PNAS
www.pnas.org

Supplementary Information for

ADP-Ribose and Analogues bound to the DeMARylating Macrodomain from the Bat Coronavirus HKU4

Robert G. Hammond, Norbert Schormann, Robert Lyle McPherson, Anthony K. L. Leung, Champion C.S. Deivanayagam, and Margaret A. Johnson

Corresponding author: Margaret A. Johnson
Email: margaret.a.johnson2@gmail.com

This PDF file includes:

Supplementary Methods

Protein Expression and Purification

The genetic sequence of the HKU4 macrodomain encoding the residues 307-478 of HKU4 nsp3 was amplified using PCR and cloned into the pET15b-TEV NESG vector (DNASU, clone #EvNO00336943) (1). Protein overexpression was achieved by transforming the plasmid into *E. coli* BL21(DE3) *pLysY* strain (New England BioLabs, C3010I). A colony was chosen and cultured overnight at 37°C and diluted into 700 mL of LB medium at a ratio of 1:20. The large culture was grown to an optical density (OD_{600}) of 0.65 and induced with 1 mM isopropyl β-D-1thiogalactopyranoside (IPTG) at 18 °C for 18 h. Culture was harvested via centrifugation for 5 min intervals at 6,000 x g and wet cell paste was stored at -80°C. Cell pellet was resuspended in lysis buffer (20 mM Tris pH 8.0, 300 mM NaCl, 20 mM Imidazole, 3 mM DTT, 1 Pierce Protease Inhibitor tablets, EDTA-Free (Thermo Fisher, A32965)/ 50 mL of buffer, 0.5 mL of Triton X-100/ 50 mL of buffer) and lysed via sonication on ice for 3 minutes (Pulse On: 15 sec, Pulse Off: 45 Sec). Lysed culture was clarified by centrifugation for 30 min at 20,000 x g. Supernatant was removed and filtered with .22 µm filter unit before chromatography. Supernatant was loaded onto a 5 mL HisTrap FF Ni-NTA column (GE Healthcare, 17525501) (Equilibration Buffer: 20 mM Tris pH 8.0, 300 mM NaCl, 20 mM Imidazole, 3 mM DTT) and eluted along a 10 column volume gradient (Elution Buffer: 20 mM Tris pH 8.0, 300 mM NaCl, 500 mM Imidazole, 3 mM DTT). Protein fractions were collected and dialyzed against equilibration buffer to remove the high imidazole concentration. The protein sample was further dialyzed against urea buffer (20 mM Tris pH 8.0, 1 M NaCl, 20 mM imidazole, 3 mM DTT, and 4 M urea) and back to equilibration buffer to remove bound nucleotide to the protein. 25 mg of purified TEV protease (Invitrogen, 12575015) was added to the protein and loaded on the 5 mL Ni-NTA HisTrap FF column again. Protein was collected in the flow through and concentrated down to 10 mL. Purification was completed with size exclusion chromatography by using a HiLoad 26/600 Superdex 75pg (GE Healthcare, 28989334) column (Gel Filtration Buffer: 20 mM Tris pH 8.0, 300 mM NaCl, 0.5 mM TCEP). All purification steps were conducted at 4 °C.

Site-Directed Mutagenesis

Macrodomain mutants were generated by mutating three amino acids (Ala326Ile, Gly351Leu, and Ile434Ala) using the QuikChange II site-directed mutagenesis kit (Agilent Technologies, 200522). Mutant plasmids were transformed, expressed, and purified under the same conditions as wild type protein.

Isothermal Titration Calorimetry

All ITC experiments were conducted on the MicroCal PEAQ ITC instrument. Protein and ligand samples were extensively dialyzed in ITC buffer before each experimental run (Binding Buffer: 20 mM Tris pH 8.0, 50 mM NaCl, 0.5 mM TCEP). Each run used 2 µL of the titrant for a total 40 injections at 25 °C with a spacing of 150 sec at a reference power of 10 kcal/mol. All titrations used 350 µL of 250 µM protein and 70 µL of 5 mM ligand as the titrant concentration. A326I, G351L, and I434A mutant titrations were recorded using the same parameters. All titrations were run in triplicate. Buffer-only titrations were conducted before and after each ITC run to determine if the conditions were similar between each sample, detect any potential contamination or sample-dependent differences, and determine whether buffer to buffer dilution was consistent between each run. None of the buffer dilution steps produced significant enthalpies, and ligand into buffer values were subtracted from each respective protein-ligand titration.

In silico peptide binding analysis with Autodock Vina

ADP-ribose crystal structures without bound ADP-ribose were used as target proteins for Autodock Vina docking (2). Modified human PARP-1 at position E488 and E491 were saved as .pdb files and converted to pdbqt format as input for the docking run. The active site boundaries included

the entire protein surface. An 18-residue sequence of human PARP-1 (487-505; sequence ^{487}A**E**PV**E**VVAPRGKSGAALS^{505}) was selected.

Demodification assay

To prepare the MARylated substrate, ^{32}P-labeled NAD$^+$ (Perkin Elmer BLU023X) was incubated with PARP10^{CD} according to the method of McPherson *et al.* (3, 4). PARP-10 catalytic domain was auto-MARylated by incubating with 15 µCi ^{32}P NAD$^+$ for 30 min at 30 °C in automodification buffer (20 mM Tris·HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 20 mM β-mercaptoethanol). To remove additional ^{32}P NAD$^+$ from the sample, modified PARP10 catalytic domain was desalted using Micro Bio-Spin column (Bio-Rad) in demodification buffer (25 mM Tris·HCl pH 7.0, 200 mM NaCl, 20 mM β-mercaptoethanol) (3, 4). 1 µL of HKU4 macrodomain WT and mutants (1 mg/mL) were added to equimolar amounts of modified PARP10 substrate for 1 h at 37 °C. 5 µL of 4X running buffer were added to each reaction mixture and inactivated by heating to 95 °C. 20 µL of sample mixtures were loaded onto 10 well 15% SDS-PAGE gels (Invitrogen), analyzed after Coomassie stain, and visualized via autoradiography. Human MacroD2 and buffer only reaction mixtures were conducted as controls. All experiments were carried out in triplicate.

Structure Determination and Data Collection

Co-Crystallization: In the initial experiments the protein of the macrodomain (in 20mM Tris pH 8.0, 150mM NaCl, 5mM DTT) was incubated at 10 mg/ml (0.54 mM) at a 20:1 molar (ligand:protein) ratio with three ligands (ADPR (Sigma Aldrich A0752), NAD$^+$ (Sigma Aldrich N0632), ADP-Glc (Sigma Aldrich A0627). Crystallization conditions using the hanging drop technique varied between 20 and 25% PEG3350 in Hepes buffer at pH 7.0 and 7.5, and different protein solution/reservoir ratios in the drop. Subsequently, to decrease precipitation and to reduce nucleation, we used a lower protein concentration (6.6 mg/ml or 0.36 mM) and a lower ligand:protein ratio (10:1). The three protein-ligand complexes all crystallized in space group P1 (a=34 Å, b=42 Å, c=60 Å; α=73º, β=88º, γ=89º) with two molecules in the asymmetric unit.

Data Collection: Data for the NAD$^+$ complex were collected on our home source equipped with a Rigaku MicroMax 007HF X-ray generator and a DECTRIS Pilatus 200k silicon pixel detector. Data for the ADPR and ADPG complexes were collected at the Advanced Photon Source (APS) on the SER-CAT 22ID beam line at a wavelength of 1.0 Å using a Mar300HS CCD detector. For data processing we used a combination of HKL 3000, XDS, Xia2, and Mosfilm.

Refinement: Initial phasing using molecular replacement (Phaser in CCP4) utilized the protein sequence and the structure of the MERS-CoV macrodomain (PDB code 5DUS) as search model (5-7). This structure was refined and used subsequently for all current data sets and structures. Refinement was carried out by a combination of Refmac in CCP4 and Phenix (5, 7). Mogul ligand restraints from the Grae web server (http://grade.globalphasing.org) were used during refinement (8). All occupancies (ligands included) were kept at 1.0 and not refined. For model building we used Coot (9). Figures were created with PyMol. Electrostatic surface representations were generated by the PDB2PQR server and illustrated using the APBS Pymol plugin (10, 11). Validation was carried out employing Phenix, QualityCheck (https://smb.slac.stanford.edu/jcsg/QC/), and the PDB Validation Suite (https://validate-rcsb-2.wwpdb.org/) (12-14). The validation was based on quality scores for geometry, clashes, accuracy of the 3D fold, quality of the data, and fit of the structural model to the data (electron density).

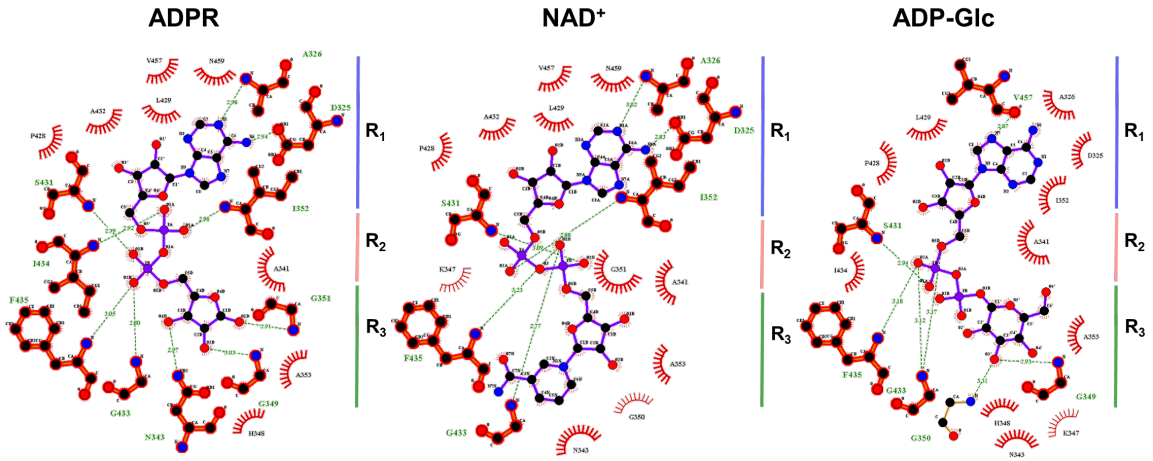
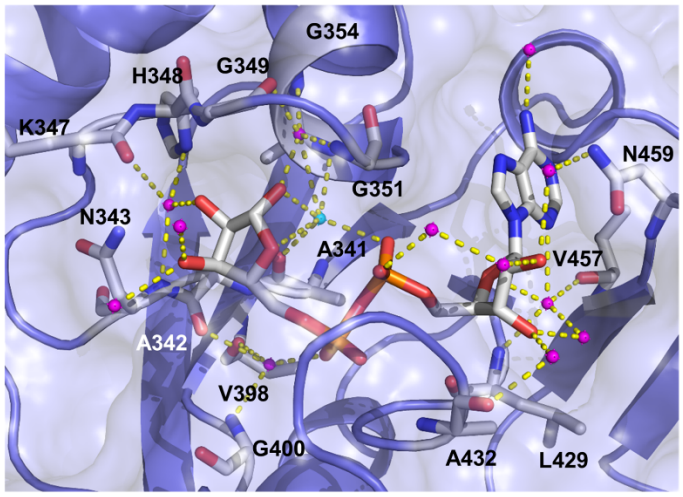
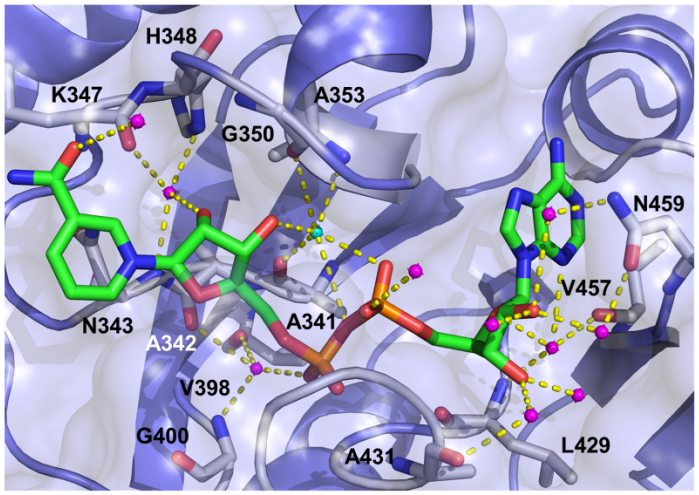


Figure S1. Description of interactions present in the HKU4 macrodomain binding site. Ligplot+ representation of each protein-ligand complex: ADP-ribose (left), NAD^+ (center), and ADP-Glc (right). Each ligand and amino acid is displayed as ball-and-stick models by atom type: carbon (black), nitrogen (blue), and oxygen (red). Hydrogen bonds are displayed as green dashes labeled with their corresponding bond lengths (Å). Hydrophobic interactions are labeled as red eyelashes.

A


G354
H348
G349
K347
N459
G351
N343
A341
V457
A342
V398
G400
A432
L429

B


H348
K347
A353
G350
N459
V457
N343
A341
A342
V398
G400
A431
L429

C

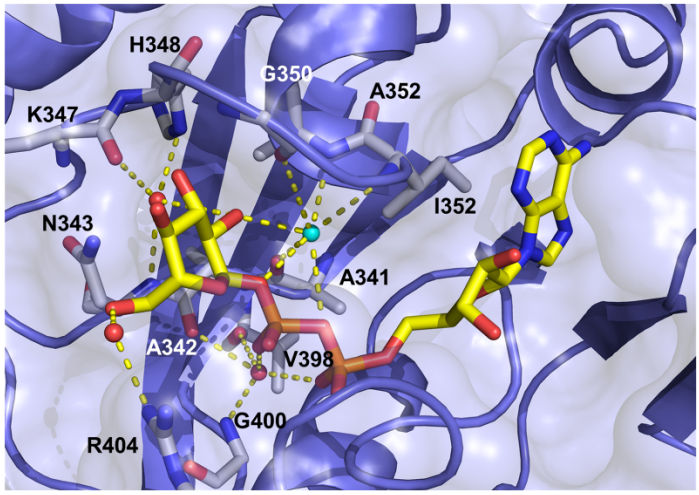

H348
G350
A352
K347
I352
N343
A341
A342
V398
G400
R404

Figure S2. Structures of ligands in the HKU4 macrodomain binding site. Binding cavity displayed with semitransparent surface overlaying protein backbone shown in cartoon display. Waters are shown as spheres (pink) and the hydrogen bond interactions are shown as dashes (yellow). Residues that form water-mediated interactions are shown as sticks and labeled. Ligand and amino acid are displayed as red (oxygen), orange (phosphorus), and blue (nitrogen). A) Complex of bound ADPR. Carbon atoms are shown as white. B) Complex of bound NAD^+. Carbon atoms are shown as green. C) Complex of ADP-Glc. Carbon atoms are shown as yellow.

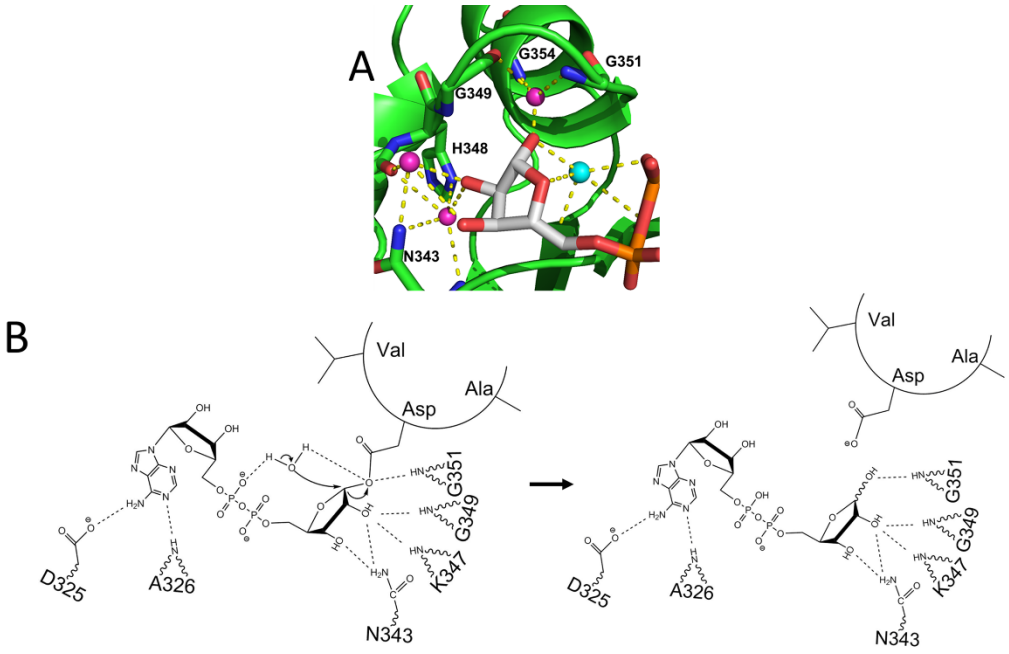
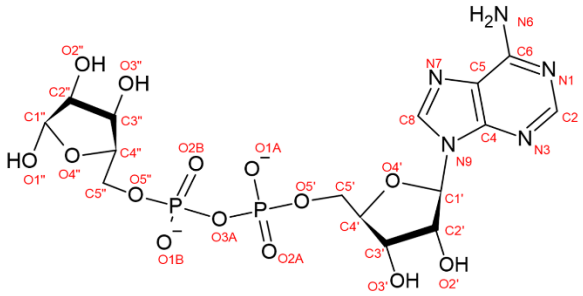
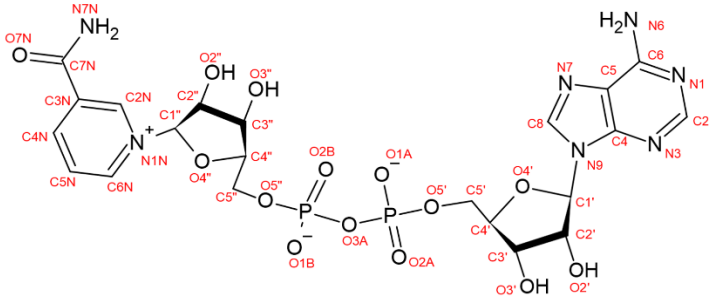


Figure S3. A) Close view of distal ribose from ADP-ribose. Water molecules involved in water-mediated interactions between the protein and ligand are shown as pink spheres. The proposed catalytic water is colored aqua. Interacting residues are labeled and hydrogen bond interactions are shown as yellow dashes. B) A proposed deMARylation mechanism for the HKU4 macrodomain. The sphere with attached MARylated aspartate represents an example substrate protein.

A



B



C

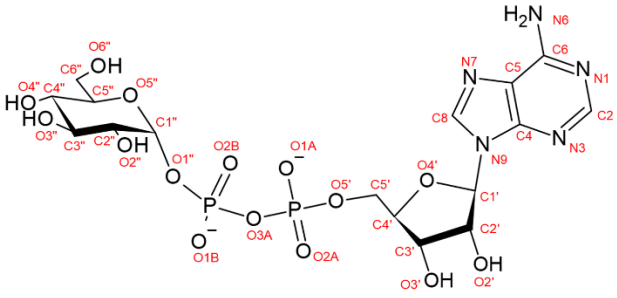


Figure S4. Structural formulas of each ligand used in complex with the HKU4 macrodomain. Each atom is labeled in red. A) ADP-ribose; B) NAD^+; C) ADP-glucose

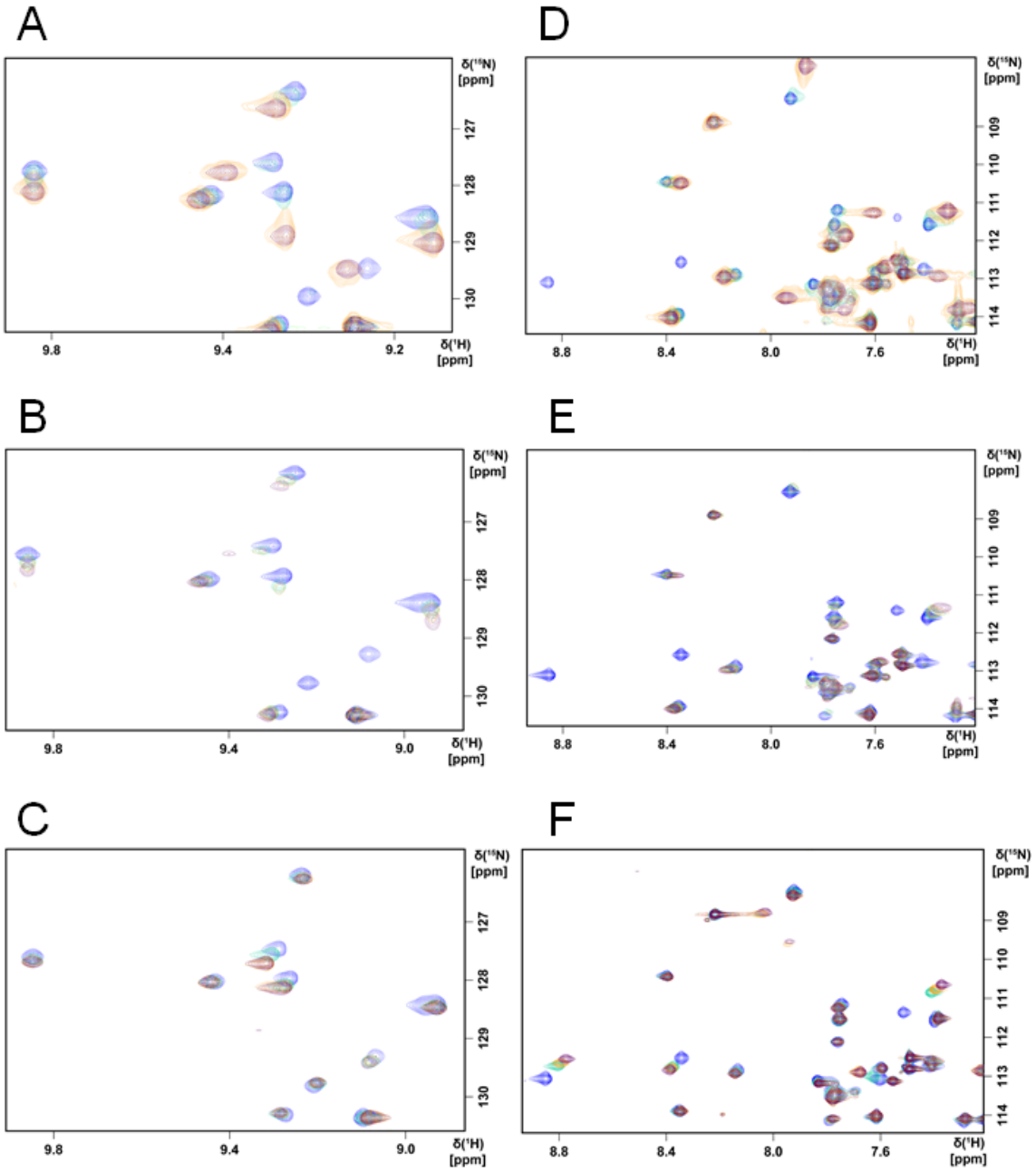


Figure S5. 2D [^{15}N,^{1}H]-HSQC spectra overlay of HKU4 macrodomain titration with ADP-ribose (A and D), NAD^{+} (B and E), and ADP-glucose (C and F). ADP-ribose titration: 0 µM (blue), 250 µM (green), 750 µM (yellow), 1.25 mM (maroon). NAD^{+} and ADP-glucose titration: 0 µM (blue), 8 mM (green), 10 mM (yellow), 1.25 mM (maroon). The HKU4 macrodomain concentration was 500 µM in all experiments.

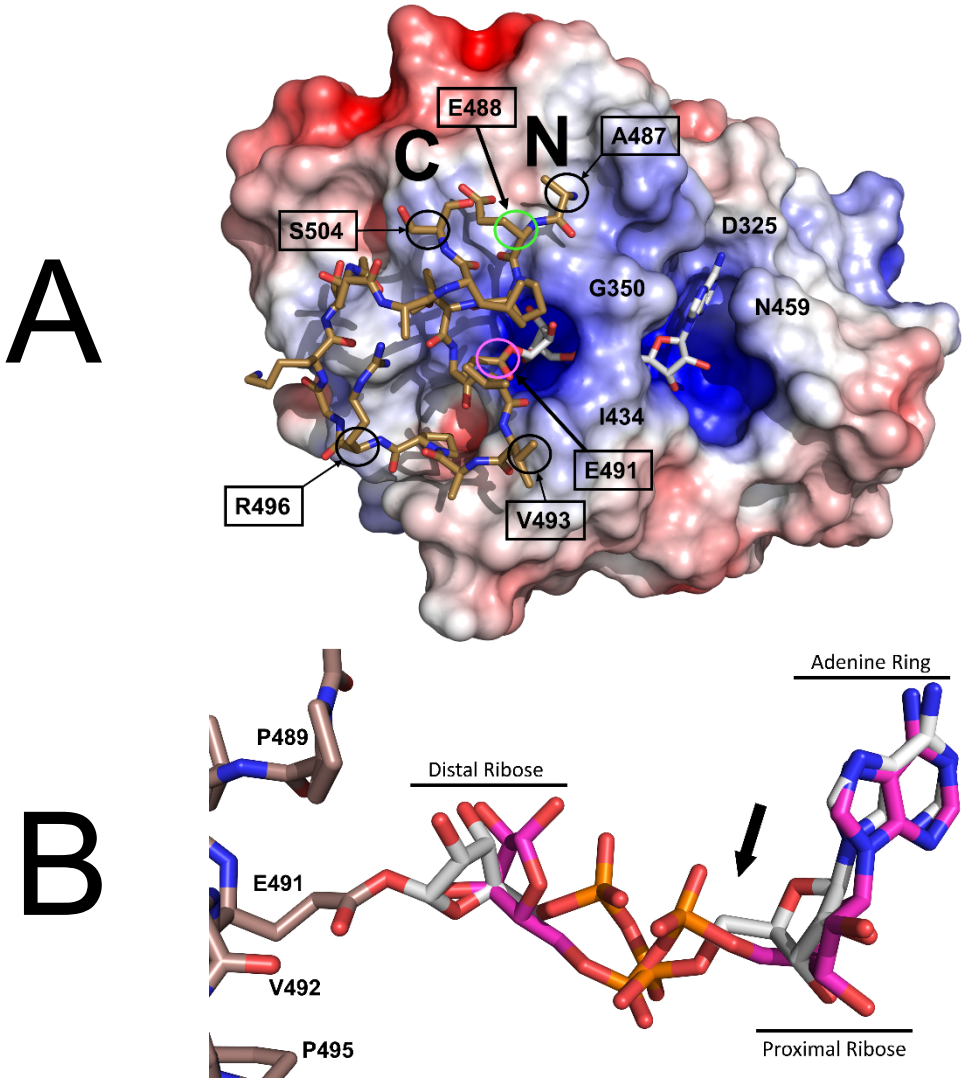


Figure S6. Docking output of HKU4 macrodomain in complex with ADP-ribose. A) Electrostatic surface representation of HKU4 macrodomain. Acidic regions are shown in red, basic regions are shown in blue, and neutral regions are shown in white. Select residues from the macrodomain (unboxed) and from the PARP-1 peptide (boxed) are indicated. The N-terminus (N) and C-terminus (C) are labeled. The two MARylation sites, E488 (green) and E491 (pink) are indicated. B) Overlay of bound ADP-ribose from crystal structure and modified E491 PARP-1 peptide. The different O4'-C4-C5-O5' torsion angles are indicated by an arrow (-89.5° in the ADPR complex and 177.7° in the docked substrate). The adenine ring, proximal ribose, and distal ribose are labeled. Select residues of the PARP-1 peptide are labeled.

Table S1. Ligand B-factors of the macrodomain complexes.

ADPR		NAD^+		ADP-Glucose	
Atom	B factor	Atom	B factor	Atom	B factor
N1	10.23	PA	17.89	PB	39.9
C2	9.75	O1A	17.65	O1B	41.4
N3	9.56	O2A	16.79	O2B	40.87
C4	9.07	O5B	18.04	O3B	40.11
C5	11.09	C5B	17.17	PA	26.48
C6	11.15	C4B	16.95	O1A	27.3
N6	12.98	O4B	17.08	O2A	23.89
N7	11.41	C3B	16.6	O3A	34.67
C8	10.64	O3B	17.03	O5D	33.14
N9	8.53	C2B	16.69	C5D	38.16
C1'	9.42	O2B	16.16	C4D	42.13
C2'	9.8	C1B	16.82	O4D	45.5
O2'	13.03	N9A	16.84	C3D	44.21
C3'	9.13	C8A	17.04	O3D	41.77
O3'	11.34	N7A	17.2	C2D	46.18
O4'	9.22	C5A	17.05	O2D	48.01
C4'	9.15	C6A	17.31	C1D	47.57
C5'	8.58	N6A	17.51	N9	49.79
O5'	7.44	N1A	16.51	C8	51.53
PA	6.63	C2A	17.15	N7	52.26
O1A	7.11	N3A	17.03	C5	52.55
O2A	7.64	C4A	16.81	C6	52.47
O3A	6.8	O3	16.8	N6	52.75

PB	5.28	PN	16.69	N1	52.66
O1B	5.37	O1N	15.07	C2	52.1
O2B	5.59	O2N	15.3	N3	52.37
O5D	5.94	O5D	19.32	C4	51.69
C5D	6.92	C5D	22.27	C1'	43.92
O4D	8.36	C4D	25.25	C2'	43.98
O1D	9.14	O4D	28.01	C3'	44.71
C1D	8.76	C3D	27.48	C4'	44.68
O2D	8.73	O3D	27.63	C5'	43.91
C2D	8.1	C2D	29.21	C6'	42.3
O3D	8.94	O2D	30.02	O2'	46.52
C3D	7.68	C1D	29.72	O3'	46.41
C4D	7.15	N1N	33.39	O4'	43.92
		C2N	36.16	O5'	44.59
		C3N	37.67	O6'	41.48
		C7N	39.56		
		O7N	40.79		
		N7N	40.29		
		C4N	36.94		
		C5N	35.77		
		C6N	34.65		
Mean	8.8		23.0		43.9
N1	12.28	PA	22.63	PB	37.57
C2	12.29	O1A	22.24	O1B	39.54
N3	11.34	O2A	22.39	O2B	39.51
C4	11.34	O5B	22.26	O3B	39.3

C5	12.29	C5B	21.44	PA	26.63
C6	12.79	C4B	20.48	O1A	26.87
N6	15.29	O4B	20.41	O2A	24.61
N7	13.15	C3B	20.15	O3A	34.48
C8	12.24	O3B	19.43	O5D	34.28
N9	10.9	C2B	20.5	C5D	40.53
C1'	11.2	O2B	21.54	C4D	45.55
C2'	11.91	C1B	20.3	O4D	49.11
O2'	14.58	N9A	20.39	C3D	48.5
C3'	11.41	C8A	20.38	O3D	49.95
O3'	13.18	N7A	20.97	C2D	50.61
O4'	10.82	C5A	20.49	O2D	52.09
C4'	9.93	C6A	20.69	C1D	51.93
C5'	10.73	N6A	20.29	N9	54.78
O5'	9.58	N1A	20.26	C8	56.78
PA	8.4	C2A	20.1	N7	57.5
O1A	9.76	N3A	20.15	C5	58.32
O2A	10.05	C4A	20.61	C6	59.12
O3A	7.99	O3	20.03	N6	58.74
PB	6.7	PN	19.3	N1	59.19
O1B	6.43	O1N	17.76	C2	58.83
O2B	7.3	O2N	17.99	N3	59.13
O5D	7.53	O5D	22.66	C4	57.45
C5D	8.1	C5D	28.58	C1'	39.15
O4D	10.37	C4D	32.08	C2'	39.85
O1D	12.18	O4D	36.71	C3'	39.7
C1D	11.12	C3D	34.72	C4'	38.7

O2D	11.32	O3D	36.74	C5'	38.25
C2D	11.06	C2D	36.75	C6'	37.04
O3D	10.92	O2D	37.23	O2'	42.79
C3D	10.36	C1D	38.42	O3'	43.6
C4D	9	N1N	41.55	O4'	39.68
		C2N	43.42	O5'	38.73
		C3N	44.32	O6'	31.97
		C7N	45.78		
		O7N	46.64		
		N7N	45.85		
		C4N	44		
		C5N	42.87		
		C6N	42.03		
Mean	10.7		28.0		44.8

SI References

1. T. B. Acton *et al.*, Robotic cloning and Protein Production Platform of the Northeast Structural Genomics Consortium. *Methods Enzymol* **394**, 210-243 (2005).
2. O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-461 (2010).
3. R. L. McPherson *et al.*, ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence. *Proc Natl Acad Sci U S A* **114**, 1666-1671 (2017).
4. L. Palazzo *et al.*, ENPP1 processes protein ADP-ribosylation in vitro. *Febs j* **283**, 3371-3388 (2016).
5. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
6. C. C. Cho, M. H. Lin, C. Y. Chuang, C. H. Hsu, Macro Domain from Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Is an Efficient ADP-ribose Binding Module: CRYSTAL STRUCTURE AND BIOCHEMICAL STUDIES. *J Biol Chem* **291**, 4894-4902 (2016).
7. W. F. Vranken *et al.*, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005).
8. I. J. Bruno *et al.*, Retrieval of crystallographically-derived molecular geometry information. *J Chem Inf Comput Sci* **44**, 2133-2144 (2004).
9. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
10. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665-667 (2004).
11. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041 (2001).
12. P. V. Afonine *et al.*, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367 (2012).
13. S. Gore *et al.*, Validation of Structures in the Protein Data Bank. *Structure* **25**, 1916-1927 (2017).
14. M. A. Elsliger *et al.*, The JCSG high-throughput structural biology pipeline. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**, 1137-1142 (2010).