Accepted Manuscript

Adenosine analogs bearing phosphate isosteres as human MDO1 ligands

Yuezhou Zhang, Mikael Jumppanen, Mirko M. Maksimainen, Samuli Auno, Zulfa Awol, Léo Ghemtio, Harikanth Venkannagari, Lari Lehtiö, Jari Yli-Kauhaluoma, Henri Xhaard, Gustav Boije af Gennäs
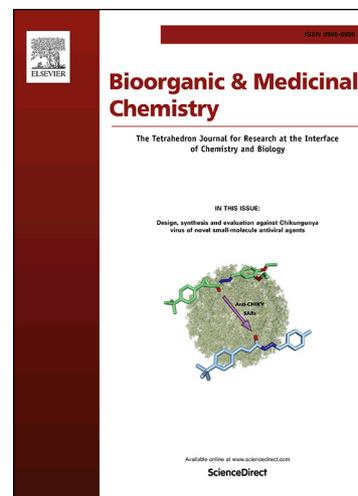
Please cite this article as: Zhang, Y., Jumppanen, M., Maksimainen, M.M., Auno, S., Awol, Z., Ghemtio, L., Venkannagari, H., Lehtiö, L., Yli-Kauhaluoma, J., Xhaard, H., Gennäs, G.B.a., Adenosine analogs bearing phosphate isosteres as human MDO1 ligands, *Bioorganic & Medicinal Chemistry* (2018), doi: https://doi.org/10.1016/j.bmc.2018.02.006

Adenosine analogs bearing phosphate isosteres as human MDO1 ligands

Leave this area blank for abstract info.

Yuezhou Zhang [a,†], Mikael Jumppanen [a,†], Mirko M. Maksimainen [b], Samuli Auno [a], Zulfa Awol [a], Léo Ghemtio [c], Harikanth Venkannagari [b], Lari Lehtiö [b], Jari Yli-Kauhaluoma [a], Henri Xhaard [a] and Gustav Boije af Gennäs [a,*]

[a] Drug Research Program, Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, FI-00014 University of Helsinki, Finland
[b] Faculty of Biochemistry and Molecular Medicine, Biocenter Oulu, FI-90014 University of Oulu, Finland
[c] Drug Research Program, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, FI-00014 University of Helsinki, Finland

Adenosine analogs bearing phosphate isosteres as human MDO1 ligands

Yuezhou Zhang [a,†], Mikael Jumppanen [a,†], Mirko M. Maksimainen [b], Samuli Auno [a], Zulfa Awol [a], Léo Ghemtio [c], Harikanth Venkannagari [b], Lari Lehtiö [b], Jari Yli-Kauhaluoma [a], Henri Xhaard [a] and Gustav Boije af Gennäs [a, *]

[a] *Drug Research Program, Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, FI-00014 University of Helsinki, Finland*

[b] *Faculty of Biochemistry and Molecular Medicine, Biocenter Oulu, FI-90014 University of Oulu, Finland*

[c] *Drug Research Program, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, FI-00014 University of Helsinki, Finland*

ARTICLE INFO

ABSTRACT

The human *O*-acetyl-ADP-ribose deacetylase MDO1 is a mono-ADP-ribosylhydrolase involved in the reversal of post-translational modifications. Until now MDO1 has been poorly characterized, partly since no ligand is known besides adenosine nucleotides. Here, we synthesized thirteen compounds retaining the adenosine moiety and bearing bioisosteric replacements of the phosphate at the ribose 5'-oxygen. These compounds are composed of either a squaryldiamide or an amide group as the bioisosteric replacement and/or as a linker. To these groups a variety of substituents were attached such as phenyl, benzyl, pyridyl, carboxyl, hydroxy and tetrazolyl. Biochemical evaluation showed that two compounds, one from both series, inhibited ADP-ribosyl hydrolysis mediated by MDO1 in high concentrations.

1. Introduction

Macrodomain proteins contain at least one copy of an α/β globular domain of about 25 kDa that is evolutionarily conserved among eukaryotes, prokaryotes and viruses.[1] Most of macrodomain proteins studied to date have been shown to function as ADP-ribose-binding modules.[2] Humans have eleven macrodomain proteins according to the Pfam 26.0 database: macrodomain-containing proteins 1-3 (MDO1-3), five members of the poly(ADP-ribose) polymerase superfamily, and three H2A histone family members.[3]

The function(s) of human MDO1 (as well as its homologs MDO2 and MDO3) are poorly known. MDO1 is localized in the mitochondria, suggesting a contribution on ATP generation, signaling, cellular differentiation and cellular growth.[4] MDO1 has been shown to function as a mono-ADP-ribosylhydrolase and a deacetylase of *O*-acetyl- ADP-ribose.[5, 6] Indirectly, MDO1 may thus be implicated in gene silencing.[7] Human MDO1 is associated with multiple cancers, such as ovarian carcinoma,[8] breast carcinoma,[9] acute leukemia[10] and prostate cancer.[11] This makes the human MDO1 protein a promising new target for potential therapeutic applications.[12] As only a few synthetic ligands have been reported, the endogenous ligand ADP-ribose should serve as a starting point for ligand-based design.

The crystal structure of human MDO1 has been solved in the *apo* state (PDB code: 2X47).[6] MDO1 is closest to MDO2 (PDB code: 4IQY)[7] with a sequence identity of 57% and a root mean square deviation of 1.61 Å over 220 Cα atoms. MDO1 and MOD2 share essentially the same binding site for ADP-ribose, with two conservative changes only, Val271[MDO1] to Ile189[MDO2]; Phe272[MDO1] to Tyr190[MDO2]. Modeling the bound mode of ADP-ribose in the *apo* MDO1 nonetheless requires a change in the side-chain conformation of Phe306[MDO1] to make it similar to Phe224[MDO2]. Affinities towards ADP-ribose are comparable for these proteins with a K_D of 0.9 μM for MDO1 and a K_D of 0.15 μM for MDO2.[4, 13]

The first objective of this study is to explore modifications to the phosphate-ribose part of ADP-ribose. Phosphate-containing compounds typically show poor bioavailability, are not physiologically stable, show low permeability, and their syntheses may be challenging.[14, 15] The second objective is to explore the binding pocket for a location that is able to accommodate large and flat fluorescent groups such as fused tricyclic aromatic rings. Such compounds, exemplified by fluorescein, would be useful fluorescent probes in cell-based studies of MDO1.

We report the design and synthesis of two series of adenosine derivatives targeting the ADP-ribose binding pocket of MDO1.

Implementing an isosteric replacement strategy, adenosine was coupled to unsymmetrically substituted squaryldiamides and various amide linkers. Squaryldiamides possess two sp^2 oxygen atoms that carry a double partial negative charge through their delocalized electrons, and have been reported to act as good mimics of the phosphate moiety.[16, 17] An amide group acts also as a promising phosphate replacement, as exemplified by structural isosteres found in the Protein Data Bank.[18] Amides have also been shown to be successful mimics of phosphodiester internucleoside linkages in RNA[19] and they are synthetically feasible.

2. Results

2.1. Ligand-based design

In the first series of derivatives, we designed eight asymmetrically substituted squaryldiamides with different lengths, shapes and polarities (Table 1). We chose four benzylic (compounds **21**, **24**, **26** and **28**), aliphatic (compound **23**), aryl (compounds **22** and **25**), and heteroaromatic (compound **27**) substituents. In the second series of five derivatives, adenosine analogs with an amide linker were designed (compounds **31**, **33**-**34**, **38** and **39**, Table 2). To investigate the replacement of the distal phosphate of ADP-ribose with a hydroxy or carboxyl group, three of these compounds were designed as derivatives of L-serine (compounds **31** and **34**) and L-aspartic acid (compound **33**). The carboxyl group of aspartic acid is similar to the phosphate group especially in terms of geometry and pK_a (pK_a of 3-5 for carboxylic acid vs. 1.54 and 6.31 for the pK_{a1} and pK_{a2} of phosphate).[20] Compounds **31** and **33** are Fmoc-protected ones tested to explore the ability of the binding pocket to accommodate large and flat fluorescent groups. This could be employed in the development of fused tricyclic fluorescent probes for future screening studies of MDO1. Two compounds contained a tetrazolyl group (compounds **38** and **39**) and were designed since tetrazole is a classical phosphate bioisosteric replacement[21] with a pK_a around 4.90.[22] The compound **39** with the 1H proton of the tetrazole moiety substituted by a phenyl group was designed to study the effect of an acidic proton of the 1H-tetrazole moiety for an enzymatic activity.

2.2. Synthesis

The compounds of the first series, the squaryldiamides **21**-**28**, were synthesized in three reaction steps starting from diethyl squarate (**4**) and two different amines (Scheme 1). First, the key building block, acetonide-protected 5'-amino-5'-deoxyadenosine (**3**) was synthesized. We treated 2',3'-O-isopropylideneadenosine (**1**) first with diphenylphosphoryl azide (DPPA) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU), and subsequently with sodium azide and 1,4,7,10,13-pentaoxacyclopentadecane (15-crown-5) in 1,4-dioxane under reflux to yield the acetonide-protected 5'-azido-5'-deoxyadenosine (**2**) in 86% yield. Next, the azido intermediate (**2**) was converted to the desired key building block (**3**) by using a modified Staudinger reaction in the presence of polymer-bound triphenylphosphine.[20] The acetonide-protected 5'-amino-5'-deoxyadenosine (**3**) was obtained in 70% yield.

We started the synthesis of target compounds by treating diethyl squarate (**4**) with various aliphatic, aromatic and benzylic amines in ethanol to give eight differently substituted alkoxy amino squarates (**5**-**12**) in 25-92% yields. Only mono-substituted products were formed in this reaction, which was consistent with the literature.[23] In a similar fashion, the acetonide-protected 5´-amino-5´-deoxyadenosine (**3**) was reacted with alkoxy amino squarates **5**-**12**. This transformation produced the acetonide-

protected and unsymmetrically 3,4-substituted cyclobut-3-ene-1,2-diones **13**-**20** that were subsequently deprotected in the presence of aqueous trifluoroacetic acid (TFA) to give a set of eight different 3-(5'-amino-5'-deoxyadenosine)-4-(alkylamino)cyclobut-3-ene-1,2-diones (**21**-**28**) in 45-90% yields.

The synthesis of the amide series was started by coupling amine **3** with *N*-hydroxysuccinimide (OSu)-activated ester and Fmoc- and *tert*-butyl-protected amino acids (Fmoc-Ser(*t*Bu)-OSu and Fmoc-Asp(O*t*Bu)-OSu, respectively) to give compounds **29** and **30** in 70 and 75% yields, respectively (Scheme 2). Deprotection of the acetonide- and the *tert*-butyl groups of compounds **29** and **30** was performed in the presence of TFA in H$_2$O/CH$_2$Cl$_2$ to give the Fmoc-protected amino acid derivatives **31** and **33** in 70 and 90% yields, respectively. The serine derivative **34** was obtained in 27% yield after deprotection of the Fmoc group with diethylamine (DEA) in *N*,*N*-dimethylformamide (DMF).

In the syntheses of the two tetrazole analogs **38** and **39**, amine **3** was first reacted with either 1H-tetrazole-5-acetic acid or [(1-phenyl-1H-tetrazol-5-yl)thio]acetic acid using *N*,*N*,*N'*,*N'*-tetramethyl-*O*-(1H-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU) as a coupling reagent to give the amides **36** and **37** in 80 and 85% yields, respectively (Scheme 3). Subsequently, the acetonide-protecting groups of amides **36** and **37** were removed with TFA/H$_2$O/CH$_2$Cl$_2$ (5:1:4) to give the products **38** and **39** in 83 and 77% yields, respectively.

2.3. Biological activities

We first attempted to measure quantitatively the binding affinities of the set of compounds, together with those of benchmark nucleotides (ADP-ribose, ADP, AMP), using isothermal titration calorimetry. We also tried differential scanning fluorimetry, where the stabilizing effect of compounds on temperature-dependent denaturation is measured by the unspecific binding of a hydrophobic reporter dye to the protein core.[24] Molecules that are better stabilizers are thought to be tighter binders. Unfortunately, as a result of these two experiments, no detection could be made in ITC and only very small ΔT_m could be obtained even for the ADP and AMP controls (less that one °C unit). For the endogenous substrate ADP-ribose to the human MDO1, the K_D measured using isothermal titration calorimetry is in the order of 0.9 μM[25] and the associated ΔT_m measured in this study was about 5 °C. Thus, the results suggest weaker affinities than for the endogenous substrate ADP-ribose, in the low to high μM if not mM. Overall, these findings are not surprising: it is unlikely that an optimal fit of the phosphate replacement in the binding pocket, associated with an affinity as good as that of ADP-ribose, will be found on a first try for our compounds.

We therefore resorted to use a biochemical assay, where presence of mono-biotin-ADP-ribosylated SRPK2 is followed qualitatively using a western blot (Figure 1). In this setup the negative control is the mono-biotin-ADP-ribosylated SRPK2 alone, whereas the positive control is composed of SPRK2 and MDO1 without an inhibitor. Inhibited MDO1 cannot efficiently hydrolyze the modification from SRPK2 resulting in stronger signal than the positive control. The analysis revealed that **28** and **31**, and to a lesser extent **27**, inhibit the catalytic activity of MDO1. Compound **28** is more efficient inhibitor than **31** based on the visual comparison of their relative intensities.

However, we could not confirm the results obtained from western blot using the Alphascreen-based activity assay for mono-ADP-ribosyl hydrolyzing macrodomains.[26] The assay utilizes the same principle as the western blot and measures hydrolysis of ADP-ribosyl group from SRPK2. Dose-response

measurements did not show any significant inhibition for compounds **28** and **31** against MDO1 or MDO2 (Supplementary Material S1) even at the highest compound concentration 6.6 mM. This indicates that the ADP-ribose analogs lacking the

distal ribose only show low potencies against macrodomains similar to ADP and AMP.

**Scheme 1.** Synthesis of the squaryldiamide derivatives.[a]
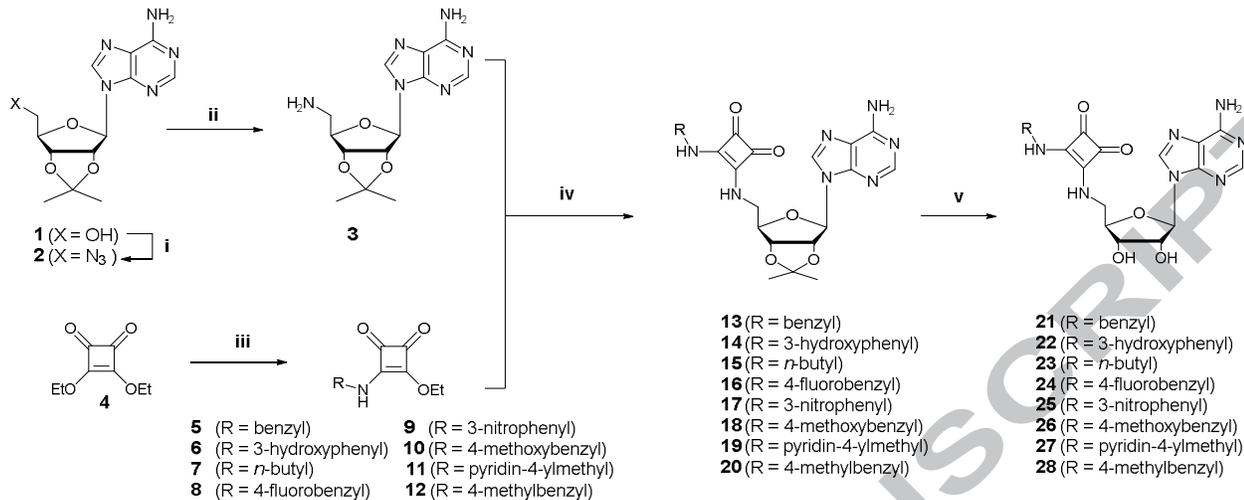
a Reagents and conditions: i) DPPA, DBU, NaN₃, 15-crown-5, 1,4-dioxane, rt to 110 °C, overnight, 86%; ii) triphenylphosphine (polymer-bound), THF, H₂O, rt, 4 h, 70%; iii) amine, EtOH, 0 °C to rt, 2 h, 25-92%; iv) DIPEA, EtOH, rt, 5 h to overnight, 37-94%; v) TFA/H₂O (4:1), rt, 5 h, 45-90%.

Scheme 2. Synthesis of the amino acid derivatives.[a]
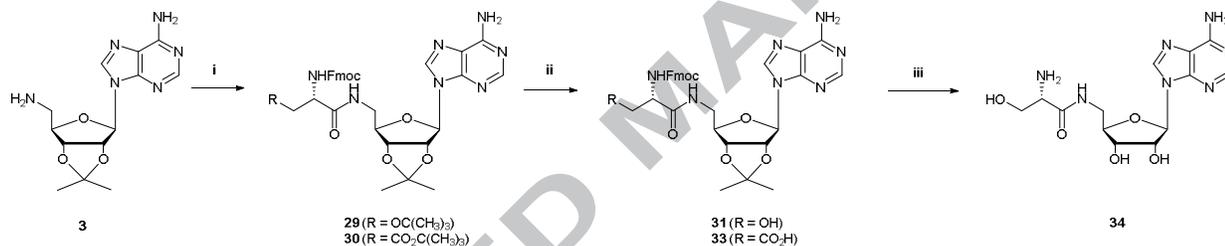
a Reagents and conditions: i) Fmoc-Ser(*t*Bu)-OSu or Fmoc-Asp(O*t*Bu)-OSu, MeOH, rt, 2 h, 70-75%; ii) TFA/H₂O/CH₂Cl₂ (4-5:1:1-4), rt, 5-24 h, 70-90%; iii) DEA, DMF, rt, 5 h, 27%.

Scheme 3. Synthesis of the tetrazole derivatives.[a]
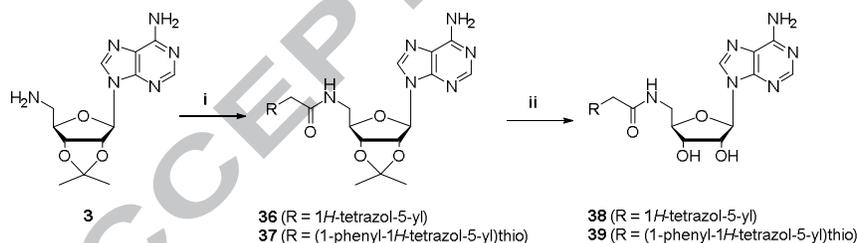
a Reagents and conditions: i) 1*H*-tetrazole-5-acetic acid or [(1-phenyl-1*H*-tetrazol-5-yl)thio]acetic acid, HBTU, DIPEA, DMF rt, 4 h, 80-85%; ii) TFA/H₂O/CH₂Cl₂ (5:1:4), rt, 4-5 h, 77-83%.
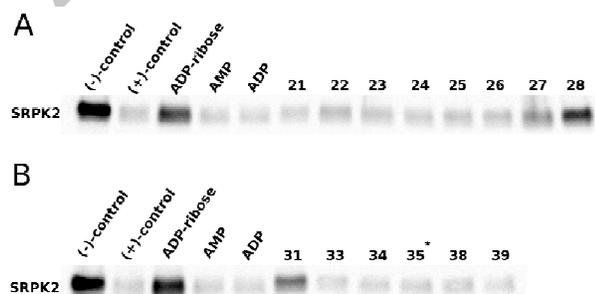
Figure 1. Western blot analysis of the inhibition efficiency for MDO1 in the hydrolysis of mono-biotin-ADP-ribosylation from SRPK2 (**A**) compounds **21-28** (**B**) compounds **31-34,** and **38-39**. Compounds **21**, **26-28** and **33-39** were tested at 1 mM concentrations, and compounds **22-25** and **31** were tested at 0.8 mM concentrations. Compound **35** is not described in this manuscript.[27]

Table 1. Inhibition of ADP-ribosyl hydrolysis and docking scores for the compounds of the squaryldiamide series prepared as shown Scheme 1.[a]
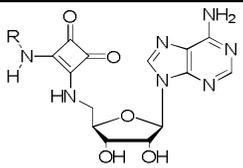
Compound	R	Docking score	Inhibition
AMP	-	-10.376	-
ADP	-	-11.029	-
ADP-ribose	-	-14.566	+++
21		-10.203	-
22		-8.972	-
23		-10.017	-
24		-9.918	-
25		-10.171	-
26		-10.692	-
27		-10.082	+
28		-11.239	+++
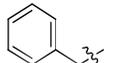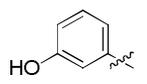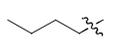
[a] Readout: -, no reading; +, weak reading; +++, high reading. Compounds were tested at 0.8-1 mM concentrations. Docking scores were calculated using the XP scoring function (Schrödinger Ltd) The lower the value, the more favorably the compound is predicted to bind.
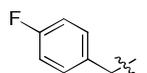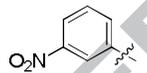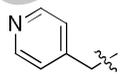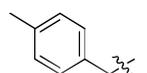
2.4. Docking simulations of ADP-ribose, compound 28 and compound 31.

As part of the design process, all the compounds were prospectively docked to the three-dimensional structure of MDO1. We hypothesized that the crystal structure of MDO2 can be used

Table 2. Inhibition of ADP-ribosyl hydrolysis and docking scores for the compounds of the amino acid series (Schemes 2 and 3).[a]
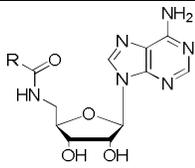
Compound	R	Docking score	Inhibition
31		-10.043	++
33		-11.902	-
34		-10.678	-
38		-9.973	-
39		-9.894	-
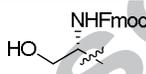
[a] Readout: -, no reading; ++, moderate reading. Compounds were tested at 0.8-1 mM concentrations. Docking scores were calculated using the XP scoring function (Schrödinger Ltd) The lower the value, the more favorably the compound is predicted to bind.

to derive the binding mode of ADP-ribose to MDO1. This is especially supported by the near-identical binding sites. Since the binding site of the apo structure is occluded, the conformation of the side chains Val181 and Phe306 had to be modified to accommodate nucleotide ligands, and this was done by reconstructing a homology model of MDO1 based on MDO2 (Supplementary Material S2). Redocking ADP-ribose to MDO1 led to a very similar binding mode as in MDO2 (Figure 2A and Figure 2B) but only after a distance restraint had forced the adenine ring onto its binding site. The docking procedure validated nonetheless that in theory no large clashes occurred at the binding site and that hydrogen-bonding groups of the protein were geometrically s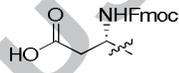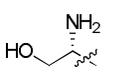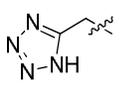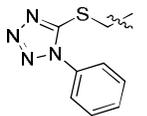uitably positioned, i.e. that the compounds were potentially inhibitors of MDO1. ADP-ribose obtained clearly the best score with a value of -14.566), while the docking scores of the second most active compounds (ADP and 28) were also among the best. We next focused on the analysis of the most active compounds, 28 and 31 (Figure 2C and 2D).

For compounds **28** and **31**, the adenine ring and the ribose were constrained to occupy a site equivalent to that predicted for ADP-ribose (Figure 2). The adenine ring is sandwiched between the planar phenyl group of Phe306 (π-π interaction) and the isopropyl side-chain of Val183. The amino group of adenine (N^6) is donating a hydrogen bond to the Asp160 carboxyl group. In

addition, the N^1 nitrogen from the Watson-Crick edge of adenine accepts a hydrogen bond from the main-chain nitrogen of Ile161. The hydrogen of the ribose O3' hydroxy group donates a hydrogen bond to the side-chain of Thr269 in many complexes.

We next examined which groups are docked in place of the phosphate groups of ADP-ribose. In MDO2, the negative charge
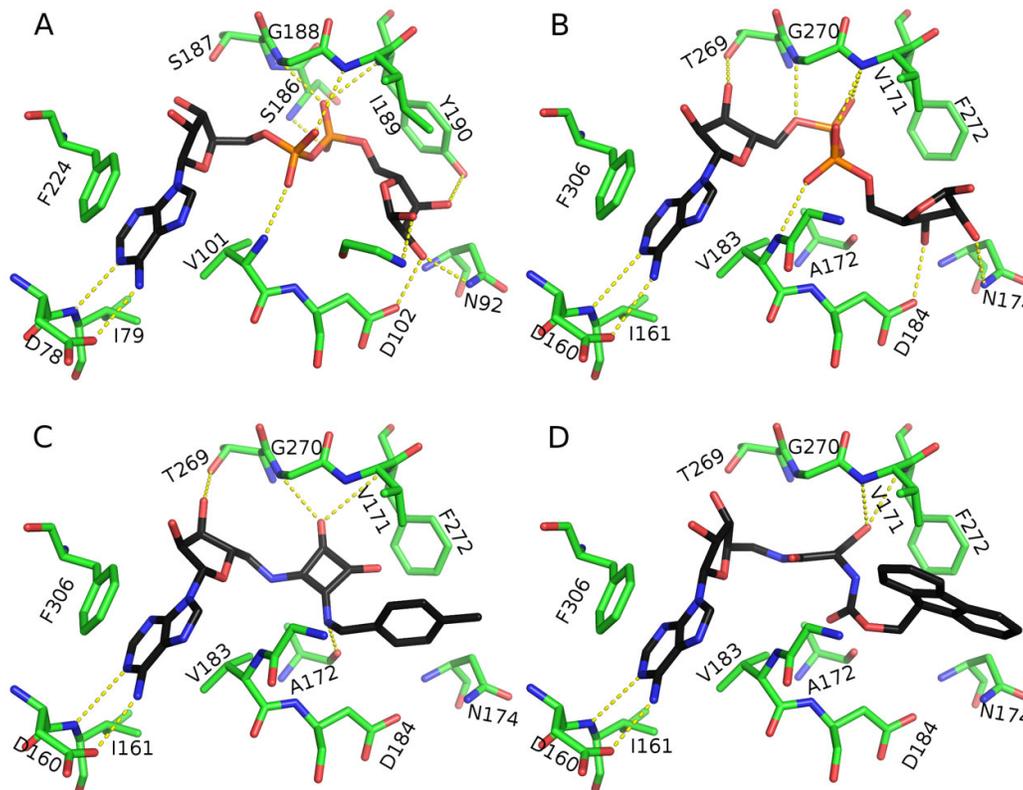

Figure 2. The complex of human MDO2 with (**A**) ADP-ribose (PDB file: 4IQY) and predicted binding modes of modeled human MDO1 in complex with (**B**) **ADP-ribose**, (**C**) compound **28**, and (**D**) compound **31**. Carbon, oxygen, and nitrogen atoms are shown in black, red, and blue, respectively. The carbon atoms of binding site residues are shown in green. Putative hydrogen bonds are depicted as dashed lines in yellow.

of the proximal phosphate is neutralized by the positively charged macrodipole at the N-terminus of helix 8. This macrodipole is generated by the sum of the dipolar contributions from the helical backbone. In addition, on the top of this helix, the main chain NH of Val101 in MDO2 donates a hydrogen bond to an oxygen atom of the proximal phosphate (Figure 2A); in MDO1 the NH of the equivalent Val183 interacts with this time the distal ADP-ribose phosphate, probably reflecting docking uncertainties (Figure 2B). For neither **28** nor **31** the amide nitrogen bonds with Gly182 and Val183 is satisfied in the docked poses. This may reflect docking uncertainties, or water molecules that do play important role in molecular recognition and may bridge these hydrogen-bonding groups, although they are not included in the docking simulation. On the other hand, it is possible that a next generation of compounds could be designed that contain a hydrogen-bond acceptor pointing towards the main chain NH groups of Val183 and Gly182. Instead, for both **28** and **31** new interactions were created with Ser268 and Thr269. In particular, an oxygen atom of the linker was found at close distance of the Val171 main-chain nitrogen, indicating the possibility of a hydrogen bond.

The third site of interest is that of the distal ribose. The ribose especially benefits from interaction with Asp160 in MDO1 and its equivalent Asp78 in MDO2. An equivalent interaction is not possible for **28** and **31** that are devoid of oxygen or nitrogen atoms in this region of the binding site. In contrast, an especially

prominent interaction is stacking on Phe272. Explaining the selectivity of **28** over its analogs **21**, **24**, **26** and **27** is highly speculative, but it may be that the methyl groups at the *para* position create additional hydrophobic interactions.

3. Conclusions

In conclusion, we designed and synthesized novel ADP-ribose analogs targeted to MDO1 by replacing the phosphates with bioisosteres and a set of aromatic substituents in place of the distal ribose. As a result, two inhibiting compounds were found, providing early indication about the SAR of MDO1 compounds. The results suggested that both a squaryldiamide and an amide linker were accepted by the binding site. A 4-methyl substituent in the phenyl group of **28** provided additional potency perhaps through interaction with an aromatic phenylalanine in the binding site and could be used as a starting point for design of more potent compounds. The next generation of derivatives could also have an additional hydrogen-bond acceptor pointing towards the main chain NH of Val183 and Gly182 to gain activity. Based on compound **31** the ADP-ribose binding groove may accommodate a bulky fluorophore group which could be a way for generation of tool compounds for studies on biological roles of the macrodomains. As the distal ribose is a key feature of the ADP-ribose, proper interactions need to be designed at this region of the binding pocket in order to achieve more potent inhibitors.

4. Experimental

4.1. Materials and general procedures

All reagents were commercially available. They were acquired from Fluka (Buchs, Switzerland) and Sigma-Aldrich (Schnelldorf, Germany) and were used without further purification. All reactions in anhydrous solvents were conducted in oven-dried glassware under argon. Thin-layer chromatography (TLC) was performed using Silica Gel 60 F_{254} (Merck) and Silica Gel 60 NH_2 F_{254}s aluminum sheets (Merck), visualized by UV illumination and stained with ninhydrin in EtOH (1.5% w/v). Column chromatography was performed with an automated high performance flash chromatography Biotage Sp4-system or with a Isolera Spektra One-system (Uppsala, Sweden) using a 0.1-mm path length flow cell UV-detector/recorder module (fixed wavelength 254 nm) for the Sp4-system or a variable UV-VIS (200-800 nm) photodiode array detector for the Isolera Spektra One-system, and the indicated mobile phase. The ^1H and ^{13}C NMR spectra were recorded on a Varian Mercury-VX 300 (Agilent Technologies, Santa Clara, California, USA) or Bruker Ascent 400 (Bruker Corporation, Billerica, Massachusetts, USA) spectrometer as solutions in $CDCl_3$, DMSO-d_6 and CD_3OD. Chemical shifts (δ) are reported as parts per million (ppm) relative to the solvent peak ($CDCl_3$ 7.26 and 77.16 ppm, DMSO-d_6 2.50 and 39.52 ppm, CD_3OD 3.31 and 49.00 ppm). Multiplicities of peaks are represented by s (singlet), d (doublet), t (triplet), q (quartet), quintet (qn), and m (multiplet). HPLC–MS analyses were performed to determine purity of all tested compounds. Purity of all tested compounds was > 95%. Mass spectra were measured on a Bruker Daltonics Esquire-HPLC spectrometer, with XTerra MS RP18 column (4.6 mm × 30 mm, 2.5 μm). High resolution mass spectra (HRMS) were measured on a Waters Synapt G2 (Waters Corporation, Milford, Massachusetts, USA) and reported for the molecular ions $[M+H]^+$.

4.2. Synthesis procedures

Procedures for the synthesis and analysis of the remaining compounds are available in the Supplementary Material.

4.2.1. 2',3'-O-Isopropylidene-5'-azido-5'-deoxyadenosine (2)

To a solution of 2′,3′-O-isopropylideneadenosine (1.00 g, 3.25 mmol) in dry 1,4-dioxane (10 mL), DPPA (1.40 g, 1.48 mL, 6.5 mmol) and DBU (1.34 g, 1.50 mL, 9.8 mmol) were added dropwise. The reaction mixture was stirred at rt for 47 h. Sodium azide (1.06 g, 16.3 mmol) and 15-crown-5 (0.527 g, 0.064 mL, 0.326 mmol) were added to the reaction mixture. The resulting mixture was irradiated in a microwave reactor at 110 °C for 4 h (1 bar, 600 rpm) and cooled to rt. Filtration of the reaction mixture and evaporation of the filtrate gave a residue that was redissolved in $CHCl_3$ (90 mL). The organic phase was washed with water (3 × 15 mL) and brine (3 × 15 mL). The organic phase was dried with anhydrous Na_2SO_4, filtrated and evaporated to give a brown oil. The crude product was purified by column chromatography on silica gel (EtOAc/EtOH/n-heptane 3:1:4) yielding the azide 2 as an amorphous yellow solid (0.928 g, 86% yield) with some 15-crown-5 as an impurity. The product was used in the next step without further purification. R_f = 0.33 ($CHCl_3$/ EtOH, 20:1). ^1H NMR (300 MHz, $CDCl_3$) δ_{ppm} 8.32 (s, 1H), 7.90 (s, 1H), 6.23 (br s, 2H), 6.10 (d, J = 2.2 Hz, 1H), 5.45 (dd, J = 6.3, 2.2 Hz, 1H), 5.04 (dd, J = 6.3, 3.3 Hz, 1H), 4.41–4.31 (m, 1H), 3.55 (dd, J = 5.7, 2.9 Hz, 2H), 1.60 (s, 3H), 1.37 (s, 3H). ^{13}C NMR (75.4 MHz, $CDCl_3$) δ_{ppm} 156.0, 153.3, 149.3,

140.0, 120.4, 114.9, 90.8, 85.8, 84.2, 82.2, 52.5, 27.3, 25.5; HRMS-ESI m/z calcd for $C_{13}H_{17}N_8O_3$ 333.1418 $[M+H]^+$, found 333.1424. ^1H NMR spectral data is consistent with those previously reported.[28]

4.2.2. 2',3'-O-Isopropylidene-5'-amino-5'-deoxyadenosine (3)

Polymer-bound triphenylphosphine (loading: 1.4 mmol/g, 4.29 g, 6.02 mmol) was suspended in anhydrous THF (60 mL). The mixture was left to stand for 5 min, then a solution of the azide 2 (1.00 g, 3.01 mmol) in anhydrous THF (2 mL) was added. The suspension was stirred at rt for 4 h. Water (0.540 mL, 30.1 mmol) was added to the reaction mixture and it was stirred at rt overnight. The resin was washed with THF, the combined solutions were dried with anhydrous Na_2SO_4, filtrated, and evaporated. The crude product was purified by column chromatography on silica gel ($CHCl_3$/MeOH 10:1 → 4:1) to give the compound 3 as a white solid (645 mg, 2.11 mmol, 70% yield). R_f = 0.25 ($CHCl_3$ /MeOH, 1:4). ^1H NMR (300 MHz, $CDCl_3$) δ_{ppm} 8.34 (s, 1H), 7.91 (s, 1H), 6.03 (d, J = 3.1 Hz, 1H), 5.68 (br s, 2H), 5.46 (dd, J = 6.5, 3.1 Hz, 1H), 5.04 (dd, J = 6.4, 3.4 Hz, 1H), 4.27 (dd, J = 9.5, 4.3 Hz, 1H), 3.05 (dd, J = 13.4, 4.6 Hz, 1H), 2.98 (dd, J = 13.4, 5.8 Hz, 1H), 1.91 (br s, 2H), 1.62 (s, 3H), 1.39 (s, 3H); ^{13}C NMR (75.4 MHz, $CDCl_3$) δ_{ppm} 155.8, 153.4, 150.3, 140.2, 120.7, 114.8, 91.0, 87.5, 83.9, 82.1, 44.0, 27.5, 25.6; HRMS-ESI m/z calcd for $C_{13}H_{19}N_6O_3$ 307.1513 $[M+H]^+$, found 307.1516.

4.2.3. 3-Ethoxy-4-[(4-fluorobenzyl)amino]cyclobut-3-ene-1,2-dione (8)

4-Fluorobenzylamine (156 μL, 171 mg, 1.36 mmol) was added at 0 °C to a solution of 3,4-diethoxycyclobut-3-ene-1,2-dione (0.20 mL, 2300 mg, 1.36 mmol) in EtOH (10 mL). The reaction mixture was stirred at rt for 2 h. The solvents were removed in vacuo and the residue was purified by column chromatography on silica gel (n-heptane/EtOAc 1:1) to give compound 8 as a white solid (189 mg, 0.76 mmol, 56% yield). R_f = 0.46. ^1H NMR (400 MHz, $CDCl_3$) δ_{ppm} 7.33-7.27 (m, 2H), 7.11–6.91 (m, 2H), 5.72 (br s, 1H), 4.77 (br s, 3H), 4.57 (br s, 1H), 1.45 (t, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ_{ppm} 189.5, 183.1, 177.9, 172.3, 162.7 (d, J = 247.3 Hz), 132.9, 129.7, 116.07 (d, J = 21.6 Hz), 70.1, 48.1, 16.0; HRMS-ESI m/z calcd $C_{13}H_{13}FNO_3$ 250.0879 $[M+H]^+$, found 250.0880.

4.2.4. 3-Ethoxy-4-[(3-nitrophenyl)amino]cyclobut-3-ene-1,2-dione (9)

3-Nitroaniline (85 mg, 0.68 mmol) was added at 0 °C to a solution of 3,4-diethoxycyclobut-3-ene-1,2-dione (0.10 mL, 120 mg, 0.68 mmol) in EtOH (3 mL). The reaction mixture was stirred at rt for 5 h. The solvents were removed in vacuo and the residue was purified by column chromatography on silica gel (n-hexane/EtOAc 1:1) to give compound 9 as an orange solid (126 mg, 0.48 mmol, 71% yield). R_f = 0.29 ($CHCl_3$/MeOH, 10:1). ^1H NMR (300 MHz, DMSO-d_6) δ_{ppm} 11.06 (br s, 1H), 8.32 (t, J = 2.2 Hz, 1H), 7.92 (ddd, J = 8.1, 2.2, 0.9 Hz, 1H), 7.78 (ddd, J = 8.2, 2.3, 1.0 Hz, 1H), 7.64 (t, J = 8.2 Hz, 1H), 4.80 (q, J = 7.1 Hz, 2H), 1.44 (t, J = 7.1 Hz, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 187.4, 184.2, 178.9, 169.3, 148.3, 139.4, 130.5, 125.3, 118.0, 113.6, 69.9, 15.5; HRMS-ESI m/z calcd $C_{12}H_{11}N_2O_5$ 263.0662 $[M+H]^+$, found 263.0668.

4.2.5. *3-Ethoxy-4-[(4-methoxybenzyl)amino]cyclobut-3-ene-1,2-dione (10)*

4-Methoxybenzylamine (178 μL, 1.36 mmol) was added at 0 °C to a solution of 3,4-diethoxycyclobut-3-ene-1,2-dione (0.20 mL, 240 mg, 1.36 mmol) in EtOH (5 mL). The reaction mixture was stirred at rt for 2 h. The solvents were removed *in vacuo*, and the residue was purified by column chromatography on silica gel (*n*-hexane/EtOAc 1:1) to give compound **10** as a white solid (291 mg, 1.17 mmol, 86% yield). R_f = 0.41 (*n*-hexane/EtOAc, 1:1). ^{1}H NMR (300 MHz, CHCl$_3$) δ_{ppm} 7.24–7.02 (m, 3H), 6.91–6.81 (m, 2H), 4.75 (br s, 2H), 4.52 (br s, 2H), 3.78 (s, 3H), 1.42 (t, J = 7.5 Hz, 1H); HRMS-ESI *m/z* calcd for $C_{14}H_{16}NO_4$ 262.1001 [M+H]$^+$, found 262.1079.

4.2.6. *3-[4-Fluorobenzyl)amino]-4-(2′,3′-O-isopropylidene-5′-amino-5′-deoxyadenosine)cyclobut-3-ene-1,2-dione (16)*

To a solution of compound **8** (179 mg, 0.72 mmol) in EtOH (10 mL) was added DIPEA (0.74 mL, 0.79 mmol) and **3** (241 mg, 0.79 mmol). The reaction mixture was stirred at rt for 6 h. The solvent was removed *in vacuo*, and the residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 10:1) to give **16** as a white solid (306 mg, 0.6 mmol, 84% yield). R_f = 0.73 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (400 MHz, DMSO-d_6) δ_{ppm} 8.33 (s, 1H), 8.14 (s, 1H), 7.89 (br s, 1H), 7.34 (s, 4H), 7.17 (t, J = 8.6 Hz, 2H), 6.19 (d, J = 2.8 Hz, 1H), 5.43 (dd, J = 6.6, 2.8 Hz, 1H), 5.03 (dd, J = 6.4, 3.4 Hz, 1H), 4.66 (d, J = 6.6 Hz, 2H), 4.37–4.17 (m, 1H), 4.00–3.70 (m, 2H), 1.54 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ_{ppm} 182.7, 182.5, 167.9, 167.6, 161.4 (d, J = 243 Hz), 156.1, 152.9, 148.8, 139.8, 135.2, 129.6 (d, J = 8.4 Hz), 119.1, 115.4 (d, J = 21.2 Hz), 113.7, 88.7, 85.0, 83.1, 81.2, 46.0, 45.1, 27.0, 25.3; HRMS-ESI *m/z* calcd $C_{24}H_{25}FN_7O_5$ 510.1901 [M+H]$^+$, found 510.1901.

4.2.7. *3-(2′,3′-O-Isopropylidene-5′-amino-5′-deoxyadenosine)-4-[(3-nitrophenyl)amino]cyclobut-3-ene-1,2-dione (17)*

To a solution of compound **9** (64 mg, 0.31 mmol) in EtOH (3 mL) was added DIPEA (0.060 mL, 0.33 mmol) and **3** (0.10 g, 0.33 mmol). The reaction mixture was stirred at rt for 5 h. The solvent was removed *in vacuo*, and the residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 10:1) to give compound **17** as a yellow solid (155 mg, 0.30 mmol, 90%). R_f = 0.46 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 9.95 (br s, 1H), 8.35 (s, 2H), 8.15 (s, 1H), 7.82 (ddd, J = 8.1, 2.2, 0.9 Hz, 1H), 7.70 (d, J = 8.1 Hz, 1H), 7.58 (t, J = 8.1 Hz, 1H), 7.30 (s, 2H), 6.24 (d, J = 2.6 Hz, 1H), 5.47 (dd, J = 6.3, 2.7 Hz, 1H), 5.09 (dd, J = 6.3, 3.7 Hz, 1H), 4.33 (dt, J = 8.0, 4.2 Hz, 1H), 4.04 (dd, J = 14.0, 4.8 Hz, 1H), 3.87 (dd, J = 13.9, 7.3 Hz, 1H), 1.56 (s, 3H), 1.34 (s, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 184.3, 180.5, 169.5, 163.0, 156.1, 152.8, 148.7, 148.6, 140.3, 140.0, 130.6, 123.9, 119.2, 116.6, 113.8, 112.3, 88.6, 84.9, 83.1, 81.0, 45.4, 27.0, 25.3; HRMS-ESI *m/z* calcd for $C_{23}H_{23}N_8O_7$ 523.1684 [M+H]$^+$, found 523.1689.

4.2.8. *3-(2′,3′-O-Isopropylidene-5′-amino-5′-deoxyadenosine)-4-[4-methoxybenzyl)amino]cyclobut-3-ene-1,2-dione (18)*

To a solution of compound **10** (64 mg, 0.31 mmol) in EtOH (3 mL) was added DIPEA (0.060 mL, 0.33 mmol) and **3** (0.10 g, 0.33 mmol). The reaction mixture was stirred at rt for 5 h. The solvent was removed *in vacuo*, and the residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 10:1) to give compound **18** as a white solid (150 mg, 0.30 mmol, 91% yield). R_f = 0.35 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 8.30 (s, 1H), 8.17 (s, 1H), 7.65 (br s, 1H), 7.31 (br s, 2H), 7.22 (d, J = 8.5 Hz, 2H), 6.90 (d, J = 8.6 Hz, 2H), 6.19 (d, J = 2.7 Hz, 1H), 5.42 (dd, J = 6.3, 2.6 Hz, 1H), 5.01 (dd, J = 6.3, 3.5 Hz, 1H), 4.60 (d, J = 6.0 Hz, 2H), 4.30–4.20 (m, 1H), 4.01–3.81 (m, 1H), 3.73 (s, 3H), 1.54 (s, 3H), 1.32 (s, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 183.0, 182.6, 168.1, 167.8, 159.0, 156.3, 153.2, 149.0, 140.3, 130.9, 129.2, 119.3, 114.3, 89.0, 83.4, 81.4, 70.1, 55.4, 46.6, 45.4, 27.3, 25.5; HRMS-ESI *m/z* calcd $C_{25}H_{28}N_7O_6$ 522.2096 [M+H]$^+$, found 522.2101.

4.2.9. *3-(2′,3′-O-Isopropylidene-5′-amino-5′-deoxyadenosine)-4-[(4-methylbenzyl)amino]cyclobut-3-ene-1,2-dione (20)*

To a solution of **3** (0.10 g, 0.33 mmol) in EtOH (3 mL) was added DIPEA (59 μL, 0.33 mmol) and **12** (0.080 mg, 0.33 mmol). The reaction mixture was stirred at rt for 5 h and evaporated *in vacuo*. Purification of the crude product by column chromatography on silica gel (CHCl$_3$/MeOH 10:1→ 4:1) gave compound **20** as a white solid (140 mg, 0.28 mmol, 85%). R_f = 0.29 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 8.31 (s, 1H), 8.17 (s, 1H), 7.64 (br s, 1H), 7.37 (br s, 3H), 7.25–7.08 (m, 4H), 6.19 (s, 1H), 5.49–5.35 (m, 1H), 5.01 (dd, J = 6.2, 3.4 Hz, 1H), 4.74–4.50 (m, 2H), 4.32–4.17 (m, 1H), 3.90 (br s, 1H), 3.80–3.63 (m, 1H), 2.27 (s, 3H), 1.54 (s, 3H), 1.32 (s, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 182.7, 182.5, 167.5, 156.2, 152.9, 148.8, 139.9, 136.6, 135.8, 129.2, 127.5, 119.2, 113.7, 88.7, 85.2, 83.1, 81.2, 46.6, 45.2, 27.1, 25.3, 20.7; HRMS-ESI *m/z* calcd $C_{25}H_{28}N_7O_5$ 506.2146 [M+H]$^+$, found 506.2152.

4.2.10. *3-(5′-Amino-5′-deoxyadenosine)-4-[(4-fluorobenzyl)amino]cyclobut-3-ene-1,2-dione (24)*

A suspension of **16** (0.281 g, 0.55 mmol) in TFA/water (4:1, 10 mL) was stirred at rt for 6 h. The solvent was evaporated *in vacuo*, and the residue was co-evaporated with toluene to remove any residual TFA and H$_2$O. The remaining residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 4:1) to give compound **24** as a light green solid (194 mg, 0.41 mmol, 75% yield). R_f = 0.57 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (400 MHz, DMSO-d_6) δ_{ppm} 8.33 (s, 1H), 8.15 (s, 1H), 8.10 (br s, 1H), 7.80 (br s, 1H), 7.41–7.31 (m, 2H), 7.29 (s, 2H), 7.21-7.10 (m, 2H), 5.91 (d, J = 5.9 Hz, 1H), 5.57 (d, J = 6.1 Hz, 1H), 5.38 (d, J = 4.9 Hz, 1H), 4.73-4.59 (m, , 3H), 4.20-4.12 (m, 1H), 4.06–3.97 (m, 1H), 3.94 (br s, 1H), 3.85–3.72 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ_{ppm} 182.6, 182.5, 168.0, 167.5, 161.5 (d, J = 243.0 Hz), 158.0 (q, J = 30.9 Hz) 156.1, 152.7, 149.4, 139.8, 135.2, 129.8 (d, J = 8.2 Hz), 119.2, 117.3 (q, J = 300.0 Hz), 115.4 (d, J = 21.4 Hz) 87.4, 83.6, 72.8, 70.8, 46.0, 45.6; HRMS-ESI *m/z* calcd $C_{21}H_{21}FN_7O_5$ 470.1588 [M+H]$^+$, found 470.1589. TFA

4.2.11. *3-(5′-Ammonium-5′-deoxyadenosine)-4-[(3-nitrophenyl)amino]cyclobut-3-ene-1,2-dione trifluoroacetate (25)*

A suspension of **17** (88 mg, 0.16 mmol) in TFA/water (4:1, 3 mL) was stirred at rt for 5 h. The solvent was evaporated *in vacuo*, and the residue was co-evaporated with toluene to remove any residual TFA and H$_2$O. The remaining residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 4:1) to give compound **25** as a yellow solid (63 mg, 0.13 mmol, 66% yield). R_f = 0.48 (CHCl$_3$/MeOH, 4:1). ^{1}H NMR (300 MHz,

DMSO-d_6) δ_{ppm} 10.44 (br zs, 1H), 8.43 (br s, 1H), 8.35 (br s, 1H), 8.22 (br s, 1H), 8.13 (s, 1H), 7.86–7.72 (m, 2H), 7.59 (t, J = 8.1 Hz, 1H), 7.24 (br s, 2H), 5.94 (d, J = 5.6 Hz, 1H), 5.56 (d, J = 5.9 Hz, 1H), 5.38 (d, J = 5.1 Hz, 1H), 4.71 (q, J = 5.5 Hz, 1H), 4.32–4.21 (m, 1H), 4.14–3.99 (m, 2H), 3.97–3.83 (m, 1H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 184.4, 180.4, 169.7, 163.0, 157.9 (q, J = 30.9 Hz), 156.1, 152.7, 149.4, 148.6, 140.6, 139.9, 130.7, 123.9, 119.2, 117.3 (q, J = 300.2 Hz), 116.6, 112.3, 87.6, 83.3, 72.8, 70.8, 45.9; HRMS-ESI m/z calcd $C_{20}H_{19}N_8O_7$ 483.1377 [M+H]$^+$, found 483.1378.

4.2.12. 3-(5′-Amino-5′-deoxyadenosine)-4-[(4-methoxybenzyl)amino]cyclobut-3-ene-1,2-dione (26)

A suspension of **18** (0.10 g, 0.20 mmol) in TFA/water (4:1, 3 mL) was stirred at rt for 5 h. The solvent was evaporated *in vacuo*, and the residue was co-evaporated with toluene to remove any residual TFA and H_2O. The remaining residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 4:1) to give compound **26** as a white solid (0.080 g, 0.17 mmol, 85% yield). R_f = 0.38 (CHCl$_3$/MeOH, 4:1). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 8.31 (s, 1H), 8.15 (s, 1H), 7.86 (br s, 1H), 7.57 (br s, 1H), 7.27 (s, 2H), 7.22 (d, J = 8.3 Hz, 2H), 6.89 (d, J = 8.3 Hz, 2H), 5.91 (d, J = 5.9 Hz, 1H), 5.52 (d, J = 5.9 Hz, 1H), 5.33 (d, J = 4.2 Hz, 1H), 4.73–4.64 (m, 1H), 4.61 (d, J = 5.9 Hz, 2H), 4.19–4.12 (m, 1H), 4.04–3.97 (m, 1H), 3.98–3.88 (m, 1H), 3.83–3.74 (m, 1H), 3.72 (s, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 182.6, 182.4, 167.9, 167.4, 158.6, 156.1, 152.7, 149.4, 139.8, 130.8, 128.9, 119.2, 114.0, 87.4, 83.6, 72.8, 70.8, 55.1, 46.2, 45.5; HRMS-ESI m/z calcd $C_{22}H_{24}N_7O_6$ 482.1788 [M+H]$^+$, found 482.1785.

4.2.13. 3-(5′-Amino-5′-deoxyadenosine)-4-[(4-methylbenzyl)amino]cyclobut-3-ene-1,2-dione (28)

A suspension of **20** (0.10 g, 0.20 mmol) in TFA/water (4:1, 3 mL) was stirred at rt for 5 h. The solvent was evaporated *in vacuo*, and the residue was co-evaporated with toluene to remove any residual TFA and H_2O. The remaining residue was purified by column chromatography (CHCl$_3$/MeOH 4:1, KP-NH column) to give compound **28** as solid (80 mg, 0.17 mmol, 85% yield). R_f = 0.18 (CH$_2$Cl$_2$/MeOH, 4:1, NH$_2$ TLC). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 8.32 (s, 1H), 8.16 (s, 1H), 7.74 (br s, 1H), 7.45 (br s, 1H), 7.33 (s, 2H), 7.16 (s, 4H), 5.92 (d, J = 5.9 Hz, 1H), 5.57 (br s, 1H), 5.39 (br s, 1H), 4.75–4.54 (m, 3H), 4.22–4.10 (m, 1H), 4.08–3.90 (m, 2H), 3.88–3.69 (m, 1H), 2.26 (s, 3H); ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 182.7, 182.6, 167.8, 167.5, 156.1, 152.8, 149.4, 139.9, 136.6, 135.8, 129.2, 127.5, 119.3, 87.5, 83.7, 72.8, 70.9, 46.6, 45.6, 20.7; HRMS-ESI m/z calcd $C_{22}H_{24}N_7O_5$ 466.1839 [M+H]$^+$, found 466.1840.

4.2.14. (9H-Fluoren-9-yl)methyl [(S)-1-[[[(3aS,4S,6S,6aS)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetra-hydrofuro[3,4-d][1,3]dioxol-4-yl]methyl]amino]-3-(tert-butoxy)-1-oxopropan-2-yl]carbamate (29)

To a solution of **3** (0.10 g, 0.33 mmol) in anhydrous MeOH (10 mL) was added Fmoc-L-Ser(tBu)-OSu (160 mg, 0.33 mmol). The reaction mixture was stirred at rt for 2.5 h under argon atmosphere. The solvents were removed *in vacuo*, and the residue was purified by column chromatography on silica gel (EtOAc/EtOH/n-heptane 3:1:4) to give compound **29** as a white solid (154 mg, 0.23 mmol, 70% yield). R_f = 0.65 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (300 MHz, CDCl$_3$) δ_{ppm} 8.49 (s, 1H), 8.11 (s,

1H), 7.82 (s, 1H), 7.72 (d, J = 7.6 Hz, 2H), 7.53 (t, J = 6.7 Hz, 2H), 7.36 (t, J = 7.6 Hz, 2H), 7.30–7.21 (m, 2H), 5.93–5.86 (m, 3H), 5.79 (d, J = 7.3 Hz, 1H), 5.35 (dd, J = 6.2, 4.1 Hz, 1H), 4.89 (dd, J = 6.2, 2.5 Hz, 1H), 4.44 (s, 1H), 4.37 (d, J = 7.0 Hz, 2H), 4.15 (t, J = 7.1 Hz, 1H), 3.92 (s, 1H), 3.74 (dd, J = 8.9, 3.9 Hz, 1H), 3.52 (d, J = 7.5 Hz, 1H), 3.48 (s, 2H), 1.61 (s, 3H), 1.35 (s, 3H), 1.10 (s, 9H); ^{13}C NMR (75.4 MHz, CDCl$_3$) δ_{ppm} 171.2, 156.0, 153.5, 149.2, 144.1, 143.9, 141.4, 140.3, 127.8, 127.8, 127.2, 127.1, 125.2, 121.0, 120.1, 114.8, 92.0, 84.2, 83.2, 81.9, 74.0, 67.3, 62.3, 55.5, 47.3, 41.6, 27.7, 27.5, 25.6.; HRMS-ESI m/z calcd $C_{35}H_{42}N_7O_7$ 672.3146 [M+H]$^+$, found 672.3146.

4.2.15. tert-Butyl (S)-3-[[[(9H-fluoren-9-yl)methoxy]carbonyl]amino]-4-[[[(3aS,4S,6S,6aS)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl]methyl]amino]-4-oxobutanoate (30)

To a solution of **3** (84 mg, 0.27 mmol) in anhydrous MeOH (10 mL) was added Fmoc-L-Asp(OtBu)-OSu (140 mg, 0.27 mmol). The reaction mixture was stirred at rt for 2 h. The solvents were removed in reduced pressure and the residue was purified by column chromatography on silica gel (CHCl$_3$/MeOH 10:1) to give compound **30** as a white solid (160 mg, 0.22 mmol, 75% yield). The product was used in the next step without further purification. R_f = 0.50 (CHCl$_3$/MeOH, 10:1). ^{1}H NMR (300 MHz, CDCl$_3$) δ_{ppm} 8.56 (s, 1H), 8.43 (br s, 1H), 7.81 (s, 1H), 7.72 (d, J = 7.6 Hz, 2H), 7.54 (d, J = 7.4 Hz, 2H), 7.36 (t, J = 7.7 Hz, 2H), 7.31–7.19 (m, 2H), 6.11 (br s, 2H), 5.98 (d, J = 8.7 Hz, 1H), 5.83 (d, J = 4.5 Hz, 1H), 5.35 (t, J = 5.4 Hz, 1H), 4.89 (dd, J = 6.4, 2.3 Hz, 1H), 4.63 (br s, 1H), 4.45 (s, 1H), 4.38 (d, J = 7.3 Hz, 2H), 4.17 (t, J = 7.1 Hz, 1H), 4.03 (br s, 1H), 3.39 (d, J = 14.7 Hz, 1H), 2.90–2.73 (m, 2H), 2.10 (br s, 1H), 1.60 (s, 3H), 1.40 (s, 9H), 1.35 (s, 3H); ^{13}C NMR (75.4 MHz, CDCl$_3$) δ_{ppm} 171.2, 170.3, 156.1, 156.0, 153.6, 149.1, 143.9, 143.8, 141.4, 141.4, 140.3, 127.8, 127.2, 125.2, 120.9, 120.1, 114.7, 92.3, 83.8, 82.8, 81.8, 81.7, 67.3, 51.9, 47.3, 41.5, 38.3, 28.2, 27.6, 25.5; HRMS-ESI m/z calcd $C_{36}H_{42}N_7O_8$ 700.3095 [M+H]$^+$, found 700.3093.

4.2.16. (9H-Fluoren-9-yl)methyl [(S)-1-[[[(2S,3S,4S,5S)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetra-hydrofuran-2-yl]methyl]amino]-3-hydroxy-1-oxopropan-2-yl]carbamate (31)

A solution of **29** (94 mg, 0.14 mmol) in TFA/H$_2$O/CH$_2$Cl$_2$ (4:1:1, 5 mL) was stirred at rt for 24 h. The solvent was evaporated *in vacuo*, and the residue was co-evaporated with toluene to remove residual TFA and H$_2$O. The crude product was purified by column chromatography on silica gel (CH$_2$Cl$_2$/MeOH 0→15%) to give **31** as a white solid (56 mg, 0.10 mmol, 70% yield). R_f = 0.21 (CHCl$_3$/ MeOH, 10:1). ^{1}H NMR (400 MHz, DMSO-d_6) δ_{ppm} 8.44 (t, J = 5.9 Hz, 1H), 8.35 (s, 1H), 8.28 (s, 1H), 8.25 (d, J = 3.1 Hz, 1H), 7.88 (d, J = 7.5 Hz, 2H), 7.73 (t, J = 7.4 Hz, 2H), 7.50–7.26 (m, 6H), 5.84 (d, J = 6.6 Hz, 1H), 5.33 (br d, J = 55.9 Hz, 1H), 4.89 (br s, 1H), 4.67 (t, J = 5.8 Hz, 1H), 4.38–3.97 (m, 6H), 3.69–3.48 (m, 4H), 3.42–3.31 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ_{ppm} 170.5, 156.0, 151.1, 149.0, 143.9, 143.8, 140.9, 140.7, 127.6, 127.1, 125.3, 120.1, 119.4, 87.8, 83.8, 72.8, 71.1, 65.8, 61.7, 57.3, 46.6, 41.1; HRMS-ESI m/z calcd for $C_{28}H_{30}N_7O_7$ 576.2207 [M+H]$^+$, found 576.2205.

4.2.17. (S)-3-[[[(9H-Fluoren-9-yl)methoxy]carbonyl]amino]-4-[[[(2S,3S,4S,5S)-5-(6-amino-9H-purin-9-yl)-3,4-

dihydroxytetrahydrofuran-2-yl]methyl]amino]-4-oxobutanoic acid (33)

A mixture of **30** (120 mg, 0.17 mmol) in TFA/H$_2$O/CH$_2$Cl$_2$ (5:1:4, 5 mL) was stirred at rt for 5 h. The solvent was removed *in vacuo* and the resulting residue was co-evaporated with toluene to remove any residual TFA and H$_2$O. The crude product was purified by column chromatography on silica gel (MeOH/CHCl$_3$ 4:1) to give compound **33** as a white solid (110 mg, 0.19 mmol, 90% yield). The product was used in the next step without further purification. R_f = 0.63 (MeOH/CHCl$_3$, 4:1). ^{1}H NMR (300 MHz, DMSO-d_6) δ_{ppm} 12.27 (br s, 1H), 8.38 (app t, 1Hz), 8.32 (s, 1H), 8.25 (s, 1H), 7.88 (d, J = 7.5 Hz, 2H), 7.70 (d, J = 7.4 Hz, 2H), 7.41 (t, J = 7.4 Hz, 2H), 7.36–7.25 (m, 4H), 5.83 (d, J = 6.5 Hz, 1H), 5.38 (br s, 1H), 5.21 (br s, 1H), 4.73–4.65 (m, 1H), 4.46–4.37 (m, 1H), 4.30–4.15 (m, 4H), 4.04–3.97 (m, 2H), 3.58–3.26 (m, 2H), 2.67 (dd, J = 16.4, 5.1 Hz, 1H), 2.56 (d, J = 8.9 Hz, 1H).; ^{13}C NMR (75.4 MHz, DMSO-d_6) δ_{ppm} 171.6, 171.1, 155.8, 154.1, 150.0, 148.8, 143.8, 143.7, 141.2, 140.6, 127.6, 127.0, 125.3, 120.0, 119.3, 87.9, 83.7, 72.9, 71.1, 65.8, 51.5, 46.6, 41.2, 36.4; HRMS-ESI *m/z* calcd for C$_{29}$H$_{30}$N$_7$O$_8$ 604.2156 [M+H]$^+$, found 604.2164.

4.2.18. (S)-2-Amino-N-[[(2S,3S,4S,5S)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl]-methyl]-3-hydroxypropanamide (34)

A mixture of **31** (0.037 g, 0.064 mmol) and DEA (7.5 µL, 0.07 mmol) in DMF (0.50 mL) was stirred at rt for 5 h. The solvent was removed *in vacuo* to give a residual syrup that was purified by column chromatography (CHCl$_3$/MeOH 0→15%, KP-NH column) to give compound **34** as white crystals (6.0 mg, 0.017 mmol, 27 % yield). ^{1}H NMR (400 MHz, CD$_3$OD) δ_{ppm} 8.28 (s, 1H), 8.25 (s, 1H), 5.94 (d, J = 5.9 Hz, 1H), 4.80 (t, J = 5.7 Hz, 1H), 4.23 (dd, J = 5.5, 3.7 Hz, 1H), 4.20–4.15 (m, 1H), 3.79 (dd, J = 14.2, 5.2 Hz, 1H), 3.74–3.63 (m, 2H), 3.50 (dd, J = 14.2, 4.2 Hz, 1H), 3.45 (t, J = 5.6 Hz, 1H); ^{13}C NMR (101 MHz, CD$_3$OD) δ_{ppm} 175.9, 157.5, 153.9, 150.5, 142.0, 120.9, 90.6, 85.1, 74.7, 72.7, 65.5, 58.0, 42.2; HRMS-ESI *m/z* calcd for C$_{13}$H$_{20}$N$_7$O$_5$ 354.1526 [M+H]$^+$, found 354.1526.

4.3. Biochemical assays

The expression construct for human MDO1 (residues 58-325) with an N-terminal 6×His tag and a thioredoxin fusion tag followed by a TEV cleavage site (pNH-TrxT) was a generous gift from the Structural Genomics Consortium, Oxford. DNA for MDO2 was a kind gift from Dr. Ahola (University of Helsinki, Finland). MDO2 (residues 7-243) was cloned to pNIC-Bsa4 vector also containing a TEV cleavage site. MDO1 and MDO2 were expressed by transforming the plasmids into *E. coli* Rosetta2 (DE3) cells. The purification protocol was similar for both enzymes. Glycerol stock was used to inoculate small 5-mL pre cultures grown overnight in Luria broth (LB) media which were then used to inoculate two large 1-L cultures containing Terrific broth (TB) auto induction media with trace elements (Formedium, UK). The media was additionally supplemented with 8 g/L glycerol. The cultures were allowed to grow up to OD$_{600}$ of 1.5 at 37 °C with shaking after which the temperature was reduced to 18 °C and the cultures were allowed to grow overnight (14-16 h). The cells were then collected by centrifugation (5500×g, 25 min, 4 °C), resuspended in lysis buffer (50 mM HEPES, pH 7.5, 500 mM NaCl, 10% glycerol, 10 mM imidazole, and 0.5 mM TCEP; 1.5 mL/g of cells) and stored at -20 °C.

The cell suspension was thawed, 10 µg/mL DNAase, 0.1 mM Pefabloc, and 0.25 mg lysozyme was added followed by incubation on ice for 5 min. The cells were disrupted by sonication and the lysate was cleared by centrifugation (35000×g, 30 min, 4 °C). The supernatant was filtered using a 0.45-µM filter syringe before it was loaded to two 1-mL HisTrap HP columns (GE Healthcare, UK), which were pre-charged with Ni^{2+} and pre-equilibrated with binding buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 10 mM imidazole, and 0.5 mM TCEP). The bound protein was washed with 5 column volumes (10 mL) of buffer supplemented with 25 mM imidazole and eluted with buffer containing 350 mM imidazole. The eluted protein fraction was concentrated and diluted in size exclusion buffer (20 mM HEPES, pH 7.5, 250 mM NaCl, 5% glycerol, and 1 mM TCEP) to a final volume of 10 mL to bring the imidazole concentration down to 50 mM. TEV protease was added (1:30 molar ratio) and the sample was incubated overnight at 4 °C. The cleaved protein was run again through a 1-mL HisTrap HP column (GE Healthcare, UK). The flow-through was collected and concentrated to perform size exclusion chromatography using Superdex S-75 column (GE Healthcare, UK). The collected fractions were pooled, concentrated and divided into small aliquots which were flash frozen in liquid nitrogen and stored at -80 °C. Protein purity was analyzed with an SDS-PAGE (Supplementary Material S3).

Compounds were tested as potential MDO1 inhibitors using a biochemical assay. A mono-biotin-ADP-ribosylated SRPK2 was used as a substrate for the hydrolysis as it can be robustly ADP-ribosylated by ARTD10.[29] A mixture of 2 µM substrate, 2 µM MDO1 and 0.8-1 mM compound was incubated in 25 mM HEPES pH 7.5, 100 mM NaCl for 16 hours. The reactions were stopped by adding 2× Laemmli buffer (Bio-Rad, USA) and incubating at 95 °C for 5 min. The samples were analyzed by SDS-PAGE and transferred onto a nitrocellulose membrane (Whatman, UK), which was blocked using 1% casein in 1× TBS buffer (Bio-Rad, USA). Finally, the samples were visualized with 1:15000 streptavidin conjugated horseradish peroxidase (PerkinElmer, USA) using a chemiluminescent substrate (WesternBright^{TM} ECL, Advansta Corporation, USA).

Dose-response measurements were performed using the previously reported Alphascreen-based activity assay for mono-ADP-ribosyl hydrolyzing macrodomains.[26] Compounds were measured in quadruplicates from 6.6 mM to 0.66 µM using half-logarithmic dilutions. After transfer of compounds to the assay plate, MDO1 and MDO2 (final concentration 800 nM for both) were added to the plate followed by the addition of a mono-ADP-ribosylated SRPK2 (final concentration 50 nM). The plate was incubated for 1 h 20 min at room temperature, followed by the addition of the acceptor and donor bead mixture (final concentration 5 µg/ml) and an incubation of 3 h at room temperature in dark. The luminescence was read with a Tecan Infinite M1000 Pro plate reader using the AlphaScreen detection module with 600 ms excitation and 300 ms integration time.

4.4. Molecular modeling

Preliminary dockings by us and others (see Supplementary Material S4) has shown that the MDO1 structure in its *apo* form is not able to accommodate ligands. We thus relied on a three-dimensional structure where the binding site had been made to mimic the close homolog MDO2 bound to ADP-ribose. For this purpose, the human MDO1 protein was remodeled using the online webserver SWISS-MODEL.[30] The best model was built based on human MOD2 structure (PDB file: 4IQY) with 56.33% sequence identity. Before docking, the modeled human MDO1 protein structure (PDB code: 2X47) was handled with protein

preparation wizard (Schrödinger Ltd). The protein was assigned atom type, side chain protonation states, and hydrogen network using PROPKA at pH 7.0. The OPLS3 force field was applied to minimize the protein. Glide receptor grids were generated by defining a 10 Å box localized at the centroid of ADP-ribose defined using the superimposed ADP-ribose-MDO2 protein complex. H-bond constraint to Asp160 was applied during grids generation, which will guarantee adenine stacked by Phe306 and Val183.

All ligands listed in table **1** and **2** were drawn in 2D Sketcher, then converted into 3D. The possible ionization states of ligand were generated at pH 7.0+/-2.0 by LigPrep Suite.

The molecular docking experiments were carried out with Maestro suite Glide and the Glide XP protocol.[31] Default parameters were used. Van der Waals scale factor was 1.0, and partial charge cutoff was 0.25. For each ligand, the maximum 10 docking poses were output. The one with lowest energy was suggested as most likely binding pose.

Acknowledgments

Supplementary Material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.201X. XX.XXX. These data include the synthesis procedures of compounds **5-7**, **11-15**, **19, 21-23**, **27**, **36-39**, and the NMR spectra of compounds as well as the abbreviations. Figures: Figure S1, Dose-response measurements; Figure S2, three-dimensional fold and structural differences between MDO2 and MDO1; Figure S3, purification of recombinant MDO1; Figure S4, docking to *apo* MDO1.

References and notes

1.   Till S, Ladurner AG. Sensing NAD metabolites through macro domains. *Front Biosci.* 2009;14:3246-3258.
2.   Han W, Li X, Fu X. The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutation Research/Reviews in Mutation Research.* 2011;727(3):86-103.
3.   Feijs KL, Forst AH, Verheugd P, Lüscher B. Macrodomain-containing proteins: regulating new intracellular functions of mono (ADP-ribosyl) ation. *Nature Reviews Molecular Cell Biology.* 2013;14(7):443-451.
4.   Malet H, Coutard B, Jamal S, et al. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *Journal of virology.* 2009;83(13):6534-6545.
5.   Rosenthal F, Feijs KL, Frugier E, et al. Macrodomain-containing proteins are new mono-ADP-ribosylhydrolases. *Nature structural & molecular biology.* 2013;20(4):502-507.
6.   Chen D, Vollmar M, Rossi MN, et al. Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J Biol Chem.* 2011;286(15):13261-13271.
7.   Jankevicius G, Hassler M, Golia B, et al. A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nature structural & molecular biology.* 2013;20(4):508-514.
8.   Tian L, Wu Z, Zhao Y, et al. Differential induction of LRP16 by liganded and unliganded estrogen receptor α in SKOV3 ovarian carcinoma cells. *J Endocrinol.* 2009;202(1):167-177.
9.   Zhao P, Lu Y, Han W. Clinicopathological significance and prognostic value of leukemia‐related protein 16 expression in invasive ductal breast carcinoma. *Cancer science.* 2010;101(10):2262-2268.
10.   Imagama S, Abe A, Suzuki M, et al. LRP16 is fused to RUNX1 in monocytic leukemia cell line with t (11; 21)(q13; q22). *European journal of haematology.* 2007;79(1):25-31.
11.   Yang J, Zhao Y, Wu Z, et al. The single-macro domain protein LRP16 is an essential cofactor of androgen receptor. *Endocrine-related cancer.* 2009;16(1):139-153.
12.   Vyas S, Chang P. New PARP targets for cancer therapy. *Nature Reviews Cancer.* 2014;14(7):502-509.
13.   Karras GI, Kustatscher G, Buhecha HR, et al. The macro domain is an ADP‐ribose binding module. *The EMBO journal.* 2005;24(11):1911-1920.
14.   Yao Z-J, King CR, Cao T, et al. Potent inhibition of Grb2 SH2 domain binding by non-phosphate-containing ligands. *J Med Chem.* 1999;42(1):25-35.
15.   Wagner GK, Pesnot T, Field RA. A survey of chemical methods for sugar-nucleotide synthesis. *Natural product reports.* 2009;26(9):1172-1194.
16.   Seio K, Miyashita T, Sato K, Sekine M. Synthesis and properties of new nucleotide analogues possessing squaramide moieties as new phosphate isosters. *Eur J Org Chem.* 2005;2005(24):5163-5170.
17.   Storer RI, Aciro C, Jones LH. Squaramides: physical properties, synthesis and applications. *Chem Soc Rev.* 2011;40(5):2330-2346.
18.   Zhang Y, Borrel A, Ghemtio L, et al. Structural Isosteres of Phosphate Groups in the Protein Data Bank. *Journal of Chemical Information and Modeling.* 2017;57(3):499-516.
19.   Mutisya D, Selvam C, Lunstad BD, et al. Amides are excellent mimics of phosphate internucleoside linkages and are well tolerated in short interfering RNAs. *Nucleic Acids Res.* 2014;42(10):6542-6551.
20.   Elliott TS, Slowey A, Ye Y, Conway SJ. The use of phosphate bioisosteres in medicinal chemistry and chemical biology. *MedChemComm.* 2012;3(7):735-751.
21.   Lassalas P, Gay B, Lasfargeas C, et al. Structure property relationships of carboxylic acid isosteres. *J Med Chem.* 2016;59(7):3183-3203.
22.   Matthew A. Tetrazolone as an acid bioisostere: application to marketed drugs containing a carboxylic acid. *Organic & biomolecular chemistry.* 2016;14(39):9343-9347.
23.   Niewiadomski S, Beebeejaun Z, Denton H, Smith TK, Morris RJ, Wagner GK. Rationally designed squaryldiamides–a novel class of sugar-nucleotide mimics? *Organic & biomolecular chemistry.* 2010;8(15):3488-3499.
24.   Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols.* 2007;2(9):2212-2221.
25.   Neuvonen M, Ahola T. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. *J Mol Biol.* 2009;385(1):212-225.
26.   Haikarainen T, Maksimainen MM, Obaji E, Lehtiö L. Development of an Inhibitor Screening Assay for Mono-ADP-Ribosyl Hydrolyzing Macrodomains Using AlphaScreen Technology. *SLAS DISCOVERY: Advancing Life Sciences R&D.* 2017:1-9.
27.   After the western blot analysis of the compounds had been performed, we recognized that compound **35** (the Fmoc-deprotected derivative of compound **33**) contained small amount of impurities. Purity determinations showed impure compound even after extensive purifications of the compound, and therefore, we made a late stage decision to remove it from the manuscript.
28.   Ikeuchi H, Meyer ME, Ding Y, Hiratake J, Richards NG. A critical electrostatic interaction mediates inhibitor recognition by human asparagine synthetase. *Biorg Med Chem.* 2009;17(18):6641-6650.
29.   Venkannagari H, Verheugd P, Koivunen J, et al. Small-Molecule Chemical Probe Rescues Cells from Mono-ADP-Ribosyltransferase ARTD10/PARP10-Induced Apoptosis and Sensitizes Cancer Cells to DNA Damage. *Cell chemical biology.* 2016;23(10):1251-1260.
30.   Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014;42(W1):W252-W258.
31.   Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein– ligand complexes. *J Med Chem.* 2006;49(21):6177-6196.