Electronic Supplementary Material (ESI) for Chemical Communications.

Supplementary Information

Bio-coordination of bismuth in *Helicobacter pylori* revealed by immobilized metal affinity chromatography

Yuchuan Wang,‡[a] Cheuk-Nam Tsang,‡[a] Feng Xu,[b] Pak-Wing Kong,[a] Ligang Hu,[a] Junwen Wang,[b] Ivan Keung Chu,[a] Hongyan Li[a] and Hongzhe Sun[a,]*

[a] Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, P. R. China. E-mail: hsun@hku.hk.

[b] Centre for Genomic Sciences & Department of Biochemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, P. R. China.

‡ These authors contributed equally to this work.

Supplementary materials and methods

Bacterial culture and protein extraction

H. pylori 26695 (American type culture collection, ATCC 700392) glycerol stock was recovered from -80°C and cultured on Columbia agar plates supplemented with 7% laked horse blood (Oxoid) and selective antibiotics (Oxoid). The agar plates were incubated at 37°C for 3 days under microaerophilic conditions (5% CO_2, 4% O_2 and 91% N_2) generated by CampyGen (Oxoid). Cells were then transferred into Brucella broth (BD) supplemented with 0.2% β-cyclodextrin (Sigma) and cultured with 105 rpm agitation at 37°C. Liquid cultures grown to mid-log phase (OD_{600} = 0.5) were harvested by centrifugation at 4,000 g, and washed with ice-cold PBS buffer for three times. Cell pellets were resuspended in ice-cold Buffer A (20 mM Hepes, 500 mM NaCl, 1 mM TCEP, 0.1% TritonX-100, pH 7.2), then lysed by sonication and ultracentrifuged (15,000 g) to remove cell debris.

Isolation of Bi-binding peptides by IMAC

A home-made Bi-NTA column (Qiagen) with a bed volume of 0.5 mL was applied for extracting Bi-binding peptides in *H. pylori* 26695. The column was washed thoroughly and equilibrated with 20 column volumes (C.V.) of Buffer A and then loaded with 2 mg of filtered *H. pylori* protein extracts. On-column protein digestion was performed with 1 mg/mL trypsin in 20 mM Tris-HCl buffer (pH 7.2) at 37°C for 1 hour. The metal unbound fragments and excessive proteolytic enzymes were washed off with 20 C.V. of Buffer A. Metal-binding peptides were finally eluted with 6 C.V. of Buffer A supplemented with 50 mM EDTA.

Bi-binding peptides from the IMAC eluent were desalted and pre-concentrated using C-18 reverse phase chromatography prior to LC-MS/MS analysis. To minimize peptide oxidation and modification, eluted peptides were acidified and desalted by reverse phase chromatography immediately after IMAC separation. The eluent was acidified with 1% TFA (Fluka) for 10 min in ice bath before loading on Sep Pak classic C18 cartridge (Waters) for purification. The peptides eluted using a three-step-gradient of 20%, 40% and 60% ACN in 0.1% TFA were combined and lyophilized overnight. Experiments were done in three biological replicates.

Bi-binding peptide analysis and data processing

The dried peptides were re-dissolved in 5% ACN/0.1% formic acid in water and fractionated by capillary-flow reversed phase liquid chromatography (RPLC), and analyzed online with a QSTAR XL Q-TOF mass spectrometer (AB sciex; Foster City, CA) installed with software Analyst QS 1.1. The acquisition parameters were as follows: nanospray voltage, 3000 V; DP, 60 V; FP, 220 V; DP2, 25 V; collision gas (CG), 5; GS1, 0 psi; GS2, 0 psi; CUR, 20. CID spectra were acquired in the IDA mode with scan cycles set to perform a 1-second full scan over a *m/z* range of 400-1600, followed by four 1-second MS/MS scan of the four most abundant peaks which exceed 10 counts and carry charge state between +2 to +4 within 100-1500 *m/z*. Dynamic exclusion time of acquired ions was set at 120 seconds.

MS/MS spectra were searched against theoretical spectra generated from the sequences of *H. pylori* 26695 in Uniprot database using the Paragon algorithm in the ProteinPilot 2.0 software (Applied Biosystems, Framingham, MA). Trypsin was set as the enzyme used. Precursor mass accuracy and product ion mass accuracy were pre-determined using QSTAR electrospray in the software. The identified peptides from the Paragon algorithm were grouped into minimal non-redundant protein sets by the ProGroup algorithm. For protein identification, a minimal unused ProtScore of 2.0 was required, determined from the percent confidence of a single (99% confidence) or a set of identified peptides.

Determination of binding constants of bismuth towards single amino acids

The binding constants of bismuth towards 20 single amino acids were determined with nano isothermal titration calorimetry (Nano ITC, TA instruments). The amino acid solutions were directly prepared from powder reagents (Sigma) in PIPES buffer (10 mM PIPES, 100 mM NaCl, pH 6.4), and diluted to appropriate concentrations (30-50 μM) supplemented with 0.3 mM TCEP. The metal titrant (Bi-NTA, 2 mM) was dissolved in water. Typically, 100 μL of metals were titrated into 1 mL of amino acid solutions over 100 s with a 5 min interval between each injection. Twenty injections were made in total. The background signals were subtracted from the raw ITC data to account for the dilution heat, which was obtained from the titration of the identical metal titrant only to the buffer solution in the same cell. The ITC data were fitted with a simple one-site binding model.

Functional enrichment analysis

In order to functional categorize IMAC-isolated Bi-binding proteins in *H. pylori*, we performed GO (gene ontology) enrichment analysis to identify statistically enriched GO terms by using an in-house enrichment tool programmed in Python. GO annotations of *H. pylori* 26695 proteins were obtained from the Uniprot database,[1] Blast searching[2] and Ortholog detection[3] using *Pseudomonas aeruginosa* (strain PA7) and *Escherichia coli* (strain ATCC 8739 / DSM 1576 / Crooks) as model organisms. The enrichment *p*-values were calculated on the basis of hypergeometric distribution, using the numbers of *H. pylori* 26695 proteins as the background list. The Benjamini-Hochberg method was applied to adjust *p*-values for false discovery rate (FDR) control. The enriched GO terms were selected with a significance threshold of *p*-values < 0.05 and FDR < 0.05.

Biological network analysis

The protein-protein interaction (PPI) networks of *H. pylori* 26695 were constructed by combining the high confidence (combined score > 700) interactions of *H. pylori* 26695 extracted from String database,[4] the yeast two-hybrid screened high quality PPIs[5, 6] and a set of literature-curated binary protein interactions in *H. pylori*.[6-9] The PPI networks were visualized with Cytoscape v2.8.3.

Supplementary Figures

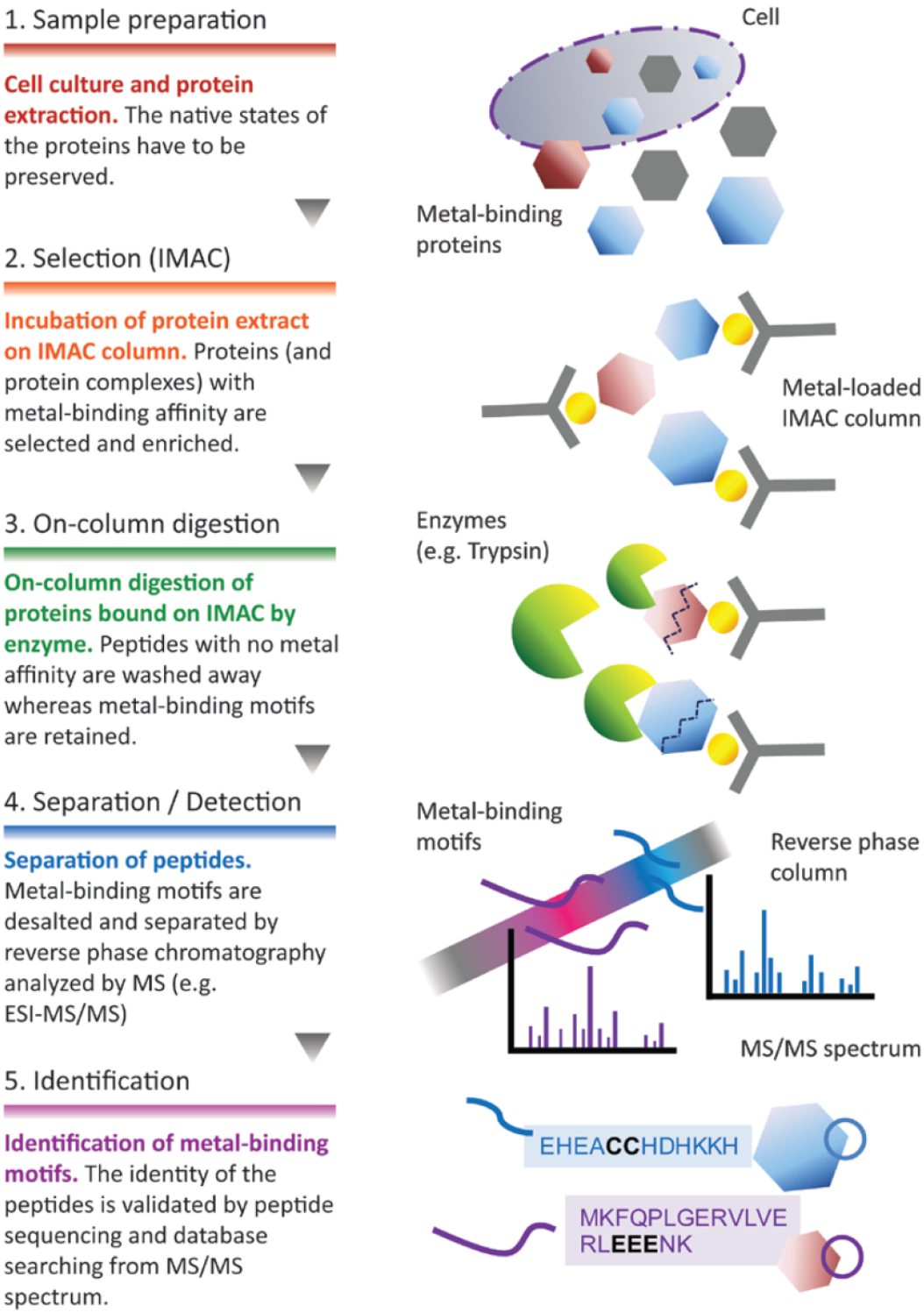

Fig. S1 Mining metal-binding motifs by IMAC-based metalloproteomics (Adapted from ref.[10]).

Fig. S2 Binding constants of Bi(III) with 20 single amino acids determined by isothermal titration calorimetry. It is noted that that Bi(III) possesses a much higher affinity towards cysteine residue than any other amino acids, in agreement with IMAC-based peptide analysis (Fig. 1A), indicating that Bi(III) utilizes a cysteine-oriented manner to interact with the protein targets in *H. pylori*.

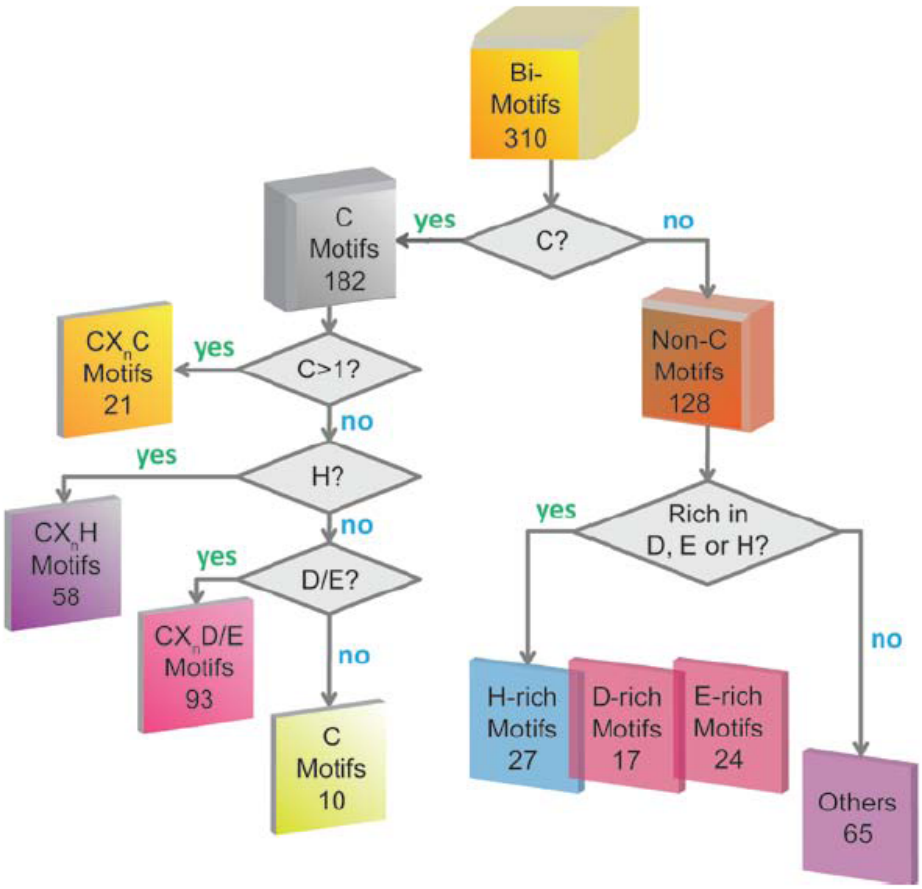

Fig. S3 Classification of Bi-binding peptides. Bi-binding peptides are classified into 2 major categories and 8 sub-categories according to the occurrence of four putative Bi-binding amino acids Cys (C), His (H), Glu (E) and Asp (D).

Structures of *H. pylori* proteins with Bi-binding motifs identified

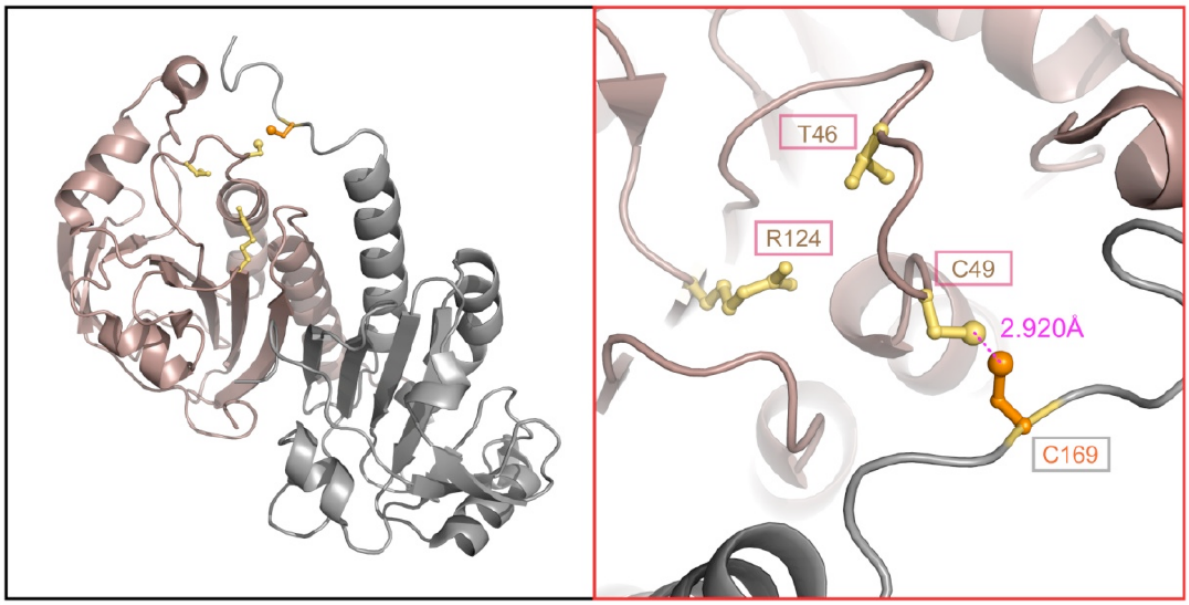

Fig. S4 Crystal structure of *H. pylori* thioredoxin reductase (TsaA, AhpC, HP1563; PDB 1ZOF). The dimeric form of *Hp*TsaA is shown on the left with two monomers displayed in gray and deep red respectively. The catalytic triad formed by C49 (peroxidatic cysteine), T46 and R124 are highlighted in yellow and C169, identified from the Bi-binding peptide (161)HFEEHGEV**C**PAGW(173), is shown in orange. C49 and C169 belonging to different subunits are linked by a disulfide bridge with bond length 2.9 Å in the crystal structure.

AhpC is highly expressed in *H. pylori* as the third-most abundant cellular protein and acts as a peroxide reductase and also a molecular chaperone for prevention of protein misfolding under oxidative stress.[11] It functions as a dimer, or a pentamer of dimers, and carries one conserved resolving cysteine which is found in the Bi-binding peptide (161)HFEEHGEV**C**PAGW(173).[12] Cys169 initiates the recycle of the oxidized peroxidatic cysteine by forming an intersubunit disulphide bond that is subjected to be reduced by thioredoxin-1 (HP0824) and the replacement of the resolving Cys with Ser retards the rate of reduction.[13]

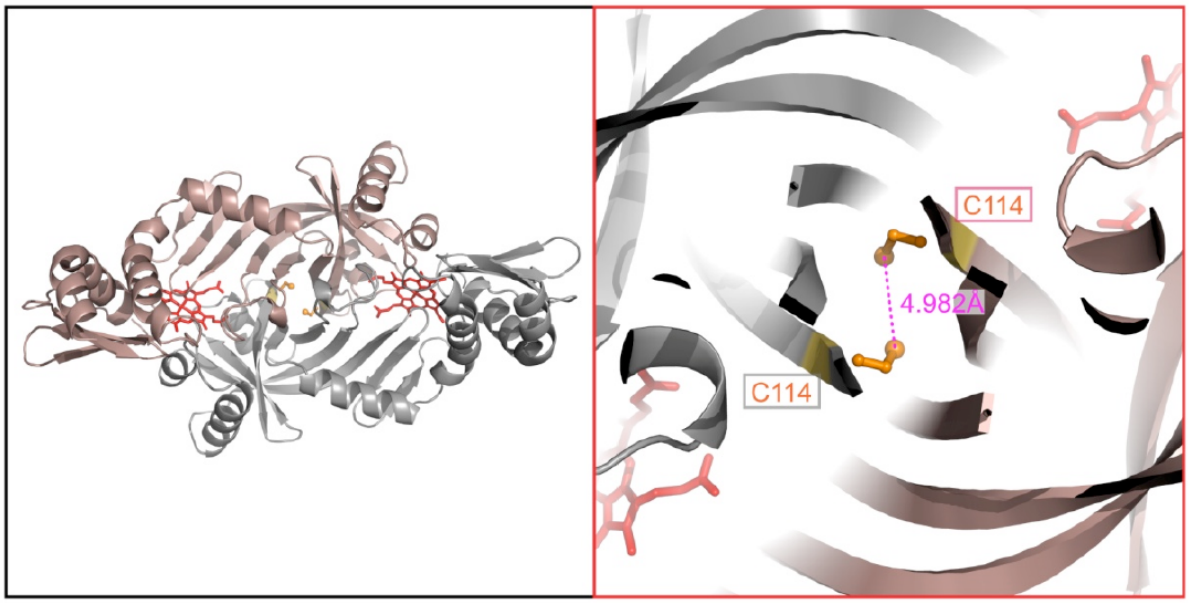

Fig. S5 Crystal structure of *H. pylori* heme iron utilization protein (HugZ, HP0318; PDB 3GAS). C114 is identified from the Bi-binding peptide (104)ATLHPNGHVV**C**(114). C114 residues from the two monomers are separated by 5.0 Å at the dimer interface.

HugZ contains a Bi-binding peptide (104)ATLHPNGHVV**C**(114) at the dimerization interface where the distance between the two intersubunit C144 (Sγ) is 5.0 Å.[14] HugZ utilizes heme as a source of iron by catalyzing the oxidative cleavage of porphyrin ring to generate free Fe^{2+}, CO and biliverdin in the presence of electron donor such as NADPH-cytochrome P450 reductase, which is crucial for the growth of *H. pylori* under iron-restricted condition.[15] Interference on dimer formation is presumptively influential to the function of HugZ, since the active site pockets, where the two hemes are accommodated, are located at the crevice of the packed monomers.[14]

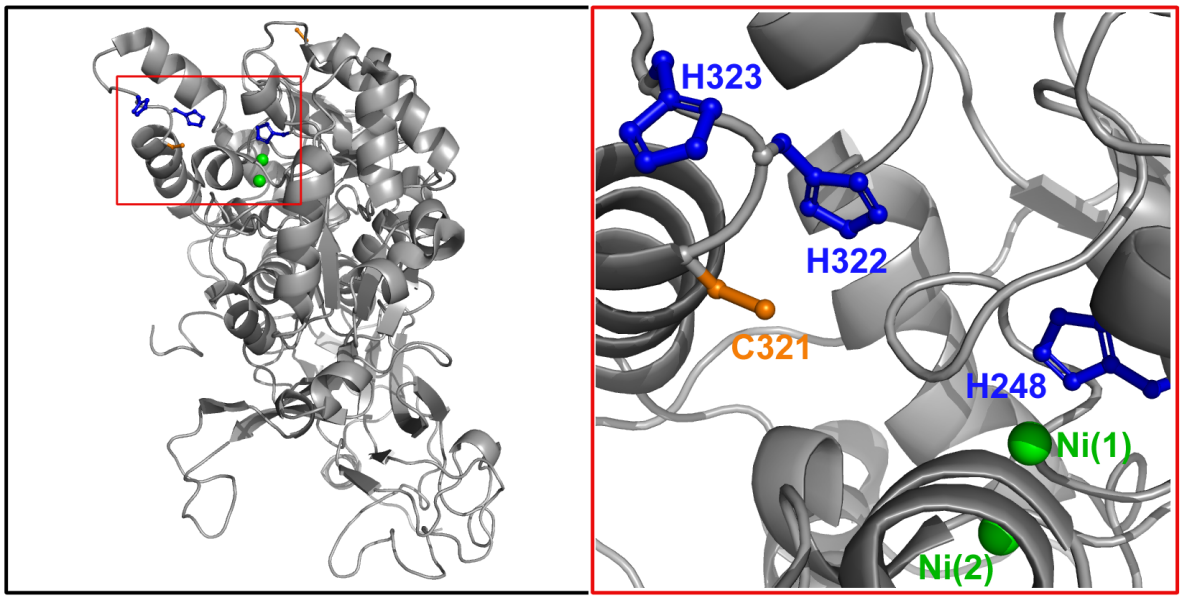

Fig. S6 Crystal structure of *H. pylori* urease subunit beta (UreB, HP0072, PDB 1E9Y). C321, H322 and H323 are identified from the Bi-binding peptide (307)TVNTEAEHMDMLMV**C**H**H**LDK(326). The flag region consisting of C321 and H322 are adjacent to the catalytic site with two Ni ions (shown in green).

Urease subunit beta in couple with UreA forms a dodecameric $(\alpha_3\beta_3)_4$ supramolecular assembly which is able to neutralize gastric acid from the surroundings by catalyzing the hydrolysis of urea to generate ammonia.[16] Catalytic site in UreB contains two nickel ions with different coordination number and geometry: Ni(1): distorted square pyramidal; Ni(2): pseudo octahedral.[17] In addition, crystal structure of UreB homolog from *B. pasteurii* revealed that the hydrolysis of urea may involve both of the nickel ions in which unsaturated Ni(1) orientates urea by coordinating its more nucleophilic carbonyl group, Ni(2) stabilizes one of the amide nitrogen of urea and the bi-Ni center activates a water molecule to OH^-.[18] Two Bi-binding motifs (241)YDVQVAI**H**TDTLNEAG**C**VEDTMAAIAGR(278) and (307)TVNTEAEHMDMLMV**C**H**H**LDK(326) with close proximity to the metallocenter are identified. The former contains Cys257 but the residue is situated far away from His248 that coordinates with Ni(1); the latter overlaps with the flap region which serves as an entrance gate to the active site of urease. In *K. aerogenes*, mutagenesis of Cys319 (Cys321 in *H. pylori*) by Ala retained nearly half of its enzymatic activity, whereas C319S, C319D and C319Y variants only exhibited 4.5%, 0.03% and no urease activity respectively, suggesting that Cys321 may not be crucial in the catalytic processes but its induction of conformational changes upon steric hindrance in the flag region can be influential.[19, 20] Moreover, C319A mutant of jack bean urease also shows a 19-fold reduced sensitivity towards Bi(EDTA), further supporting that Cys321 in *H. pylori* is a target site for Bi-based drugs which will lead to enzyme inactivation.[21]

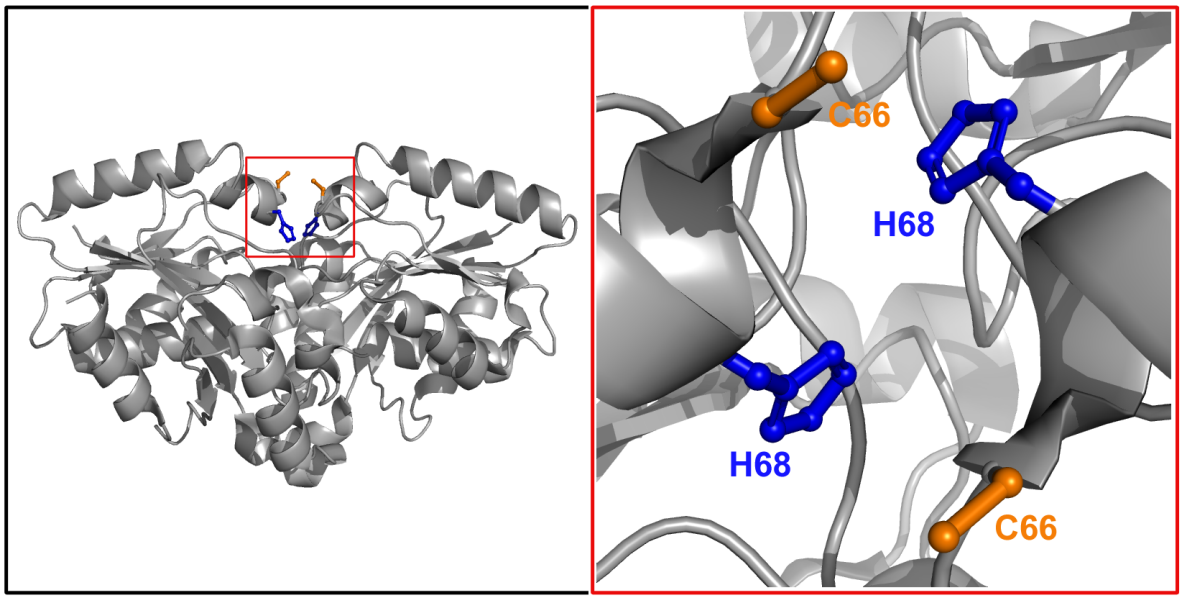

Fig. S7 Crystal structure of *H. pylori* urease accessory protein UreG (HP0068, PDB 4HI0). The dimeric form of *Hp*UreG is shown on the left.

UreG is one of the accessory proteins which are responsible for the activation of urease. In cooperation with UreF and UreH, the GTPase activity of UreG can utilize ammonia bicarbonate to promote the carbamylation of Lys in the active site of apo-urease, where the cofactor Ni^{2+} ions are concomitantly inserted by UreE.[22-24] The identified CXH motif located on the dimer interface of *Hp*UreG is involved in *Hp*UreE recognition, as revealed by a structure model of UreE-UreG complex.[25]

Structures of *H. pylori* protein homologues with Bi-binding motifs identified

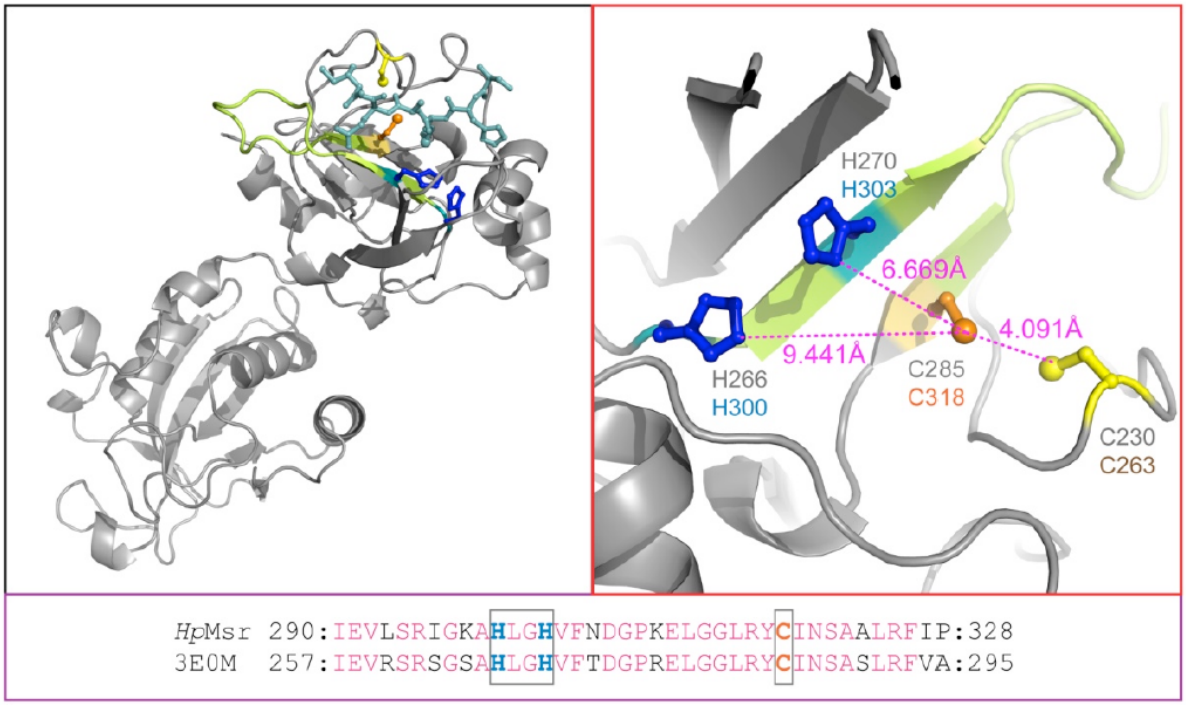

Fig. S8 Crystal structure of *H. pylori* peptide methionine sulfoxide reductase (MsrAB, HP0224) homologue from *S. pneumonia* (PDB 3E0M) and its active site. Active residue C285(Sγ) (aligned C318) is situated at the center of H266(Nε), H270(Nδ), and C230(Sγ) with distances of 9.4, 6.7 and 4.1 Å respectively.

MsrAB in *H. pylori* is a membrane-associated fusion protein composed of two commonly separated enzymes MsrA (N-terminus) and MsrB (C-terminus), which protects cells from oxidative damages by repairing proteins with oxidized methionine residues, i.e., the *S* and *R* isomers of Met sulfoxides.[26] The stereospecificity, substrate affinity and the reactivity of the two subunits are different but they share a similar three-step catalytic mechanism.[27] The MsrB subunit was identified to interact with bismuth through two consecutive motifs (299)A**H**LG**H**VFNDGPKELGGLR(316) and (317)Y**C**INSAALR(325), which form the substrate-binding pocket.[27] His300 and His303 can stabilize the Met-*R*-SO by H-bonds and deprotonated catalytic Cys318 attacks the sulfur atom of the sulfoxide (or in the form of sulfonium ion $^{+}$S-OH).[28, 29] The interdisulfide bond formed in the transition state is reduced by the recycling cysteine (Cys263) whilst an intradisulfide bond is formed between C318 and C263 and the reduced substrate is released with a loss of H_2O.[29] Similar to AhpC, the active site can finally be regenerated by the Trx1, instead of Trx2.[30] Interacting partners of MsrAB in *H. pylori* includes Trx1 and several peroxide-treated Met-rich proteins such as chaperone GroEL, ROS detoxification enzyme catalase (KatA) and site-specific recombinase (SSR).[30] In the absence of MsrAB, hydrophobic Met residues are converted into hydrophilic Met sulphoxide and the conversion may lead to structural changes that

decrease or ruin the biological activities of the proteins. For instance, the catalase activity of KatA in *H. pylori msr* mutant strain reduces by 50% under 10% O_2 compared to the wild-type strain, due to an inability of the Met residues of KatA to be repaired.[30]

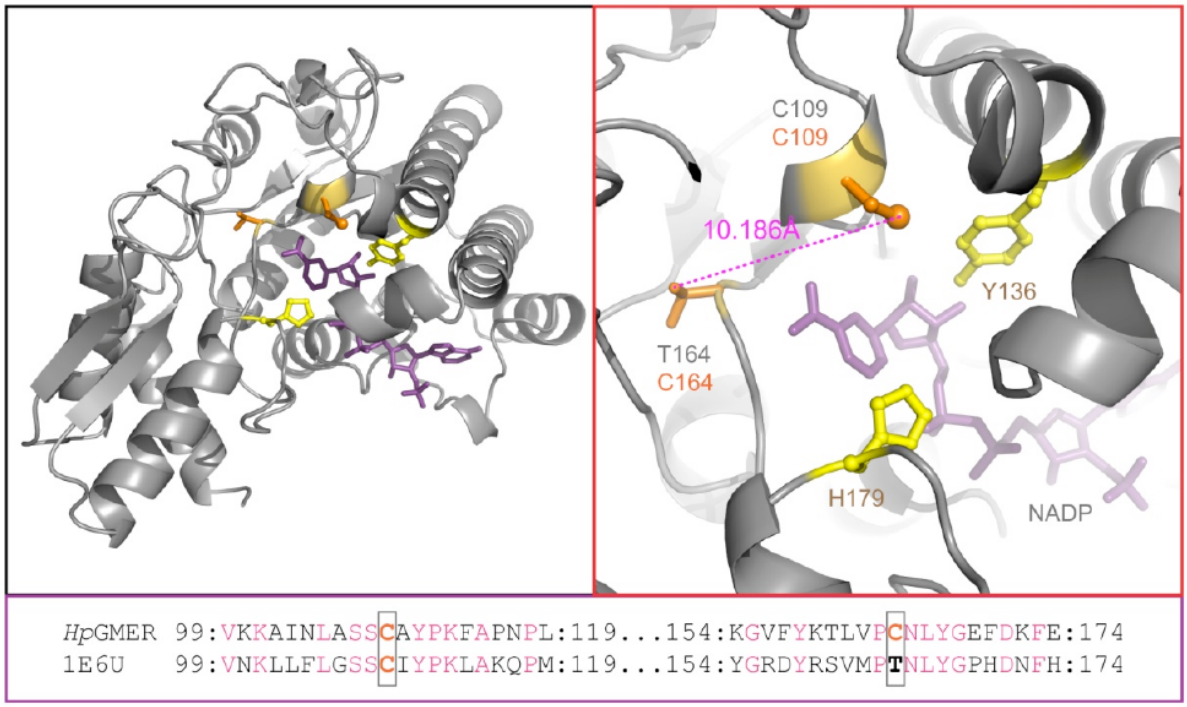

Fig. S9 Crystal structure of *H. pylori* GDP-L-fucose synthase (Fcl, HP0045) homologue from *E. coli* (PDB 1E6U) and its active site. C109 (orange), Y136 and H179 are active residues which are close to the NADP-binding domain (purple). The distance between C109(Sγ) and T164(Oγ) (aligned C164) is 10.2 Å.

GDP-L-fucose synthase (GMER) is involved in the last step of biosynthesis of GDP-L-fucose, the substrate of fucosyl transferase, for generating L-fucose-containing glycoconjugates (e.g., cell-surface antigens) which are essential for cell-cell interaction.[31] GMER acts as a bifunctional protein in the conversion of GDP-4-keto-6-deoxy-D-mannose to GDP-L-fucose – the configurations of C_3 and C_5 in the hexose ring of the substrate are changed from *R* to *S* and the carboxyl group on C_4 is subsequently reduced to hydroxyl group by NADPH.[32] The conserved Cys109 from the identified Bi-binding peptide (102)AINLASS**C**AYPK(123) functions as a base that facilitates the deprotonation of C_3-H and C_5-H in the first step of epimerization and the enzymatic activities of Cys109Ala and Cys109Ser GMER mutants are reduced by 7000- and 540-fold respectively.[31, 32]

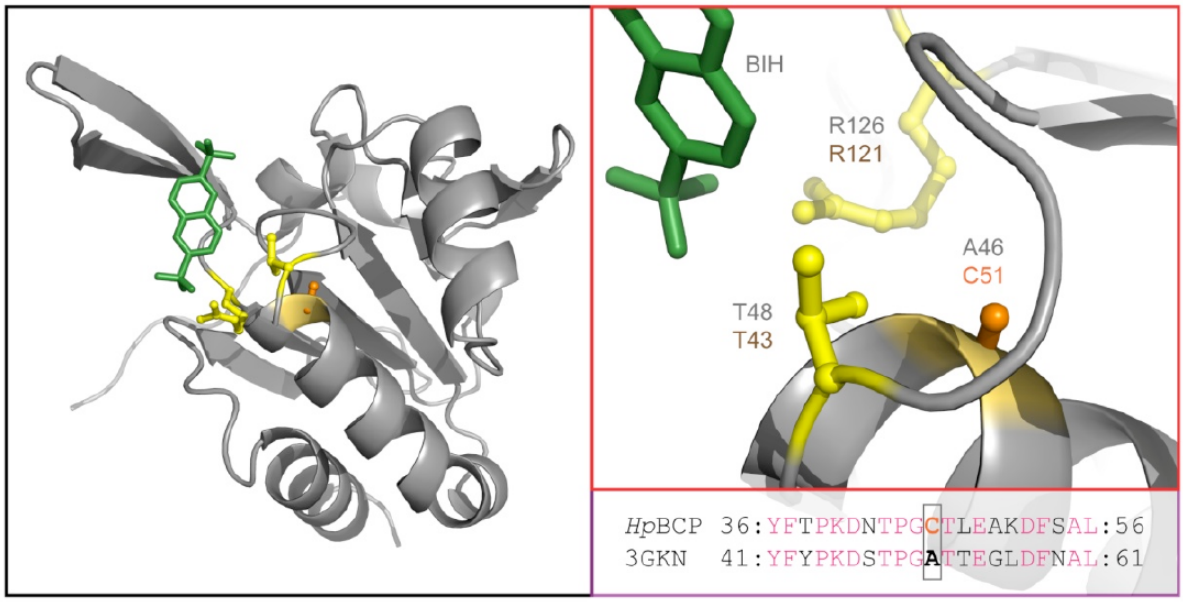

Fig. S10 Crystal structure of *H. pylori* putative peroxiredoxin bcp (Bcp, HP0136) homologue from *X. campestris* (PDB 3GKN) and its active site. A46 (aligned C51), T48 and R121 forms the catalytic triad whereas the trapped intermediate ligand BIH (naphthalene-2,6-disulfonic acid) is filled in green.

BCP and AhpC belong to two different sub-groups of peroxiredoxins (Prx) that prevent oxidative damages originated from peroxides such as hydrogen peroxide (H_2O_2) and organic peroxides by catalyzing the reduction reaction: $ROOH + 2e^- + 2H^+ \rightarrow ROH + H_2O$.[33] The thio peroxidase activity of BCP homolog in *E. coli* shows higher substrate preference towards linoleic acid hydroperoxide (R = unsaturated fatty acid) compared with H_2O_2 and *t*-butyl hydroperoxide (R = alkyl group).[34] Cys46 identified from the Bi-binding peptide (37)FYPKDNTPG**C**TLEAKDF(51) is a redox-active peroxidatic cysteine that reacts with peroxide by nucleophilic substitution (in the form of Cys-S^-) and forms cysteine sulfenic acid (Cys-SOH).

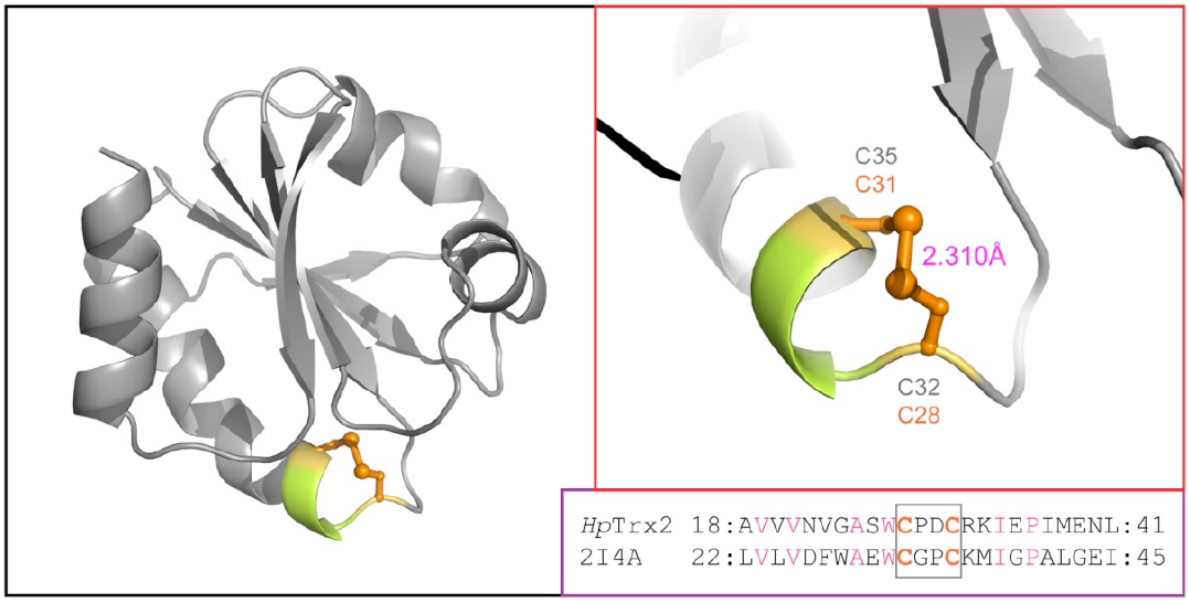

Fig. S11 Crystal structure of *H. pylori* thioredoxin (Trx2, HP1458) homologue from *A. aceti* thioredoxin (PDB 2I4A) and the aligned CXXC motif. *Hp*Trx2 C28 and C31 are identified from the Bi-binding peptide (14)IAHQAVVVNVGASW**C**PD**C**R(32). In the illustrated protein structure, the disulfide bond with 2.3 Å is formed between C32 and C35 (aligned C28 and C31).

Trx2, together with thioredoxin Trx1 (HP0824) and thioredoxin reductase (TrxB, HP0825) form the basis of the thioredoxin system in *H. pylori* which utilized the reducing potential obtained from NADPH to reduce disulfide bonds, including those from oxidized reductase such as Bcp, TagD and AhpC.[13, 35] *H. pylori* strain double mutant in the *trxA1* and *trxA2* genes was very sensitive to the oxidative and nitrosative stresses. Surprisingly, the thioredoxin system is not necessary for viability of the bacterium.[36] Bi-IMAC captured Trx2 *via* its bi-Cys motif (14)IAHQAVVVNVGASW**C**PD**C**R(32) in which Cys28 and Cys31 are catalytic residues responsible for thio-disulfide exchange.[13] *H. pylori* deprived of Trx2 is more sensitive to the exposure to 21% (v/v) O_2 (air), H_2O_2, cumene hydroperoxide and O_2^- despite that the effects are generally less obvious than that of Trx1. However, the specific downstream functions of Trx2 remain unclear.

Table S1 Lists of Bi-binding peptides in *H. pylori* 26695 identified by IMAC.

Locus tag	Gene name	Peptide sequence *	Measured m/z	Calculated m/z	Motif pattern
HP0010	GroEL, HspB	**E**I**E**LS**C**PVANMGAQLVK	1800.8434	1800.9114	C
HP0011	GroES, HspA	LV**E**RL**EEE**NK	1257.6516	1257.6565	E-rich
		EH**E**A**CC**H**D**H**K	1207.3879	1207.4498	HXXCCHXH
HP0019	HP0019	NN**E**LQLL**C**FR	1248.6239	1248.6285	C
HP0030	HP0030	I**D**LIS**C**RP**D**NFG**E**VWAK	1962.0818	1961.9669	C
HP0033	ClpA, HP0033	SVIKV**EE**PSK**E**A**C**Y	1580.7471	1580.7756	C
		IFAQA**E**AISLVSNAIKIQ**HC**GLSAK	2611.4492	2611.4155	HC
		KA**C**V**D**LTS**DY**	1113.4619	1113.5012	–
		K**H**P**HC**LLLL**DE**I**E**K	1688.8901	1686.9127	HXHC
HP0045	Fcl	TLVP**C**NL**Y**G**E**F**D**KF**EE**K	2031.0693	2030.9656	C
		YCEYVSA**E**K	1090.4534	1090.4641	C
		AINLASS**C**A**Y**PK	1236.6290	1236.6173	C
HP0068	UreG	IGV**C**GPVGSGK	972.5051	972.5062	C
		IIGVETGG**C**P**H**TAIR	1522.7665	1522.7926	CXH
HP0071	UreI	SAL**H**PTAPV**E**GA**ED**IAQVS**H**	2028.1979	2027.9912	$HX_{15}H$
HP0072	UreB, HpuB	**YD**VQVAI**H**T**D**TLN**E**AG**C**V**ED**T MAAIAGR	2962.2957	2962.3801	HX_8C
		EVTSKPANKVSLAQLFSIF	2078.1172	2078.1411	–
		TVNT**EA**E**H**M**D**MLMV**CHH**L**D**K	2353.0591	2353.0322	HX_6CHH
		LG**D**T**D**LIA**E**V**EHDY**TI**Y**G**EE**LK	2522.1958	2522.2065	H, D-rich, E-rich
		TM**H**TF**H**T**E**GAGGG**H**AP**D**IIK	2075.0693	2075.9846	$HXXHX_7H$
HP0073	UreA, HpuA	NVG**D**RPVQIGS**H**F**H**FF**E**VNR	2355.1548	2355.1509	HXH

HP0083	RpsI	VWLTPGKG**E**LSIN**E**QSLNQWLGG**HE**AIK	3103.4919	3103.6091	H
HP0084	RplM	**H**RPF**Y**TPNV**DC**G**D**FVVVINANK	2504.2061	2504.2271	HX_9C
		SKTLQ**E**ML**E**KAP**E**KL**YH**LAVR	2482.2842	2483.3569	H
		L**ED**K**EY**FT**H**SG**Y**FGSTK	2007.9124	2007.9214	H
HP0086	Mqo	VAIV**E**K**C**SKLAQISSSAK	1861.0173	1861.0343	C
		R**YE**SFK**E**IFVGL**EE**F**D**KQK	2391.1694	2391.2112	E-rich
		GL**D**LG**C**LPVAGSF	1247.6316	1247.6219	C
		DRTKRKL**E**LG**E**KKI**C**T**H**KGITF	2600.4788	2600.4585	CXH
		GANGI**D**R**HE**NIIG**H**G**Y**	1721.7909	1721.8234	HX_5H
		NMTPSPGATS**C**L	1176.4568	1177.5107	C
HP0088	RpoD	FGLL**DDESD**R	1165.5048	1165.5251	D-rich
HP0109	DnaK	IV**D**RNGA**C**AI**E**ISGK	1544.7620	1544.7981	C
		TNLN**E**N**D**AN**E**IQNAINALK**DC**V KN**D**NATK	3186.3801	3186.5212	C
HP0110	GrpE	**H**GI**E**GI**EC**L**EE**F**D**P**H**F**H**NAIMQVK	2792.2820	2792.3052	HX_3CX_6HXH
HP0124	InfC	KVK**H**AR**E**FI**E**SNK**H**VKF	2096.1941	2096.1643	HX_9H
		EAL**H**IAQNLGL**D**LVLISASAKPPV**C**K	2699.6099	2699.5044	$HX_{20}C$
		VQTMMQ**D**LANP**EK**E**P**KT**E**GR	2301.1118	2301.1094	–
HP0125	RpmI	**HVHH**TNA**H**SVMSLL**C**R	1840.9128	1840.8937	$HXHHX_3HX_6C$
		GSAFKS**H**ILTK	1187.6486	1187.6663	H
HP0126	RplT	AK**E**QL**E**R	872.4550	872.4716	–
HP0129	HP0129	**EE**GL**D**L**HDC**A**CE**GPF**HDHE**R	2307.9744	2307.9272	$HXCXCX_4HXH$
		GKKPS**HH**K**H**	1054.5311	1054.5785	HHXH, H-rich
HP0130	HP0130	QQQALQQ**E**F**E**K	1375.8878	1375.6732	–
HP0136	Bcp	NVILL**CDED**KKAANL**Y**	1820.9319	1820.9342	C

		IINTQGVL**E**K**C**F**Y**	1526.7339	1526.7803	C
		F**Y**PK**D**NTPG**C**TL**E**AK**D**F	1945.0101	1944.8928	C
HP0137	HP0137	NIIKV**E**F**ED**LV**EE**Y**K**H**FQVLNK	2733.3812	2733.4377	H, E-rich
		LLSLITLN**C**IILLKK**E**SIVR	2281.3760	2281.4170	C
HP0138	HP0138	GIQAQ**E**T**D**LG**E**LIIQLIN**EH**PV**H**I VVPAI**HK**	3423.8772	2423.8879	$HXXHX_6H$, H-rich
		LNAA**YEEE**P**E**KLNAIAR	1930.0457	1929.9796	E-rich
HP0139	HP0139	**HDY**L**E**LF**E**G**H**A**E**FNMVK**D**F**C**SR	2686.3442	2686.1943	HX_8HX_9C
HP0153	RecA	A**E**I**D**G**D**MG**D**Q**H**VGLQAR	1810.8323	1810.8268	H, D-rich
		TTLSL**H**IIA**EC**QK	1455.7925	1455.7755	HX_4C
HP0154	Eno	A**CE**NVNSVIK	1075.5417	1075.5332	C
HP0160	HcpD	GSG**CH**NVAVM**YY**TGK	1585.7191	1585.7018	CH
		YTGKGVPK**D**L**D**KAIS**YY**	1917.0349	1916.9884	–
		QNG**E**QALNL**Y**K	1276.6719	1276.6411	–
HP0175	HP0175	QRNPNF**D**F**D**KLK**E**K**E**K**E**ALI**D**QAIR	3044.5911	3044.6042	D-rich
HP0176	Fba	**ESCE**KAVKAGF	1167.5514	1167.5594	C
HP0177	Efp	TF**H**AG**D**K**CEE**PNLV**E**K	1815.8577	1815.8462	HX_4C
HP0182	LysS	TG**E**LSI**H**AL**E**F**H**ILSK	1793.9615	1793.9675	HX_4H
HP0197	MetK	TW**E**KTNKA**EE**IKAFF	1840.8962	1840.9359	–
		SSA**E**L**EKC**VKSVF	1425.7144	1425.7174	C
		NLVASGV**CD**K	1005.5557	1004.4961	C
		SGK**D**PSKV**D**RSAA**Y**	1479.7138	1479.7318	–
		II**ERD**QKAKVA**CE**TL	1715.8866	1715.9240	C
		GGS**CPH**GGGAF	945.3709	945.3763	CXH
		A**CKETE**TL	893.4190	893.4164	C
HP0199	HP0199	G**E**LMLI**CD**TSGK	1265.6934	1265.5996	C

HP0200	RpmF	**D**GTWKLP**HH**INKFTK**EY**	2113.0659	2113.0745	HH
HP0210	HtpG	GKF**E**IS**EC**VK	1138.5118	1138.5692	C
		TL**E**LNPN**H**AILQK	1488.7991	1489.8253	H
HP0211	HcpA	KKG**C**KSSVK**E**A**CD**ALK**E**L	1936.0615	1936.0122	CX_7C
		NAQGTAK**DE**KQAV**E**NF	1748.8125	1748.8329	–
		SKA**CE**LK**D**GRG**CY**	1428.6491	1428.6489	CX_7C
		EKA**CD**LK**D**SPG**C**INAG**Y**	1782.8108	1782.7916	CX_7C
HP0217	HP0217	GVIGFN**DCDD**GSK	1325.6344	1325.5558	C
HP0221	NifU, HP0221	GLR**DD**P**D**TPAVPGQK	1564.8085	1564.7845	D-rich
HP0222	HP0222	M**E**KT**E**NT**DE**TRLR	1621.7758	1621.7729	–
HP0224	MsrA, MsrAB	KA**DE**VIV**DDD**K	1245.6616	1245.6089	D-rich
		A**H**LG**H**VFN**D**GPK**E**LGGLR	1916.0371	1916.0017	HXXH
		YCINSAALR	1009.5146	1009.5015	C
HP0232	HP0232	GQSGFFGSSRPTSSPAVSSGTR	2156.0391	2156.0247	–
HP0247	RhpA	TV**C**V**Y**GGQSVK	1139.5548	1139.5645	C
HP0264	ClpB	IQL**E**AQF**E**N**E**K**E**VFK**E**ISR	2336.2122	2336.2012	E-rich
HP0296	RplU	ALV**E**LV**E**VLAVSK**E**GKLS**C**GKP FVNGAK	2884.6135	2884.6096	C
		EGKLS**C**GKPFVNGAKI**E**A**E**VIN**E**GR	2644.5730	2644.3643	C
HP0297	RpmA	M**H**PGNNVGMGK**DH**TL**Y**ALI**D**G VVKF**EH**K**D**R	3377.6629	3377.6760	$HX_{10}HX_{13}H$, H-rich, D-rich
HP0318	HP0318	ATL**H**PNG**H**VV**C**	1146.5303	1146.5604	HX_3HXXC
		FGQV**HH**A**E**NVAFK	1482.7402	1482.7368	HH
		NAI**E**L**C**QSV**E**KT**HD**LK	1940.0264	1940.0037	CX_6H
		FI**E**RGA**E**F**D**KAF**D**SFI**E**K	2148.0317	2148.0527	–
HP0331	MinD	G**E**PVIRT**DCE**SAKA**Y**QR	1921.9868	1921.9316	C

		IL**C**LPLIGIIP**E**D**HH**IISATNK	2409.2942	2409.3455	$CX_{10}HH$
		N**C**NLSQALIT**D**K	1318.6843	1318.6552	C
HP0349	PyrG	GISSSSIATLLQ**HC**N**Y**QVSILK	2361.2091	2361.2363	HC
		LQSPNPIIL**D**FIKSALSKS	2070.3621	2070.1726	–
		LG**EYEC**E**E**IMPNSLL**E**K	1866.8744	1866.8744	C
HP0390	Tpx	SVISM**D**LPFSQGQI**C**GA**E**GIK	2177.9614	2179.0654	C
		VTFK**EE**T**Y**QL**E**GK	1570.8160	1570.7878	–
HP0399	RpsA	IGQ**E**IT**C**VVVAI**E**KSNNK	1944.0939	1944.0350	C
		S**H**SSLKN**D**AN**H**IGKR	1661.0383	1662.8550	HX_8H
HP0404	HP0404	IIQG**E**IP**C**SK	1086.5854	1086.5743	C
HP0417	MetG	VL**C**P**DC**LR	917.4855	917.4463	CXXC
HP0470	PepF, HP0470	IPF**EE**QNLS**EEE**ILALL**H**NPKR	2618.3625	2618.3704	H, E-rich
HP0491	RpmB	KGPMIGN**H**VS**H**ANNK	1603.0129	1602.8049	HXXH
		SIKIQL**DD**GTTK	1318.7970	1317.7140	–
HP0514	RplI	AG**E**V**CE**VK**D**G**Y**GNNFLIANQK	2268.1138	2268.0845	C
HP0547	CagA, Cag26	I**D**RLNQIASGLGGVGQAAGFPLKR	2437.5355	2437.3555	–
HP0551	RpmE	GI**HPEY**IP**C**KVT**C**VTSGK	1928.9850	1928.9666	HX_5CX_3C
		I**D**ISSF**CH**PF**Y**TGS**D**K	1815.8109	1815.8138	CH
HP0554	HP0554	MK**E**F**D**KA**C**TF	1218.5018	1218.5413	C
HP0561	FabG	SRFGSVVNVASIIG**E**R	1689.9613	1689.9163	–
		T**ED**F**HH**VI**D**NNLTSAFIG**C**R	2288.0481	2288.0645	$HHX_{12}C$
HP0562	RpsU	KQT**D**RNLVVT**EC**R	1560.8016	1560.8042	C
		EG**D**AF**DE**A**Y**R	1171.4811	1171.4781	–
HP0564	HP0564	GTNQ**D**SSIN**CD**SSSR	1569.6284	1569.6326	C
HP0589	OorA	AAI**E**VG**C**R	817.4187	817.4116	C
HP0590	OorB	MSS**Y**VN**C**NTV**H**TT**H**GR	1805.8057	1805.7937	CX_3HXXH

HP0599	HylB, HP0599	VF**C**GLAKL**DH**VVFK	1574.8657	1574.8643	CX_6H
		EVNL**Y**QSLLNL**C**L**HE**GFVGIK	2389.2578	2388.2512	CXH
		NNL**Y**GMVFGLNSF**D**ITS**H**KN**C**R	2529.1685	2529.1892	HXXC
		NTS**C**VG**EYH**K	1136.5004	1136.4921	CX_4H
HP0618	Adk	SVI**E**V**E**VS**E**NTAK**E**R	1688.8551	1688.8582	E-rich
HP0631	HydA, HP0631	RG**H**F**D**AG**E**FV**EH**FG**DE**NAK	2161.0774	2160.9614	HX_8H
HP0639	QueC	V**C**LVGF**DY**AQK	1241.6011	1241.6115	C
HP0695	HyuA, HP0695	GQVVIPVRQ**EE**VKVAVK**E**LL**E** AGAK	2688.8452	2688.5537	E-rich
		V**YSE**SA**C**SP**E**LGFSVTGVIMR	2231.0312	2231.0603	C
		G**Y**SPS**D**FV**C**FS**Y**GGAGPV**H**T**Y**G **Y**T**E**GLGFK	3162.2598	3162.4070	CX_9H
		DSGL**D**TVSVT**DCH**	1347.5900	1347.5613	CH
		GM**E**VGLI**C**NK	1062.4757	1062.5201	C
		AIVI**C**LLQS**H**K	1223.6543	1223.7061	CX_4H
		SAFG**C**A**C**A**D**F**EY**R	1436.5536	1436.5667	CXC
HP0696	HP0696	A**HK****D**IGV**CSE**F	1204.5653	1204.5547	HX_5C
		EFK**E**R**D**PIG**YE**KMF	1787.8467	1787.8552	–
		FG**D**G**Y**MIT**C**R	1161.4769	1161.4948	C
HP0720	HP0720	GLNVSV**C**SG**D**TR	1206.5546	1206.5663	C
		KK**DEHC**SPK	1070.5275	1070.5178	HC
		TKGLS**DEE**IKAK	1317.7317	1317.7140	–
HP0721	HP0721	GNM**ED**M**DDD**FGLRS**C**K	1831.7058	1831.7175	C
		HGKK**HD**K**DHDD**K**DHDHHDED HS**D**K**H**	3063.3270	3063.2952	$HX_3HX_3HX_4HXHHX_3HX_3H$, H-rich, D-rich
HP0742	PrsA, Prs	**D**VILV**DD**MI**D**TAGTI**C**K	1820.8483	1820.8900	C
		VQK**CD**KITTL	1147.5991	1147.6271	C

HP0776	RpoZ	KLV**D**IAIR**E**IA**E**GKI**D**I**D**RI**DE**RN	2793.5608	2793.5349	D-rich
HP0777	PyrH	**E**I**CE**S**Y**I**Y**R	1174.5039	1174.5328	C
HP0779	AcnB	RP**H**AI**DE**VFIGS**C**MTNIG**H**FR	2399.1118	2399.1626	HX_9CX_5H
HP0795	Tig	**EC**VPSVGV**E**VPN**EE**K	1613.8110	1613.7607	C
HP0824	TrxA	I**C**KVNT**DE**Q**EE**LSAK	1705.8318	1705.8192	C
HP0853	YheS, HP0853	TIGN**E**AL**EC**E**C**NISK	1519.7584	1519.7189	C
HP0857	GmhA	QSV**H**LLI**E**TL**E**NQGK	1707.8937	1707.9155	H
		ILI**C**GNGGSAS**D**AQ**H**FAA**E**LTGR	2287.1072	2287.1016	$CX_{10}H$
HP0866	GreA	I**CA**E**L**KQLK	1044.5472	1044.6001	C
HP0891	VdlD	VGNTS**CE**VGIKVL	1317.7506	1317.6963	C
		S**ED**IKTR**E**IT**H**TNS**CY**	1895.9136	1895.8683	HX_3C
HP0893	HP0893	ALNLT**E**NG**Y**TI**EEE**R**E**ILAR	2333.1887	2333.1863	E-rich
		ARAK**D**TKNRL**C**F	1421.7551	1421.7561	C
		AL**E**KLK**E**IQAQKQR	1681.9878	1681.9839	–
HP0900	HypB	F**C**VV**E**G**D**LQTNR	1379.6725	1379.6504	C
		M**CA**D**A**VIISK	1049.5328	1049.5249	C
HP0902	HP0902	M**E**VV**H**FL**E**GV**C**F**EK**	1665.7753	1665.7894	HX_5C
HP0919	CarB	AQTA**C**FNPIK	1091.4875	1091.5433	C
HP0953	HP0953	VAQK**E**FGSV**C**ALR	1406.7495	1406.7340	C
HP0954	RdxA	GL**DSC**IIGGF**D**PL	1305.7056	1305.6274	C
		IA**C**LIALGK	900.5319	900.5466	C
HP0958	HP0958	KKK**EE**LV**E**KT**E**PKI**Y**	1861.0475	1861.0560	E-rich
HP0978	FtsA	KI**C**AIVA**E**FK	1120.6106	1120.6316	C
		SA**C**AKAGL**D**N**D**K**H**IL	1554.7827	1554.7825	CX_9H
HP0979	FtsZ	NASTT**EC**Y**R**EV**DD**VLVR	1969.0845	1968.9211	C
HP1021	HP1021	**H**FSST**C**T**D**T**E**LANK	1552.6707	1552.6827	HX_4C

HP1024	CbpA, HP1024	IPIGV**EE**G**E**KIR	1338.7665	1338.7507	–
		SRGPS**ED**L**DD**ILSSIFGK	1934.9886	1934.9585	D-rich
HP1027	Fur	VW**C**K**EC**Q**ESEC**	1342.6102	1342.4991	CX_2CX_4C
		SGT**H**LSP**EE**IT**H**SIRQK**D**KNTSISSV**Y**R	3169.6125	3169.6116	HX_7H
		ENFI**C**VL**E**TSK	1281.6714	1281.6274	C
		II**E**FA**D**P**E**I**E**NR	1444.6987	1444.7198	–
		L**E**TL**E**SIL**E**R	1201.6318	1201.6554	–
HP1043	HP1043	LSIQ**EY**E**K**EAIR**H**F	1761.8720	1761.9050	H
HP1045	AcsA	AL**E**NNA**C**PSV**E**K	1273.5818	1273.5973	C
HP1048	InfB	A**EC**A**E**LG**Y**NPV**D**WGG**EHE**FIPVSAK	2717.2646	2717.2432	$CX_{13}H$
		VGTIAG**C**VVS**D**GVIAR	1515.8049	1515.8079	C
HP1098	HcpC	AS**YD**KA**CD**LK**D**SPG**C**F	1718.7039	1718.7280	CX_7C
		KKG**C**KLGAKGA**CD**ILKQL	1873.0552	1873.0641	CX_7C
		SKA**CE**L**E**NGGG**C**F	1313.5909	1313.5380	CX_7C
HP1123	SlyD	**D**FSAT**H**VMV**DY**N**H**PLAGK	2001.0606	2000.9414	HX_6H
HP1125	Omp18, HP1125	GETKPK**C**AQKTR**EC**Y	1740.7786	1740.8287	CX_6C
HP1138	ParB	GLA**D**IFP**E**IN**E**V**YE**QGL**YE**R	2354.1096	2354.1431	E-rich
HP1147	RplS	IFPF**YSE**SLASV**E**VLR	1855.9099	1855.9720	–
		GNGV**D**KTF**C**VRK	1322.6970	1322.6765	C
		TRTQ**Y**F**E**GV**C**IAIR	1655.7837	1655.8453	C
		YIQQF**ED**AQLK**D**K	1624.8472	1624.8097	–
HP1151	RpsP	**D**GGWI**E**SIG**YY**NPLS**E**PK**D**IK	2380.1729	2380.1587	–
		MS**E**RV**E**KLSQKA	1420.6748	1420.7344	–
HP1152	Ffh	VLL**C**A**CD**LQR	1132.5508	1132.5732	CXC
HP1158	ProC	ALSVI**E**SFGN**C**VR	1393.8417	1393.7024	C
		ERP**E**MII**E**QI**C**TPK	1685.8236	1685.8480	C

HP1179	DeoB	A**C**FNNLA**D**SN**D**R	1338.6470	1338.5623	C
HP1189	Asd	TL**H**GF**C**VA**D**QLR	1358.6913	1358.6765	HXXC
HP1190	HisS	FVSL**HH**QTLGMPFKR	1796.9891	1796.9508	HH
		NM**H**SGMQLN**C**LSFLK	1721.7876	1721.8052	HX_6C
HP1195	FusA	IG**E**V**HD**GAATM**D**WM**E**Q**E**K**E**R	2332.0234	2331.0259	H, E-rich
		DTLTG**D**TL**C**D**E**K	1309.5948	1309.5708	C
HP1196	RpsG	MMV**D**RLAN**E**LM**D**AAS**D**K	1908.8856	1908.8744	D-rich
		AFNK**IEE**KSG**E**KGI**E**VF**E**KAL**E**R	2650.4304	2650.3967	E-rich
HP1197	RpsL	KKTKSPALV**EC**PQRRGV**C**TRV**Y**	2518.3127	2518.3625	CX_6C
		LTSKF**E**VIS**Y**IPG**E**G**H**NLQ**EH**SIVLVR	3064.6082	3064.6345	HX_4H
HP1198	RpoBC	FK**D**IGT**CE**K**C**GVAIT**H**SK	1935.9948	1935.9547	$CXXCX_5H$
		ITL**EY**AG**CE**FGK	1329.6953	1329.6274	C
HP1200	RplJ	ALLI**CDY**KGLSVR	1449.7665	1449.8014	C
HP1201	RplA	ASFP**EE**KIK**E**NML**E**LVK	2004.2743	2004.0602	E-rich
HP1203	NusG	FIG**E**NKKPTPLS**EADI**G**H**IL**E**K	2435.3042	2435.3059	H
HP1204	RpmG	KF**C**PR**E**NK**H**TL**H**K**E**IK	2007.1300	2007.0836	CX_5HXXH
		TNT**E**KL**E**LK	1074.5947	1074.5920	–
		CSDCEDIN**Y**STTK	1475.5460	1475.5723	CXXC
HP1205	Tuf, TufA	GMVL**C**KPGSITP**H**K	1466.8027	1466.7738	CX_7H
		HYA**H**V**DC**PG**H**A**DY**VK	1710.7862	1710.7573	HXXHXXCXXH
		GITIATS**H**I**EYE**T**E**NR	1832.9380	1832.8905	–
		DYDNI**D**NAP**EE**K**E**R	1706.7026	1706.7384	D-rich
		AL**EE**AKAGNVG**E**WG**E**K	1686.8461	1686.8213	E-rich
HP1223	HP1223	MV**YDD**T**CD**KKN	1330.5490	1330.5533	C
HP1227	Cy553, HP1227	SKIVNMMS**E**A**E**I**E**K**D**LM**D**FK	2357.0906	2357.1316	–
HP1236	HP1236	**E**GF**H**LFW**EE**R**C**N**D**LS**H**QILNK	2614.3621	2614.2388	HX_6CX_4H

		KLPVGS**DEYE**LVF**E**R	1779.8798	1779.9043	–
HP1237	CarA	GVLGV**C**GV**D**TR	1074.5741	1074.5492	C
HP1242	HP1242	**H**NQL**DDD**IKTA**E**QQNAS**D**A**E**V S**H**MKK	2951.4248	2951.3679	$HX_{21}H$, D-rich
		LKLK**DE**I**H**SMII**EY**R	1887.0477	1887.0288	H
HP1244	RpsR	T**E**AKISFI**DY**K**D**L**D**MLK**H**	2166.1165	2166.1030	H, D-rich
		YCK**Y**T**E**AK	1004.4599	1004.4637	C
HP1257	PyrE	ILV**CED**IITTGK	1303.7494	1303.7057	C
HP1246	RpsF	KA**E**PSMIV**E**L**E**RL**Y**	1676.8657	1676.8807	–
		ILKPTLV**EEE**IKSKI**E**F	2015.2754	2015.1554	E-rich
		EK**HE**KT**EH**T**HSHH**T**EEA**E**SVGS**HSE**	2878.2849	2878.2502	$HX_4HXHXHHX_9H$, H-rich, E-rich
HP1266	Nqo3, HP1266	VLN**D**F**YH**NPI**C**GAGR	1675.8099	1674.7937	HX_3C
HP1285	HP1285	N**C**IK**Y**TP**E**TW**D**KF**E**K	1900.9351	1900.9030	C
		DKKPAVIL**D**L**DE**TVLNTF**DY**AG**Y**LIK	2953.6316	2953.5688	D-rich
		K**EC**VSPITR	1031.5654	1031.5433	C
		SVK**YH**QQSA**E**IR	1443.6820	1444.7423	H
HP1292	RplQ	S**YI**E**K**LTTAAR	1251.7308	1251.6823	–
		NLAIALI**EH**NKI**E**TGI**Y**K	2039.3724	2039.1415	H
HP1293	RpoA	**C**FN**C**L**D**KIGIK	1250.6206	1250.6151	CXXC
HP1294	RpsD	TS**DY**GLQLK**E**K	1280.6588	1280.6613	–
HP1295	RpsK	STP**Y**AAQQAV**E**SALSK	1650.8274	1649.8260	–
		DITPLP**H**NG**C**RPPK	1544.7375	1544.7770	HXXC
HP1296	RpsM	V**HE**LS**ED**EVSSIAK	1541.7742	1541.7573	H
HP1297	RpmJ	MKVRPSVKKM**CD**N**C**KIIK	2134.2693	2134.1611	CXXC
		GVIRVI**C**ATPK	1155.6388	1155.6798	C
HP1301	RplO	GF**E**GGQQPLQR	1215.5895	1215.5996	–

HP1302	RpsE	M**EE**INR**EE**FQ**E**VVVNIGR	2190.0693	2190.0740	E-rich
HP1303	RplR	**Y**AQAI**DD**VKQSTIT**H**I**D**GRKM	2389.1802	2389.2061	H, D-rich
		A**EE**LKKAGI**E**RAV**Y**	1575.8410	1575.8621	–
HP1304	RplF	SRIGKRII**E**IPSSVQASV**E**GSKL	2453.3743	2453.3965	–
		KNSK**E**K**H**EL**E**T**H**	1478.7579	1478.7478	HX_4H
HP1305	RpsH	RQNVGG**E**VL**C**SIW	1459.6763	1459.7274	C
HP1306	RpsZ, RpsN	**C**GRP**H**SV**Y**R	1073.5156	1073.5189	CX_3H
HP1308	RplX	TSQVVV**E**G**C**K	1048.5062	1048.5223	C
HP1309	RplN	**Y**ASVGSVIVASVK	1278.7920	1278.7184	–
		LNVA**D**NSGAK**E**IM**C**IK	1704.8201	1704.8539	C
		F**DD**NAAVIL**D**AK	1290.6638	1290.6456	D-rich
HP1310	RpsQ	VG**D**FVSAI**EC**R	1194.5676	1194.5703	C
HP1311	RpmC	SIK**E**L**EE**LL**H**AKKA**E**LF**E**LR	2395.2583	2395.3474	H, E-rich
HP1312	RplP	TKIVT**C**E**S**EN**E**I**Y**	1527.7366	1527.7126	C
HP1313	RpsC	**H**PQA**D**AQLAA**E**NVATQL**E**K	2033.0432	2033.0178	H
		GIQF**E**KK**EE**AK**EE**R**E**PR	2100.9316	2102.0757	E-rich
HP1314	RplV	KPTS**H**VFV**E**VA**E**GK**E**MK	1914.0016	1914.9873	H
		SALIVS**C**R	847.4296	847.4586	C
HP1315	RpsS	NV**H**NGRVFIPV**Y**IT**E**N**H**VG**Y**	2327.1631	2327.1812	$HX_{13}H$
HP1316	RplB	TGTSG**H**PVSPWGTPAKG**Y**K	1927.1335	1926.9589	H
		YILS**EC**MASVGVVGN**ED**FINVS IGKAGR	2927.3159	2927.4521	C
		GSAMNPV**DH**P**H**GGG**E**GK	1645.6707	1645.7267	HXH
HP1317	RplW	GVLVVQTAQNVTK	1355.8047	1355.7772	–
		INSLKQ**E**GKVK	1242.7312	1242.7296	–
HP1318	RplD	AGSITSPVFVGGGVS**H**GATNNR	2084.0823	2084.0398	H

		LAL**E**Y**A**L**EE**K	1177.6232	1177.6230	–
HP1319	RplC	M**E**FLVQKIGMSR	1437.7391	1437.7472	–
		TI**D**ANSTPVTLLK	1371.7830	1371.7610	–
HP1325	FumC	LASGPR**C**GLG**E**L	1171.5675	1171.6019	C
		NI**HC**ASGI**E**PNR**E**KI**DY**	1958.0114	1957.9316	HC
HP1334	HP1334	**E**R**EE**L**HEH**LISKP**C**LALR	2172.1633	2172.1475	$HXHX_5C$
		IFF**YDC**KPLK	1272.6573	1272.6577	C
HP1335	MnmA, TrmU	GLAL**DH**ALKLG**CE**KIATG**HY**	2109.0796	2109.1040	HX_5CX_6H
		IKNAQKA**CE**F	1150.5356	1150.5804	C
HP1345	Pgk	**D**LASL**CD**VFVN**D**AFGTS**H**R	2066.1521	2065.9526	$CX_{11}H$
		ILP**C**F**E**VL**D**K	1175.5662	1175.6260	C
		IR**E**SLPTIQ**YC**I**D**NK	1791.8684	1791.9189	C
		FLRG**EEE**N**DE**NLAK	1662.8130	1662.7849	E-rich
HP1379	Lon	**H**ALKPS**E**V**E**IS**HEC**LK	1818.0001	1818.9298	$HX_{10}HXC$
HP1385	Fbp	SGSMVA**D**V**HH**VLVKKGGMF	1998.1230	1998.0179	HH
		LASQ**HE**A**H**IL**E**K**Y**	1536.7208	1537.7889	HXXH
HP1430	Rnj	T**D**NNQNN**E**N**HE**NSS**E**NSKA**DE**MR	2676.2700	2676.0703	H, E-rich
HP1441	PpiA	VIAGFVAQGG**C**P**Y**GTGTGGPG**H**R	2158.0557	2158.0378	$CX_{10}H$
HP1458	HP1458	IA**H**QAVVVNVGASW**C**P**DC**R	2038.2634	2037.9877	$HX_{11}CXXC$
		GKV**E**FFKVSF**DE**SQ**D**LK**E**SLGIR	2657.5583	2657.3701	–
HP1483	UbiE	GLRNVV**E**RQ**E**ALK**E**FF	1934.0929	1934.0374	–
		RLV**D**VA**C**GTG**D**ML	1348.7601	1348.6479	C
		A**CE**H**A**FLFL**E**NK	1420.6879	1420.6809	CXH
		AIKK**CEE**L**E**NKASF	1608.7720	1608.8181	C
HP1496	RplY	SKKRISV**EC**AP**EH**LP**DHY**	2108.0754	2108.0474	CX_3HX_3H

HP1554	RpsB	L**H**IPIVAPL**D**TN**C**D**P**D**L**V**DY**PIPGN**DD**AIR	3270.5605	3270.6230	$HX_{10}C$
		DLL**E**CGV**H**FG**H**QTR	1610.7328	1610.7623	CXXHXXH
HP1555	Tsf	TIAQVVA**DC**SK**E**WN**DD**LK	2034.0880	2033.9278	C
		ETLALIA**E**I**E**K**D**N**EE**AKR	2071.3164	2071.0798	E-rich
		DLT**D**AGMM**DC**K	1198.4769	1198.4668	C
HP1563	TsaA	**H**F**EEH**G**E**V**C**PAGW	1496.6093	1496.6143	HX_3HX_3C
		HAVIN**D**LPLGR	1203.6980	1203.6724	H
		AT**H**QGVA**EY**LK	1215.5597	1215.6248	H
HP1576	MetN	LIN**C**L**E**RPSSG**E**VLVNGVNLLK	2366.2827	2366.2991	C
HP1588	HP1588	VTIPA**C**IVVATLR	1354.9545	1354.8007	C
		GEGVL**Y**K**E**IL**CD**V**CD**KLK	2022.1926	2022.0343	CXXC
		SL**EE**M**DDEE**VK**E**M**CDE**LSIK	2369.9202	2371.0117	C

* Amino acids with Bi-binding abilities are highlighted in bold.

Table S2 Amino acid compositions in Bi-binding peptides

Amino acids	Frequency in proteins (%)[37]	Amino acid compositions (%)		Changes in occurrence (%) ‡
		Hp proteome †	Bi-binding peptides in *Hp*	
A (Ala)	7.7	6.8	6.9	**1.8**
C (Cys)	2.0	1.1	4.2	**273.8**
D (Asp)	5.2	4.9	5.7	**15.3**
E (Glu)	6.2	7.2	9.9	**38.4**
F (Phe)	4.0	5.3	3.6	-32.0
G (Gly)	7.4	5.7	7.2	**26.9**
H (His)	2.3	2.1	3.8	**78.9**
I (Ile)	5.3	7.2	6.9	-4.7
K (Lys)	5.9	9.0	7.7	-2.9
L (Leu)	8.5	11.2	7.7	-31.3
M (Met)	2.4	2.2	1.5	-30.8
N (Asn)	4.3	5.8	4.6	-21.4
P (Pro)	5.1	3.3	3.2	-3.0
Q (Gln)	4.1	3.7	2.8	-24.7
R (Arg)	5.1	3.5	3.4	-0.9
S (Ser)	6.9	6.7	6.1	-10.2
T (Thr)	5.9	4.3	4.3	-0.01
V (Val)	6.6	5.6	6.8	**21.2**
W (Trp)	1.4	0.7	0.4	-49.0
Y (Tyr)	3.2	3.7	2.4	-34.0

† The background *Hp* proteome was obtained by theoretical trypsin digestion of the whole *H. pylori* 26695 proteome performed in the website of *MS-Digest* (UCSF).

‡ The changes in amino acid occurrence were compared between the theoretical trypsin digested *H. pylori* proteome and the identified Bi-binding peptides. The amino acids that are more abundant in Bi-binding peptides are highlighted in bold.

Table S3 Enriched GO Biological Processes associated with putative Bi-binding proteins identified by IMAC.

GO ID	Bi-protein Ratio	Background Ratio	*p*-value	FDR	GO term name	Proteins associated *
GO:0006412	55/166	98/1555	3.41×10^{-31}	4.99×10^{-29}	translation	P0A0X4 | P55972 | P55973 | P55975 | P56002 | P56003 | P56004 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56126 | P56127 | P56455 | P64275 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637
GO:0044267	68/166	163/1555	5.93×10^{-29}	4.34×10^{-27}	cellular protein metabolic process	O25668 | O25982 | P0A0R3 | P0A0X4 | P42383 | P48285 | P55970 | P55972 | P55973 | P55975 | P55994 | P55995 | P56002 | P56003 | P56004 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56112 | P56116 | P56126 | P56127 | P56420 | P56455 | P56460 | P64275 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | P71404

GO:0019538	70/166	187/1555	2.14×10^{-26}	1.18×10^{-24}	protein metabolic process	O24875 | O25216 | O25668 | O25982 | P0A0R3 | P0A0X4 | P42383 | P48285 | P55970 | P55972 | P55973 | P55975 | P55994 | P55995 | P56002 | P56003 | P56004 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56112 | P56116 | P56126 | P56127 | P56420 | P56455 | P56460 | P64275 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | P71404
GO:0009059	67/166	179/1555	4.36×10^{-25}	2.13×10^{-23}	macromolecule biosynthetic process	O24886 | O25528 | O25668 | O25671 | O25684 | O25806 | P0A0X4 | P55972 | P55973 | P55975 | P55976 | P55993 | P55994 | P56001 | P56002 | P56003 | P56004 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56126 | P56127 | P56455 | P56460 | P60325 | P64275 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637

GO:0009058	85/166	384/1555	3.80×10^{-15}	1.19×10^{-13}	biosynthetic process	O24886 | O25009 | O25087 | O25116 | O25286 | O25356 | O25528 | O25577 | O25668 | O25671 | O25684 | O25686 | O25773 | O25801 | O25806 | O25835 | O25936 | O26017 | P0A0X4 | P55972 | P55973 | P55975 | P55976 | P55993 | P55994 | P56001 | P56002 | P56003 | P56004 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56104 | P56016 | P56126 | P56127 | P56162 | P56184 | P56195 | P56455 | P56460 | P60325 | P64275 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | Q09066

GO:0044237	114/166	636/1555	1.60×10^{-14}	4.70×10^{-13}	cellular metabolic process	O24886 | O24913 | O24950 | O25087 | O25098 | O25116 | O25286 | O25356 | O25524 | O25528 | O25560 | O25577 | O25629 | O25668 | O25671 | O25684 | O25686 | O25773 | O25801 | O25806 | O25825 | O25835 | O25883 | O25893 | O25982 | O26017 | O26096 | P0A0R3 | P0A0X4 | P14916 | P42383 | P42445 | P48285 | P55970 | P55972 | P55973 | P55975 | P55976 | P55993 | P55994 | P55995 | P56001 | P56002 | P56003 | P56004 | P56005 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56097 | P56104 | P56016 | P56019 | P56112 | P56116 | P56126 | P56127 | P56154 | P56162 | P56184 | P56185 | P56195 | P56418 | P56420 | P56455 | P56460 | P60325 | P64275 | P64382 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | P66928 | P69996 | P71404 | Q09066
GO:0006457	11/166	18/1555	4.90×10^{-7}	1.02×10^{-5}	protein folding	O25668 | O25982 | P0A0R3 | P21762 | P42383 | P55970 | P55994 | P56112 | P56116 | P56420 | P56003
GO:0045454	4/166	4/1555	5.63×10^{-4}	9.88×10^{-3}	cell redox homeostasis	O25151 | O25996 | P55979 | P66928
GO:0009166	12/166	43/1555	1.07×10^{-3}	1.67×10^{-2}	nucleotide catabolic process	O25098 | O25524 | O25560 | O26096 | P55972 | P55995 | P56002 | P56003 | P56005 | P56097 | P56195 | Q09066

GO:0044248	19/166	88/1555	1.11×10^{-3}	1.67×10^{-2}	cellular catabolic process	O25098 | O25524 | O25560 | O25686 | O25835 | O26096 | P14916 | P48285 | P55972 | P55995 | P56002 | P56003 | P56005 | P56097 | P56154 | P56184 | P56195 | P69996 | Q09066
GO:0006184	7/166	17/1555	1.36×10^{-3}	1.86×10^{-2}	GTP catabolic process	O25560 | P55972 | P56002 | P56003 | P56005 | P56097 | Q09066
GO:0006351	8/166	22/1555	1.40×10^{-3}	1.86×10^{-2}	transcription, DNA-templated	O25671 | O25684 | O25806 | P55976 | P55993 | P56001 | P60325 | P64275
GO:0009116	19/166	92/1555	1.84×10^{-3}	2.28×10^{-2}	nucleoside metabolic process	O25098 | O25116 | O25356 | O25524 | O25560 | O25577 | O25835 | O26096 | P55972 | P55995 | P56002 | P56003 | P56005 | P56097 | P56104 | P56016 | P56162 | P56460 | Q09066
GO:0006414	5/166	9/1555	1.91×10^{-3}	2.28×10^{-2}	translational elongation	P55975 | P56002 | P56003 | P56004 | P64275
GO:0006091	14/166	60/1555	2.38×10^{-3}	2.75×10^{-2}	generation of precursor metabolites and energy	O24913 | O25686 | O25825 | O25883 | O26017 | P0A0R3 | P48285 | P55970 | P56026 | P56109 | P56154 | P56418 | P56460 | P66928
GO:0055114	24/166	131/1555	2.46×10^{-3}	2.77×10^{-2}	oxidation-reduction process	O24913 | O24950 | O24951 | O25011 | O25087 | O25151 | O25286 | O25311 | O25312 | O25348 | O25608 | O25686 | O25773 | O25801 | O25825 | O25856 | O25883 | P21762 | P55970 | P55979 | P56184 | P56418 | P56460 | P66928
GO:0009164	11/166	43/1555	3.41×10^{-3}	3.56×10^{-2}	nucleoside catabolic process	O25098 | O25524 | O25560 | O26096 | P55972 | P55995 | P56002 | P56003 | P56005 | P56097 | Q09066

GO:0009117	23/166	128/1555	3.74×10^{-3}	3.73×10^{-2}	nucleotide metabolic process	O25098 | O25116 | O25524 | O25560 | O25577 | O25801 | O25835 | O26096 | P55970 | P55972 | P55995 | P56002 | P56003 | P56005 | P56097 | P56104 | P56106 | P56162 | P56184 | P56195 | P66269 | P71404 | Q09066

* Proteins are listed by Uniprot ID.

Table S4 Enriched GO Molecular Functions associated with putative Bi-binding proteins identified by IMAC.

GO ID	Bi-protein Ratio	Background Ratio	*p*-value	FDR	GO term name	Proteins associated
GO:0003735	45/166	54/1555	6.31×10^{-37}	2.77×10^{-34}	structural constituent of ribosome	P0A0X4 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637
GO:0019843	30/166	39/1555	1.92×10^{-22}	7.65×10^{-21}	rRNA binding	P0A0X4 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56026 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56039 | P56041 | P56043 | P56045 | P56046 | P56047 | P56049 | P56078 | P66119 | P66185 | P66449 | P66459 | P66572 | P66609 | P66621
GO:0003723	42/166	90/1555	2.07×10^{-19}	7.56×10^{-18}	RNA binding	O25893 | P0A0X4 | P55972 | P55973 | P55975 | P56002 | P56003 | P56004 | P56005 | P56008 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56026 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56039 | P56041 | P56043 | P56045 | P56046 | P56047 | P56049 | P56078 | P56127 | P56185 | P64275 | P66119 | P66185 | P66449 | P66459 | P66572 | P66609 | P66621

GO:0003676	54/166	218/1555	5.37×10^{-11}	1.18×10^{-9}	nucleic acid binding	O25029 | O25665 | O25671 | O25684 | O25758 | O25806 | O25893 | P0A0X4 | P42445 | P55972 | P55973 | P55975 | P55993 | P55995 | P56001 | P56002 | P56003 | P56004 | P56005 | P56008 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56026 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56039 | P56041 | P56043 | P56045 | P56046 | P56047 | P56049 | P56078 | P56126 | P56127 | P56185 | P60325 | P64275 | P66119 | P66185 | P66449 | P66459 | P66572 | P66609 | P66621
GO:0046914	21/166	81/1555	3.76×10^{-5}	7.17×10^{-4}	transition metal ion binding	O25009 | O25216 | O25356 | O25528 | O25560 | O25577 | O25668 | O25671 | O25748 | O25825 | P0A0R3 | P14916 | P55970 | P55994 | P56021 | P56058 | P56109 | P56195 | P66185 | P69996 | Q09066
GO:0051082	5/166	8/1555	5.29×10^{-4}	9.68×10^{-3}	unfolded protein binding	O25668 | P21762 | P42383 | P55994 | P56116
GO:0005515	17/166	67/1555	9.41×10^{-4}	1.59×10^{-2}	protein binding	O25116 | O25356 | O25528 | O25668 | O25671 | O25801 | P21762 | P42383 | P48285 | P55970 | P55994 | P56001 | P56116 | P56420 | P56460 | P71404 | Q09066
GO:0003746	5/166	8/1555	1.07×10^{-3}	1.67×10^{-2}	translation elongation factor activity	P55975 | P56002 | P56003 | P56004 | P64275
GO:0000049	7/166	17/1555	1.36×10^{-3}	1.86×10^{-2}	tRNA binding	O25893 | P0A0X4 | P56020 | P56029 | P56041 | P56127 | P66609
GO:0003924	7/166	17/1555	1.36×10^{-3}	1.86×10^{-2}	GTPase activity	O25560 | P55972 | P56002 | P56003 | P56005 | P56097 | Q09066

GO:0046983	4/166	5/1555	1.51×10^{-3}	1.95×10^{-2}	protein dimerization activity	O25801 | P55970 | P56001 | Q09066
GO:0042802	7/166	18/1555	1.92×10^{-3}	2.28×10^{-2}	identical protein binding	O25116 | O25671 | P48285 | P55970 | P56420 | P56460 | Q09066
GO:0001882	35/166	219/1555	2.95×10^{-3}	3.24×10^{-2}	nucleoside binding	O24875 | O25029 | O25098 | O25116 | O25356 | O25524 | O25577 | O25629 | O25668 | O25686 | O25835 | O25893 | O26096 | P0A0R3 | P42383 | P42445 | P55972 | P55994 | P55995 | P56002 | P56003 | P56005 | P56097 | P56104 | P56106 | P56116 | P56126 | P56127 | P56154 | P56184 | P56185 | P56455 | P56460 | P71404 | Q09066
GO:0000166	40/166	263/1555	3.61×10^{-3}	3.69×10^{-2}	nucleotide binding	O24875 | O25029 | O25087 | O25098 | O25116 | O25286 | O25356 | O25524 | O25577 | O25629 | O25668 | O25686 | O25801 | O25835 | O25893 | O26096 | P0A0R3 | P42383 | P42445 | P55970 | P55972 | P55994 | P55995 | P56002 | P56003 | P56005 | P56097 | P56104 | P56106 | P56116 | P56126 | P56127 | P56154 | P56184 | P56185 | P56455 | P56460 | P66119 | P71404 | Q09066
GO:0030554	31/166	192/1555	4.23×10^{-3}	4.13×10^{-2}	adenyl nucleotide binding	O24875 | O25029 | O25098 | O25116 | O25356 | O25524 | O25577 | O25629 | O25668 | O25686 | O25835 | O25893 | O26096 | P0A0R3 | P42383 | P42445 | P55970 | P55994 | P55995 | P56104 | P56106 | P56116 | P56126 | P56127 | P56154 | P56184 | P56185 | P56455 | P56460 | Q09066

Table S5 Enriched GO Cellular Components associated with putative Bi-binding proteins identified by IMAC.

GO ID	Bi-protein Ratio	Background Ratio	*p*-value	FDR	GO term name	Proteins associated
GO:0005840	45/166	56/1555	1.29×10^{-35}	2.83×10^{-33}	ribosome	P0A0X4 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637
GO:0044444	53/166	92/1555	9.48×10^{-31}	1.04×10^{-28}	cytoplasmic part	O25116 | O25883 | P0A0X4 | P48285 | P56005 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56116 | P56418 | P56420 | P56460 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637

GO:0005737	88/166	290/1555	1.22×10^{-26}	7.67×10^{-25}	cytoplasm	O25116 | O25528 | O25668 | O25671 | O25773 | O25801 | O25883 | O25893 | O25936 | P0A0R3 | P0A0X4 | P14916 | P21762 | P42383 | P42445 | P48285 | P55970 | P55972 | P55973 | P55975 | P55993 | P55994 | P55995 | P56002 | P56003 | P56004 | P56005 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56097 | P56104 | P56106 | P56116 | P56126 | P56127 | P56154 | P56184 | P56185 | P56195 | P56418 | P56420 | P56455 | P56460 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | P69996 | P71404 | Q09066

GO:0005622	90/166	338/1555	1.34×10^{-22}	5.89×10^{-21}	intracellular	O24864 | O25116 | O25348 | O25528 | O25668 | O25671 | O25773 | O25801 | O25883 | O25893 | O25936 | P0A0R3 | P0A0X4 | P14916 | P21762 | P42383 | P42445 | P48285 | P55970 | P55972 | P55973 | P55975 | P55993 | P55994 | P55995 | P56002 | P56003 | P56004 | P56005 | P56008 | P56009 | P56010 | P56011 | P56013 | P56018 | P56020 | P56021 | P56023 | P56026 | P56028 | P56029 | P56030 | P56031 | P56032 | P56034 | P56035 | P56036 | P56038 | P56039 | P56041 | P56042 | P56043 | P56044 | P56045 | P56046 | P56047 | P56049 | P56050 | P56052 | P56054 | P56058 | P56078 | P56097 | P56104 | P56106 | P56116 | P56126 | P56127 | P56154 | P56184 | P56185 | P56195 | P56418 | P56420 | P56455 | P56460 | P66119 | P66142 | P66185 | P66217 | P66269 | P66449 | P66459 | P66572 | P66609 | P66621 | P66637 | P69996 | P71404 | Q09066
GO:0015935	7/166	7/1555	1.01×10^{-6}	2.11×10^{-5}	small ribosomal subunit	P0A0X4 | P56009 | P56010 | P56011 | P56026 | P66572 | P66609
GO:0015934	5/166	6/1555	2.11×10^{-4}	4.23×10^{-3}	large ribosomal subunit	P56029 | P56030 | P56039 | P56047 | P56054

References

1. C. UniProt, *Nucleic Acids Res.*, 2013, **41**, D43-D47. DOI: 10.1093/nar/gks1068.
2. A. Conesa, S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon and M. Robles, *Bioinformatics*, 2005, **21**, 3674-3676.
3. M. Remm, C. E. V. Storm and E. L. L. Sonnhammer, *J. Mol. Biol.*, 2001, **314**, 1041-1052.
4. A. Franceschini, D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering and L. J. Jensen, *Nucleic Acids Res.*, 2013, **41**, D808-D815. DOI: 810.1093/nar/gks1094.
5. J.-C. R. Rain, L. Selig, H. D. Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schachter, Y. Chemama, A. Labigne and P. Legrain, *Nature*, 2001, **409**, 211-215.
6. R. Häuser, A. Ceol, S. V. Rajagopala, R. Mosca, G. Siszler, N. Wermke, P. Sikorski, F. Schwarz, M. Schick, S. Wuchty, P. Aloy and P. Uetz, *Mol. Cell. Proteomics*, 2014, **13**, 1318-1329.
7. P. Voland, D. L. Weeks, E. A. Marcus, C. Prinz, G. Sachs and D. Scott, *Am. J. Physiol. Gastrointest. Liver Physiol.*, 2003, **284**, G96-G106.
8. T. Cheng, H. Li, X. Yang, W. Xia and H. Sun, *Metallomics*, 2013, **5**, 804-807.
9. X. Yang, H. Li, T. Cheng, W. Xia, Y. T. Lai and H. Sun, *Metallomics*, 2014, **6**, 1731-1736.
10. C.-N. Tsang, R. Ge and H. Sun, in *Biological Chemistry of Arsenic, Antimony and Bismuth*, ed. H. Sun, John Wiley & Sons, Ltd., 2011, ch. 15, pp. 353-376.
11. M. H. Chuang, M. S. Wu, W. L. Lo, J. T. Lin, C. H. Wong and S. H. Chiou, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 2552-2557.
12. E. Papinutto, H. J. Windle, L. Cendron, R. Battistutta, D. Kelleher and G. Zanotti, *Biochim. Biophys. Acta*, 2005, **1753**, 240-246.
13. L. M. Baker, A. Raudonikiene, P. S. Hoffman and L. B. Poole, *J. Bacteriol.*, 2001, **183**, 1961-1973.
14. Y. Hu, F. Jiang, Y. Guo, X. Shen, Y. Zhang, R. Zhang, G. Guo, X. Mao, Q. Zou and D. C. Wang, *J. Biol. Chem.*, 2011, **286**, 1537-1544.
15. Y. Guo, G. Guo, X. Mao, W. Zhang, J. Xiao, W. Tong, T. Liu, B. Xiao, X. Liu, Y. Feng and Q. Zou, *BMC Microbiol.*, 2008, **8**, 226.
16. N.-C. Ha, S.-T. Oh, J. Y. Sung, K. A. Cha, M. H. Lee and B.-H. Oh, *Nat. Struct. Biol.*, 2001, **8**, 505-509.
17. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani, *J. Biol. Inorg. Chem.*, 2000, **5**, 110-118.
18. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani, *Structure*, 1999, **7**, 205-216.
19. P. R. Martin and R. P. Hausinger, *J. Biol. Chem.*, 1992, **267**, 20024-20027.
20. M. A. Pearson, L. O. Michel, R. P. Hausinger and P. A. Karplus, *Biochemistry*, 1997, **36**, 8164-8172.
21. L. Zhang, S. B. Mulrooney, A. F. Leung, Y. Zeng, B. B. Ko, R. P. Hausinger and H. Sun, *BioMetals*, 2006, **19**, 503-511.
22. A. Soriano and R. P. Hausinger, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 11140-11144.
23. F. Biagi, F. Musiani and S. Ciurli, *J. Biol. Inorg. Chem.*, 2013, **18**, 571-577.
24. X. Yang, H. Li, T. P. Lai and H. Sun, *J. Biol. Chem.*, 2015, DOI: 10.1074/jbc.M114.632364.
25. M. Bellucci, B. Zambelli, F. Musiani, P. Turano and S. Ciurli, *Biochem. J.*, 2009, **422**, 91-100.

26. P. Alamuri and R. J. Maier, *Mol. Microbiol.*, 2004, **53**, 1397-1406.
27. Y. K. Kim, Y. J. Shin, W. H. Lee, H. Y. Kim and K. Y. Hwang, *Mol. Microbiol.*, 2009, **72**, 699-709.
28. S. Boschi-Muller, A. Gand and G. Branlant, *Arch. Biochem. Biophys.*, 2008, **474**, 266-273.
29. N. Brot and H. Weissbach, *Biopolymers*, 2000, **55**, 288-296.
30. P. Alamuri and R. J. Maier, *J. Bacteriol.*, 2006, **188**, 5839-5850.
31. C. Rosano, A. Bisso, G. Izzo, M. Tonetti, L. Sturla, A. De Flora and M. Bolognesi, *J. Mol. Biol.*, 2000, **303**, 77-91.
32. S. T. B. Lau and M. E. Tanner, *J. Am. Chem. Soc.*, 2008, **130**, 17593-17602.
33. Z. A. Wood, E. Schroder, J. R. Harris and L. B. Poole, *Trends Biochem. Sci.*, 2003, **28**, 32-40.
34. W. Jeong, M.-K. Cha and I.-H. Kim, *J. Biol. Chem.*, 2000, **275**, 2924-2930.
35. G. Wang, P. Alamuri and R. J. Maier, *Mol. Microbiol.*, 2006, **61**, 847-860.
36. S. L. Comtois, *Microbiology*, 2003, **149**, 121-129.
37. D. Whitford, *Proteins: Structure and Function*, John Wiley & Sons, 2003.