Discovery and characterization of highly potent and selective allosteric USP7 inhibitors


Published in:
Nature Chemical Biology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
This work is made available online in accordance with the publisher’s policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Discovery and Characterisation of Highly Potent & Selective Allosteric USP7 Inhibitors.

Authors: Gerald Gavory¹, Colin R O’Dowd¹, Matthew D Helm¹, Jakub Flasz², Elias Arkoudis², Anthony Dossang², Caroline Hughes¹, Eamon Cassidy¹, Keeva McClelland¹, Ewa Odrzywol¹, Natalie Page¹, Oliver Barker¹, Hugues Miel¹, Timothy Harrison¹,²*.

¹Almac Discovery Ltd, Centre for Precision Therapeutics, 97 Lisburn Road, Belfast, Northern Ireland BT9 7AE, United Kingdom. ²Centre for Cancer Research and Cell Biology, Queen’s University, Belfast, Northern Ireland BT9 7AE, United Kingdom.

*Corresponding Author: timothy.harrison@almacgroup.com
ABSTRACT

Given the importance of USP7 in oncogenic pathways, the identification of USP7 inhibitors has attracted considerable interest. Despite substantial efforts however, the development of validated deubiquitinase (DUB) inhibitors which exhibit drug-like properties and a well-defined mechanism of action has proved particularly challenging. In this article, we describe the identification, optimisation and detailed characterisation of highly potent (IC$_{50} < 10$ nM) selective USP7 inhibitors together with their less active, enantiomeric counterparts. We also disclose for the first time co-crystal structures of a human DUB enzyme complexed with small molecule inhibitors which reveal a previously undisclosed allosteric binding site. Finally, we report the identification of cancer cell lines hyper-sensitive to USP7 inhibition (EC$_{50} < 30$ nM) and demonstrate equal or superior activity in these cell models compared to clinically relevant MDM2 antagonists. Overall, these findings demonstrate the tractability and druggability of DUBs and provide important tools for additional target validation studies.
Keywords: 

Ubiquitin 
Deubiquitinase 
Ubiquitin-Specific Protease 
USP7 
p53 
Mdm2
INTRODUCTION

Over the past three decades, protein ubiquitination has emerged as an important post-translational modification with roles in a plethora of cellular processes\(^1\).

Dysregulation of the ubiquitin proteasome system (UPS) has been implicated in the pathogenesis of multiple human diseases including cancer, immune and inflammatory-related medical conditions\(^2,3\). Ubiquitin (Ub) is conjugated to protein substrates through the concerted action of E1, E2 and E3 enzymes and removed by deubiquitinating enzymes (DUBs)\(^4\). The approval and clinical success of the proteasome inhibitor Velcade\(^\circledR\) (bortezomib) and its successors has validated the UPS as a viable target for therapeutic intervention\(^5\). Manipulation of the UPS pathway upstream of the proteasome therefore presents new opportunities for the development of novel therapeutics with potentially enhanced specificity and reduced toxicity profiles\(^6\).

The DUB family comprises 103 members sub-divided into 6 classes\(^7,8\). USPs, which are cysteine proteases, represent the largest sub-family of DUBs with more than 50 family members reported\(^9\). USPs catalyse the removal of ubiquitin from target substrates thus preventing their induced-degradation by the proteasome and/or regulating their activation and subcellular localization, and as such represent an emerging and attractive target class\(^10-12\). Although the co-crystal structure of a viral DUB in complex with a non-covalent inhibitor has been reported\(^13\), to date there have been no reports of a co-crystal structure of a human DUB in complex with a small molecule inhibitor.
Amongst all USPs, Ubiquitin Specific Protease 7 (USP7) has attracted the most attention due to its involvement in multiple oncogenic pathways\textsuperscript{14}. Of particular importance, USP7 plays a key role in regulating the stability of MDM2, an oncoprotein and E3 ligase which promotes the proteasomal degradation of the tumor suppressor p53\textsuperscript{15, 16, 17}. Beyond MDM2, USP7 has also been implicated in the regulation of several other proteins linked to cancer\textsuperscript{14, 18-23}, immunotherapy\textsuperscript{24}, glucose metabolism\textsuperscript{25} and viral infections\textsuperscript{26-28}.

To date, several reports have described attempts to identify potent and selective USP7 inhibitors\textsuperscript{29-36}, however this has proved challenging with reported inhibitors generally exhibiting low (i.e. micromolar) potency, often coupled with sub-optimal physicochemical properties, poor stability and/or low selectivity\textsuperscript{29, 35, 37}. In addition, redox properties of ligands can confound the interpretation of biological data\textsuperscript{38}.

Herein, we describe the discovery of highly potent (nM) and selective inhibitors of USP7. We demonstrate that lead compounds inhibit USP7 in a non-covalent and reversible manner, and confirm direct binding to the target by orthogonal methods including the first reported high resolution co-crystal structures of USP7 in complex with small molecule inhibitors. These inhibitors possess excellent in vitro selectivity for USP7 and potently engage endogenous USP7 at concentrations which are similar to their biochemical potencies. The observed target engagement translates into degradation of MDM2, stabilisation of p53 and induction of p21 in multiple cell lines. Finally, cell lines which are hyper-sensitive to USP7 inhibition have been identified, thus providing new insights and directions for the potential development of USP7 inhibitors.
RESULTS

Identification of small-molecule USP7 inhibitors

Due to the difficulties in identifying high quality hit matter for DUBs using HTS\(^3^9\), we decided to utilize fragment-based methods for the identification of USP7 inhibitors\(^4^0\).

We used Surface Plasmon Resonance (SPR) for primary hit finding and screened 1946 fragments against the catalytic domain of USP7, followed by the synthesis, purchase and testing of additional fragments. From this work, Fragments A and B were identified as USP7 binders (Table 1). To complement our fragment-based approaches, we also synthesised and profiled examples of all disclosed USP7 inhibitors\(^2^9,4^1\), and by combining features of these published molecules with USP7 binding fragments (e.g. Fragment B, *manuscript in preparation*) we were ultimately able to generate novel compounds with reasonable affinity for USP7 as exemplified by 1 (IC\(_{50} = 13.1\) \(\mu\)M, Table 1 and *Supplementary Figure 1a*). Compound 1 is chemically stable, has high aqueous solubility (>200 \(\mu\)M at pH\(_{7.4}\)) and is devoid of redox liabilities which may lead to the non-selective and spurious inhibition of USP7 (Table 1)\(^4^2\). We also confirmed inhibition of USP7 in a second assay format (IC\(_{50} = 15.1\) \(\mu\)M, Table 1) using a non-isopeptide linked substrate (Ub-Rh110). In addition, the binding of 1 to USP7 was confirmed by an orthogonal biophysical technique (SPR, \(K_d = 7.7\) \(\mu\)M, *Supplementary Figure 1b* and Table 1). Our SPR data also indicated 1 was a reversible binder \(k_{\text{off}} = 0.25\) s\(^{-1}\) which was further confirmed using a high-dilution biochemical assay (*Supplementary Figure 1c*). Collectively, these data confirm 1 as a genuine, reversible inhibitor of USP7 and a potentially tractable starting point suitable for further optimization.
Co-crystal structure of USP7 bound with an inhibitor

A key breakthrough in the development of potent and selective USP7 inhibitors was the finding that the introduction of a chiral methyl group ((R) stereochemistry) at the benzylic position of the phenethylamide chain increased USP7 inhibitory activity by up to 40-fold. Thus, 2 (IC$_{50}$ = 0.3 μM, Table 1) represented the first sub-μM USP7 inhibitor from this series. Importantly in the context of identifying negative control compounds, we observed a significant difference in activity between 2 and its enantiomer ent-2 (IC$_{50}$ = 46.0 μM, Table 1).

The X-ray crystal structure of 2 in complex with USP7 was subsequently solved at high-resolution (2.2 Å, Figure 1a). The protein was observed to be in a conformation similar to the USP7 apo structure (PDB: 1NB8), with a disrupted catalytic triad and an overall Cα RMSD of 1.03 Å (see Supplementary Figure 2). The crystal structure also revealed key conformational changes that enable ligand binding: Phe409 adopts a novel side chain conformation not seen in either the apo or ubiquitin-conjugated USP7 structures available in the PDB (1NB8 and 1NBF respectively). This side chain movement creates a hydrophobic cavity occupied by the phenyl ring of 2 (Figure 1b).

The loop between Gly458 and Gly463 also changes conformation (see Supplementary Figure 3), positioning His461 between the ligand and the catalytic cysteine. The ligand therefore sits in an exo-site ca. 5.5 Å away from the catalytic cysteine Cys223 (as measured from the carbonyl group) and partially protrudes into the channel normally occupied by the C-terminal tail of ubiquitin (Figure 1c). The folded conformation of the ligand is reinforced by allylic 1,3-strain between the CH
of the chiral centre and the phenyl ring. Our analysis also indicates that 2 adopts a folded conformation in which the phenyl ring is packed against the piperidine ring, creating a potential CH-π intra-molecular interaction between a piperidine C-3 axial hydrogen and the aromatic group (Figure 1b). The co-crystal structure also provided a potential explanation at the molecular level for the difference in activity between 2 and its enantiomer (ent-2). Inversion of the ligand stereo centre in the structure leads to a direct steric clash with His461 (Supplementary Figure 4). Together, this study reveals the first reported co-crystal structure of a DUB enzyme bound to a small molecule inhibitor in a novel site situated 5.5 Å from the catalytic cysteine (see Supplementary Figures 5a-d for further details).

Optimisation of 2 provides highly potent USP7 inhibitors

Using the co-crystal structure to facilitate structure-based design, we were able to further optimise the potency of compound 2 by modifying the core heterocycle and building off the heterocyclic ring into the pocket containing Gln351. This led to 3 (IC50 = 6.0 nM; Table 1) and provided a further ca. 50-fold increase in potency relative to 2. In order to confirm the determinants of binding, we solved the high-resolution (1.7 Å) crystal structure of a close analogue of 3 (compound 4, IC50 = 22.0 nM, Figure 2a) in complex with USP7. This co-crystal structure shows a high degree of similarity to the previously described complex of USP7 with 2 (Figure 1a). The Ca RMSD is 0.35 Å and the same side chain movements which create the hydrophobic pocket and the same pattern of hydrogen bonds are observed in both structures. The structure reveals how the benzylic amine extension to the heterocyclic core allows the amine
to form additional hydrogen bonds with Gln351 and two adjacent bridging waters
while N-1 of the pyrazole can form a hydrogen bond with Phe409 instead of the
pyrimidone carbonyl, leading to a significant increase in overall potency (Figure 2a).
The catalytic triad is also misaligned. An overlay of the co-crystal structures of 2 and
4 in complex with USP7 is shown in Figure 2b (See also Supplementary Figures 6a-c
for further details).

Biochemical and biophysical profiling of 3 and ent-3.
The detailed biochemical and biophysical profiling of 3 is compiled in Table 1. 3 is
free of redox liabilities, is highly soluble and is 400-fold more active than its
enantiomer ent-3 (IC$_{50}$ = 2.4 µM; Supplementary Figure 7a, Table 1). 3 was also
confirmed to be highly potent in an alternative assay format using the Ub-Rh110
substrate (IC$_{50}$ = 1.5 nM; Table 1). Characterisation by SPR revealed tight binding of 3
to USP7 (K$_d$ = 2.0 nM; Supplementary Figure 7b; Table 1). Under similar conditions,
ent-3 was confirmed to be a much weaker binder (K$_d$ = 5.0 µM). Subsequent kinetic
analysis by SPR demonstrated that the potency gain obtained during compound
optimisation was driven by an increase in target site occupancy as evidenced by the
slower $k_{off}$ value for 3 ($k_{off} = 0.0004$ s$^{-1}$; $t_{1/2} = 28.4$ min) relative to 1 ($k_{off} = 0.25$ s$^{-1}$;
$t_{1/2} = 2.7$ s).

Determination of a non-competitive mode of inhibition
Kinetic experiments were subsequently performed to characterise the mode of
inhibition of USP7 by these inhibitors using the most potent compound 3 as a
representative example (Supplementary Figures 8a, b). Lineweaver-Burk analysis
revealed that 3 acts as a non-competitive inhibitor. The apparent $K_m$ value for Ub-AMC was calculated as 1.9 $\mu$M and the inhibitory constant for 3 derived from this analysis ($K_i = 5.5$ nM) was in excellent agreement with the IC$_{50}$ value obtained previously (IC$_{50} = 6.0$ nM; Table 1). An Eadie-Hofstee analysis of the initial Michaelis-Menten dataset was performed, which confirmed the non-competitive mode of inhibition (Supplementary Figure 8c). In a separate experiment, USP7 was assayed with increasing concentration of the Ub-AMC substrate. Under the conditions of this experiment, the IC$_{50}$ values for 3 were independent of the substrate concentration further supporting the non-competitive mode of inhibition (average IC$_{50} = 1.0$ nM, see Supplementary Figure 9).

3 is highly selective against DUBs, proteases and kinases

One long standing question in DUB research has related to the degree of selectivity achievable when targeting this enzyme class$^{43}$. To address this question, we next assessed the selectivity profile of our inhibitors against a panel of DUBs (n=39; DUBprofiler™). In these experiments, 3 was screened at a fixed concentration of 100 $\mu$M and showed no significant activity (<20% inhibition) against 21 USPs and 17 other DUBs across the various sub-family members (Figure 3). We extended our screening to alternative target classes which indicated no significant off-targeting effect at 10 $\mu$M against a panel of 63 proteases or 49 representative kinases (Supplementary Figures 10a and b). Together these data demonstrate that high selectivity for inhibition of USP7 can be achieved versus DUBs and other target classes (>1000-fold).

3 Demonstrates highly potent target engagement in cells
We then investigated whether 3 was capable of interacting with endogenous USP7. Cells were treated with increasing concentration of inhibitor, lysed and the ubiquitin-propargylamine (Ub-PA) probe was then added. 3 efficiently competed with the Ub-PA probe in a concentration-dependent manner (EC$_{50}$ = 49 nM, Figure 4a). Only a minimal (<10-fold) cellular drop-off in the activity of 3 was observed relative to the biochemical activity (IC$_{50}$ = 6.0 nM) consistent with the non-competitive mode of inhibition demonstrated previously. Ent-3 had no significant effect on target engagement (EC$_{50}$ > 50 µM; Figure 4a). Using a similar protocol, we assessed the selectivity of 3 in cells and demonstrated lack of target engagement against USP47 (EC$_{50}$ > 50 µM; Figure 4b), the closest homologue of USP7. Off-target activity against USP47 has been previously reported for USP7 inhibitors$^{31,34}$. A homology model of USP47 was created based on our USP7 co-crystal structures (Supplementary Figure 11a). Sequence alignment indicated that 6 residues within 5.0 Å of the compound binding site are different in USP47 relative to USP7 (Supplementary Figure 11b) providing a molecular basis for this selectivity. We extended our cellular investigations beyond USP47 to other USPs and again showed no effect on target engagement at concentrations up to 50 µM (Figure 4b). This data is consistent with the lack of measurable biochemical inhibition against these USPs (IC$_{50}$ > 100 µM; Figure 3). Together, these studies demonstrate good cellular permeability, potent target engagement with endogenous USP7 and excellent selectivity in cells against selected DUBs including USP47.

3 increases p53 and decreases MDM2 total levels in cells
We next evaluated the effect of 3 on downstream pathway components of USP7 signalling. HCT116 cells were treated with increasing concentrations of 3 for 2 h and lysed for western blotting analysis probing for key components of the USP7/MDM2 axis. Whilst the total USP7 protein levels remained steady, levels of p53 increased in a concentration-dependent manner (Figure 5a). The levels of p21 protein mirrored this increase in p53 indicating that the transcriptional activity of p53 was also restored. Importantly, these effects on p53 and p21 were independent of a genomic stress response as evidenced by the absence of p53 phosphorylation on Serine 15. Concomitantly, the levels of the MDM2 oncoprotein decreased in a concentration dependent manner. As shown in Figure 5b, these observations were also extended to the MCF7 cell line. In both cases, reduced levels of MDM2 could be rescued by the addition of the proteasome inhibitor MG132. The reduced levels of MDM2 following treatment with 3 contrasted markedly with the significant increase in MDM2 total protein levels induced by SAR405838 (Figures 5a-b); a characteristic of MDM2 antagonists caused by the feedback loop between p53 and MDM2. By comparison, ent-3 had no significant effect on the total levels of p53, p21 or MDM2 (Figures 5a-b).

We next performed a quantitative analysis of ubiquitinated MDM2 levels using the commercial MDM2 whole cell lysate kit from MSD in the MDM2 amplified osteosarcoma cell line SJSA-1. Our data confirmed a potent effect of 3 on the total levels of ubiquitinated MDM2 (EC_{50} = 55.0 nM, Figure 5c, Table 1) with a significantly reduced activity for ent-3 (EC_{50} > 20 µM). Altogether, these observations are consistent with the expected mechanism of action of a USP7 inhibitor, in which a
decrease in the levels of the oncogene MDM2 restores both the levels and transcriptional activity of the tumor suppressor p53\textsuperscript{15,46}.

Identification of cell lines hyper-sensitive to 3

Having demonstrated target engagement and pathway modulation, we next assessed the anti-proliferative effects of 3. To this end, we profiled 3 in a panel of cancer cell lines and identified a subset of lines hyper-sensitive to USP7 inhibition including the acute lymphoblastic leukemia cell line RS4;11. Follow-up studies using a viability assay performed at 72 h confirmed acute sensitivity to 3 (EC\textsubscript{50} = 2.0 nM, Supplementary Table 3 and Supplementary Figure 12a). A time-course analysis by western blotting (8 to 72 h post treatment) revealed a strong and sustained effect on the total levels of p53 which was mirrored by the induction of p21 (Supplementary Figure 12b). We also observed a strong, time-dependent apoptotic response in this cell line as evidenced by both PARP and caspase 3 cleavage. Interestingly, the apoptotic response was time-shifted compared to treatment with the MDM2 antagonist SAR405838. Under identical experimental conditions, ent-3 had no effect on p53 and p21 levels or apoptosis over this time-course and only had a weak impact on cell growth (EC\textsubscript{50} = 2.2 µM). The acute sensitivity of RS4;11 cells to 3 is in line with the potent target engagement we observed in this cell line (EC\textsubscript{50} = 6.0 nM, Supplementary Figure 12c).

Hyper-sensitivity to 3 could be extended to solid tumor cell lines including prostate (LNCaP, EC\textsubscript{50} = 29.0 nM, Supplementary Figure 13a; Supplementary Table 3). In this cell line, we again observed potent target engagement (EC\textsubscript{50} = 7.0 nM) with excellent
selectivity in cells over 4 other USP’s including USP47 (Supplementary Figures 13c and d), which translated into a significant increase in ubiquitination of MDM2 protein following treatment with 3 (but not ent-3, Supplementary Figure 13e). As described previously for the RS4;11 cell line, no significant effect on cell growth was observed following treatment with ent-3 (EC₅₀ = 9.3 µM). Intriguingly, we observed that the modulation of p53 and p21 was less marked in this cell line than in RS4;11 cells and that the effect was typically observed at an earlier time point (8 h for p53 and up to 48 h for p21; Supplementary Figure 13b). In addition, we could not detect any significant degree of apoptosis in this cell line (as judged by the lack of caspase 3 and PARP cleavage) following treatment with 3 up to 72 h. Similar observations have been made following treatment with the MDM2 antagonist SAR405838 (Supplementary Figure 13b). This data is consistent with previously published studies using MDM2 antagonists in the LNCaP cell line where caspase-3 independent cell death has been described⁴⁵. Further studies will be required to elucidate the mechanism(s) of cell death in this cell line.

Benchmarking the activity of 3 against MDM2 antagonists

To further investigate the potential therapeutic relevance of USP7 inhibitors, we next performed benchmarking experiments versus several MDM2 antagonists. Included in these studies were nutlin-3a, RG7112 and SAR405838. Both the RS4;11 and LNCaP cell lines have been previously shown to be highly responsive to MDM2 antagonists⁴⁴,⁴⁵. Under identical experimental conditions, 3 compared very favourably to all three MDM2 antagonists (Supplementary Figures 14a-b). For example, 3 was found to be 70-fold more potent than nutlin-3a, 30-fold more potent
than RG7112 and 16-fold more potent than SAR405838 in the RS4;11 cell line (Supplementary Table 3). A similar ranking was observed in the LNCaP cell line. It will be important in future studies to further understand and delineate both the commonalities and differences between USP7 inhibitors and MDM2 antagonists. For instance, we have already observed that the osteosarcoma cell line SJSA-1 is sensitive to MDM2 antagonists (EC$_{50}$ = 100.0 nM and 320.0 nM for SAR405838 and RG7112 respectively) but resistant to treatment with 3 (EC$_{50} >$ 20 µM, Supplementary Figure 14c and Supplementary Table 3). Interestingly, under identical experimental conditions to those in which we observed full target engagement in the RS4;11 and LNCaP cell lines, only partial target engagement was observed in the SJSA-1 cell line even at high USP7 inhibitor concentrations (Supplementary Figure 14d). The above observations are encouraging and further mechanistic studies aimed at fully understanding the dependence on the p53 gene status and at precisely defining the scope of application of these USP7 inhibitors are warranted and will be reported at a later stage.

Discussion

DUBs have emerged as an attractive target class for the development of first in class medicines with potential for high therapeutic impact. However, despite 15 years of intense research, DUBs have proved largely refractory to drug discovery efforts. The development of genuine DUB inhibitors combining drug-like properties with a well-defined mode of inhibition thus remains a high priority for the research community.
USP7 has attracted special attention due to its established connections to known oncogenic pathways and other disorders\(^{14}\). The prospect of simultaneously restoring the cellular levels and function of p53 whilst promoting the degradation of the oncogenic E3 ligase MDM2 is particularly appealing. USP7 inhibitors have been described previously in literature publications, however many of these scaffolds have typically lacked high potency, stability, selectivity and/or a well-defined mechanism of action\(^{29,36,37}\). To the best of our knowledge the data reported herein represents the first published examples of highly potent (nM), selective and reversible small molecule DUB inhibitors with a clearly defined mode of binding to USP7 underpinned by high-resolution X-Ray structural information.

We prosecuted hit-finding using a combination of fragment-based screening and rational design. Using our UbiPlex™ assay platform, the early chemical hit matter identified (e.g. 1) was optimized leading to the identification of 2 and 3 (Table 1). We demonstrated that these inhibitors are highly soluble, chemically stable and free of redox recycling activity (Table 1), which can otherwise lead to the non-specific and/or spurious inhibition of DUBs. Biochemical activity was demonstrated in two different assay formats and binding was orthogonally validated using SPR (Supplementary Figures 1a-b, 7a-b; Table 1). Finally, X-ray structural studies provided unambiguous binding validation of this series of inhibitors to USP7. Analysis of these high-resolution (1.7-2.3 Å) structures revealed that the inhibitors bind outside of the active site in a previously undisclosed allosteric pocket situated 5.5 Å from the catalytic cysteine. Our analysis also revealed a binding mode with a folded conformation of the ligand, stabilised by an intra-molecular CH-π interaction.
between a piperidine CH and the phenyl group of the phenylbutanamide side-chain (Figures 1b; 2a-b). The ligand-bound structures are reminiscent of the apo structure of USP7 in which the catalytic triad remains misaligned and in a non-functional, catalytically incompetent state (Supplementary Figure 2). We postulate that a potential mechanism of action of these molecules may be to prevent the alignment of the catalytic triad within the active site of USP7 and by partially protruding into the channel which normally accommodates the C-terminus of ubiquitin, the ligands may additionally induce a local distortion of the Ub tail, thus precluding catalysis. These two mechanisms are not necessarily mutually exclusive and could act in concert. Furthermore, both models are consistent with the non-competitive mode of inhibition demonstrated during the course of our studies (Supplementary Figure 8).

Structure-based design was used to further optimise our early leads culminating in the identification of our most potent inhibitor 3 which consistently exhibited IC₅₀ values in the single digit nM range (Table 1). High selectivity (>10,000-fold) against other members of the USP and DUB families was demonstrated (Figure 3). The high specificity of 3 for USP7 was confirmed in cells as evidenced by the lack of target engagement against USP47 (up to 50 µM), the closest USP7 homologue (Figure 4b) ¹, ³⁴. Under similar conditions, 3 proved to be highly potent against USP7 in this cellular target engagement assay (EC₅₀ = 49.0 nM, Figure 4a). We extended our off-target screening to alternative target classes which indicated no significant off-targeting effect at 10 µM against a panel of 63 proteases and 49 kinases (Supplementary Figures 10a, b respectively).
In order to explore how potentially relevant the structural information may be in designing inhibitors for other USPs, we performed a comparison of the key residues forming the exosite for Compounds 2 and 4 with the equivalent residues from the published crystal structures of other USP’s (Supplementary Figure 15). Interestingly, the key residues involved in binding are conserved in each of the USP’s. However, we observe significant alterations in certain sidechain conformations for some residues and target specific features which may preclude binding of molecules such as 2 and 4, consistent with their excellent selectivity profiles. However the relatively high degree of conservation in the residues forming the binding site suggests that there may be an opportunity to utilise these ligands for the design of inhibitors of other USPs, and this remains an area of ongoing investigation.

With potent, selective chemical probes in hand, concentration dependent pathway modulation was subsequently demonstrated leading to increased protein levels of p53 and reduced levels of the oncoprotein MDM2. In addition, we observed a concomitant induction in p21 protein levels (Figures 5a-b). The absence of p53 phosphorylation on Ser15 following treatment with 3 together with the absence of any effect on protein levels (of p53, MDM2 or p21) following treatment with ent-3 demonstrates the non-genotoxic, on-target nature of these effects. These observations were extended to multiple cell lines (Supplementary Figures 12b and 13b) and taken together are consistent with the expected mechanism of action of a USP7 inhibitor. Next, we assessed the anti-proliferative effects of these compounds and in so doing identified cell lines hyper-sensitive (EC50 < 30 nM) to USP7 inhibition
which included both haematological (RS4;11) and solid tumor cell lines (LNCaP) (Supplementary Figures 12a and 13a). A robust time- and concentration-dependent apoptotic induction was observed in RS4;11 cells as evidenced by PARP and caspase 3 cleavage. Consistent with previously published studies, no marked apoptotic response was detected in the LNCaP cell line using either USP7 inhibitor 3 or the highly potent MDM2 antagonist SAR405838. Ent-3 had no marked effect on apoptosis and exerted no significant anti-proliferative effect in either of these two cell lines (Table 2, Supplementary Figures 12a-b, 13a-b). The marked differences in behaviours between 3 and ent-3 reflects the value and utility of enantiomeric pairs as chemical probes to confirm on-target effects during target validation studies.

Finally, we benchmarked the anti-proliferative activity of 3 against established MDM2 antagonists. Profiled head-to-head, 3 was more potent than all three MDM2 antagonists in the RS4;11 and LNCaP cell lines (Supplementary Figures 14a-b). Taken together, these benchmarking studies represent an important proof-of-concept for USP7 inhibitors and justify further cellular and in vivo profiling. Carefully-designed in vivo efficacy and toxicity studies will then be required to fully define the therapeutic relevance, potential therapeutic index and scope of pharmacologically targeting USP7. Emphasis will be placed in future studies on assessing the p53-dependence and unveiling the markers of sensitivity to USP7 inhibitors as well as in delineating the overlap and differences in activity when compared to MDM2 antagonists. In this regard, potential differences have already been observed as evidenced by the lack of
response of the SJSA-1 cell line to 3 compared to the high sensitivity to SAR405838 and RG7112 (Supplementary Figure 14c, Supplementary Table 3).

In conclusion, we report herein the first highly potent (nM), reversible and selective inhibitors of USP7 which exhibit a well-characterised binding mode and modulate the total levels of downstream targets of USP7 including MDM2, p53 and p21. As part of these studies, we have demonstrated the ligandability and tractability of USP7 as a target and reported the first co-crystal structures of a DUB with high affinity small molecule inhibitors. We have additionally identified cell lines hyper-sensitive to USP7 inhibitors and demonstrated potency in cells comparable or superior to that of clinically relevant MDM2 antagonists. These findings represent important milestones in terms of feasibility and proof-of-concept studies, and we anticipate that these enantiomeric chemical probes will be of direct relevance and utility to the USP7 research community and more broadly to all those aiming to identify and develop innovative DUB therapeutics.

ACKNOWLEDGEMENTS

The authors thank collaborators at Ubiquigent (Dundee, UK), Boston Biochem (Boston, US) and Beactica (Uppsala, Sweden) for their contributions, as well as Crelux GmbH (Martinsried, Germany) for solving the crystal structures, and Prof. C. Scott and Dr James Burrows (Queen’s University, Belfast) for helpful discussions and for providing the HA-Ub plasmid respectively. This study was supported by the Almac Group, the European Regional Development Fund and Invest Northern Ireland (Grant RD1010668).
Author Contributions

TH conceived the concept and directed the research. CROD and GG helped conceive and develop the concept and designed and supervised medicinal chemistry and biology experiments. MDH, EA, JF and HM carried out the design, synthesis and characterization of compounds. AD performed SPR experiments. CH, KMcc, EO, EC, AD and NP carried out compound screening, target validation and biochemical and cellular profiling studies. OB carried out computational modelling and structural analysis. TH and GG wrote the manuscript with input from other authors.

Competing Financial Interests Statement

The authors declare no competing financial interest.
References


FIGURE LEGENDS

Figure 1: High-resolution X-ray co-crystal structure of USP7 in complex with 2 (PDB accession number 5N9R). a, Overall structure of USP7 in complex with 2 (2.2 Å resolution). The catalytic triad is highlighted in green. Compound 2 is shown in an orange space filling representation. Ubiquitin aldehyde (from PDB structure 1NBF) is included as a magenta ribbon for reference (position determined by sequence and structural alignment with 1NBF). b, Ribbon and stick representation showing the key interactions between USP7 and 2. The intramolecular CH-π interaction between piperidine C-3 axial hydrogen and the aromatic group is also shown. c, Co-crystal structure of USP7 with 2 overlaid with the C-terminal region of ubiquitin (gold) illustrating how 2 partially protrudes into the channel normally occupied by the Ub tail creating a steric clash with the ubiquitin C-terminus.

Figure 2: High-resolution X-ray crystal structure of USP7 in complex with 4 (PDB accession number 5N9T). a, Structure solved at 1.7 Å resolution. Ribbon and stick representation showing the key interactions between USP7 and 4 including the interaction with Q351, the adjacent water molecules and the various potential hydrogen bonds (predicted interaction energy magnitude ≥ 0.5 kcal/mol, distance cutoff 4.5 Å) between 4, F409 and R408. b, Overlay of the structures of Compounds 2 and 4 in complex with USP7.
Figure 3: Selectivity profile of 3 against a panel of DUBs. a, Representative selectivity profile of 3 against a panel of 21 USPs and 17 other DUBs (DUBprofiler^TM, Ubiquigent). Screening was performed at a fixed concentration of 100 µM against all other members of the USP/DUB family. All data reported as the mean of 2 independent experiments.

Figure 4: Target engagement and selectivity profile of 3 against selected USPs in cells. a, Cells (HCT116) were treated with 3 and ent-3 for 2 h (as indicated), lysed and the ubiquitin-propargylamine (Ub-PA) probe was added. Samples were subsequently analysed by western blotting probing for USP7 (quantitative analysis shown from 2 independent experiments). b, No target engagement (HCT116 cells) was detected up to 50 µM against the most closely related USP7 family member (USP47). Similar results were obtained when probing against additional non-related USPs including USP4, 11, 19. (+) and (−) signs represent the presence or absence of the Ub-PA probe. Concentration (Conc.). Full blots and cut membranes are shown in Supplementary Figures 16a, b.

Figure 5: Treatment of cells with 3 caused non-genotoxic stabilisation of p53 and decreased levels of MDM2. a, HCT116 cells were treated with 3, ent-3 and SAR405838 for 2 h and lysed for western blotting analysis probing for USP7, p53, pSer15-p53, p21, MDM2 and β-actin as indicated. b, Repeat of (a) in the MCF7 (breast adenocarcinoma) cell line. Western blotting analysis was performed as
described previously. The proteasome inhibitor MG132 was used at a concentration of 10 µM. Full blots and cut membranes are shown in Supplementary Figures 16c, d.

c, Quantitative determination of ubiquitinated MDM2 in SJSJ-1 cell line. Cells were treated with increasing concentration of inhibitors for 24 h (as indicated). Assay and analysis was performed as recommended by the manufacturer’s instructions (MSD).

Signal (A.U.) represents the ratio of ubiquitinated MDM2 normalised to total levels of MDM2. Data reported as the mean of at least three independent experiments with standard deviations.

**TABLES**

**Table 1: Chemical structures and compiled data for USP7 binding fragments and inhibitors.**

<table>
<thead>
<tr>
<th>Assays</th>
<th>1</th>
<th>2</th>
<th>ent-2</th>
<th>3</th>
<th>ent-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redox a</td>
<td>free</td>
<td>free</td>
<td>free</td>
<td>free</td>
<td>free</td>
</tr>
<tr>
<td>Kinetic solubility (µM)</td>
<td>&gt;200</td>
<td>&gt;200</td>
<td>&gt;200</td>
<td>&gt;200</td>
<td>&gt;200</td>
</tr>
<tr>
<td>USP7 IC50 (µM) – FP assay</td>
<td>13.1±0.8</td>
<td>0.3±0.1</td>
<td>46.0±5.6</td>
<td>0.006±0.002</td>
<td>2.4±0.3</td>
</tr>
<tr>
<td>USP7 IC50 (µM) – Ub-Rh110</td>
<td>15.1±1.4</td>
<td>0.3±0.07</td>
<td>&gt;50</td>
<td>0.0015±0.001</td>
<td>1.9±0.4</td>
</tr>
<tr>
<td>SPR Kd (µM)</td>
<td>7.7±0.1</td>
<td>2.6±0.1</td>
<td>&gt;50</td>
<td>0.002±0.001</td>
<td>5.0±0.1</td>
</tr>
<tr>
<td>SPR k_on (s⁻¹)</td>
<td>0.25±0.01</td>
<td>0.03±0.001.</td>
<td>0.47±0.01</td>
<td>0.0004±0.0001</td>
<td>0.09±0.01</td>
</tr>
<tr>
<td>DUB selectivityc</td>
<td>-</td>
<td>0/18</td>
<td>-</td>
<td>0/38</td>
<td>-</td>
</tr>
<tr>
<td>Target engagement EC50 (µM) b</td>
<td>&gt;50</td>
<td>0.85±0.15</td>
<td>&gt;50</td>
<td>0.049±0.025</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Ub-MDM2 EC50 (µM) b</td>
<td>&gt;20</td>
<td>2.6</td>
<td>&gt;20</td>
<td>0.055±0.015</td>
<td>&gt;20</td>
</tr>
</tbody>
</table>

*Chemical structures, biochemical, biophysical profiling of USP7 inhibitors. Upper panel: chemical structures of the inhibitors (as indicated). Primary hit finding was performed by SPR using the Beactica fragment library. See Online Methods Section and main text for more details on these assays. Enantiomer (ent); a redox free: free of redox cyclic activity when tested up to 200 µM; fluorescence polarisation (FP). b 50 and 20 µM were the maximum concentration used. -- not determined. c number of DUBs showing >20% inhibition at 100µM / total number of DUBs tested. Data reported is the mean of at least 3 independent experiments ± standard deviations.
ONLINE METHODS

Materials and reagents

All reagents and chemicals were purchased from Sigma-Aldrich unless otherwise stated. The MDM2 antagonists nutlin-3a (Tocris, #3984), SAR405838 (MedchemExpress; MI-773, #HY-17493) and RG7112 (MedchemExpress, #HY-10959) were purchased from commercial suppliers as indicated and used with no further purification. All inhibitors were prepared as 10 mM DMSO stocks for cell culture experiments and stored in a controlled environment using the MultiPod system. CellTiter-Glo® was purchased from Promega (#G7571). The ubiquitin-propargylamine (Ub-PA) probe was purchased from UbiQ (#UbiQ-057). Unless otherwise stated, all other reagents were obtained from commercial sources and used without further purification.

Compound synthesis and characterisation

Compound synthesis and characterisation data are included in the Synthetic Procedures Section.

Cell lines and culture conditions

All cell lines were obtained from the American Type Culture Collection (ATCC), authenticated by STR profiling (Promega) and shown to be mycoplasma-free using the MycoAlert mycoplasma detection (Lonza; LT07-318). For growth, cells were maintained at 37°C in a humidified atmosphere with 5% CO₂. HCT116 (colorectal) cells were cultured in McCoy's 5A supplemented with 10% (v/v) FBS, 1% (v/v)
Penicillin/streptomycin, 1% (v/v) L-Glutamine. RS4;11 (acute lymphoblastic leukemia), LNCaP (prostate) and SJSA-1 (osteosarcoma) cells were cultured in RPMI supplemented with 10% (v/v) FBS, 1% (v/v) Penicillin/streptomycin, 1% (v/v) L-Glutamine. MCF7 (breast) cells were cultured in Eagle’s Minimum Essential Medium supplemented with 10% (v/v) FBS, 0.01 mg/mL human recombinant insulin and 1% (v/v) Penicillin/streptomycin. Medium and supplements were purchased from Life Technologies except where indicated.

Biochemical assay and reversibility studies

USP7 activity was monitored in a fluorescence polarisation (FP) homogeneous assay using the isopeptide ubiquitin-Lys-tetramethylrhodamine substrate (Ub-TMR; U-558, Boston Biochem). Full-length USP7 was purchased from Boston Biochem (His6-USP7FL, E-519). Unless otherwise stated, all other reagents were purchased from Sigma. Enzymatic reactions were conducted in black flat bottom polystyrene 384-well plates (Nunc) in 30 μL total volume. USP7 (2.5 nM, 10 μL) was incubated in assay buffer containing 50 mM HEPES (pH 7.2), 150 NaCl, 0.5 mM EDTA, 5 mM DTT, 0.05% BSA (w/v), 0.05% CHAPS in the presence or absence of inhibitor (10 μl). Inhibitors were stored as 10 mM DMSO stocks in an inert environment (low humidity, dark, low oxygen, room temperature) using the Storage Pod System and serial dilutions were prepared in buffer just prior to the assay (typically from 200 to 0.001 μM in 10 dp curve). Following incubation at room temperature for 30 min, the enzymatic reactions were initiated by dispensing the Ub-TMR substrate (250 nM, 10 μL). FP was measured every 15 min over a period of 90 min (within the linear range of the assay) using a Synergy 4 plate reader (BioTek) exciting at 530 nm and measuring the
amount of parallel and perpendicular light at 575 nm. The FP signal was subsequently normalised to the no compound control (i.e. DMSO). Analysis and IC$_{50}$ values were derived using GraphPadPrism (GraphPad Software, Inc, La Jolla, CA; four-parameter logistic function). All data presented as mean ± s.d. (n >=3). For the 100x dilution assay, USP7 was pre-incubated for 30 min with 10x the inhibitor IC$_{50}$ value followed by 100x dilution. USP7 activity was assessed 15 min post-dilution as described above. Ubiquitin, iodoacetamide (#A3221; Sigma-Aldrich) and H$_2$O$_2$ were used at a final pre-dilution concentration of 150 µM, 1.0 mM and 0.001 % solution respectively.

**Redox assay**

The property of small molecule to oxidise the target was assessed using the method described by Lor et al$^{42}$. Briefly, compounds at a final assay concentration of 200 µM were incubated for 30 min in buffer containing 50 mM HEPES (pH 7.5), 50 mM NaCl, plus 5 µM Resazurin (#R7017; Sigma-Aldrich) and 50 µM DTT, in a black low volume 384-well Greiner plate. Readout was performed with Excitation of 560 nm and Emission of 590 nm using a Synergy 4 plate reader (BioTek). Data was normalised to the no DTT control.

**SPR experiments**

Residues S207 to E560 of the USP7 catalytic domain (CD) was expressed in E.coli cells using a p10T7-2 expression vector with an N-terminus His6-GST tag cleavable with TEV protease. The purified protein stored in a PBS solution was immobilised on a
COOH₅ SensiQ chip by amine coupling at a concentration of 150 µg.ml⁻¹ and a pH = 4.5 Na-Acetate solution to achieve a Rmax of approximately 100 RU, for a 500 g.mol⁻¹ small molecule under saturation conditions (i.e. above 10-fold Kd). The SensiQ Pioneer surface plasmon resonance (SPR) system was equilibrated with a HBS-EP buffer solution (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05 % P20 surfactant) with the addition of 5 % DMSO and 2 mM DTT on the day of the experiment. All samples were tested in the same buffer solution with 5% DMSO and 2 mM DTT with an association injection time of 75 µl.min⁻¹ and a 60 s dissociation time under a 3 % sucrose gradient. Data was analysed using Qdat v 2.5.2.12. from SensiQ Technologies, Inc. unless otherwise stated.

**Mode of inhibition**

Compound 3 at final assay concentrations ranging from 25 µM to 0.78 nM was pre-incubated for 30 min with 5 nM of full-length USP7 (2.5 nM final assay concentration). Five microliters of ubiquitin-aminocumarin (Ub-AMC; # 60-0116-050; Ubiquigent) was added at varying concentrations ranging from 1.5 nM to 25 µM to a final assay volume of 15 µl in low volume black 384-well plates (Greiner # 784076). Fluorescence intensity (ex = 350 nm, em = 445 nm) was measured over 90 min and initial linear rates were calculated. Global fit analysis for competitive models were performed using GraphPadPrism (GraphPad Software, Inc, La Jolla, CA; Global fit Michaelis-Menten enzyme inhibition). IC₅₀ values were also determined at increasing concentration of Ubiquitin-AMC representing the substrate concentration at 1-fold, 10-fold and 50-fold the Kₘ of ubiquitin-AMC for USP7 using a Synergy 4 plate reader.
(BioTek). Data analysis was performed using GraphPad Prism (GraphPad Software, Inc, La Jolla, CA; four-parameter logistic function).

**F-Test**

Global fitting models for each mode of inhibition; competitive, non-competitive and uncompetitive, were subject to an F-Test statistical analysis to further confirm the mode of inhibition. The F-Test compares the fit between two different equations. The lower \( \chi^2 \) indicates the fit between two models is significantly better. The \( \chi^2 \) values for the competitive, non-competitive and uncompetitive models were respectively; 1352.6, 572.8 and 3307.1. A significant probability should be lower than 0.05. The non-competitive model gives the best statistical fit with the lowest \( \chi^2 \) and a probability of \( 1.0 \times 10^{-12} \) in accordance with the visual interpretation. The results of the F-Test analysis are summarised in **Supplementary Table 2**.

**Protein Production, Crystallisation, data collection and structure determination**

The USP7 catalytic domain (residues 207-560) genetically fused with a C-terminal hexa-histidine tag was expressed in E.coli. BL21 cells were transformed with the corresponding expression plasmid and grown in Terrific broth (TB) and protein expression induced with 0.25 mM IPTG overnight at 16°C. After harvesting by centrifugation, cell pellets were resuspended in Lysis Buffer (40 mM TRIS-HCl, 500 mM NaCl, 1 mM AEBSF, 2 mM TCEP, 5 mM Imidazole, 0.1% Tween 20, pH 7.5) and lysed by sonication on ice. The soluble fraction was then loaded directly onto an IMAC column (5 mL HisTrap HP) pre-equilibrated with Lysis Buffer and the protein eluted with IMAC Buffer B (40 mM TRIS-HCl, 500 mM NaCl, 1 mM AEBSF, 2 mM
TCEP, 300 mM Imidazole, 0.1% Tween 20, pH 7.5). Fractions containing the desired protein were pooled and buffer exchanged by disalysis (MWCO 8,000-10,000 Da) against anion exchange (AEX) Buffer A (20 mM TRIS-HCl, 30 mM NaCl, 1 mM EDTA, 4 mM DTT, pH 8.0). The protein was then loaded onto a YMC-BioPro ion exchange column (15 x 120, 7.4 ml) pre-equilibrated with AEX Buffer A and eluted over 30 CV with a gradient of 0-50% AEX Buffer B (20 mM TRIS-HCl, 1M NaCl, 1 mM EDTA, 4 mM DTT, pH 8.0). Fractions were analysed by SDS-PAGE and those containing the desired protein were pooled and then further purified by SEC (HighLoad Superdex 75 column) using a running buffer of 10 mM TRIS-HCl, 100 mM NaCl, 4 mM DTT, pH 8. SEC fractions were analysed by SDS-PAGE and the pure fractions pooled and concentrated (Vivaspin column, MWCO 12KDa) to 5.3 mg/ml as measured by UV. A280nm.

Crystals of USP7 in complex with inhibitors were grown by hanging drop vapour diffusion. For the USP7/2 complex, a 14.2 mg/ml solution of USP7 (10 mM TRIS-HCl, 100 mM NaCl, 4mM TCEP) at pH 8.0 was pre-incubated with a 5.9-fold molar excess of 2 (50 mM in DMSO) for 3 h. 0.7 µl of the protein solution was then mixed with 0.7 µl of reservoir solution containing 100 mM TRIS-HCl (pH 7.75), 200 mM Li2-Sulfate, 29% (w/v) PEG4000 and equilibrated at 4°C over 0.4 ml of reservoir solution. For the USP7/4 complex, a 14.2 mg/ml solution of USP7 (10 mM TRIS-HCl, 100 mM NaCl, 4mM TCEP) at pH 8.0 was pre-incubated with an 8.9-fold molar excess of 4 (150 mM in DMSO) for 3 h. 2 µl of the protein solution was then mixed with 2 µl of reservoir solution (100 mM TRIS-HCl (pH 8.5), 200 mM Li2-Sulfate, 28% (w/v) PEG 4000) and equilibrated at 20°C over 0.4 ml of reservoir solution.
Diffraction data at 2.23 Å resolution for a USP7/2 crystal was collected with a Bruker MicroStar rotating anode generator equipped with Osmic optics. The structure was solved via molecular replacement using the PDB entry 1NB8 as a template. Iterative manual modelling in Coot and refinement using REFMAC5 resulted in the final model. 97% of backbone torsions for the final model are within the Ramachandran favoured regions, with 3% in the allowed regions. Diffraction data at 1.7 Å resolution for a USP7/4 crystal was collected at the ESRF synchrotron radiation source, id30a1, Grenoble. The structure was solved via molecular replacement using the USP7/2 structure as a template. Iterative manual modelling in Coot and refinement using REFMAC5 resulted in the final model. 96% of backbone torsions for the final model are within the Ramachandran favoured regions, with 4% in the allowed regions. The crystallography data collection and refinement statistics are provided in Supplementary Table 1. Composite omit maps were calculated in Phenix; ligand atoms were removed and mFo-DFc omit maps generated using the phenix.composite_omit_map tool and are shown in Supplementary Figures 5c and 6b. The atomic co-ordinates and structure factors for the USP7/2 and USP7/4 complex structures have been deposited in the PDB under accession codes 5N9R and 5N9T respectively.

Computational Chemistry

Images were created using Molecular Operating Environment (MOE, 2016.0802; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal,
Ab-initio calculations

Ab-initio quantum mechanical calculations to determine the relative strain induced by rotation of the phenyl ring were performed in Jaguar. In order to simplify the calculation and eliminate the effect of interactions between the phenyl ring and the piperazine, the ligand from the USP7/compound 2 complex was truncated and reproduced in both the (R-) and (S-) stereoisomers (Supplementary Figure 4). The geometries of both the (R-) and (S-) forms of the truncated ligand were optimized in Jaguar using a B3LYP/6-31G** basis set and the default optimisation settings. The dihedral angle between the phenyl ring and the methyl carbon attached to the adjacent chiral centre was manually adjusted to reflect the orientation of the phenyl ring observed in the crystal structure (37.7° between the phenyl and methyl for the active (R-) stereoisomer and 151.3° for the inactive (S-) stereoisomer). Single point energy calculations were then performed for each stereoisomer in Jaguar, using the B3LYP/6-31G** basis set and default settings.

Selectivity assays

Selectivity assays were performed against all DUBs available in the DUBprofiler™ panel (n=39 in total; Ubiquigent). Screening was performed at a fixed inhibitor
concentration of 100 µM. Data generated is displayed as a percentage inhibition of
total enzyme activity for each enzyme. Of note, USP47 was not part of the
DUBprofiler™ panel at Ubiquigent and so was not tested biochemically. Selectivity
against the protease (n=63) and kinase (SelectScreen; n=49) panels were performed
at Reaction Biology and Life Technologies respectively using a fixed inhibitor
concentration of 10 µM. Data generated is displayed as a percentage inhibition of
total enzyme activity for each enzyme. All data presented as mean ± s.d. (n >=2).

**Target engagement assay**

HCT116, LNCaP or RS4;11 cells were treated with vehicle (DMSO) or USP7 inhibitors
for 2 h. Following incubation, cells were washed extensively thrice with 1x PBS and
harvested in TE lysis buffer containing 50 mM TRIS-HCl (pH7.4), 150 mM NaCl, 5 mM
MgCl₂, 0.5 mM EDTA, 0.5% NP40, 10% Glycerol, 2 mM DTT and clarified cell lysates
(40 µg) incubated with the ubiquitin-propargylamine probe (Ub-PA; 8 µg/ml final
concentration) in assay buffer containing 50 mM TRIS-HCl (pH7.6), 5 mM MgCl₂, 250
mM Sucrose, 0.5 mM EDTA, 2 mM DTT for 30 min. The reaction was terminated by
the addition of LDS sample buffer (Life Technologies) and heated to 70°C. Samples
were then analysed by western blotting using the Cell Signalling anti-USP7 Ab
(#4833; 1/1000 dilution). EC₅₀ values were determined upon densitometry analysis.
Band intensities were quantified using ImageJ software where the upper
bands (USP7-Ub) and lower bands (USP7) were calculated as a percentage of the
 corresponding DMSO controls (-/+ Ub-PA) and values were then normalised to the
sum of the lower and upper bands for each concentration. Selectivity was assessed
using the Cell Signalling anti-USP4 (#2651, 1:1000), anti-USP11 (#GTX101446, 1:1000), anti-USP30 antibodies (#ab189518; 1:1000) and the Santa Cruz anti-USP47 antibody (#sc-100633, 1:500).

Cellular activity assay and western blotting

In order to assess the cellular activity of the compounds, HCT116 or MCF7 cells were treated with compounds or vehicle (DMSO) for a period of 2 h. Cells were subsequently harvested, lysed in radioimmuno-precipitation (RIPA) buffer containing 50 mM TRIS-HCl (pH 7.6), 150 mM NaCl, 1 mM EDTA, 1.0% NP40, 0.25% Na-deoxycholate and supplemented with a phosphatase (PhosSTOP, Roche) and protease inhibitor cocktail tablet (complete Mini, Roche). For the timecourse experiments, LnCaP and RS4;11 cells were treated with inhibitors as indicated and harvested at 8, 24, 48 and 72 h. Western blotting analyses were carried out using antibodies purchased from Cell Signalling; anti-USP7 (#4833; 1:1000 dilution), anti-Ser15p53 (#9284; 1:1000), anti-PARP (#9542, 1:1000), anti-cleaved Caspase 3 (#9664; 1:1000), Santa Cruz; anti-p53 (#sc-263; 1:500), Millipore; anti-MDM2 (#OP46; 1:200), anti-p21 (#05-345; 1:1000), Sigma; anti-β-actin (#A5316; 1:2000), HRP conjugated anti-rabbit (#A0545; 1:5000) and HRP conjugated anti-mouse (#A9917; 1:5000).

MDM2 ubiquitination assay (MSD)

SJSA-1 cells were seeded in 96 well plate format and treated with vehicle (DMSO) or inhibitors as indicated for 24 h. 7 dp titration curves were typically used from 10 pM to 20 µM in 1 log unit increments. Cells were then lysed and the total protein
concentrations determined with the BCA protein kit assay (ThermoFischer Scientific; #23227). Detection of the ubiquitinated and total MDM2 protein levels was performed using the MDM2 whole cell lysate kit as recommended by the manufacturer’s instructions (# K15168D; Meso Scale Discovery). Data was acquired on a MesoQuickPlex SQ120 reader. Results were normalised to the total MDM2 protein concentration and EC\textsubscript{50} values derived using GraphPadPrism (GraphPad Software, Inc, La Jolla, CA; four-parameter logistic function). Data presented as mean ± s.d. (n >=3).

**MDM2 ubiquitination levels (western blot)**

LNCaP cells were seeded in 100 mm dishes to achieve a confluency of 70-90% the following day. 10 µg of HA-tagged ubiquitin plasmid was transfected per plate using Lipofectamine 2000 (ThermoFischer), at a Lipofectamine:DNA ratio of 1:4. Media was replenished after 6 h and cells were harvested 48 h post-transfection. Cells were treated with 3 and ent-3 for 45 min to 1.5 h (as indicated). Protein G Sepharose beads (GE Healthcare) were stored in a 25% slurry in ethanol at 4\textdegree C. Immediately prior to use, 0.5 mL was aliquoted, centrifuged at 1000 x g and washed twice in 1 mL lysis buffer. Beads were then suspended in 1 mL RIPA lysis buffer and incubated with 50 µg anti-MDM2 (#Ab-1, Merck) 2 h at 4\textdegree C. The beads were then pelleted by centrifugation, resuspended in fresh lysis buffer and aliquoted between lysate samples (750 µg protein). The volume of each sample was normalised with lysis buffer and samples were allowed to incubate for 4 h. Beads were then washed 4 times in 1 mL lysis buffer in a succession of spin-resuspension cycles and boiled in 50
μL LDS sample buffer to dissociate bound proteins, prior to SDS-PAGE. Samples were analysed by western blotting using the anti-HA antibody (Thermo Scientific, 1:2000).

**Cell proliferation assay**

Cells were seeded in 96 well plate format (typically 2500 cells/well for MCF7, LNCaP, SJSA-1 and 5000 cells/well for the RS4;11 line) and treated after 24 h with increasing concentrations of compound (typically in a 9 dp curve ranging from 31 μM to 1.25 nM). Cell viability was assessed by CellTiter-Glo® using a Synergy 4 plate reader (BioTek) after 72 h (RS4;11) or 6 days (LNCaP, SJSA-1) as recommended by the manufacturer’s instructions (Promega; # G7571). Analysis and EC₅₀ values were derived using GraphPadPrism (GraphPad Software, Inc, La Jolla, CA; four-parameter logistic function). Data presented as mean ± s.d. (n >=3).

**Data Availability Statement**

Structural data that support the findings of this study have been deposited in the RCSB Protein Data Bank (PDB, [www.rcsb.org](http://www.rcsb.org)) with the accession numbers 5N9R and 5N9T. The authors declare that all other data supporting the findings of this study are available within the paper (and its Supplementary Information files).

**Statistics and reproducibility**

The data reported in this study represent the mean and standard deviation of at least 3 independent experiments as specified in the Figure legend (unless otherwise stated; **Figure 3**). For the mode of inhibition study, an F-test statistical analysis was
performed as described in details in the Online methods and Supplementary Table

2. 2.
SUPPLEMENTARY INFORMATION - Figures

Supplementary Figure 1

a

![Graph showing IC50 = 13.1 ± 0.8 μM](image)

b

![Graph showing Kd = 7.7 ± 0.1 μM](image)

c

![Bar chart showing activity recovery](image)
Supplementary Figure 1: Biochemical characterisation and binding validation of the reversible USP7 inhibitor 1. a, Biochemical inhibition of USP7 by 1. Activity was monitored using the full-length protein and the Ub-TMR isopeptide substrate. b, Representative orthogonal binding validation. Binding of 1 was confirmed by SPR with representative titration sensorgrams shown for inhibitor concentration of 0, 0.8, 4.0, 20.0 and 100.0 μM (data summarised in Table 1). c, Reversible binding demonstrated by high-dilution assay. Assay controls included ubiquitin, iodoacetamide and H₂O₂. DMSO: no compound control using the same DMSO final concentration for treatment with 1. Data reported as the mean of at least 3 independent experiments with standard deviations.

Supplementary Figure 2

Supplementary Figure 2: USP7 catalytic site and misaligned triad (H464, C223 and D481). Inter-residue distances are shown in Å. Left: USP7-Ub bound structure (1NBF), showing the properly aligned triad; middle: USP7 co-crystal structure with 2, showing a misaligned triad (the co-crystal structure with 4 has an essentially identical arrangement, with 0.27 Å pairwise RMSD between the triad residues); right: USP7 apo structure (1NB8), showing a similarly misaligned catalytic triad.
Supplementary Figure 3: Diagram showing the conformational change of the loop between Gly458 and Gly463 in response to binding of 2. The loop folds in towards the ubiquitin binding channel positioning His461 between the ligand and the catalytic cysteine. This results in a movement of 4.94 Å as indicated.
Supplementary Figure 4: Rationale for the enantioselective binding of 2 (vs ent-2) to USP7. Left: co-crystal structure of USP7 with 2. Middle: ent-2 cannot bind in an analogous conformation to 2 due to a steric clash with H461 and internal strain induced by movement of the phenyl ring further out of the plane of the C-H bond from the chiral centre. Right: Rotation of the phenyl ring cannot completely ameliorate the internal strain due to a steric clash with K420. Areas of steric clash are indicated by yellow circles and dotted lines. See main text and Online Methods section for details. The truncated ligand 7 used in the ab-initio strain calculations is shown below.

![Chemical Structure](image)
Supplementary Figure 5

a

b

Legend:
- polar
- acidic
- basic
- greasy
- proximity
- contour
- sidechain acceptor
- sidechain donor
- backbone acceptor
- backbone donor
- solvent residue
- metal complex
- H arene-H
- arene-arene
- arene-cation
- solvent contact
- metal/ion contact
- ligand
- receptor
- exposure
Supplementary Figure 5: Electron density, 2D plot, omit and surface charge maps for 2 in complex with USP7. a, 2mFo-DFc 2σ difference map of 2 in complex with USP7. b, 2D interaction plot of 2 in complex with USP7. c, mFo-DFc 3σ composite omit map for 2 in complex with USP7. d, Map showing the surface charge around the binding pocket for 2 in complex with USP7.

Supplementary Figure 6

a
Supplementary Figure 6: 2D plot, omit and surface charge maps for 4 in complex with USP7. a, 2D interaction plot of 4 in complex with USP7. b, mFo-DFc 3σ
composite omit map for 4 in complex with USP7. c, Map showing the surface charge around the binding pocket for 4 in complex with USP7.

**Supplementary Figure 7**

a

![Graph showing activity as a function of log [compound] (M).]

b

![Graph showing response (RU) over time (s).]

3 $IC_{50} = 0.006 \pm 0.002 \mu M$

ent-3 $IC_{50} = 2.4 \pm 0.3 \mu M$

3 $K_d = 0.002 \pm 0.001 \mu M$

ent-3 $K_d = 5.0 \pm 0.1 \mu M$
Supplementary Figure 7: Development of highly potent USP7 inhibitors. a, Biochemical inhibition of USP7 by 3 and its enantiomer ent-3 (>400-fold difference in potency). Biochemical activity was monitored using the full-length protein and the Ub-TMR isopeptide substrate. b, Orthogonal characterisation and binding validation by SPR. Representative sensorgrams for 3 and ent-3 are shown (data summarised in Table 1). Data reported as the mean of at least 3 independent experiments with standard deviations.
Supplementary Figure 8: Characterisation of the mode of inhibition of USP7 by 3. a, Representative Michaelis-Menten plot. Inhibition was performed as indicated using varying concentrations of Ub-AMC (from 0 to 25 µM) and inhibitor concentration (from 0 to 25 nM, as indicated). R-square values for fitting were >0.95 in all cases. b, Lineweaver-Burk plot (plotted using data from A). c, Eadie-Hofstee plot of USP7 inhibition by 3. Results from A were plotted as $V/[S]$ against $V$ where $V$ is the reaction velocity for the inhibitor concentrations and $S$ is the Ub-AMC concentrations (as indicated). The slope of the plots (parallel lines indicative of unchanged $K_M$ values) reveal a non-competitive mode of inhibition. R-square values for fitting were >0.95 in all cases. F-Test statistical analysis shows the lowest Chi$^2$ of 572.9 for the non-competitive model and the best probability of 1.72E-12 (see Supplementary Table 2).
Supplementary Figure 9: Biochemical inhibition of USP7 by 3 with increasing concentration of substrate. Biochemical activity was monitored using the full-length USP7 protein and the Ub-AMC substrate using the concentration range as indicated. Analysis and IC_{50} values were derived using GraphPad Prism (GraphPad Software, Inc, La Jolla, CA; four-parameter logistic function). Data represents the mean of 2 independent experiments.

Supplementary Figure 10
Supplementary Figure 10: Selectivity profiles of Compound 3 against proteases and kinases. a, Representative selectivity profile against the protease family. 3 was screened at 10 µM fixed concentration against a panel of 63 representative proteases across all family members including cysteine proteases (Reaction Biology).

b, Representative selectivity profile against a panel of kinases. 3 was screened at 10 µM against a panel of 49 representative kinases (SelectScreen®, Life technologies). Data reported as the mean of at least two independent experiments with standard deviations.
Supplementary Figure 11

a

![Image](image_url)

b

<table>
<thead>
<tr>
<th>USP7 residues</th>
<th>USP47 corresponding residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln351</td>
<td>Pro328</td>
</tr>
<tr>
<td>Met407</td>
<td>Lys391</td>
</tr>
<tr>
<td>Met410</td>
<td>Asp394</td>
</tr>
<tr>
<td>Asp459</td>
<td>Ser498</td>
</tr>
<tr>
<td>Asn460</td>
<td>Ala499</td>
</tr>
<tr>
<td>His461</td>
<td>Ala500</td>
</tr>
</tbody>
</table>
Supplementary Figure 11: Binding selectivity of inhibitor for USP7 vs the closest homologue USP47. a, Homology model of USP47 with the residues in the binding site which differ between USP47 and USP7 highlighted, leading to severe steric clashes with the USP7 ligands. The USP47 homology model was created in MOE. Residues 188-564 of Uniprot sequence Q96K76 were aligned to the co-crystal structure of USP7 and 4. Fifteen main chain models were created, with 3 samples for each sidechain at 310 K, using the AMBER12:EHT force field. The bound ligand from the co-crystal structure was not included in the model and N-terminal and C-terminal outgaps were not modelled. Energy minimisation was applied to intermediate models and the final model, to a gradient of 1.0 and 0.5, respectively. The Protonate3D function was performed on the final model before minimisation. b, Table showing the 6 USP7 residues in the binding site situated within 5.0 Å of the ligand which differ in USP47 (as shown).

Supplementary Figure 12

a
<table>
<thead>
<tr>
<th></th>
<th>0.1 µM</th>
<th>0.5 µM</th>
<th>1.0 µM</th>
<th>0.1 µM</th>
<th>0.5 µM</th>
<th>1.0 µM</th>
<th>0.1 µM</th>
<th>0.5 µM</th>
<th>1.0 µM</th>
<th>0.1 µM</th>
<th>0.5 µM</th>
<th>1.0 µM</th>
<th>0.1 µM</th>
<th>0.5 µM</th>
<th>1.0 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMOSO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ent-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAM66533B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **USP7**
- **p53**
- **p21**
- **MDM2**
- **cl. caspase 3**
- **PARP**
- **cl. PARP**
- **β-actin**

---

**Ub-PA**

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>USP7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP7-Ub</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Compound 3**

**ent-3**

**Vinculin**

---

**Densitometry (%)**

- **EC50 = 0.006 µM**
- **EC50 = 37.3 µM**
Supplementary Figure 12: Characterisation of effects of 3 in RS4;11 cell line

a, Anti-proliferative activity of 3 and ent-3 in the RS4;11 cell line. Cellular viability was measured 72 h post treatment using the CellTiter-Glo® assay (Promega). Viability data reported as the mean of at least 3 independent experiments with standard deviations. EC₅₀ values compiled in Supplementary Table 3. b, Representative western blotting experiment in RS4;11 demonstrating the time-dependent apoptosis induction following treatment with the inhibitors (as indicated). c, Target engagement of 3 and ent-3 in RS4;11 cells. Cells were treated with 3 and ent-3 for 2 h (as indicated), lysed and the ubiquitin-propargylamine (Ub-PA) probe was added. Samples were subsequently analysed by western blotting probing for USP7. Densitometry and quantitative analyses were performed as described previously. Full blots and cut membranes are shown in Supplementary Figure 16e.

Supplementary Figure 13

a
Supplementary Figure 13: Characterisation of effects of 3 in LNCaP cell line. a, Anti-proliferative effects demonstrated in solid tumor cell line (LNCaP). Cellular viability
was measured after 6 days as described previously. Viability data reported as the
mean of at least 3 independent experiments with standard deviations. b,
Representative western blotting analysis in the LNCaP cell line. Cells were treated
with inhibitor as indicated. EC<sub>50</sub> values compiled in Supplementary Table 3. c, Target
engagement: cells were treated with 3 and ent-3 for 2 h (as indicated), lysed and the
ubiquitin-propargylamine (Ub-PA) probe was added. Samples were subsequently analysed
by western blotting probing for USP7. Densitometry analysis was performed for the upper
(USP7-Ub) and lower (USP7) bands. Signal was subsequently normalised to the combined
densitometry values for the upper and lower bands after background subtraction
(quantitative analysis shown from 2 independent experiments). d, No target engagement
was detected up to 50 µM against the most closely related USP7 family member (USP47).
Similar results were obtained when probing against additional non-related USPs including
USP4, 19, 28. (+) and (−) signs represent the presence or absence of the Ub-PA probe.
Concentration (Conc.). e, Increased levels of MDM2 ubiquitination following treatment with
3. HA-tagged ubiquitin (HA-Ub) was expressed in LNCaP cells. Cells were treated with 1.0 µM
of compound or vehicle (as indicated) and MG132 (20 µM). Cells were lysed after 45 min or
1.5 h and MDM2 was immunoprecipitated from all samples. Western blotting analysis was
performed using an antibody against the HA-tag (upper panel) or MDM2 (lower panel). Full
blots and cut membranes shown in Supplementary Figures 16f-h.
Supplementary Figure 14

a

b

c
Supplementary Figure 14: Direct benchmarking of 3 to MDM2 antagonists. a, RS4;11 cells were treated with USP7 inhibitors or the MDM2 antagonists nutlin-3a, RG7112 and SAR405838 (as indicated). Cellular viability was measured after 72 h using CellTiter-Glo® (Promega) as described previously. b, Repeat of (a) in the LNCaP prostate cell line. Viability was assessed 6 days post treatment as described above. c, Differential activity between MDM2 antagonists and USP7 inhibitor in the osteosarcoma SJSA-1 cell line (as indicated). Viability was assessed 6 days post treatment as described before. Data reported as the mean of three independent experiments with standard deviations. All EC₅₀ values compiled in Supplementary Table 3. d, Target engagement in the SJSA-1 cell line. Full blots and cut membranes
are shown in **Supplementary Figure 16i**. The assay was performed as described previously for the HCT116 and LNCaP cell lines.

**Supplementary Figure 15**

![Supplementary Figure 15: Comparison of the key residues forming the site for Compounds 2 and 4 with the equivalent residues from published crystal structures of other USPs. USP7 is shown in salmon; other USPs (USP14, PDB 2AYN; USP8, PDB 2GFO; USP2, PDB 2IBI; USP4, PDB 2Y6E; USP21, PDB 3I3T; USP5, PDB 3IHP; USP2a, PDB 3NHE; USP18, 5CHT; USP46, PDB 5CVM; USP12, PDB 5K16) in cyan and grey. 2 is shown in orange. The key binding residues involved in binding which are indicated on the diagram are conserved in each of the USP’s. Significant alterations in sidechain conformation are observed for some residues, particularly F409, which is observed in two different positions for the published structures.](image-url)
depending on whether the USP is in an apo state or in complex with ubiquitin.

However neither of these clusters overlay precisely with the position of F409 in the current structure. Y514 also adopts a novel conformation in the current structure when in complex with the inhibitors reported here, compared to that observed in other published USP crystal structures. The position of D295 deviates from that of the equivalent residues in other USPs. This is due to the unusual activation mechanism of USP7, in which a conformational change in the switching loop (residues 285-291) is required for ubiquitin binding. Compounds 2 and 4 bind to the inactivated state of USP7, in which this conformational change has not occurred.

**Supplementary Figure 16**

![Compound 3](image1)

![ent-3](image2)
**Supplementary Figure 16: Full western blots or cut membranes.**

a, Target engagement in HCT116 cell line (from Figure 4a).
b, Selectivity profile against various DUBs in HCT116 cell line (from Figure 4b).
c, Western blotting in HCT116 cell line corresponding to Figure 5a.
d, Western blotting in MCF7 cell line corresponding to Figure 5b.
e, Western blotting in RS4;11 cell line corresponding to Supplementary Figure 12b (timepoint as indicated).
f, Western blotting in LNCaP cell line corresponding to Supplementary Figure 13b (timepoint as indicated).
g, Target engagement in LNCaP cell line using 3 and ent-3 corresponding to Supplementary Figure 13c.
h, Selectivity profile against various DUBs in LNCaP cell line.
corresponding to Supplementary Figure 13d. i, Target engagement using 3 in SJSA-1 cell line corresponding to Supplementary Figure 14d.

SUPPLEMENTARY INFORMATION - Tables

Supplementary Table 1: Crystallographic data collection and refinement statistics (molecular replacement)

<table>
<thead>
<tr>
<th>Data collection</th>
<th>USP7/2, 5N9R</th>
<th>USP7/4, 5N9T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>$P_{2_1}$</td>
<td>$P_{2_1}$</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$, $b$, $c$ (Å)</td>
<td>74.94, 66.94, 81.09</td>
<td>75.51, 67.62, 81.05</td>
</tr>
<tr>
<td>$\alpha$, $\beta$, $\gamma$ (°)</td>
<td>90.00, 105.09, 90.00</td>
<td>90.00, 105.90, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.00-2.23 (2.29-2.25)</td>
<td>29.92-1.73 (1.82-1.73)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.06 (0.31)</td>
<td>0.08 (0.65)</td>
</tr>
<tr>
<td>$I/\sigma(I)$</td>
<td>16.4 (4.5)</td>
<td>5.7 (1.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.2 (99.4)</td>
<td>94.3 (94.3)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.5 (3.4)</td>
<td>2.2 (2.2)</td>
</tr>
</tbody>
</table>

| Refinement      |              |              |
| Resolution (Å)  | 30.00-2.23 (2.29-2.23) | 29.92-1.73 (1.82-1.73) |
| No. reflections | 35975 (1848) | 73492 (3826) |
| $R_{work}$ / $R_{free}$ | 0.148 / 0.212 | 0.175 / 0.232 |
| No. atoms       |              |              |
| Protein         | 5763         | 5813         |
| Ligand/ion      | 97           | 115          |
| Water           | 528          | 880          |
| $B$-factors     |              |              |
| Protein         | 64.54        | 43.62        |
| Ligand/ion      | 58.54        | 40.11        |
| Water           | 64.08        | 52.40        |
| R.m.s. deviations |            |              |
| Bond lengths (Å) | 0.009        | 0.011        |
| Bond angles (°) | 1.378        | 1.475        |

* Data were collected from a single crystal for each structure. Values in parentheses are for the highest-resolution shell.
**Supplementary Table 2: F-Test statistical analysis.** See Online Methods section and main text for more details.

<table>
<thead>
<tr>
<th>Inhibition Mode</th>
<th>Analysis 1</th>
<th>Analysis 1</th>
<th>Analysis 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-competitive Inhibition</td>
<td>Non-competitive Inhibition</td>
<td>Competitive Inhibition</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inhibition Mode</th>
<th>Analysis 2</th>
<th>Analysis 2</th>
<th>Analysis 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Competitive Inhibition</td>
<td>Uncompetitive Inhibition</td>
<td>Uncompetitive Inhibition</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Number of data points:</td>
<td>76</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>Chi2 Value (1)</td>
<td>572.881</td>
<td>572.881</td>
<td>3307.05</td>
</tr>
<tr>
<td>Chi2 Value (2)</td>
<td>3307.05</td>
<td>1352.63</td>
<td>1352.63</td>
</tr>
<tr>
<td>F statistic:</td>
<td>5.77267</td>
<td>2.36111</td>
<td>2.4449</td>
</tr>
<tr>
<td>Probability:</td>
<td>1.72E-12</td>
<td>0.000313566</td>
<td>0.00017992</td>
</tr>
</tbody>
</table>

**Supplementary Table 3: Anti-proliferative activity of USP7 inhibitors and benchmarking to MDM2 antagonists**

<table>
<thead>
<tr>
<th>Cell lines / EC50 (nM)</th>
<th>3</th>
<th>nutlin-3a</th>
<th>RG7112</th>
<th>SAR405838</th>
<th>ent-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS4;11</td>
<td>2.0±0.9</td>
<td>140±38</td>
<td>62.0±15.0</td>
<td>32.0±10.0</td>
<td>2200±730</td>
</tr>
<tr>
<td>LNCaP</td>
<td>29.0±4.0</td>
<td>620±160</td>
<td>170.0±25.0</td>
<td>64.0±25</td>
<td>9300±2900</td>
</tr>
<tr>
<td>SJSA-1</td>
<td>&gt;20000</td>
<td>960±45</td>
<td>320±43</td>
<td>100±25</td>
<td>&gt;20000</td>
</tr>
</tbody>
</table>

See Online Methods section and main text for more details. Enantiomer (ent-). Data reported is the mean of at least 3 independent experiments ± standard deviations.
SUPPLEMENTARY NOTE – Synthetic Procedures

Abbreviations and Acronyms

aq: aqueous; Boc: tert-butyloxycarbonyl; DCM: dichloromethane; br: broad; d: doublet; DIPEA: diisopropylethylamine; DME: dimethoxyethane; DMF: N,N-dimethylformamide; DMSO: dimethylsulfoxide; EDC: N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride; EtOAc: ethyl acetate; PE: petroleum ether 40/60; ee: enantiomeric excess; ESI: electrospray ionization; h: hour; HATU: N-[(dimethylamino)-1H-1,2,3-triazolo-[4,5-b]pyridin-1-ylmethylene]-N-methylmethanaminium hexafluorophosphate N-oxide; HPLC: high pressure liquid chromatography; LC: liquid chromatography; LCMS: liquid chromatography mass spectrometry; M: molar; m/z: mass-to-charge ratio; MeCN: acetonitrile; MeOH: methanol; min: minutes; MS: mass spectrometry; m: multiplet; NMR: nuclear magnetic resonance; q: quartet; quint: quintet; Rt: retention time; RT: room temperature; s: singlet; SPhos: 2-dicyclohexylphosphino-2′,6′-dimethoxybiphenyl; TFA: trifluoroacetic acid; THF: tetrahydrofuran; t: triplet; v/v: volume per unit volume; Walphos SL-W008-2:((S)-1-((S)-2-[2-(dicyclohexylphosphino)phenyl]ferrocenyl)ethylbis[3,5-bis(trifluoromethyl)phenyl]phosphine;

General Experimental Conditions

Solvents and reagents
Common organic solvents that were used in reactions (e.g. THF, DMF, DCM, and methanol) were purchased anhydrous from Sigma-Aldrich® in Sure/Seal™ bottles and were handled appropriately under nitrogen. Water was deionised using an Elga PURELAB Option-Q. All other solvents used (i.e. for work-up procedures and purification) were generally HPLC grade and were used as supplied from various commercial sources. Unless otherwise stated, all starting materials used were purchased from commercial suppliers and used as supplied.

**Microwave synthesis**

Microwave experiments were carried out using a Biotage Initiator™ Eight system.

**Flash chromatography**

Purification of compounds by flash chromatography was achieved using a Biotage Isolera Four system using the stated cartridges.

**NMR spectroscopy**

$^1$H NMR spectra were recorded at ambient temperature using a Bruker Avance (300 MHz), Bruker Avance III (400 MHz) or Bruker Ascend (500 MHz) spectrometer. All chemical shifts (δ) are expressed in ppm. Residual solvent signals were used as an internal standard and the characteristic solvent peaks were corrected to the reference data outlined in *J. Org. Chem.*, 1997, 62, p7512-7515; in other cases, NMR solvents contained tetramethylsilane, which was used as an internal standard.

**Liquid Chromatography Mass Spectrometry (LCMS)**
Liquid Chromatography Mass Spectrometry (LCMS) experiments to determine retention times (RT) and associated mass ions were performed using the following methods:

**Method A:** The system consisted of an Agilent Technologies 6130 quadrupole mass spectrometer linked to an Agilent Technologies 1290 Infinity LC system with UV diode array detector and autosampler. The spectrometer consisted of an electrospray ionization source operating in positive and negative ion mode. LCMS experiments were performed on each sample submitted using the following conditions: LC Column: Agilent Eclipse Plus C18 RRHD, 1.8 µm, 50 x 2.1 mm maintained at 40 °C. Mobile phases: A) 0.1% (v/v) formic acid in water; B) 0.1% (v/v) formic acid in acetonitrile.

<table>
<thead>
<tr>
<th>Gradient Time (min)</th>
<th>Flow (mL/min)</th>
<th>%A</th>
<th>%B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.5</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>1.80</td>
<td>0.5</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2.20</td>
<td>0.5</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2.50</td>
<td>0.5</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>3.00</td>
<td>0.5</td>
<td>80</td>
<td>20</td>
</tr>
</tbody>
</table>

**Method B:** The system consisted of an Agilent Technologies 6140 single quadrupole mass spectrometer linked to an Agilent Technologies 1290 Infinity LC system with UV diode array detector and autosampler. The spectrometer consisted of a multimode ionization source (electrospray and atmospheric pressure chemical ionizations)
operating in positive and negative ion mode. LCMS experiments were performed on
each sample submitted using the following conditions: LC Column: Zorbax Eclipse
Plus C18 RRHD, 1.8 µm, 50 x 2.1 mm maintained at 40 °C. Mobile phases: A) 0.1%
(v/v) formic acid in water; B) 0.1% (v/v) formic acid in acetonitrile.

<table>
<thead>
<tr>
<th>Gradient Time (min)</th>
<th>Flow (mL/min)</th>
<th>%A</th>
<th>%B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>1.0</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>1.80</td>
<td>1.0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2.20</td>
<td>1.0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2.21</td>
<td>1.0</td>
<td>95</td>
<td>5</td>
</tr>
<tr>
<td>2.50</td>
<td>1.0</td>
<td>95</td>
<td>5</td>
</tr>
</tbody>
</table>

Nomenclature

Unless otherwise indicated, the nomenclature of structures was determined using
the ‘Convert Structure to Name’ function of ChemBioDraw Ultra 12.0.2
(CambridgeSoft/PerkinElmer).

Compound 1

Step 1: 1-(3-Phenylpropanoyl)piperidin-4-one (8):
A solution of tert-butyl 4-oxopiperidine-1-carboxylate (5 g, 25.1 mmol) in DCM (25 mL) and TFA (9.67 mL, 125 mmol) was stirred at RT for 24 h before the solvent was removed in vacuo and the product dried under high vacuum. To a stirred suspension of the TFA salt in dry DCM (125 mL) was added DIPEA (13.2 mL, 75.0 mmol) before 3-phenylpropanoic acid (4.52 g, 30.1 mmol), EDC (6.26 g, 32.6 mmol) and DMAP (0.307 g, 2.51 mmol) were added. The reaction mixture was stirred at RT for 18 h, diluted with DCM (150 mL) and washed with saturated NaHCO₃(aq) (250 mL). The aqueous layer was further extracted with DCM (75 mL) before the combined organic phases were washed with 3% HCl(aq) (150 mL) and brine (150 mL). The combined organic phases were passed through a Biotage phase separator, concentrated in vacuo and purified by flash chromatography (Biotage KP-Sil 100 g cartridge, 0-100% EtOAc in PE) to give the title compound as pale yellow oil. LCMS (Method A): Rₜ = 0.89 min, m/z = 232 [M+H]+. ¹H NMR (300 MHz, CDCl₃): δ 7.37 – 7.15 (m, 5H), 3.89 (t, J = 6.4 Hz, 2H), 3.66 (t, J = 6.3 Hz, 2H), 3.04 (dd, J = 8.5, 6.8 Hz, 2H), 2.73 (dd, J = 8.4, 6.9 Hz, 2H), 2.44 (t, J = 6.4 Hz, 2H), 2.26 (t, J = 6.3 Hz, 2H). ¹³C NMR (75 MHz, CDCl₃): δ 206.87, 171.14, 141.06, 128.71, 128.59, 126.51, 44.24, 41.15, 40.94, 40.92, 35.01, 31.70.

**Step 2: 3-Phenyl-1-(1-oxa-6-azaspiro[2.5]octan-6-yl)propan-1-one (9):**
To a solution of trimethylsulfonium iodide (4.59 g, 22.5 mmol) in dry DMSO (20 mL) was added a 60% dispersion of NaH in mineral oil (0.899 g, 22.5 mmol). The resulting mixture was stirred at RT for 1 h before a solution of 1-(3-phenylpropanoyl)piperidin-4-one (2.08 g, 8.99 mmol) in dry DMSO (10 mL) was added. The reaction mixture was stirred at 50 °C for 2 h before it was allowed to cool to RT and quenched by the addition of water (100 mL). The resulting mixture was extracted with Et₂O (3 x 50 mL), the combined organic extracts were washed with brine, dried over Na₂SO₄, concentrated *in vacuo*, and the product was purified by flash chromatography (Biotage KP-Sil 50 g cartridge, 0-60% EtOAc in PE) to give the title compound (1.41 g, 64%) as a colourless oil. LCMS (Method A): $R_t = 1.07$ min, $m/z = 246$ [M+H]+. ¹H NMR (300 MHz, CDCl₃): $\delta$ 7.38 – 7.09 (m, 5H), 4.21 – 4.05 (m, 1H), 3.70 – 3.53 (m, 1H), 3.42 (tdd, $J = 13.5, 10.0, 3.6$ Hz, 2H), 3.08 – 2.89 (m, 2H), 2.77 – 2.55 (m, 4H), 1.81 (ddd, $J = 14.2, 10.0, 4.5$ Hz, 1H), 1.68 (ddd, $J = 13.7, 8.1, 4.5$ Hz, 1H), 1.49 – 1.29 (m, 2H). ¹³C NMR (75 MHz, CDCl₃): $\delta$ 170.77, 141.37, 128.65, 128.58, 126.34, 57.04, 53.84, 44.29, 40.55, 35.23, 33.64, 32.79, 31.73.

**Step 3:** 7-Bromo-3-((4-hydroxy-1-(3-phenylpropanoyl)piperidin-4-yl)methyl)thieno[3,2-d]pyrimidin-4(3H)-one (1):

![Chemical Structure]

A mixture of 3-phenyl-1-(1-oxa-6-azaspiro[2.5]octan-6-yl)propan-1-one (206 mg, 0.840 mmol), 7-bromothieno[3,2-d]pyrimidin-4(3H)-one¹ (176 mg, 0.763 mmol) and Cs₂CO₃ (298 mg, 0.916 mmol) in DMF (8 mL) was heated at 80 °C for 16 h. Upon cooling
to RT, the mixture was diluted with saturated NH$_4$Cl(aq) (40 mL) and extracted with DCM (3 x 30 mL) using a Biotage phase separator. The combined organic phases were concentrated in vacuo and the residue was purified by flash chromatography (Biotage KP-NH 11 g cartridge, 0-30% MeOH in DCM) to give the title compound (204 mg, 56%) as a colourless solid. LCMS (Method A): $R_f$ = 1.22 min (purity >98% at 254 nm), $m/z$ = 476, 478 [M+H$^+$]. $^1$H NMR (500 MHz, DMSO-$d_6$): $\delta$ 8.39 (s, 2H), 7.33 – 7.10 (m, 5H), 4.96 (s, 1H), 4.11 – 4.06 (m, 1H), 4.06 – 3.96 (m, 2H), 3.67 – 3.59 (m, 1H), 3.25 – 3.16 (m, 1H), 2.93 – 2.85 (m, 1H), 2.80 (t, $J$ = 7.7 Hz, 2H), 2.67 – 2.53 (m, 2H), 1.49 – 1.32 (m, 4H). $^{13}$C NMR (126 MHz, DMSO-$d_6$): $\delta$ 169.58, 156.66, 153.40, 151.22, 141.44, 132.49, 128.37, 128.17, 125.78, 121.88, 108.30, 69.21, 53.64, 40.78, 36.92, 34.82, 34.14, 33.86, 30.87. HRMS (TOF MS ES+): $m/z$ [M + H$^+$] Calcd for C$_{21}$H$_{23}$N$_3$O$_3$SBr 476.0644, found 476.0642.

**Compound 2**

Step 1: tert-Butyl 4-((7-bromo-4-oxothieno[3,2-d]pyrimidin-3(4H)-yl)methyl)-4-hydroxypiperidine-1-carboxylate (10):
A mixture of tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (640 mg, 3.00 mmol), 7-bromothieno[3,2-d]pyrimidin-4(3H)-one1 (578 mg, 2.50 mmol) and Cs₂CO₃ (978 mg, 3.00 mmol) in DMF (8.3 mL) was heated at 80 °C for 16 h. Upon cooling to RT, the mixture was diluted with saturated NH₄Cl(aq) (40 mL) and extracted with DCM (3 x 30 mL) using a Biotage phase separator. The combined organic phases were concentrated in vacuo and the residue was purified by flash chromatography (GraceResolv silica 80 g cartridge, 0-100% EtOAc in cyclohexane) to give the title compound (867 mg, 78%) as a pale yellow solid. LCMS (Method A): Rᵣ = 1.31 min, m/z = 466, 468 [M+Na]+. ¹H NMR (300 MHz, DMSO-d₆): δ 8.40 (s, 2H), 4.95 (s, 1H), 4.11 – 3.95 (m, 2H), 3.76 – 3.54 (m, 2H), 3.14 – 2.91 (m, 2H), 1.57 – 1.27 (m, 4H), 1.39 (s, 9H).

¹³C NMR (75 MHz, CDCl₃): δ 157.99, 154.72, 154.01, 149.91, 131.82, 122.92, 109.26, 79.80, 70.44, 55.50, 39.46, 34.98, 28.48.

Step 2: 7-Bromo-3-((4-hydroxypiperidin-4-yl)methyl)thieno[3,2-d]pyrimidin-4(3H)-one (11):

A solution of tert-butyl 4-((7-bromo-4-oxothieno[3,2-d]pyrimidin-3(4H)-yl)methyl)-4-hydroxypiperidine-1-carboxylate (250 mg, 0.563 mmol) was stirred in DCM (3 mL) and TFA (3 mL) for 20 min before the reaction was purified using a 10 g SCX-2 cartridge (10% MeOH in DCM then 20% 7 M NH₃ in MeOH in DCM) to give the title compound (189 mg, 98%) as a colourless solid. LCMS (Method A): Rᵣ = 0.36 min, m/z = 344, 346 [M+H]+. ¹H NMR (400 MHz, CDCl₃): δ 8.30 (s, 1H), 7.78 (s, 1H), 4.11 (s, 2H), 2.95 – 2.85
(m, 4H), 1.70 – 1.58 (m, 2H), 1.56 – 1.46 (m, 2H). $^{13}$C NMR (101 MHz, CDCl$_3$): δ 158.07, 154.07, 150.04, 131.75, 123.03, 109.35, 70.63, 55.50, 42.03, 36.30.

Step 3: (R)-7-Bromo-3-((4-hydroxy-1-(3-phenylbutanoyl)piperidin-4-yl)methyl)thieno[3,2-d]pyrimidin-4(3H)-one (2):

![Chemical Structure]

DIPEA (0.359 mL, 2.06 mmol) was added to a suspension of 7-bromo-3-((4-hydroxypiperidin-4-yl)methyl)thieno[3,2-d]pyrimidin-4(3H)-one (177 mg, 0.514 mmol), (R)-3-phenylbutanoic acid (101 mg, 0.617 mmol) and HATU (235 mg, 0.617 mmol) in DCM (10.3 mL). After 10 min, the reaction was quenched by the addition of saturated NaHCO$_3$(aq) (30 mL) and the mixture was extracted with DCM (3 x 40 mL) using a Biotage phase separator. The combined organic phases were concentrated in vacuo and the residue was purified by flash chromatography (GraceResolv silica 40 g cartridge, 0-100% EtOAc in cyclohexane then 0-30% MeOH in EtOAc) to give the title compound (134 mg, 53%) as a colourless solid. LCMS (Method A): RT = 1.28 min (purity >95% at 254 nm), $m/z$ = 490, 492 [M+H]$^+$. $^1$H NMR (500 MHz, DMSO-d$_6$): δ 8.57 – 8.16 (m, 2H), 7.46 – 7.03 (m, 5H), 5.00 – 4.88 (m, 1H), 4.13 – 3.87 (m, 3H), 3.71 – 3.58 (m, 1H), 3.25 – 3.08 (m, 2H), 2.91 – 2.77 (m, 1H), 2.66 – 2.52 (m, 2H), 1.58 – 1.13 (m, 7H).

$^{13}$C NMR (126 MHz, DMSO-d$_6$): δ 169.08 + 169.06 (conformers), 156.67 + 156.62 (conformers), 153.38, 151.23 + 151.20 (conformers), 146.66 + 146.54 (conformers), 132.49, 128.18 + 128.15 (conformers), 126.88 + 126.85 (conformers), 125.93 + 125.89
(conformers), 121.87, 108.30, 69.20 + 69.15 (conformers), 53.61, 41.01 + 40.90
(conformers), 40.22 + 40.19 (conformers), 36.87, 36.21 + 35.99 (conformers), 34.93 +
34.80 (conformers), 34.25 + 34.12 (conformers), 22.06 + 21.86 (conformers). HRMS
(TOF MS ES+): m/z [M + H]^+ Calcd for C_{22}H_{25}N_{3}O_{3}SBr 490.0800, found 490.0793.

**Compound 5 (ent-2)**

Compound ent-2 was made by an identical method as 2 starting from 7-Bromo-3-((4-
hydroxypiperidin-4-yl)methyl)thieno[3,2-d]pyrimidin-4(3H)-one but using (S)-3-
phenylbutanoic acid in the coupling step.

**Compound 3**

![Chemical structure of Compound 3](image)

**Step 1: 3-Bromo-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (12):**

![Chemical structure of 3-Bromo-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one](image)

Bromine (3.34 mL, 64.7 mmol) was added to a suspension of 2-methyl-2H-
pyrazolo[4,3-d]pyrimidin-7(6H)-one\textsuperscript{2} (3.24 g, 21.6 mmol) in AcOH (21.6 mL) in a
reaction tube. The tube was sealed and the mixture was heated at 95 °C for 18 h
before being cooled to RT and 1:1 EtOH/Et₂O (100 mL) was added. Saturated sodium thiosulfate (50 mL) was added and once the colour had dissipated, the product was isolated by filtration. The resulting precipitate was washed with water (3 x 50 mL) and dried under high vacuum at 75 °C to give the title compound (4.80 g, 97%) as a pale yellow solid. LCMS (Method A): Rₜ = 0.40 min, m/z = 229, 231 [M+H]+. ¹H NMR (500 MHz, DMSO- d₆): δ 12.02 (br. s, 1H), 7.84 (d, J = 3.5 Hz, 1H), 4.07 (s, 3H).


Step 2: tert-Butyl 4-((3-bromo-2-methyl-7-oxo-2H-pyrazolo[4,3-d]pyrimidin-6(7H)-yl)methyl)-4-hydroxypiperidine-1-carboxylate (13):

A suspension of tert-butyl 1-oxa-6-azaspiro[2.5]octane-6-carboxylate (5.59 g, 26.2 mmol), 3-bromo-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (3 g, 13.1 mmol) and Cs₂CO₃ (4.69 g, 14.4 mmol) in DMF (44 mL) was heated at 80 °C for 18 h. Upon cooling to RT, the reaction was quenched by the addition of saturated NH₄Cl(aq) (200 mL) and the mixture was extracted with EtOAc (3 x 50 mL). The combined organic phases were passed through a Biotage phase separator, concentrated in vacuo and the residue was purified by flash chromatography (GraceResolv silica 220 g cartridge, 0-100% EtOAc in cyclohexane then 0-15% MeOH in EtOAc) to give the title compound (4.22 g, 73%) as a pale yellow foam. LCMS (Method A): Rₜ = 1.06 min, m/z = 464, 466 [M+Na]+. ¹H NMR (500 MHz, CDCl₃): δ 7.93 (s, 1H), 4.18 – 3.98 (m, 2H), 4.12 (s, 3H), 3.95 – 3.75 (m, 2H), 3.22 – 3.05 (m, 2H), 1.68 – 1.48 (m, 4H), 1.43 (s, 9H).
\[ \delta 157.49, 154.78, 147.18, 136.10, 135.65, 109.25, 79.78, 70.73, 55.36, 39.63, 39.38, 35.13, 28.55 \]

**Step 3:** 3-Bromo-6-((4-hydroxypiperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (14):

A solution of tert-butyl 4-((3-bromo-2-methyl-7-oxo-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-yl)methyl)-4-hydroxypiperidine-1-carboxylate (10 g, 22.6 mmol) in DCM (75 mL) and TFA (37.5 mL) was stirred for 5 min before the reaction mixture was purified using 2 x 70 g SCX-2 cartridges (1:1 DCM/MeOH then 1:1 DCM/7 M in NH₃ in MeOH) to give the title compound (7.69 g, 99%) as a colourless foam. LCMS (Method A): Rᵣ = 0.238 min, m/z = 342, 344 [M+H]+. ¹H NMR (500 MHz, DMSO-d₆): δ 8.06 (s, 1H), 4.87 (s, 1H), 4.08 (s, 3H), 3.96 (s, 2H), 2.78 (dd, J = 8.1, 3.6 Hz, 4H), 1.51 (dt, J = 14.7, 7.6 Hz, 2H), 1.36 (dt, J = 13.0, 3.8 Hz, 2H). ¹³C NMR (126 MHz, DMSO-d₆): δ 165.96, 155.76, 148.16, 135.65, 135.02, 108.82, 68.91, 53.16, 40.64, 34.23.

**Step 4:** (R)-3-Bromo-6-((4-hydroxy-1-(3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (15):

Page 12 of 21
Using a pressure equalized dropping funnel, DIPEA (40.8 mL, 234 mmol) was added
dropwise over 35 min to a solution of 3-bromo-6-((4-hydroxypiperidin-4-yl)methyl)-
2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (20 g, 58.4 mmol), (R)-3-
phenylbutanoic acid (10.6 g, 64.3 mmol) and HATU (24.5 g, 64.3 mmol) in DCM (784
mL). After 2 h, the reaction was quenched by the addition of saturated NaHCO3(aq) (1
L) and the mixture was extracted with DCM (3 x 100 mL). The combined organic
phases were passed through a Biotage phase separator, concentrated in vacuo and
the residue was purified by flash chromatography (Biotage KP-Sil 340 g cartridge, 0-
100% EtOAc in cyclohexane then 0-30% MeOH in EtOAc) to give the title compound
(26.8 g, 94%) as colourless solid. LCMS (Method A): Rt = 1.06 min, m/z = 488, 490
[M+H]+. 1H NMR (500 MHz, DMSO-d6): δ 8.07 – 7.97 (m, 1H), 7.34 – 7.08 (m, 5H),
4.90 – 4.80 (m, 1H), 4.08 (s, 3H), 4.08 – 3.84 (m, 3H), 3.69 – 3.58 (m, 1H), 3.26 – 3.10
(m, 2H), 2.91 – 2.81 (m, 1H), 2.65 – 2.52 (m, 2H), 1.54 – 1.22 (m, 4H), 1.20 (d, J = 7.0
Hz, 3H). 13C NMR (126 MHz, DMSO-d6): δ 169.07 + 169.05 (conformers), 155.78 +
155.73 (conformers), 148.07, 146.67 + 146.54 (conformers), 135.65, 135.04, 128.19
+ 128.15 (conformers), 126.86 + 126.85 (conformers), 125.95 + 125.88 (conformers),
108.84, 69.28 + 69.23 (conformers), 53.10, 41.03 + 40.91 (conformers), 40.22, 39.31,
36.88, 36.20 + 35.97 (conformers), 34.96 + 34.83 (conformers), 34.21 + 34.09
(conformers), 22.05 + 21.86 (conformers).

Steps 5 and 6: (R)-3-(4-(Aminomethyl)phenyl)-6-((4-hydroxy-1-(3-
phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-
one (3):
A 1 L round bottom flask was charged with (R)-3-bromo-6-((4-hydroxy-1-(3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (15.6 g, 31.9 mmol), SPhos (786 mg, 1.92 mmol), Pd(OAc)$_2$ (215 mg, 0.958 mmol), (4-(((tert-butoxycarbonyl)amino)methyl)phenyl)boronic acid (12.0 g, 47.9 mmol) and K$_3$PO$_4$ (13.6 g, 63.8 mmol) and was degassed by evacuating and refilling the flask with N$_2$ three times using a Schlenk manifold. Under a N$_2$ atmosphere, $n$-butanol (128 mL) was added before the flask was evacuated and refilled with N$_2$ three times using a Schlenk manifold. The reaction mixture was heated at 100 °C for 2 h before being cooled to RT and concentrated in vacuo. The residue was partitioned between water (350 mL) and EtOAc (250 mL). The phases were separated and the aqueous phase was extracted using EtOAc (2 x 50 mL). The combined organic phases were dried over Na$_2$SO$_4$, concentrated in vacuo and the residue was purified by flash chromatography (Biotage KP-Sil 340 g cartridge, 0-100% EtOAc in cyclohexane then 0-20% MeOH in EtOAc), the impure fractions were re-purified using a GraceResolv silica 80 g cartridge, 0-100% EtOAc in cyclohexane then 0-20% MeOH in EtOAc) to give (R)-tert-butyl 4-((6-((4-hydroxy-1-(3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-7-oxo-6,7-dihydro-2H-pyrazolo[4,3-d]pyrimidin-3-yl)benzylcarbamate (15.4 g, 78%) as a colourless foam. LCMS (Method A): $R_f = 1.37$ min, $m/z = 615$ [M+H]$^+$. A solution of (R)-tert-butyl 4-((4-hydroxy-1-
(3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-7-oxo-6,7-dihydro-2H-
pyrazolo[4,3-d]pyrimidin-3-yl)benzylcarbamate (8.73 g, 14.2 mmol) in DCM (14 mL)
and TFA (14 mL) was stirred for 90 min before the reaction was purified using 4 x 10
g SCX-2 cartridges in parallel (1:10 MeOH in DCM then 1:3 7 M in NH₃ in MeOH in
DCM). The basic phases were combined and concentrated to give the crude product
(4.89 g). The DCM/MeOH phases were concentrated and the residue was purified
using 4 x 10 g SCX-2 cartridges in parallel (1:10 MeOH in DCM then 1:3 7 M in NH₃ in
MeOH in DCM) and the basic phases were combined and concentrated to give the
crude product (2.62 g). The combined crude products (7.51 g) were purified by flash
chromatography (GraceResolve silica 120 g cartridge, 0-30% dilute NH₃ in MeOH in
DCM) to give the title compound (6.9 g, 94%) as a colourless solid after freeze drying.

LCMS (Method A): Rₜ = 0.71 min (purity >99% at 254 nm), m/z = 515 [M+H]⁺. ¹H NMR
(500 MHz, DMSO-d₆): δ 8.03 – 7.92 (m, 1H), 7.69 – 7.43 (m, 4H), 7.30 – 7.21 (m, 4H),
7.19 – 7.12 (m, 1H), 4.87 (s, 1H), 4.10 (s, 3H), 4.07 – 3.86 (m, 3H), 3.81 (s, 2H), 3.70 –
3.61 (m, 1H), 3.28 – 3.12 (m, 2H), 2.93 – 2.83 (m, 1H), 2.66 – 2.52 (m, 2H), 2.31 (br. s,
2H), 1.57 – 1.23 (m, 4H), 1.21 (d, J = 6.9 Hz, 3H). ¹³C NMR (126 MHz, DMSO-d₆): δ
169.07 + 169.05 (conformers), 156.24 + 156.19 (conformers), 147.27, 146.68,
146.56, 145.04, 136.00, 134.59 + 134.41 (conformers), 129.20, 128.20 + 128.16
(conformers), 127.37, 126.86, 125.95 + 125.89 (conformers), 125.31, 69.35 + 69.30
(conformers), 52.94, 45.29, 41.07 + 40.96 (conformers), 40.25 + 40.23 (conformers),
39.51, 36.92, 36.20 + 35.99 (conformers), 35.02 + 34.91 (conformers), 34.27 + 34.14
(conformers), 22.06 + 21.88 (conformers). HRMS (TOF MS ES+): m/z [M + H]⁺ Calcd
for C₂₉H₃₅N₆O₃ 515.2771, found 515.2773.
**Compound 6 (ent-3)**

Compound ent-3 was made by an identical method as 3 starting from 3-Bromo-6-((4-hydroxypiperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one but using (S)-3-phenylbutanoic acid in the coupling step.

**Compound 4**

![Compound 4](image)

**Step 1: (E)-4,4,4-Trifluoro-3-phenylbut-2-enoic acid (16):**

LiOH (132 mg, 5.5 mmol) was added to solution of (E)-ethyl 4,4,4-trifluoro-3-phenylbut-2-enoate (J. Fluorine Chem. 2013, 152, 56) (1.22 g, 5 mmol) in THF (10 mL) and water (5 mL) at RT. After 1 h, the pH of the reaction mixture was adjusted to pH 4 by the addition of 1 M HCl(aq) and the mixture was extracted with DCM (3 x 10 mL) using a Biotage phase separator. The combined organic phases were concentrated and the product was dried *in vacuo* to give (E)-4,4,4-trifluoro-3-phenylbut-2-enoic acid (1.06 g, 98%) as a colourless solid. LCMS (Method A): R<sub>T</sub> = 1.27 min, m/z = 215 [M-H]. ¹H NMR (400 MHz, DMSO-<i>d</i>₆): δ 13.21 (s, 1H), 7.48 – 7.39 (m, 3H), 7.31 (dd, J
= 6.6, 2.9 Hz, 2H), 6.84 (q, $J = 1.5$ Hz, 1H). $^{19}$F NMR (376 MHz, DMSO-$d_6$): $\delta$ -65.64. $^{13}$C NMR (101 MHz, DMSO-$d_6$): $\delta$ 165.33, 137.24 (q, $J = 30.0$ Hz), 130.70, 129.33, 128.59, 128.41, 127.63 (q, $J = 5.4$ Hz), 122.87 (q, $J = 274.5$ Hz).

Step 2: (R)-4,4,4-Trifluoro-3-phenylbutanoic acid (17):

A suspension of bis(norbornadiene)rhodium(I) tetrafluoroborate (30.5 mg, 81.8 $\mu$mol) and Walphos SL-W008-2 (77 mg, 81.8 $\mu$mol) in MeOH (100 mL) in a 300 mL glass reactor autoclave was degassed with N$_2$. After 30 min, a solution was obtained and (E)-4,4,4-trifluoro-3-phenylbut-2-enoic acid (900 mg, 4.16 mmol) was added. The reaction was stirred under 5.8 bar H$_2$ for 24 h before being concentrated in vacuo to give the title compound (1 g, >100%) as a colourless solid. This material was used without further purification. LCMS (Method A): RT = 1.24 min, $m/z = 217 \ [M-H]^-$.

$^{1}$H NMR (500 MHz, CDCl$_3$): $\delta$ 7.40 – 7.28 (m, 5H), 3.87 (pd, $J = 9.3$, 4.9 Hz, 1H), 3.07 (dd, $J = 16.8$, 4.9 Hz, 1H), 2.93 (dd, $J = 16.8$, 9.7 Hz, 1H). Spectral data matches that reported in the literature. A 2 M TMS-diazomethane solution in hexanes (0.110 mL, 0.220 mmol) was added to a solution of (R)-4,4,4-trifluoro-3-phenylbutanoic acid (24 mg, 0.110 mmol) in MeOH (20 µL) and toluene (0.2 mL). After 1 hour, the reaction mixture was purified directly by chromatography (10 x 0.75 cm silica plug, 0-30% Et$_2$O in pentane) to give (R)-methyl 4,4,4-trifluoro-3-phenylbutanoate (22 mg, 86%) as a pale yellow oil. The Et$_2$O/pentane were removed carefully to avoid loss of material.
The ee was measured using a Diacel OJ-H column (250 x 4.6 mm, 5 micron) and 1% iPrOH in hexane as the mobile phase. The racemic sample gave 2 peaks: $R_T = 23.220 \text{ min (49.9021\%)}$ and $R_T = 26.365 \text{ min (50.0979\%)}$. This sample gave the same 2 peaks: $R_T = 23.401 \text{ min (0.6863\%)}$ and $R_T = 26.099 \text{ min (99.3137\%)}$. This corresponds to an ee measurement for the acid that is >98%. The acid was assigned as $R$ by comparison with literature in which the authors use the same OJ stationary phase under which conditions the (R)-enantiomer has the longer retention time.

Step 3: (R)-3-Bromo-6-((4-hydroxy-1-(4,4,4-trifluoro-3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (18):

![Chemical Structure]

DIPEA (0.224 mL, 1.28 mmol) was added to a suspension of 3-bromo-6-((4-hydroxypiperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (132 mg, 0.385 mmol), (R)-4,4,4-trifluoro-3-phenylbutanoic acid (70 mg, 0.32 mmol) and HATU (122 mg, 0.32 mmol) in DCM (20 mL). After 16 h, the reaction was quenched by the addition of saturated NaHCO$_3(aq)$ (20 mL) and the mixture was extracted with DCM (3 x 20 mL) using a Biotage phase separator. The combined organic phases were concentrated in vacuo and the residue was purified by flash chromatography (Biotage KP-NH 11 g cartridge, 0-100% EtOAc in cyclohexane then 0-30% MeOH in EtOAc) to give the title compound (142 mg, 82%) as a colourless solid. LCMS (Method B): $R_T = 1.19 \text{ min, } m/z 542, 544 [M+H]^+$. $^1$H NMR (500 MHz, DMSO-d$_6$): $\delta$

8.07 – 7.97 (m, 1H), 7.49 – 7.20 (m, 5H), 4.88 (s, 1H), 4.15 – 4.04 (m, 1H), 4.08 (s, 3H),
4.00 – 3.83 (m, 3H), 3.80 – 3.71 (m, 1H), 3.29 – 3.11 (m, 2H), 3.00 – 2.77 (m, 2H), 1.64 – 1.12 (m, 4H). $^{13}$C NMR (126 MHz, DMSO-$d_6$): $\delta$ 166.60 + 166.50 (conformers), 155.78 + 155.72 (conformers), 148.08 + 148.06 (conformers), 135.66, 135.05 + 135.03 (conformers), 134.81 – 134.60 (m, due to conformers and coupling with CF$_3$), 129.16, 128.37 + 128.34 (conformers), 128.02 + 127.98 (conformers), 108.86, 69.25 + 69.15 (conformers), 53.15 + 53.06 (conformers), 45.65 – 44.83 (m, due to conformers and coupling with CF$_3$), 40.79, 37.33 + 37.24 (conformers), 34.82 + 34.76 (conformers), 34.05 + 33.99 (conformers), 31.44. Signal for NCH$_3$ is under DMSO and the CF$_3$ is not observed.

Steps 4 and 5: (R)-3-(4-(Aminomethyl)phenyl)-6-((4-hydroxy-1-(4,4,4-trifluoro-3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (4):

A mixture of (R)-3-bromo-6-((4-hydroxy-1-(4,4,4-trifluoro-3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-2H-pyrazolo[4,3-d]pyrimidin-7(6H)-one (250 mg, 0.461 mmol), (4-(((tert-butoxycarbonyl)amino)methyl)phenyl)boronic acid (231 mg, 0.922 mmol), Pd(PPh$_3$)$_4$ (53 mg, 0.046 mmol), K$_3$PO$_4$ (391 mg, 1.84 mmol), 1,4-dioxane (3 mL) and water (1 mL) in a reaction tube was degassed by bubbling N$_2$ for 20 min. The reaction tube was sealed and the reaction was heated at...
130 °C under microwave irradiation for 30 min. The reaction mixture was diluted 
with saturated NH₄Cl(aq) (40 mL) and extracted with DCM (3 x 30 mL) using a Biotage 
phase separator. The combined organic phases were concentrated *in vacuo* and the 
residue was purified by flash chromatography (40 g GraceResolv silica, 0-100% EtOAc 
in cyclohexane then 0-10% MeOH in EtOAc) to give (R)-tert-butyl 4-((6-((4-hydroxy-1-
(4,4,4-trifluoro-3-phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-7-oxo-6,7-dihydro-
2H-pyrazolo[4,3-d]pyrimidin-3-yl)benzylcarbamate (200 mg, 65%) as a pale yellow 
solid. LCMS (Method A): R₉ = 1.47 min, *m/z* 669 [M+H]⁺.

A solution of (R)-tert-butyl 4-((6-((4-hydroxy-1-(4,4,4-trifluoro-3-
phenylbutanoyl)piperidin-4-yl)methyl)-2-methyl-7-oxo-6,7-dihydro-2H-pyrazolo[4,3-
d]pyrimidin-3-yl)benzylcarbamate (200 mg, 0.300 mmol) in DCM (2 mL) and TFA (1 
ml) was stirred at RT for 15 min. The reaction mixture was purified using a 10 g SCX-
2 cartridge (MeOH then 7 M in NH₃ in MeOH). The basic phases were combined, 
concentrated *in vacuo* and the residue was purified by flash chromatography 
(GraceResolv silica 24 g cartridge, 0-100%, EtOAc in cyclohexane then 0-25% MeOH 
in EtOAc; then Biotage KP-NH 11 g cartridge, 0-100%, EtOAc in cyclohexane then 0-
10% MeOH in EtOAc) to give the title compound (90 mg, 52%) as a colourless solid 
after lyophilization. LCMS (Method B): R₉ = 0.81 min (purity >99% at 254 nm), *m/z* 
569 [M+H]⁺. ¹H NMR (500 MHz, DMSO-d₆): δ 8.02 – 7.93 (m, 1H), 7.70 – 7.60 (m, 2H), 
7.57 – 7.44 (m, 2H), 7.43 – 7.38 (m, 2H), 7.38 – 7.28 (m, 3H), 4.88 (s, 1H), 4.26 – 4.02 
(m, 1H), 4.10 (s, 3H), 4.02 – 3.85 (m, 3H), 3.81 (s, 2H), 3.80 – 3.71 (m, 1H), 3.26 – 3.11 
(m, 2H), 3.02 – 2.80 (m, 2H), 2.43 (s, 2H), 1.61 (td, J = 12.7, 12.1, 4.3 Hz, 0.5H), 1.44 – 
1.32 (m, 2.5H), 1.27 – 1.17 (m, 1H). ¹³C NMR (126 MHz, DMSO-d₆): δ 166.60 + 166.50 
(conformers), 156.24 + 156.18 (conformers), 147.27 + 147.25 (conformers), 145.22,
136.02, 134.85 – 134.24 (m, due to conformers and coupling with CF₃), 129.19, 129.16, 128.37, 128.34, 128.28, 128.22, 128.02, 127.99, 127.34, 127.23, 126.07, 126.04, 125.26, 69.30 + 69.21 (conformers), 52.98 + 52.89 (conformers), 45.46 – 44.90 (m, due to conformers and coupling with CF₃), 45.35, 40.83, 39.39, 37.36 + 37.28 (conformers), 34.89 + 34.80 (conformers), 34.09 + 34.04 (conformers), 31.44.

The signals between 130 – 125 ppm are complex due to the conformers and the CF₃ signals. HRMS (TOF MS ES+): m/z [M + H]+ Calcd for C₂₉H₃₂N₆O₃F₃ 569.2488, found 569.2509.

References

1 Bae, I. H. et al. PCT Int. Appl. (2013), WO 2013100632.