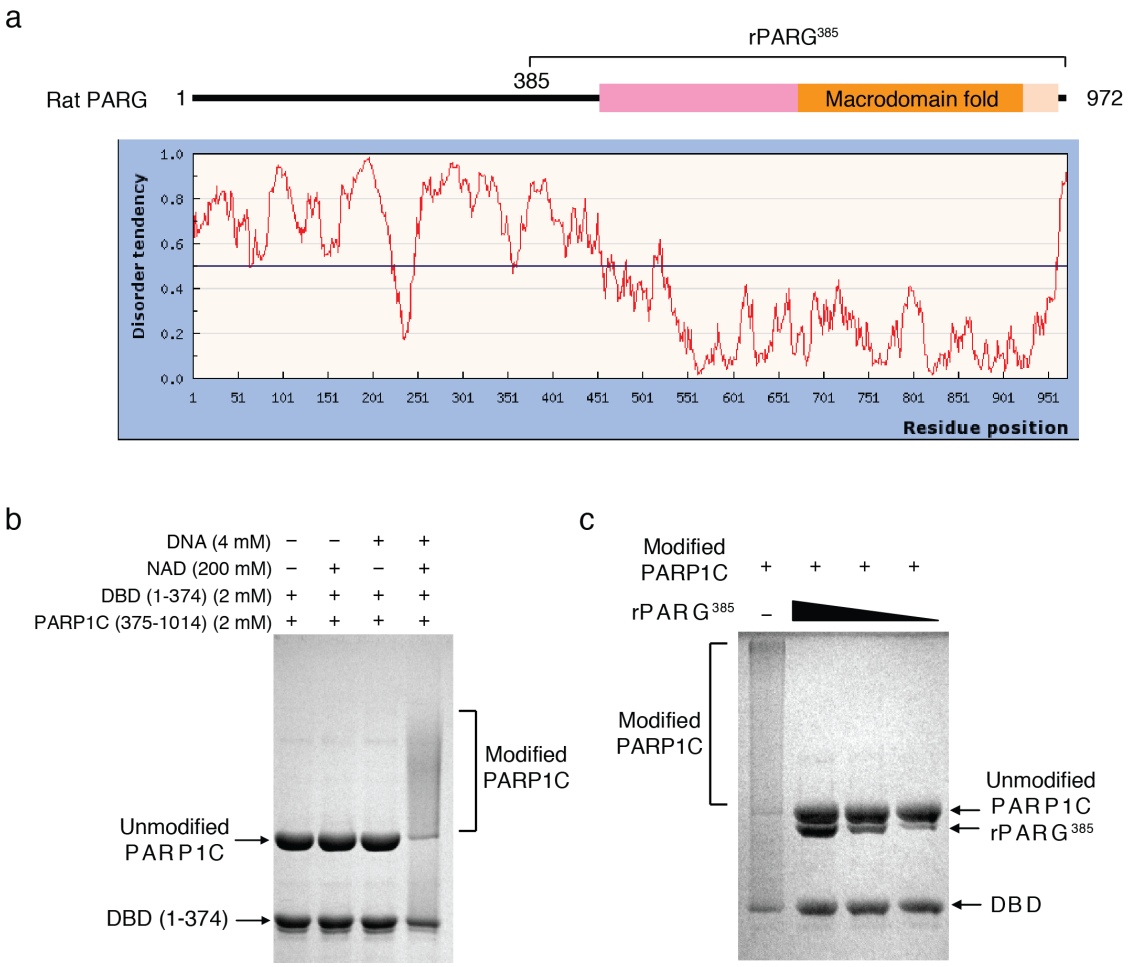
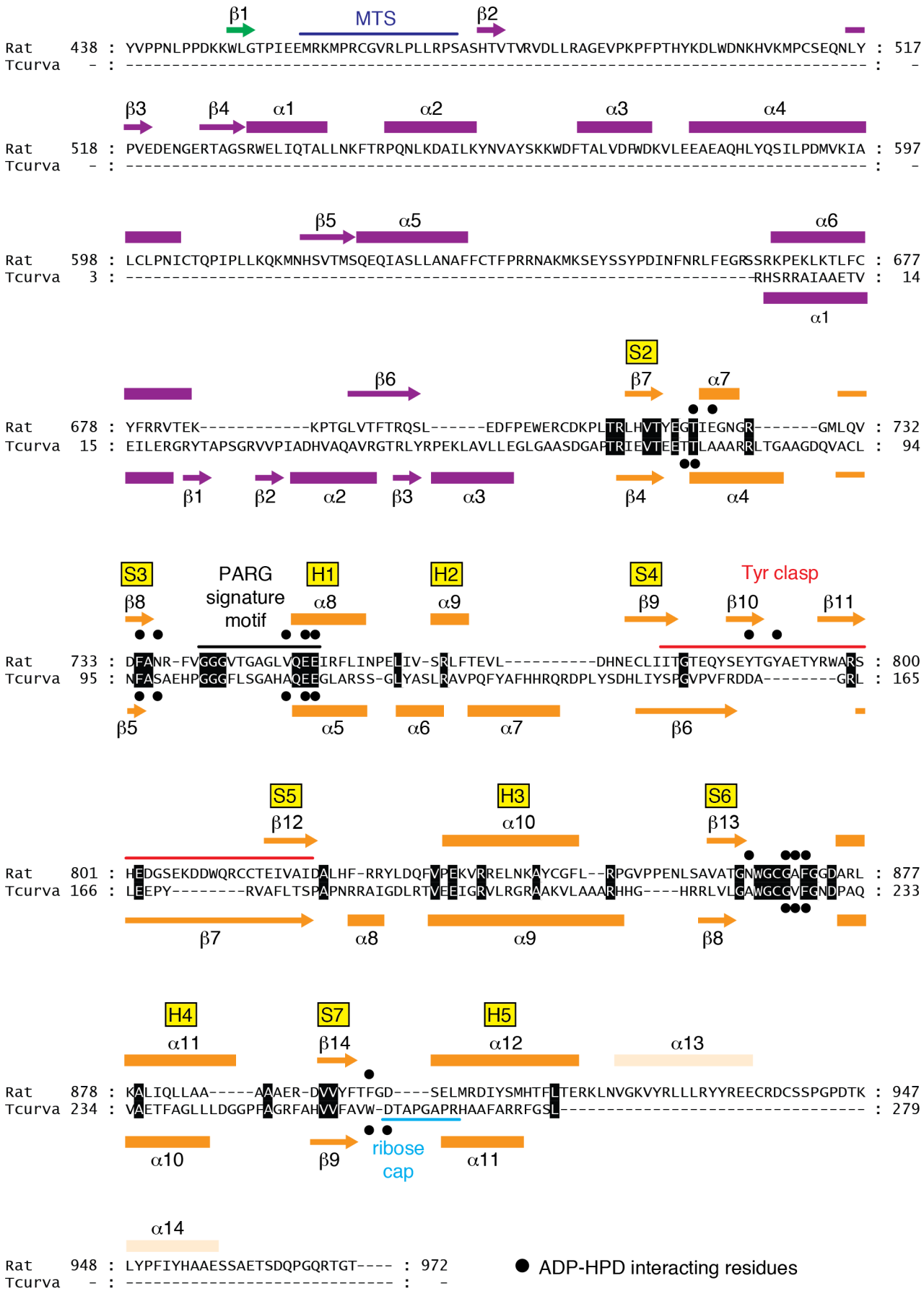
Supplementary information

Structure of mammalian poly (ADP-ribose) glycohydrolase reveals a flexible tyrosine clasp as a unique substrate-binding element

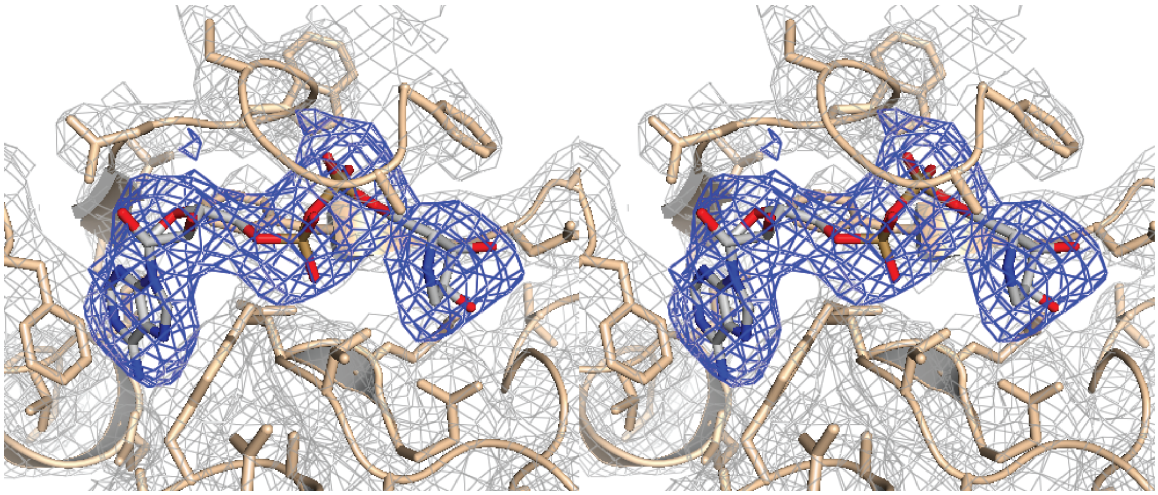
In-Kwon Kim, James R. Kiefer, Chris M. Ho, Roderick A. Stegeman, Scott Classen, John A. Tainer, and Tom Ellenberger



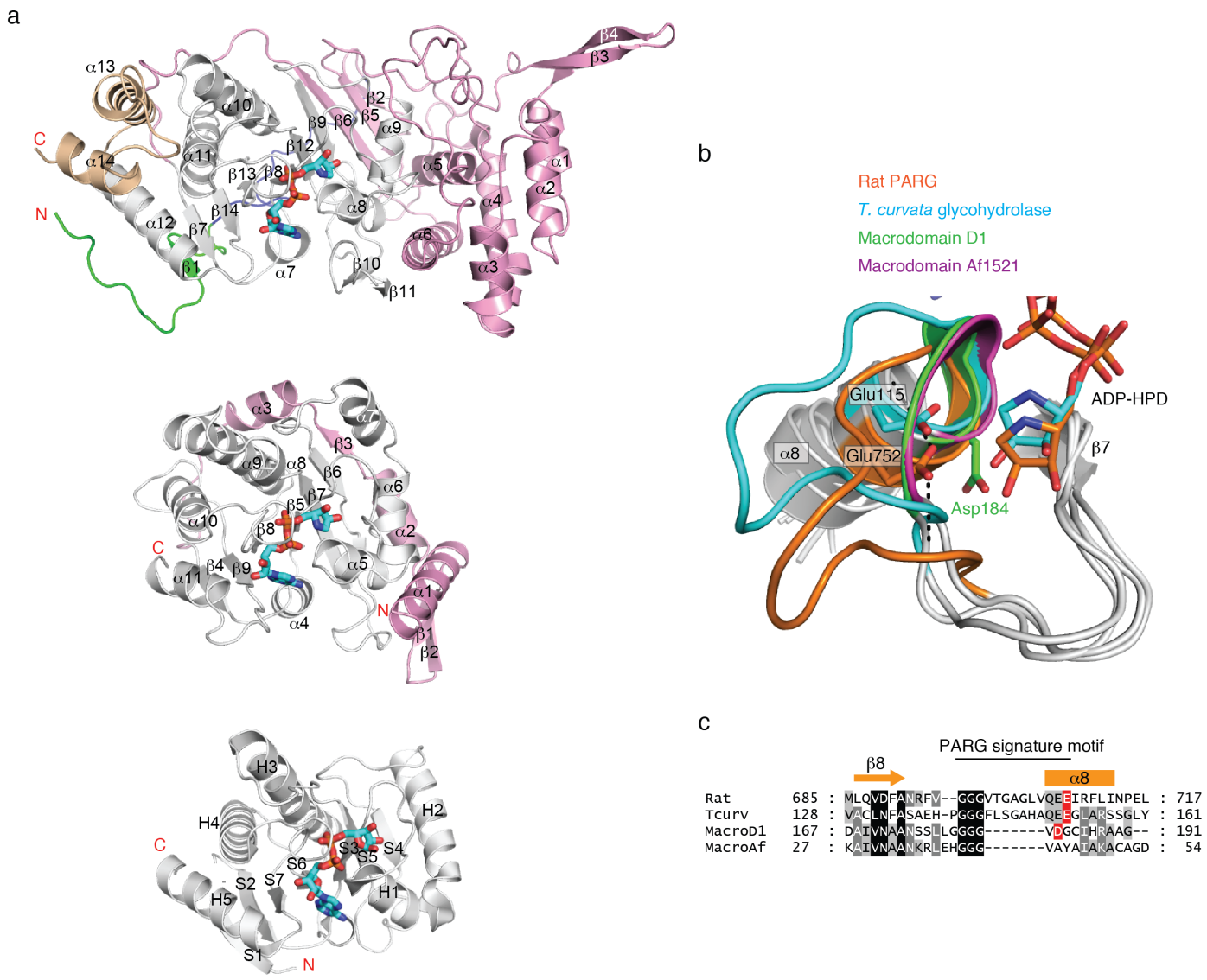
Supplementary Figure 1 rPARG^{385} is catalytically active. (**a**) The order-disorder prediction from IUPRED[1] indicates disorder in the N-terminal A-domain of PARG (residues 1–456)[2], whereas the rPARG^{385} protein that was crystallized (**Figure 1**) was predicted to be well-ordered. (**b**) PARP1 activation assay. A C-terminal fragment spanning the automodification, WGR, and catalytic domains of PARP1 (PARP1C, residues 375–1014) is automodified when incubated with the PARP1 DNA-binding domain (DBD, residues 1–374) and DNA in the presence of NAD^+. The PARylated PARP1 protein runs as a smear on SDS-PAGE. (**c**) The crystallized rPARG^{385} catalytic domain reverses the PARylation of PARP1, causing the protein to migrate as a discrete band on SDS-PAGE.


β1
MTS
β2
Rat 438 : YVPPNLPPDKKWLGTPIEEMRKMPRCGVRLPLLRPSASHTVTVRVDLLRAGEVPKPFPTHYKDLWDNKHVKMPCSEQNLY : 517
Tcurva - : ---------- :
β3
β4
α1
α2
α3
α4
Rat 518 : PVEDENGERTAGSRWELIQTALLNKFTRPQNLKDAILKYNVAYSKKWDFTALVDFWDKVLEEAEAQHLYQSILPDMVKIA : 597
Tcurva - : ---------- :
β5
α5
α6
Rat 598 : LCLPNICTQPIPLLKQKMNHSVTMSQEQIASLLANAFFCTFPRRNAKMKSEYSSYPDINFNRLFEGRSSRKPEKLKTLFC : 677
Tcurva 3 : ----------RHSRRAIAAETV : 14
α1
S2
β6
β7
α7
Rat 678 : YFRRVTEK----------KPTGLVTFTRQSL------EDFPEWERCDKPLTRLHVTYEGTIEGNGR-------GMLQV : 732
Tcurva 15 : EILERGRYTAPSGRVVPIADHVAQAVRGTRLYRPEKLAVLLEGLGAASDGAPTRIEVTEETTLAAARRLTGAAGDQVACL : 94
β1
β2
α2
β3
α3
β4
α4
S3
PARG signature motif
H1
H2
S4
Tyr clasp
β8
α8
α9
β9
β10
β11
Rat 733 : DFANR-FVGGGVTGAGLVQEEIRFLINPELIV-SRLFTEVL----------DHNECLIITGTEQYSEYTGYAETYRWARS : 800
Tcurva 95 : NFASAEHPGGGFLSGAHAQEEGLARSS-GLYASLRAVPQFYAFHHRQRDPLYSDHLIYSPGVPVFRDDA--------GRL : 165
β5
α5
α6
α7
β6
S5
H3
S6
β12
α10
β13
Rat 801 : HEDGSEKDDWQRCCTEIVAIDALHF-RRYLDQFVPEKVRRELNKAYCGFL--RPGVPPENLSAVATGNWGCCAFGGDARL : 877
Tcurva 166 : LEEPY--------RVAFLTSPAPNRRAIGDLRTVEEIGRVLRGRAAKVLAAARHHG----HRRLVLGAWGCCVFGNDPAQ : 233
β7
α8
α9
β8
H4
S7
H5
α11
β14
α12
α13
Rat 878 : KALIQLLAA-----AAAER-DVVYFTFGD----SELMRDIYSMHTFLTERKLNVGKVYRLLLRYYREECRDCSSPGPDTK : 947
Tcurva 234 : VAETFAGLLLDGGPFAGRFAHVVFAVW-DTAPGAPRHAAFARRFGSL---------- : 279
α10
β9
ribose cap
α11
α14
Rat 948 : LYPFIYHAAESSAETSDQPGQRTGT---- : 972
Tcurva - : ---------- : -
● ADP-HPD interacting residues

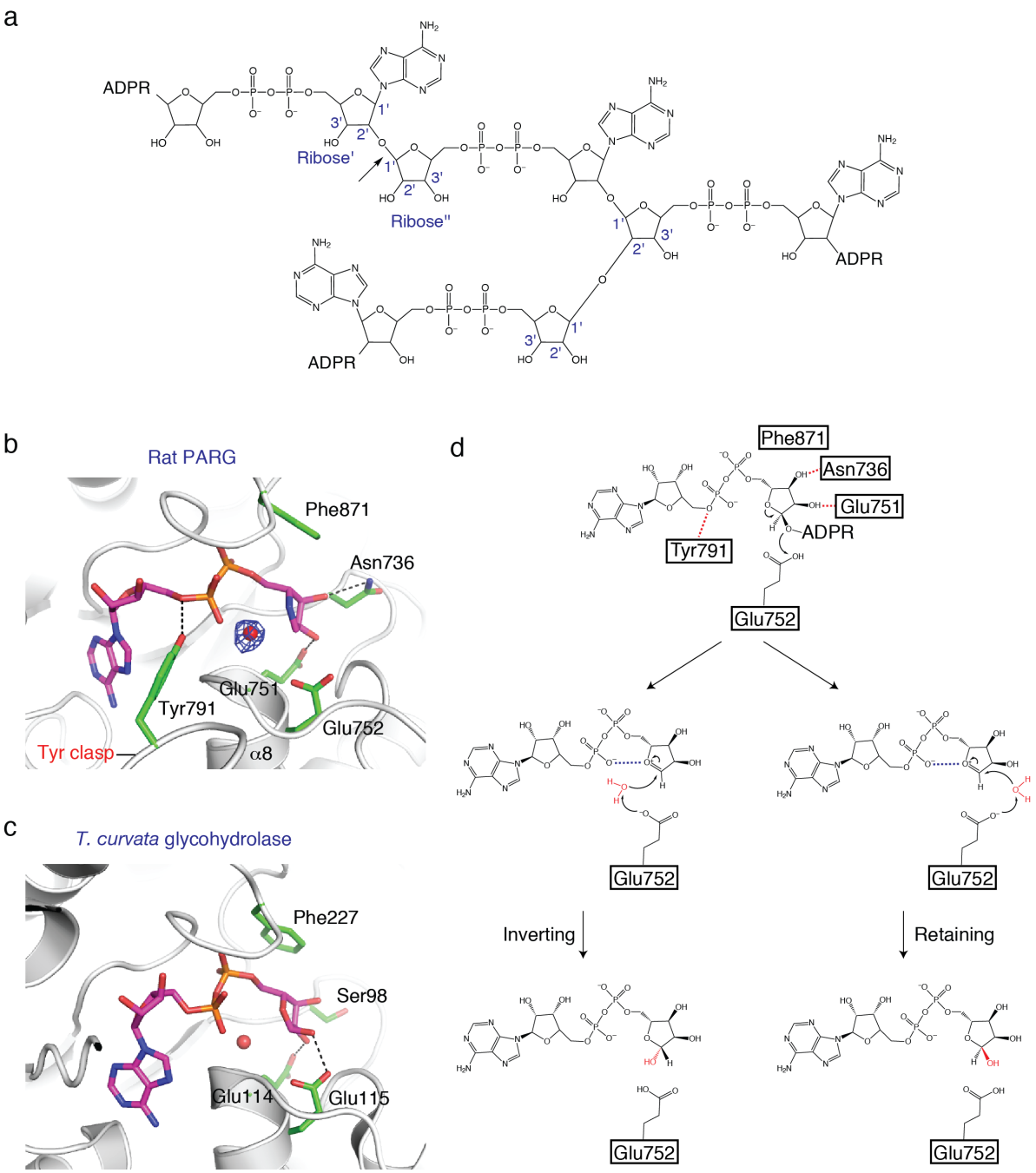
Supplementary Figure 2 Structure-based alignment of rat PARG and *T. curvata* glycohydrolase. The conserved strands (S2, S3, etc.) and helices (H1, H2, etc.) of a canonical macro domain fold are shown with the sequences of rat PARG and *T. curvata* glycohydrolase aligned by superposition of the crystal structures. Residues in contact with the ADP-HPD ligand (•) are primarily located within the connecting segments joining strands S2, S3, S6, and S7 of the macro domain fold. A polypeptide segment in the *T. curvata* glycohydrolase, which we term the ribose cap (cyan), completely blocks the 2′-OH of the adenosine ribose and restricts its activity to exo-glycosidic cleavage of PAR polymers[3].



Supplementary Figure 3 Stereo view of the electron density for the ADP-HPD ligand. Difference electron density map ($F_o - F_c$) for the ADP-HPD contoured at 3.0 σ (blue) is shown with $2F_o - F_c$ electron density around the ligand (contoured at 1.0 σ, grey). The ligand-bound rPARG^{385} structure was solved by molecular replacement using the apo-rPARG^{385} structure as a search model, and phases were calculated prior to the addition of ADP-HPD to the model. The orientation is similar to Figure 3b.



Supplementary Figure 4 Structural comparisons of rat PARG, *T. curvata* glycohydrolase and other macrodomains. (**a**) Structural comparisons of (top) rPARG[385], (middle) *T. curvata* glycohydrolase[3] (PDB accession 3SII), and (bottom) the macrodomain Af1521[4] (PDB accession 2BFQ) reveal the expanded structure of the PARG macro domain fold. Additional N-terminal (pink) and C-terminal (wheat) helical bundles pack against the conserved macrodomain core (white) of the rat PARG protein. The canonical secondary structural elements of the macrodomain are shown with alphanumeric labels in macrodomain Af1521 (bottom). The actual secondary structures assigned by DSSP[5] and labeled with Greek letter prefixes (β7, α7, etc.) are shown for rPARG[385] (top) and *T. curvata* glycohydrolase (middle)**.** (**b-c**) Structural (**b**) and sequence (**c**) comparison of the glycine-rich loop of rat PARG (orange), *T. curvata* glycohydrolase[3] (cyan, PDB accession 3SII), macrodomain D1[6] (green, PDB accession 2X47) and macrodomain Af1521[4] (purple, PDB accession 2BFQ). Catalytic residues are shown as a stick in panel (**b**) and colored red in panel (**c**). The PARG signature motif emanates from the glycine-rich loop and precisely orients the catalytic Glu752 residue towards C1′ of ribose″ moiety. Hydrogen bonds that position the Glu752 side chain and involve neighboring mainchain nitrogen atoms from Gly742 and Val749 of PARG signature motif are shown as dashed lines.



Supplementary Figure 5 Proposed reaction scheme for mammalian PARG. (**a**) The poly(ADP-ribose) (PAR) substrate is a branched, repeating polymer of ADP-ribose units linked by α(1″-2′) O-glycosidic linkages with irregular branching via α(1‴-2″) O-glycosidic linkages. PARG cleaves the α(1″-2′) O-glycosidic glycosidic linkages (arrows), releasing oligo-(ADP-ribose) chains and ADP-ribose. (**b**) A bound water molecule located close to Glu752 in the unliganded rPARG^{385} structure, is in a similar position as a bound water in the *T. curvata* glycohydrolase complexed to ADP-ribose[3] (**c**). We have superimposed the ADP-HPD ligand from the 3.0 Å resolution complex structure onto the 1.95 Å structure of unliganded rPARG^{385} showing the bound water. This water can serve as the nucleophile attacking the anomeric C1′ of ribose″ as previously proposed[3]. The position of Glu752 with respect to the substrate analogue could support a water attack from either side of the ribose ring, generating either the ADP-α-ribose″ (retaining mechanism) or the ADP-β-ribose″ (inverting mechanism) as shown in panel (**d**).

Supplementary Table 1 Crystallographic data statistics

	apo rPARG^{385}			rPARG^{385}–ADP-HPD complex
Data collection				
Space group		$C222_1$		$C222_1$
Cell dimensions				
a, *b*, *c* (Å)		126.5, 199.5, 50.5		130.8, 196.0, 163.5
α, β, γ (°)		90.0, 90.0, 90.0		90.0, 90.0, 90.0
	Native	*Peak*	*Inflection*	
Wavelength	1.12712	0.97954	0.97968	1.12712
Resolution (Å)	30–1.95	30–2.25	30–2.25	30–3.0
R_{sym} (%)	5.7 (44.1)	6.7 (46.2)	6.5 (44.4)	5.5 (65.5)
$I / \sigma I$	21.1 (1.7)	29.2 (3.9)	29.6 (4.0)	31.9 (3.7)
Completeness (%)	98.1 (83.4)	100.0 (100.0)	100.0 (100.0)	90.1 (98.7)
Redundancy	6.2 (3.6)	7.6 (6.4)	7.6 (6.5)	8.8 (9.2)
Refinement				
Resolution (Å)	30–1.95			30–3.0
No. reflections	43903			35972
R_{work} / R_{free}	18.0 / 21.7			24.6 / 27.4
No. atoms				
Protein	4,151			11,847
Ligand/ion	–			105
Water	203			–
B-factors				
Protein	32.1			49.7
Ligand/ion	–			78.4
Water	38.2			–
R.m.s deviations				
Bond lengths (Å)	0.014			0.012
Bond angles (°)	1.62			1.40

*Values in parentheses are for highest-resolution shell.
Each dataset was collected from a single crystal.

SUPPLEMENTARY REFERENCES

1. Dosztanyi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433-4 (2005).
2. Patel, C.N., Koh, D.W., Jacobson, M.K. & Oliveira, M.A. Identification of three critical acidic residues of poly(ADP-ribose) glycohydrolase involved in catalysis: determining the PARG catalytic domain. *Biochem J* **388**, 493-500 (2005).
3. Slade, D. et al. The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* **477**, 616-20 (2011).
4. Karras, G.I. et al. The macro domain is an ADP-ribose binding module. *EMBO J* **24**, 1911-20 (2005).
5. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-637 (1983).
6. Chen, D. et al. Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J Biol Chem* **286**, 13261-71 (2011).