Accepted Manuscript

Post-HTS case report and structural alert: promiscuous 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one actives verified by ALARM NMR

Jayme L. Dahlin, J. Willem M. Nissink, Subhashree Francis, Jessica M. Strasser, Kristen John, Zhiguo Zhang, Michael A. Walters

Please cite this article as: Dahlin, J.L., Nissink, J.W.M., Francis, S., Strasser, J.M., John, K., Zhang, Z., Walters, M.A., Post-HTS case report and structural alert: promiscuous 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one actives verified by ALARM NMR, *Bioorganic & Medicinal Chemistry Letters* (2015), doi: http://dx.doi.org/10.1016/j.bmcl.2015.08.020

Post-HTS case report and structural alert: promiscuous 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one actives verified by ALARM NMR

Jayme L. Dahlin[a,b], J. Willem M. Nissink[c], Subhashree Francis[d], Jessica M. Strasser[d], Kristen John[d], Zhiguo Zhang[e] and Michael A. Walters[d*]

[a] Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

[b] Medical Scientist Training Program, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

[c] AstraZeneca, Oncology Innovative Medicines, Cambridge, CB4 0WG, UK

[d] Institute for Therapeutics Discovery and Development, University of Minnesota, Minneapolis, MN 55414, USA

[e] Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

[*] Corresponding author. Tel.: +1 612 626 6864; E-mail address: mwalters@umn.edu (M.A. Walters).

Conflicting interests | The authors hereby declare no conflicting interests pertaining to the material in this manuscript.

Abbreviations | Acetyl-CoA, acetyl coenzyme A; ALARM NMR, a La assay to detect reactive molecules by nuclear magnetic resonance; Asf1, anti-silencing function 1; AZ, AstraZeneca; BME, β-mercaptoethanol; BSA, bovine serum albumin; CoA, coenzyme A; CPM, *N*-[4-(7-diethylamino-4-methylcoumarin-3-yl)phenyl]maleimide; dH3-H4; *Drosophila* histone H3-H4; DMSO, dimethyl sulfoxide; DNA, deoxyribonucleic acid; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GSH, glutathione; H3K27, histone H3 lysine 27; H3K27ac, histone H3 lysine 27 acetylation; H3K56, histone H3 lysine 56; H3K56ac, histone H3 lysine 56 acetylation; HAT, histone acetyltransferase; HMQC, heteronuclear multiple quantum coherence; HPLC, high-performance liquid chromatography; HRMS, high-resolution mass spectrometry; HRP-PR, horseradish peroxidase-phenol red; HTS, high-throughput screen or high-throughput screening; IC$_{50}$, half maximal inhibitory concentration; IPTG, isopropyl β-D-1-

thiogalactopyranoside; log*D*, distribution coefficient; log*P*, partition coefficient; *m/z*, mass-to-charge ratio; LRMS-ESI, low-resolution mass spectrometry-electrospray ionization; MeCN, acetonitrile; MeOH, methanol; MS, mass spectrometry; NMR, nuclear magnetic resonance; PAINS, pan-assay interference compounds; pBSF, negative log of binomial survivor function; REOS, Rapid Elimination Of Swill; Rtt109, regulator of Ty1 transposition 109; SAR, structure-activity relationship; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SIR, structure-interference relationship; TFA, trifluoroacetic acid; UPLC, ultra-performance liquid chromatography; Vps75, vacuolar protein sorting 75.

ABSTRACT

Despite its wide use, not every high-throughput screen (HTS) yields chemical matter suitable for drug development campaigns, and seldom are 'go/no-go' decisions in drug discovery described in detail. This case report describes the follow-up of a 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one active from a cell-free HTS to identify small-molecule inhibitors of Rtt109-catalyzed histone acetylation. While this compound and structural analogs inhibited Rtt109-catalyzed histone acetylation *in vitro*, further work on this series was halted after several risk mitigation strategies were performed. Compounds with this chemotype had a poor structure-activity relationship, exhibited poor selectivity among other histone acetyltransferases, and tested positive in a β-lactamase counter-screen for chemical aggregates. Furthermore, ALARM NMR demonstrated compounds with this chemotype grossly perturbed the conformation of the La protein. In retrospect, this chemotype was flagged as a 'frequent hitter' in an analysis of a large corporate screening deck, yet similar compounds have been published as screening actives or chemical probes versus unrelated biological targets. This report − including the decision-making process behind the 'no-go' decision − should be informative for groups engaged in post-HTS triage and highlight the importance of considering physicochemical properties in early drug discovery.

- 3 -
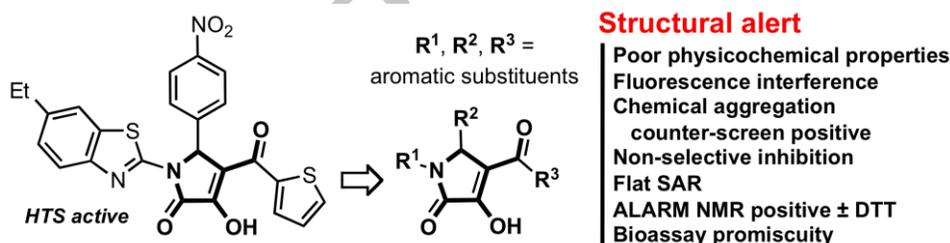
High-throughput screening (HTS) has become an indispensible part of both academic and industrial early drug discovery. Identifying useful chemical matter for drug development campaigns is the goal for most real and virtual HTS. However, not every discovery campaign is a 'success' by this metric. Sometimes HTS fails to identify any confirmed actives, which can be the result of assay design, the chemical library employed, and/or the druggability of the target. Other points where discovery campaigns can fail to progress include the post-HTS triage and the hit-to-lead phases. There are many reasons why projects fail to progress past these critical stages: inability to confirm primary assay activity by orthogonal assays, poor lead-like properties, target non-selectivity, or synthetic inaccessibility, to name a few[1]. However, despite their potential utility, rarely are the details of the decision-making process behind 'go/no-go' decisions described in the scientific literature[2-4].

Rtt109 is a fungal histone acetyltransferase (HAT) that catalyzes the acetylation of histone H3 lysine 56 (H3K56ac)[5]. Rtt109 and H3K56ac are crucial for replication-coupled nucleosome assembly in yeast, and fungi with defects in Rtt109 functioning or H3K56 acetylation show defects in cell growth and proliferation, genomic stability, and resistance to genotoxins[6]. Rtt109 forms a stable complex with the histone chaperone Vps75. *In vivo*, the histone chaperone Asf1 is essential for H3K56ac, and the Asf1-H3-H4 complex is considered the physiological substrate for the Rtt109-Vps75 complex[6]. Importantly, there is no known mammalian Rtt109 homolog, which has led to the hypothesis that selective, potent inhibitors of Rtt109-catalyzed histone acetylation can function as minimally-toxic antifungal agents[7-10].

Our group reported the results of a cell-free fluorometric HTS to identify small-molecule inhibitors of Rtt109-catalyzed histone acetylation[11]. This assay utilized the thiol-scavenging probe *N*-[4-(7-diethylamino-4-methylcoumarin-3-yl)phenyl]maleimide (CPM) to quantify the amount of free coenzyme A (CoA) byproduct produced by the Rtt109 HAT reaction. The primary assay screened 225K compounds, and 1.5K primary actives were initially identified. The post-HTS triage utilized substructure filters to remove pan-assay interference compounds (PAINS), a series of compounds with substructures that are associated with promiscuous bioactivity and assay interference, often due to non-specific thiol reactivity[12-15]. These compounds are generally considered nonviable lead compounds. The post-HTS triage also included several counter-screens to remove assay interference compounds. The screening cascade identified only three non-PAINS compounds with confirmed activity suitable for more detailed follow-up studies. Many of the compounds initially discarded were thiol-reactive, and compounds with several

prominent chemotypes were shown to form covalent adducts with biological thiols including CoA, glutathione (GSH), and protein cysteines[16].

In this case report, we describe the follow-up for one of the three compounds surviving our initial triage, compound **1a**. Work on this compound was eventually halted, though only after considerable time and resources were spent investigating it as a potential lead compound. Despite its promising initial *in vitro* potency and confirmed activity in two orthogonal assays, compound **1a** and close analogs failed to show selectivity towards Rtt109 or any meaningful structure-activity relationship (SAR). Several lines of experimental evidence including ALARM NMR suggest the promiscuity of compound **1a** is due to its tendency to form chemical aggregates and engage in non-specific compound-protein interactions. Importantly, we describe the rationale for its eventual dismissal as well as the accompanying experimental processes that ultimately drove this decision. We discuss strategies that, in retrospect, could have led to a quicker 'no-go' decision. It is hoped this case report can enhance the process of making 'go/no-go' decisions, highlight the scholarly importance of investigating the causes behind assay promiscuity, and guide better data- and medicinal chemistry-driven decisions earlier in the discovery process before consuming increasingly precious resources.

As described above, our group screened approximately 225K small-molecules for inhibitors of Rtt109-catalyzed histone acetylation[11]. During an extensive triage, the approximately 1.5K actives from the primary HTS were triaged by cheminformatics and several counter-screens to yield only three active compounds before the commencement of more in-depth studies such as the ones described in this report. One of these three remaining compounds, named **1a**, and its associated 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one chemotype **1** ('chemical structural motif') are the subjects of this case report. In the primary HTS, it had a single-point percent inhibition of 47% at 10 µM final compound concentration. Notably, this production run of the HTS did not include detergent in the assay buffer. Several descriptors were calculated for compound **1a**, indicating it did not posses many 'lead-like' properties (Table 1)[17, 18]. However, the paucity of active compounds from the HTS (that were not PAINS) led us to follow-up on this compound, despite its lack of desirable lead-like traits including calculated physicochemical properties. Notably, compound **1a** was not flagged as a PAINS in our cheminformatics filters[13]. It showed promising activity at low micromolar concentrations *in vitro* in early dose-response confirmation experiments (Figure 1A). These follow-up experiments were performed with a standard amount of non-ionic detergent (0.01% Triton X-100), as including detergents is now a

standard strategy to mitigate micelle formation in HTS. These data demonstrated that the assay readout for compound **1a** in the HTS format was dose-dependent, and that the primary HTS result was not the result of a random error[19-21].

The activity of compound **1a** was quickly re-confirmed with a commercially obtained sample that was re-purified in-house by HPLC prior to re-testing (Figure 1A). To note, the original library sample of compound **1a** was of commercial origin with an unspecified synthetic route. Our in-house purification was performed to mitigate the risk of bioactive compound impurities[22, 23]. We did not attempt enantiomeric separation on the samples for compound **1a** or its analogs, as this project was in the early stage and the chemotype **1** would likely under facile epimerization in solution. As additional quality-control, we confirmed the structural identity and purity of the commercial sample to avoid follow-up on a sample with bioactive impurities or an incorrectly assigned chemical structure (Figures S1-S5)[24-26]. We note the structure of compound **1a** allows for tautomerization. Ab initio calculations (B3LYP/6-31G**; gas-phase) predict compound **1a** exists primarily in the tautomeric form depicted throughout this text. However, calculations (SM6/B3LYP/6-31G**//B3LYP/6-31G**) that also take into account solvation do not rule out some fraction of the molecules existing as other tautomers in solution (Table S1).

Importantly, the inhibitory activity of compound **1a** was reconfirmed in an orthogonal slot blot assay, a Western blot that utilizes antibodies to specific histone modifications to directly quantify the acetylated histone products (H3K56ac) in reaction aliquots, rather than the CoA by-product of the HAT reaction (Figure 1B). These data confirmed that compound **1a** was capable of inhibiting Rtt109-catalyzed histone acetylation *in vitro* in both the primary CPM-based HTS assay (using samples from two independent sources) and an orthogonal antibody-based assay for the commercial re-supply sample.

Although they are relatively straightforward and inexpensive, we and others have demonstrated CPM-based thiol-scavenging HTS assays are prone to a variety of compound-mediated interference mechanisms[9, 11, 16, 27]. This includes fluorescence interference, either by fluorescent compounds that can give false-negative readouts, or fluorescence quenchers, which can lead to false-positive readouts[28]. Another source of assay interference are reactive compounds that form adducts with the assay reagents. One example is nucleophilic compounds, which can form adducts with the electrophilic CPM to form fluorescent adducts, leading to either false-positive or false-negative readouts depending on the fluorescent properties of the adducts. Another

concerning interference mechanism is through electrophilic compounds, which can (1) trap thiol by-products of the enzymatic reaction, leading to a false-positive readout, or (2) react with protein residues, which most often leads to off-target effects and bioassay promiscuity.

As part of the Rtt109 post-HTS triage, we tested compound **1a** in a series of counter-screens for the aforementioned interference types. In the context of the CPM-based readout, compound **1a** did not form fluorescent adducts with any of the assay reagents (CPM, acetyl-CoA, CoA) nor was it fluorescent in the assay buffer by itself (data not shown). One concern with fluorescence-based assays is fluorescence interference[28-33]. Indeed, compound **1a** was able to quench the fluorescence readout fluorescence quenching counter-screen, albeit at relatively high compound concentrations (Figure 2B). Consistent with a quenching mechanism, absorption spectra of compound **1a** in assay buffer showed the compound absorbed light in the UV and violet regions (Figure 2B). Fluorescence emission sweeps of compound **1a** in assay buffer did not show appreciable fluorescence when the compound was excited at 405 nm (data not shown). These data show that compound **1a** is capable of absorbing the assay excitation wavelength, but not capable of significantly fluorescing in the same wavelength as the fluorescent CPM-CoA adduct when excited at 405 nm. Lastly, we tested compound **1a** for evidence of thiol-trapping in another counter-screen and found it showed a nearly identical interference pattern as the quenching counter-screen (Figure 2C). Given that both the thiol-trapping and quenching counter-screen share the same readouts, one can deduce that the decrease in readout in the thiol-trapping counter-screen is due to fluorescence interference and not CoA-trapping. To summarize to this point, the slot blot data demonstrate compound **1a** can inhibit Rtt109-catalyzed histone acetylation *in vitro*, whereas the CPM assay readout for compound **1a** reflects a mixture of true inhibition and assay interference. With this admittedly less-than-ideal profile − and given the scarcity of compounds identified in our HTS that could inhibit Rtt109-catalyzed histone acetylation *in vitro* − we still chose to cautiously continue more experiments with compound **1a**.

Compound stability is another important consideration throughout the drug discovery process[34]. Previously, we and others reported classes of compounds that were not stable to assay conditions[16, 35]. One class, *p*-hydroxyarylsulfonamides, decomposed to form a reactive intermediate (likely a quinone) capable of forming covalent adducts with biological thiols including protein cysteines residues[16]. Therefore, we assessed the stability of compound **1a**. Unremarkably, samples of compound **1a** incubated in our HTS-like conditions did not show gross evidence of instability in the assay buffer when analyzed over time by UPLC-MS (Figure

S6). These data demonstrated compound **1a** was grossly stable to the assay conditions and that its inhibitory activity *in vitro* was not due to a degradation product detectable by UPLC-MS.

To reconcile the observed activity in the CPM-based HTS assay and the orthogonal slot blot assay with the assay interference, compound **1a** was also tested in a second orthogonal assay. This method utilizes the radiolabeled substrate [^{3}H]-acetyl-CoA and quantifies the amount of [^{3}H]-acetate incorporated into histone substrates. Importantly, this readout is not subject to the interference in the CPM-based method and has an improved signal:noise ratio. However, its throughput is much lower, it is more costly, and it requires additional safety precautions and disposal procedures[36]. Compound **1a** demonstrated reproducible, low micromolar activity (IC$_{50}$ = 6 ± 2 μM; Table 2). During the commercial re-supply of compound **1a**, we also purchased several close structural analogs that were also commercially available in an attempt to establish preliminary evidence of SAR ('SAR-by-commerce'). One of the favorable attributes of this series was the number of analogs commercially available. For example, 612 structural analogs with at least 80% 2-D similarity were available from a popular vendor at one point (eMolecules, accessed 2 December 2014). Furthermore, this chemical scaffold appeared amenable to chemical analog expansion by straightforward synthetic routes[37-42]. To this end, we obtained commercial analogs **1b**-**1u** and **2a**-**2l** for further testing versus Rtt109-Vps75. In our experience, one major drawback to SAR-by-commerce (compared to performing in-house synthesis) is that the choice of analogs is limited by commercial availability. The plan at the time was to generate more preliminary data with readily available commercial compounds, which if the data showed promise, could then be used to justify a more resource-intensive medicinal chemistry effort.

The structural analogs **1b**-**1p** all showed low micromolar IC$_{50}$ values in the same radiolabeled orthogonal HAT assay (Table 2). Notably, compounds with an *N*-alkyl substituent such as **1s**-**1u** were inactive (Table 2). We additionally tested a small number of compounds from the AZ corporate library, compounds **1v**-**1z**. Of this subset, only compound **1z** showed any activity versus Rtt109-Vps75. The other AZ compounds, which had an *N*-alkyl substituent (**1v**-**1y**), were inactive versus Rtt109-Vps75. We also tested the most structurally-similar 4-aroyl-1,5-disubstituted-3-amino-2*H*-pyrrol-2-ones (chemotype **2**, compounds **2a**-**2l**) that were commercially available. All of these analogs failed to inhibit Rtt109-catalyzed histone acetylation at less than 125 μM final compound concentrations (Table 3). These latter data suggested that the 3-hydroxy component was an important structural component for inhibiting Rtt109 activity *in vitro*. While we were confident that **1a** and its close analogs with chemotype **1** could inhibit

Rtt109-Vps75 HAT activity, the lack of an interpretable SAR at several structural positions raised our suspicions about non-specific inhibition[12, 43].

With its activity observed in one primary and two orthogonal assays, compound **1a** was subsequently assessed for redox-activity, a widely reported source of promiscuous bioactivity[44-46]. Using a horseradish peroxidase-phenol red (HRP-PR) surrogate assay for H_2O_2 production[47], compound **1a** was found to be only slightly redox-active at relatively high compound concentrations (i.e. 160 µM), independent of DTT in the HAT assay buffer (Figure S7). Compounds **1l** (+ DTT only), **2c**, **2d,** and **2j** showed similar low-level redox-activity at high compound concentrations (data not shown). All the other structural analogs of compound **1a** were inactive in this assay both in the presence and absence of DTT (data not shown).

Generally consistent with these results at low micromolar compound concentrations, 8/8 (100%) of compound **1a** structural analogs were inactive for H_2O_2 production in a previously reported HTS for redox-active compounds (Table S3)[44]. In this screen, compounds were tested at 10 µM final concentrations. Along with our observations that Rtt109-Vps75, p300 and Gcn5-Ada2-Ada3 are not overly sensitive to H_2O_2 in our standard *in vitro* assay conditions[48] and that the production of H_2O_2 by compound **1a** was only observed at concentrations an order of magnitude greater than its anti-HAT IC_{50} values, the results of this counter-screen demonstrate that compound **1a** is highly unlikely to inhibit HAT activity *in vitro* by the production of H_2O_2.

In parallel, we also assessed compound **1a** and chemical analogs for promiscuous enzymatic inhibition via chemical aggregation using another surrogate enzymatic assay based on the β-lactamase AmpC, a structurally unrelated enzyme whose activity is sensitive to compound aggregates[49]. Compound **1a** inhibited AmpC activity in the absence of detergent at low micromolar compound concentrations (10 µM), but this inhibition could be attenuated by the inclusion of detergent (0.01% Triton X-100) in the buffer (Figure 3A). We also tested several structural analogs that were active versus Rtt109-Vps75, compounds **1b**-**1f**, all of which showed inhibition patterns consistent with chemical aggregate formation (Figure 3A). We also tested several analogs that were inactive against Rtt109-Vps75, compounds **1q**-**1t** and **2a**-**2l**, and found that all of them did not significantly inhibit AmpC regardless of detergent status (data not shown). These data suggest that the 3-hydroxy group is an important component of AmpC inhibition, and that substitution with a 3-amino group likely prevents AmpC inhibition in the

absence of detergent. These data also show alkyl substitutions at the N1 position of chemotype **1** (i.e. compounds **1s** and **1t**) appear to prevent AmpC inhibition in the absence of detergent.

We note, however, that 24/24 (100%) of compound **1a** analogs were not flagged as potential chemical aggregators in an independent HTS using AmpC (Table S3). This observed difference between the two similar assays may result from the differences in assay buffers (Tris-HCl, pH 8.0 versus sodium phosphate, pH 7.0)[50] or perhaps more subtle inter-assay differences such as the timing of the assay steps and/or the nature of intra-well solution mixing, both of which can be crucial parameters in this assay.

Since compound **1a** was flagged by the aggregation counter-screen, we reasoned structural analogs might also act as chemical aggregators that could contribute to enzymatic inhibition in our assay systems, even though we had included detergent in our post-HTS assays to mitigate this possibility. We therefore investigated the effect of detergent on compound **1a** using the [^{3}H]-acetyl-CoA-based HAT assay, given the multiple sources of interference associated with the CPM-based assay. The omission of standard amounts of detergent (0.01% Triton X-100) and BSA (50 ng/μL) led to only a 2-fold more potent inhibition for compound **1a** versus Rtt109-Vps75, while higher concentrations of BSA (which may also mitigate the impact of chemical aggregators in biochemical assays[51]) with standard amounts of Triton X-100 also did not lead to large differences in the IC$_{50}$ values for compound **1a** (Figure 3B). Surprisingly, the inclusion of 10-fold higher concentrations of Triton X-100 led to an approximate 10-fold less potent inhibition of Rtt109-Vps75 activity by compound **1a** (Figure 3B).

Analysis of dose-response curves is another important consideration in HTS triage and can provide clues about the molecular mechanism(s) of target engagement such as cooperativity[52, 53]. A promising sign was that compound **1a** and all its analogs showed complete dose-responses (i.e. 0 to 100 percent inhibition for inhibitors) with two horizontal asymptotes in both the HTS assay and the orthogonal radiolabeled HAT assay. Steeper Hill slopes have also been associated with anomalous binding behavior such as aggregation[53-55]. Consistent with this phenomenon, compound **1a** had steeper Hill slopes in follow-up assays versus three different HATs (Table 4). In addition, examination of the slot blot results demonstrated a sharp cut-off in activity, consistent with a steeper Hill slope (Figure 1B).

While IC$_{50}$ values are by themselves an imperfect measure of selectivity[56], one of our initial

project criteria for progression to a 'hit-to-lead' (H2L) phase was a compound with at least 5-fold selectivity by IC_{50} value (but preferably greater) for Rtt109 versus other HATs. In parallel with the aggregation and redox-activity screens, we tested compounds **1a**-**1h** versus two other HATs using the [^{3}H]-acetyl-CoA-based HAT assay. We performed a selectivity counter-screen versus (1) p300, the human HAT that catalyzes H3K56ac and shares only gross tertiary structure with yeast Rtt109, and (2) Gcn5-Ada2-Ada3, a yeast HAT that is implicated in yeast RCNA and acetylates N-terminal H3 residues like H3K27 and H3K9[57, 58]. Compound **1a** inhibited both p300 and Gcn5-Ada2-Ada3 at low micromolar concentrations with no apparent selectivity towards Rtt109 (Table 2). Likewise, analogs **1b**-**1h** showed similar IC_{50} values and failed to show any convincing selectivity towards Rtt109-Vps75, demonstrating this compound series inhibited HATs non-selectively (Table 2).

Given the lack of any observed selectivity in the HAT counter-screen, we examined whether the observed anti-HAT activity for chemotype **1** was isolated to HATs, or represented even broader target promiscuity as the AmpC counter-screen suggested. Therefore, we tested the ability of compound **1a** and its chemical analogs **1b**-**1f** to modulate the conformation of the human La antigen, a protein completely unrelated from Rtt109. This assay, ALARM NMR, has been useful for identifying thiol-reactive compounds, as it contains several cysteines that when covalently modified, induce characteristic peak shifts and line broadening in nearby residues[59, 60]. Importantly, this method also utilizes an orthogonal detection method (2D ^{1}H-^{13}C HMQC) for assaying target engagement.

As expected for a 'frequent hitter' and not a simple, non-specific HAT inhibitor, compound **1a** perturbed the La protein conformation as evidenced by the signal decreases relative to DMSO controls, while the inactive analog compound **2i** did not perturb the La protein conformation when tested (Figure 4). The inclusion of DTT in the assay buffer did not attenuate the line broadening for compound **1a**, which is consistent with a non thiol-reactive mechanism of action in this experimental setting[59, 60]. Similar results were observed for compounds **1b**-**1f** (data not shown). This non-specific interaction is consistent with the lack of any apparent *in vitro* selectivity for chemotype **1** against a panel of HATs. Another possibility is H_2O_2 production by compound **1a** and analogs[59], but this explanation is less likely in our setting given the low levels of redox-activity observed only at relatively high compound concentrations in our HRP-PR counter-screen.

We have previously shown certain compound classes can interfere with assay readouts by multiple chemical mechanisms[16]. One of the main mechanisms causing non-specific bioactivity is electrophilic reaction of compounds with cysteines on proteins[59-61]. Other biologically important thiols include CoA and glutathione (GSH), the latter of which can be used as a surrogate assay for thiol-reactivity[62, 63]. Since compound **1a** was ALARM NMR positive with and without DTT in the reaction mixture, thiol-reactivity could not be clearly established. Therefore, we next sought to determine whether compound **1a** was thiol-reactive by alternate assays. We first incubated compound **1a** with either CoA or GSH in assay buffer and analyzed reaction aliquots by UPLC-MS. However, we were unable to observe any compound-thiol adducts (data not shown). Using the radiolabeled HAT assay versus Rtt109, we also assessed whether the addition of DTT in the assay buffer could attenuate the inhibitory activity of compound **1a** against Rtt109. There was no significant difference between the IC_{50} values for compound **1a** in either the presence or absence of DTT and BSA (p = 0.61, n = 3; Figure S8). By contrast, the positive control CPM probe showed a significant decrease in IC_{50} when DTT and BSA were omitted from the reaction mixture. These results suggest that compound **1a** is not a thiol-reactive chemotype in our assay conditions. However, we note a related compound has shown evidence of time-dependent and irreversible inhibition using kinetic studies, although isolation of a covalent adduct was not performed[64]. These data leave open the possibility this series may be reactive under certain conditions.

Certain compounds can react non-enzymatically with protein lysine side chains[65]. As such, we explored this possibility for compound **1a**. However, we did not observe any detectable amine-compound **1a** adducts by UPLC-MS when compound **1a** was incubated with either $N\alpha$-acetyl-L-lysine methyl ester or N-butylamine. Furthermore, the addition of excess N-butylamine did not attenuate the perturbation of the La protein conformation by compound **1a** when tested by ALARM NMR (data not shown).

We also performed additional experiments to characterize the basic mechanism of Rtt109-Vps75 inhibition by compound **1a** using the [^{3}H]-acetyl-CoA HAT assay. One hallmark of reactive compounds (and also slow, tight binders) is time-dependent inhibition. Therefore, we assessed whether the enzymatic inhibition was time-dependent by varying the amount of time compound **1a** was allowed to pre-incubate with Rtt109-Vps75 prior to initiating the HAT reaction[56]. We found no significant difference in the IC_{50} values of compound **1a** between 5, 15 and 30 min pre-incubation time (Figure 5A). This result demonstrates Rtt109-Vps75 inhibition by

compound **1a** is not enhanced by prolonged pre-incubation times.

Another important distinction is whether a compound exerts its mechanism of action by reversible or irreversible means. Making this distinction is important for interpreting experiments and performing structure-based drug design. Compounds that covalently react with their target generally show irreversible target modification, whereas compounds not covalently reacting with their target show reversible modification. Therefore, we also examined for evidence of reversible/irreversible inhibition of Rtt109-Vps75 *in vitro* HAT activity by compound **1a**. We performed a modified jump dilution experiment[56, 66, 67]. We observed that Rtt109-Vps75 activity was restored upon dilution, which is consistent with a reversible type of enzymatic inhibition (Figure 5B). These data suggest compound **1a**, and most likely its other structural analogs, inhibit Rtt109-Vps75 activity *in vitro* by a non-specific, reversible process. This is consistent with our other data, including the positive AmpC counter-screen, the non-specific compound-protein interactions observed by ALARM NMR, and the lack of observed thiol reactivity in multiple assays.

In a recent report, we advocated for examinations of bioactive compounds in the context of previous studies[1]. As such, we examined the scientific literature for reports of bioactivity for compound **1a** and close structural analogs. The number of publications and patents invoking screening 'hits' similar to compound **1a** has grown exponentially in the past decade (Figure 7). We found that related compounds have been published as active versus targets and/or systems unrelated to our screening target, Rtt109. These include both real and virtual screens, with several of the reported targets being protein-protein interactions[37-40, 42, 64, 68-84].

We also examined repositories of HTS data for indication of bioassay and target promiscuity for compound **1a** and structural analogs. There were no reports of compound **1a** in any PubChem bioassay. However, analysis of PubChem records showed several analogs of compound **1a** were active in a variety of bioassays (Figure 6)[85-87]. Furthermore, we also utilized an open-source tool for predicting bioactivity promiscuity, BadApple (Bioactivity data associative promiscuity pattern learning engine). We found that compound **1a** and several related substructures were predicted to be highly promiscuous (Table 4).

Finally, we examined the HTS records from AZ for evidence of assay promiscuity in an industrial setting[88]. These records include over 1M compounds tested in up to hundreds of HTS

campaigns, with wide ranges of assay types. Mining of this screening deck demonstrated that the compound class of 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-ones showed an increased incidence of promiscuous bioactivity profiles, indicating their frequent-hitting behavior is not exclusive to academic screening campaigns or screening libraries (Table 6). Looking more closely at the substitution patterns of compounds in the corporate data set and their alignment with frequent-hitter behavior in historical assay data, we found that aromatic substitution at all three positions poses the highest risk of such behavior. These data suggest that aliphatic substitutions at any of the three positions lessen the risk of anomalous behavior, as does substitution of the hydroxyl. These observations align closely with our Rtt109 structure-activity data (Figure 8), suggesting that the empirical observations for the smaller set of compounds in this paper apply to the larger class as well.

It remains unclear what properties modulate the indiscriminate binding behavior. Properties of the class, in particular of the poly-aromatic examples, are predominantly non-lead-like, with most compounds in this report exhibiting high lipophilicity. Modification of the structure with aliphatic groups or *O*-substitution will modulate properties including fluorescence, lipophilicity, and acidity of the hydroxyl group, but no clear relationship could be found between promiscuous behavior and predicted pK_a, lipophilicity ($\log P$, $\log D$), or polar surface area (Nissink, unpublished observations).

This case report describes the follow-up of a 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one active, compound **1a**, from a cell-free HTS to identity small-molecule inhibitors of Rtt109-catalyzed histone acetylation. This compound showed publication-quality low micromolar IC_{50} values in a set of three orthogonal *in vitro* assays. However, during subsequent experiments, it became clear that this compound and its structural analogs have several problematic features. Due to the conjugated nature of its poly-aromatic structure, compound **1a** is able to interfere with the fluorescence-based readout of our CPM-based HTS, and conceivably other optical-based assays as well. Compound **1a** showed evidence of redox-activity, but only at relatively high compound concentrations. Compound **1a** does not appear to be thiol reactive in our system, as shown by the CoA-CPM counter-screen, the lack of detectable adducts by UPLC-MS, and the additional mechanistic experiments for time-dependence and reversibility. While compound **1a** was not flagged as a PAINS, in retrospect it was flagged by the 'Rapid Elimination of Swill' (REOS) filters our group periodically employs (Table 1)[89].

Compound **1a** was flagged for its aggregation tendencies in a β-lactamase counter-screen for chemical aggregators. A teaching point is the importance of detergent concentration, as the IC_{50} values of compound **1a** varied by at least an order of magnitude depending on the amount of Triton X-100 present (Figure 3B). While we and others have used 0.01% Triton X-100 as our default detergent concentration, our data suggest even higher concentrations of detergent can further attenuate the potency of certain compounds like compound **1a**. Therefore, if non-specific inhibition or chemical aggregation is suspected, it may be worth testing the effect of additional detergent on the bioactivity of a test compound. One possible explanation for this observation is that aggregate formation may not be completely attenuated at this standard detergent concentration for chemotype **1**, or that this additional detergent disrupts some non-aggregate-related compound-protein interactions. There are also examples of screening compounds forming bioactive aggregates in the presence of detergent. For instance, the compound titled '1541' assembles into non-globular nanofibrils capable of activating procaspase-3[90]. These and other explanations for the assay behavior of compound **1a** and analogs are possible, and are the subject of ongoing studies.

Only after performing extensive studies did our project team abandon further work on this scaffold. A summary of the pros and cons of this scaffold is provided (Table 7). Notable arguments supporting the utility of compound **1a** and its analogs include its low micromolar *in vitro* IC_{50} values, its activity in several orthogonal assays, its synthetic accessibility and its lack of identifiable redox-activity or thiol-reactivity in our hands. Counter-arguments include its poor calculated physicochemical properties, its lack of selectivity, its tendency to form chemical aggregates in our assay buffer, its perturbation of the unrelated ALARM NMR protein and the bioassay promiscuity of related compounds. One reason this compound was pursued, despite its lack of lead-like characteristics, was because so few compounds were identified in our HTS. Our experience with this compound argues against pursuing compounds with few lead-like characteristics and known liabilities. Of course, we note ours is only a case report and care should be taken to draw broader conclusions, yet it is clear that progression of such compounds with adverse properties incurs a higher risk of failure.

Our experiments and cheminformatics analyses provide several insights into the SAR/SIR of chemotype **1** (Figure 8). For example, the 3-hydroxy position is critical for *in vitro* activity in multiple assays, including anti-HAT activity, ALARM NMR and an AmpC chemical aggregation counter-screen. Strikingly, all 3-amino analogs that we tested were inactive in these assays.

Next, no dramatic differences were observed with regard to anti-HAT activity of the furan and thiophene analogs connected to the C4-position, suggesting these two groups, and possibly other aromatic groups, are interchangeable at this position. The types and positions on the phenyl ring connected at the C5-position were also interchangeable. Similarly, the heteroaromatic substituents anchored at the N1-position appeared interchangeable. Direct *N*-alkyl substituents at this position showed reduced potency (e.g. compounds **1s**-**1y**), with the exception of compound **1i**.

Though N1-alkyl substitution was predicted to have an effect on acidity of the 3-hydroxy group (median predicted pK_a 3.8 for poly-aromatic compounds; 5.1 for N1-alkyl substituted compounds in the sets E and C from Table 6), acidity of the group was not predictive of anomalous behavior in assays, nor could a relationship be identified with other properties like solubility or lipophilicity. Subsequent measurement of aliphatic-substituted compounds **1w** and **1x**, and the aromatically-substituted compound **1z** suggested that their acidity was even more similar than predicted, with experimental pK_a values of 3.3 and 3.6 for **1w** and **1x**, and 3.0 for **1z** (Nissink, unpublished results).

Work on this compound could have been halted earlier with a better-designed post-HTS system and project management. In hindsight, a more thorough literature and database search prior to performing more labor-intensive *in vitro* assays could have prevented further work on this series. We should also have more strongly weighted the contribution of calculated physicochemical properties in our triage. We also encountered some opposition from members of the project team who had more experience with the biological sciences and were less familiar with medicinal chemistry and HTS. Last but not least, there was also some reluctance to cease work on a project many years in the making; an example of a sunk cost conundrum. In several projects we have encountered hesitancy from non-chemists to make key project decisions based on structural arguments alone, and it was only after generating experimental data that these members of the project could be convinced otherwise. This is in contrast with best practice in corporate screening environments where historical data are available and used extensively to triage HTS outcome upfront and provide annotation of hits[48, 88, 89, 91]. Better communication at the start of the project between members of the chemistry and biology arms of the project would likely have led to earlier agreement on a 'no-go' decision.

Our experience with chemotype **1** leads us to make several recommendations. It may be useful

to include this substructure in future HTS cheminformatics filters, such as those used for flagging PAINS and other liable substructures. Annotation through use of such filters as a first step could highlight issues for scaffolds, or at least alert researchers to the chemotype so that wasteful follow-up can be prevented. Use of such information as an annotation together with a wider 'body of evidence' for or against a compound is preferred, as application of such alerts for the purpose of elimination could lead to 'throwing the baby out with the bathwater'. Some 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-ones may actually represent true hits and *bona fide* lead compounds, and data mining of the wider class in a corporate collection suggests that incidence of anomalous behavior is elevated, but not across all of the class. There appears to be one example of this scaffold forming a crystal structure with its proposed target[68]. Based on the body of work pertaining to chemotype **1**, we believe investigators should carefully consider the wisdom of working with this particular compound class in an early drug discovery setting, unless there is confirmation of useful target engagement (plus selectivity) and ways can be envisaged to move away from the 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one chemotype.

The observation that compounds with chemotype **1** are active against a variety of biological targets in multiple assay systems should cause concern and instigate closer inspection if these compounds show up as actives. Of course, the counter-argument is that specificity may be attainable with sufficient medicinal chemistry optimization. This is certainly feasible, but it may not be without expending considerable resources (including financial, animal, and personnel costs) or a high likelihood of failure. One may also invoke this scaffold as 'privileged' but the overwhelming majority of reports using this scaffold do not provide evidence of specific and therapeutically-useful target engagement such as protein-ligand crystal structures. Because of the concerns raised in this report, we believe those choosing to work with this scaffold should carefully evaluate such compounds for selectivity, stability, redox-activity, aggregation, and thiol-reactivity in their specific systems. Adverse observations increase the level of risk that a compound will fail, and a researcher should consider the cost of taking on that risk.

Our experience with this scaffold supports the use of ALARM NMR and other triage tools in the post-HTS setting. Compound **1a** and some of its analogs were ALARM NMR-positive independent of DTT concentration. This finding supports another (simultaneous) benefit of performing ALARM NMR in HTS triage, as it can help identify compounds that likely non-specifically modulate the conformation of the La antigen. In our case, compounds that can

modulate a protein like the La antigen, a target that is completely unrelated to Rtt109, should raise suspicions about their utility as potential therapeutics or chemical probes. Employing such a strategy to gauge off-target modulation is likely broadly applicable to HTS triage.

In summary, the 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one chemotype **1** discussed in this report has been shown to have poor properties including indiscriminate binding behavior, and our experience highlights the importance of considering historical knowledge, physicochemical properties and a suitable triaging cascade of post-HTS assays to identify truly promising leads. We also recommend the more frequent publication of 'no-go' compound decisions, especially where the compounds are not recognized by known substructure filters[92].
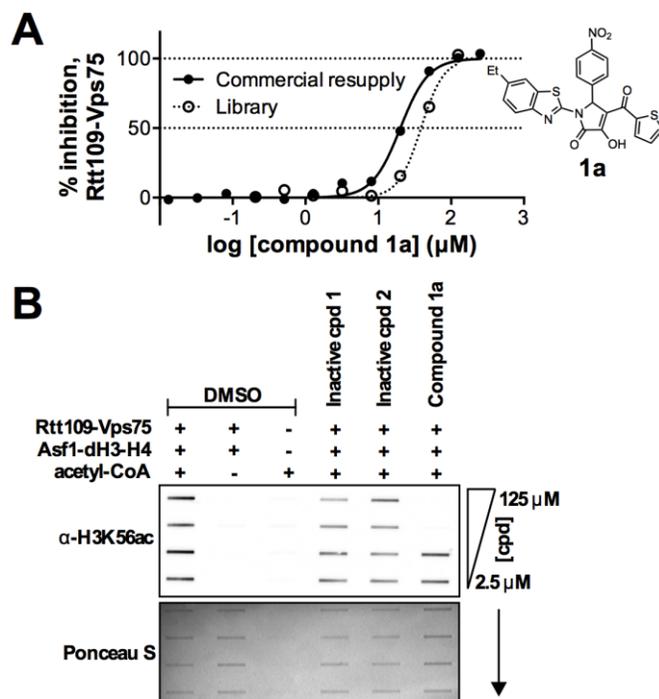
Figure 1 | Inhibitory activity of a 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one screening compound against Rtt109-catalyzed histone acetylation *in vitro*. (A) Dose-response curves for compound **1a** in library cherry-pick (mean for two replicates) and commercial resupply compound samples (mean ± SD for three replicates). Shown is the chemical structure of compound **1a**. (B) HAT inhibition of compound **1a** in the orthogonal slot blot assay using the commercial resupply sample.
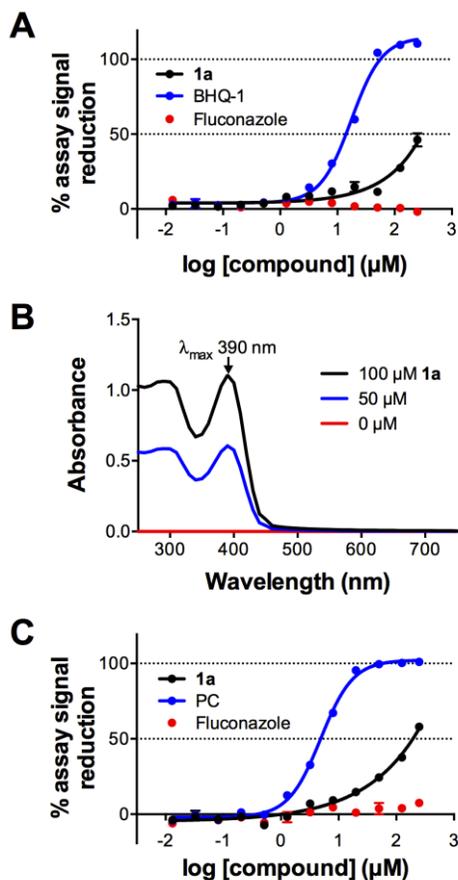
Figure 2 | CPM-based assay interference counter-screens for compound 1a. (A) Compound **1a** weakly interferes with the HTS assay in a fluorescence quenching counter-screen. BHQ-1 and fluconazole are positive and negative quenching controls, respectively. (B) Absorption profile for compound **1a** in assay buffer. (C) Compound **1a** does not interfere with the HTS assay readout via CoA-trapping (n.b. compare to panel 2A). PC (positive control; 4-((9*H*-purin-6-yl)thio)-7-bromo-5-nitrobenzo[*c*][1,2,5]thiadiazole) and fluconazole are positive and negative thiol-trapping controls, respectively[16]. Data are mean ± SD for three replicates.
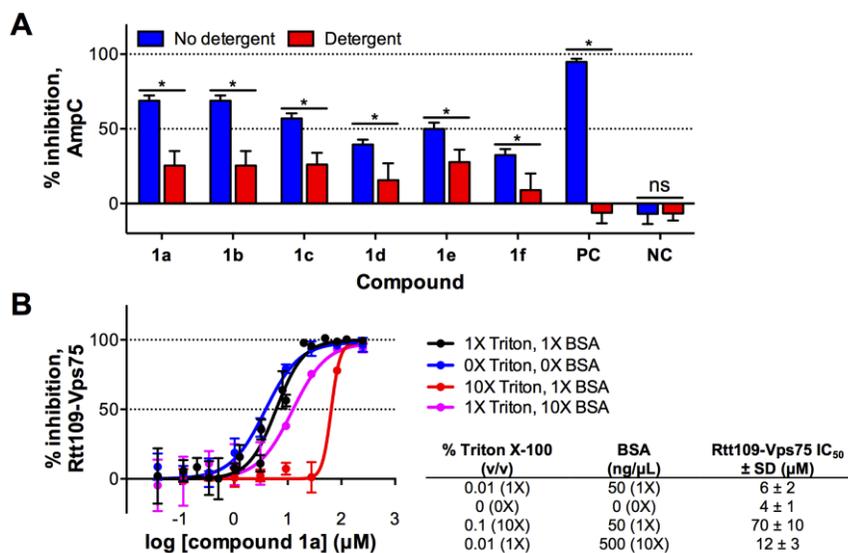
Figure 3 | Assessment of compound 1a and related analogs for chemical aggregation. (A) β-lactamase counter-screen for chemical aggregators. Compounds were tested for inhibition of the β-lactamase AmpC ± non-ionic detergent (0.01% Triton X-100) in Tris buffer pH 8.0. Rottlerin and lidocaine were included as positive control (PC) and negative control (NC) compounds, respectively. Data are mean ± SD for three replicates. (B) Inhibition of compound **1a** against Rtt109-Vps75 as a function of Triton X-100 and BSA concentration using the [^{3}H]-acetyl-CoA-based HAT assay in the presence of 5 mM DTT. Data are mean ± SD for three replicates. Data shown for compound **1a** under standard conditions (1X detergent and BSA) is taken from Table 2.
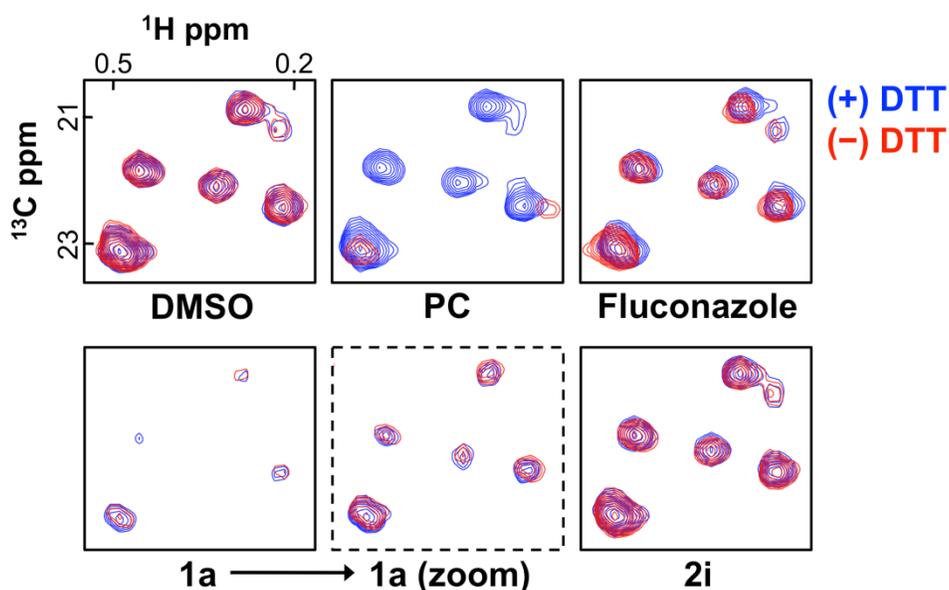
Figure 4 | Effect of the 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one active (compound 1a) on La protein conformation using ALARM NMR. Shown are 2D ^{1}H-^{13}C HMQC spectra of selected ^{13}C-labelled methyl groups for compound **1a** as tested by ALARM NMR. Compounds were incubated with the La protein in either the presence (blue) or absence (red) of 20 mM DTT. 2-Chloro-1,4-naphthoquinone (PC) and fluconazole are shown as positive and negative compound controls, respectively. A scaled-up view of compound **1a** is shown ('1a zoom'). Note the spectra with and without DTT are nearly identical. Compound **2i**, an inactive analog, was also included as a negative control compound. Signals were normalized to DMSO controls. Shown are representative results from one of three independent experiments.
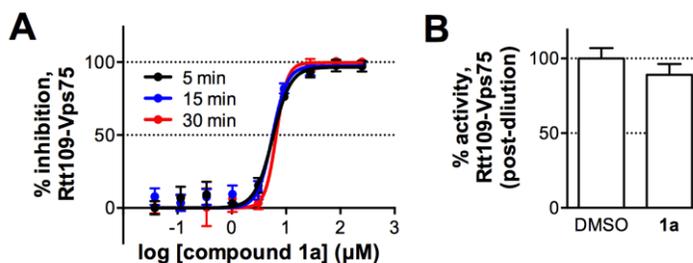
Figure 5 | Additional mechanistic experiments for compound 1a. (A) Dependence of incubation time on activity of compound **1a** against Rtt109-Vps75 activity. Compound **1a** was allowed to pre-incubate with Rtt109-Vps75 for 5, 15 or 30 min prior to initiating the HAT reaction with [^{3}H]-acetyl-CoA. Data are mean ± SD for three replicates. (B) Dilution-based experiment to examine reversibility of Rtt109-Vps75 inhibition by compound **1a**. Data are mean ± SD for nine replicates.

CID 3146565
37/638 bioactive
33 protein targets

CID 2950564
36/629 bioactive
24 protein targets

CID 2910886
26/690 bioactive
23 protein targets

CID 4905842
36/507 bioactive
28 protein targets

CID 3152781
25/402 bioactive
18 protein targets

CID 2939274
32/548 bioactive
28 protein targets

CID 2918794
19/453 bioactive
10 protein targets
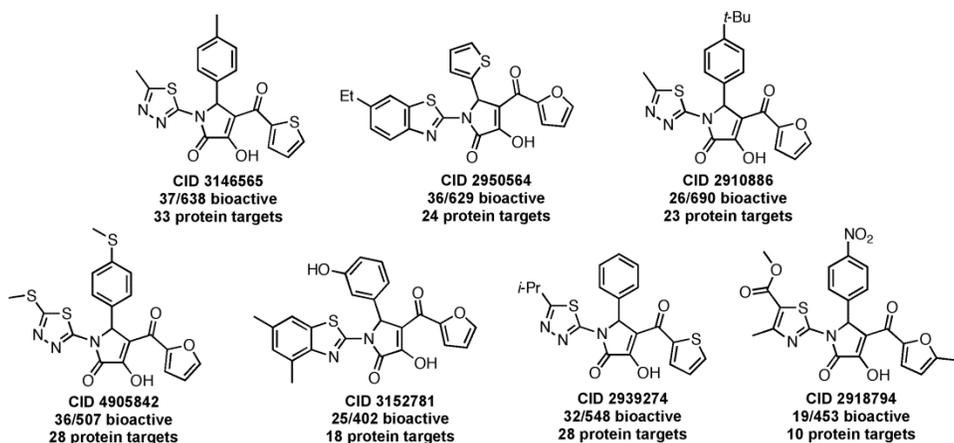
Figure 6 | Selected examples of bioactive 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-ones in PubChem. The compounds shown are structurally similar to compound **1a** and are active in multiple PubChem bioassays versus a variety of protein targets. There is no PubChem bioactivity data reported for compound **1a**. Data accessed 2 December 2014.
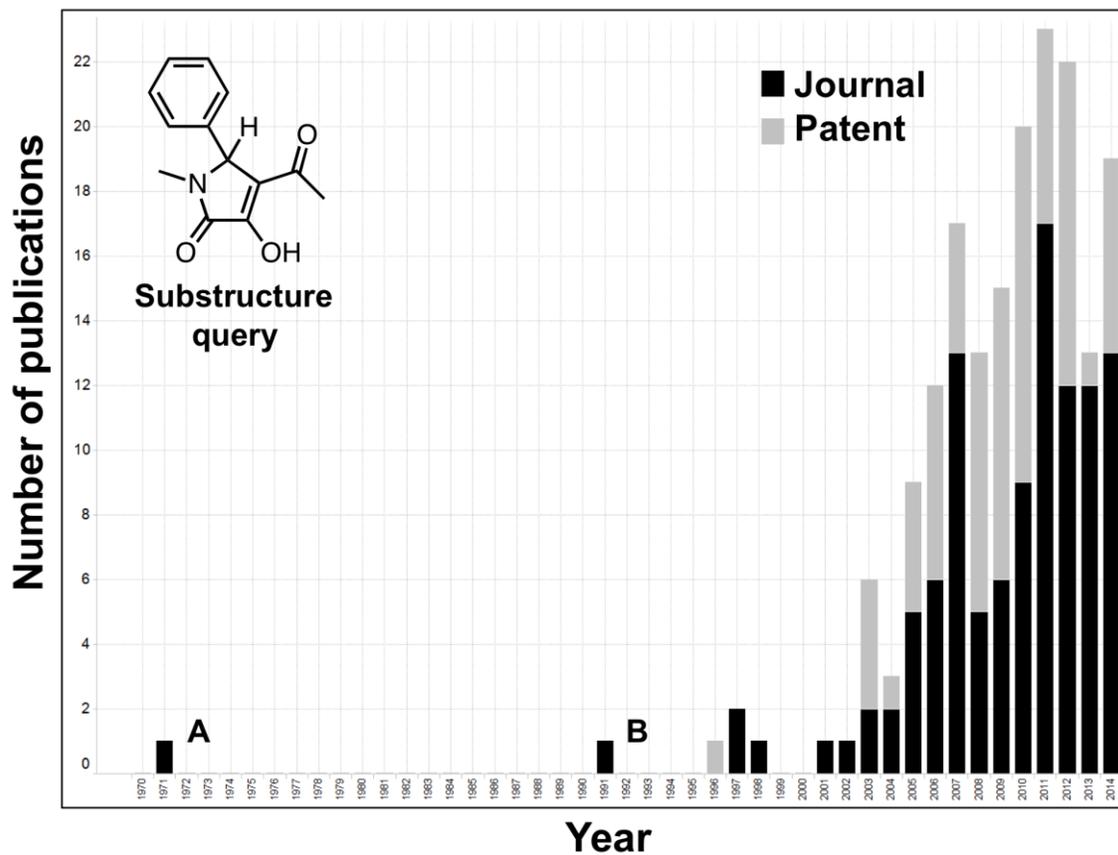
Figure 7 | Publication trend of compound 1a substructure. Shown is the annual number of publications in which compounds containing a compound **1a** substructure were reported as being used in 'biological studies'. The data was obtained from SciFinder using default substructure searching methods. **A** and **B** refer to the first reported synthesis and subsequent bioactivity measurements for this substructure, respectively[93, 94]. Data accessed 29 December 2014.
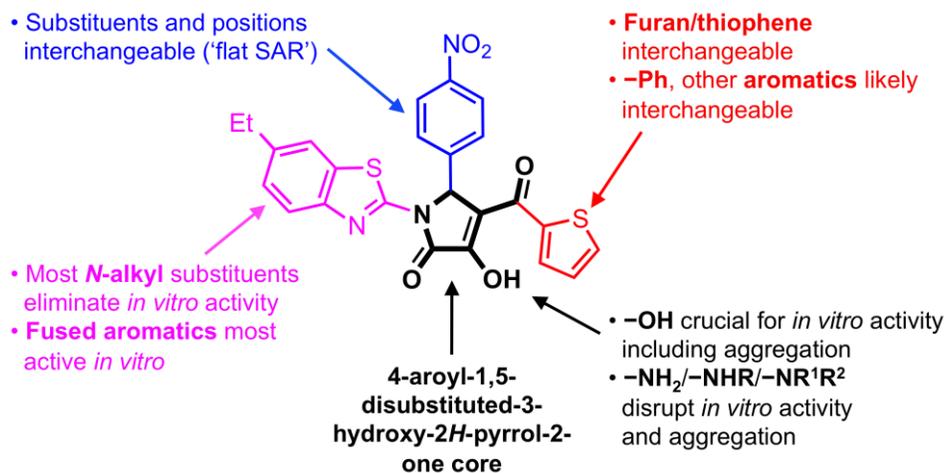
Figure 8 | Summary of *in vitro* structure-activity/structure-interference relationships for chemotype 1. Shown is compound **1a**, the parent compound originally identified in a cell-free fluorometric HTS for inhibitors of Rtt109-catalyzed histone acetylation.

| Descriptor	Value	Remarks
Molecular weight	491.5	Pass[95]
cLog*P*	5.4	Fail[95]
H-bond donors	1	Pass[95]
H-bond acceptors	6	Pass[95]
Rotatable bonds	6	Pass[96]
Stereocenters	1	–
tPSA	172	High | Fail[96]
Aromatic rings	4	–
Fsp^3	0.13	Calculated in Pipeline Pilot 8.5
PAINS	No	Calculated in Canvas 2.0
REOS	Yes	Canvas 2.0 (-NO$_2$)[89]
Tier	7	Least favorable[92]

Table 1 | Selected descriptors for compound 1a.
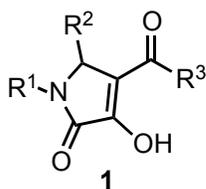
ID	R¹	R²	R³	Rtt109-Vps75 IC_{50} (µM)	p300 IC_{50} (µM)	Gcn5-Ada2-Ada3 IC_{50} (µM)
1a	6-Et-benzothiazol-2-yl	4-NO₂-phenyl	thiophen-2-yl	6 ± 2[a]	9 ± 5[b]	7 ± 2[b]
1b	6-Et-benzothiazol-2-yl	3-NO₂-phenyl	thiophen-2-yl	8.5 (6 to 12)	14 (5 to 40)	22 (10 to 47)
1c	6-Et-benzothiazol-2-yl	3-OH-phenyl	thiophen-2-yl	7.5 (6.6 to 8.6)	12 (6 to 19)	12 (9 to 16)
1d	6-Et-benzothiazol-2-yl	4-methylphenyl	thiophen-2-yl	9.1 (5.4 to 15)	16 (10 to 26)	10 (7 to 15)
1e	6-Et-benzothiazol-2-yl	4-Cl-phenyl	thiophen-2-yl	9.5 (4.7 to 19)	12 (7 to 20)	14 (7 to 27)
1f	6-Et-benzothiazol-2-yl	4-NO₂-phenyl	furan-2-yl	22 (16 to 29)	25 (16 to 41)	4.2 (2 to 10)
1g	6-Et-benzothiazol-2-yl	4-i-Pr-phenyl	thiophen-2-yl	34 (25 to 45)	29 (20 to 40)	7 (3 to 15)
1h	6-Et-benzothiazol-2-yl	4-F-phenyl	thiophen-2-yl	29 (4 to 100)	14 (3 to 80)	21 (10 to 42)
1i	furan-2-ylmethyl	4-OBn-phenyl	thiophen-2-yl	18 (13 to 24)	–	–
1j	6-methylpyridin-2-yl	4-OPh-phenyl	thiophen-2-yl	20 (12 to 33)	–	–
1k	6-F-benzothiazol-2-yl	4-NO₂-phenyl	thiophen-2-yl	11 (8.9 to 15)	–	–
1l	5-methylbenzothiazol-2-yl	4-NO₂-phenyl	thiophen-2-yl	5.0 (4.7 to 5.4)	–	–
1m	6-Et-benzothiazol-2-yl	4-i-Pr-phenyl	thiophen-2-yl	4.2 (3.0 to 5.9)	–	–
1n	3-methylisoxazol-5-yl	3-OPh-phenyl	thiophen-2-yl	43 (19 to 94)	–	–
1o	4,7-dimethylbenzothiazol-2-yl	3,4-diCl-phenyl	benzofuran-2-yl	4.6 (2.3 to 9.1)	–	–
1p	4,7-dimethylbenzothiazol-2-yl	4-OH-phenyl	benzofuran-2-yl	4.0 (2.4 to 8.5)	–	–
1q	1,3,4-thiadiazol-2-yl	2-F-phenyl	thiophen-2-yl	Inactive	–	–
1r	1,3,4-thiadiazol-2-yl	2-Cl-phenyl	thiophen-2-yl	Inactive	–	–
1s	pyridin-3-ylmethyl	3-OPh-phenyl	thiophen-2-yl	Inactive	–	–
1t	pyridin-3-ylmethyl	4-i-Pr-phenyl	thiophen-2-yl	Inactive	–	–

1u				Inactive	–	–
1v				Inactive	–	–
1w				Inactive	–	–
1x				Inactive	–	–
1y				Inactive	–	–
1z				80 (54 to 120)	–	–
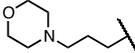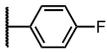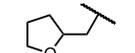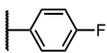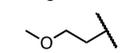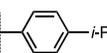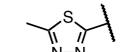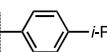
^a IC₅₀ values shown are means ± SD of six independent experiments.

^{a} IC$_{50}$ values shown are means ± SD of six independent experiments.
^{b} IC$_{50}$ values shown are means ± SD of three independent experiments.

Table 2 | SAR-by-commerce: inhibitory activity of compound 1a and structural analogs (chemotype 1) using an *in vitro* [^{3}H]-acetyl-CoA HAT assay. Except for the data for compound **1a**, the data are IC$_{50}$ values with 95% confidence intervals in parentheses. Compounds **1v**-**1z** were obtained from the AZ corporate library.
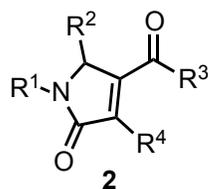
ID	R^1	R^2	R^3	R^4	Rtt109-Vps75 IC_{50} (μM)	p300 IC_{50} (μM)	Gcn5-Ada2-Ada3 IC_{50} (μM)
2a	-Ph	-Ph	(thiophene)	(NH-m-tolyl)	Inactive	Inactive	Inactive
2b	(pyridine)	-Ph	-Ph	(HN-phenyl-OEt)	Inactive	Inactive	Inactive
2c	-Ph	-Ph	-Ph	(morpholine)	Inactive	Inactive	Inactive
2d	-Ph	(fluorophenyl)	-Ph	-NHPh	Inactive	Inactive	Inactive
2e	-Ph	-Ph	(thiophene)	(NH-o-tolyl)	Inactive	Inactive	Inactive
2f	(pyridine)	-Ph	-Ph	(HN-p-tolyl)	Inactive	Inactive	Inactive
2g	-Ph	(nitrophenyl)	-Ph	-NHPh	Inactive	Inactive	Inactive
2h	-Ph	-Ph	(furan)	(HN-p-tolyl)	Inactive	Inactive	Inactive
2i	-Ph	-Ph	(thiophene)	-NHPh	Inactive	Inactive	Inactive
2j	-H	-Ph	-Ph	(HN-p-tolyl)	Inactive	Inactive	Inactive
2k	-Ph	(bromophenyl)	-Ph	-NHPh	Inactive	Inactive	Inactive
2l	-Ph	-Ph	(p-tolyl)	(morpholine)	Inactive	–	–

Table 3 | Inactive structural analogs of compound 1a. Compounds **2** were purchased from commercial vendors and then tested for inhibition of Rtt109-Vps75 activity using the [^{3}H]-acetyl-CoA-based HAT assay. Compounds **2a**-**2k** were purified by standard RP-HPLC procedures and were also tested for inhibition of p300 and Gcn5-Ada2-Ada3 activity *in vitro* using same HAT assay. All the compounds were 'Inactive' (IC_{50} values > 125 μM) versus these HATs. All purified compounds had acceptable purities (> 95% by ELS and UV 254 nm) and parent ions consistent with the vendor-supplied structures (LRMS-ESI) when analyzed by UPLC-MS.

ID	Hill slope, Rtt109-Vps75	Hill slope, p300	Hill slope, Gcn5-Ada2-Ada3
1a	2.0 ± 0.6^{a}	1.5 ± 0.3^{b}	1.8 ± 0.4^{b}

[a] IC_{50} values shown are means ± SD of six independent experiments.
[b] IC_{50} values shown are means ± SD of three independent experiments.

Table 4 | Hill slopes of compound 1a using an *in vitro* [^{3}H]-acetyl-CoA HAT assay. The calculations are based on the data from Table 2 for compound **1a**).

Structural query	Remarks	pScore	BadApple advisory
	Parent compound	1463	**High**
	–	689	**High**
	–	944	**High**
	–	506	**High**
	–	678	**High**
	–	176	Moderate
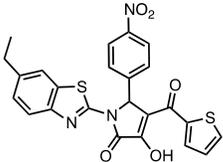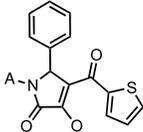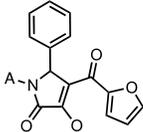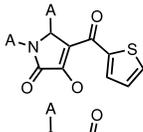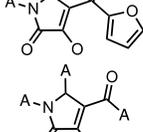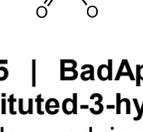
Table 5 | BadApple bioactivity promiscuity analysis of compound 1a, 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-ones and related substructures. A, any atom. pScore: 0-100, low advisory for promiscuity; 101-300, moderate advisory; > 300, high advisory. Accessed 1 June 2014 (http://pasilla.health.unm.edu/tomcat/badapple/badapple).

Query ID	Substructure query[a]	Description	N[b]	N_{data}[b]	A[b] (pBSF > 2)	P(a ≥ A|p=0.06)[c]
A		Unsubstituted hydroxyl and monosubstituted at R^2 position	537	483	22	0.93
B		Compounds substituted at hydroxyl position; no or disubstitution at R^2 position (no monosubstituted examples found)	28	27	0	1.00
C		Aliphatic non-ring linker at position R^1	413	371	1	1.00
D		Aromatic group at position R^1	104	93	21	**< 0.001**
E		Aromatic groups at position R^1, R^2, R^3	85	75	20	**< 0.001**
F		Aromatic groups at position R^1, R^3; aliphatic at position R^2 Note – both compounds are disubstituted at R^2 (no mono-substituted examples found)	2	2	0	1.00
G		Aromatic groups at position R^1, R^2; aliphatic at position R^3	19	18	1	0.67
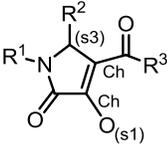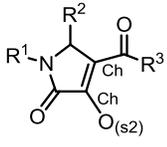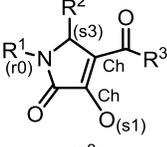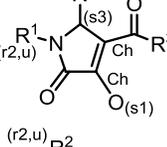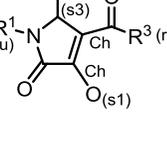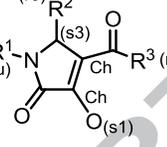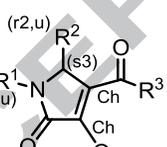
[a] Structure annotations: s*n*, number of nonhydrogen substituents (e.g. "s2"); r*n*, number of connected ring bonds (e.g. "r2"); u, aromatic bonds.
[b] N is the number of compounds having the substructure, and N_{data} designates the subset of compounds for which a pBSF score had been derived. This is dependent on the availability of HTS screening data. A is the count of compounds with a pBSF score ≥ 2.
[c] Cumulative binomial probability of seeing A or more compounds with a pBSF score in a set of N_{data} compounds when the expected incidence is 0.06. A very low chance (**bolded**) suggests that the observed count is unexpected, i.e. the set of compounds shows an unexpectedly high incidence of anomalous binders. Expected incidence of anomalous binders is 6% (averaged over all compounds with data in the AZ collection).

Table 6 | 3-hydroxy-pyrrolidin-2-ones bioassay promiscuity analysis in an industrial HTS setting. Structure-property-relationships-by-data suggest that aromatic substitutions at all three positions R^1, R^2 and R^3 lead to an increased incidence of anomalous behavior. Data patterns suggest that aliphatic substitution at any of the three positions, or substitution of the hydroxyl alleviates this behavior.

'Go'	'No-go'
Active in HTS	Lack of lead-like qualities
Dose-response confirmed in HTS	Lack of interpretable SAR ('flat SAR')
Activity confirmed in slot blot orthogonal asasy	No apparent selectivity between relevant HATs
Activity re-confirmed by commercial re-supply	Positive in aggregation counter-screen (β-lactamase)
Activity confirmed by radiolabeled HAT assay (second orthogonal assay)	Poor aqueous solubility[a]
Low micromolar activity in multiple *in vitro* assays	Fluorescent interference
Minimal redox-activity in counter-screen (HRP-PR)	ALARM NMR positive ± DTT
Synthetically and commercially accessible analogs	Promiscuous bioactive analogs in PubChem
Stable in assay conditions	Publications with analogs active versus unrelated targets
No thiol adducts detected (UPLC-MS)	Frequent hitters in corporate HTS setting

[a] <50 ug/mL solubility in assay buffer (precipitation observed at higher concentrations).

Table 7 | Factors behind a 'go/no-go' decision. The attributes of compound **1a** and its 4-aroyl-1,5-disubstituted-3-hydroxy-2*H*-pyrrol-2-one chemotype **1** were divided into factors that were generally favorable ('go') or unfavorable ('no-go') for project progression.

Supplementary data

Supplementary data associated with this article can be found in the online version. Files containing these data include: (1) Supporting Information, which contains materials and methods, characterization data for compound **1a**, Figures S1-S8, Tables S1-S3, and author contributions; (2) a CSV file containing SMILES, InChI, InChIKey and activity data for compounds **1a**-**1z** and **2a**-**2l**; and (3) a corresponding MOL file.

References and Notes

1. Dahlin, J. L.; Walters, M. A. *Future Med. Chem.* **2014**, *6*, 1265.

2. Wipf, P.; Arnold, D.; Carter, K.; Dong, S.; Johnston, P. A.; Sharlow, E.; Lazo, J. S.; Huryn, D. *Curr. Top. Med. Chem.* **2009**, *9*, 1194.

3. Huryn, D. M.; Smith, A. B. *Curr. Top. Med. Chem.* **2009**, *9*, 1206.

4. Devine, S.; Mulcair, M.; Debono, C.; Leung, E.; Nissink, J.; Lim, S.; Chandrashekaran, I.; Vazirani, M.; Mohanty, B.; Simpson, J.; Baell, J.; Scammells, P.; Norton, R.; Scanlon, M. *J. Med. Chem.* **2015**, *58*, 1205.

5. Han, J.; Zhou, H.; Horazdovsky, B.; Zhang, K.; Xu, R.; Zhang, Z. *Science* **2007**, *315*, 653.

6. Dahlin, J. L.; Chen, X.; Walters, M. A.; Zhang, Z. *Crit. Rev. Biochem. Mol. Biol.* **2014**, *50*, 31.

7. Dahlin, J. L.; Kottom, T. J.; Han, J.; Zhou, H.; Walters, M. A.; Zhang, Z.; Limper, A. H. *Antimicrob. Agents Chemother.* **2014**, *58*, 3650.

8. Wurtele, H.; Tsao, S.; Lépine, G.; Mullick, A.; Tremblay, J.; Drogaris, P.; Lee, E.-H.; Thibault, P.; Verreault, A.; Raymond, M. *Nat. Med.* **2010**, *16*, 774.

9. Lopes da Rosa, J.; Bajaj, V.; Spoonamore, J.; Kaufman, P. D. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2853.

10. Lopes da Rosa, J.; Boyartchuk, V. L.; Zhu, L. J.; Kaufman, P. D. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1594.

11. Dahlin, J. L.; Sinville, R.; Solberg, J.; Zhou, H.; Francis, S.; Strasser, J.; John, K.; Hook, D. J.; Walters, M. A.; Zhang, Z. *PLOS ONE* **2013**, *8*, e78877.

12. Baell, J. B. *Future Med. Chem.* **2010**, *2*, 1529.

13. Baell, J. B.; Ferrins, L.; Falk, H.; Nikolakopoulos, G. *Aust. J. Chem.* **2013**, *66*, 1483.

14. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53*, 2719.

15. Baell, J.; Walters, M. A. *Nature* **2014**, *513*, 481.

16. Dahlin, J. L.; Nissink, J. W. M.; Strasser, J. M.; Francis, S.; Zhou, H.; Zhang, Z.; Walters, M. A. *J. Med. Chem.* **2015**, *58*, 2091.

17. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discov. Today* **2003**, *8*, 876.

18. Rishton, G. M. *Drug Discov. Today* **2003**, *8*, 86.

19. Malo, N.; Hanley, J. A.; Cerquozzi, S.; Pelletier, J.; Nadon, R. *Nat. Biotech.* **2006**, *24*, 167.

20. Zhang, J. H.; Chung, T. D.; Oldenburg, K. R. *J. Comb. Chem.* **2000**, *2*, 258.

21. Gubler, H.; Schopfer, U.; Jacoby, E. *J. Biomol. Screen.* **2013**, *18*, 1.

22. Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N.-E.; Scott, N.; Petersen, A.; Gabriel, S.; Vu, B.; George, K. M.; Narayanan, A.; Li, S. H.; Qian, H.; Beatini, N.; Niu, L.; Gan, Q.-F. *ACS Med. Chem. Lett.* **2013**, *4*, 197.

23. Tarzia, G.; Antonietti, F.; Duranti, A.; Tontini, A.; Mor, M.; Rivara, S.; Traldi, P.; Astarita, G.; King, A.; Clapper, J. R.; Piomelli, D. *Ann. Chim.* **2007**, *97*, 887.

24. Jacob, N. T.; Lockner, J. W.; Kravchenko, V. V.; Janda, K. D. *Angew. Chem. Int. Ed.* **2014**, *53*, 6628.

25. Park, J.-G.; Kahn, J. N.; Tumer, N. E.; Pang, Y.-P. *Sci. Rep.* **2012**, *2*, 631.

26. Kenseth, J. R.; Coldiron, S. J. *Curr. Opin. Chem. Biol.* **2004**, *8*, 418.

27. Chung, C. C.; Ohwaki, K.; Schneeweis, J. E.; Stec, E.; Varnerin, J. P.; Goudreau, P. N.; Chang, A.; Cassaday, J.; Yang, L.; Yamakawa, T.; Kornienko, O.; Hodder, P.; Inglese, J.; Ferrer, M.; Strulovici, B.; Kusunoki, J.; Tota, M. R.; Takagi, T. *Assay Drug Dev. Technol.* **2008**, *6*, 361.

28. Simeonov, A.; Jadhav, A.; Thomas, C. J.; Wang, Y.; Huang, R.; Southall, N. T.; Shinn, P.; Smith, J.; Austin, C. P.; Auld, D. S.; Inglese, J. *J. Med. Chem.* **2008**, *51*, 2363.

29. Gribbon, P. *Drug Discov. Today* **2003**, *8*, 1035.

30. Turek-Etienne, T. C.; Small, E. C.; Soh, S. C.; Xin, T. A.; Gaitonde, P. V.; Barrabee, E. B.; Hart, R. F.; Bryant, R. W. *J. Biomol. Screen.* **2003**, *8*, 176.

31. Vedvik, K. L.; Eliason, H. C.; Hoffman, R. L.; Gibson, J. R.; Kupcho, K. R.; Somberg, R. L.; Vogel, K. W. *Assay Drug Dev. Technol.* **2004**, 2, 193.

32. Gul, S.; Gribbon, P. *Expert Opin. Drug Discov.* **2010**, *5*, 681.

33. Jadhav, A.; Ferreira, R. S.; Klumpp, C.; Mott, B. T.; Austin, C. P.; Inglese, J.; Thomas, C. J.; Maloney, D. J.; Shoichet, B. K.; Simeonov, A. *J. Med. Chem.* **2010**, *53*, 37.

34. Di, L.; Kerns, E. H. *Chem. Biodivers.* **2009**, *6*, 1875.

35. Reed, D.; Shen, Y.; Shelat, A. A.; Arnold, L. A.; Ferreira, A. M.; Zhu, F.; Mills, N.; Smithson, D. C.; Regni, C. A.; Bashford, D.; Cicero, S. A.; Schulman, B. A.; Jochemsen, A. G.; Guy, R. K.; Dyer, M. A. *J. Biol. Chem.* **2010**, *287*, 10786.

36. Berndsen, C. E.; Denu, J. M. *Methods* **2005**, *36*, 321.

37. Gein, V. L.; Bobyleva, A. A.; Levandovskaya, E. B.; Odegova, T. F.; Vakhrin, M. I. *Pharm. Chem. J.* **2012**, *46*, 23.

38. Gein, V. L.; Fedorova, N. L.; Levandovskaya, E. B.; Syropyatov, B. Y.; Voronina, É. V.; Danilova, N. V.; Kovaleva, M. Y. *Pharm. Chem. J.* **2011**, *45*, 355.

39. Gein, V. L.; Mar'yasov, M. A.; Silina, T. A.; Makhmudov, R. R. *Pharm. Chem. J.* **2014**, *47*, 539.

40. Gein, V. L.; Platonov, V. S.; Voronina, É. V. *Pharm. Chem. J.* **2004**, *38*, 316.

41. Gao, H.; Sun, J.; Yan, C.-G. *Beilstein J. Org. Chem.* **2013**, *9*, 2934.

42. Zhuang, C.; Miao, Z.; Zhu, L.; Dong, G.; Guo, Z.; Wang, S.; Zhang, Y.; Wu, Y.; Yao, J.; Sheng, C.; Zhang, W. *J. Med. Chem.* **2012**, *55*, 9630.

43. Baell, J. *ACS Med. Chem. Lett.* **2015**, *6*, 229.

44. Soares, K. M.; Blackmon, N.; Shun, T. Y.; Shinde, S. N.; Takyi, H. K.; Wipf, P.; Lazo, J. S.; Johnston, P. A. *Assay Drug Dev. Technol.* **2010**, *8*, 152.

45. Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P. *J. Med. Chem.* **2007**, *50*, 2385.

46. Johnston, P. A. *Curr. Opin. Chem. Biol.* **2011**, *15*, 174.

47. Johnston, P. A.; Soares, K. M.; Shinde, S. N.; Foster, C. A.; Shun, T. Y.; Takyi, H. K.; Wipf, P.; Lazo, J. S. *Assay Drug Dev. Technol.* **2008**, *6*, 505.

48. Pearce, B. C.; Sofia, M. J.; Good, A. C.; Drexler, D. M.; Stock, D. A. *Journal of Chemical Information and Modeling* **2006**, *46*, 1060.

49. Feng, B. Y.; Shoichet, B. K. *Nat. Protoc.* **2006**, *1*, 550.

50. Frenkel, Y. V.; Clark, A. D.; Das, K.; Wang, Y.-H.; Lewi, P. J.; Janssen, P. A. J.; Arnold, E. *J. Med. Chem.* **2005**, *48*, 1974.

51. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712.

52. Inglese, J.; Auld, D. S.; Jadhav, A.; Johnson, R. L.; Simeonov, A.; Yasgar, A.; Zheng, W.; Austin, C. P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11473.

53. Prinz, H. *J. Chem. Biol.* **2012**, *5*, 1.

54. Prinz, H. *J. Chem. Biol.* **2010**, *3*, 37.

55. Shoichet, B. K. *J. Med. Chem.* **2006**, *49*, 7274.

56. Copeland, R. A. *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists*; New York: Wiley-Interscience, 2005.

57. Tang, Y.; Holbert, M.; Wurtele, H.; Meeth, K.; Rocha, W.; Gharib, M.; Jiang, E.; Thibault, P.; Verreault, A.; Cole, P.; Marmorstein, R. *Nat. Struct. Mol. Biol.* **2008**, *15*, 738.

58. Burgess, R. J.; Zhou, H.; Han, J.; Zhang, Z. *Mol. Cell.* **2010**, *37*, 469.

59. Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; Lerner, C. G.; Chen, J.; Hajduk, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 217.

60. Huth, J. R.; Song, D.; Mendoza, R. R.; Black-Schaefer, C. L.; Mack, J. C.; Dorwin, S. A.; Ladror, U. S.; Severin, J. M.; Walter, K. A.; Bartley, D. M.; Hajduk, P. J. *Chem. Res. Toxicol.* **2007**, *20*, 1752.

61. Rishton, G. *Drug Discov. Today* **1997**, *2*, 382.

62. Schebb, N. H.; Faber, H.; Maul, R.; Heus, F.; Kool, J.; Irth, H.; Karst., U. *Anal. Bioanal. Chem.* **2009**, *394*, 1361.

63. Epps, D. E.; Taylor, B. M. *Anal. Biochem.* **2001**, *295*, 101.

64. Hou, X.; Li, R.; Li, K.; Yu, X.; Sun, J.-P.; Fang, H. *J. Med. Chem.* **2014**, *57*, 9309.

65. Baeza, J.; Smallegan, M. J.; Denu, J. M. *ACS Chem. Biol.* **2015**, *10*, 122.

66. Copeland, R. A. *Anal. Biochem.* **2003**, *320*, 1.

67. Foley, T. L.; Rai, G.; Yasgar, A.; Daniel, T.; Baker, H. L.; Attene-Ramos, M.; Kosa, N. M.; Leister, W.; Burkart, M. D.; Jadhav, A.; Simeonov, A.; Maloney, D. J. *J. Med. Chem.* **2014**, *57*, 1063.

68. Brown, C. S.; Lee, M. S.; Leung, D. W.; Wang, T.; Xu, W.; Luthra, P.; Anantpadma, M.; Shabman, R. S.; Melito, L. M.; MacMillan, K. S.; Borek, D. M.; Otwinowski, Z.; Ramanan, P.; Stubbs, A. J.; Peterson, D. S.; Binning, J. M.; Tonelli, M.; Olson, M. A.; Davey, R. A.; Ready, J. M.; Basler, C. F.; Amarasinghe, G. K. *J. Mol. Biol.* **2014**, *426*, 2045.

69. Zimmerman, S. S.; Khatri, A.; Garnier-Amblard, E. C.; Mullasseril, P.; Kurtkaya, N. L.; Gyoneva, S.; Hansen, K. B.; Traynelis, S. F.; Liotta, D. C. *J. Med. Chem.* **2014**, *57*, 2334.

70. Jia, J.; Xu, X.; Liu, F.; Guo, X.; Zhang, M.; Lu, M.; Xu, L.; Wei, J.; Zhu, J.; Zhang, S.; Zhang, S.; Sun, H.; You, Q. D. *PLOS ONE* **2013**, *8*, e59315.

71. Xu, Z.; Chen, Y.-W.; Battu, A.; Wilder, P.; Weber, D.; Yu, W.; MacKerell, A. D. J.; Chen, L.-M.; Chai, K. X.; Johnson, M. D.; Lin, C.-Y. *J. Med. Chem.* **2011**, *54*, 7567.

72. Hirayama, K.; Aoki, S.; Nishikawa, K.; Matsumoto, T.; Wada, K. *Bioorg. Med. Chem.* **2007**, *15*, 6810.

73. Yangthara, B.; Mills, A.; Chatsudthipong, V.; Tradtrantip, L.; Verkman, A. S. *Mol. Pharm.* **2007**, *72*, 86.

74. Arnold, L. A.; Estebanez-Perpina, E.; Togashi, M.; Jouravel, N.; Shelat, A.; McReynolds, A. C.; Mar, E.; Nguyen, P.; Baxter, J. D.; Fletterick, R. J.; Webb, P.; Guy, R. K. *J. Biol. Chem.* **2005**, *280*, 43048.

75. Dayam, R.; Al-Mawsawi, L. Q.; Neamati, N. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6155.

76. Tanrikulu, Y.; Nietert, M.; Scheffer, U.; Proschak, E.; Grabowski, K.; Schneider, P.; Weidlich, M.; Karas, M.; Gobel, M.; Schneider, G. *ChemBioChem* **2007**, *8*, 1932.

77. Gein, V. L.; Yushkov, V. V.; Splina, T. A.; Gein, L. F.; Yatsenko, K. V.; Shevtsova, S. G. *Pharm. Chem. J.* **2008**, *42*, 255.

78. Rose, R.; Erdmann, S.; Bovens, S.; Wolf, A.; Rose, M.; Hennig, S.; Waldmann, H.; Ottmann, C. *Angew. Chem. Int. Ed.* **2010**, *49*, 4129.

79. Vidović, D.; Busby, S. A.; Griffin, P. R.; Schürer, S. C. *ChemMedChem* **2011**, *6*, 94.

80. Floquet, N.; Richez, C.; Durand, P.; Maigret, B.; Badet, B.; Badet-Denisot, M.-A. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1966.

81. Bayry, J.; Tchilian, E. Z.; Davies, M. N.; Forbes, E. K.; Draper, S. J.; Kaveri, S. V.; Hill, A. V. S.; Kazatchkine, M. D.; Beverley, P. C. L.; Flower, D. R.; Tough, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10221.

82. Davies, M. N.; Bayry, J.; Tchilian, E. Z.; Vani, J.; Shaila, M. S.; Forbes, E. K.; Draper, S. J.; Beverley, P. C. L.; Tough, D. F.; Flower, D. R. *PLOS ONE* **2009**, *4*, e8084.

83. Nevin, D. K.; Peters, M. B.; Carta, G.; Fayne, D.; Lloyd, D. G. *J. Med. Chem.* **2012**, *55*, 4978.

84. Starosylaa, S. A.; Volynetsa, G. P.; Lukashova, S. S.; Gorbatiuka, O. B.; Golubb, A. G.; Bdzholaa, V. G.; Yarmoluk, S. M. *Bioorg. Med. Chem.* **2015**, *23*, 2489.

85. Canny, S. A.; Cruz, Y.; Southern, M. R.; Griffin, P. R. *Bioinformatics* **2009**, *28*, 140.

86. Han, L.; Wang, Y.; Bryant, S. H. *Bioinformatics* **2009**, *25*, 2251.

87. Hu, Y.; Bajorath, J. *AAPS J.* **2013**, *15*, 808.

88. Nissink, J. W. M.; Blackburn, S. *Future Med. Chem.* **2014**, *6*, 1113.

89. Walters, W. P.; Namchuk, M. *Nat. Rev. Drug Discov.* **2003**, *2*, 259.

90. Zorn, J.; Wille, H.; Wolan, D.; Wells, J. *J. Am. Chem. Soc.* **2011**, *133*, 19630.

91. Metz, J. T.; Huth, J. R.; Hajduk, P. J. *J. Comput. Aided Mol. Des.* **2007**, *21*, 139.

92. Cox, P. B.; Gregg, R. J.; Vasudevan, A. *Bioorg. Med. Chem.* **2012**, *20*, 4564.

93. Merchant, J. R.; Hakim, M. A.; Pillay, K. S.; Patell, J. R. *J. Med. Chem.* **1971**, *14*, 1239.

94. Gein, V. L.; Shumilovskikh, E. V.; Andreichikov, Y. S.; Saraeva, R. F.; Korobchenko, L. V.; Vladyko, V. G.; Boreko, E. I. *Khim.-Farm. Zh.* **1991**, *25*, 37.

95. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3.

96. Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615.

Graphical abstract