



- (51) International Patent Classification:
C12Q 1/37 (2006.01) *C12P 21/06* (2006.01)
- (21) International Application Number:
PCT/US2016/036934
- (22) International Filing Date:
10 June 2016 (10.06.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/173,554 10 June 2015 (10.06.2015) US
62/264,834 8 December 2015 (08.12.2015) US
- (71) Applicant: **THE JOHNS HOPKINS UNIVERSITY**
[US/US]; Technology Ventures, 100 N. Charles Street, 5th
Floor, Baltimore, MD 21201 (US).
- (72) Inventors: **LEUNG, Anthony, K.**; 10107 Carillon Court,
Ellicott City, MD 21042 (US). **GABELLI, Sandra, B.**;
2933 Excelsior Springs Ct., Ellicott City, MD 21042 (US).
DANIELS, Casey, M.; 17801 Bowie Mill Road,
Derwood, MD 20855 (US).
- (74) Agents: **CORLESS, Peter, F.** et al.; Mintz Levin Cohn
Ferris Glovsky And Popeo, P.C., One Financial Center,
Boston, MA 02111 (US).
- (81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR IDENTIFYING ADP-RIBOSYLATED SITES BY MASS SPECTROMETRY

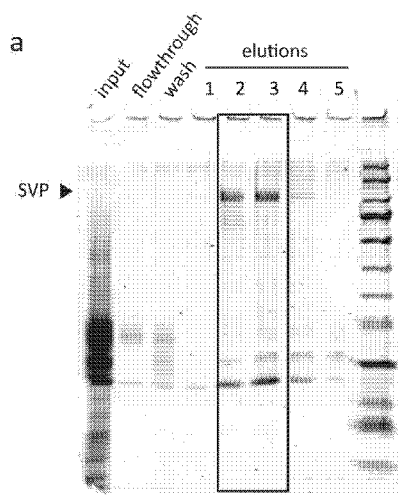


Figure 1A

(57) Abstract: The present disclosure relates to the use of a nudix hydrolase (RppH or NudT16) to cleave poly ADP-riboseylated proteins and leave a tag (phosphoribose) to be analyzed by mass spectrometry to identify post-translational modification (PTM) profiles of proteins. The disclosure relates to methods for identifying an ADP- ribosylation profile in hydrolyzed proteins using mass spectrometry.



Published:

- *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*

COMPOSITIONS AND METHODS FOR IDENTIFYING ADP- RIBOSYLATED SITES BY MASS SPECTROMETRY

5 RELATED APPLICATIONS

This application claims the benefit of priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No. 62/173,554, filed June 10, 2015, and Provisional Application No. 62/264,834, filed on December 8, 2015, both of which are incorporated herein by reference in their entireties.

10

STATEMENT OF RIGHTS TO INVENTION MADE UNDER FEDERALLY SPONSORED RESEARCH

This work was supported by the following grants: National Institutes of Health Grant Nos. R01-GM104135, 5T32CA009110, and P50 CA062924. The government
15 has certain rights in the invention.

FIELD OF THE DISCLOSURE

The present disclosure relates to the use of nudix hydrolases and mass spectrometry to detect sites of post-translational modifications (PTMs). In particular,
20 the disclosure relates to methods for detecting disease-relevant PTMs by detecting ADP-ribosylation using mass spectrometry.

BACKGROUND OF THE DISCLOSURE

Poly(ADP-ribose) and mono(ADP-ribose) (collectively referred to as ADP-
25 ribose or ADPr) are posttranslational modifications (PTMs) important in all major cellular processes. Poly(ADP-ribose) is particularly well known for the critical role it plays in DNA repair, and consequently, cancer. Though ADP-ribosylation is therapeutically important, study of this PTM has been limited by a lack of mass spectrometry based proteomic tools for identifying the amino acid residues carrying
30 this modification. Consequently, there is an urgent need in the art for improved

methods of detection of PTMs, and in particular, for improved methods of detecting PTMs relevant to diseases such as, for example, cancer.

SUMMARY OF THE DISCLOSURE

5 The present disclosure features methods for identifying the post translational modifications in proteins. In particular, the disclosure features methods for identifying ADP-ribosylation in proteins by using hydrolases to tag sites of ADP-ribosylation in proteins for subsequent analysis and/or detection with mass spectrometry.

10

DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this disclosure belongs. The following references provide one of skill with a general
15 definition of many of the terms used in this disclosure: Singleton *et al.*, *Dictionary of Microbiology and Molecular Biology* (2nd ed. 1994); *The Cambridge Dictionary of Science and Technology* (Walker ed., 1988); *The Glossary of Genetics*, 5th Ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); and Hale & Marham, *The Harper Collins Dictionary of Biology* (1991). As used herein, the following terms have the meanings
20 ascribed to them unless specified otherwise.

By “agent” is meant any small compound, antibody, nucleic acid molecule, or polypeptide, or fragments thereof.

By “alteration” is meant an increase or decrease. An alteration may be by as little as 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, or by 40%, 50%, 60%, or even by as
25 much as 75%, 80%, 90%, or 100%. An alteration may be a change in sequence relative to a reference sequence or a change in expression level, activity, or epigenetic marker (e.g., promoter methylation).

By “biologic sample” is meant any tissue, cell, fluid, or other material derived from an organism.

30 By “cancer” (also called neoplasia, dysplasia, malignant tumor, and/or malignant neoplasia) is meant a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body (e.g., metastasize).

Not all tumors are cancerous; benign tumors do not spread to other parts of the body. There are over 100 different known cancers that affect humans.

In this disclosure, “comprises,” “comprising,” “containing” and “having” and the like may have the meaning ascribed to them in U.S. Patent law and may mean
5 “includes,” “including,” and the like; “consisting essentially of” or “consists essentially” likewise has the meaning ascribed in U.S. Patent law and the term is open-ended, allowing for the presence of more than that which is recited so long as basic or novel characteristics of that which is recited is not changed by the presence of more than that which is recited, but excludes prior art embodiments.

10 By “control” is meant a standard or reference condition. For example, the methylation level present at a promoter in a neoplasia may be compared to the level of methylation present at that promoter in a corresponding normal tissue.

“Detect” refers to identifying the presence, absence or amount of the analyte to be detected.

15 By “diagnostic” is meant any method that identifies the presence of a pathologic condition or characterizes the nature of a pathologic condition (e.g., a neoplasia). Diagnostic methods differ in their sensitivity and specificity. While a particular diagnostic method may not provide a definitive diagnosis of a condition, it suffices if the method provides a positive indication that aids in diagnosis.

20 By “hydrolase” is meant an enzyme that catalyzes the hydrolysis of a chemical bond. In some embodiments, hydrolases (and their homologs) of the present disclosure are isolated from prokaryotes, eukaryotes, bacteria, yeast, xenopus, zebrafish, hydra, nematode, drosophila, fruit fly, mammalian, mouse, sheep, bovine, equine, porcine, caprine, ovine, canine, feline, lupine, vulpine, murine, primate, and
25 human species.

The phrase “in combination with” is intended to refer to all forms of administration that provide a de-methylating agent, or the methods of the instant disclosure (e.g. methods of detection of methylation) together with a second agent, such as a chemotherapeutic agent, or a de-methylating agent, where the two are
30 administered concurrently or sequentially in any order.

By “isolated polynucleotide” is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the

nucleic acid molecule of the disclosure is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule that is transcribed from a DNA molecule, as well as a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

By “PARP inhibitor” is meant an agent that inhibits enzymes from the poly ADP ribose polymerase family (e.g., PARP). A PARP inhibitor may include, but is not limited to, any of the following agents:

Compound	Smile Code	Vendor	Catalog number
1	<chem>c12c(nc(c3cccc3)[nH]c1=O)ccc2</chem>	Asinex	BAS 05269446
2	<chem>c12c(c3CCCC3s1)c(-O)[nH]cn2</chem>	Chemdiv	1306-0033
3	<chem>c12c([nH]cnc1cccc2)=O</chem>	Chemdiv	C201-0583
4	D ^P Q <chem>c1ccc2c(c1)C(NCC2OCCCCN1CCCCC1)=O</chem>	Calbiochem	300270
5	NU1025 <chem>c1cc(c2c(c1)c([nH]c1c2cc1)O)O</chem>	Calbiochem	493800
6	PJ34 <chem>c1ccc2c1c([nH]c1c2cc1)NC(CN(C)C)=O,Cl</chem>	Calbiochem	528150
7	EB-47 <chem>c1cc(c2c(c1)C(NC2=O)NC(=O)CN1CCN(CC1)C([C@@H]1O[C@@H]([C@@H]([C@@H]1O))n1c2c(nc1c(ncn2)N))=O</chem>	Calbiochem	324473
8	6(5H)-Phenanthridinone <chem>c1cc2c3c([nH]c1c2cc1)O)ccc3</chem>	Sigma	299634
9	DiQ ₁ 1,5-Isoquinolinediol <chem>c1ccc2c1c(ccc2O)O</chem>	Sigma	I138
10	<chem>c1ccc2c1[nH]c([nH]c2=O)=O</chem>	Sigma	142026
11	DR2313 <chem>C1SCC2c1c(nc([nH]2)C)=O</chem>	Alexis biochemicals	DR2313
12	INH2BP <chem>o1c(ccc2c1ccc2)N=O</chem>	Alexis biochemicals	ALX-270-278
13	3-Methyl-5-AIQ <chem>c1ccc2c(c1N)cc([nH]c2=O)C</chem>	Alexis biochemicals	ALX-270-450
14	5-AIQ <chem>c1cc(c2c(c1)c([nH]c2=O)N,Cl</chem>	Calbiochem	164300
15	TiQ-A <chem>c1ccc2c(c1)c([nH]c1c2cc1)=O</chem>	Calbiochem	612100
16	KU0058948 <chem>c1ccc2c(c1)c([nH]c2=O)Cc1cc(c(cc1)F)C(N1CCNCCC1)=O</chem>	KuDos (gift)	KU0058948
17	4-Amino-1,8-naphthalimide <chem>c1ccc2c3c1c(ccc3c([nH]c2=O)=O)N</chem>	Sigma	A0966
18	1-Isoquinolind, 1SQ <chem>c1ccc2c(c1O)ccc2</chem>	Sigma	55435
19	3-Aminobenzamide, 3-AB <chem>c1(ccc(c1)C(N)=O)N</chem>	Sigma	A0788
20	Vitamin B3 (nicotinamide) <chem>O=C(N)c1cccnc1</chem>	Sigma	N3376
21	<chem>O=c1[nH]cnc2sc3CCCC3c12</chem>	ENAMINE	T0514-1628
22	<chem>O=c1[nH]ncc2cccc12</chem>	ENAMINE	T0509-6940
23	<chem>OC(=O)CC1nc2sc3CCCC3c2c(-O)[nH]1</chem>	ENAMINE	T5621369
24	<chem>Cc1cccc2c(-O)[nH]c(CSc3ncccn3)nc12</chem>	ENAMINE	T5636531
25	<chem>CCNC(=O)CCc1nc2sc3CCCC3c2c(-O)[nH]1</chem>	ENAMINE	T5823447
26	<chem>Cc1csc1SCc2nc3c(C)ccc3c(-O)[nH]2)n1</chem>	ENAMINE	T5875616
27	<chem>n1ccccc1SCc1[nH]c(-O)c2c(n1)nc2c1cccc1</chem>	ENAMINE	T5810060
28	<chem>COc1ccc(cc1)NC(=O)c1[nH]c1c(-O)c2c1cccc2</chem>	ENAMINE	T5689927

29	<chem>n1[nH]c(=O)[nH+](c(=O)c1Sc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T0511-0211
30	<chem>c1cccc2c1c([nH]c(n2)C(=O)N)N=C/c1ccc(F)cc1)=O</chem>	ENAMINE	T0512-3494
31	<chem>ClC(Cl)(Cl)C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0515-7484
32	<chem>c1ccnc(O)c1C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0517-3776
33	<chem>COc1cccc(OC)c1C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0517-4507
34	<chem>[nH+](cccc1CNC(=O)CCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T0519-7725
35	<chem>c1cccc2c1c([nH]c(n2)Sc1cc(=O)oc2ccc3c(c12)ccc3)=O</chem>	ENAMINE	T5214418
36	<chem>c1cccc2c1c([nH]c(n2)CSc1nnc2n1ccc2)=O</chem>	ENAMINE	T5220503
37	<chem>O1CCC(C@@H)1(CNCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5266916
38	<chem>c1cccc(OC)c1NC(=O)CSc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5282221
39	<chem>O=C([O-])[C@H](C)(NC(=O)CCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5301904
40	<chem>O1CCC(C@H)1(CNC(=O)c1c2c(n[nH]1)ccc2</chem>	ENAMINE	T5302202
41	<chem>c1cc(F)ccc1CNC(=O)CN(CC)Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5405085
42	<chem>c1cc(F)cc([N+])([O-])c1NC(=O)C(N@@H+)(C)(Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5406525
43	<chem>c1cccc(F)c1CNC(=O)c1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5433140
44	<chem>C1COCCN1CNC(=O)Cc1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5433913
45	<chem>CCGn1c(=O)[nH]nc1Sc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5435488
46	<chem>NC(=O)CCS(=O)(=O)CCN(C)Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5443746
47	<chem>COc1ccc(cc1OC)CN(Cc1ccsc1)C(=O)c1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5458928
48	<chem>c1cccc2c1c(n[nH]c2=O)C(=O)N[C@H](C)(c1ccc(cc1)n1ccc1)</chem>	ENAMINE	T5475708
49	<chem>Gn1nc(ccc1=O)C(=O)N(CCO)Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5510865
50	<chem>n1cccc1[C@@H](C)(NC(=O)CCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5513304
51	<chem>CCOC(=O)NC(=O)CN(CC)Cc1nc2c(c([nH]1)=O)sc2</chem>	ENAMINE	T5342377
52	<chem>c1cs2c1c(nc([nH]2)CN(CCO)Cc1[nH]c(=O)c2c(n1)sc2)=O</chem>	ENAMINE	T5233261
53	<chem>O=C(/1NC(=O)NC(=O)C1=C/Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T5390770
54	<chem>NC(=O)NC(=O)CN(C)Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5434894
55	<chem>c1cccc(c1C#N)O[C@@H](C)(c1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5442715
56	<chem>c1cccc2c1c([nH]c(n2)CCC(=O)N[C@H]1(CCN@@H+)(C1)(Cc1cccc1)))=O</chem>	ENAMINE	T5524834
57	<chem>COc1ccc(cc1)CCNC(=O)CSc1[nH]c(=O)c2c(n1)nc2c1cc(C)ccc1</chem>	ENAMINE	T5526850
58	<chem>o1cccc1CNC(=O)c1c2c(n[nH]1)ccc2</chem>	ENAMINE	T5531146
59	<chem>c1cs2c1nc([nH]c2=O)C(N@@H+)(CCC(C@H)1(c1ccsc1)))</chem>	ENAMINE	T5533330
60	<chem>c1cccc(c1C#N)NC(=O)C@H(C)(N(C)Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5540724
61	<chem>c1cc(F)cc(F)c1[C@@H](C)(NC(=O)Cc1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5546152
62	<chem>CCOc1cccc(c1O)[C@H]1(NC(=O)c2c(N1)ccc2</chem>	ENAMINE	T5639683
63	<chem>c1cccc2c1C(N[C@H](N2)(c1c(F)cc(Br)cc1))=O</chem>	ENAMINE	T5642602
64	<chem>COc1c(O)c(ccc1)[C@@H]1(NC(=O)c2c(N1)ccc2)OC</chem>	ENAMINE	T5699977
65	<chem>c1scoc1C(=O)OCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5707050
66	<chem>COc(=O)c1ccc(cc1)NC(=O)CCc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5707911
67	<chem>c1ccc(Br)cc1NC(=O)Cc1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5717474
68	<chem>c1cccc(O)c1NC(=O)Cc1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5723652
69	<chem>c1ccc(OC)cc1NC(=O)c1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5725081
70	<chem>c1cc(O)ccc1C(=O)N(C[C@@H]1(CCCO1))Cc1[nH]c(=O)c2c(n1)ccc2</chem>	ENAMINE	T5731587
71	<chem>c1cccc1NC(=O)Cc1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5734903
72	<chem>c1cccc1NC(=O)c1n[nH]c(=O)c2c1ccc2</chem>	ENAMINE	T5742485

73	<chem>c1cccc(C)C1[C@@H](C)(NC(=O)Cc1n[nH]c(=O)c2c1cccc2)</chem>	ENAMINE	T5748611
74	<chem>c1ccc2c1c[nH]c(n2)SCC(=O)N[C@@H]1(CCCc2c1cccc2)=O</chem>	ENAMINE	T5757803
75	<chem>c1cccc(F)c1CN(C)Cc1nc2c(c[nH]1)=O)sc2</chem>	ENAMINE	T5773275
76	<chem>n1cnen1c1ccc(cc1)[C@@H](C)(N(C)Cc1nc2c(c[nH]1)=O)sc2</chem>	ENAMINE	T5774012
77	<chem>c1cc(O)ccc1C1(CC1)C(=O)N(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5799957
78	<chem>N1CCN(C1=O)C(=O)C[N@H](C)(Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5807539
79	<chem>c1cccc2c1c[nH]c(n2)[C@@H](C)(N(C)Cc1cc(O)C(c1)OC)=O</chem>	ENAMINE	T5812575
80	<chem>c1cccc1[C@@H](C)(NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5817800
81	<chem>C[NH+](C)Cc1c(ccc1)CN(C=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5820448
82	<chem>c1ccc(O)Cc1CN(C=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5820522
83	<chem>c1cccc2c1c(c[nH]2)C(=O)N(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5823618
84	<chem>c1cc(F)ccc1[C@@H](C)(N(C)Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5826883
85	<chem>c1cc(O)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)sc2</chem>	ENAMINE	T5826869
86	<chem>c1cccc1C(=O)N(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5827773
87	<chem>Fc1c(F)ccc(c1)N(C=O)Cn1c(=O)[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5828779
88	<chem>c1cccc2c1c(n[nH]c2=O)C(=O)N(Cc1cc2cc1)cc(c2)OC</chem>	ENAMINE	T5829176
89	<chem>N#Cc1cc(ccc1)NC(=O)CSc1[nH]c(=O)c2c(n1)sc2</chem>	ENAMINE	T5829810
90	<chem>c1cccc2c1c(n[nH]c2=O)C(=O)N(Cc1ccc(cc1)C[NH+](CC)CC</chem>	ENAMINE	T5831807
91	<chem>c1cccc2c1c[nH]c(n2)C[N@@H+](Cc2c(C)sc2)=O</chem>	ENAMINE	T5836432
92	<chem>c1cccc2c1c[nH]c(n2CC(=O)N[C@@H](C)(c1c(F)cc(F)cc1)=O)=O</chem>	ENAMINE	T5840584
93	<chem>c1ccc(F)cc1NC(=O)C[N@H+](C)(Cc1nc2c(c[nH]1)=O)sc2</chem>	ENAMINE	T5844481
94	<chem>O1COc2c1cc(cc2)CN(C=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5847624
95	<chem>c1cccc2c1C(N[C@@H](N2)(c1cc(O)c2c1)OCCO2)=O</chem>	ENAMINE	T5853381
96	<chem>C1CN(C)Cc2c1nc(s2)NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5855515
97	<chem>c1cc(F)ccc1OC[C@@H](O)(CN(C)Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5863205
98	<chem>Nc1nc(n[nH]1)SCc1[nH]c(=O)c2c(n1)c(C)ccc2</chem>	ENAMINE	T5867921
99	<chem>c1cccc2c1c[nH]c(n2)CCC(=O)N(C)Cc1c(ccc1)N1CCOCC1=O</chem>	ENAMINE	T5904510
100	<chem>c1cccc2c1c(n[nH]c2=O)CC(=O)N(Cc1ccc[nH+](c1)N1CCOCC1</chem>	ENAMINE	T5904604
101	<chem>c1cc([N+](=O)=O)ccc1[C@@H](O)(CN(C=O)Cc1n[nH]c(=O)c2c1cccc2)</chem>	ENAMINE	T5918237
102	<chem>Cc1cc(c(cc1)O)C(=O)c1cc2c(nc1)[nH]c(=O)c1c2cccc1</chem>	ENAMINE	T5949414
103	<chem>c1ccc([O-])c1C(=O)c1cc2c(nc1)[nH]c(=O)c1c2cccc1</chem>	ENAMINE	T5949445
104	<chem>c1cc(F)cc(F)c1[C@@H](C)(N(C=O)CCC1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5970736
105	<chem>c1cccc2c(=O)[nH]c3c(c12)ccc(n3)c1cccc1</chem>	ENAMINE	T5974626
106	<chem>Cc1cc(C(F)(F)F)nc2[nH]c(=O)c3c(c12)ccc3</chem>	ENAMINE	T5989689
107	<chem>n1c(c2cc(S(N(CCO)C)=O)=O)c(cc2)Cc2c(ccc2)c([nH]1)=O</chem>	ASINEX_GOLD	BAS 04375852
108	<chem>n1c(c2cc(S(=O)(=O)NCCN3CCOCC3)c(cc2)Cc2c(ccc2)c([nH]1)=O</chem>	ASINEX_GOLD	BAS 06635491
109	<chem>c1cc(F)ccc1C(=O)N(Cc1c(O)cccc1)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ASINEX_PLATINUM	ASN 05812848
110	<chem>c1cccc1CCNS(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	ASINEX_PLATINUM	ASN 09801991
111	<chem>CCc1ccc(cc1)c1n[nH]c(=O)c2c1cccc2</chem>	ASINEX_GOLD	BAS 06805602
112	<chem>n1c(c2cccc2c[nH]1)=O)Cc1ccc(cc1)CC</chem>	INTERBIOSCREEN	STOCK15-03944
113	<chem>n1c(c2cccc2c[nH]1)=O)c1cc(N)c(cc1)C</chem>	INTERBIOSCREEN	STOCK45-10309
114	<chem>O1COc2c1cc(c(c2)Cl)[C@@H]1(NC(=O)c2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-38708
115	<chem>O1COc2c1cc(cc2)[C@@H]1(NC(=O)c2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-31658
116	<chem>c12c3cccc1cccc2c(=O)[nH]c3=O</chem>	INTERBIOSCREEN	STOCK25-08660

117	<chem>COc1c(O)c(cc1OC)[C@@H]1N(C(=O)C2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-34270
118	<chem>O1COc2c1cc(cc2)c1[nH]c(=O)c2c(n1)cccc2</chem>	INTERBIOSCREEN	STOCK55-79756
119	<chem>[nH+]1c[nH]cc1CCN(C(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-28405
120	<chem>c1cccc2c1c[nH]c2=O)CC(=O)NCC(=O)Nc1cc(ccc1)NC(=O)C</chem>	INTERBIOSCREEN	STOCK65-34326
121	<chem>C1COCCN1C(=O)CNC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-34374
122	<chem>C1COCCN1C(=O)[C@@H](C)(N(C(=O)Cc1n[nH]c(=O)c2c1cccc2)</chem>	INTERBIOSCREEN	STOCK65-38709
123	<chem>c1cccc2c1c([nH]c(n2)SCC(=O)NCCc1c[nH]c2c1cccc2)=O</chem>	INTERBIOSCREEN	STOCK65-43755
124	<chem>[nH+]1cccc1CNC(=O)CCNC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-46880
125	<chem>n1c(c2cccc2c([nH]1)=O)Cc1c(cc1cc1)Cl</chem>	KEYORGANICS	1E-57
126	<chem>[N+](c1c2c(c(=O)[nH]cc2)ccc1)([O-])=O</chem>	KEYORGANICS	10B-099
127	<chem>FC(F)(F)c1cc(ccc1)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	KEYORGANICS	2F-3015
128	<chem>FC(F)(F)c1cc(ccc1)NCC1n[nH]c(=O)c2c1cccc2</chem>	KEYORGANICS	12B-089
129	<chem>COc1ccc(cc1)c1n[nH]c(=O)c2c1cccc2</chem>	KEYORGANICS	MS-1309
130	<chem>c1cccc2c1c([nH]c(n2)[C@@H](C)(N(C2+)[Cc1ccc(Cl)cc1]))=O</chem>	KEYORGANICS	MS-3000
131	<chem>c1cccc(Cl)c1CNCc1[nH]c(=O)c2c(n1)cccc2</chem>	KEYORGANICS	MS-3205
132	<chem>n1c(c2cccc2c([nH]1)=O)c1cc([N+](O-)=O)c(cc1)NCCCN(C)C</chem>	VITAS-M LABORATORY	STK238640
133	<chem>c1(cc(=O)c2c(cc1)cccc2)c1cccc1</chem>	VITAS-M LABORATORY	STK164205
134	<chem>c1([nH]c2ccc(S(NC3cccc3)(=O)=O)cc2c2c1cccc2)=O</chem>	VITAS-M LABORATORY	STK272141
135	<chem>c12c3c(cccc3c([nH]n1)=O)c1c2cccc1</chem>	CHEMBRIDGE	6259889
136	<chem>c12c3c(n[nH]c(c3ccc2NC2CCCC2)=O)c2c(c1=O)cccc2</chem>	CHEMBRIDGE	7456760
137	<chem>n1c(c2cc(S(N3CCN(CC3)Cc3cccc3)(=O)=O)c(cc2)C)c2c(c(=O)[nH]1)cccc2</chem>	CHEMBRIDGE	7669944
138	<chem>c1([nH]c2ccc(S(N3CC3)(=O)=O)cc2c2c1cccc2)=O</chem>	CHEMBRIDGE	9040441
139	<chem>c12c3cc(S(=O)(=O)N)cc1cccc2c(=O)[nH]c3=O</chem>	CHEMBRIDGE	7650155
140	<chem>FC(F)(F)c1cc(ccc1)NC(=O)CSc1[nH]c(=O)c2c(n1)cccc2</chem>	CHEMBRIDGE	7663670
141	<chem>Cc1ccc(cc1)NS(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	7841182
142	<chem>O1COc2c1cc(cc2)CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMBRIDGE	7959335
143	<chem>C1COCCN1S(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	9026017
144	<chem>CN(C)S(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	9030160
145	<chem>c12c(c(=O)[nH]c(n1)C)c1CCCC1s2</chem>	CHEMDIV	3909-7439
146	<chem>c12[nH]c3CCCC3c2cccc1C(=O)N</chem>	CHEMDIV	5348-0038
147	<chem>c1([nH]c(Cl)cc2c1cccc2)=O</chem>	CHEMDIV	3909-9376
148	<chem>c12c3c(c(Cl)cccc3c([nH]n1)=O)c(c1c2cccc1)=O</chem>	CHEMDIV	8011-2789
149	<chem>c1([nH]c2ccc([N+](O-)=O)cc2c2c1cccc2)=O</chem>	CHEMDIV	R507-0010
150	<chem>c1cc(F)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)CCCC2</chem>	CHEMDIV	D365-0030
151	<chem>COc1ccc(cc1)NC(=O)CSc1[nH]c(=O)c2c(n1)CCCC2</chem>	CHEMDIV	D365-0199
152	<chem>c1cccc(c1C(=O)N)NC(=O)CSc1[nH]c(=O)c2c(n1)cccc2</chem>	CHEMDIV	F019-0833
153	<chem>c1cccc2c1c([nH]c(n2)CSc1nc(cc([nH]1)=O)C)=O</chem>	CHEMDIV	8016-1228
154	<chem>c1cc(F)ccc1CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0248
155	<chem>c1ncccc1CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0339
156	<chem>CC[NH+](CC)CCCN(C(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0477
157	<chem>c1ncccc1Cc1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-3059
158	<chem>c1cccc(Cl(F)F)F)c1C(=O)NCC1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	G786-2295
159	<chem>c1cc(Cl)ccc1CCC(=O)NCC1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	G786-2308
160	<chem>C1(c2c(c(O)ccc2)CCN1)=O</chem>	LIFE CHEMICAL	F2679-0001

161		<chem>n1c2c(C(=O)N)cccc2nc2c1cccc2</chem>	LIFE CHEMICAL	F3031-0010
162		<chem>c1cc(Br)ccc1SCc1[nH]c(=O)c2c(n1)cccc2</chem>	LIFE CHEMICAL	F2551-0013
163		<chem>c1cc(O)ccc1NC(=O)CSc1nc2c(c([nH]1)=O)CCCC2</chem>	LIFE CHEMICAL	F3116-0089
164		<chem>c1cccc2c1C(N[C@@H]1(N2Cc2c(C1)cccc2))=O</chem>	LIFE CHEMICAL	F3284-0013
165		<chem>COc1ccc(cc1)C[C@@H]1(c2c(CN3[C@@H]1(NC(=O)c1c3cccc1))cccc2)</chem>	LIFE CHEMICAL	F3097-0628
166		<chem>c1cc(O)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)CCCC2</chem>	LIFE CHEMICAL	F3308-0716
167		<chem>c1cccc1CN(C(=O)CSc1[nH]c(=O)c2c(n1)n(nc2)c1cccc1</chem>	LIFE CHEMICAL	F2085-0030
168		<chem>OCCNC(=O)C(=O)NCc1n[nH]c(=O)c2c1cccc2</chem>	LIFE CHEMICAL	F2092-0051
169		<chem>c1cccc1CN(C(=O)COc1cccc2c1CCNC2=O</chem>	LIFE CHEMICAL	F2679-0060
170		<chem>O=C1NC(=O)c2c1c(ccc2)NC(=O)CCS(=O)(=O)c1ccc(cc1)OC</chem>	LIFE CHEMICAL	F2814-1523
171		<chem>Fc1cccc2c1C(N[C@@H](N2)(c1c(O)ccc(c1)[N+](O-)=O)=O</chem>	MAYBRIDGE	BT8 04585
172		<chem>c1cc(O)ccc1SCc1n[nH]c(=O)c2c1cccc2</chem>	MAYBRIDGE	KM 09200
173		<chem>c1cc(cc(c1)C(N)=O)O</chem>	KaironKern	K181
174		<chem>c1(c2c(ccc1)ncc(n2)c1ccc(cc1)O)C(=O)N</chem>	Alexis biochemicals	ALX-270-486
175	ABT-888	<chem>O=C(N)c1c2nc([nH]c2ccc1)C11(NCCC1)C</chem>	Alexis biochemicals	270-444-M005
176		<chem>c1(c(c2cccc2oc1=O)Nc1cccc1)NC(=O)C</chem>	ChemBridge	5351069
177		<chem>c1(c(c2cccc2oc1=O)NCCCN1CCN(CC1)C)N</chem>	ChemBridge	5244706
178		<chem>c1(n(c(nm1)SCc1[nH]c(=O)c2c(n1)cccc2)CC)c1ccc(cc1)Cl</chem>	ChemBridge	7965028
179		<chem>c12nc(N3C(C)CCCC3)n(c1c(=O)[nH]c(n2C)=O)Cc1cccc1</chem>	ChemBridge	7996603
180		<chem>c12nc(N3C(C)CCCC3)n(c1c(=O)[nH]c(n2C)=O)Cc1cccc1</chem>	ChemBridge	7367553
181		<chem>c12c(nc(C3CC3)nc1O)CCSC2</chem>	Maybridge	RF03876
182	XAV939	<chem>c1(c2ccc(c1F)F)cc2nc2c(CSCC2)c(n1)O</chem>	Maybridge	RF03920
183		<chem>c12c(nc(c3cccc3)nc1O)CCSC2</chem>	Maybridge	RF03877
184	Olaparib	<chem>Fc1c(C(=O)N2CCN(C(=O)C3CC3)CC)cc(Gc2n[nH]c(=O)c3c2ccc3)cc1</chem>	JS Research Chemicals	QAS 763113-22-0
185	Rucaparib	<chem>Fc4cc2c1c(nc(c1CCNC2=O)c3ccc(cc3)CN)C4</chem>	Axon MedChem	Axon 1529

By “sensitivity” is meant the percentage of subjects with a particular disease that are correctly detected as having the disease. For example, an assay that detects 98/100 of carcinomas has 98% sensitivity.

By “specificity” is meant the percentage of subjects without a particular disease who test negative.

By “marker” is meant any protein or polynucleotide having an alteration in expression level or activity that is associated with a disease or disorder.

The term “neoplasm” or “neoplasia” as used herein refers to inappropriately high levels of cell division, inappropriately low levels of apoptosis, or both. A neoplasm creates an unstructured mass (e.g., a tumor), which may be either benign or malignant. For example, cancer is a neoplasia. Examples of cancers include, without limitation, lung carcinoma, small cell lung carcinoma, non-small cell lung carcinoma, leukemias (e.g., acute leukemia, acute lymphocytic leukemia, acute myelocytic

leukemia, acute myeloblastic leukemia, acute promyelocytic leukemia, acute myelomonocytic leukemia, acute monocytic leukemia, acute erythroleukemia, chronic leukemia, chronic myelocytic leukemia, chronic lymphocytic leukemia), polycythemia vera, lymphoma (Hodgkin's disease, non-Hodgkin's disease),

5 Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors such as sarcomas and carcinomas (e.g., fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic

10 cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor,

15 cervical cancer, uterine cancer, testicular cancer, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, schwannoma, meningioma, melanoma, neuroblastoma, and retinoblastoma). Lymphoproliferative disorders are also considered to be proliferative diseases.

20 The phrase "nucleic acid" as used herein refers to an oligonucleotide, nucleotide, polynucleotide, or to a fragment of any of these, to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent a sense or antisense strand, peptide nucleic acid (PNA), or to any DNA-like or RNA-like material, natural or synthetic in origin. As will be understood

25 by those of skill in the art, when the nucleic acid is RNA, the deoxynucleotides A, G, C, and T are replaced by ribonucleotides A, G, C, and U, respectively.

The term "nucleic acid sequence" refers to a succession of letters that indicate the order of nucleotides within a DNA (e.g., GACT) or RNA (e.g., GACU) molecule. By convention, sequences are usually presented from the 5' end to the 3' end. For

30 DNA, the sense strand is used. Because nucleic acids are normally linear (unbranched) polymers, specifying the sequence is equivalent to defining the covalent structure of the entire molecule. For this reason, the nucleic acid sequence is also termed the primary structure.

The term “gene” refers to a segment of deoxyribonucleic acid that encodes a polypeptide including the upstream and downstream regulatory sequences.

Specifically, the term gene includes the promoter region upstream of the gene.

The term “promoter” or “promoter region” refers to a minimal sequence
5 sufficient to direct transcription or to render promoter-dependent gene expression that is controllable for cell-type specific or tissue-specific gene expression, or is inducible by external signals or agents. Promoters may be located in the 5' or 3' regions of the gene. In general, a promoter includes, at least, 50, 75, 100, 125, 150, 175, 200, 250, 300, 400, 500, 750, 1000, 1500, or 2000 nucleotides upstream of a given coding
10 sequence (e.g., upstream of the coding sequence for genes). One of skill in the art will appreciate that a promoter location may vary outside these parameters for some genes, and also that some genes may comprise more than one promoter (e.g., multiple tissue specific promoters).

By “protein” or “polypeptide” or “peptide” is meant any chain of more than
15 two natural or unnatural amino acids, regardless of post-translational modification (e.g., glycosylation or phosphorylation), constituting all or part of a naturally-occurring or non-naturally occurring polypeptide or peptide, as is described herein.

The term “peptide sequence” or “amino acid sequence”, is the order in which amino acid residues, connected by peptide bonds, lie in the chain in peptides and
20 proteins. The sequence is generally reported from the N-terminal end containing free amino group to the C-terminal end containing free carboxyl group. Peptide sequence is often called protein sequence if it represents the primary structure of a protein.

A “purified” or “biologically pure” nucleic acid or protein is sufficiently free of other materials such that any impurities do not materially affect the biological
25 properties of the protein or cause other adverse consequences. That is, a nucleic acid or peptide of this disclosure is purified if it is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Purity and homogeneity are typically determined using analytical chemistry techniques, for
30 example, polyacrylamide gel electrophoresis or high performance liquid chromatography. The term “purified” may denote that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. For a protein that may be subjected to modifications, for example, phosphorylation or glycosylation, different

modifications may give rise to different isolated proteins, which may be separately purified.

Similarly, by “substantially pure” is meant a nucleotide or polypeptide that has been separated from the components that naturally accompany it. Typically, the
5 nucleotides and polypeptides are substantially pure when they are at least 60%, 70%, 80%, 90%, 95%, or even 99%, by weight, free from the proteins and naturally-occurring organic molecules with they are naturally associated.

The term “sample” as used herein refers to any biological or chemical mixture for use in the method of the disclosure. The sample may be a biological sample. The
10 biological samples are generally derived from a patient, preferably as a bodily fluid (such as tumor tissue, lymph node, sputum, blood, bone marrow, cerebrospinal fluid, phlegm, saliva, or urine) or cell lysate. The cell lysate may be prepared from a tissue sample (e.g. a tissue sample obtained by biopsy), for example, a tissue sample (e.g. a tissue sample obtained by biopsy), blood, cerebrospinal fluid, phlegm, saliva, urine, or
15 the sample may be cell lysate. In preferred examples, the sample is one or more of blood, blood plasma, serum, cells, a cellular extract, a cellular aspirate, tissues, a tissue sample, or a tissue biopsy. In preferred embodiments, the sample is from esophageal tumor cells, tissue or origin.

By “reduces” is meant a negative alteration of at least 10%, 25%, 50%, 75%,
20 or 100%.

By “reference” is meant a standard or control sample, state, or condition.

A “reference sequence” is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of or the entirety of a specified sequence; for example, a segment of a full-length cDNA or gene sequence, or the
25 complete cDNA or gene sequence. For polypeptides, the length of the reference polypeptide sequence will generally be at least about 16 amino acids, preferably at least about 20 amino acids, more preferably at least about 25 amino acids, and even more preferably about 35 amino acids, about 50 amino acids, or about 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will
30 generally be at least about 50 nucleotides, preferably at least about 60 nucleotides, more preferably at least about 75 nucleotides, and even more preferably about 100 nucleotides or about 300 nucleotides or any integer thereabout or therebetween.

By “*EcRppH*” is meant a polypeptide having the following sequence:

MIDDDGYRPN VGIVICNRQG QVMWARRFGQ HSWQFPQGGI
 NPGESAEQAM YRELFEEVGLSRKDVRILAS TRNWLRYKLP KRLVRWDTKP
 VCIGQKQKWFLQLVSGDAE INMQTSSTPE FDGWRWVSYW YPVRQVVSFK
 5 RDVYRRVMKE FASVVM SLQE NTPKPQNASA YRRKRG, which is translated
 from the genomic sequence (called nudH/ygdP, which may be located at:
 (www)ncbi.nlm.nih.gov/nuccore/NC_000913.3?report=fasta&from=2968447&to=29
 68977)

ACCTCTTTTACGTCGATAAGCAGATGCGTTTTGTGGTTTTGGCGTATTTTCC
 10 TGCAGTGACATCACCACACTCGCGAACTCTTTCATTACCCTACGGTAGACA
 TCACGTTTAAATGACACCACCTGTCTGACCGGATAACCAGTAACTTACCCAT
 CGCCAGCCGTCAAACCTCTGGTGTACTGCTGGTTTGCATATTGATTTCTGCA
 TCGCCGCTCACCAGCTGCAAGAGAAACCATTTTTGTTTTTGGCCGATACAA
 ACCGGCTTCGTGTCCCAACGCACCAAACGTTTCGGTAATTTGTAGCGCAAC
 15 CAGTTACGCGTTGAAGCAAGGATTCTGAACGTCTTTGCGGCTTAATCCTACT
 TCTTCAAACAATTCACGGTACATCGCCTGCTCTGCGGATTCTCCGGGGTTG
 ATTCCGCCTTGCGGAAATTGCCAGGAGTGCTGACCAAATCGCCGGGGCCCA
 CATTACCTGCCCCCTGGCGATTACAAATCACGATACCTACGTTTGGGCGGTA
 GCCATCGTCATCAATCAC

20 By “*HsNudT16*” is meant a polypeptide having the following sequence:

MAGARRLELGEALALGSGWRHACHALLYAPDPGMLFGRIPRLYAILMQMRF
 DGRLGFPGGFVDTQDRSLEDGLNRELREELGEAAAFRVERTDYRSSHVGS
 PRVVAHFYAKRLTLEELLA VEAGATRAKDHGLEVLGLVRVPLYTLRDGVGG
 LPTFLENSFIGSAREQLLEALQDLGLLQSGSISGLKIPAAH

25 The term “subject” as used herein is meant to include vertebrates, preferably a
 mammal. Mammals include, but are not limited to, humans, camels, horses, goats,
 sheep, cows, dogs, cats, and the like.

Ranges provided herein are understood to be shorthand for all of the values
 within the range. For example, a range of 1 to 50 is understood to include any
 30 number, combination of numbers, or sub-range from the group consisting 1, 2, 3, 4, 5,
 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29,
 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 as
 well as all intervening decimal values between the aforementioned integers such as,

for example, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, and 1.9. With respect to sub-ranges, “nested sub-ranges” that extend from either end point of the range are specifically contemplated. For example, a nested sub-range of an exemplary range of 1 to 50 may comprise 1 to 10, 1 to 20, 1 to 30, and 1 to 40 in one direction, or 50 to 40, 50 to 30, 50 to 20, and 50 to 10 in the other direction.

As used herein, the terms “treat,” “treating,” “treatment,” and the like refer to reducing or ameliorating a disorder and/or symptoms associated therewith. It will be appreciated that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be completely eliminated.

Unless specifically stated or obvious from context, as used herein, the term “or” is understood to be inclusive. Unless specifically stated or obvious from context, as used herein, the terms “a,” “an,” and “the” are understood to be singular or plural.

Unless specifically stated or obvious from context, as used herein, the term “about” is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About may be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from context, all numerical values provided herein are modified by the term about.

The recitation of a listing of chemical groups in any definition of a variable herein includes definitions of that variable as any single group or a combination of listed groups. The recitation of an embodiment for a variable or aspect herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

Any compositions or methods provided herein may be combined with one or more of any of the other compositions and methods provided herein.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-1G show an improved protocol for the purification of snake venom phosphodiesterase I (SVP) for the digestion of protein-conjugated PAR. SVP was affinity purified on a blue sepharose column (boxed fractions in Figure 1A) producing a protease-contaminated product (Figure 1B and Figure 1C, proteolysis is evident by the loss of protein bands following exposure to SVP). The addition of

SVP results in protein degradation within lysates (Figure 1B) and of purified PARP1 (Figure 1C). The rightmost section in Figure 1B indicates the size of SVP from the pooled fraction from (Figure 1A). For (Figure 1B) 1 mg of whole cell lysate was exposed to 50 pmoles of SVP for 2 hrs at 37°C, pH 7. For (Figure 1C) 10 pmoles of PARP1 was exposed to 5 pmoles of SVP for 2 hrs at 37°C, pH 7. Size exclusion chromatography (Figure 1D and Figure 1E) results in a purified SVP fraction (Figure 1F) that removed the contaminating protease activity (Figure 1G). Top sections in Figure 1C and Figure 1G show the autoradiographs of PARylated PARP1 and the bottom sections show the corresponding coomassie staining of the gels. Reaction conditions for (Figure 1G) were identical to those of (Figure 1C). ***SDS-PAGE well, **SDS-PAGE interface between stacking and resolving gels, *native size of PARP1; rep = replicate.

Figures 2A-2I show Nudix ADPrases are ineffective against protein-conjugated ADP-ribose. Figure 2A, 2B and 2C are autoradiographs showing ³²P-labeled mono- or poly(ADP-ribose) conjugated to PARP1 following exposure to canonical sugar hydrolases: *EcNudF*, *EcYfcD* and *EcNudE* and *Bd3179*. For (Figure 2A), 5 pmoles of PARylated PARP1 was exposed to 10 pmoles of hydrolase for 2 hrs at 37°C; for (Figure 2B), 100 pmoles of hydrolases were utilized, while in (Figure 2C) 5 pmoles of PARP1 E988Q, a mutant only capable of synthesizing mono(ADP-ribose), was exposed to 100 pmoles of hydrolases for 2 hrs at 37°C. Ribbon diagrams show the structure of each of the enzymes used: (Figure 2D) *EcNudF*/ADPrase (PDB ID 1KHZ) (Gabelli et al. Nat. Struc. Bio. 8, 467-472, (2001)), (Figure 2E) *EcYfcD* (PDB ID 2FKB), (Figure 2F) *EcNudE* (PDB ID 1VHG) and (Figure 2G) *Bd3179* (PDB ID 5C7Q) (Wolff et al. PLoS pathogens 11, (2015)). Figure 2H shows a surface representation of *EcNudF* with modeled PAR polymer based on the complex with a nonhydrolyzable ADPr (Gabelli, S. B. et al. Biochem. 41, 9279-9285 (2002)). Figure 2I shows a surface representation of *EcNudF* with the modeled PAR polymer, the inset shows how the terminal ribose is protected and buried within the enzyme. The white arrowhead indicates where protein conjugation would occur on ADPr. ***SDS-PAGE well, **SDS-PAGE interface between stacking and resolving gels, *native size of PARP1, ^co-purified PARP1 protein fragments.

Figure 3 shows sizing chromatography of the Nudix enzyme *EcRppH*. *EcRppH* size exclusion chromatography uses 50 mM Tris-HCl pH 7.5, 1 mM EDTA,

0.30 M NaCl at a flow rate of 2.0 mL/min; peak at 259.4 mL on a HiPrep 26/60 Superdex 200 prep grade. Molecular weight standards thyroglobulin Mr 670,000, peak 140.6 mL; γ -globulin Mr 158,000, peak 184.5 mL; ovalbumin Mr 44,000, peak 226.2 mL; myoglobin Mr 17,000, peak 253.4 mL; vitamin B12 Mr 1350, peak 297.3.

5 **Figures 4A-4D** show single domain Nudix hydrolases *EcRppH* and *Dr1184* show activity against protein-conjugated ADPr. Figure 4A and Figure 4B show autoradiographs showing the ^{32} P-labeled PAR (Figure 4A) or MAR (Figure 4B) conjugated to PARP1 following exposure to Nudix hydrolases: *ArORF147*, *Pa3470*, *Dr1184*, and *EcRppH*. For (Figure 4A), 5 pmoles of PARylated PARP1 was exposed to either 20 or 100 pmoles of hydrolase for 2 hrs at 37°C. For (Figure 4B), 5 pmoles of MARylated PARP1 E988Q mutant was exposed to 100 pmoles of hydrolase for 2 hrs at 37°C. Figure 4C and Figure 4D show surface representations of *Dr1184*/CoAse (teal, PDB ID 1NQY) (Figure 4C) and *EcRppH* (green, PDB ID 4S2Y) (Figure 4D) with PAR modeled into their active sites based on the orientation of capped mRNA bound to *EcRppH* in the reported structure (Vasilyev & Serganov. J. Bio. Chem. 290, 9487-9499 (2015)). A dark shadow shows the possible orientation of the attached PARylated protein. ***SDS-PAGE well, **SDS-PAGE interface between stacking and resolving gels, *native size of PARP1, ^ co-purified PARP1 protein fragments.

20 **Figures 5A-5H** show *HsNudT16* degrades protein-conjugated ADPr. Figure 5A shows a ribbon model of the structure of *Dr1184*/CoAse (teal, PDB ID 1NQY). Figure 5B shows a ribbon model of the structure of *EcRppH* (green, PDB ID 4S2Y). Figure 5C shows a ribbon model of *HsNudT16* (PDB ID 2xSQ), Figure 5D shows a surface representation of *HsNudT16* with PAR modeled (sticks) and a protein depicted at the conjugation site of ADPr. Figure 5E shows a structural alignment between *HsNudT16* and *EcRppH*, while Figure 5F shows a structural alignment of *HsNudT16*, *EcRppH* and *Dr1184*. The grey hexagon in Figures 5A, Figure 5B and Figure 5E show the helix that moves upon substrate binding. Figures 5G and Figure 5H show the removal of ^{32} P-labeled ADPr from 5 pmoles of PARylated (Figure 5G) or MARylated (Figure 5H) PARP1 by increasing amounts of SVP (0.1, 1 and 10 pmoles) or *Dr1184*, *EcRppH* or *HsNudT16* (50, 250 and 750 pmoles). ***SDS-PAGE well, **SDS-PAGE interface between stacking and resolving gels, *native size of PARP1.

Figures 6A-6C shows treatment of PARylated PARP1 protein with SVP, *EcRppH* and *HsNudT16* results in the identification of phosphoribosylation sites by mass spectrometry. Figure 6A shows E491 (red, identified as 212 Dalton mass shift) on PARP1 was identified as a PARP1 automodification site following SVP, *EcRppH* and *HsNudT16* treatment. Figure 6B shows E190 and K192 were identified as PARP1 automodification sites on PARP1 following *EcRppH* digestion. Figure 6C shows identification of R452 as a PARP1 automodification site following *HsNudT16* treatment.

Figure 7A and 7B show alignment of Nudix enzymes with activity against protein-conjugated PAR reveals a novel 3_{10} helix. Figure 7A shows alignment of the sequences of *EcRppH*, *HsNudT16* and *Dr1184*. The red background indicates sequence identity, red letters sequence indicates conservation of charge, and white background indicates non-homologous residues. The secondary structure of each enzyme is shown in blue. The box indicates a short 3_{10} helix not observed in other Nudix enzymes such as the ADPRases. The secondary structure elements are more conserved between the two enzymes with activity towards PAR and MAR (*EcRppH* and *HsNudT16*) compared to *Dr1184* that may only hydrolyze PAR. Figure 7B shows surface representation of the model of *EcRppH* with PAR modeled and minimized using the software package MOE. The diphosphate bond of the terminal ADP-ribose attached to a protein is in the active site, well positioned to be hydrolyzed and released as a phosphoribosylated protein product. The scissile bond is marked with an arrow.

Figure 8 shows an alignment of NudT16 enzyme sequences across human, bovine, sheep, murine, and xenopus species.

Figures 9A-9C show alignments of RppH enzyme sequences across bacterial species. Figure 9A depicts alignments of the N terminal sequences corresponding to the $\beta 1$ - $\eta 1$ structures for bacterial species. Figure 9B depicts alignments of the sequences corresponding to the $\beta 5$ - $\eta 3$ structures for bacterial species. Figure 9C depicts alignments of the C terminal sequences corresponding to the $\beta 8$ - $\alpha 3$ structures for bacterial species.

DETAILED DESCRIPTION OF THE DISCLOSURE

The present disclosure features methods that are useful for identifying sites of post-translational modification (PTMs) such as, for example, ADP-ribosylation. The disclosure is based, at least in part, upon the discovery that ADP-ribosylation of
5 certain proteins may be detected by using specific nudix hydrolases in combination with mass spectrometry. For example, the present disclosure provides methods of detecting sites of ADP-ribosylated tagged proteins generated with *Ec* (and other bacterial) RppH nudix hydrolases. According to the techniques herein, Nudix hydrolases (e.g., *Ec*RppH, *Hs*NudT16, other RppH from bacteria, and the like) may
10 be used to digest ADP-ribosylated proteins in preparation for analysis with mass spectrometric techniques. Based on analyses described herein, other sequence or structural homologs of RppH single domain Nudix hydrolases in other species may be used for this purpose (including mammalian homologues such as *Hs*NudT16) as explained in Fig. 5E and 5F. In another aspect, the present disclosure provides an
15 improved protocol for the purification of snake venom phosphodiesterase I (SVP).

Study of ADP-ribosylation (ADPr) has been limited by a lack of mass spectrometry-based proteomic tools. Recent work has demonstrated the potential of a tag-based pipeline in which the ADPr monomer or polymer is simplified down to its phosphoribose protein attachment site, leaving a 212 Dalton tag at the site of
20 modification. While the pipeline has been proven effective by multiple groups, a barrier to application has become evident: the enzyme used to transform ADPr into phosphoribose — snake venom phosphodiesterase (SVP) from the rattlesnake *Crotalus adamanteus* — must be purified from venom (no recombinant expression/purification scheme has been shown) and may be tedious to isolate away
25 from all contaminating phosphatases and proteases. According to the techniques herein, an improved protocol for successfully purifying SVP for use in this pipeline is outlined, as well as alternatives to SVP such as, for example, hydrolases from the Nudix hydrolase super family: mammalian Nudt16 (e.g., human Nudt16, product of the gene NUD16) and *E. coli* RppH. Importantly, expression and purification
30 schemes for these hydrolase enzymes have been proven to result in large, high quality yields of functional enzyme. The techniques herein demonstrate the utility in identifying ADP-ribosylation sites on Poly(ADP-ribose) Polymerase 1 (PARP1) by

the combined use of hydrolases and mass spectrometry. A structure-based rationale for the ability of these nudix hydrolases, but not canonical nudix ADPrases, to degrade protein-conjugated ADPr is also offered.

As described above, the *Escherichia coli* RppH (and other sequence or structural homologs of RppH in any species) enzyme may be used for the purpose of studying ADP-ribosylation using mass spectrometry. *Escherichia coli* RppH is an enzyme that belongs to the Nudix hydrolase superfamily, a group of enzymes that catalyzes hydrolysis of **N**ucleoside **D**iphosphates linked to other moieties (“**X**”). It was not previously shown that *E. coli* RppH may hydrolyze protein-conjugated ADP-ribose. Since this enzyme and its derivatives may be purified in large quantities, these findings allow its use in mass spectrometry to identify ADP-ribosylation sites. Derivatives that make the enzymes more efficient and stable may be made by modification based on the structural data described herein. In addition, enzymes from similar classes could also potentially be used for this purpose.

ADP-ribosylation

ADP-ribosylation refers to the transfer of the ADP-ribose group from NAD⁺ to target proteins post-translationally, either attached singly as mono(ADP-ribose) (MAR) or in polymeric chains as poly(ADP-ribose) (PAR). ADP-ribosylation is a post-translational modification (PTM) implicated in a number of disease states, including cancer, diabetes, and a range of neuropathologies (Curtin, N. J. & Szabo, C. Mol. aspects of med. 34, 1217-1256 (2013)). This protein modification is synthesized by ADP-ribosyl transferases, commonly known as poly(ADP-ribose) polymerases (PARPs), which transfer the ADP-ribose (ADPr) group from NAD⁺ to protein acceptor amino acids in monomeric (mono(ADPr), MAR) and/or polymeric (poly(ADPr), PAR) form (Hottiger et al. Trends Biochem. Sci. 35, 208-219 (2010)), (Vyas et al. Nat. comm. 5, 4426 (2014)). Notably, at least ten Phase III clinical trials for PARP inhibitors are ongoing for cancer patients, with one drug already approved by the European Medicine Agency as well as the Food and Drug Administration in the US. The inhibitor has already shown efficacy in ovarian, prostate, breast, pancreas, lymphoma and many other solid or blood cancers. Identification of specific amino acid acceptors of ADPr group(s), and therefore characterization of the cellular role played by this important protein modification, has been hampered by the lack of a

robust, universal method for identifying ADP-ribosylation sites in the proteome. This need has lately been addressed by three different methods (Zhang et al. Nat. Methods 10, 981-984 (2013)) (Daniels et al. J. Proteome Res. doi:10.1021/pr401032q (2014)) (Chapman et al. J. Proteome Res. doi:10.1021/pr301219h (2013)) (Hengel et al. J. Am. Soc. Mass Spec. 20, 477-483 (2009)) (Oetjen et al. FEBS J. 276, 3618-3627 (2009)) (Tao et al. Biochem. 48, 11745-11754 (2009)) (Messner, S. et al. Nuc. Acids Res. 38, 6350-6362 (2010)) (Rosenthal, F. et al. Methods Mol. Bio. 780, 57-66 (2011)) (Hengel & Goodlett. Internat. J. Mass Spect. 312, 114-121(2012)) (Palazzo, L. et al. Biochemical J., doi:10.1042/BJ20141554 (2015)) (Daniels et al. Mol. cell 58, 911-924 (2015)), all of which involve the removal of any ADPr subunits beyond the protein-proximal monomer, followed by identification of the 'tag' left behind at the ADPr conjugation site, as reviewed in Daniels et al. Mol. Cell 58, 911-924 (2015).

One of these methods, hydrolysis of MAR/PAR down to its phosphoribose attachment site, relies upon the pyrophosphatase activity of snake venom phosphodiesterase I (SVP) from *Crotalus adamanateus*, which may be purchased in a partially purified form that requires further purification for use against ADP-ribosylated proteins (Vyas, S. et al. Nat. comm. 5, 4426 (2014)), (Daniels et al. J. of Proteome Res., doi:10.1021/pr401032q (2014)), (Chapman et al. J. Proteome Res. doi:10.1021/pr301219h (2013)). Unfortunately, this complicated purification scheme ultimately results in a high level of prep-to-prep variability, likely due to the inherently variable protein source (snake venom) as well as the number of purification steps involved. In an effort to identify a more reliable tool for the degradation of protein-conjugated MAR/PAR to phosphoribose, described herein is the characterization of candidate enzymes from the Nudix hydrolase superfamily.

The Nudix hydrolase superfamily catalyzes hydrolysis of Nucleoside Diphosphates linked to other moieties ("X"). Most Nudix families contribute to cellular 'housekeeping' through the breakdown of a wide range of nucleoside diphosphate derivatives (Bessman et al. J. Bio. Chem. 271, 25059-25062 (1996)). One of these diphosphate containing compounds is ADPr (Mildvan, A. S. et al. Arch. Biochem. Biophys. 433, 129-143 (2005)) (Tong et al. J. Biol. Chem. 284, 11256-11266 (2009)) (Gabelli et al. Nat. Struc. Bio. 8, 467-472 (2001)) (Dunn et al. J. Biol. Chem. 274, 32318-32324 (1999)), a molecule which is known to accumulate in cells with potentially cytotoxic effects by: (1) altering calcium entry into cells via channel

gating, thus affecting membrane depolarization (Perraud, A. L. et al. Nat. 411, 595-599 (2001)), (2) serving as a neurotransmitter in primate and murine colons (Durnin et al. J. Physiol. 590, 1921-1941 (2012)), and (3) spontaneously modifying proteins (Jacobson et al. Mol. Cell. Biochem. 138, 207-212 (1994)), potentially altering

5 intracellular post-translational signaling. Without Nudix hydrolase activity, free ADPr would amass during the breakdown of PAR (Miwa & Sugimura. J. Biol. Chem. 246, 6362-6364 (1971)), (Steffen & Pascal. EMBO J. 32, 1205-1207 (2013)), as a side product of tRNA synthesis (Shull et al. Nuc. acids Res. 33, 650-660 (2005)), following NAD⁺ glycohydrolysis (Dolle et al. FEBS J. 280, 3530-3541 (2013)),

10 following deacetylation of O-acetyl-ADPr (Peterson, F. C. et al. J. Biol. Chem. 286, 35955-35965 (2011)), or through the breakdown of the signaling molecule cyclic ADPr (Kim et al. Science 261, 1330-1333 (1993)). Accordingly, ADPr degrading Nudix enzymes are broadly conserved, with humans possessing at least six distinct ADPr pyrophosphatases (ADPrases) responsible for hydrolyzing ADPr to AMP and

15 phosphoribose (Gabelli et al. Nat. Struct. Bio. 8, 467-472, (2001), McLennan, A. G. Cellular and molecular life sciences : CMLS 63, 123-143 (2006)).

According to the techniques herein, Nudix hydrolases may replace SVP for generating phosphoribose tags at ADP-ribosylation sites in the proteomics pipeline described above. Bacterial Nudix hydrolases were screened for comparable

20 hydrolysis activity of protein-conjugated ADPr. These efforts lead to the identification of the RNA 5' pyrophosphohydrolase (RppH) from *Escherichia coli* (*EcRppH*) and other homologs as capable of degrading protein-conjugated ADPr, including both MAR and PAR. From a structural and biological perspective, this finding was unexpected as *EcRppH* is an RNA decapping enzyme, and not an

25 ADPrase (Deana et al. Nature 451, 355-358 (2008)). In this disclosure, a structure-based rationale for the inability of Nudix ADPrases to degrade protein-conjugated ADPr, in contrast to Nudix RNA decapping enzymes is provided. Furthermore, the use of both *EcRppH* and *HsNudT16* in the identification of protein ADP-ribosylation sites by mass spectrometry is demonstrated herein.

30 Mono(ADP-ribose) may be detected by mass spectrometry directly, poly(ADP-ribose) may be detected by cleaving PAR into MAR using the enzyme poly(ADP-ribose) glycohydrolase (PARG). Alternatively a method was proposed by

Zhang et al in 2013 wherein MAR and PAR were chemically removed, leaving a mass aberration of 15 daltons at the vacated attachment site.

Nudix Hydrolase

5 The Nudix family is a protein family of phosphohydrolases. These hydrolases use water-mediated catalysis to break a phosphate bond in their substrate to create two products. Substrates hydrolysed by Nudix enzymes comprise a wide range of organic pyrophosphates, including nucleoside di- and triphosphates, dinucleoside and diphosphoinositol polyphosphates, nucleotide sugars and RNA caps, with varying
10 degrees of substrate specificity. There are two components to the Nudix family: the so-called Nudix fold of a beta sheet with alpha helices on each side and the Nudix motif which contains catalytic and metal-binding amino acids. The Nudix motif is GXXXXX[E/D]XXXXXXX[X]REUXEEXG[U/Y] where U is Isoleucine, Leucine, or Valine and X is any amino acid and Y is E for NudT16. This forms a b-strand
15 followed by a helix motif (or loop-helix-loop) which contains the catalytic amino acids; specifically the glutamates that coordinate the catalytic metal. In other words, for NudT16 the motif is GXXXXXDXXXXXXXXREUXEEXGY while for RppH, the motif is GXXXXXEXXXXXXXXXREUXEEXGY. Nudix family enzymes include Dcp2 of the decapping complex, ADP-ribose diphosphatase, MutT, ADPRase, Ap4A,
20 RppH, and many others.

Mass Spectrometry

 Mass spectrometry (MS) is an analytical chemistry technique that helps identify the amount and type of chemicals present in a sample by measuring the mass-
25 to-charge ratio and abundance of gas-phase ions. A mass spectrum is a plot of the ion signal as a function of the mass-to-charge ratio. The spectra are used to determine the elemental or isotopic signature of a sample, the masses of particles and of molecules, and to elucidate the chemical structures of molecules, such as peptides and other chemical compounds. Mass spectrometry works by ionizing chemical compounds to
30 generate charged molecules or molecule fragments and measuring their mass-to-charge ratios. In a typical MS procedure, a sample, which may be solid, liquid, or gas, is ionized, for example by bombarding it with electrons. These ions are then separated according to their mass-to-charge ratio, typically by accelerating them and

subjecting them to an electric or magnetic field: ions of the same mass-to-charge ratio will undergo the same amount of deflection. The ions are detected by a mechanism capable of detecting charged particles, such as an electron multiplier. Results are displayed as spectra of the relative abundance of detected ions as a function of the mass-to-charge ratio. The spectrum that results is referred to as the MS1. In order to obtain an MS2, a single intact molecular ion is selected from the MS1 spectrum, 'trapped' in the ion trap, fragmented (e.g., through collision with an inert gas), and the mass-to-charge ratios of these fragments is then determined and plotted on the MS2 spectrum, which allows the researcher to sequence the peptide including its modifications (e.g., ADPr). The atoms or molecules in the sample may be identified by correlating known masses to the identified masses or through a characteristic fragmentation pattern.

A mass spectrometer consists of at least: an ion source and a mass analyzer, serving as a detector. The ionizer converts a portion of the sample into ions. There is a wide variety of ionization techniques, depending on the phase (solid, liquid, gas) of the sample and the efficiency of various ionization mechanisms for the unknown species. The most basic mass spectrometer (1) ionizes the molecules (e.g., for peptides this is typically by electrospray ionization, ESI), and (2) measures the mass to charge ratio of these ions, typically through Fourier transformation of the frequency signal. In order to obtain an MS2 there is typically an ion trap (to filter out the mass-to-charge ratio of interest) and a fragmentation method such as, for example, collision induced dissociation/CID (wherein molecules are smashed into inert ions, as described above) that takes place in the ion trap. The differences in masses of the fragments allow the mass analyzer to sort the ions by their mass-to-charge ratio. The detector measures the value of an indicator quantity and thus provides data for calculating the abundances of each ion present.

Mass spectrometry has both qualitative and quantitative uses. These include identifying unknown compounds, determining the isotopic composition of elements in a molecule, and determining the structure of a compound by observing its fragmentation. Other uses include quantifying the amount of a compound in a sample or studying the fundamentals of gas phase ion chemistry (the chemistry of ions and neutrals in a vacuum). As an analytical technique, it possesses distinct advantages such as: increased sensitivity over most other analytical techniques because the

analyzer, as a mass-charge filter, reduces background interference; excellent specificity from characteristic fragmentation patterns to identify unknowns or confirm the presence of suspected compounds; information about molecular weight; information about the isotopic abundance of elements; and temporally resolved chemical data. One disadvantage of the method is that it often fails to distinguish between optical and geometrical isomers and the positions of a substituent in o-, m- and p- positions in an aromatic ring. Also, its scope is limited in identifying hydrocarbons that produce similar fragmented ions.

Derivatization strategies for mass spectrometric analysis are commonplace and have been reviewed previously [Knapp, D. R. *Methods Enzymology* 1990, 193, 314-329.; Anderegg, R. J. *Mass Spectrom. Rev.* 1988, 7, 395-424.; Roth, K. D. W., Huang, Z-H., Sadagopan N, and Watson J. T. *Mass Spectrom. Rev.* 1998, 17, 255-274.; Sadagopan, N. and Watson J. T. *J. Am. Soc. Mass. Spectrom.* 2001, 12, 399-409.; Jones, M. B., Jeffrey, W. A., Hansen, H. F., Pappin, D. J. C. *Rapid Commun. Mass Spectrom.* 1994, 8, 737-42.; Spengler, B., Luetzenkirchen, F., Metzger, S., Chaurand, P., Kaufmann, R., Jeffery, W., Bartlett- Jones, M. and Pappin, D. J. C. *Int. J. Mass Spectrom. Ion Proc.* 1997, 169/170, 127-140.; Keogh, T., Lacey, M. P., and Youngquist, R. S. *Rapid. Commun. Mass Spectrom.* 2000, 14, 2348.].

Analysis of the digested peptides may be by any mass spectrometry-based method that allows high-throughput multiplexed analysis. Mass spectrometry is a sensitive and accurate technique for identifying molecules. Generally, mass spectrometers have two main components, an ion source for the production of ions and a mass-selective analyzer for measuring the mass-to-charge ratio of ions, which may then be converted into a measurement of mass for these ions. Several ionization methods are known in the art and described herein.

Different mass spectrometry methods, for example, quadrupole mass spectrometry, ion trap mass spectrometry, time-of-flight mass spectrometry and tandem mass spectrometry may utilize various combinations of ion sources and mass analyzers which allows for flexibility in designing customized detection protocols. Furthermore, a mass spectrometer may be programmed to select ions of a particular mass for transmission into the mass spectrometer while blocking other ions. The ability to precisely control the movement of ions in a mass spectrometer allows for greater options in detection protocols which may be advantageous when a large

number of peptides, for example, from a multiplex experiment, are being analyzed. Mass spectrometry methods are well known in the art (see Burlingame et al. Anal. Chem. 70:647 R-716R (1998); Kinter and Sherman, Protein Sequencing and Identification Using Tandem Mass Spectrometry Wiley-Interscience, New York (2000)). The basic processes associated with a mass spectrometry method are the generation of gas-phase ions derived from the sample, and the measurement of their mass-to-charge ratio. When coupled with liquid chromatography, mass spectrometry technology exists by which several thousands of peptide or protein species may be separated, detected and quantified in a single operation.

10 The mass spectrometry may be preceded by a chromatography step. New chromatography based methods for the identification of the proteins contained in complex mixtures without the need for separation of the mixture into individual protein components are available. A separation step may also be used to remove salts, enzymes, or other buffer components. Several methods well known in the art, such as
15 chromatography, gel electrophoresis, or precipitation, may be used to suitably purify the sample prior to the introduction to the mass spectrometer. For example, reversed phase chromatography may be used to remove salt from a sample as it may separate peptides based on their hydrophobicity when the buffer is gradually changed over from aqueous to organic. The choice of separation method may depend on the
20 amount of a sample. For example, when small amounts of sample are available or a miniaturized apparatus is used, a micro-chromatography separation step may be used. In addition, whether a separation step is desired, and the choice of separation method, may depend on the detection method used. For example, the efficiency of matrix-assisted laser desorption/ionization and electrospray ionization may be improved by
25 removing salts from a sample. For example, salts may absorb energy from the laser in matrix-assisted laser desorption/ionization and result in lower ionization efficiency.

 Any type of mass spectrometer may be used with the methods and systems described herein, including, but not limited to, spectrometers capable of liquid chromatography-mass spectrometry (LC/MS), or liquid chromatography-tandem mass
30 spectrometry (LC/MS/MS). Exemplary spectrometers useful in connection with the methods disclosed herein include, among others, the Thermo Q-Exactive series, Thermo Orbitrap series, Thermo LTQ series, Thermo TSQ series, AB SCIEX models 4000, 5500, 5600; Waters Xevo series; or Agilent 6490.

Any of the methods disclosed herein may be further automated by use of a robotic device known in the art. Specifically, the steps of derivatization and proteolytic digestion may be automated by use of a robotic device known in the art. In other embodiments, the steps of derivatization and proteolytic digestion may be
5 parallelized by use of multi-chamber reaction vessels that are compatible with the robotic device.

Neoplasia/Cancer

In some embodiments, methods of the present disclosure may be used to
10 detect or screen for ADP-ribosylation associated neoplasia or cancer. Neoplasia comprises a large family of diseases that involve abnormal cell growth with the potential to invade or spread to other parts of the body. They form a subset of neoplasms. A neoplasm or tumor is a group of cells that have undergone unregulated growth, and will often form a mass or lump, but may be distributed diffusely. Six
15 characteristics are typical of cancer: self-sufficiency in growth signaling; insensitivity to anti-growth signals; evasion of apoptosis; enabling of a limitless replicative potential; induction and sustainment of angiogenesis; and activation of metastatic invasion of tissue. The progression from normal cells to cells that may form a discernible mass to outright cancer involves multiple steps known as malignant
20 progression.

The methods described herein are useful in detecting and/or treating various types of malignancies and/or tumors, e.g., non-Hodgkin's lymphoma (NHL), acute lymphocytic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), chronic myelogenous leukemia (CML), multiple myeloma (MM),
25 breast cancer, ovarian cancer, head and neck cancer, bladder cancer, melanoma, colorectal cancer, pancreatic cancer, lung cancer, leiomyoma, leiomyosarcoma, glioma, and glioblastoma. Solid tumors include, e.g., breast tumors, ovarian tumors, lung tumors, pancreatic tumors, prostate tumors, melanoma tumors, colorectal tumors, lung tumors, head and neck tumors, bladder tumors, esophageal tumors, liver tumors,
30 and kidney tumors.

Neuropathologies

In some embodiments, methods of the present disclosure may be used to detect or screen for ADP-ribosylation associated neuropathologies. A number of neuropathologies have been associated with aberrations in ADP-ribosylation. Experimental studies indicate that over-activation of the DNA repair protein poly(ADP-ribose) polymerase (PARP) in response to oxidative damage to DNA may cause cell death due to depletion of NAD⁺. Oxidative damage to DNA and other macromolecules has been reported to be increased in the brains of patients with Alzheimer's disease. There is enhanced PARP activity in Alzheimer's disease and pharmacological interventions aimed at inhibiting PARP may have a role in slowing the progression of the disease.

PARP over-activation under pathologic conditions including traumatic brain injury (TBI) results in cell death. Furthermore, PARP inhibitors block nitric acid neurotoxicity, protect against post-stroke inflammation, and provide neuroprotection following stroke; PARylated proteins may be found in the cerebrospinal fluid following TBI. Intra-mitochondrial poly-ADP-ribosylation occurs following excitotoxic and oxidative injury. As the effects of PARP activation on mitochondrial respiration appear regulated by poly(ADP-ribose) glycohydrolase, a direct effect of poly-ADP-ribosylation on electron transport chain function is suggested.

Evidence from *in vitro* studies on non-neuronal cells in culture have shown that when fully activated by free radical-induced DNA damage, PARP depletes cellular NAD⁺ and consequently adenosine triphosphate (ATP) levels, and that this depletion is associated with a cell death that may be prevented by PARP inhibitors. An involvement of PARP in the control of brain energy metabolism during neurotoxic insult supports the participation of PARP in MPTP-induced neurotoxicity and suggests that PARP inhibitors might be beneficial in the treatment of Parkinson's disease. In fact, PARP levels may be increased in neurons of Parkinson's patients.

Animal and *in vitro* studies suggest that overactivation of poly(ADP-ribose) polymerase (PARP) in response to oxidative DNA damage makes a substantial contribution to cell death after brain ischaemia. Global brain ischaemia due to cardiac arrest induces a rapid increase in the amount of neuronal and glial PARP that may be detected by immunohistochemistry. Global brain ischaemia causes accumulation of poly(ADP-ribose) after cardiac arrest. The distribution of cells with accumulation of poly(ADP-ribose) corresponds in general to regions of ischaemic damage or

immediately adjacent neocortex. There is a role for PARP in post-ischaemic necrosis. There is a potential for reducing ischaemic brain damage by the use of PARP inhibitors.

Neuropathologies include neurodegeneration; the umbrella term for the
5 progressive loss of structure or function of neurons, including death of neurons. In some embodiments, neuropathological or neurodegenerative diseases including amyotrophic lateral sclerosis (PARP is activated in spinal cord oligodendrocytes of ALS patients), Parkinson's, Alzheimer's (nuclear proteins from Alzheimer's patients are highly PARylated), multiple sclerosis (PARP inhibition attenuates demyelination
10 and oligodendrocyte depletion in MS lesions), and Huntington's occur as a result of neurodegenerative processes impacted by ADP-ribosylation.

Several neurodegenerative diseases are classified as proteopathies because they are associated with the aggregation of misfolded proteins. One such protein is alpha-synuclein which may aggregate to form insoluble fibrils in pathological
15 conditions characterized by Lewy bodies, such as Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy. Alpha-synuclein is the primary structural component of Lewy body fibrils. In addition, an alpha-synuclein fragment, known as the non-Abeta component (NAC), is found in amyloid plaques in Alzheimer's disease. Another protein is tau; when hyperphosphorylated, tau protein is the main
20 component of neurofibrillary tangles in Alzheimer's disease. Also, beta amyloid is the major component of senile plaques in Alzheimer's disease.

Parkinson's disease (PD) and Huntington's disease are both late-onset and associated with the accumulation of intracellular toxic proteins. Diseases caused by the aggregation of proteins are known as proteinopathies, and they are primarily
25 caused by aggregates in the following structures: cytosol (e.g. Parkinson's & Huntington's); nucleus (e.g. Spinocerebellar ataxia type 1 (SCA1)); endoplasmic reticulum (ER), (as seen with neuroserpin mutations that cause familial encephalopathy with neuroserpin inclusion bodies); and extracellularly excreted proteins - amyloid- β in Alzheimer's disease.

30 Alzheimer's disease is characterized by loss of neurons and synapses in the cerebral cortex and certain subcortical regions. This loss results in gross atrophy of the affected regions, including degeneration in the temporal lobe and parietal lobe, and parts of the frontal cortex and cingulate gyrus. Alzheimer's disease has been

hypothesized to be a protein misfolding disease (proteopathy), caused by accumulation of abnormally folded A-beta and tau proteins in the brain. Plaques are made up of small peptides, 39–43 amino acids in length, called beta-amyloid (also written as A-beta or A β). Beta-amyloid is a fragment from a larger protein called amyloid precursor protein (APP), a transmembrane protein that penetrates through the neuron's membrane. APP is critical to neuron growth, survival and post-injury repair. In Alzheimer's disease, an unknown process causes APP to be divided into smaller fragments by enzymes through proteolysis. One of these fragments gives rise to fibrils of beta-amyloid, which form clumps that deposit outside neurons in dense formations known as senile plaques.

Parkinson's disease manifests as bradykinesia, rigidity, resting tremor and posture instability. Parkinson's disease is a degenerative disorder of the central nervous system. It results from the death of dopamine-generating cells in the substantia nigra, a region of the midbrain; the cause of cell-death is unknown.

The mechanism by which the brain cells in Parkinson's are lost may consist of an abnormal accumulation of the protein alpha-synuclein bound to ubiquitin in the damaged cells. The alpha-synuclein-ubiquitin complex cannot be directed to the proteosome. This protein accumulation forms proteinaceous cytoplasmic inclusions called Lewy bodies. The latest research on pathogenesis of disease has shown that the death of dopaminergic neurons by alpha-synuclein is due to a defect in the machinery that transports proteins between two major cellular organelles — the endoplasmic reticulum (ER) and the Golgi apparatus. Certain proteins like Rab1 may reverse this defect caused by alpha-synuclein in animal models.

Diabetes

In some embodiments, methods of the present disclosure may be used to detect or screen for ADP-ribosylation associated diabetes. Diabetes mellitus (DM), commonly referred to as diabetes, is a group of metabolic diseases in which there are high blood sugar levels over a prolonged period. Symptoms of high blood sugar include frequent urination, increased thirst, and increased hunger. If left untreated, diabetes may cause many complications. Acute complications include diabetic ketoacidosis and non-ketotic hyperosmolar coma. Serious long-term complications include cardiovascular disease, stroke, chronic kidney failure, foot ulcers, and damage

to the eyes. Diabetes is due to either the pancreas not producing enough insulin or the cells of the body not responding properly to the insulin produced. There are three main types of diabetes mellitus. Type 1 DM results from the pancreas's failure to produce enough insulin. This form was previously referred to as “insulin-dependent diabetes mellitus” (IDDM) or “juvenile diabetes”. The cause is unknown. Type 2 DM begins with insulin resistance, a condition in which cells fail to respond to insulin properly. As the disease progresses a lack of insulin may also develop. This form was previously referred to as “non-insulin-dependent diabetes mellitus” (NIDDM) or “adult-onset diabetes”. The primary cause is excessive body weight and not enough exercise. Gestational diabetes is the third main form and occurs when pregnant women without a previous history of diabetes develop a high blood-sugar level.

Diabetic patients frequently suffer from retinopathy, nephropathy, neuropathy and accelerated atherosclerosis. The loss of endothelial function precedes these vascular alterations. Activation of poly(ADP-ribose) polymerase (PARP) is an important factor in the pathogenesis of endothelial dysfunction in diabetes. Destruction of pancreatic islet cells have been shown to induce hyperglycemia, intravascular oxidant production, DNA strand breakage, PARP activation and a selective loss of endothelium-dependent vasodilation. PARP inhibition prevents pancreatic islet cell lysis. Also, a PARP1 knockout mouse is completely protected from diabetes. Furthermore, PARP inhibition reverses endothelial dysfunction, as well as microvascular complications from diabetes.

The following examples are offered by way of illustration, not by way of limitation. While specific examples have been provided, the above description is illustrative and not restrictive. Any one or more of the features of the previously described embodiments may be combined in any manner with one or more features of any other embodiments in the present disclosure. Furthermore, many variations of the disclosure will become apparent to those skilled in the art upon review of the specification. The scope of the disclosure should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the appended claims along with their full scope of equivalents.

It should be appreciated that the disclosure should not be construed to be limited to the examples that are now described; rather, the disclosure should be

construed to include any and all applications provided herein and all equivalent variations within the skill of the ordinary artisan.

The practice of the present disclosure employs, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are well within the purview of the skilled artisan. Such techniques are explained fully in the literature, such as, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984); "Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller and Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991). These techniques are applicable to the production of the polynucleotides and polypeptides of the disclosure, and, as such, may be considered in making and practicing the disclosure. Particularly useful techniques for particular embodiments will be discussed in the sections that follow.

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the assay, screening, and therapeutic methods of the disclosure, and are not intended to limit the scope of what the inventors regard as their disclosure.

EXAMPLES

EXAMPLE 1: An improved protocol for obtaining phosphodiesterase I purified from *Crotalus adamanteus* involves both affinity purification and size exclusion chromatography

Snake venom phosphodiesterase I (SVP) from *Crotalus adamanteus* was shown to degrade PAR nearly 50 years ago (Nishizuka et al. J. Bio. Chem. 242, 3164-3171 (1967)) and has since proven a valuable tool for the degradation of PAR into its linear, branching and terminal subunits, a technique that yields quantitative information regarding the molecular structure of the intact polymer (Kawaichi et al. J.

Bio. Chem. 256, 9483-9489 (1981)), (Desmarais et al. Biochim. et biophys. acta 1078, 179-186 (1991)). However, the utility of this enzyme is determined by the purification scheme employed to isolate it from the large number of proteases as well as phosphatases and nucleotidases present in the *C. adamanteus* venom (Rokyta et al
5 BMC genomics 13, 312, (2012)). Oka et al. successfully isolated the phosphodiesterase activity of commercially available SVP away from the contaminating phosphatase and 5'-nucleotidase activity through affinity purification using blue sepharose, a molecule which mimics NAD⁺ and therefore interacts with the active domain of SVP (Oka et al. Biochem. Biophys. Res. Comm. 80, 841-848
10 (1978)). The results from a simplified version of this method are shown in Figure 1A, where 150 mM Potassium Phosphate pH 7.5 was used as a single step elution off of a blue sepharose column. This purification scheme paved the way for development of the quantitative method mentioned above to determine the structure of the intact polymer, but did not address the need to eliminate contaminating protease activity.
15 This protease activity may be problematic when using SVP to hydrolyze protein-conjugated ADPr, either MAR or PAR, for the purpose of creating a phosphoribose 'tag' at the otherwise ADP-ribosylated amino acid residue (Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014)), (Chapman et al. J. Proteome Res. doi:10.1021/pr301219h (2013)), (Hengel & Goodlett. Internat. J. Mass Spect. 312,
20 114-121 (2012)). Such protease activity is demonstrated in Figure 1B wherein a complex mixture of proteins is exposed to blue sepharose purified SVP, resulting in the degradation of the target proteins and the appearance of SVP along with its co-purified proteins. This proteolytic activity is further shown against purified, ³²P-PARYlated PARP1 (both native and denatured) in Figure 1C. In order to separate the
25 115 kD SVP from the major contaminating proteins (<30 kD), the blue sepharose purified product was subjected to size exclusion chromatography, yielding a simple mixture of the various glycolytic forms of SVP, based on the knowledge that most secreted proteins are glycosylated (Cao, J. et al. J. Proteome Res. 8, 662-672, (2009)), including those found in snake venom (Soares & Oliveira. Prot. Pep. Lett. 16, 913-
30 919 (2009)) (Figure 1D, 1E, and 1F). When tested against ³²P-PARYlated PARP1 as in Figure 1C, this highly pure form of SVP displayed phosphodiesterase activity without apparent proteolytic activity (Figure 1G). Similar results were seen against whole cell lysate, allowing for use of this enzyme for ADP-ribosylation site

identification by mass spectrometry (Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014)).

While the pipeline presented here is an effective method for isolating SVP from snake venom, the complexity of the purification scheme, along with the lot-to-lot variability observed from commercial sources that serve as the input for this purification, could be greatly improved upon by the availability of a recombinant, stable enzyme that could be reliably expressed, purified, and scaled to meet the often large material demands of proteomic pipelines. For this reason, the activity of SVP was compared with that of relatively small, stable, and well characterized Nudix hydrolases, which are predicted to cleave protein conjugated ADPr with similar specificity.

EXAMPLE 2: Nudix ADPrases do not hydrolyze protein-conjugated ADP-ribose

Nudix ADPrases are responsible for the breakdown of free ADPr into its phosphoribose and adenosine monophosphate subunits, thus modulating the levels of free ADPr ((Gabelli et al. Nat. Struct. Bio. 8, 467-472, (2001)), (Gabelli, S. B. et al. Biochem. 41, 9279-9285 (2002)). Nudix ADPrases were tested for hydrolase activity against protein-conjugated MAR and PAR: PARP1, an enzyme known to autoPARylate in the presence of NAD⁺, was exposed to ³²P-labeled NAD⁺ producing either ³²P-labeled PARylated (on WT PARP1) or MARylated (on the catalytically deficient PARP1 E988Q mutant) proteins to serve as substrates for hydrolysis by candidate Nudix enzymes or the positive control, SVP (Figure 2A, 2B, 2C). As Figure 2B shows, the various Nudix enzymes tested did not significantly hydrolyze ADPr from its conjugated proteins when compared to the positive control SVP. From a structural perspective, the lack of activity towards ADPr could be explained by the dimeric structure of ADPrases, where each dimer is formed by monomers of an N-terminal β -sheet domain and a C-terminal Nudix domain (Figures 2D, 2E, 2F, 2G). The N-terminal domains are swapped, creating two active sites where both monomers contribute to substrate recognition (each active site is composed of the N-terminal β -sheet of one monomer and the C-terminal Nudix domain of the other monomer). As shown in Figures 2H and Figure 2I, ADPr is nested in the active site of the ADPrase EcNudF so that the 1'-hydroxyl of the terminal ribose group is completely buried by

the protein dimer (white arrowhead in Figure 2I), preventing conjugation to another ADPr group (or a protein). This explanation could likely be extended to the other three nucleoside sugar hydrolases tested in this study as they display the same quaternary arrangement and have a high structural homology with a pairwise root mean square deviation ranging from 0.9 to 2.0 Å (Figure 2D, 2E, 2F, 2G; rmsd calculated with SSM (Krissinel & Henrick, Acta cryst. Sec. D, Biol. cryst. 60, 2256-2268, (2004))).

10 **EXAMPLE 3: Single domain Nudix hydrolases are capable of hydrolyzing protein-conjugated ADPr**

In order to determine whether Nudix enzymes with active sites more open to fit the target ADPr group bound to either a PAR polymer or protein, Nudix enzymes representative of families that are not swapped dimers and lack N- or C-terminal domain insertions were studied. Enzymes with just the Nudix fold should have a more open active site. Four Nudix enzymes known to be monomeric by gel filtration (Xu et al. J. Biol. Chem. 278, 37492-37496, (2003)) (e.g. RppH as shown in Figure 3) were chosen. Two of them (*Dr1184/CoAse* from *Deinococcus radiodurans* and *EcRppH* from *Escherichia coli*) degraded the ³²P-labeled PAR on the model protein PARP1 (Figure 4A), while only *EcRppH* showed slight activity against ³²P-labeled MAR at the tested concentration (Figure 4B). Structural analysis and modeling revealed that the active site within these enzymes could accommodate protein-conjugated ADPr (Figure 4C and Figure 4D), as opposed to the dimeric ADPrases (c.f. Figure 2D, 2E, 2F, 2G).

25 **EXAMPLE 4: Both *EcRppH* and *HsNudT16* may degrade protein-conjugated PAR and MAR to a phosphoribose tag for mass spectrometry**

A recent study by Palazzo et al (Palazzo, L. et al. Biochemical J. doi:10.1042/BJ20141554 (2015)) has revealed that *HsNudT16*, a human Nudix (deoxy)inosine diphosphatase (Abolhassani, N. et al. Nuc. Acids Res. 38, 2891-2903 (2010)), which is also known to decap small nucleolar RNAs (Ghosh et al. Mol. Cell 13, 817-828 (2004)) as well as cytoplasmic mRNAs (Song et al. Mol. cell 40, 423-432 (2010)), has the ability to degrade protein-conjugated ADPr. As shown in Figure 5A-5F *HsNudT16* has a high structural similarity to both *EcRppH* and *Dr1184* (which

showed activity against protein-conjugated ADPr, see Figure 4A and Figure 4B) and also possesses an open active site which would allow for the target ADPr to be conjugated to a protein or additional ADPr unit(s). Based on these similarities, the activity against protein-conjugated ADPr would be comparable for all three Nudix enzymes. To test this, ³²P-PARylated or MARylated PARP1 was exposed to increasing amounts of SVP, *EcRppH*, *Dr1184* or *HsNudT16*. As shown via autoradiographs in Figure 5G and Figure 5H, both *HsNudT16* and *EcRppH* are able to hydrolyze both protein-conjugated PAR and MAR. For this reason, both *EcRppH* and *HsNudT16* are potential tools for the transformation of ADPr to a phosphoribose protein tag.

To validate that *EcRppH* and *HsNudT16* are degrading protein-conjugated ADPr down to its phosphoribose attachment site, 60 pmoles of automodified WT PARP1 was treated with 120 pmoles of SVP, 3 nmoles *EcRppH* or 3 nmoles of *HsNudT16* (either 2x or 50x molar excess over PARP1, as per the activity assays shown in Figures 5A-5H) before digesting the proteins to peptides and subjecting them to phosphoenrichment on an IMAC matrix (Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014)). As shown in Table 1, phosphoribosylation sites were confidently identified by mass spectrometry on PARP1 following treatment by each enzyme, demonstrating that all three enzymes are hydrolyzing the ADPr modification down to its phosphoribose attachment site. Notable differences were observed between the site identifications that were achieved following the use of the three enzymes — only two of the 14 phosphoribosylated peptide species were found in all samples, and all enzymes produced at least one unique peptide species. Taken together, these data suggest that these enzymes have an inherent bias for PTM structure, location, or length and therefore may be used in parallel to achieve greater coverage of the ADP-ribosylated proteome. Representative spectra from these runs have been annotated in Figure 6, further confirming the evidence of phosphoribose modifications on these peptides.

Table 1: Phosphoribosylated PARP1 peptides identified by LC-MS/MS following hydrolysis of PARylated PARP1 by SVP, *EcRppH* and *HsNudT16*.

Sequence	pR Probabilities	# of pR	Mass error [ppm]	PEP			Intensity		
				SVP	<i>EcRppH</i>	<i>HsNudT16</i>	SVP	<i>EcRppH</i>	<i>HsNudT16</i>
1 66-HPDVEVDGFSE*LR-78		1	0.39	3.2E-06		2.7E-03	7.2E+05		9.5E+05
2 143-MVDPSEKPOLGMIDR-156		2	0.29	4.8E-06		1.4E-05	7.9E+05		5.5E+06
3 166-NRSELGFPEYSASQLK-182		1	1.05			4.0E-03			2.3E+07
4 183-GFSLLATD*K-192		2	-0.84	1.0E-03	9.2E-03		2.5E+05	3.8E+06	
5 208-RKGDEVDGVDEAK-221		1	0.39			2.0E-02	4.0E+05		3.0E+06
6 426-ASLQ*ISTK-433		1	0.64	1.2E-03			5		
7 448-ANIRVSEDFLODVSASTK-463		1	0.04			2.1E-09			2.3E+06
468-SLQELFLAHILSPWGAEVK-486		1	1.32			1.1E-03			4.8E+06
8 468-SLQELFLAHILSPWGAEVK-486		1	0.02		1.4E-02			7.6E+05	
468-SLQELFLAHILSPWGAEVK-486		1	0.87			2.1E-03			1.0E+06
9 487-AE*PV*VVAPR-496		1;2	0.18	1.2E-03	1.7E-05	6.9E-03	1.5E+07	1.0E+07	1.6E+07
10 487-AEPVEVAP*G*-498		1	0.72	2.4E-03		5.2E-03	8.0E+05		9.7E+05
11 638-FYPLE*IDYQDDEAVK-653		1	-0.76	4.7E-09	4.9E-09	6.8E-03	1.0E+06	1.9E+06	3.2E+06
12 637-KFYPLEIDYQDDEAVK-653		1	-1.11	1.6E-06			8.3E+05		
13 685-AMVEYEDLQK-695		1	0.19	1.4E-09			5.4E+05		
14 688-DPIDVNYE*K*-796		1;2	0.30	4.7E-05			2.8E+05		
pR probabilities: 100% = 100% 75% - 99% = 75% 25% - 74% = 25% 1% - 24% = 1% 0% = 0%									

5

Table 1 depicts phosphoribosylated peptides reported with their corresponding posterior error probabilities (PEP) and summed intensities from 60 pmoles of the model protein PARP1 after the conversion of ADPr to phosphoribose by 120 pmoles of SVP, 3 nmoles of *EcRppH* or 3 nmoles of *HsNudT16*. All elements of the table are found in the pR (DEKRC)Sites.txt output table from MaxQuant. Summed intensities include all forms of the peptide (for example, both the singly and doubly modified forms of peptides 9 and 14), PEPs are the most confident reported for each peptide, and mass errors [ppm] are the largest mass errors reported for each peptide identification included in Table 1.

15

EXAMPLE 5: Materials and Methods**PARP1 is mutagenized to the E988Q catalytically deficient mutant**

The vector pET28 6xHis-PARP1 was a gift from Dr. John Pascal and served as the template for mutagenesis into the mono (ADP-ribose) restricted mutant of PARP1, E988Q. The vector was mutagenized by combining the vector with the following reagents: 1x Pfu reaction buffer (Agilent), 0.5 ng/uL pET28 His-PARP1 template, 2.5 ng/uL primers (Forward: GACACCTCTCTACTATATAACCAGTACATTGTCTATGATATTGC, Reverse: GCAATATCATAGACAATGTACTGGTTATATAGTAGAGAGGTGTC), 0.2 mM dNTPs (Life Technologies), 1 µl (10 units/µl) of PfuTurbo DNA polymerase (Agilent). The polymerase chain reaction (PCR) method was conducted under the following conditions: 95°C for 30 seconds (1 cycle), 95°C for 30 seconds/55°C for 60 seconds/68°C for 17 minutes (12 cycles), 68°C for 51 minutes (1 cycle). The template was digested with 1 µl (10 units/µl) of DpnI restriction enzyme (New England Biolabs) for 90 minutes at 37°C. 5µL of the digested DNA was transferred to one tube of SoloPack Gold Supercompetent Cells (Stratagene) and incubated on ice for 30 minutes, placed in a 42°C water bath for 30 seconds and then placed on ice for 2 minutes. Afterwards, 250 µL of pre-heated SOC medium (Quality Biological) was added to the reaction and incubated for 1 hr at 37°C shaking at 250 rpm. Cells were plated on LB-Kanamycin plates and incubated at 37°C overnight. Colonies were sequenced for validation.

Expression and purification of wild-type (WT) & E988Q HisPARP1

Wild type and E988Q His-PARP1 were expressed and purified, as described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014).

Purification of SVP

SVP was purified, as previously described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014), from starting material obtained from United States Biological, catalog number P4072, lot number L14030507 C14062702. Briefly, SVP powder from a vial was dissolved into 1 mL of loading buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 10% glycerol) and then loaded onto a pre-equilibrated 1 mL HiTrap blue sepharose column (GE, 17-0412-01), washed with 5 column volumes of

loading buffer and then 5 column volumes of elution buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 10% glycerol, 150 mM potassium phosphate). Desired fractions were pooled, dialyzed against loading buffer. Samples were then further dialyzed into size exclusion chromatography buffer (10 mM Tris-HCl pH 7.3, 50 mM NaCl, 15 mM
5 MgCl₂, 1% glycerol) and resolved over a SuperDex 200/10/300 GL (GE Healthcare) using an ÄKTA FPLC (GE, 18-1900-26); desired fractions were pooled and stored at -80 °C.

Assessment of contaminating proteolysis activity in SVP prep

10 For whole cell lysate, 1 mg of proteins from HeLa whole cell lysate was denatured in 8M Urea (Sigma-Aldrich) 50 mM Tris pH 7 for 10 minutes at 37°C before being reduced in 1 mM Tris- (2-Carboxyethyl)phosphine (Life Technologies) for 10 minutes and then alkylated in 2 mM 2- chloroacetamide (Sigma-Aldrich) for 10 minutes in the dark. Samples were then diluted to a final concentration of 1M Urea,
15 50 mM NaCl (Sigma-Aldrich), 15 mM MgCl₂ (Quality Biological) and 0.2M Tris pH 7.3. For each sample, 5 µg of SVP was added and incubated for 2 hours at 37°C. Samples were run on an in-house 6-10% SDS-PAGE gel and transferred to a nitrocellulose membrane. Total protein was visualized by ProAct membrane stain (Amresco) per the manufacturer's instructions. For purified recombinant His-PARP1,
20 1 µg of His-PARP1 was automodified, as described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014), and then switched into the same buffer used for the whole cell lysate (described above) with or without 1M urea. For each µg of PARP1, 500 ng of SVP was used and digestion was carried out at 37°C for 2 hours. Samples were run on in-house 6-10% SDS-PAGE gels and total protein was visualized by
25 SimplyBlue Safe Stain (Life Technologies) per the manufacturer's instructions. PAR labeled with ³²P was visualized on a phosphor-screen (GE,BAS-III 2040) followed by imaging on a Typhoon FLA7000 (GE Healthcare Life Sciences).

Expression and purification of Nudix hydrolases

30 The expression and purification have been previously published for the following hydrolases: *E. coli* NudF (EcNudF/ADPRase/UniProtKB Q93K97) (Gabelli et al. Nat. Struct. Bio. 8, 467-472, (2001)), *E. coli* NudE (EcNudE/ORF186/UniProtKB J7QI36) (O'Handley et al. J. Bio. Chem. 273, 3192-

- 3197 (1998)), *Bdellovibrio Bacteriovorus* (Bd3179/BdNDPSase/UniProtKB Q6MIH8), *Deinococcus radiodurans* Dr1184 (Dr1184/ UniProtKB Q9RV46) (Kang, L. W. et al. J. Bact. 185, 4110-4118 (2003)), *E.coli* RppH (EcRppH/ORF176/UniProtKB P0A776 (Bessman, M. J. et al. J. Bio. Chem. 276, 37834-37838 (2001)), *Agrobacterium tumefaciens* ORF147 (AtORF147/UniProtKB Q7CX66) and *Pseudomonas aeruginosa* Pa3470 (Pa3470/UniProtKB Q9HYD6)(Xu et al. J. Biol. Chem. 278, 37492-37496, (2003)). Homo sapiens NudT16 (HsNudT16/UniProtKB Q96DE0) was expressed and purified as described in the methods from the Structural Genomics Consortium
- (<http://www.thesgc.org/structures/3cou>). For *EcRppH*, the last step of purification yielded $\geq 90\%$ homogeneity before being loaded onto a HiLoad 26/60 Superdex 200 prep grade gel filtration column (GE Healthcare), equilibrated in gel filtration buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl); fractions containing $\geq 95\%$ homogeneity, as determined by SDS-PAGE, were pooled. *EcYfcD* was purified using the method described for *EcADPRase* (Gabelli et al. Nat. Struct. Bio. 8, 467-472, (2001)).

Structural analysis of Selected Nudix hydrolases

- Selected Nudix nucleotide sugar hydrolases were used to test their activity against protein-conjugated ADPr: *E. coli* ADPRase, *Mycobacterium tuberculosis* ADPRase, Homo sapiens ADPRase, *E. coli* GDPMK, and *Bdellovibrio bacteriovorus* HD100 NDPSase. The structures were structurally aligned using SSM36 and rendered with PyMOL (Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1 (2010)). PAR was constructed and minimized in MOE software package (Chemical Computing Group, Montreal, Canada) (Molecular Operating Environment (MOE) 2013.08. Chemical Computing Group Inc. (2015)). For the sugar hydrolases, PAR was modeled in the active site taking into account the binding preference observed in the structures in complex with sugar nucleoside derivatives (Gabelli et al. Nat. Struct. Bio. 8, 467-472, (2001)).

- Selected known single domain Nudix enzymes were structurally aligned using SSM (Krissinel & Henrick, Acta cryst. Sec. D, Biol. cryst. 60, 2256-2268, (2004)) and rendered using PyMOL (Molecular Operating Environment (MOE) 2013.08. Chemical Computing Group Inc. (2015)). PAR was modeled in the active site using a 'template guide', i.e. the mRNA present in the structure of the complex of *EcRppH*

(PDB ID 4S2X)(Vasilyev & Serganov. J. Bio. Chem. 290, 9487-9499 (2015)) and the IMP present in the structure of *HsNudT16* (PDB ID 2XSQ) (Tresaugues, L. et al. PloS one 10, e0131507, doi:10.1371/journal.pone.0131507 (2015)). Models of the complex structures were built using the selected structures and the PAR. Modeling steps were performed in MOE (Molecular Operating Environment (MOE) 2013.08. Chemical Computing Group Inc. (2015)). Hydrogen atoms and charges were added automatically to the protein. The binding site was defined as all protein atoms within 4.5 angstrom of any atom of bound ligand. A steepest descent energy minimization was used with a MMFF94x forcefield.

10

Automodification of WT and E988Q PARP1

Automodification was performed, as described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014), with the following changes: both WT and E988Q PARP1 were incubated with 0.6 μ M (1.85 kBq/ μ L, 37 kBq/sample) ^{32} P-NAD⁺ for 10 minutes at room temperature, following which WT PARP1 was incubated with 1 mM NAD⁺ (non-radioactive) for 10 minutes at room temperature to allow for polymer elongation.

15

Hydrolysis of protein-conjugated ADPr to phosphoribose

For comparative analyses by SDS PAGE: 5 pmoles of 6xHis-PARP1 WT or E988Q mutant were exposed to a hydrolase (various enzymes and amounts) in hydrolysis buffer (50 mM Tris-HCl pH 7.0 (Thermo Scientific), 150 mM NaCl (Sigma-Aldrich), 15 mM MgCl₂ (Quality Biological), 1 mM 3-aminobenzamide (Sigma Aldrich)) for two hours at 37°C.

20

For comparative analyses by liquid chromatography-tandem mass spectrometry (LC-MS/MS): 60 pmoles of 6xHis-PARP1 was exposed to 0.1 nmoles of SVP or 3 nmoles of nudix hydrolase (*EcRppH* or *HsNudT16*) in hydrolysis buffer for two hours at 37°C.

25

Protein digestions for LC-MS/MS analysis

Proteins were denatured in 8M Urea (Sigma Aldrich) 50 mM Tris-HCl pH 7 for 10 minutes at 37°C before being reduced in 1 mM Tris-(2-Carboxyethyl) phosphine (Life Technologies) for 10 min and then alkylated in 2 mM 2-

30

chloroacetamide (Sigma-Aldrich) for 10 minutes in the dark. Samples were diluted to: 1M Urea, 50 mM NaCl (Sigma-Aldrich), 15 mM MgCl₂ (Quality Biological), 0.2 M Tris-HCl pH 7 (7.3 at room temperature), and 1 mM CaCl₂ (Sigma-Aldrich). LysC (Wako) and Trypsin (Promega) were added at in a 1:50 enzyme:substrate ratio.

5

Phosphoenrichment of phosphoribosylated peptides

The ion metal affinity chromatography (IMAC) method, as described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014), was used.

10 **LC-MS/MS analysis of phosphoribosylated peptides**

Peptides were separated on a Thermo-Dionex RSLCNano UHPLC instrument with ~10 cm x 75 micron ID fused silica capillary columns with ~10 micron tip opening made in-house with a laser puller (Sutter) and packed with 3 micron reversed phase C18 beads (Reprosil-C18.aq, 120 Angstroms, Dr. Maisch). A 90 min gradient of 10-35% B at 200 nL/min using 0.1% acetic acid as solvent A and solvent B of 0.1% acetic acid, 99.9% acetonitrile was applied. MS data was collected with a Thermo Orbitrap Elite. Data-dependent analysis was applied using Top5 selection and fragmentation was induced by CID and HCD. Profile mode data was collected in all scans. MS analysis parameters were as previously described in Daniels et al. J. Proteome Res., doi:10.1021/pr401032q (2014).

20

Database search of MS/MS spectra for peptide and protein identification

Raw files were analyzed by MaxQuant version 1.5.3.8 using protein, peptide and site FDRs of 0.01 and a score minimum of 40 for modified peptides, 0 for unmodified peptides; delta score minimum of 17 for modified peptides, 0 for unmodified peptides. Sequences were searched against an in-house database containing the proteins of interest as well as Uniprot *Escherichia coli* BL21 DE3 database (definitions updated October 15th, 2014). MaxQuant search parameters: Variable modifications included Oxidation (M), Acetylation (Protein N-term), carbamidomethyl (C), phosphorylation (STY) and phosphoribosylation (DEKRC). Max labeled amino acids was 3, max missed cleavages was 2, enzyme was Trypsin/P, and max charge was 7.

25

30

The present disclosure has revealed new substrates for the RNA decapping Nudix hydrolases *EcRppH* and *HsNudT16*, as well as the CoAse *Dr1184*. In an attempt to determine the structural basis for this unique protein-conjugated ADPr hydrolase activity, these three proteins were aligned with respect to their known structures and identified a short, novel 3_{10} helix downstream from the helix of the Nudix signature sequence (Figure 7A, red box), suggesting a possible significance to these structures for the degradation of protein-conjugated ADPr. However, the structure is not known of the other two single domain Nudix hydrolases tested which lack the activity of interest: *AtORF147* and *Pa3470* (c.f. Figure 4A and Figure 4B). Moreover, the sequence homology between *EcRppH/HsNudT16/Dr1184* and *AtORF147/Pa3470* is not high beyond the Nudix signature sequence, so it is difficult to predict if the presence or absence of such short helices are the determining structural factor to facilitate the hydrolysis of protein-conjugated ADPr. Modeling of a trimeric PAR in the binding site of *EcRppH* positions the ribose towards the front of the enzyme allowing ample space for a protein to bind (Figure 7B). Notably, the bond to be cleaved to leave a phosphoribose attached to a protein is in an ideal position to be hydrolyzed by the residues of the catalytic site of the Nudix enzyme (Figure 7B). This model gives compelling structural arguments to the biochemistry data collected.

Mass spectrometry based proteomics represents the gold standard for the study of posttranslational modifications, and the field of ADP-ribosylation benefits from increased access to the suite of proteomic tools which have been developed for other PTMs such as phosphorylation, acetylation and ubiquitylation. The advent of tag-based approaches for identifying ADP-ribosylation sites has begun to provide access to these tools, but adoption has been relatively low due to technical difficulties which accompany the current methods. The disclosure described herein streamlines one of the most promising methods for ADPr site identification: the simplification of PAR or MAR to its phosphoribose attachment. Recombinant *HsNudT16* and *EcRppH* is synthesized and purified from *E. coli*, allowing for low-cost, high yield production which may be performed in most proteomic laboratories. Furthermore, since the structure is known for both of these enzymes, it is possible to predict mutations and truncations which could increase the enzyme's activity towards protein conjugated ADPr; for example, introducing mutations which further open up the active site to

allow larger ADP-ribosylated substrates access. The disclosure described herein will aid in the mass spectrometry based proteomic discovery of the ADP-ribosylated proteome, and researchers will find these tools provide them with a reliable and robust method for identifying tags of protein ADP-ribosylation sites on substrates of interest.

What is claimed is:

1. A method of generating phosphoribose tags on an adenosine diphosphate (ADP)-ribosylated protein comprising:
 - obtaining a sample including the ADP-ribosylated protein;
 - incubating the sample with a hydrolase; and
 - generating one or more phosphoribose tags on the ADP-ribosylated protein.
2. The method of claim 1, wherein incubating further comprises digesting the sample.
3. The method of claim 2, wherein digesting further comprises:
 - denaturing the ADP-ribosylated protein; and
 - incubating the denatured ADP-ribosylated protein with a protease.
4. The method of claim 1, wherein the hydrolase is a Nudix hydrolase.
5. The method of claim 4, wherein the Nudix hydrolase is selected from the group consisting of *Dr1184*, *RppH*, *Nudt16*, and combinations thereof.
6. The method of claim 4, wherein the Nudix hydrolase comprises a GXXXXXX[E/D]XXXXXXXX[X]REUXEEXG[U/Y] motif.
7. The method of claim 1, wherein the spectrometric technique is selected from the group consisting of mass spectrometry (MS), liquid-chromatography – MS, gas chromatography – MS, ion mobility – MS, and capillary electrophoresis - MS.
8. The method of claim 3, wherein the protease is a serine protease.
9. The method of claim 8, wherein the serine protease is LysC or trypsin.
10. The method of claim 1, further comprising:
 - detecting the one or more phosphoribose tags on the ADP-ribosylated protein.

11. The method of claim 10, wherein detecting occurs by a spectrometric technique selected from the group consisting of mass spectrometry (MS), liquid-chromatography – MS, gas chromatography – MS, ion mobility – MS, capillary electrophoresis - MS.
12. A method of identifying sites of adenosine diphosphate (ADP)-ribosylation on a protein, comprising:
 - obtaining a sample including one or more proteins having ADP-ribosylation;
 - incubating the sample with a hydrolase;
 - digesting the hydrolyzed sample; and
 - detecting one or more ADP-ribosylation sites using a spectrometric technique.
13. The method of claim 12, wherein the hydrolase is a Nudix hydrolase.
14. The method of claim 13, wherein the Nudix hydrolase is selected from the group consisting of *Dr1184*, *RppH*, *Nud16*, as well as the sequence and structural homologs of *RppH* in any species, and combinations thereof.
15. The method of claim 14, wherein the Nudix hydrolase comprises a GXXXXX[E/D]XXXXXXXX[X]REUXEEXG[U/Y] motif.
16. The method of claim 12, wherein the spectrometric technique is selected from the group consisting of mass spectrometry (MS), liquid-chromatography – MS, gas chromatography – MS, ion mobility – MS, capillary electrophoresis - MS.
17. The method of claim 12, wherein digesting further comprises:
 - denaturing the ADP-ribosylated protein.
18. The method of claim 17, wherein digesting further comprises incubating the denatured ADP-ribosylated protein with a protease.
19. The method of claim 18, wherein the protease is a serine protease.

20. The method of claim 19, wherein the serine protease is LysC or trypsin.
21. A method of detecting a disease associated with adenosine diphosphate (ADP)-ribosylation, comprising:
- obtaining a sample from a subject;
 - incubating the sample with a hydrolase;
 - digesting the hydrolyzed sample;
 - detecting one or more ADP-ribosylation sites using a spectrometric technique;
- and
- identifying the disease associated with ADP-ribosylation.
22. The method of claim 21, wherein the sample is selected from the group consisting of blood, plasma, tissue, saliva, urine, sputum, mucous, vitreous fluid, lymphatic fluid, lung aspirant, bile, stomach fluid, sweat, secretions, and stool.
23. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of cancer, diabetes, colitis, arthritis, reperfusion injury, transplantation, and neuropathology.
24. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of stroke, traumatic brain injury, Parkinson's diseases, Meningitis, and hypoglycaemia.
25. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of myocardial infarction, cardiopulmonary bypass, transplantation, ischaemic cardiomyopathy, aortic banding-induced heart failure, diabetic cardiomyopathy, myocardial failure, and ageing-associated heart failure.
26. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of Diabetic endothelial dysfunction, hypertension, ageing, balloon angioplasty, and endothelial injury.

27. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of Interstitial pulmonary fibrosis, acute respiratory distress syndrome, hyperoxic lung injury, and asthma.

28. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of Uveitis; retinal ischemia/reperfusion; diabetic retinopathy; and optic nerve transection.

29. The method of claim 21, wherein the disease associated with ADP-ribosylation is selected from the group consisting of cochlear ischemia, acoustic trauma; diabetic neuropathy; carvenous nerve injury; endotoxic and septic shock; and thoraco-abdominal ischemia/reperfusion.

30. A kit for detecting ADP-ribosylation comprising a Nudix hydrolase, hydrolysis buffer, salts, PARP inhibitor, urea, and denaturing buffers.

31. The kit of claim 30, wherein the denaturing buffers are selected from the group consisting of urea, guanidinium, sodium dodecyl sulfate, and trichloroacetic acid.

32. The kit of claim 30, wherein the hydrolase is a Nudix hydrolase.

33. The kit of claim 32, wherein the Nudix hydrolase is selected from the group consisting of NudF, NudE, Bd3179/BD NDPSases, *Dr1184*, RppH, Orf147, Nud16, YfcD, Pa3470, and combinations thereof.

34. The kit of claim 30, wherein the Nudix hydrolase comprises a GXXXXX[E/D]XXXXXXXX[X]REUXEEXG[U/Y] motif.

35. The kit of claim 30, wherein the PARP inhibitor is any one or more of the following PARP inhibitors:

Compound		Smile Code	Vendor	Catalog number
1		<chem>c12c(nc(c3ccccc3)[nH]c1=O)ccc2</chem>	Asinex	BAS 05269446
2		<chem>c12c(c3CCCCc3s1)c(=O)[nH]cn2</chem>	Chemdiv	1306-0033
3		<chem>c12c([nH]cnc1cccc2)=O</chem>	Chemdiv	C201-0583

4	DPQ	<chem>c1ccc2c(c1)C(NCC2OCCCC1CCCCC1)=O</chem>	Calbiochem	300270
5	NU1025	<chem>c1cc(c2c(c1)c([nH]c(n2)C)=O)O</chem>	Calbiochem	493800
6	PJ34	<chem>c1ccc2c1c([nH]c1c2cc(cc1)NC(CN(C)C)=O)=O,Cl</chem>	Calbiochem	528150
7	EB-47	<chem>c1cc(c2c(c1)C(NC2)=O)NC(=O)CN1CCN(CC1)C([C@@H]1O[C@@H]1[C@@H]1(C@H)1(O)))(n1c2c(nc1)c(ncn2)N))=O</chem>	Calbiochem	324473
8	6(5H)-Phenanthridinone	<chem>c1cc2c3c([nH]c(c2cc1)=O)cccc3</chem>	Sigma	299634
9	DiQ, 1,5-Isoquinolinediol	<chem>c1ccc2c1c(ccc2O)O</chem>	Sigma	1138
10		<chem>c1ccc2c1[nH]c([nH]c2=O)=O</chem>	Sigma	142026
11	DR2313	<chem>C1SCC2c1c(nc([nH]2)C)=O</chem>	Alexis biochemicals	DR2313
12	INH2BP	<chem>O1c(ccc2c1ccc(c2)N)=O</chem>	Alexis biochemicals	ALX-270-278
13	3-Methyl-5-AIQ	<chem>c1ccc2c(c1N)cc([nH]c2=O)C</chem>	Alexis biochemicals	ALX-270-450
14	5-AIQ	<chem>c1cc(c2c(c1)c([nH]c2)=O)N,Cl</chem>	Calbiochem	164300
15	TiQ-A	<chem>c1ccc2c(c1)c([nH]c1c2ccs1)=O</chem>	Calbiochem	612100
16	KU0058948	<chem>c1ccc2c(c1)c(n[nH]c2=O)Cc1cc(c(cc1)F)C(N1CCNCCC1)=O</chem>	KuDos (gift)	KU0058948
17	4-Amino-1,8-naphthalimide	<chem>c1ccc2c3c1c(ccc3c([nH]c2=O)=O)N</chem>	Sigma	A0966
18	1-Isoquinolind, ISQ	<chem>c1ccc2c(c1O)ccc2</chem>	Sigma	55435
19	3-Aminobenzamide, 3-AB	<chem>c1(cccc1)C(N)=O,N</chem>	Sigma	A0788
20	Vitamin B3 (nicotinamide)	<chem>O=C(N)c1ccccc1</chem>	Sigma	N3376
21		<chem>O=c1[nH]cnc2sc3CCCC3c12</chem>	ENAMINE	T0514-1628
22		<chem>O=c1[nH]ncc2ccccc12</chem>	ENAMINE	T0509-6940
23		<chem>OC(=O)CCc1nc2sc3CCC3c2c(=O)[nH]1</chem>	ENAMINE	T5621369
24		<chem>Cc1cccc2c(=O)[nH]c(CSc3ncccc3)nc12</chem>	ENAMINE	T5636531
25		<chem>CCN(C(=O)CCc1nc2sc3CCCC3c2c(=O)[nH]1</chem>	ENAMINE	T5823447
26		<chem>Cc1csc(CSc2nc3c(C)cccc3(=O)[nH]2)n1</chem>	ENAMINE	T5875616
27		<chem>n1ccccc1SCc1[nH]c(=O)c2c(n1)nc2c1cccc1</chem>	ENAMINE	T5810060
28		<chem>COc1ccc(cc1)NC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5689927
29		<chem>n1[nH]c(=O)[nH]c(=O)c1SCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T0511-0211
30		<chem>c1cccc2c1c([nH]c(n2)C(=O)N)N=C/c1ccc(F)cc1=O</chem>	ENAMINE	T0512-3494
31		<chem>ClC(Cl)(Cl)C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0515-7484
32		<chem>c1ccnc(Cl)c1C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0517-3776
33		<chem>COc1cccc(OC)c1C(=O)Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T0517-4507
34		<chem>[nH+]1ccccc1CNC(=O)CCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T0519-7725
35		<chem>c1cccc2c1c([nH]c(n2)SCc1cc(=O)oc2ccc3c(c12)cccc3)=O</chem>	ENAMINE	T5214418
36		<chem>c1cccc2c1c([nH]c(n2)CSc1nnc2n1cccc2)=O</chem>	ENAMINE	T5220503
37		<chem>O1CCC(C@@H)1(CN(Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5266916
38		<chem>c1cccc(OC)c1NC(=O)CSc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5282221
39		<chem>O=C([O-])[C@@H](C)NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5301904
40		<chem>O1CCC(C@@H)1(CN(C(=O)c1c2c(n[nH]1)cccc2)</chem>	ENAMINE	T5302202
41		<chem>c1cc(F)ccc1CNC(=O)CN(CC)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5405085
42		<chem>c1cc(F)cc([N+])([O-])=O)c1NC(=O)C(Nc1c2c(n1)cccc2)</chem>	ENAMINE	T5406525
43		<chem>c1cccc(F)c1CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5433140
44		<chem>C1COCCN1CNC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5433913
45		<chem>CCN1c(=O)[nH]nc1SCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5435488
46		<chem>NC(=O)CCS(=O)(=O)CCN(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5443746
47		<chem>COc1ccc(cc1OC)CN(Cc1cccc1)C(=O)c1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5458928

48	<chem>c1cccc2c1c[nH]c2=O)C(=O)N[C@H](C)(c1ccc(cc1)n1cccc1)</chem>	ENAMINE	T5475708
49	<chem>Cn1nc(ccc1=O)C(=O)N(CCOC)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5510865
50	<chem>n1cccc1[C@H](C)(NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5513304
51	<chem>CCOC(=O)NC(=O)CN(CC)Cc1nc2c(c[nH]1)=O)sc2</chem>	ENAMINE	T5342377
52	<chem>c1cs2c1c(nc[nH]2)CN(CCO)Cc1[nH]c(=O)c2c(n1)sc2)=O</chem>	ENAMINE	T5233261
53	<chem>O=C/1NC(=O)NC(=O)C1=C/Nc1cccc2c1C(NC2=O)=O</chem>	ENAMINE	T5390770
54	<chem>NC(=O)NC(=O)CN(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5434894
55	<chem>c1cccc(c1C#N)O[C@H](C)(c1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5442715
56	<chem>c1cccc2c1c[nH]c(n2)CCC(=O)N[C@H]1(Cc1Nc[nH]1(C1)(Cc1cccc1)))=O</chem>	ENAMINE	T5524834
57	<chem>COc1ccc(cc1)CCNC(=O)CSc1[nH]c(=O)c2c(n1)n(nc2)c1cc(C)ccc1</chem>	ENAMINE	T5526850
58	<chem>o1cccc1CN(C(=O)c1c2c(n[nH]1)cccc2</chem>	ENAMINE	T5531146
59	<chem>c1cs2c1nc[nH]2=O)C(N@H+)(CCC[C@H]1(c1cccs1))</chem>	ENAMINE	T5533330
60	<chem>c1cccc(c1C#N)NC(=O)[C@H](C)(N(C)Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5540724
61	<chem>c1cc(F)cc(F)c1[C@H](C)(N(C(=O)Cc1n[nH]c(=O)c2c1cccc2)</chem>	ENAMINE	T5546152
62	<chem>CCOc1cccc(c1C)[C@H]1(NC(=O)c2c(N1)cccc2)</chem>	ENAMINE	T5639683
63	<chem>c1cccc2c1C(N[C@H](N2)(c1c(F)cc(Br)cc1))=O</chem>	ENAMINE	T5642602
64	<chem>COc1c(O)c(ccc1)[C@H]1(NC(=O)c2c(N1)cccc2)OC</chem>	ENAMINE	T5699977
65	<chem>c1cccc1C(=O)OCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5707050
66	<chem>COC(=O)c1ccc(cc1)NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5707911
67	<chem>c1ccc(Br)cc1NC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5717474
68	<chem>c1cccc(O)c1NC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5723652
69	<chem>c1ccc(OC)cc1NC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5725081
70	<chem>c1cc(O)ccc1C(=O)N(C[C@H]1(CCCO1))Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5731587
71	<chem>c1cccc1NC(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5734903
72	<chem>c1cccc1NC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5742485
73	<chem>c1cccc(C)c1[C@H](C)(NC(=O)Cc1n[nH]c(=O)c2c1cccc2)</chem>	ENAMINE	T5748611
74	<chem>c1cs2c1c[nH]c(n2)SCC(=O)N[C@H]1(CCCc2c1cccc2)=O</chem>	ENAMINE	T5757803
75	<chem>c1cccc(F)c1CN(C)Cc1nc2c(c[nH]1)=O)sc2</chem>	ENAMINE	T5773275
76	<chem>n1cccc1c1ccc(cc1)[C@H](C)(N(C)Cc1nc2c(c[nH]1)=O)sc2)</chem>	ENAMINE	T5774012
77	<chem>c1cc(O)ccc1C1(CC1)C(=O)N(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5799957
78	<chem>N1CCN(C1=O)C(=O)C(N@H+)(C)(Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5807539
79	<chem>c1cccc2c1c[nH]c(n2)[C@H](C)(N(C)Cc1cc(O)c(cc1)OC)=O</chem>	ENAMINE	T5812575
80	<chem>c1cccc1[C@H](C)(NC(=O)CCc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5817800
81	<chem>C[NH+](C)Cc1c(cccc1)CN(C(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5820448
82	<chem>c1ccc(OC)cc1CN(C(=O)Cc1n[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5820522
83	<chem>c1cccc2c1c[nH]2)C(=O)CN(C)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5823618
84	<chem>c1cc(F)ccc1[C@H](C)(N(C)Cc1[nH]c(=O)c2c(n1)cccc2)</chem>	ENAMINE	T5826883
85	<chem>c1cc(O)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)sc2</chem>	ENAMINE	T5826869
86	<chem>c1cccc1C(=O)N(CC)Cc1[nH]c(=O)c2c(n1)cccc2</chem>	ENAMINE	T5827773
87	<chem>Fc1c(F)ccc(c1)NC(=O)Cn1c(=O)[nH]c(=O)c2c1cccc2</chem>	ENAMINE	T5828779
88	<chem>c1cccc2c1c[nH]c2=O)C(=O)NCc1cc2c(cc1)cc(cc2)OC</chem>	ENAMINE	T5829176
89	<chem>N#Cc1cc(ccc1)NC(=O)CSc1[nH]c(=O)c2c(n1)sc2</chem>	ENAMINE	T5829810
90	<chem>c1cccc2c1c[nH]c2=O)C(=O)NCc1cccc(cc1)C[NH+](CC)CC</chem>	ENAMINE	T5831807
91	<chem>c1cccc2c1c[nH]c(n2)C(N@H+)(Cc2c(CC1)sc2)=O</chem>	ENAMINE	T5836432

92	<chem>c1cccc2c1c([nH]c(n2CC(=O)N[C@H](C)(c1cF)c(F)cc1)=O)=O</chem>	ENAMINE	T5840584
93	<chem>c1ccc(F)cc1NC(=O)C[N@H+](C)(Cc1nc2c(c([nH]1)=O)sc2)</chem>	ENAMINE	T5844481
94	<chem>O1COc2c1cc(cc2)CNC(=O)Cc1n[nH]c(-O)c2c1cccc2</chem>	ENAMINE	T5847624
95	<chem>c1cccc2c1C(N[C@H](N2)(c1cc(C)cc2c(c1)OCCO2))=O</chem>	ENAMINE	T5853381
96	<chem>C1CN(C)Cc2c1nc(s2)NC(=O)CCc1[nH]c(-O)c2c1cccc2</chem>	ENAMINE	T5855515
97	<chem>c1cc(F)ccc1OC[C@H](O)(CN(C)Cc1[nH]c(-O)c2c1cccc2)</chem>	ENAMINE	T5863205
98	<chem>Nc1nc([nH]1)SCc1[nH]c(-O)c2c1c1c(C)ccc2</chem>	ENAMINE	T5867921
99	<chem>c1cccc2c1c([nH]c(n2)CCC(=O)N(C)Cc1c(ccc1)N1CCOCC1)=O</chem>	ENAMINE	T5904510
100	<chem>c1cccc2c1c([nH]c2=O)CC(=O)NCc1ccc([nH+])c1N1CCOCC1</chem>	ENAMINE	T5904604
101	<chem>c1cc([N+](=O)=O)ccc1[C@H](O)(CNC(=O)Cc1n[nH]c(-O)c2c1cccc2)</chem>	ENAMINE	T5918237
102	<chem>Cc1cc(c(cc1)O)C(=O)c1cc2c(nc1)[nH]c(-O)c1c2cccc1</chem>	ENAMINE	T5949414
103	<chem>c1cccc([O-])c1C(=O)c1cc2c(nc1)[nH]c(-O)c1c2cccc1</chem>	ENAMINE	T5949445
104	<chem>c1cc(F)cc(F)c1[C@H](C)(N(C(=O)CCCc1[nH]c(-O)c2c1cccc2)</chem>	ENAMINE	T5970736
105	<chem>c1cccc2c(-O)[nH]c3c(c12)ccc(n3)c1cccc1</chem>	ENAMINE	T5974626
106	<chem>Cc1cc(C(F)(F)F)nc2[nH]c(-O)c3c(c12)cccc3</chem>	ENAMINE	T5989689
107	<chem>n1c(c2cc(S(N(CCO)C))=O)=O)c(cc2)C)c2c(ccc2)c([nH]1)=O</chem>	ASINEX_GOLD	BAS 04375852
108	<chem>n1c(c2cc(S(=O)=O)NCCN3CCOCC3)c(cc2)C)c2c(ccc2)c([nH]1)=O</chem>	ASINEX_GOLD	BAS 06635491
109	<chem>c1cc(F)ccc1C(=O)N(Cc1c(OC)ccc1)Cc1[nH]c(-O)c2c1cccc2</chem>	ASINEX_PLATINUM	ASN 05812848
110	<chem>c1cccc1CCNS(=O)(=O)c1cc2c3c(c1)cccc3c(-O)[nH]c2=O</chem>	ASINEX_PLATINUM	ASN 09801991
111	<chem>CCc1ccc(cc1)c1n[nH]c(-O)c2c1cccc2</chem>	ASINEX_GOLD	BAS 06805602
112	<chem>n1c(c2cccc2c([nH]1)=O)Cc1ccc(cc1)CC</chem>	INTERBIOSCREEN	STOCK15-03944
113	<chem>n1c(c2cccc2c([nH]1)=O)c1cc(N)c(cc1)C</chem>	INTERBIOSCREEN	STOCK45-10309
114	<chem>O1COc2c1cc(c(c2)Cl)C@H]1(NC(=O)c2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-38708
115	<chem>O1COc2c1cc(cc2)C@H]1(NC(=O)c2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-31658
116	<chem>c12c3cccc1cccc2c(-O)[nH]c3=O</chem>	INTERBIOSCREEN	STOCK25-08660
117	<chem>COc1c(OC)cc(cc1OC)C@H]1(NC(=O)c2c(O1)cccc2)</chem>	INTERBIOSCREEN	STOCK55-34270
118	<chem>O1COc2c1cc(cc2)c1[nH]c(-O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK55-79756
119	<chem>[nH+][c1[nH]cc1CCN(C(=O)Cc1n[nH]c(-O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-28405
120	<chem>c1cccc2c1c([nH]c2=O)CC(=O)NCC(=O)Nc1cc(cc1)NC(=O)C</chem>	INTERBIOSCREEN	STOCK65-34326
121	<chem>C1COCCN1C(=O)CNC(=O)Cc1n[nH]c(-O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-34374
122	<chem>C1COCCN1C(=O)C@H](C)(NC(=O)Cc1n[nH]c(-O)c2c1cccc2)</chem>	INTERBIOSCREEN	STOCK65-38709
123	<chem>c1cccc2c1c([nH]c(n2)SCC(=O)NCCc1c[nH]c2c1cccc2)=O</chem>	INTERBIOSCREEN	STOCK65-43755
124	<chem>[nH+][c1cccc1CN(C(=O)CCN(C(=O)Cc1n[nH]c(-O)c2c1cccc2</chem>	INTERBIOSCREEN	STOCK65-46880
125	<chem>n1c(c2cccc2c([nH]1)=O)Cc1c(cc(cc1)Cl)Cl</chem>	KEYORGANICS	1E-57
126	<chem>[N+](c1c2c(c(-O)[nH]cc2)ccc1)([O-])=O</chem>	KEYORGANICS	10B-099
127	<chem>Fc(F)(F)c1cc(ccc1)Cc1[nH]c(-O)c2c1cccc2</chem>	KEYORGANICS	2F-301S
128	<chem>Fc(F)(F)c1cc(ccc1)NCCc1n[nH]c(-O)c2c1cccc2</chem>	KEYORGANICS	12B-089
129	<chem>COc1ccc(cc1)c1n[nH]c(-O)c2c1cccc2</chem>	KEYORGANICS	MS-1309
130	<chem>c1cccc2c1c([nH]c(n2)C@H](C)([N+H2+])Cc1cc(C)cc1)=O</chem>	KEYORGANICS	MS-3000
131	<chem>c1cccc(Cl)c1CNCc1[nH]c(-O)c2c1cccc2</chem>	KEYORGANICS	MS-3205
132	<chem>n1c(c2cccc2c([nH]1)=O)c1cc([N+](=O)=O)c(cc1)NCCCN(C)C</chem>	VITAS-M LABORATORY	STK238640
133	<chem>c1(cc(=O)c2c(o1)ccc2)c1cccc1</chem>	VITAS-M LABORATORY	STK164205
134	<chem>c1([nH]c2ccc(SNCc3cccc3)(=O)=O)cc2c2c1cccc2)=O</chem>	VITAS-M LABORATORY	STK272141
135	<chem>c12c3c(cccc3c([nH]n1)=O)c1c2cccc1</chem>	CHEMBRIDGE	6259889

136		<chem>c12c3c[nH]c(c3ccc2NC2CCCC2)=O)c2c(c1=O)cccc2</chem>	CHEMBRIDGE	7456760
137		<chem>n1c(c2cc(S[N3CCN(CC3)C3cccc3](=O)=O)c(cc2)C)c2c(c(=O)[nH]1)cccc2</chem>	CHEMBRIDGE	7669944
138		<chem>c1([nH]c2ccc(S[N3CC3](=O)=O)cc2c2c1cccc2)=O</chem>	CHEMBRIDGE	9040441
139		<chem>c12c3cc(S(=O)(=O)N)cc1cccc2c(=O)[nH]c3=O</chem>	CHEMBRIDGE	7650155
140		<chem>FC(F)(F)c1cc(ccc1)NC(=O)CSc1[nH]c(=O)c2c(n1)cccc2</chem>	CHEMBRIDGE	7663670
141		<chem>Cc1ccc(cc1)NS(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	7841182
142		<chem>O1COC2c1cc(cc2)CN(C(=O)c1[nH]c(=O)c2c1cccc2</chem>	CHEMBRIDGE	7959335
143		<chem>C1COCN1S(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	9026017
144		<chem>CN(C)S(=O)(=O)c1cc2c3c(c1)cccc3c(=O)[nH]c2=O</chem>	CHEMBRIDGE	9030160
145		<chem>c12c(c(=O)[nH]c(n1)C)c1CCCC1s2</chem>	CHEMDIV	3909-7439
146		<chem>c12[nH]c3CCCCc3c2cccc1C(=O)N</chem>	CHEMDIV	5348-0038
147		<chem>c1([nH]c(C)cc2c1cccc2)=O</chem>	CHEMDIV	3909-9376
148		<chem>c12c3c(c(C)cccc3c([nH]n1)=O)c(c1c2cccc1)=O</chem>	CHEMDIV	8011-2789
149		<chem>c1([nH]c2ccc([N+])([O-])=O)cc2c2c1cccc2)=O</chem>	CHEMDIV	R507-0010
150		<chem>c1cc(F)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)CCCC2</chem>	CHEMDIV	D365-0030
151		<chem>COC1ccc(cc1)NC(=O)CSc1[nH]c(=O)c2c(n1)CCC2</chem>	CHEMDIV	D365-0199
152		<chem>c1cccc(c1C(=O)N)NC(=O)CSc1[nH]c(=O)c2c(n1)cccc2</chem>	CHEMDIV	F019-0833
153		<chem>c1cccc2c1c([nH]c(n2)CSc1nc(cc([nH]1)=O)C)=O</chem>	CHEMDIV	8016-1228
154		<chem>c1cc(F)ccc1CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0248
155		<chem>c1ncccc1CNC(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0339
156		<chem>CC[NH+](CC)CCCN(C(=O)c1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-0477
157		<chem>c1ncccc1Cc1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	C066-3059
158		<chem>c1cccc(C(F)(F)F)c1C(=O)NCc1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	G786-2295
159		<chem>c1cc(C)ccc1CC(=O)NCc1n[nH]c(=O)c2c1cccc2</chem>	CHEMDIV	G786-2308
160		<chem>C1(c2c(c(c(OC)ccc2)CCN1)=O</chem>	LIFE CHEMICAL	F2679-0001
161		<chem>n1c2c(C(=O)N)cccc2nc2c1cccc2</chem>	LIFE CHEMICAL	F3031-0010
162		<chem>c1cc(Br)ccc1Sc1[nH]c(=O)c2c(n1)cccc2</chem>	LIFE CHEMICAL	F2551-0013
163		<chem>c1cc(C)ccc1NC(=O)CSc1nc2c(c([nH]1)=O)CCCC2</chem>	LIFE CHEMICAL	F3116-0089
164		<chem>c1cccc2c1C(N(C@H)1(N2Cc2c(C1)cccc2))=O</chem>	LIFE CHEMICAL	F3284-0013
165		<chem>COC1ccc(cc1)C[C@H]1(c2c(CN3[C@H]1(NC(=O)c1c3cccc1))cccc2)</chem>	LIFE CHEMICAL	F3097-0628
166		<chem>c1cc(C)ccc1NC(=O)CSc1[nH]c(=O)c2c(n1)CCC2</chem>	LIFE CHEMICAL	F3308-0716
167		<chem>c1cccc1CNC(=O)CSc1[nH]c(=O)c2c(n1)n(nc2)c1cccc1</chem>	LIFE CHEMICAL	F2085-0030
168		<chem>OCCNC(=O)C(=O)NCc1n[nH]c(=O)c2c1cccc2</chem>	LIFE CHEMICAL	F2092-0051
169		<chem>c1cccc1CNC(=O)COc1cccc2c1CCNC2=O</chem>	LIFE CHEMICAL	F2679-0060
170		<chem>O=C1NC(=O)c2c1c(ccc2)NC(=O)CCCS(=O)(=O)c1ccc(cc1)OC</chem>	LIFE CHEMICAL	F2814-1523
171		<chem>Fc1cccc2c1C(N(C@H)(N2)(c1c(O)ccc(c1))N+)([O-])=O</chem>	MAYBRIDGE	BTB 04585
172		<chem>c1cc(C)ccc1Sc1n[nH]c(=O)c2c1cccc2</chem>	MAYBRIDGE	KM 09200
173		<chem>c1cc(ccc1)C(N)=O</chem>	KaironKem	K181
174		<chem>c1(c2c(ccc1)nc(cc2)c1ccc(cc1)C)C(=O)N</chem>	Alexis biochemicals	ALX-270-486
175	ABT-888	<chem>O=C(N)c1c2nc([nH]c2ccc1)C11(NCCCC1)C</chem>	Alexis biochemicals	270-444-M005
176		<chem>c1(c(c2cccc2oc1=O)Nc1cccc1)NC(=O)C</chem>	ChemBridge	5351069
177		<chem>c1(c(c2cccc2oc1=O)NCCCC1CCN(CC1)C)N</chem>	ChemBridge	5244706
178		<chem>c1(n(c(m1)SCc1[nH]c(=O)c2c(n1)cccc2)CC)c1ccc(cc1)Cl</chem>	ChemBridge	7965028
179		<chem>c12nc(N3C(C)CCCC3)n(c1c(=O)[nH]c(n2C)=O)Cc1cccc1</chem>	ChemBridge	7996603

180		<chem>c12nc(N3C(C)CCCC3)n(c1c(-O)[nH]c(n2C)=O)Cc1cccc1</chem>	ChemBridge	7367553
181		<chem>c12c(nc(C3CC3)nc1O)CCSC2</chem>	Maybridge	RF03876
182	XAV939	<chem>c1(c2ccc(c(F)F)cc2)nc2c(CSCC2)c(n1)O</chem>	Maybridge	RF03920
183		<chem>c12c(nc(c3cccn3)nc1O)CCSC2</chem>	Maybridge	RF03877
184	Olaparib	<chem>Fc1c(C(-O)N2CCN(C(-O)C3CC3)CC2)cc(Cc2n[nH]c(-O)c3c2cccc3)cc1</chem>	JS Research Chemicals	CAS 763113-22-0
185	Rucaparib	<chem>Fc4cc2c1c(nc(c1CCNC2=O)c3ccc(cc3)CNC)c4</chem>	Axon MedChem	Axon 1529

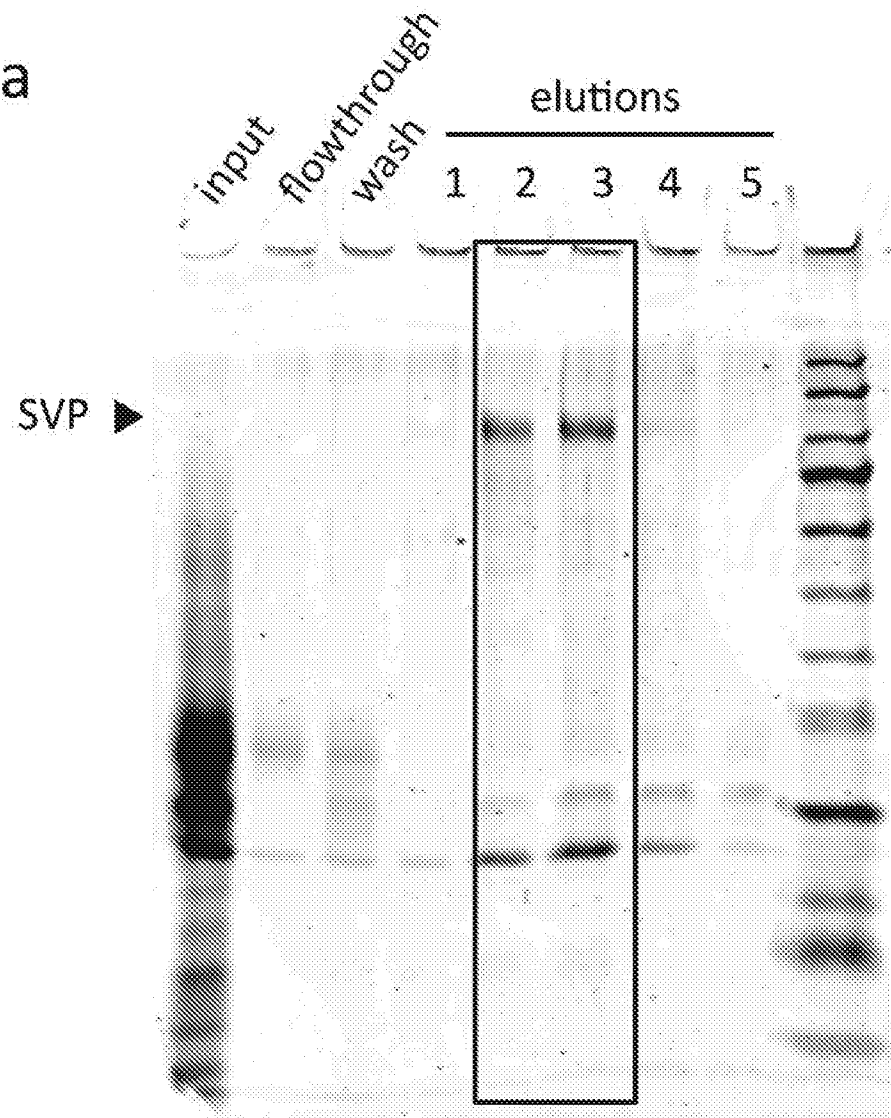


Figure 1A

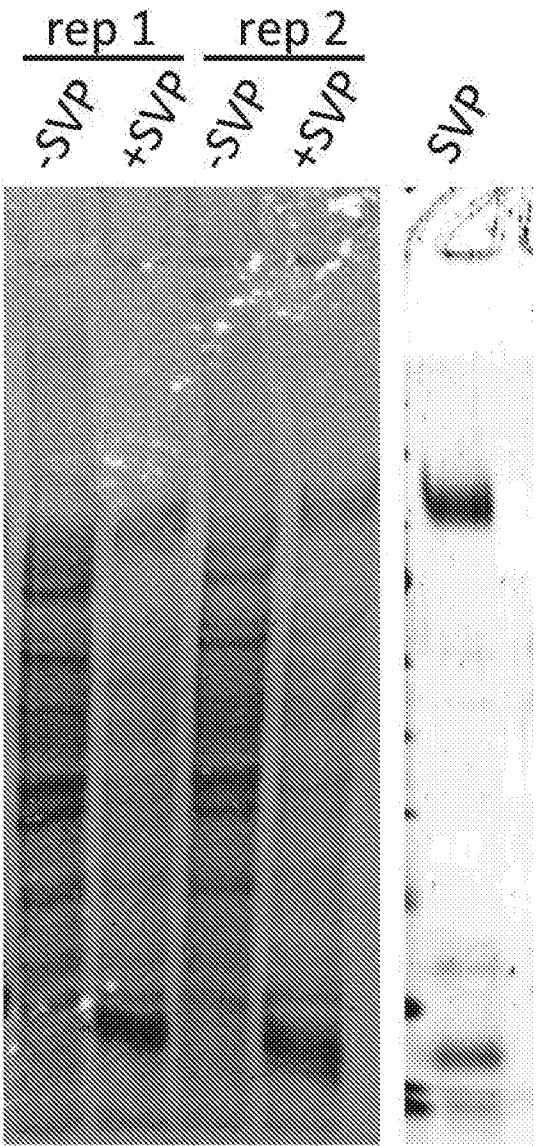


Figure 1B

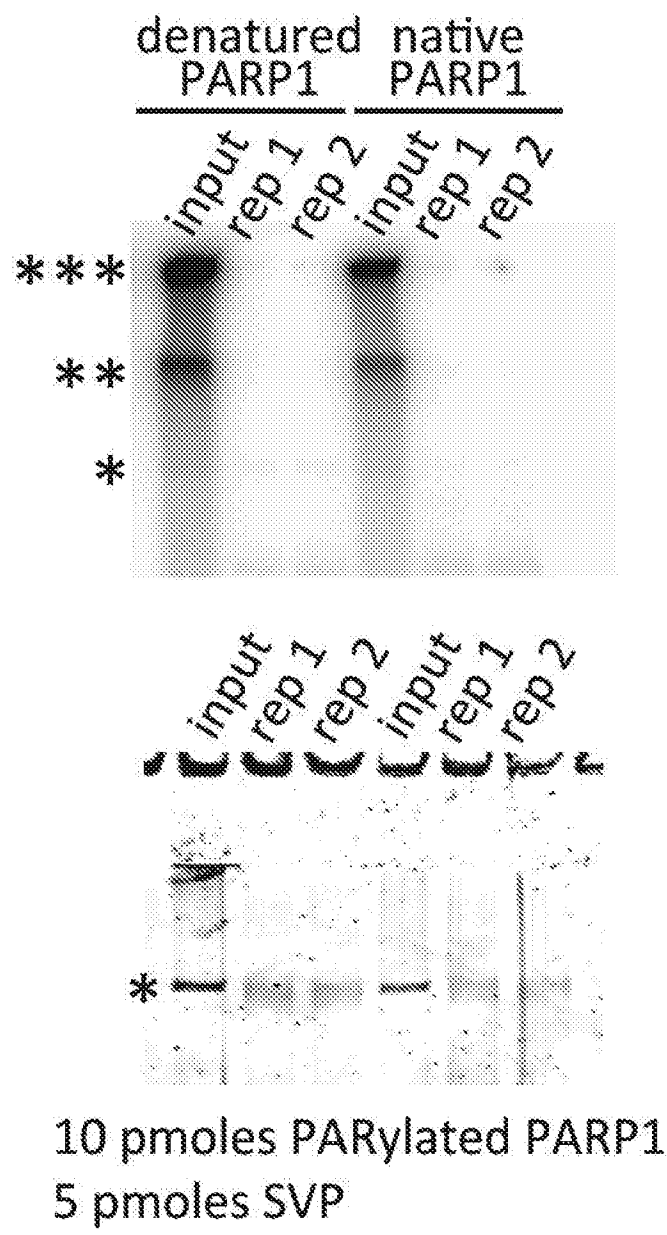


Figure 1C

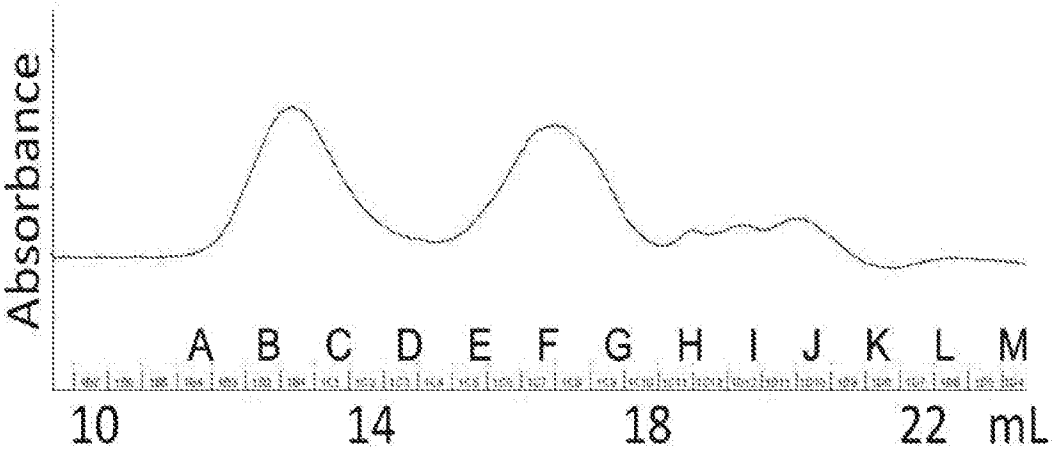


Figure 1D

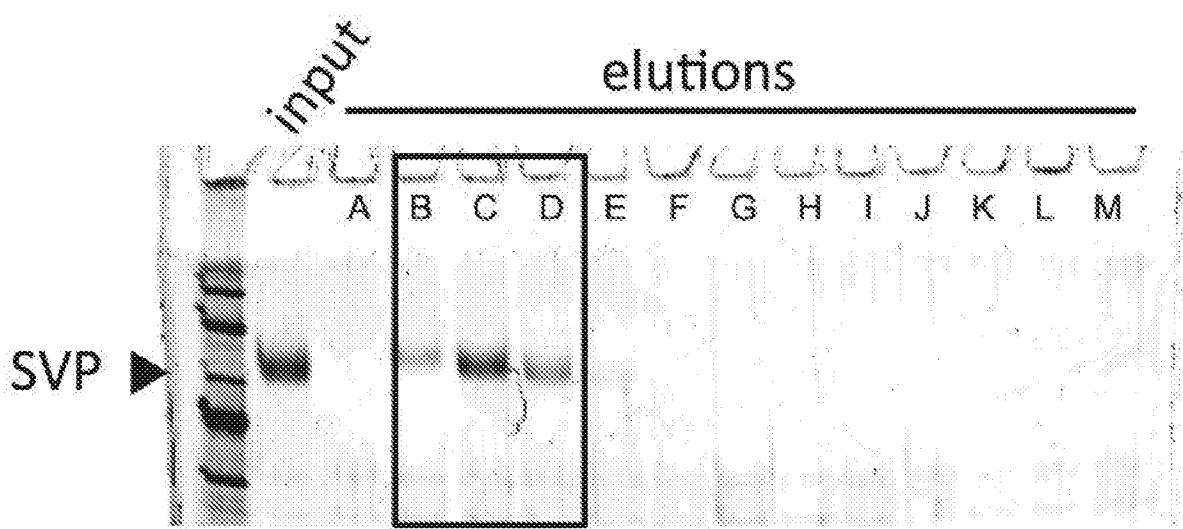
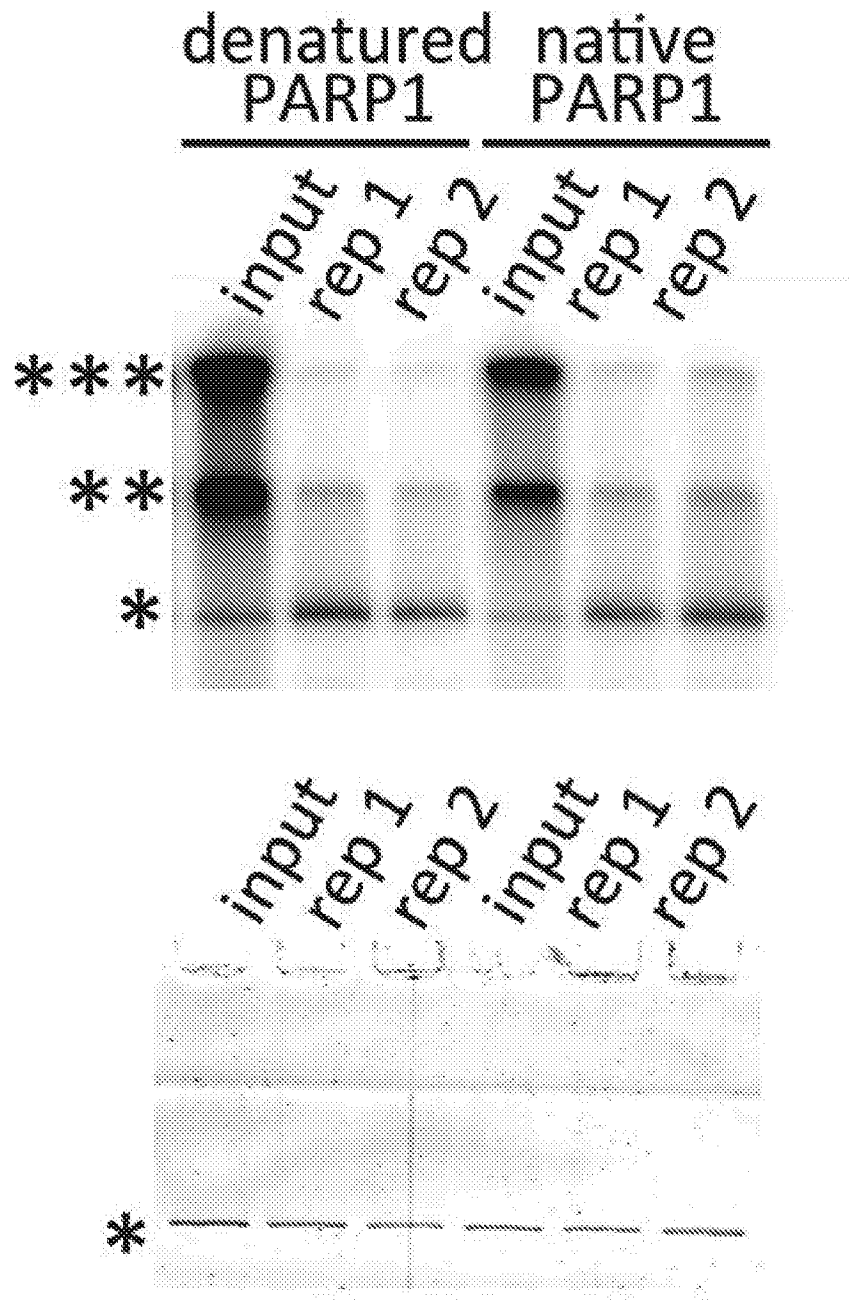


Figure 1E



Figure 1F



10 pmoles PARylated PARP1
5 pmoles SVP

Figure 1G

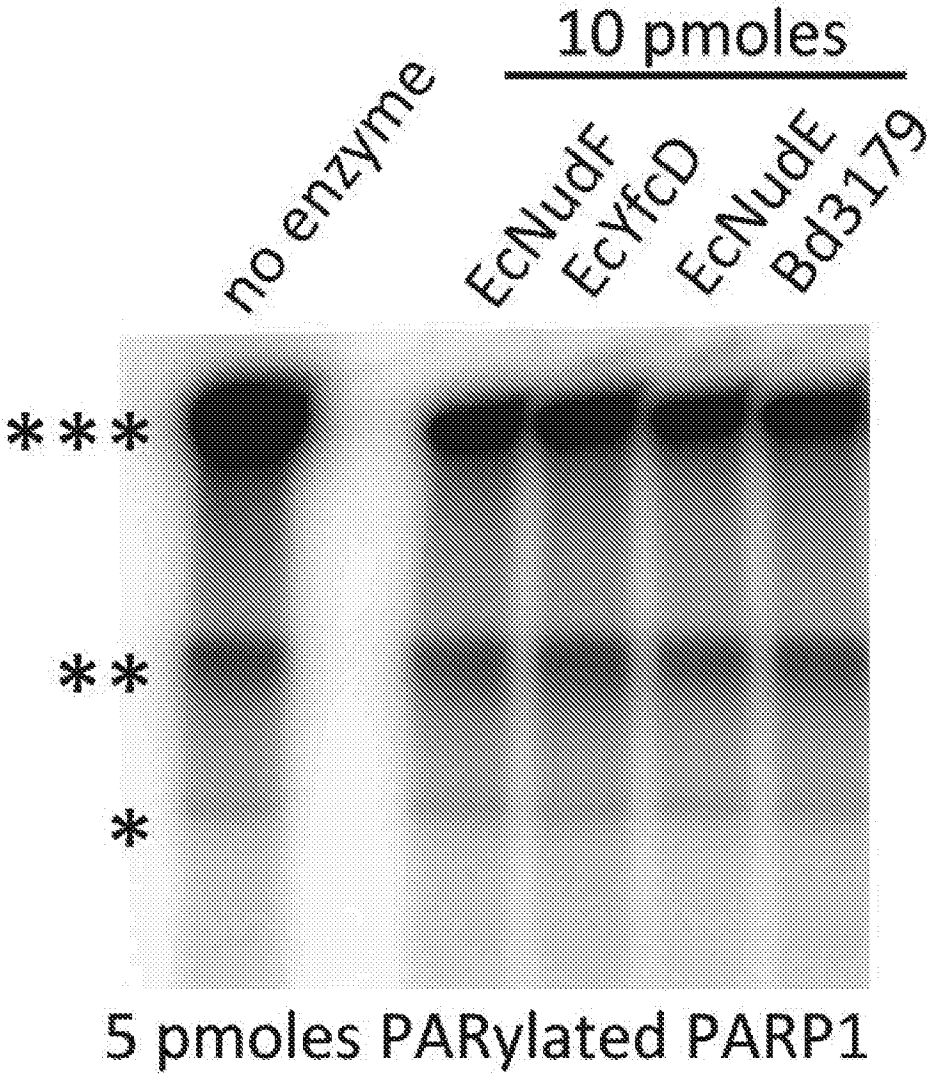
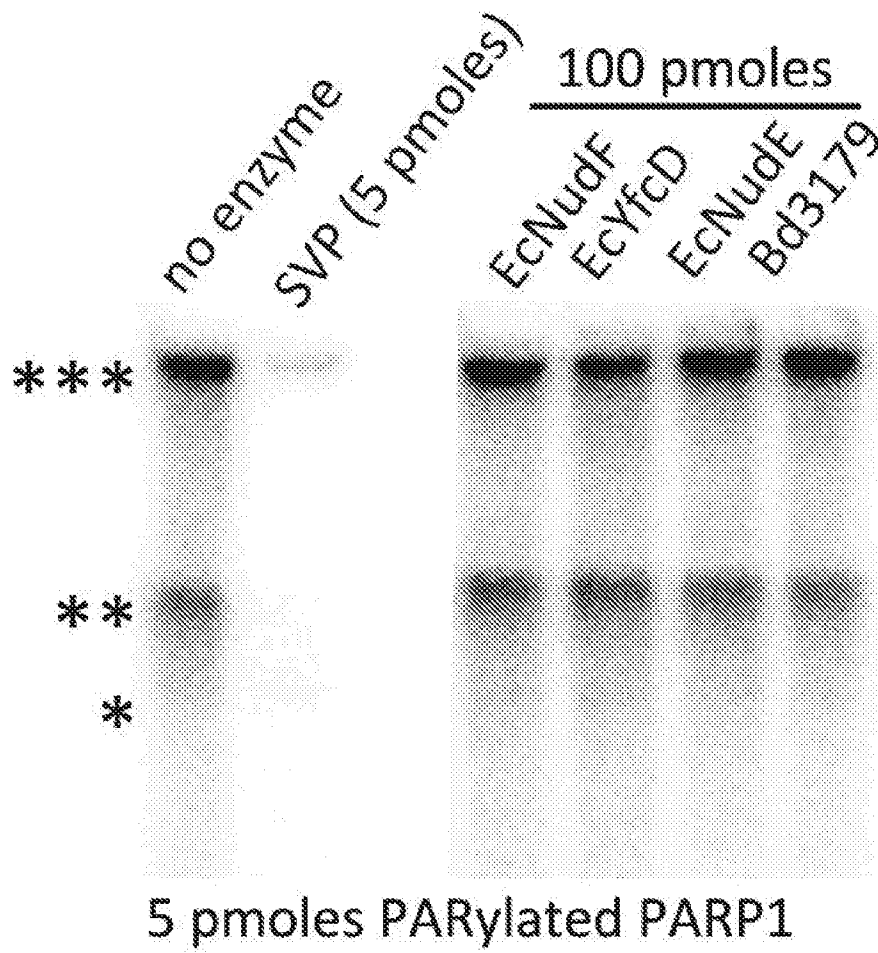
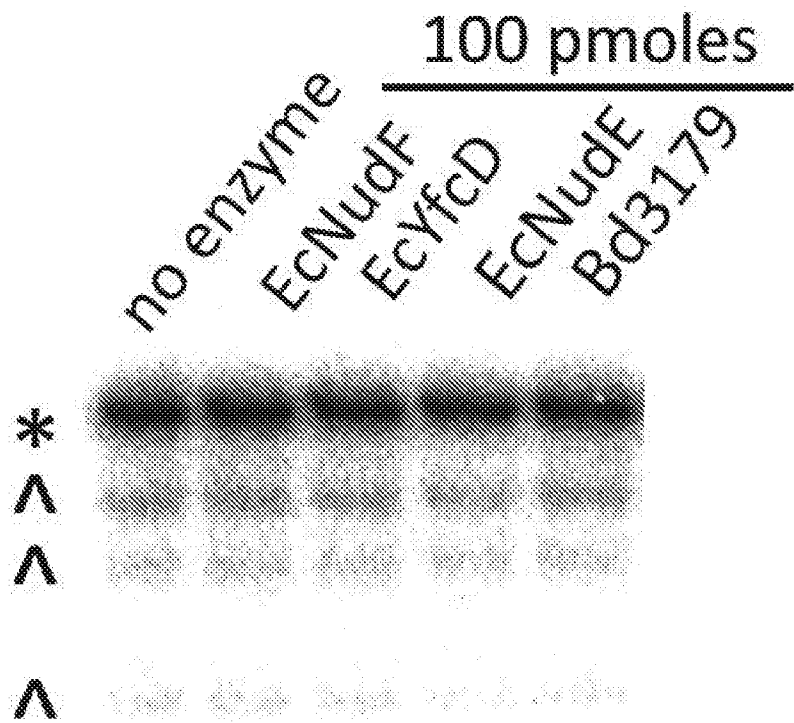


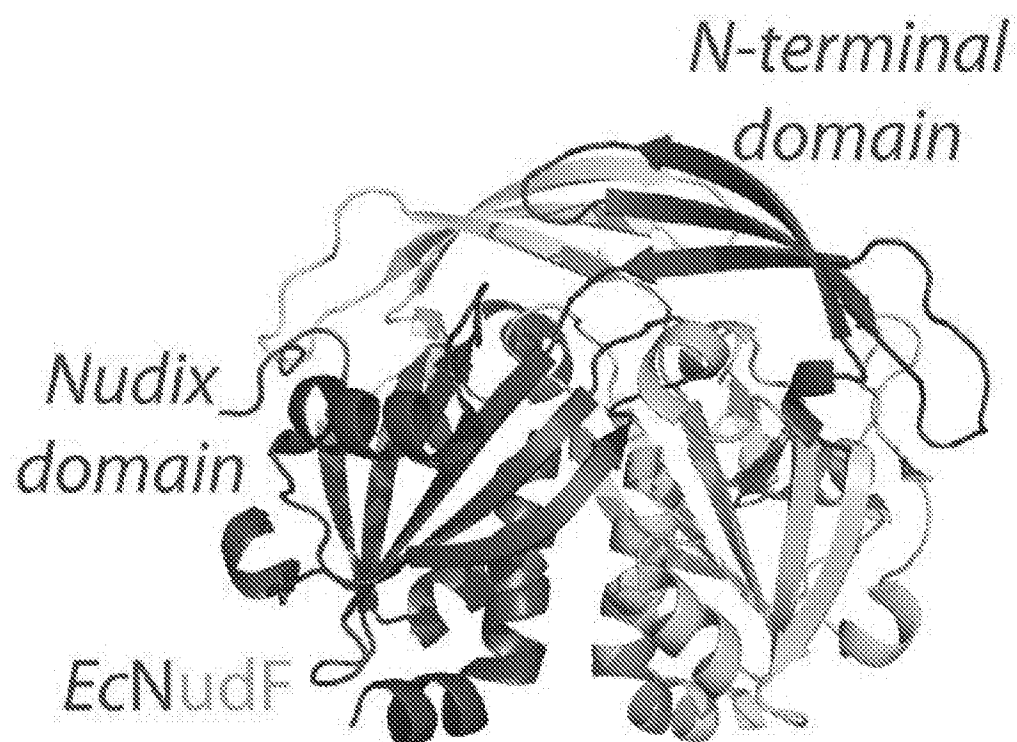
Figure 2A

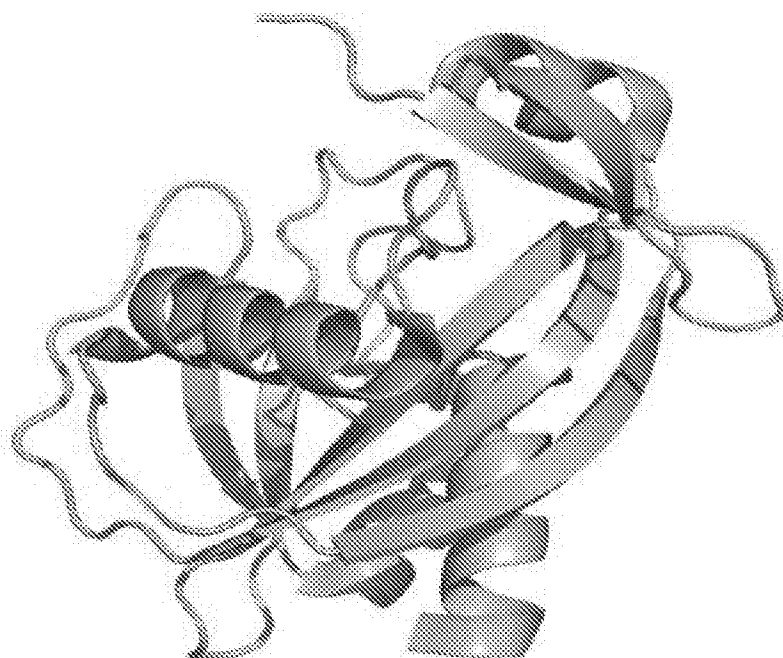
**Figure 2B**



5 pmoles MARYlated PARP1

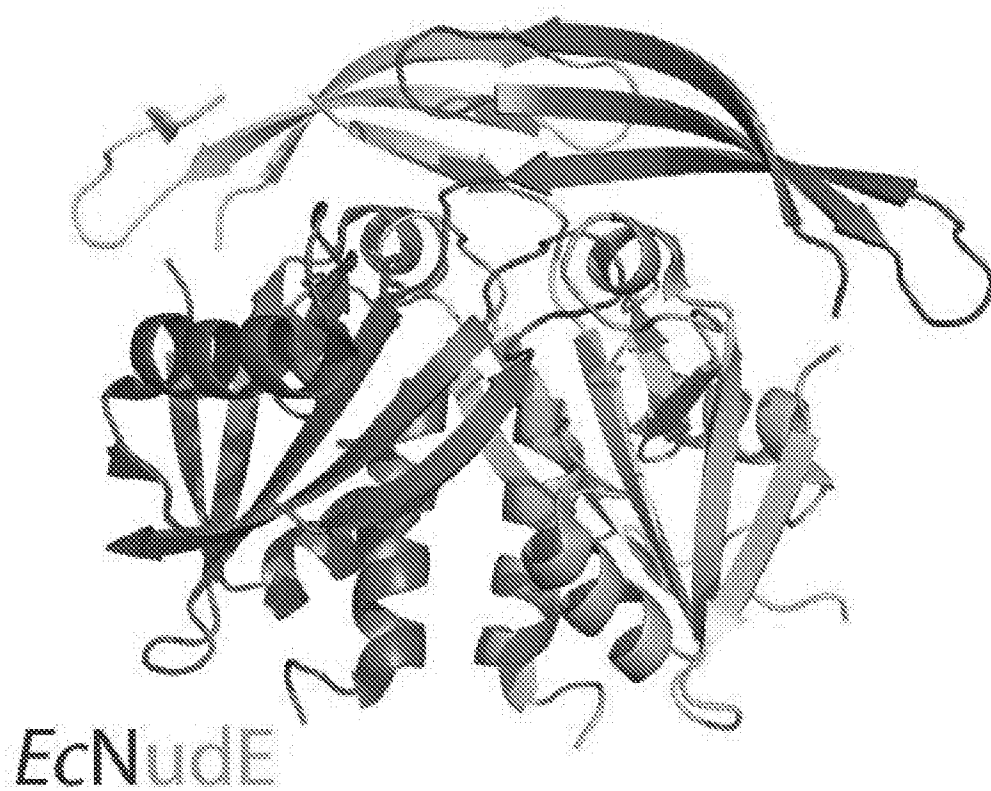
Figure 2C

**Figure 2D**



EcYfcD

Figure 2E

**Figure 2F**

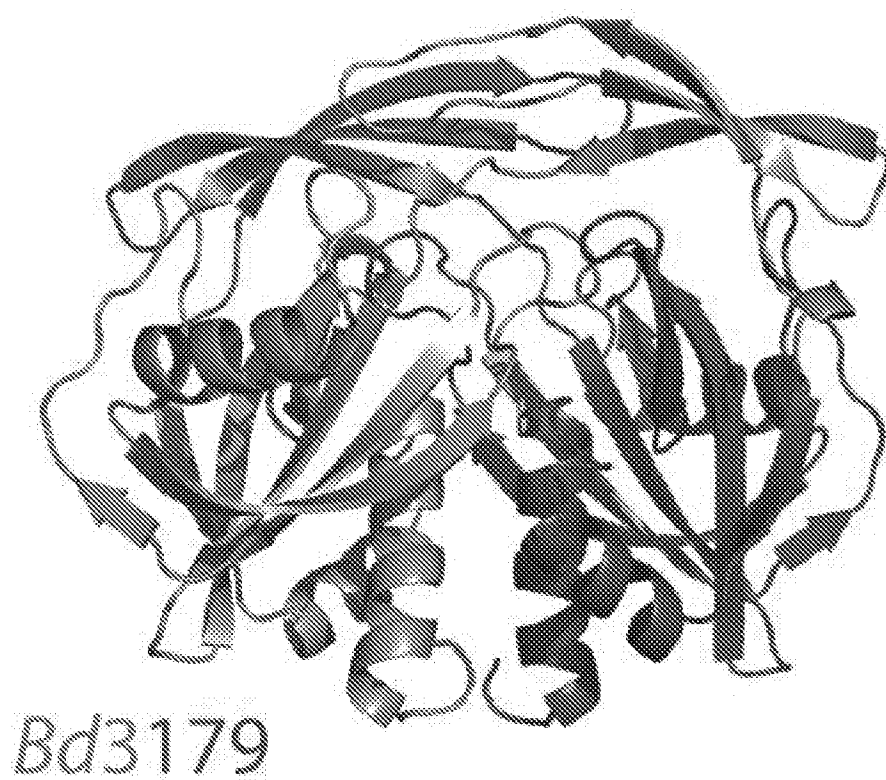
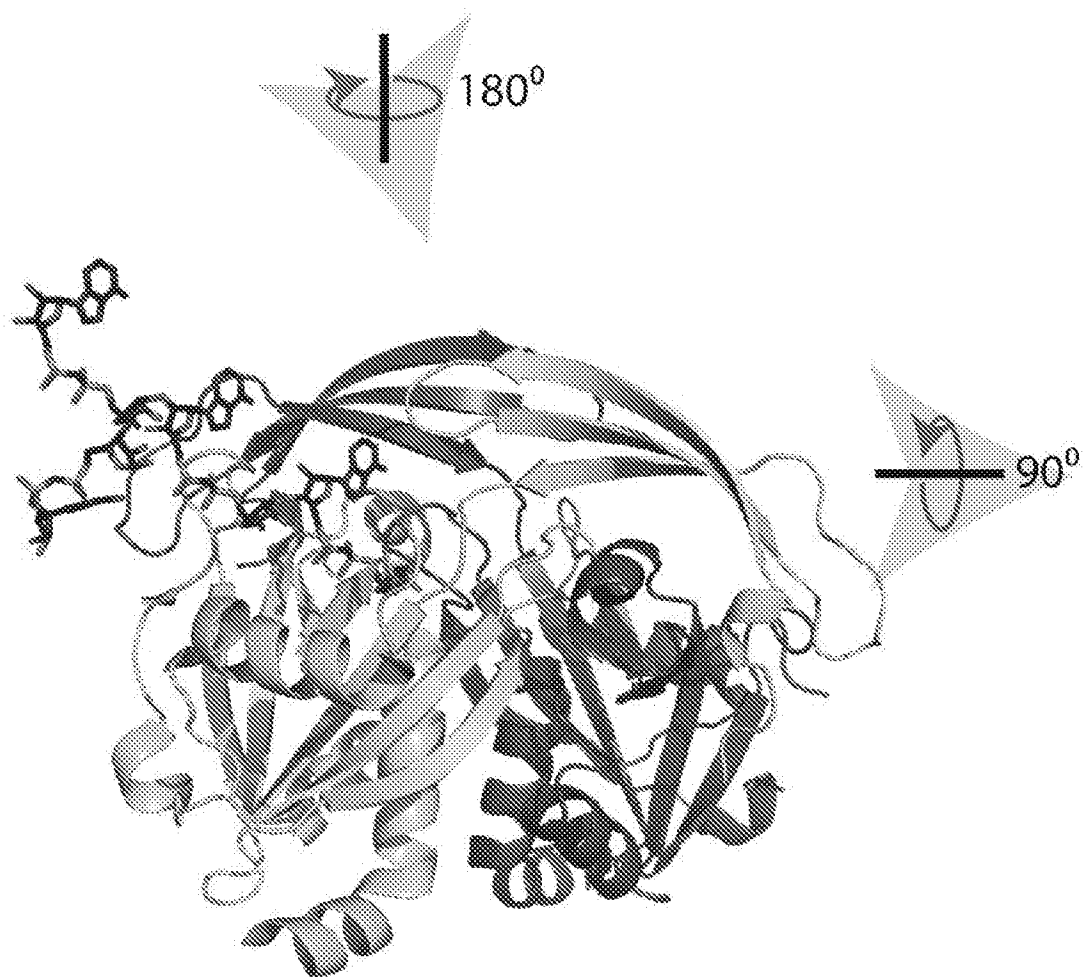
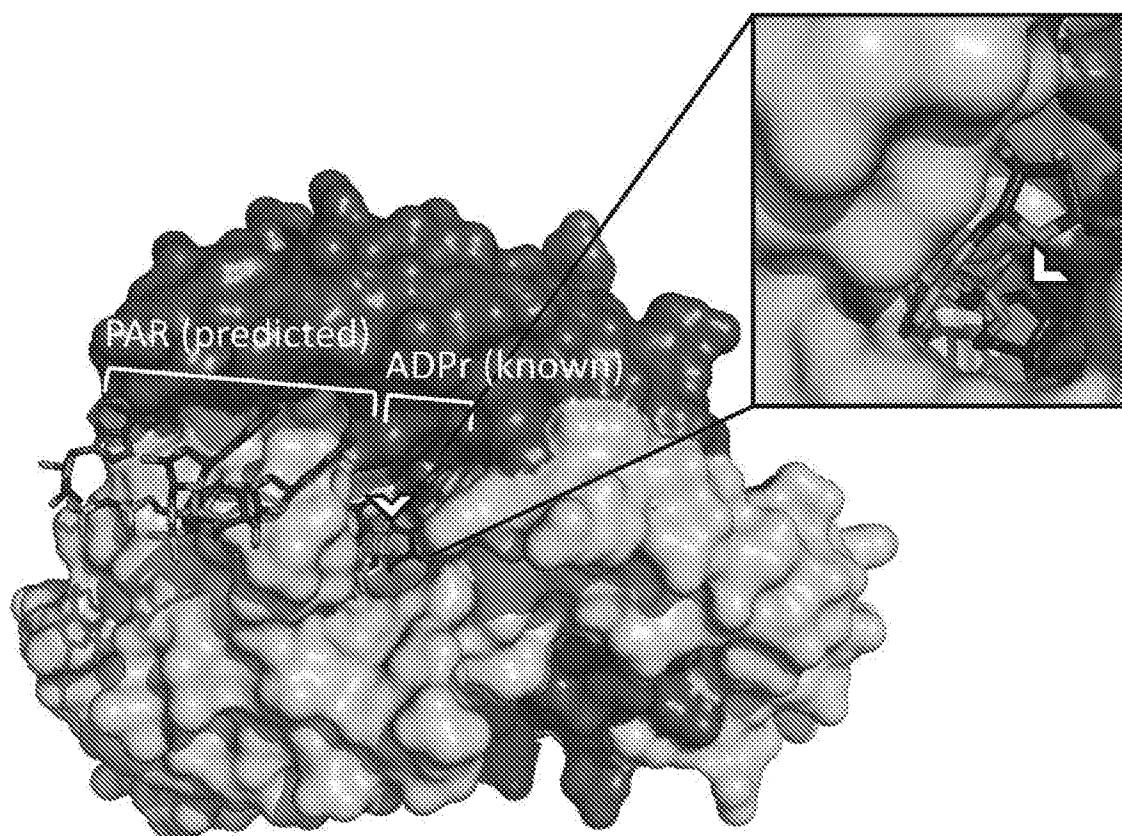
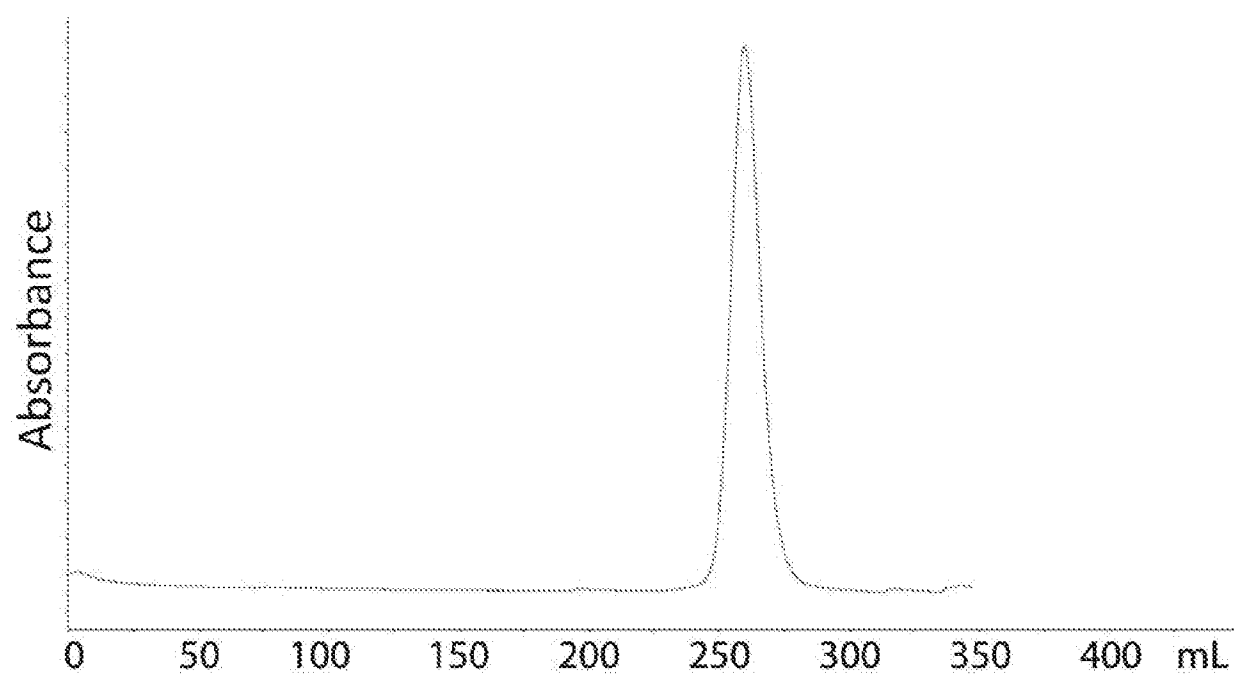


Figure 2G

**Figure 2H**

**Figure 2I**

**Figure 3**

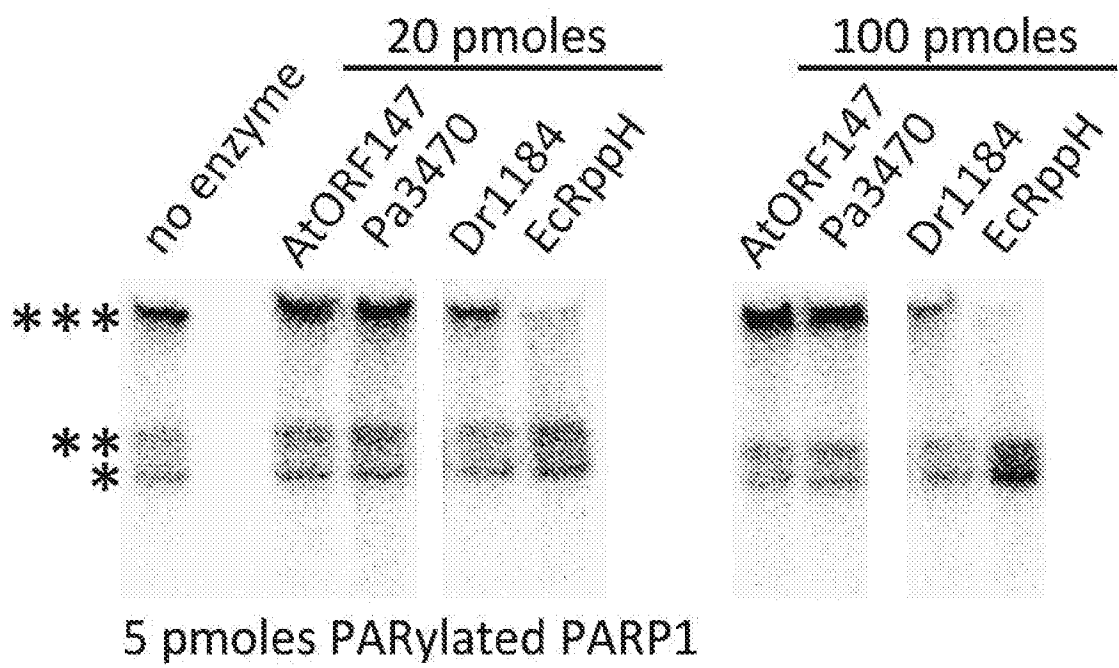


Figure 4A

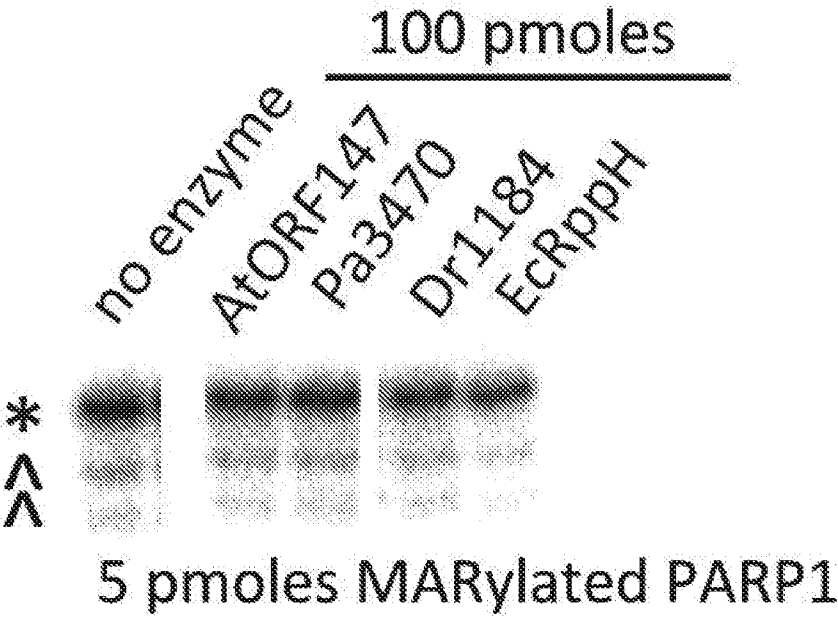


Figure 4B

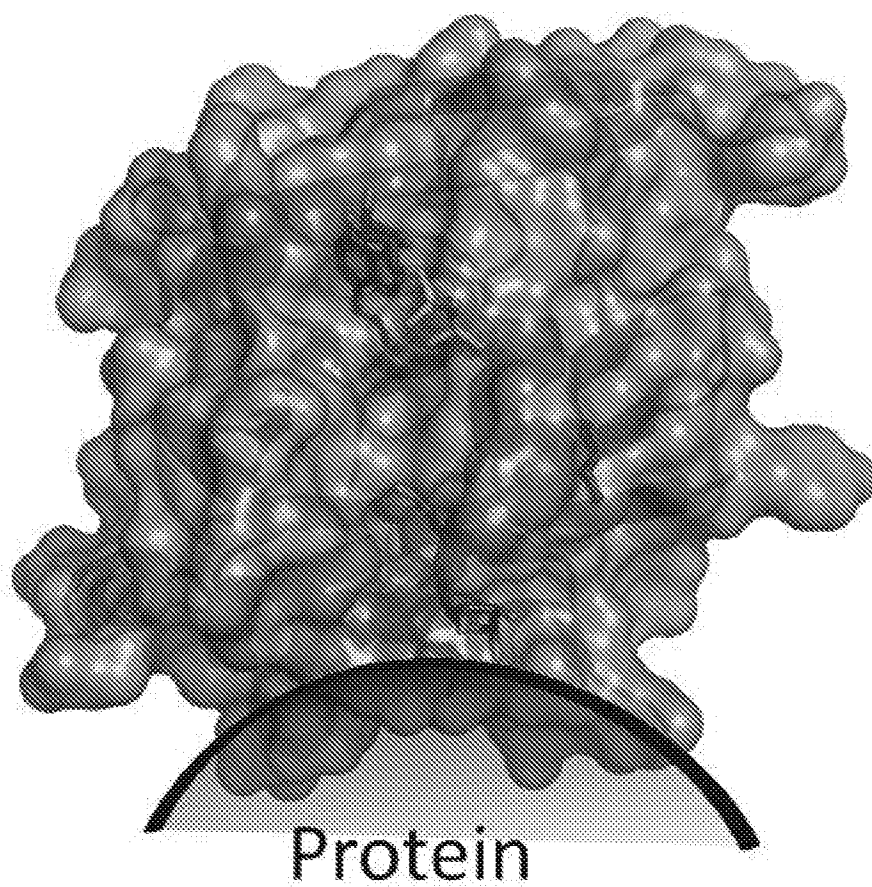


Figure 4C

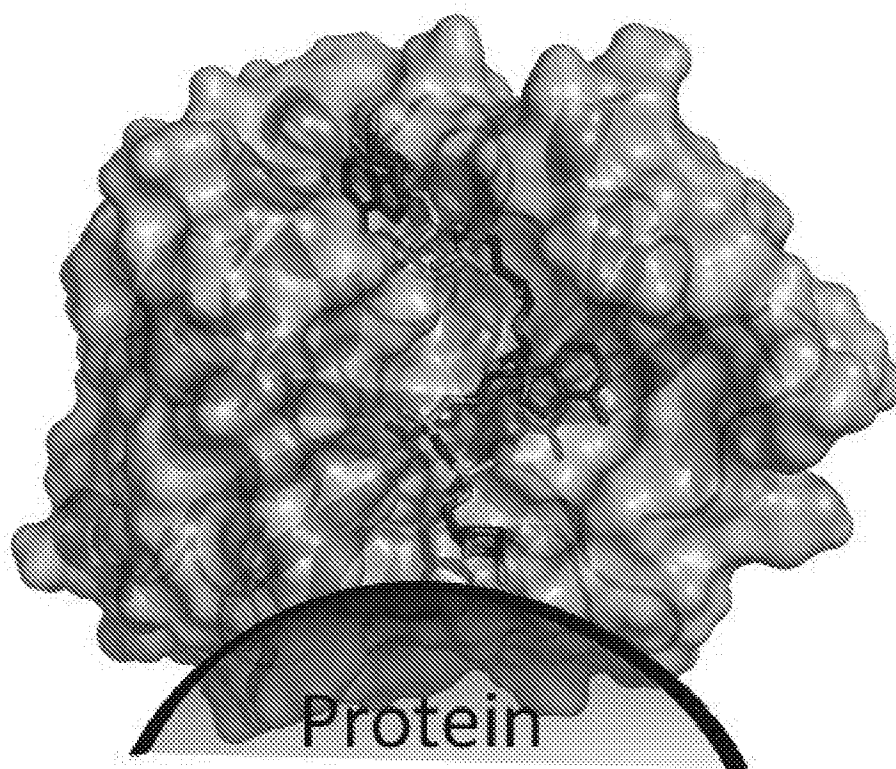


Figure 4D

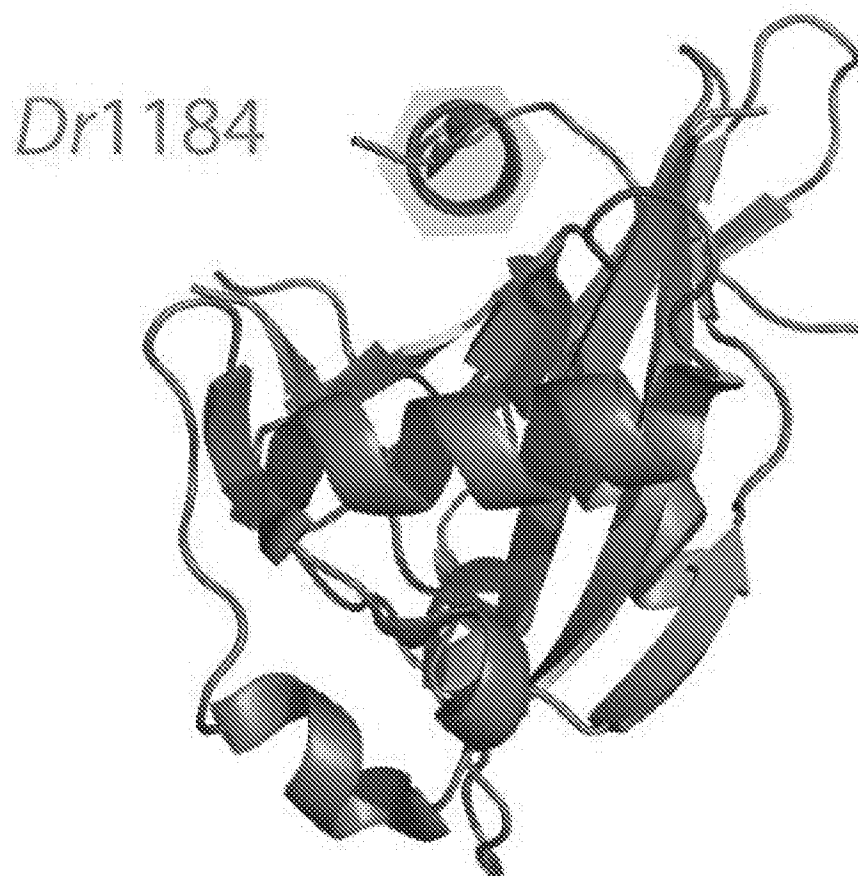
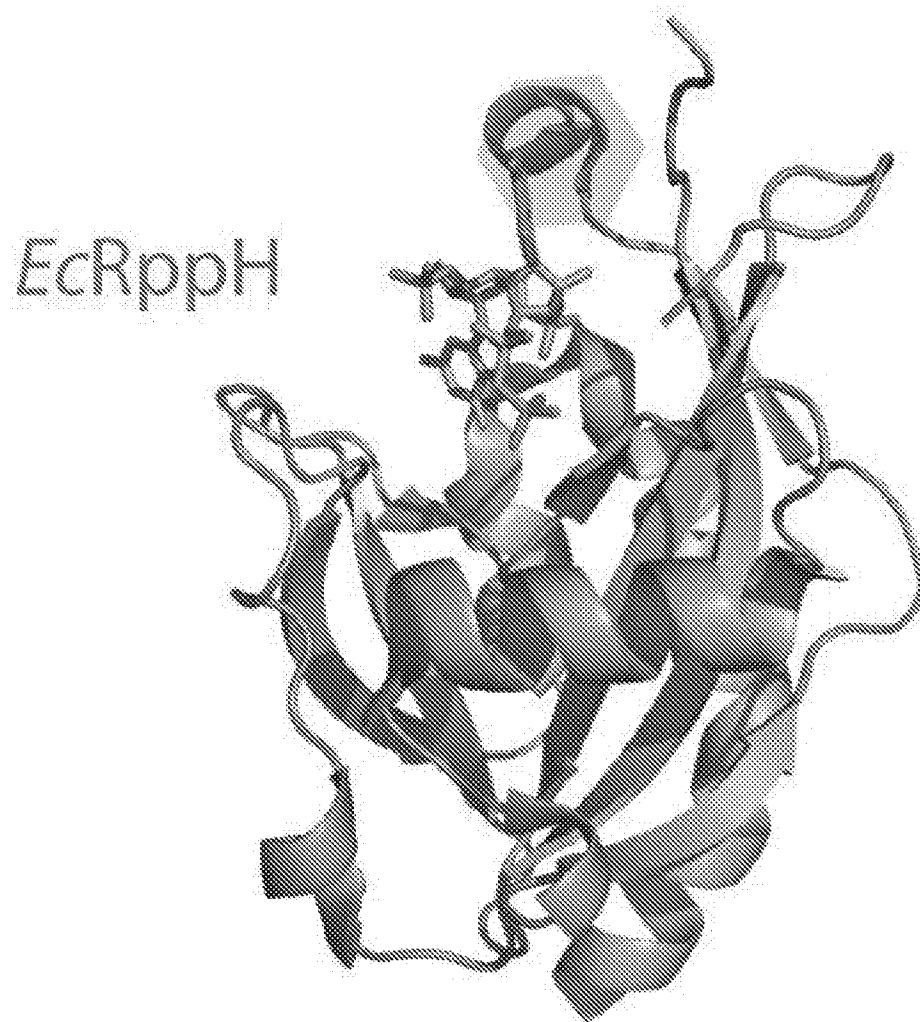


Figure 5A

**Figure 5B**

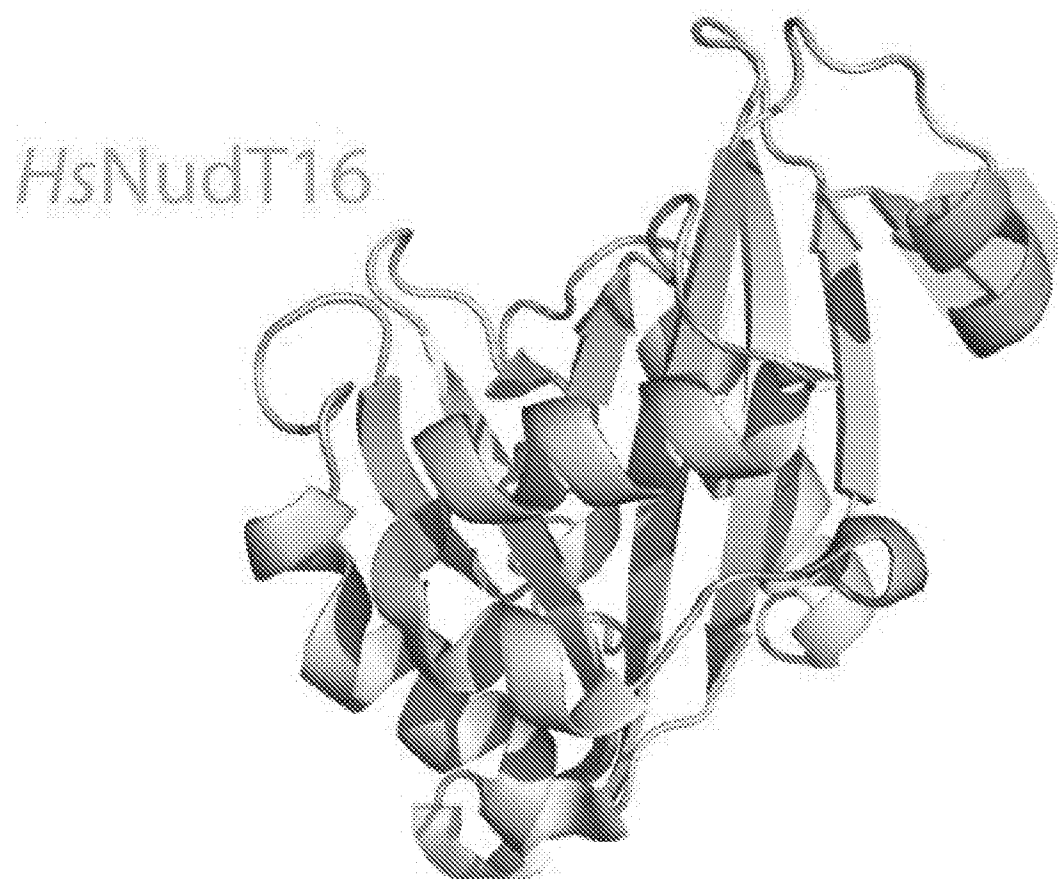


Figure 5C

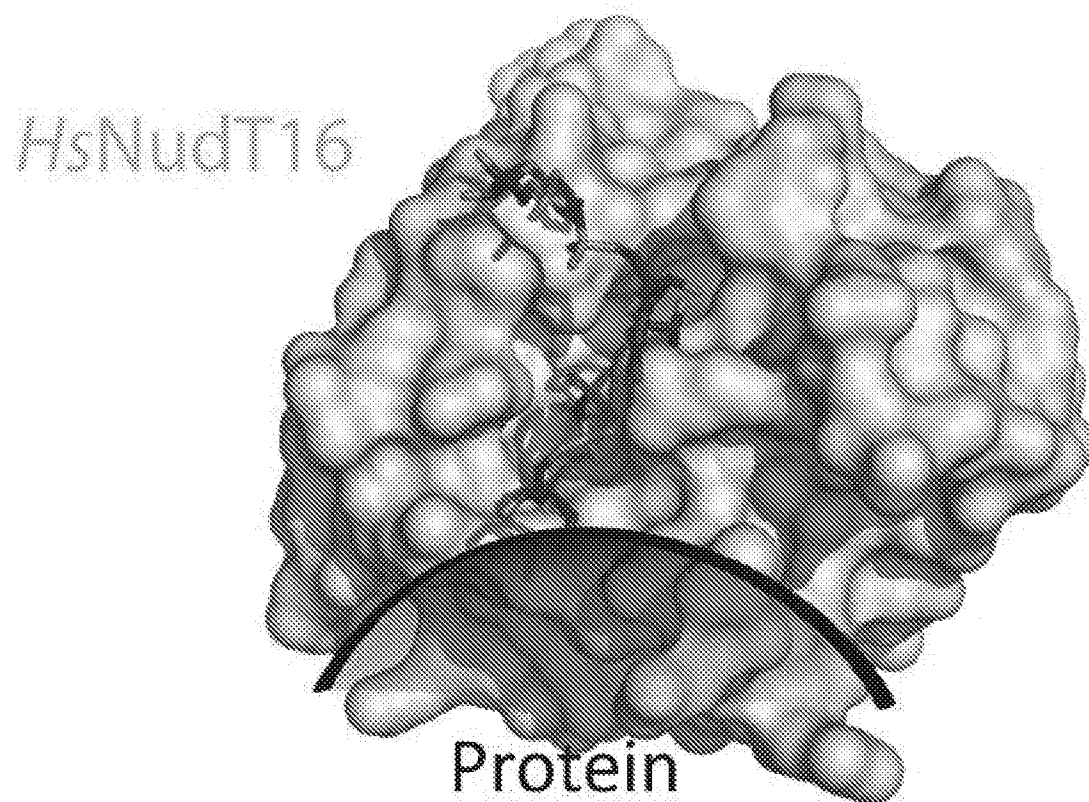
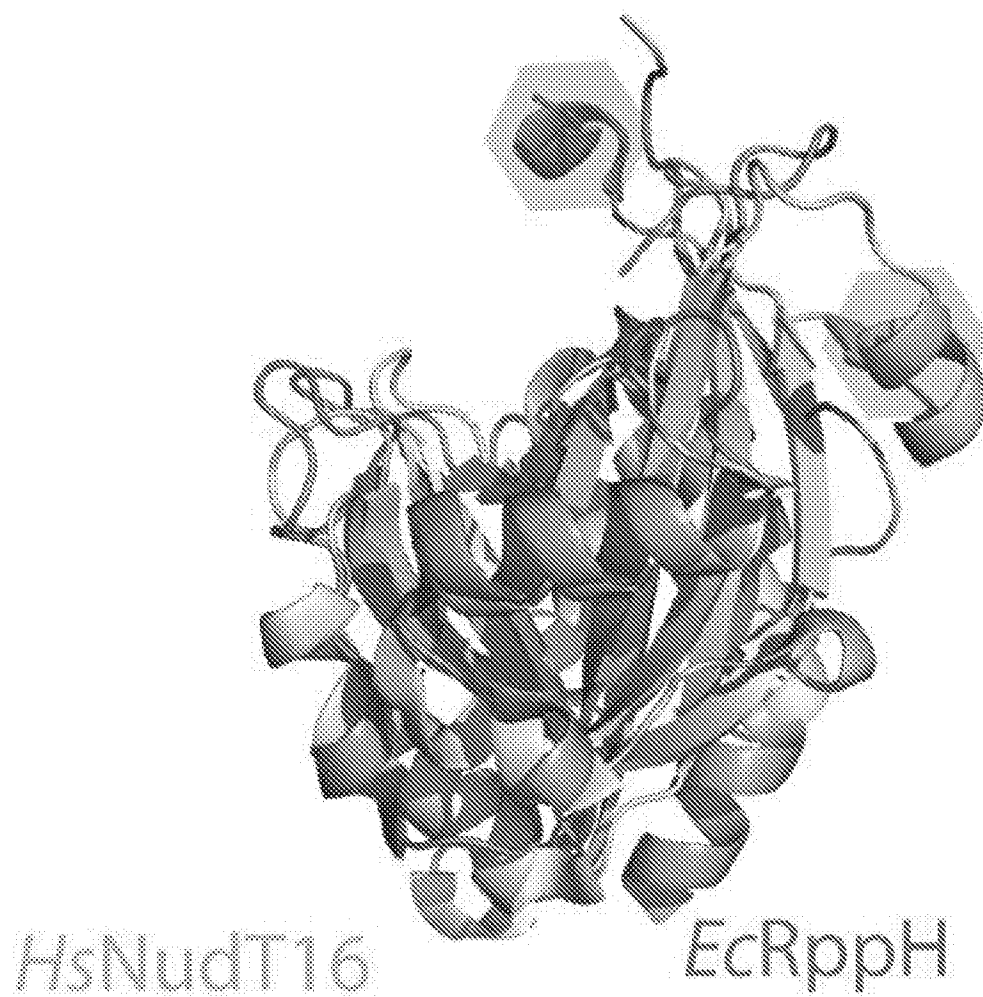
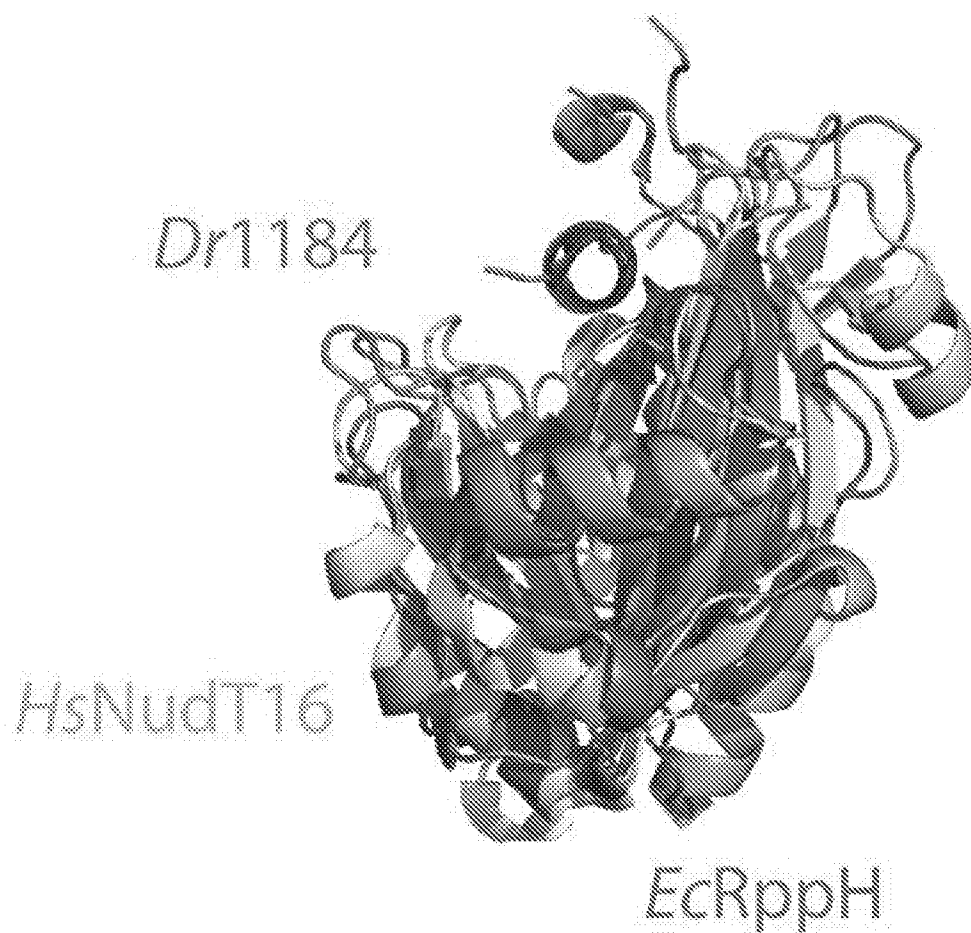


Figure 5D

**Figure 5E**

**Figure 5F**

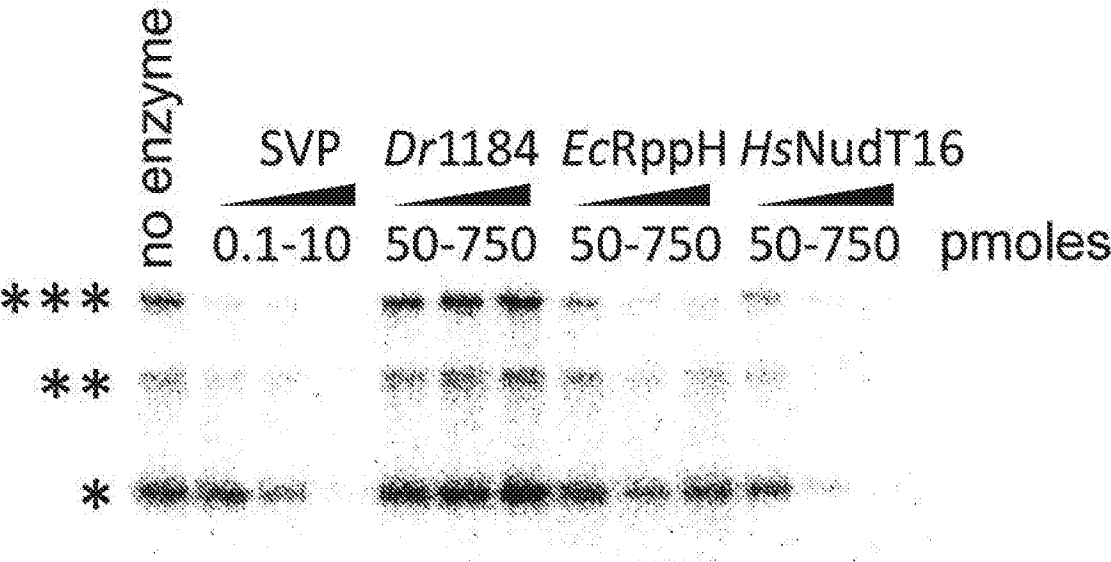


Figure 5G

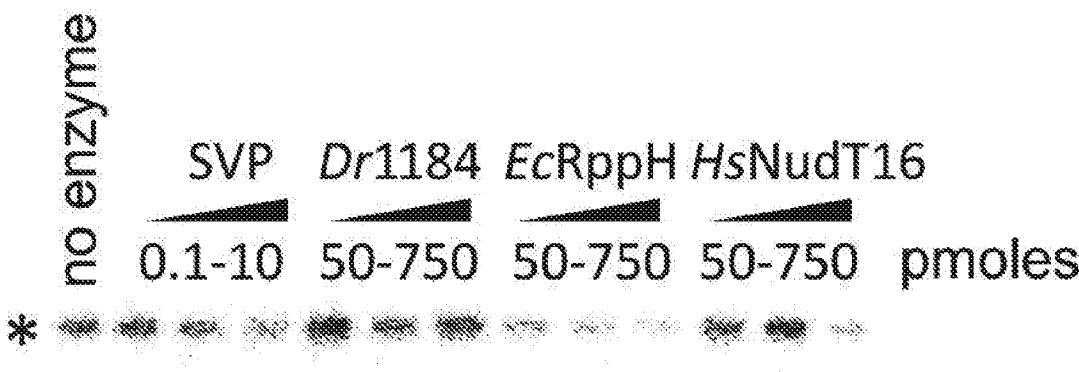


Figure 5H

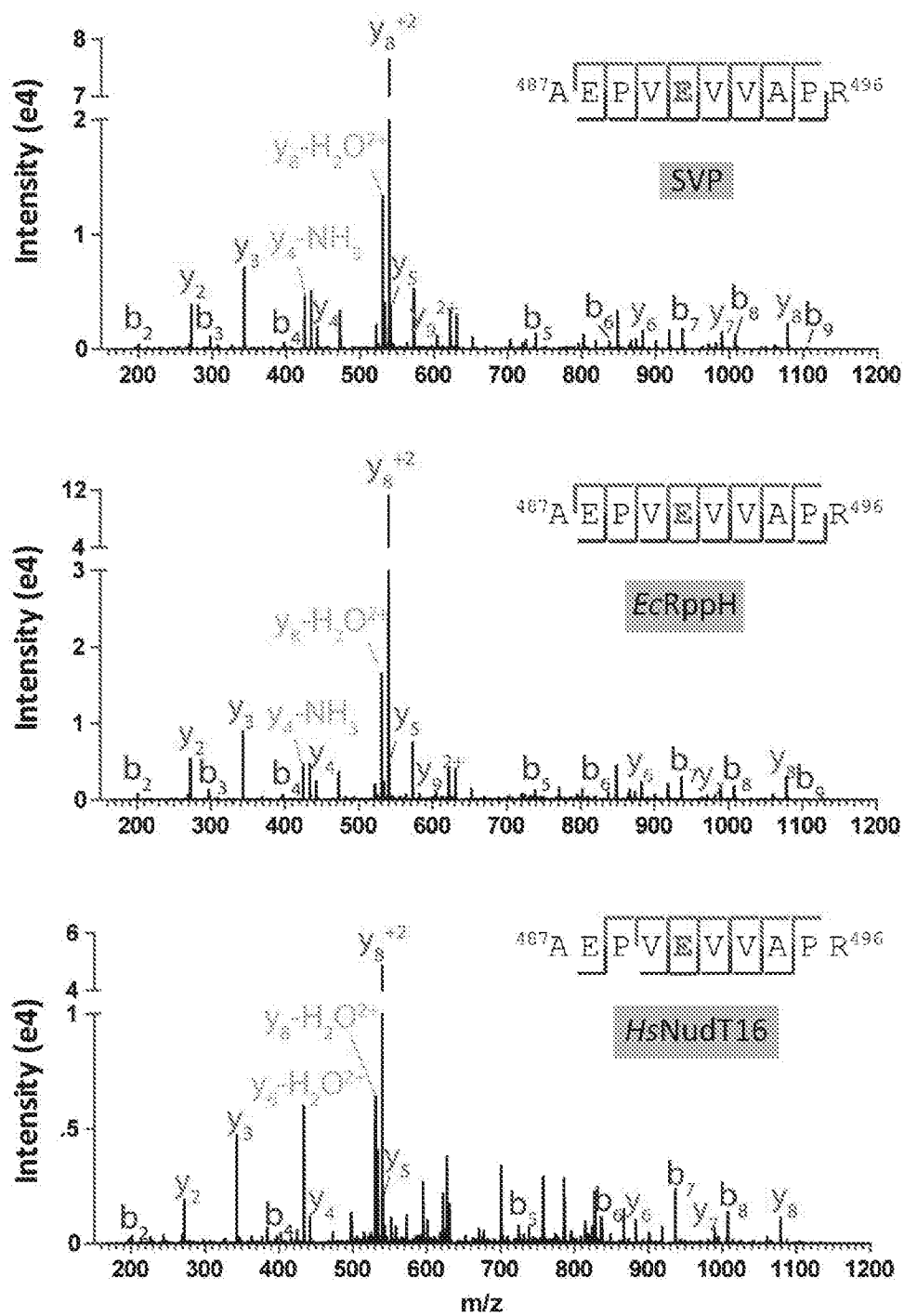


Figure 6A

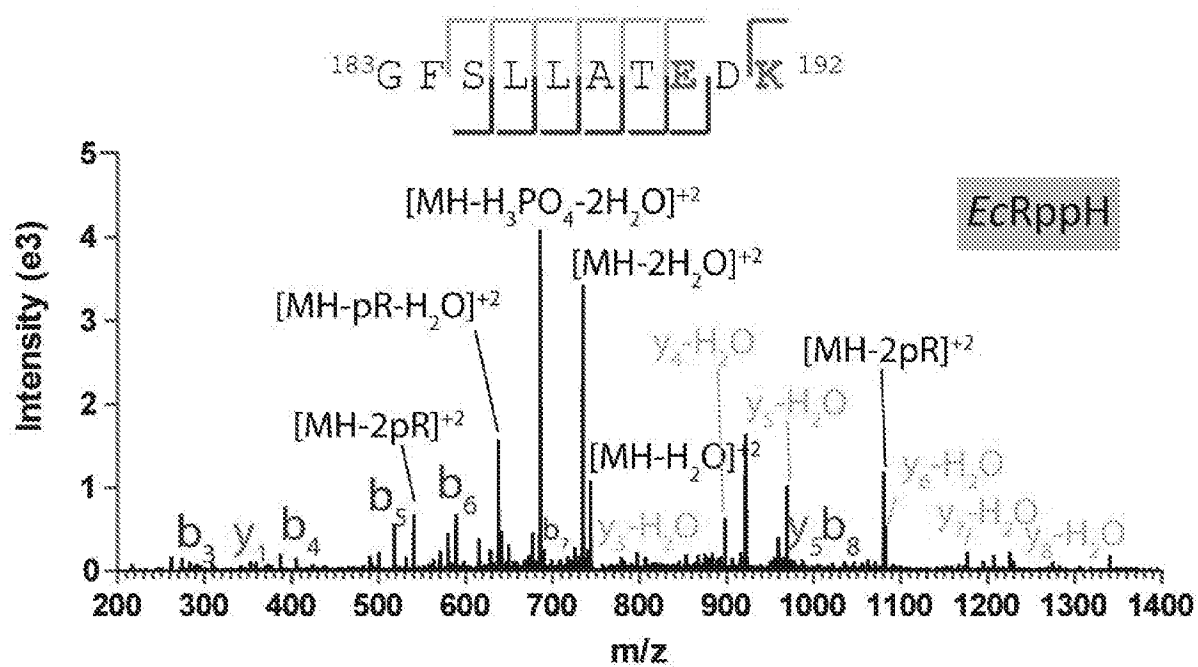


Figure 6B

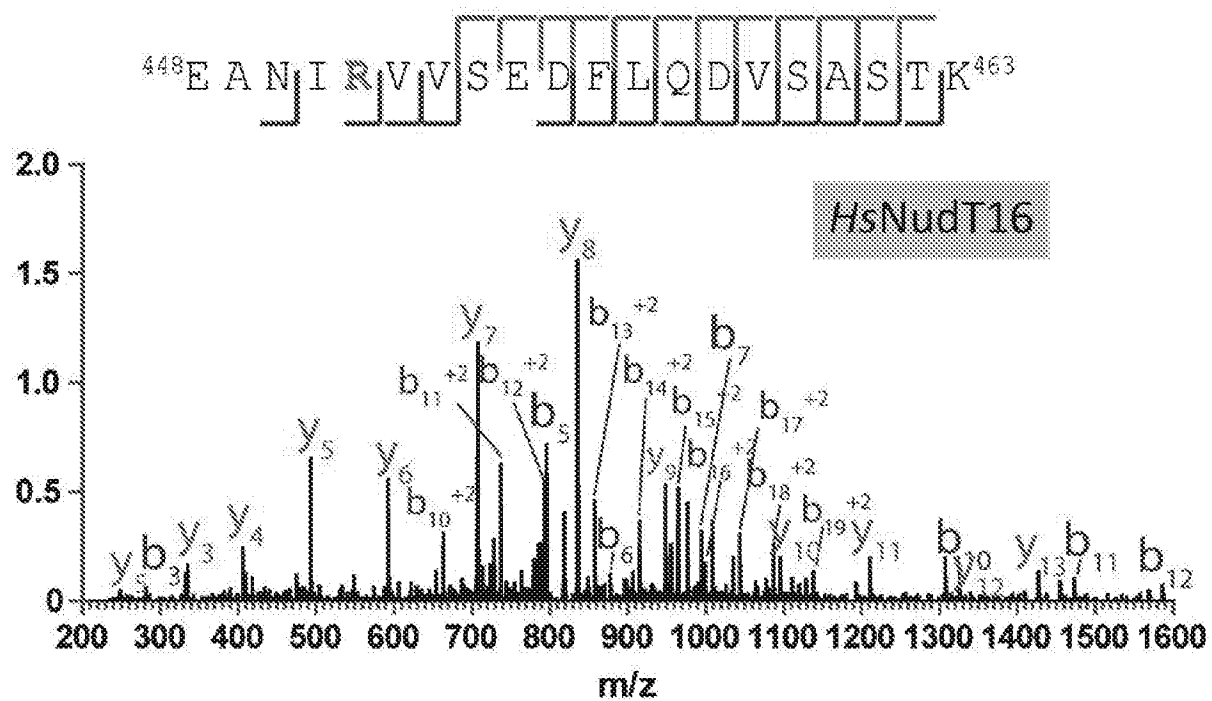


Figure 6C

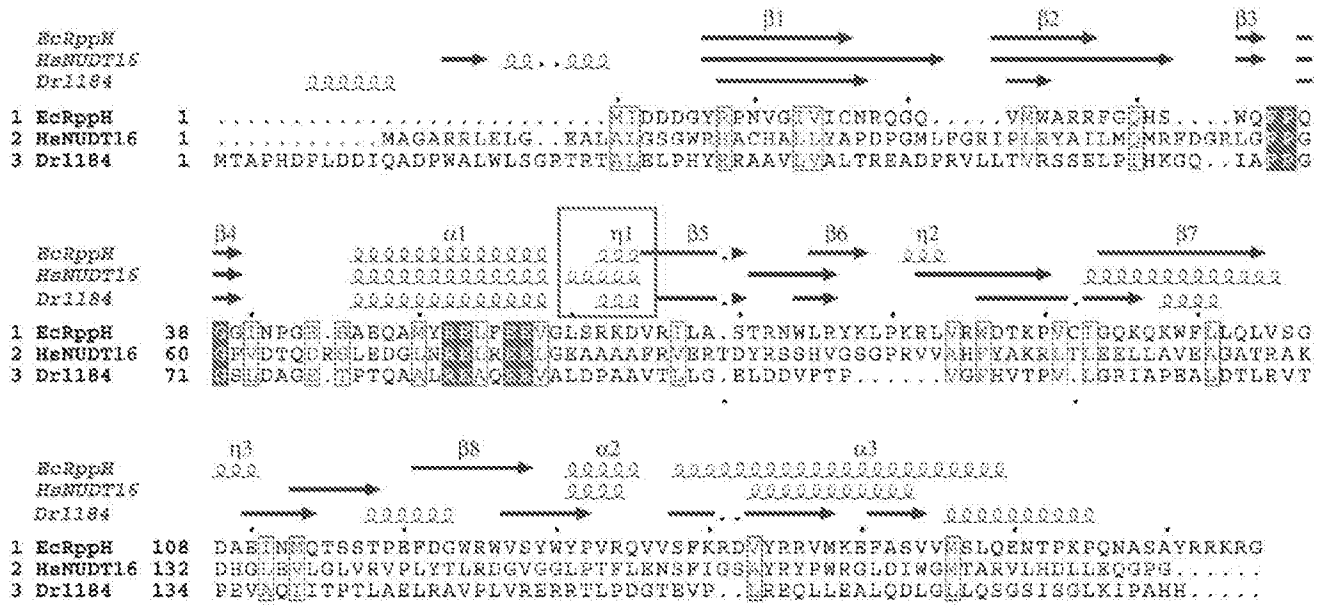
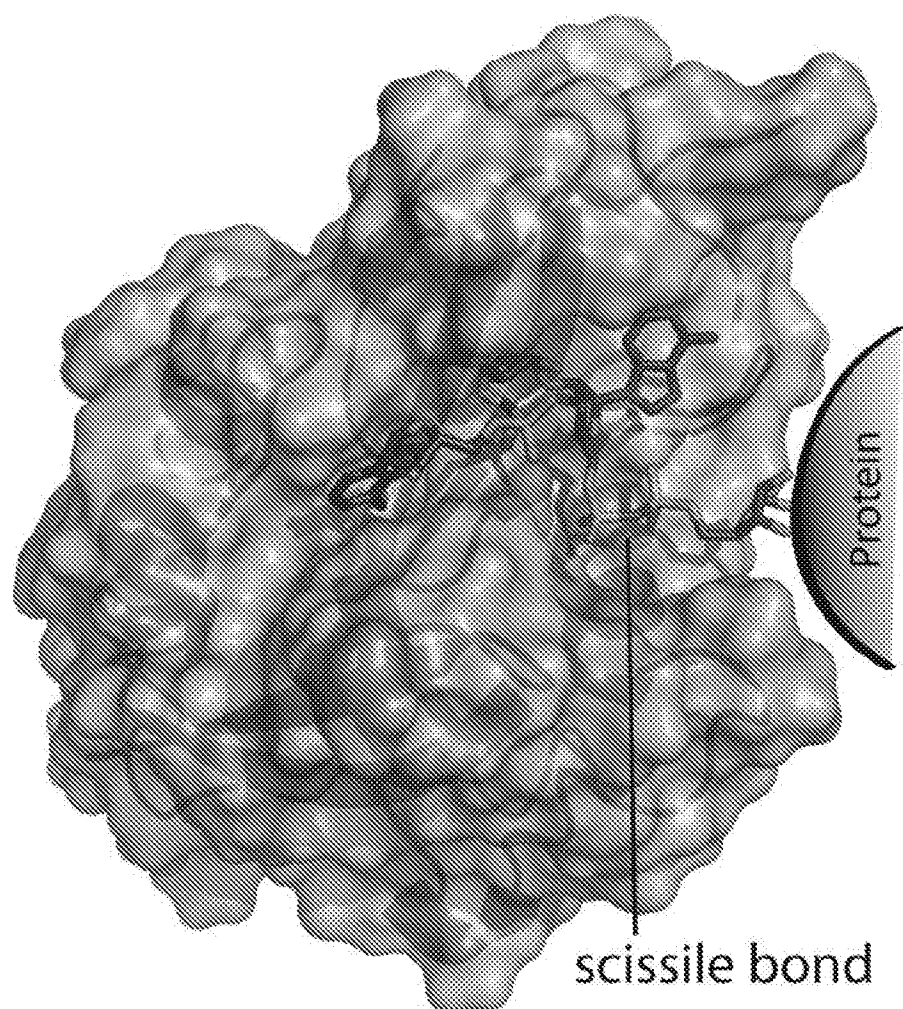


Figure 7A

**Figure 7B**

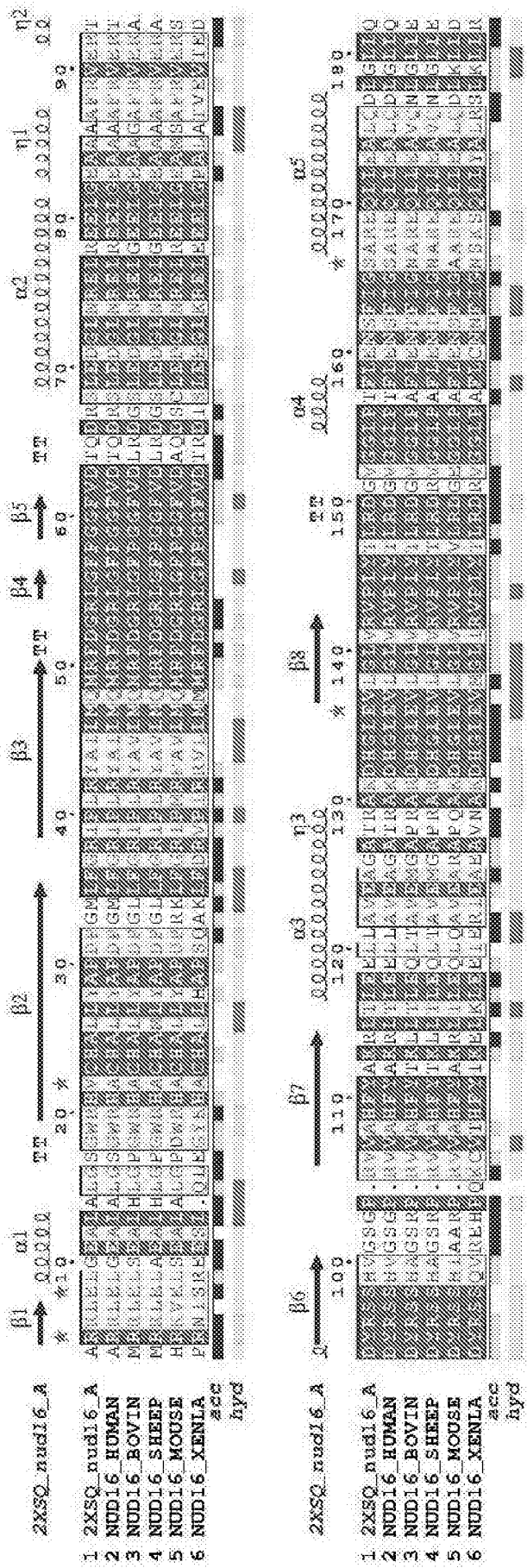


Figure 8

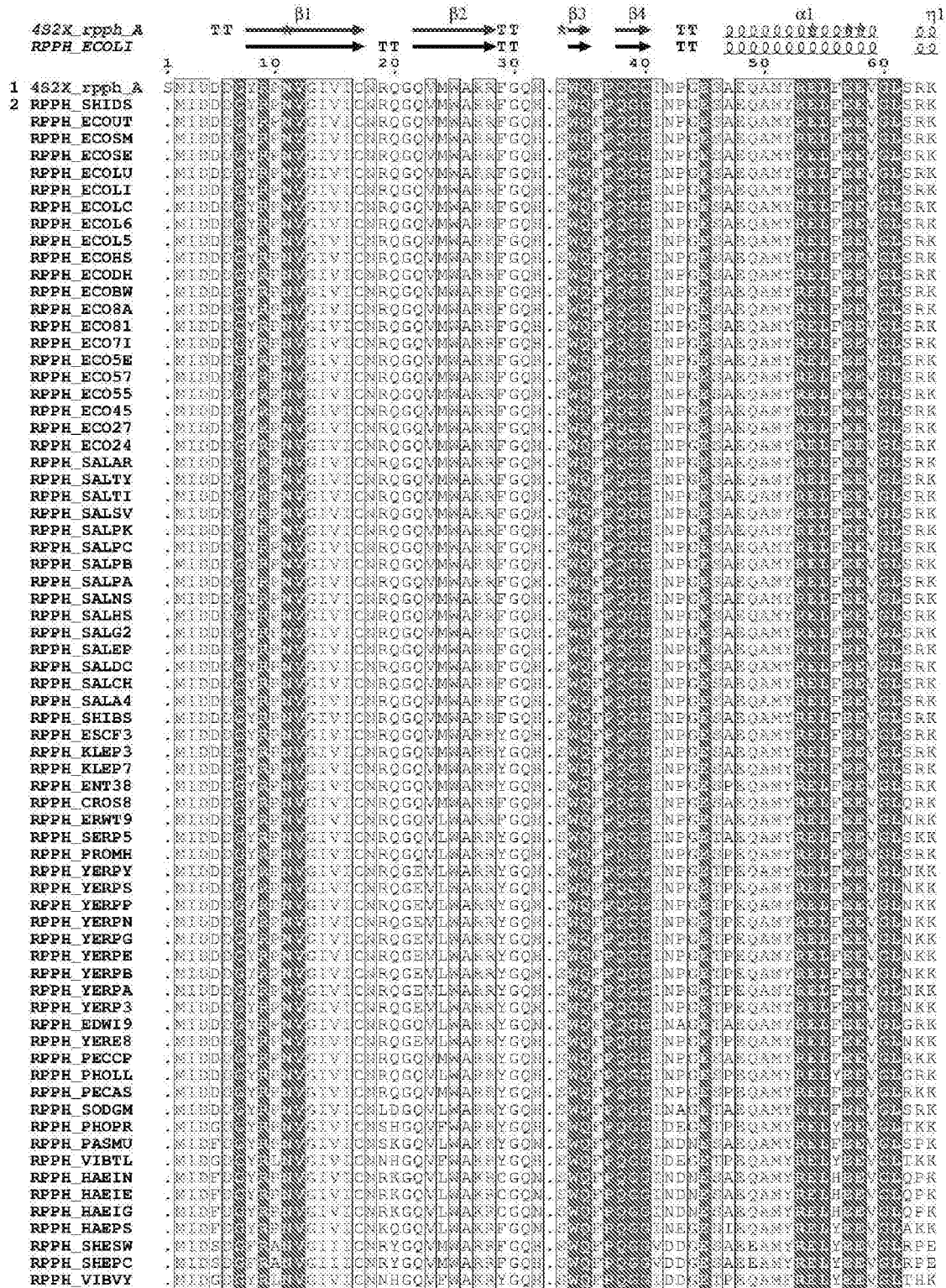


Figure 9A continued

RPPH_ACIB5	MIDRE	GIIL	ANDD	GOVL	MAKK	IG	ENAG	GGG	IQF	SE	TE	EQALF	REL	HE	LPE							
RPPH_ACIB3	MIDRE	GIIL	ANDD	GOVL	MAKK	IG	ENAG	GGG	IQF	SE	TE	EQALF	REL	HE	LPE							
RPPH_PSYWF	MIDRE	GIIL	AM	TQGO	VL	AAK	VG	EDG	GGG	GGG	IDE	SE	PLD	AMY	REL	HE	YPR					
RPPH_SACD2	MIDRE	GIIL	TD	DQGR	LM	AKK	VGG	OD	AG	GGG	IKH	NE	SE	HAL	REL	HE	CKA					
RPPH_ALKEH	MIDRE	GIIL	ANA	AAG	OV	TA	ARR	IG	ON	AG	GGG	IK	AC	TE	EE	ALF	REL	HE	APA			
RPPH_FRAP2	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_THICR	MIDRE	GIIL	IVN	KEG	KI	PM	GK	EL	Y	OD	AG	GGG	GGG	RE	NE	ET	PO	AVF	REL	HE	EPS	
RPPH_PSYCK	MIDRE	GIIL	AM	TQGO	VL	AAK	IG	ENAG	GGG	GGG	IDR	SE	PM	D	AMY	REL	HE	W	W	HPR		
RPPH_PSYA2	MIDRE	GIIL	AM	TQGO	VL	AAK	IG	ENAG	GGG	GGG	IDR	SE	PM	D	AMY	REL	HE	W	W	HPR		
RPPH_FRATM	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRATW	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRATT	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRATO	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRATH	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRATF	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_FRAT1	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	KS	ETS	GGG	GGG	VAT	SE	TP	EQ	AMY	REL	HE	RPH		
RPPH_THISH	MIDRE	GIIL	IS	MO	EB	RL	TA	AK	IG	OC	AG	GGG	IR	RD	SE	EL	D	AMY	REL	HE	GPE	
RPPH_LARHH	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	SPH	
RPPH_XANC5	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_XANAC	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_CHRVO	MIDRE	GIIL	IT	AK	NE	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_NITOC	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_METFK	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_XANOR	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_XANOM	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_STRMK	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_METCA	MIDRE	GIIL	IS	DE	GR	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_STRM5	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_NEIG1	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_NEIMF	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_NEIMB	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_LEGPL	MIDRE	GIIL	IV	DS	DR	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_LEGPH	MIDRE	GIIL	IV	DS	DR	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_LEGPC	MIDRE	GIIL	IV	DS	DR	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_LEGFA	MIDRE	GIIL	IV	DS	DR	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_NEIMA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
3 RPPH_NEIMO	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_THIDA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_METFP	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_XYLEM	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_XYLET	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_XYLEF2	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_XYLEFA	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_VARFS	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_RHOFT	MIDRE	GIIL	AM	IVIL	MR	QDR	VF	GGK	VRE	EDG	GGG	GGG	NT	DE	TP	VE	AMY	REL	HE	R	HT	LPE
RPPH_LEPCP	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_AZOSB	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_POLNA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_AROAE	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_CUPNH	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_ACIAE	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_CUPTR	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_ACIEE	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BORPA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BORPE	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BORBR	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_ACISJ	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
4 RPPH_VEREI	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_RALSO	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_DELAS	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_CUPPJ	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_HERAR	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_JANMA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_DECAR	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_RALPJ	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_POLSQ	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_RUTMC	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_POLNS	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
5 RPPH_BURCC	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BURCH	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BURCA	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	
RPPH_BURCJ	MIDRE	GIIL	IC	NR	NO	VF	GGK	VRE	ETS	GGG	GGG	IK	PC	SE	PE	AMY	REL	HE	M	W	LPO	

Figure 9A continued

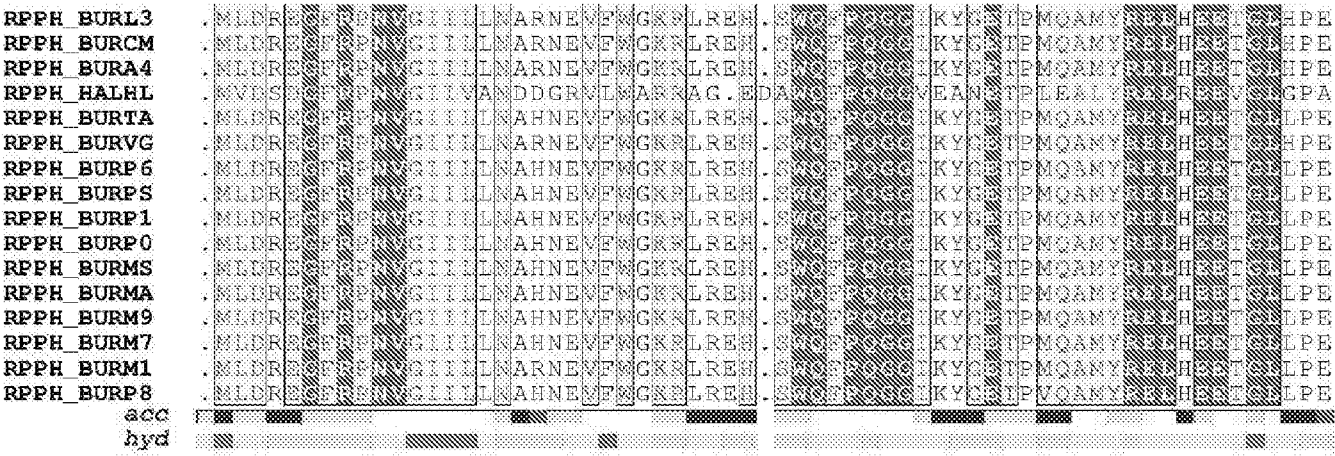


Figure 9A continued

Figure 9B

Figure 9B continued

RPPH_ACIB5	HVQIQAQEGGKLRKILHNYIENSH.....S.DFVCIQKFLIKLTA.PAKNQLNLADPPE
RPPH_ACIB3	HVQIQAQEGGKLRKILHNYIENSH.....S.DFVCIQKFLIKLTA.PAKNQLNLADPPE
RPPH_PSYWF	HVQIQAQEGGKLRKILHNYIENSH.....Q.EELCIQKFLIKLDE.PMTEHIFDPAKPE
RPPH_SACD2	DVEVVGVTGGKLRKILHNYIENSH.....K.EPKCVQKFLIKLVS.NCSAIR.DASSPAE
RPPH_ALKEH	DVEVVGVTGGKLRKILHNYIENSH.....S.RFVCIQKFLIKLVS.EKEQVQDRSERPE
RPPH_FRAP2	DVEVVGVTGGKLRKILHNYIENSH.....R.EFVCIQKFLIKLKT.SESNNI.EANDSEF
RPPH_THICR	DVRVIGRTGDLKILHNYIENSH.....S.QFVCIQKFLIKLGLS.DEEKINENTHETPE
RPPH_PSYCK	HVQIQAQEGGKLRKILHNYIENSH.....Q.YPLCIQKFLIKLDE.PMTEHIFDPAKPE
RPPH_PSYA2	HVQIQAQEGGKLRKILHNYIENSH.....Q.YPLCIQKFLIKLDE.PMTEHIFDPAKPE
RPPH_FRATM	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRATV	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRATT	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRATO	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRATH	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRATF	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_FRAT1	DVEVVGVTGGKLRKILHNYIENSH.....K.EELCIQKFLIKLKS.PESYID.DANDSPPE
RPPH_THISH	HVEVIGKTKDLKILHNYIENSH.....S.NFVCIQKFLIKLVS.DETCVREDSVQFAP
RPPH_LARHH	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_XANAC	HVEVIGKTKDLKILHNYIENSH.....E.RQVCIQKFLIKLVS.DESHKLDHDTPE
RPPH_CHRVO	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_NITOC	HVKIIGRTGDLKILHNYIENSH.....N.EELCIQKFLIKLVS.EEQDVO.NLTKPE
RPPH_METFK	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_XANOR	HVEVIGKTKDLKILHNYIENSH.....E.RQVCIQKFLIKLVS.DESHKLDHDTPE
RPPH_XANOM	HVEVIGKTKDLKILHNYIENSH.....E.RQVCIQKFLIKLVS.DESHKLDHDTPE
RPPH_STRMK	HVEVIGKTKDLKILHNYIENSH.....E.RQVCIQKFLIKLVS.DESHKLDHDTPE
RPPH_METCA	YNEIIGRTGDLKILHNYIENSH.....S.YPLCIQKFLIKLVS.TEANIR.DCSEPE
RPPH_STRM5	HVEVIGKTKDLKILHNYIENSH.....E.RQVCIQKFLIKLVS.DESHKLDHDTPE
RPPH_NEIG1	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_NEIMF	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_NEIMB	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_LEGPL	DVEVVGVTGGKLRKILHNYIENSH.....S.EELVCIQKFLIKLVS.SEQKVR.DLSDPE
RPPH_LEGPH	DVEVVGVTGGKLRKILHNYIENSH.....S.EELVCIQKFLIKLVS.SEQKVR.DLSDPE
RPPH_LEGPC	DVEVVGVTGGKLRKILHNYIENSH.....S.EELVCIQKFLIKLVS.SEQKVR.DLSDPE
RPPH_LEGFA	DVEVVGVTGGKLRKILHNYIENSH.....S.EELVCIQKFLIKLVS.SEQKVR.DLSDPE
RPPH_NEIMA	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
3 RPPH_NEIM0	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_THIDA	HVKIIGRTGDLKILHNYIENSH.....N.EELVCIQKFLIKLVS.RSDVNCIRATHPE
RPPH_METPP	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_XYLEM	HVKIIGRTGDLKILHNYIENSH.....R.SQMCIQKFLIKLVS.DESHVQ.DQSENPE
RPPH_XYLFT	HVKIIGRTGDLKILHNYIENSH.....R.SQMCIQKFLIKLVS.DESHVQ.DQSENPE
RPPH_XYLF2	HVKIIGRTGDLKILHNYIENSH.....R.SQMCIQKFLIKLVS.DESHVQ.DQSENPE
RPPH_XYLFA	HVKIIGRTGDLKILHNYIENSH.....R.SQMCIQKFLIKLVS.DESHVQ.DQSENPE
RPPH_VARPS	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_RHOFT	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_LEPCP	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_AZOSB	HVKIIGRTGDLKILHNYIENSH.....W.RNTYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_POLNA	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_AROAE	HVKIIGRTGDLKILHNYIENSH.....W.RNTYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_CUPNH	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_ACIAE	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_CUPTR	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_ACIEE	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_BORPA	HVKIIGRTGDLKILHNYIENSH.....W.RNTYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_BORPE	HVKIIGRTGDLKILHNYIENSH.....W.RNTYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_BORER	HVKIIGRTGDLKILHNYIENSH.....W.RNTYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_ACISJ	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
4 RPPH_VEREI	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_RALSO	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_DELAS	HVKIIGRTGDLKILHNYIENSH.....A.ECHYKQKFLIKLVS.HOWDEN.RATHPE
RPPH_CUPPJ	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_HERAR	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_JANMA	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_DECAR	HVKIIGRTGDLKILHNYIENSH.....W.KESYRQKFLIKLVS.RSDVNCIRATHPE
RPPH_RALPJ	HVKIIGRTGDLKILHNYIENSH.....I.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_POLSQ	HVKIIGRTGDLKILHNYIENSH.....E.EELVCIQKFLIKLVS.LSDVNCIRATHPE
RPPH_RUTMC	HVKIIGRTGDLKILHNYIENSH.....Q.EELVCIQKFLIKLVS.NENN.KINMETOVPE
RPPH_POLNS	HVKIIGRTGDLKILHNYIENSH.....E.EELVCIQKFLIKLVS.LSDVNCIRATHPE
5 RPPH_BURCC	HVKIIGRTGDLKILHNYIENSH.....V.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_BURCH	HVKIIGRTGDLKILHNYIENSH.....V.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_BURCA	HVKIIGRTGDLKILHNYIENSH.....V.ECHYKQKFLIKLVS.RSDVNCIRATHPE
RPPH_BURCJ	HVKIIGRTGDLKILHNYIENSH.....V.ECHYKQKFLIKLVS.RSDVNCIRATHPE

Figure 9B continued

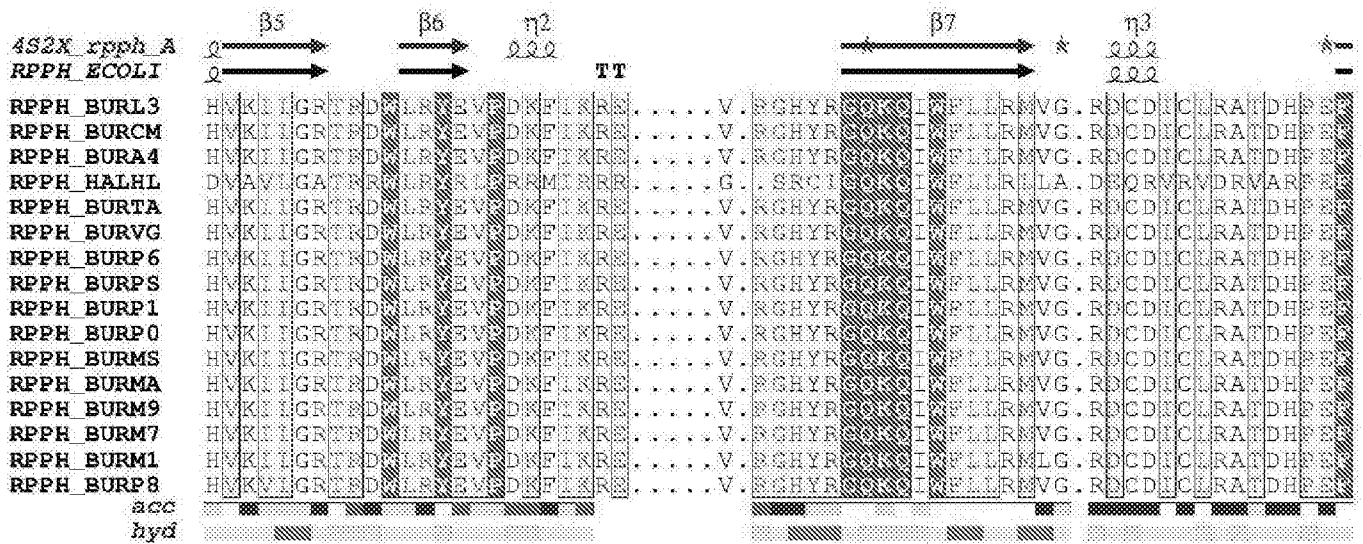


Figure 9B continued

		130	140	150	160						
1	4S2K_rpph_A	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLAA
2	RPPH_SHIDS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOUT	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOSM	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOSE	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOLU	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOLI	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOLC	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOL6	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOL5	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOHS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECODH	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECOBW	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO8A	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO81	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO7I	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO5E	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO57	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO55	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO45	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO27	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ECO24	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_SALAR	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALTY	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALTI	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALSV	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALPK	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALPC	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALPB	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALPA	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALNS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALHS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALG2	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALEP	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALDC	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALCH	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SALA4	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_SHIBS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_ESCF3	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MALQD
	RPPH_KLEP3	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLAE
	RPPH_KLEP7	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLAE
	RPPH_ENT38	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MLAPQE
	RPPH_CROS8	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MPVQE
	RPPH_ERWT9	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MPVQE
	RPPH_SERP5	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MPVQE
	RPPH_PROMH	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MPVQE
	RPPH_YERPY	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPP	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPN	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPG	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPE	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPB	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERPA	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_YERP3	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_EDWI9	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MSLQE
	RPPH_YERE8	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_PECCP	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	IN	PRILLQE
	RPPH_PHOLL	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	MPVQE
	RPPH_PECAS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ES	PRILLQE
	RPPH_SODGM	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	SG	MPVQE
	RPPH_PHOPR	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	MPVQE
	RPPH_PASMU	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	ASV	FDGAR
	RPPH_VIBTL	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AS	LAMPFKE
	RPPH_HAEIN	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AS	IFTDNP
	RPPH_HAEIE	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AS	IFTDNP
	RPPH_HAEIG	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AS	IFTDNP
	RPPH_HAEFS	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	QVLMGEQK
	RPPH_SHESW	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	TALSFTQ
	RPPH_SHEPC	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AA	TALSFTQ
	RPPH_VIBVY	EG	VS	RY	YFVRQVVS	AKDV	RR	YV	KET	AS	LAMPFRE

Figure 9C

RPPH_VIBVU
 RPPH_ALISH
 RPPH_SHESA
 RPPH_VIBCH
 RPPH_SHEON
 RPPH_HAEI8
 RPPH_SHEB9
 RPPH_SHEB8
 RPPH_SHEB5
 RPPH_SHEB2
 RPPH_VIBFM
 RPPH_VIBF1
 RPPH_SHESR
 RPPH_SHESM
 RPPH_VIBCB
 RPPH_VIBPA
 RPPH_SHEFN
 RPPH_SHEHH
 RPPH_ACTS2
 RPPH_SHEPA
 RPPH_SHEPW
 RPPH_MANSM
 RPPH_SHELP
 RPPH_HAEDU
 RPPH_ACTPJ
 RPPH_ACTP7
 RPPH_HISS2
 RPPH_HAES1
 RPPH_SHEWM
 RPPH_SHESH
 RPPH_HAMD5
 RPPH_ACTP2
 RPPH_TOLAT
 RPPH_SHEDO
 RPPH_PSEA6
 RPPH_PSEHT
 RPPH_AERHH
 RPPH_COLP3
 RPPH_AERS4
 RPPH_PSYIN
 RPPH_IDILO
 RPPH_BAUCH
 RPPH_BLOFL
 RPPH_PSEU2
 RPPH_PSESM
 RPPH_PSE14
 RPPH_PSEFS
 RPPH_PSEPW
 RPPH_PSEPK
 RPPH_PSEPG
 RPPH_PSEE4
 RPPH_PSEAE
 RPPH_PSEAB
 RPPH_PSEAB
 RPPH_PSEA8
 RPPH_PSEA7
 RPPH_PSEPF
 RPPH_PSEF5
 RPPH_PSEP1
 RPPH_BLOPB
 RPPH_HAHCH
 RPPH_PSEMY
 RPPH_AZOV2
 RPPH_PSEU5
 RPPH_CARHU
 RPPH_CELJV
 RPPH_MARMS
 RPPH_ACIA2
 RPPH_TERTT
 RPPH_ACIBY
 RPPH_ACIBT
 RPPH_ACIBS
 RPPH_ACIBC

Figure 9C continued

Figure 9C continued

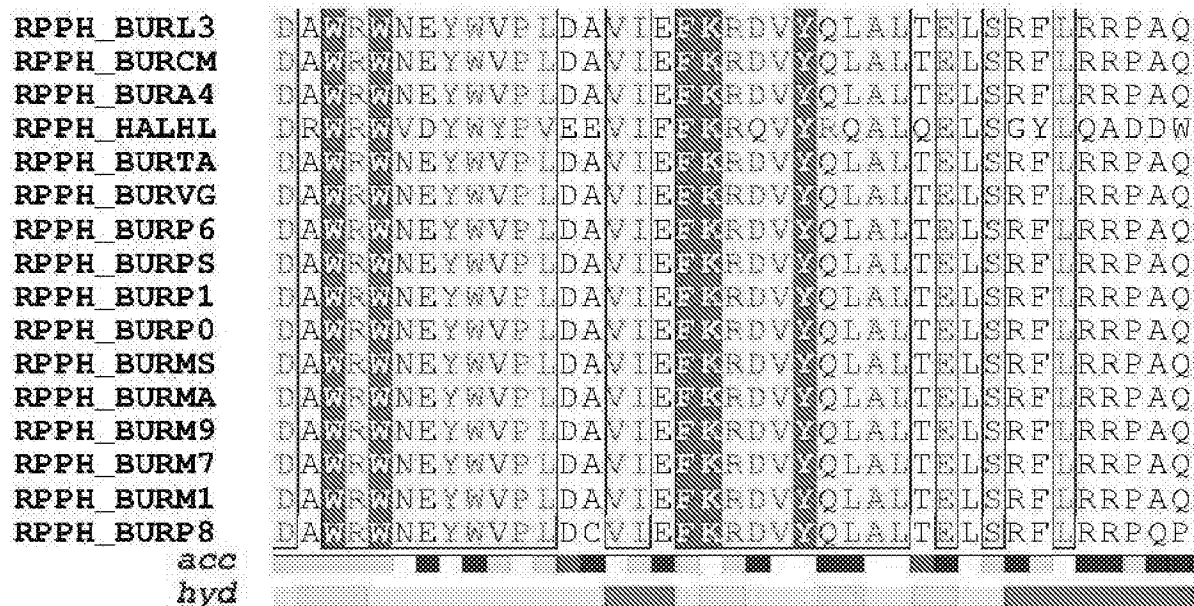


Figure 9C continued