (t, *J* = 6.8, Hz, 2H, 4'-CH₂); 2.79 (dd, *J* = 16.9, 4.7 Hz, 1H, 5-CHCH₂H_b); 2.56 (dd, *J* = 16.9, 7.0 Hz, 1H, 5-CHCH₂H_b); 2.00 (s, 3H, 8-CH₃); 1.89-1.73 (m, 1H, 1'-CH₂H_b); 1.71-1.57 (m, 1H, 1'-CH₂H_b); 1.50-1.37 (m, 20H, NHCOOC(CH₃)₃, CH₂COOC(CH₃)₃ and 3'-CH₂); 1.35-1.25 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CDCl₃): δ 172.4 (CH₂COOC(CH₃)₃); 171.5 (C-4); 170.8 (C-1); 170.5 (C-7); 156.3 (NHCOOC(CH₃)₃); 82.1 (CH₂COOC(CH₃)₃); 79.4 (NHCOOC(CH₃)₃); 52.5 (C-2); 52.3 (OCH₃); 49.5 (C-5); 40.5 (C-4'); 37.2 (5-CHCH₂); 31.8 (C-1'); 29.4 (C-3'); 28.6 (NHCOOC(CH₃)₃^a); 28.2 (CH₂COOC(CH₃)₃^a); 23.3 (C-8); 22.5 (C-2').

4.6.7. Methyl (2*S*,5*S*)-3,6-diaza-2-[4-(*tert*-butoxycarbonyl)aminobutyl]-5-[2-(*tert*-butoxycarbonyl)ethyl]-4,7-dioxooctanoate (**25**)

The ester was synthesised using **procedure C** from the commercially available (*S*) *N*-acetyl-5-(1,1-dimethylethyl)ester glutamic acid (**23**) (405 mg, 1.65 mmol), methyl (*S*)-2-amino-6-(*tert*-butoxycarbonylamino)hexanoate hydrochloride (**21**) (513 mg, 1.73 mmol), EDCI (500 mg, 2.62 mmol), HOBT (422 mg, 3.13 mmol) and DIPEA (0.30 mL, 1.72 mmol), to yield **25** (700 mg, 1.44 mmol, 87%) as a white powder, mp. 133-136°C. MS (ESI⁺), *m/z* 488 (100%) [MH⁺], 388 (30). HRMS (ESI⁺) calcd for C₂₃H₄₁N₃O₈ + H: 488.2972; found 488.2986. ¹H NMR (300 MHz, CDCl₃): δ 7.22 (d, *J* = 7.9 Hz, 1H, 3-NH); 6.82 (d, *J* = 7.5 Hz, 1H, 6-NH); 4.51-4.40 (m, 2H, 2-CH and 5-CH); 3.66 (s, 3H, OCH₃); 3.01 (t, *J* = 6.5, Hz, 2H, 4'-CH₂); 2.41-2.22 (m, 2H, 2''-CH₂); 2.06-1.84 (m, 5H, 8-CH₃ and 1''-CH₂); 1.83-1.72 (m, 1H, 1'-CH₂H_b); 1.69-1.55 (m, 1H, 1'-CH₂H_b); 1.48-1.34 (m, 20H, NHCOOC(CH₃)₃, 2''-CH₂COOC(CH₃)₃ and 3'-CH₂); 1.33-1.22 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CDCl₃): δ 173.1 (CH₂COOC(CH₃)₃); 172.6 (C-4); 171.7 (C-1); 170.7 (C-7); 156.3 (NHCOOC(CH₃)₃); 81.1 (CH₂COOC(CH₃)₃); 79.3 (NHCOOC(CH₃)₃); 52.6 (C-5); 52.5 (C-2); 52.4 (OCH₃); 40.4 (C-4'); 31.8 (C-1'); 31.7 (C-2''); 29.5 (C-3'); 28.6 (NHCOOC(CH₃)₃); 28.2 (CH₂COOC(CH₃)₃); 27.9 (C-1''); 23.2 (C-8); 22.7 (C-2').

4.6.8. (2*S*,5*S*)-3,6-Diaza-2-[4-(*tert*-butoxycarbonyl)aminobutyl]-5-(*tert*-butoxycarbonyl)methyl-4,7-dioxooctanoic acid (**26**)

The acid was synthesised using **procedure B** from the ester **24** (600 mg, 1.27 mmol) and LiOH.H₂O (180 mg, 4.29 mmol) to yield **26** (460 mg, 1.00 mmol, 79%) as white crystals, mp. 115-117°C. MS (ESI⁺), *m/z* 460 (100%) [MH⁺], 461 (50) [MD⁺], 304 (80). HRMS (ESI⁺) calcd for C₂₁H₃₇N₃O₈ + H: 460.2653; found 460.2657. ¹H NMR (300 MHz, CD₃OD): δ 4.76 (dd, *J* = 8.5, 5.5 Hz, 1H, 5-CH); 4.36 (dd, *J* = 8.7, 4.8 Hz, 1H, 2-CH); 3.02 (t, *J* = 6.7 Hz, 2H, 4'-CH₂); 2.77 (dd, *J* = 16.3, 5.5 Hz, 1H, 5-CHCH_aH_b); 2.57 (dd, *J* = 16.2, 8.5 Hz, 1H, 5-CHCH_aH_b); 1.98 (s, 3H, 8-CH₃); 1.94-1.81 (m, 1H, 1'-CH_aH_b); 1.78-1.64 (m, 1H, 1'-CH_aH_b); 1.51-1.34 (m, 22H, NHCOOC(CH₃)₃, CH₂COOC(CH₃)₃, 3'-CH₂ and 2'-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 173.8 (C-1); 172.0 (CH₂COOC(CH₃)₃); 171.8 (C-4); 170.0 (C-7); 157.3 (NHCOOC(CH₃)₃); 81.2 (CH₂COOC(CH₃)₃); 78.7 (NHCOOC(CH₃)₃); 52.4 (C-2); 50.1 (C-5); 40.0 (C-4'); 37.2 (5-CHCH₂); 31.1 (C-1'); 29.2 (C-3'); 27.6 (NHCOOC(CH₃)₃); 27.1 (CH₂COOC(CH₃)₃); 22.8 (C-8); 21.3 (C-2').

4.6.9. (2*S*,5*S*)-3,6-Diaza-2-[4-(*tert*-butoxycarbonyl)aminobutyl]-5-[2-(*tert*-butoxycarbonyl)ethyl]-4,7-dioxooctanoic acid (27**)**

The acid was synthesised using **procedure B** from the ester **25** (820 mg, 1.68 mmol), LiOH.H₂O (174 mg, 4.15 mmol) and THF/water 3:1 (60 mL) to yield the crude product, which was purified by pTLC (silica gel; 15:1 CH₂Cl₂/MeOH) to yield **27** (542 mg, 1.14 mmol, 68%) as white crystals, mp. 110-112°C. MS (ESI⁺), *m/z* 472 (100%) [M - H⁺], 398 (20). HRMS (ESI⁺) calcd for C₂₂H₃₉N₃O₈ - H: 472.2664; found 472.2666. ¹H NMR (300 MHz, CD₃OD): δ 4.42-4.32 (m, 1H, 5-CH); 3.94-3.86 (m, 1H, 2-CH); 3.02 (t, *J* = 6.6 Hz, 2H, 4'-CH₂); 2.33 (dd, *J* = 15.0, 6.9 Hz, 1H, 2''-CH_aH_b); 2.23 (dd, *J* = 15.5, 7.8 Hz, 1H, 2''-CH_aH_b); 2.09-1.91 (m, 4H, 8-CH₃ and 1''-CH_aH_b); 1.89-1.75 (m, 2H, 1''-CH_aH_b and 1'-CH_aH_b); 1.69-1.56 (m, 1H, 1'-CH_aH_b); 1.49-1.36 (m, 22H, NHCOOC(CH₃)₃, 2''-CH₂COOC(CH₃)₃, 3'-CH₂ and 2'-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 173.8 (C-1); 172.7 (CH₂COOC(CH₃)₃); 170.2 (C-4); 170.0 (C-7); 157.3 (NHCOOC(CH₃)₃); 80.6 (CH₂COOC(CH₃)₃); 78.6 (NHCOOC(CH₃)₃); 52.7 (C-5); 51.6 (C-2); 39.9 (C-4'); 31.3 (C-1'); 30.9 (C-2''); 29.2 (C-3'); 27.7 (NHCOOC(CH₃)₃); 27.3 (C-1''); 27.2 (CH₂COOC(CH₃)₃); 22.0 (C-8); 21.3 (C-2').

4.6.10. 4.6.10 Methyl 4-aminobutanoate hydrochloride (28**)**^{82,83}

Using **procedure F**, the commercially available γ -aminobutyric acid (100 mg, 0.97 mmol) and SOCl₂ (0.7 mL, 9.6 mmol), **28** (140 mg, 0.92 mmol, 95%) was obtained as an off white solid, mp. 164-166°C. MS (ESI⁺), *m/z* 118 (100%) [MH⁺]. HRMS (ESI⁺) calcd for C₅H₁₁NO₂ + H: 118.0863; found 118.0872. ¹H NMR (300 MHz, CD₃OD): δ 3.69 (s, 3H, OCH₃); 2.99 (t, *J* = 7.5 Hz, 2H, 4-CH₂); 2.49 (t, *J* = 7.2 Hz, 2H, 2-CH₂); 2.01-1.89 (m, 2H, 3-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 173.0 (C-1); 51.6 (OCH₃); 39.5 (C-4); 30.5 (C-2); 28.4 (C-3).

4.6.11. Methyl (7*S*,10*S*)-5,8,11-triaza-7-[4-(*tert*-butoxycarbonyl)aminobutyl]-10-(*tert*-butoxycarbonyl)methyl-6,9,12-trioxotridecanoate (29**)**

The ester was synthesised using **procedure C** from the acid **26** (200 mg, 0.44 mmol), the prepared amine hydrochloride **28** (70 mg, 0.46 mmol), EDCI (111 mg, 0.58 mmol), HOBT (87 mg, 0.64 mmol) and DIPEA (0.07 mL, 0.40 mmol), to yield **29** (204 mg, 0.37 mmol, 84%) as an off white powder, mp. 138-140°C. MS (ESI⁺), *m/z* 559 (100%) [MH⁺], 560 (30) [MD⁺], 459 (20). HRMS (ESI⁺) calcd for C₂₆H₄₆N₄O₉ + H: 559.3343; found 559.3358. ¹H NMR (300 MHz, CDCl₃): δ 6.99 (d, *J* = 7.95 Hz, 1H, 8-NH); 6.93-6.83 (m, 2H, 5-NH and 11-NH); 4.70-4.65 (m, 1H, 10-CH); 4.30-4.27 (m, 1H, 7-CH); 3.60 (s, 3H, OCH₃); 3.24-

3.15 (m, 2H, 4-CH₂); 3.01 (t, *J* = 6.7, Hz, 2H, 4'-CH₂); 2.71-2.68 (m, 2H, 10-CHCH₂); 2.28 (t, *J* = 7.3, Hz, 2H, 2-CH₂); 1.98 (s, 3H, 12-CH₃); 1.91-1.82 (m, 1H, 1'-CH_aH_b); 1.81-1.73 (m, 2H, 3-CH₂); 1.62-1.52 (m, 1H, 1'-CH_aH_b); 1.44-1.39 (m, 2H, 3'-CH₂); 1.39-1.32 (m, 18H, NHCOOC(CH₃)₃, CH₂COOC(CH₃)₃); 1.30-1.23 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CDCl₃): δ 173.9 (C-1); 171.5 (CH₂COOC(CH₃)₃); 171.3 (C-9); 171.1 (C-6) 170.8 (C-12); 156.3 (NHCOOC(CH₃)₃); 82.2 (CH₂COOC(CH₃)₃); 79.1 (NHCOOC(CH₃)₃); 53.5 (C-7); 51.9 (OCH₃); 50.1 (C-10); 40.7 (C-4'); 39.8 (C-4); 37.0 (10-CHCH₂); 31.6 (C-1'); 31.5 (C-2 and C-3'); 28.7 (NHCOOC(CH₃)₃); 28.2 (CH₂COOC(CH₃)₃); 24.7 (C-3); 23.3 (C-13); 22.8 (C-2').

4.6.12. Methyl (7*S*,10*S*)-5,8,11-triaza-7-[4-(*tert*-butoxycarbonyl)aminobutyl]-10-[2-(*tert*-butoxycarbonyl)ethyl]-6,9,12-trioxotridecanoate (30**)**

The ester was synthesised using **procedure C** from the acid **27** (98 mg, 0.21 mmol), the prepared amine hydrochloride **28** (33 mg, 0.22 mmol), EDCI (61 mg, 0.32 mmol), HOBT (43 mg, 0.732 mmol) and DIPEA (0.05 mL, 0.29 mmol), to yield **30** (109 mg, 0.19 mmol, 92%) as a white powder, mp. 136-139°C. MS (ESI⁺), *m/z* 573 (12%) [MH⁺], 473 (5), 142 (25), 115 (100). HRMS (ESI⁺) calcd for C₂₇H₄₈N₄O₉ + H: 573.3494; found 573.3499. ¹H NMR (500 MHz, CDCl₃): δ 7.17 (d, *J* = 7.8 Hz, 1H, 8-NH); 6.86-6.78 (m, 2H, 5-NH and 11-NH); 4.63-4.57 (m, 1H, 10-CH); 4.23-4.19 (m, 1H, 7-CH); 4.05 (bs, 1H, 4'-CH₂NH); 3.66 (s, 3H, OCH₃); 3.24-3.15 (m, 2H, 4-CH₂); 3.01 (t, *J* = 6.5, Hz, 2H, 4'-CH₂); 2.41-2.22 (m, 4H, 2-CH₂ and 2''-CH₂); 2.06-1.84 (m, 5H, 13-CH₃ and 1''-CH₂); 1.83-1.72 (m, 3H, 1'-CH_aH_b and 3-CH₂); 1.69-1.55 (m, 1H, 1'-CH_aH_b); 1.48-1.34 (m, 20H, NHCOOC(CH₃)₃, 2''-CH₂COOC(CH₃)₃ and 3'-CH₂); 1.33-1.22 (m, 2H, 2'-CH₂). ¹³C NMR (126 MHz, CDCl₃): δ 173.6 (C-1); 172.8 (CH₂COOC(CH₃)₃); 171.5 (C-9); 171.3 (C-6); 170.9 (C-12); 156.4 (NHCOOC(CH₃)₃); 81.0 (CH₂COOC(CH₃)₃); 79.0 (NHCOOC(CH₃)₃); 53.1 (C-10); 53.0 (C-7); 51.6 (OCH₃); 40.1 (C-4'); 38.8 (C-4); 32.1 (C-2''); 31.7 (C-1'); 31.3 (C-2); 29.3 (C-3'); 28.4 (NHCOOC(CH₃)₃); 28.0 (CH₂COOC(CH₃)₃); 27.6 (C-1''); 24.4 (C-3); 22.9 (C-13); 22.7 (C-2').

4.6.13. (3*S*,6*S*)-3-Acetamido-6-(4-aminobutyl)-5,8-diaza-12-methoxy-4,7,12-trioxododecanoic acid hydrochloride (8**)**

Compound **29** (70 mg, 0.12 mmol) was converted to the uncharacterised *N*-Boc deprotected trifluoroacetate salt via **procedure D**, the resulting solid was then converted, via **procedure E**, using acetonitrile as the solvent, to give the hydrochloride salt **8** (53 mg, 0.10 mmol, 84%) as a hygroscopic brown amorphous solid. MS (ESI⁺), *m/z* 403 (100%) [MH⁺], 404 (20). HRMS (ESI⁺) calcd for C₁₇H₃₀N₄O₇ + H: 403.2193; found 403.2192. [α]_D²⁵ +41.1 (c. 0.09, EtOH). ¹H NMR (300 MHz, CD₃OD): δ 4.64 (dd, *J* = 11.9, 6.1 Hz, 1H, 3-CH); 4.34-4.23 (m, 1H, 6-CH); 3.65 (s, 3H, OCH₃); 3.28-3.15 (m, 2H, 9-CH₂); 3.00-2.85 (m, 3H, 4'-CH₂ and 2-CH_aH_b); 2.84-2.70 (m, 1H, 2-CH_aH_b); 2.35 (t, *J* = 6.9 Hz, 2H, 11-CH₂); 2.04-1.86 (m, 4H, 1'-CH_aH_b and COCH₃); 1.83-1.61 (m, 5H, 1'-CH_aH_b, 3'-CH₂ and 10-CH₂); 1.54-1.39 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 175.4 (C-1); 174.2 (C-12); 174.0 (C-4); 173.7 (C-7); 173.5 (COCH₃); 54.7 (C-6); 52.2 (OCH₃); 51.8 (C-3); 40.6 (C-4'); 39.7 (C-9); 36.5 (C-2); 32.0 (C-1' and C-3'); 27.8 (C-11); 25.5 (C-10); 23.8 (COCH₃); 22.7 (C-2').

4.6.14. (4*S*,7*S*)-4-Acetamido-7-(4-aminobutyl)-6,9-diaza-13-methoxy-5,8,13-trioxotridecanoic acid hydrochloride (9**)**

Compound **30** (85 mg, 0.15 mmol) was converted to the *N*-Boc deprotected trifluoroacetate salt via **procedure D**, and the

resulting solid was then converted, *via* procedure E, using acetonitrile as the solvent, to give the hydrochloride salt **9** (50 mg, 0.11 mmol, 74%) as a hygroscopic light brown amorphous solid. MS (ESI⁺), *m/z* 417 (100%) [MH⁺], 418 (20) [MD⁺], 143 (17). HRMS (ESI⁺) calcd for C₁₈H₃₂N₄O₇ + H: 417.2349; found 417.2361. [α]_D²⁵ -67.4 (c. 0.17, EtOH). ¹H NMR (300 MHz, CD₃OD): δ 4.39-4.20 (m, 2H, 4-CH and 7-CH); 3.64 (s, 3H, OCH₃); 3.30-3.15 (m, 2H, 10-CH₂); 2.98-2.85 (m, 2H, 4'-CH₂); 2.48-2.27 (m, 4H, 2-CH₂ and 12-CH₂); 2.10-1.91 (m, 5H, 3-CH₂ and COCH₃); 1.84-1.60 (m, 6H, 1'-CH₂, 3'-CH₂ and 11-CH₂); 1.49-1.36 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 175.0 (C-1); 174.5 (C-13); 173.9 (C-5 and C-8); 172.4 (COCH₃); 54.6 (C-4); 54.2 (C-7); 52.3 (OCH₃); 41.0 (C-4'); 39.6 (C-10); 32.3 (C-2); 31.9 (C-1'); 31.8 (C-12); 27.9 (C-3 and C-3'); 25.5 (C-11); 23.8 (COCH₃); 22.7 (C-2').

4.6.15. Methyl (7*S*,10*S*)-7-(4-aminobutyl)-5,8,11-triaza-10-methoxycarbonylmethyl-6,9,12-trioxotridecanoate hydrochloride (**10**)

Using procedure F, **8** (32 mg, 0.08 mmol) and SOCl₂ (0.25 mL, 3.43 mmol), **10** (14 mg, 0.03 mmol, 90%) was obtained as an off white hygroscopic amorphous solid. MS (ESI⁺), *m/z* 417 (5%) [MH⁺], 142 (35), 102 (100). HRMS (ESI⁺) calcd for C₁₈H₃₂N₄O₇ + H: 417.2344; found 417.2351. [α]_D²⁵ +62.8 (c. 0.11, EtOH). ¹H NMR (300 MHz, CD₃OD): δ 4.61-4.52 (m, 1H, 10-CH); 4.34-4.23 (m, 1H, 7-CH); 3.68 (s, 6H, 1-COOCH₃ and 10-CHCH₂COOCH₃); 3.29-3.14 (m, 2H, 4-CH₂); 2.99-2.84 (m, 3H, 4'-CH₂ and 10-CHCH₂H_b); 2.83-2.71 (m, 1H, 10-CHCH₂H_b); 2.32 (t, *J* = 6.8 Hz, 2H, 2-CH₂); 2.07-1.88 (m, 4H, 1'-CH₂H_b and 13-CH₂); 1.82-1.62 (m, 5H, 1'-CH₂H_b, 3'-CH₂ and 3-CH₂); 1.51-1.36 (m, 2H, 2'-CH₂). ¹³C NMR (75 MHz, CD₃OD): δ 175.3 (C-1 and 10-CHCH₂COOCH₃); 174.0 (C-9); 173.0 (C-6); 172.3 (C-12); 54.3 (C-17); 52.1 (1-COOCH₃ and 10-CHCH₂COOCH₃); 51.9 (C-7); 41.0 (C-4'); 39.8 (C-4); 36.6 (10-CHCH₂); 31.9 (C-1'); 31.4 (C-3'); 28.1 (C-2); 25.6 (C-3); 23.3 (C-13); 22.4 (C-2').

4.6.16. *N*-[(2-Hydroxyethyl)-*N*-(4-nitrophenyl)]-3-aminopropanenitrile (**35**)

The flow system (outlined in Fig. 6), including the omnifit column packed with K₂CO₃ was initially primed with DMSO whilst the volcano obtained a temperature of 160 °C. Once primed, syringe pump A was fitted with a syringe containing 10 mL of 1.6 M ethanolamine in DMSO and flow rate was set to 0.25 mL/min. Syringe pump B was fitted with a syringe containing 10 mL of 1.6 M 3-butenenitrile in DMSO and flow rate was set to 0.25 mL/min. Syringe pump C was fitted with a syringe containing 20 mL of 1.8 M of 1-fluoro-4-nitrobenzene in DMSO and flow rate was set to 0.3 mL/min. Syringe pumps A and B were initialised and upon the mixture eluting from the microreactor (*t_R* = 3.5 mins, monitored at 214 nm) syringe pump C was initiated to introduce the 1.8 M solution of 1-fluoro-4-nitrobenzene into the reaction stream. The reaction stream was continued through and omitted column packed with K₂CO₃ heated to 160 °C with the eluent was collected in a Erlenmeyer flask containing 200 mL iced H₂O. Once the entirety of 10 mL solution from syringe pump A and B was delivered, both syringes were reloaded with 5 mL of DMSO and the pumps were reinitiated to flush the system. Once the system flush was complete the Erlenmeyer flask containing 200 mL iced H₂O attained room temperature and the resulting yellow precipitate was collected to afford **35** (2.73 g, 70 % overall) as a bright yellow crystalline solid (91-92 °C). IR ν 3448, br/m; 2990, m; 2867, m; 2244, m; 1739, m; 1640, w; 1470, w; 1439, w; 1366, s; 1234, w; 1140, m; 1056, s; 1035, s; 850, m; 794, w. MS (ESI⁺) *m/z* 236 (M+1, 100%); HRMS (ESI⁺) for C₁₁H₁₄N₃O₃; calculated 236.1026, found 236.1035. ¹H NMR (DMSO *d*₆) (300 MHz) δ 8.06, 2H, d, *J* = 9.6 Hz, ArH3, ArH5; 6.93, 2H, d, *J* = 9.6 Hz,

ArH2, ArH6; 4.93, 1H, t, *J* = 5.4 Hz, OH; 3.86, 2H, t, *J* = 6.6 Hz, CH₂CH₂C≡N; 3.62, 4H, bs, CH₂CH₂O; 2.85, 2H, t, *J* = 6.6 Hz, NCH₂CH₂C≡N. ¹³C NMR (DMSO *d*₆) (75 MHz) δ 153.2, ArC1; 136.8, ArC4; 126.6, ArC3, ArC5; 119.9, C≡N; 111.9, ArC2, ArC6; 58.6, CH₂CH₂C≡N; 53.3, NCH₂CH₂O; 47.3, CH₂CH₂O; 15.8, CH₂CH₂C≡N.

4.6.17. *N*-[(2-Acetyloxyethyl)-*N*-(4-nitrophenyl)]-3-aminopropionitrile (**37**)

A solution of **35** (4.40 g, 18.70 mmol), acetic anhydride (2.00 mL, 21.16 mmol), and Et₃N (2.50 mL, 43.03 mmol) was stirred at rt for 16 h before being diluted with water (200 mL). The resulting precipitate was collected and subjected to flash silica gel column chromatography (1:1 EtOAc/hexanes) to afford *N*-[(2-acetyloxyethyl)-*N*-(4-nitrophenyl)]-3-aminopropionitrile (**37**) (0.62 g, 92%) as a bright yellow crystalline solid, m.p. 98-99 °C. IR ν 2960, m; 2884, m; 2244, m; 1731, s; 1598, m; 1521, w; 1480, w; 1306, m; 1296, m; 1234, m; 1204, m; 1120, s; 1045, m; 835, w; 759, m. MS (ESI⁺) *m/z* 278 (M+1, 100%); HRMS (ESI⁺) for C₁₃H₁₆N₃O₅; calculated 278.1143, found 278.1141. ¹H NMR (CDCl₃) (300 MHz) δ 8.08, 2H, d, *J* = 9.3 Hz, ArH3, ArH5; 6.71, 2H, d, *J* = 9.3 Hz, ArH2, ArH6; 4.28, 2H, t, *J* = 5.7 Hz, CH₂CH₂O; 3.83, 2H, t, *J* = 7.2 Hz, CH₂CH₂C≡N; 3.79, 2H, t, *J* = 5.7 Hz, CH₂CH₂O; 2.72, 2H, t, *J* = 7.2 Hz, CH₂CH₂C≡N; 2.03, 3H, s, CH₃. ¹³C-NMR (CDCl₃) (75 MHz) δ 166.3, C=O; 146.8, ArC1; 133.7, ArC4; 121.8, ArC3, ArC5; 113.2, C≡N; 106.4, ArC2, ArC6; 56.4, CH₂CH₂O; 45.3, CH₂CH₂O; 42.7, CH₂CH₂C≡N; 16.3, CH₃; 11.3, CH₂CH₂C≡N.

4.6.18. *N*-[(2-Acetyloxyethyl)-*N*-(4-aminophenyl)]-3-aminopropionitrile (**31**)

A 0.05 M solution of **37** (0.62 g, 2.24 mmol) in EtOH (44.0 mL) was flowed (0.5 mL/min) through a H-cube Pro fitted with a 70 mm 10% Pd/C CatCart column which was heated to 50 °C under 50 bar with full H₂. Once the entirety of the 0.05 M solution of **37** was delivered an addition 10 mL of EtOH was passed through the system and the resulting eluent was concentrated *in vacuo* to afford **31** (0.54 g, 98%) as a light brown oil. (References^{84,85} cite **31** with no spectral data described). IR ν 3447, br/m; 2960, s; 2884, m; 2243, m; 1737, m; 1639, w; 1470, w; 1429, m; 1368, m; 1235, m; 1061, s; 1034, s; 845, w; 753, m. MS (ESI⁺) *m/z* 248 (M+1, 100%); HRMS (ESI⁺) for C₁₃H₁₈N₃O₂; calculated 248.1401, found 248.1399. ¹H NMR (CDCl₃) (500 MHz) δ 6.66, 2H, d, *J* = 8.9 Hz, ArH2, ArH6; 6.61, 2H, d, *J* = 8.9 Hz, ArH3, ArH5; 4.12, 2H, t, *J* = 5.6 Hz, CH₂CH₂O; 3.66, 2H, bs, NH₂; 3.49, 2H, t, *J* = 6.1 Hz, CH₂CH₂C≡N; 3.42, 2H, t, *J* = 6.1 Hz, CH₂CH₂O; 2.46, 2H, t, *J* = 6.8 Hz, CH₂CH₂C≡N; 2.00, 3H, s, CH₃. ¹³C NMR (CDCl₃) (125 MHz) δ 171.2, C=O; 140.2, ArC1; 139.5, ArC4; 118.9, C≡N; 118.2, ArC2, ArC6; 116.9, ArC3, ArC5; 62.2, CH₂CH₂O; 51.9, CH₂CH₂O; 49.2, CH₂CH₂CN; 21.1, CH₂CH₂CN; 16.5, CH₃.

4.6.19. *N*-[4-(2-Acetyloxyethylamino)-(2-cyanoethylamino)anilino]-3-nitrobenzamide (**6**)

A solution of **31** (0.26 g, 1.06 mmol), 3-nitrobenzoic acid (0.19 g, 1.17 mmol), DCC (0.24 g, 1.17 mmol), DMAP (0.14 g, 1.17 mmol), and DMF (10.00 mL) was stirred under an atmosphere of nitrogen for 16 h. The reaction was then diluted with water (100 mL), extracted with EtOAc (3 × 25 mL), dried (MgSO₄), concentrated *in vacuo*, and the crude subjected to flash silica gel column chromatography (2:1 EtOAc:Hex) to afford *N*-[4-(2-acetyloxyethylamino)-(2-cyanoethylamino)anilino]-3-nitrobenzamide **6** (0.33 g, 77%) as a dark yellow solid (m.p. 137-138 °C). IR ν 2960, m; 2884, m; 2244, m; 1731, s; 1680, s; 1598, m; 1521, w; 1480, w; 1306, m; 1296, m; 1234, m; 1204, m; 1120, s; 1065, m; 845, m. MS (ESI⁺) *m/z* 370 (M+1, 100%); HRMS (ESI⁺) for C₂₀H₂₁N₄O₅; calculated 397.1514, found 397.1512. ¹H

NMR (CDCl₃) (500 MHz) δ 8.69, 1H, s, ArH2; 8.36, 1H, d, J = 7.5 Hz, ArH6; 8.30, 1H, bs, NH; 8.25, 1H, d, J = 7.5 Hz, ArH4; 7.66, 1H, dd, J = 8.3, 8.3 Hz, ArH5; 7.51, 2H, d, J = 8.5 Hz, ArH3', ArH5'; 6.70, 2H, d, J = 8.5 Hz, ArH2', ArH6'; 4.23, 2H, t, J = 5.6 Hz, CH₂CH₂O; 3.72, 2H, t, J = 6.5 Hz, CH₂CH₂C≡N; 3.65, 2H, t, J = 6.5 Hz, CH₂CH₂O; 2.63, 2H, t, J = 6.5 Hz, CH₂CH₂C≡N; 2.05, 3H, s, CH₃. ¹³C NMR (CDCl₃) (125 MHz) δ 171.2, C=O; 163.5, NC=O; 148.4, ArC3; 143.9, ArC4'; 136.8, ArC1; 133.7, ArC4; 130.2, ArC5; 128.6, ArC1'; 126.3, ArC6; 123.2, ArC2', ArC6'; 122.1, ArC2; 118.5, C≡N; 113.4, ArC3', ArC5'; 61.7, CH₂CH₂O; 50.4, CH₂CH₂O; 47.9, CH₂CH₂C≡N; 21.1, CH₃; 16.3, CH₂CH₂C≡N.

4.6.20. *N*-[4-(2-Acetyloxyethylamino)-(2-cyanoethylamino)anilino]-2-chloro-4-nitrobenzamide (**11**)

Compound **11** was prepared using the procedure described for **6** using **31** (0.11, 0.44 mmol), 2-chloro-4-nitrobenzoic acid (0.13 g, 0.66 mmol), DCC (0.14 g, 0.66 mmol), DMAP (0.08 g, 0.66 mmol), and DMF (10.00 mL). The crude was subjected to flash silica gel column chromatography (1:1 EtOAc/hexanes) to afford *N*-[4-(2-acetyloxyethylamino)-(2-cyanoethylamino)anilino]-2-chloro-4-nitrobenzamide **11** (0.12 g, 62%) as an orange crystalline solid (m.p. 129-131 °C). IR ν 2965, m; 2894, m; 2240, m; 1731, s; 1680, s; 1589, m; 1521, w; 1480, w; 1306, m; 1296, m; 1234, m; 1204, m; 1120, s; 1045, m; 835, m; 759, s. MS (ESI+) m/z 431 (M+1, 100%); HRMS (ESI+) for C₂₀H₂₀ClN₄O₅; calculated 431.1124, found 431.1122. ¹H NMR (CDCl₃) (300 MHz) δ 8.34, 1H, d, J = 2.4 Hz, ArH3; 8.22, 1H, dd, J = 8.7 Hz, ArH5; 7.96, 1H, bs, NH; 7.90, 1H, d, J = 8.7 Hz, ArH6; 7.54, 2H, d, J = 8.7 Hz, ArH2', ArH6'; 6.78, 2H, d, J = 8.7 Hz, ArH3', ArH5'; 4.28, 2H, t, J = 6.0 Hz, CH₂CH₂O; 3.78, 2H, t, J = 6.9 Hz, CH₂CH₂C≡N; 3.73, 2H, t, J = 6.0 Hz, CH₂CH₂O; 2.68, 2H, t, J = 6.9 Hz, CH₂CH₂C≡N; 2.10, 3H, s, CH₃. ¹³C NMR (CDCl₃) (75 MHz) δ 171.1, C=O; 162.8, NC=O; 149.0, ArC4; 144.1, ArC4'; 141.3, ArC1; 132.3, ArC1'; 131.2, ArC6; 128.1, ArC2; 125.7, ArC3; 122.8, ArC2', ArC6'; 122.4, ArC5, 118.3, C≡N; 113.3, ArC3', ArC5'; 61.6, CH₂CH₂O; 50.4, CH₂CH₂O; 47.9, CH₂CH₂C≡N; 21.1, CH₃; 16.2, CH₂CH₂C≡N.

4.6.21. *N*-(2-Acetyloxyethyl)-*N*-(4-[2,6-dichlorobenzylideneamino]phenyl)-3-aminopropanenitrile (**12**)

A solution of **31** (0.14 g, 0.57 mmol), 2,6-dichlorobenzaldehyde (0.10 g, 0.57 mmol), 3Å molecular sieves (0.25 g), and anhydrous MeOH (15 mL) was stirred at rt for 18 h. The reaction mixture was then diluted with EtOAc (100 mL), filtered through celite, dried (MgSO₄) and concentrated *in vacuo* to afford **12** (0.13 g, 55%) as a dark yellow semi solid. IR ν 2960, s; 2884, m; 2243, m; 1737, s; 1690, m; 1470, w; 1429, m; 1368, m; 1235, m; 1061, s; 1034, s; 845, w; 768, s. MS (ESI+) m/z 404 (M+1, 100%); HRMS (ESI+) for C₂₀H₂₀Cl₂N₃O₂; calculated 404.0934, found 404.0933. ¹H NMR (CDCl₃) (500 MHz) δ 8.69, 1H, s, CH=N; 7.38, 2H, d, J = 7.5 Hz, ArH3', ArH5'; 7.32, 2H, d, J = 9.0 Hz, ArH3, ArH5; 7.25, 1H, t, J = 9.0 Hz, ArH4'; 6.78, 2H, d, J = 9.0 Hz, ArH2, ArH6; 4.27, 2H, t, J = 6.0 Hz, CH₂CH₂O; 3.77, 2H, t, J = 6.5 Hz, CH₂CH₂C≡N; 3.70, 2H, t, J = 6.0 Hz, CH₂CH₂O; 2.65, 2H, t, J = 6.5 Hz, CH₂CH₂C≡N; 2.06, 3H, s, CH₃. ¹³C NMR (CDCl₃) (125 MHz) δ 171.1, C=O; 153.3, N=CH; 145.6, ArC1; 142.0, ArC4; 135.4, ArC2', AC6'; 133.2, ArC1'; 130.6, ArC4'; 129.1, ArC3', ArC5'; 123.2, ArC3, ArC5; 118.3, C≡N; 113.4, ArC2, ArC6; 61.7, CH₂CH₂O; 50.5, CH₂CH₂O; 47.9, CH₂CH₂C≡N; 21.1, CH₃; 16.2, CH₂CH₂C≡N.

Acknowledgments

DMAP/CPG was the recipient of an ARC DECRA fellowship (DE130100513) and acknowledges the postgraduate scholarship provided the University of Wollongong and the ARC Linkage program. ND acknowledges the University of Wollongong and Aveva for a matching scholarship.

References

- Samji, H.; Cescon, A.; Hogg, R. S.; Modur, S. P.; Althoff, K. N.; Buchacz, K.; Burchell, A. N.; Cohen, M.; Gebo, K. A.; Gill, M. J.; Justice, A.; Kirk, G.; Klein, M. B.; Korthuis, P. T.; Martin, J.; Napravnik, S.; Rourke, S. B.; Sterling, T. R.; Silverberg, M. J.; Deeks, S.; Jacobson, L. P.; Bosch, R. J.; Kitahata, M. M.; Goedert, J. J.; Moore, R.; Gange, S. J. *PLoS One* **2013**, *8*, e81355/1-e81355/8, 8 pp.
- Stebbing, J.; Bower, M.; Mandalia, S.; Nelson, M.; Gazzard, B. *Clin. Exp. Immunol.* **2006**, *145*, 271-276.
- Hammer, S. M.; Eron, J. J., Jr.; Reiss, P.; Schooley, R. T.; Thompson, M. A.; Walmsley, S.; Cahn, P.; Fischl, M. A.; Gatell, J. M.; Hirsch, M. S.; Jacobsen, D. M.; Montaner, J. S. G.; Richman, D. D.; Yeni, P. G.; Volberding, P. A. *JAMA, J. Am. Med. Assoc.* **2008**, *300*, 555-570.
- Clavel, F.; Hance, A. J. *N. Engl. J. Med.* **2004**, *350*, 1023-1035.
- Lahsasni, S.; Ben Hadda, T.; Masand, V.; Pathan, N. B.; Parvez, A.; Warad, I.; Shaheen, U.; Bader, A.; Aljofan, M. *Res. Chem. Intermed.* **2015**, *41*, 5121-5136.
- Gordon, C. P.; Griffith, R.; Keller, P. A. *Med. Chem.* **2007**, *3*, 199-220.
- Delelis, O.; Carayon, K.; Saib, A.; Deprez, E.; Mouscadet, J.-F. *Retrovirology* **2008**, *5*, No pp given.
- Cherepanov, P.; Ambrosio, A. L. B.; Rahman, S.; Ellenberger, T.; Engelman, A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17308-17313.
- Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. *Nat. Chem. Biol.* **2010**, *6*, 442-448.
- Fader, L. D.; Malenfant, E.; Parisien, M.; Carson, R.; Bilodeau, F.; Landry, S.; Pesant, M.; Brochu, C.; Morin, S.; Chabot, C.; Halmos, T.; Bousquet, Y.; Bailey, M. D.; Kawai, S. H.; Coulombe, R.; LaPlante, S.; Jakalian, A.; Bhardwaj, P. K.; Wernic, D.; Schroeder, P.; Amad, M. a.; Edwards, P.; Garneau, M.; Duan, J.; Cordingley, M.; Bethell, R.; Mason, S. W.; Bos, M.; Bonneau, P.; Poupart, M.-A.; Faucher, A.-M.; Simoneau, B.; Fenwick, C.; Yoakim, C.; Tsantrizos, Y. *ACS Med. Chem. Lett.* **2014**, *5*, 422-427.
- Pannecouque, C.; Pluymers, W.; Van Maele, B.; Tetz, V.; Cherepanov, P.; De Clercq, E.; Witvrouw, M.; Debyser, Z. *Curr. Biol.* **2002**, *12*, 1169-1177.
- Patel, D.; Antwi, J.; Koneru, P. C.; Serrao, E.; Forli, S.; Kessler, J. J.; Feng, L.; Deng, N.; Levy, R. M.; Fuchs, J. R.; Olson, A. J.; Engelman, A. N.; Bauman, J. D.; Kvaratskhelia, M.; Arnold, E. J. *Biol. Chem.* **2016**, *291*, 23569-23577.
- Australian Federation of AIDS Organisations (AFAO), *HIV in Australia: 2017*. **2017**, <https://www.afao.org.au/wp-content/uploads/2017/04/HIV-in-Australia-2017.pdf>.
- Rafiei, N.; Woolley, I.; Korman Tony, M.; Chibo, D.; Gooley, M.; McMahon James, H. *AIDS* **2017**, *31*, 598-600.
- Wensing Annemarie, M.; Calvez, V.; Gunthard Huldrych, F.; Johnson Victoria, A.; Paredes, R.; Pillay, D.; Shafer Robert, W.; Richman Douglas, D. *Top. Antivir. Med.* **2014**, *22*, 642-50.
- Fenwick, C.; Amad, M. a.; Bailey, M. D.; Bethell, R.; Bos, M.; Bonneau, P.; Cordingley, M.; Coulombe, R.; Duan, J.; Edwards, P.; Fader, L. D.; Faucher, A.-M.; Garneau, M.; Jakalian, A.; Kawai, S.; Lamorte, L.; LaPlante, S.; Luo, L.; Mason, S.; Poupart, M.-A.; Rioux, N.; Schroeder, P.; Simoneau, B.; Tremblay, S.; Tsantrizos, Y.; Witvrouw, M.; Yoakim, C. *Antimicrob. Agents Chemother.* **2014**, *58*, 3233-3244, 13 pp.
- Dalton, N.; Gordon, C. P.; Boyle, T. P.; Vandegraaf, N.; Deadman, J.; Rhodes, D. I.; Coates, J. A.; Pyne, S. G.; Keller, P. A.; Bremner, J. B. *Org. Biomol. Chem.* **2016**, *14*, 6010-6023.
- Patel, P. A.; Kvaratskhelia, N.; Mansour, Y.; Antwi, J.; Feng, L.; Koneru, P.; Kobe, M. J.; Jena, N.; Shi, G.; Mohamed, M. S.; Li, C.; Kessler, J. J.; Fuchs, J. R. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 4748-4752.
- Greenwald, J.; Le, V.; Butler, S. L.; Bushman, F. D.; Choe, S. *Biochemistry* **1999**, *38*, 8892-8898.

20. Galilee, M.; Britan-Rosich, E.; Griner, S. L.; Uysal, S.; Baumgartel, V.; Lamb, D. C.; Kossiakoff, A. A.; Kotler, M.; Stroud, R. M.; Marx, A.; Alian, A. *Structure (Oxford, U. K.)* **2016**, *24*, 1936-1946.
21. Wang, J.-Y.; Ling, H.; Yang, W.; Craigie, R. *EMBO J.* **2001**, *20*, 7333-7343.
22. Chen, I. J.; Neamati, N.; MacKerell, A. D., Jr. *Curr. Drug Targets Infect. Disord.* **2002**, *2*, 217-234.
23. Bujacz, G.; Alexandratos, J.; Zhou-Liu, Q.; Clement-Mella, C.; Wlodawer, A. *FEBS Lett.* **1996**, *398*, 175-178.
24. Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Engelman, A.; Craigie, R.; Davies, D. R. *Science (Washington, D. C.)* **1994**, *266*, 1981-6.
25. Chen, J. C. H.; Krucinski, J.; Miercke, L. J. W.; Finer-Moore, J. S.; Tang, A. H.; Leavitt, A. D.; Stroud, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8233-8238.
26. Wielens, J.; Headey, S. J.; Jeevarajah, D.; Rhodes, D. I.; Deadman, J.; Chalmers, D. K.; Scanlon, M. J.; Parker, M. W. *FEBS Lett.* **2010**, *584*, 1455-1462.
27. Maignan, S.; Guilloteau, J.-P.; Zhou-Liu, Q.; Clement-Mella, C.; Mikol, V. *J. Mol. Biol.* **1998**, *282*, 359-368.
28. Le Rouzic, E.; Bonnard, D.; Chasset, S.; Bruneau, J.-M.; Chevreuril, F.; Le Strat, F.; Nguyen, J.; Beauvoir, R.; Amadori, C.; Brias, J.; Vomscheid, S.; Eiler, S.; Levy, N.; Delelis, O.; Deprez, E.; Saib, A.; Zamborlini, A.; Emiliani, S.; Ruff, M.; Ledoussal, B.; Moreau, F.; Benarous, R. *Retrovirology* **2013**, *10*, 144/1-144/21, 21 pp.
29. Goldgur, Y.; Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Craigie, R.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 9150-9154.
30. Goldgur, Y.; Craigie, R.; Cohen, G. H.; Fujiwara, T.; Yoshinaga, T.; Fujishita, T.; Sugimoto, H.; Endo, T.; Murai, H.; Davies, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13040-13043.
31. Molteni, V.; Greenwald, J.; Rhodes, D.; Hwang, Y.; Kwiatkowski, W.; Bushman, F. D.; Siegel, J. S.; Choe, S. *Acta Crystallogr., Sect. D Biol. Crystallogr.* **2001**, *D57*, 536-544.
32. Kessler, J. J.; Jena, N.; Koh, Y.; Taskent-Sezgin, H.; Slaughter, A.; Feng, L.; de Silva, S.; Wu, L.; Le Grice, S. F. J.; Engelman, A.; Fuchs, J. R.; Kvaratskhelia, M. *J. Biol. Chem.* **2012**, *287*, 16801-16811.
33. Tsiang, M.; Jones, G. S.; Niedziela-Majka, A.; Kan, E.; Lansdon, E. B.; Huang, W.; Hung, M.; Samuel, D.; Novikov, N.; Xu, Y.; Mitchell, M.; Guo, H.; Babaoglu, K.; Liu, X.; Geleziunas, R.; Sakowicz, R. *J. Biol. Chem.* **2012**, *287*, 21189-21203.
34. Feng, L.; Sharma, A.; Slaughter, A.; Jena, N.; Koh, Y.; Shkriabai, N.; Larue, R. C.; Patel, P. A.; Mitsuya, H.; Kessler, J. J.; Engelman, A.; Fuchs, J. R.; Kvaratskhelia, M. *J. Biol. Chem.* **2013**, *288*, 15813-15820.
35. Jurado, K. A.; Wang, H.; Slaughter, A.; Feng, L.; Kessler, J. J.; Koh, Y.; Wang, W.; Ballandras-Colas, A.; Patel, P. A.; Fuchs, J. R.; Kvaratskhelia, M.; Engelman, A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8690-8695, S8690/1-S8690/7.
36. Sharma, A.; Slaughter, A.; Feng, L.; Kessler, J. J.; Kvaratskhelia, M.; Jena, N.; Fuchs, J. R.; Fadel, H. J.; Poeschla, E.; Malani, N.; Male, F.; Bushman, F. D.; Wu, L. *PLoS Pathog.* **2014**, *10*, e1004171.
37. Peat, T. S.; Dolezal, O.; Newman, J.; Mobley, D.; Deadman, J. J. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 347-362.
38. Slaughter, A.; Jurado, K. A.; Deng, N.; Feng, L.; Kessler, J. J.; Shkriabai, N.; Larue, R. C.; Fadel, H. J.; Patel, P. A.; Jena, N.; Fuchs, J. R.; Poeschla, E.; Levy, R. M.; Engelman, A.; Kvaratskhelia, M. *Retrovirology* **2014**, *11*, 100/1-100/28, 28 pp.
39. Gupta, K.; Turkki, V.; Sherrill-Mix, S.; Hwang, Y.; Eilers, G.; Taylor, L.; McDanal, C.; Wang, P.; Temelkoff, D.; Nolte, R. T.; Velthuisen, E.; Jeffrey, J.; Van Duyne, G. D.; Bushman, F. D. *PLoS Biol.* **2016**, *14*, e1002584/1-e1002584/24.
40. Rhodes, D. I.; Peat, T. S.; Vandegraaff, N.; Jeevarajah, D.; Le, G.; Jones, E. D.; Smith, J. A.; Coates, J. A. V.; Winfield, L. J.; Thienthong, N.; Newman, J.; Lucent, D.; Ryan, J. H.; Savage, G. P.; Francis, C. L.; Deadman, J. J. *Antiviral Chem. Chemother.* **2011**, *21*, 155-168.
41. Wielens, J.; Headey, S. J.; Deadman, J. J.; Rhodes, D. I.; Le, G. T.; Parker, M. W.; Chalmers, D. K.; Scanlon, M. J. *ChemMedChem* **2011**, *6*, 258-61.
42. Rhodes, D. I.; Peat, T. S.; Vandegraaff, N.; Jeevarajah, D.; Newman, J.; Martyn, J.; Coates, J. A. V.; Ede, N. J.; Rea, P.; Deadman, J. J. *ChemBioChem* **2011**, *12*, 2311-2315.
43. Peat, T. S.; Rhodes, D. I.; Vandegraaff, N.; Le, G.; Smith, J. A.; Clark, L. J.; Jones, E. D.; Coates, J. A. V.; Thienthong, N.; Newman, J.; Dolezal, O.; Mulder, R.; Ryan, J. H.; Savage, G. P.; Francis, C. L.; Deadman, J. J. *PLoS One* **2012**, *7*, e40147.
44. Wielens, J.; Headey, S. J.; Rhodes, D. I.; Mulder, R.; Ryan, J. H.; Savage, G. P.; Francis, C. L.; Deadman, J. J. *PLoS One* **2012**, *7*, e40147.
45. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3WNF> **2013**, to be published.
46. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=4y1c> **2015**, to be published.
47. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=4y1d> **2015**, to be published.
48. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3WNE> **2013**, to be published.
49. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3wng> **2013**, to be published.
50. Northfield, S. E.; Wielens, J.; Headey, S. J.; Mulcair, M.; Scanlon, M. J.; Parker, M. W.; Thompson, P. E.; Chalmers, D. K. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3wnh> **2013**, to be published.
51. Zhao, X. Z.; Smith, S. J.; Maskell, D. P.; Métiouf, M.; Pye, V. E.; Fesen, K.; Marchand, C.; Pommier, Y.; Cherepanov, P.; Hughes, S. H.; Burke, T. R. *J. Med. Chem.* **2017**, *60*, 7315-7332.
52. Ribeiro, A. J. M.; Ramos, M. J.; Fernandes, P. A. *J. Am. Chem. Soc.* **2012**, *134*, 13436-13447.
53. Bhatt, H.; Patel, P.; Pannecouque, C. *Chem. Biol. Drug Des.* **2014**, *83*, 154-166.
54. Gupta, P.; Garg, P.; Roy, N. *Chem. Biol. Drug Des.* **2012**, *79*, 835-849.
55. Reddy, K. K.; Singh, S. K.; Tripathi, S. K.; Selvaraj, C. *SAR QSAR Environ. Res.* **2013**, *24*, 581-595.
56. de Carvalho, L. L.; Maltarollo, V. G.; Ferreira de Lima, E.; Weber, K. C.; Honorio, K. M.; da Silva, A. B. F. *PLoS One* **2014**, *9*, e81301/1-e81301/9, 9 pp.
57. Perryman, A. L.; Forli, S.; Morris, G. M.; Burt, C.; Cheng, Y.-H.; Palmer, M. J.; Whitby, K.; McCammon, J. A.; Phillips, C.; Olson, A. J. *J. Mol. Biol.* **2010**, *397*, 600-615.
58. Gu, W.-g.; Ip, D. T.-M.; Liu, S.-j.; Chan, J. H.; Wang, Y.; Zhang, X.; Zheng, Y.-t.; Wan, D. C.-C. *Chem.-Biol. Interact.* **2014**, *213*, 21-27.
59. Reddy, K. K.; Singh, P.; Singh, S. K. *Mol. Biosyst.* **2014**, *10*, 526-536.
60. Serrao, E.; Xu, Z.-L.; Debnath, B.; Christ, F.; Debyser, Z.; Long, Y.-Q.; Neamati, N. *Bioorg. Med. Chem.* **2013**, *21*, 5963-5972.
61. Serrao, E.; Debnath, B.; Otake, H.; Kuang, Y.; Christ, F.; Debyser, Z.; Neamati, N. *J. Med. Chem.* **2013**, *56*, 2311-2322.
62. De Luca, L.; Morreale, F.; Chimirri, A. *J. Chem. Inf. Model.* **2012**, *52*, 3245-3254.
63. Dayam, R.; Neamati, N. *Bioorg. Med. Chem.* **2004**, *12*, 6371-6381.
64. Anderson, A. C. *Chem. Biol.* **2003**, *10*, 787-797.
65. Bergner, A.; Gunther, J.; Hendlich, M.; Klebe, G.; Verdonk, M. *Biopolymers* **2002**, *61*, 99-110.
66. Sotriffer, C. A.; Ni, H.; McCammon, J. A. *J. Med. Chem.* **2000**, *43*, 4109-4117.
67. Liao, C.; Nicklaus, M. C. *ChemMedChem* **2010**, *5*, 1053-1066.
68. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639-1662.
69. Gupta, K.; Brady, T.; Dyer Benjamin, M.; Malani, N.; Hwang, Y.; Male, F.; Nolte, R. T.; Wang, L.; Velthuisen, E.; Jeffrey, J.; Van Duyne, G. D.; Bushman, F. D. *J. Biol. Chem.* **2014**, *289*, 20477-88.

70. Ovenden, S. P. B.; Yu, J.; San Wan, S.; Sberna, G.; Murray Tait, R.; Rhodes, D.; Cox, S.; Coates, J.; Walsh, N. G.; Meurer-Grimes, B. M. *Phytochemistry (Elsevier)* **2004**, *65*, 3255-3259.
71. Pemberton, I. K.; Buckle, M.; Buc, H. *J. Biol. Chem.* **1996**, *271*, 1498-506.
72. Grobler, J. A.; Stillmock, K.; Hu, B.; Witmer, M.; Felock, P.; Espeseth, A. S.; Wolfe, A.; Egbertson, M.; Bourgeois, M.; Melamed, J.; Wai, J. S.; Young, S.; Vacca, J.; Hazuda, D. J. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 6661-6666.
73. Marchand, C.; Johnson, A. A.; Karki, R. G.; Pais, G. C. G.; Zhang, X.; Cowansage, K.; Patel, T. A.; Nicklaus, M. C.; Burke, T. R., Jr.; Pommier, Y. *Mol. Pharmacol.* **2003**, *64*, 600-609.
74. Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; Miller, M. D. *Science (Washington, D. C.)* **2000**, *287*, 646-650.
75. Billich, A. *Curr. Opin. Invest. Drugs (Thomson Curr. Drugs)* **2003**, *4*, 206-209.
76. Nair, V.; Chi, G.; Ptak, R.; Neamati, N. *J. Med. Chem.* **2006**, *49*, 445-447.
77. Johnson, A. A.; Marchand, C.; Pommier, Y. *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)* **2004**, *4*, 1059-1077.
78. Chow, S. A. *Methods (San Diego)* **1997**, *12*, 306-317.
79. Shibagaki, Y.; Holmes, M. L.; Appa, R. S.; Chow, S. A. *Virology* **1997**, *230*, 1-10.
80. Tous, G.; Bush, A.; Tous, A.; Jordan, F. *J. Med. Chem.* **1990**, *33*, 1620-34.
81. Mochizuki, M.; Anjo, T.; Okada, M. *Chem. Pharm. Bull.* **1978**, *26*, 3914-19.
82. Ram, S.; Ehrenkauf, R. E. *Synthesis* **1986**, 133-5.
83. Voyksner, R. D.; Straub, R.; Keever, J. T.; Freeman, H. S.; Hsu, W. N. *Environ. Sci. Technol.* **1993**, *27*, 1665-72.
84. Ruzicka, K.; Smekal, O.; Velich, V.; Dlask, V. *Chem. Prum.* **1985**, *35*, 630-5.

Supplementary Material

Supplementary material that may be helpful in the review process should be prepared and provided as a separate electronic file. That file can then be transformed into PDF format and submitted along with the manuscript and graphic files to the appropriate editorial office.