

## Supporting Information

### Knowledge-based Characterization of Similarity Relationships in the Human Protein-Tyrosine Phosphatase Family for Rational Inhibitor Design

*Dušica Vidović, Stephan C. Schürer\**

#### Table of Contents:

1. Table S1. List of 102 PTP genes
2. Table S2. List of sequence to model template identities for 455 PTP structure models
3. Figure S1. Correlation between normalized and raw site similarities based on best models
4. Figure S2. Correlation between sequence similarity and mean and maximum site similarities of PTP gene combinations
5. Figure S3. Correlation between sequence identity and sequence similarity for PTP pairs
6. Figure S4. Distribution of sequence similarities, sequence identities and site similarities
7. Figure S5. Correlation between sequence and site similarities and their dependence on the average model template identity (for each PTP combination)
8. Figure S6. 3D PCA analysis of sequence and site similarities
9. Figure S7. Sequence similarity hierarchical clustering of classical PTPs with gene annotation
10. Figure S8. Site similarity hierarchical clustering of classical PTPs with gene annotation
11. Figure S9. 2D PCA analysis of sequence and site similarities for classical PTPs
12. Figure S10. 3D PCA analysis of sequence and site similarities for classical PTPs
13. Figure S11. Correlation between sequence and site similarities for classical PTPs
14. Figure S12. PTPN1 and PTPN3 site comparison
15. Figure S13. Correlation of calculated small molecule inhibitor activity-based PTP similarities vs. sequence and site similarities.
16. Figure S14. Hierarchical clustering of experimental  $pK_i$  values of PTPN1 inhibitors screened against a set of classical PTPs
17. Table S3. Proximity of PTPs to PTPN1 based on similarity by small molecule SAR, sequences, and sites.
18. Table S4. Sequence and site similarities of classical PTPs relative PTPN1
19. Figure S15. PTPN1 and PTPN2 catalytic site comparison

**Table S1.** List of collected PTP genes with their classification

Gene	ApprovedName	SubType	Type
PTPRA	protein tyrosine phosphatase, receptor type, A	classical receptor	Classical
PTPRB	protein tyrosine phosphatase, receptor type, B	classical receptor	Classical
PTPRC	protein tyrosine phosphatase, receptor type, C	classical receptor	Classical
PTPRD	protein tyrosine phosphatase, receptor type, D	classical receptor	Classical
PTPRE	protein tyrosine phosphatase, receptor type, E	classical receptor	Classical
PTPRF	protein tyrosine phosphatase, receptor type, F	classical receptor	Classical
PTPRG	protein tyrosine phosphatase, receptor type, G	classical receptor	Classical
PTPRH	protein tyrosine phosphatase, receptor type, H	classical receptor	Classical
PTPRJ	protein tyrosine phosphatase, receptor type, J	classical receptor	Classical
PTPRK	protein tyrosine phosphatase, receptor type, K	classical receptor	Classical
PTPRM	protein tyrosine phosphatase, receptor type, M	classical receptor	Classical
PTPRN	protein tyrosine phosphatase, receptor type, N	classical receptor	Classical
PTPRN2	protein tyrosine phosphatase, receptor type, N polypeptide 2	classical receptor	Classical
PTPRO	protein tyrosine phosphatase, receptor type, O	classical receptor	Classical
PTPRQ	protein tyrosine phosphatase, receptor type, Q	classical receptor	Classical
PTPRR	protein tyrosine phosphatase, receptor type, R	classical receptor	Classical
PTPRS	protein tyrosine phosphatase, receptor type, S	classical receptor	Classical
PTPRT	protein tyrosine phosphatase, receptor type, T	classical receptor	Classical
PTPRU	protein tyrosine phosphatase, receptor type, U	classical receptor	Classical
PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1	classical receptor	Classical
PTPN1	protein tyrosine phosphatase, non-receptor type 1	classical non-receptor	Classical
PTPN2	protein tyrosine phosphatase, non-receptor type 2	classical non-receptor	Classical
PTPN3	protein tyrosine phosphatase, non-receptor type 3	classical non-receptor	Classical
PTPN4	protein tyrosine phosphatase, non-receptor type 4	classical non-receptor	Classical
PTPN5	protein tyrosine phosphatase, non-receptor type 5	classical non-receptor	Classical
PTPN6	protein tyrosine phosphatase, non-receptor type 6	classical non-receptor	Classical
PTPN7	protein tyrosine phosphatase, non-receptor type 7	classical non-receptor	Classical
PTPN9	protein tyrosine phosphatase, non-receptor type 9	classical non-receptor	Classical
PTPN11	protein tyrosine phosphatase, non-receptor type 11	classical non-receptor	Classical
PTPN12	protein tyrosine phosphatase, non-receptor type 12	classical non-receptor	Classical
PTPN13	protein tyrosine phosphatase, non-receptor type 13	classical non-receptor	Classical
PTPN14	protein tyrosine phosphatase, non-receptor type 14	classical non-receptor	Classical
PTPN18	protein tyrosine phosphatase, non-receptor type 18	classical non-receptor	Classical
PTPN20A	protein tyrosine phosphatase, non-receptor type 20A	classical non-receptor	Classical
PTPN21	protein tyrosine phosphatase, non-receptor type 21	classical non-receptor	Classical
PTPN22	protein tyrosine phosphatase, non-receptor type 22	classical non-receptor	Classical
PTPN23	protein tyrosine phosphatase, non-receptor type 23	classical non-receptor	Classical
DUSP1	dual specificity phosphatase 1	MKPs	DUSP
DUSP2	dual specificity phosphatase 2	MKPs	DUSP
DUSP4	dual specificity phosphatase 4	MKPs	DUSP
DUSP5	dual specificity phosphatase 5	MKPs	DUSP
DUSP6	dual specificity phosphatase 6	MKPs	DUSP
DUSP7	dual specificity phosphatase 7	MKPs	DUSP
DUSP8	dual specificity phosphatase 8	MKPs	DUSP
DUSP9	dual specificity phosphatase 9	MKPs	DUSP
DUSP10	dual specificity phosphatase 10	MKPs	DUSP
DUSP16	dual specificity phosphatase 16	MKPs	DUSP
STYXL1	serine/threonine/tyrosine interacting-like 1	MKPs	DUSP
DUSP3	dual specificity phosphatase 3	Atypical DUSP	DUSP
DUSP11	dual specificity phosphatase 11	Atypical DUSP	DUSP
DUSP12	dual specificity phosphatase 12	Atypical DUSP	DUSP
DUSP13	dual specificity phosphatase 13	Atypical DUSP	DUSP
DUSP14	dual specificity phosphatase 14	Atypical DUSP	DUSP
DUSP15	dual specificity phosphatase 15	Atypical DUSP	DUSP
DUSP18	dual specificity phosphatase 18	Atypical DUSP	DUSP
DUSP19	dual specificity phosphatase 19	Atypical DUSP	DUSP

DUSP21	dual specificity phosphatase 21	Atypical DUSP	DUSP
DUSP22	dual specificity phosphatase 22	Atypical DUSP	DUSP
PTPMT1	protein tyrosine phosphatase, mitochondrial 1	Atypical DUSP	DUSP
DUSP26	dual specificity phosphatase 26	Atypical DUSP	DUSP
DUSP23	dual specificity phosphatase 23	Atypical DUSP	DUSP
DUSP28	dual specificity phosphatase 28	Atypical DUSP	DUSP
DUPD1	dual specificity phosphatase and pro isomerase domain containing 1	Atypical DUSP	DUSP
EPM2A	laforin	Atypical DUSP	DUSP
RNGTT	RNA guanylyltransferase and 5'-phosphatase	Atypical DUSP	DUSP
STYX	serine/threonine/tyrosine interacting protein	Atypical DUSP	DUSP
SSH1	slingshot homolog 1	Slingshots	DUSP
SSH2	slingshot homolog 2	Slingshots	DUSP
SSH3	slingshot homolog 3	Slingshots	DUSP
PTP4A1	protein tyrosine phosphatase type IVA	PRLs	DUSP
PTP4A2	protein tyrosine phosphatase type IVA	PRLs	DUSP
PTP4A3	protein tyrosine phosphatase type IVA	PRLs	DUSP
CDC14A	CDC14 cell division cycle 14 homolog A	CDC14s	DUSP
CDC14B	CDC14 cell division cycle 14 homolog B	CDC14s	DUSP
CDKN3	cyclin-dependent kinase inhibitor 3	CDC14s	DUSP
PTPDC1	protein tyrosine phosphatase domain containing 1	CDC14s	DUSP
PTEN	phosphatase and tensin homolog	PTENs	DUSP
TPTE2	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2	PTENs	DUSP
TPTE	transmembrane phosphatase with tensin homology	PTENs	DUSP
TNS1	tensin 1	PTENs	DUSP
TENC1	tensin like C1 domain containing phosphatase	PTENs	DUSP
TNS3	tensin 3	PTENs	DUSP
MTM1	myotubularin 1	Myotubularins	DUSP
MTMR1	myotubularin related protein 1	Myotubularins	DUSP
MTMR2	myotubularin related protein 2	Myotubularins	DUSP
MTMR3	myotubularin related protein 3	Myotubularins	DUSP
MTMR4	myotubularin related protein 4	Myotubularins	DUSP
SBF1	SET binding factor 1	Myotubularins	DUSP
MTMR6	myotubularin related protein 6	Myotubularins	DUSP
MTMR7	myotubularin related protein 7	Myotubularins	DUSP
MTMR8	myotubularin related protein 8	Myotubularins	DUSP
MTMR9	myotubularin related protein 9	Myotubularins	DUSP
MTMR10	myotubularin related protein 10	Myotubularins	DUSP
MTMR11	myotubularin related protein 11	Myotubularins	DUSP
MTMR12	myotubularin related protein 12	Myotubularins	DUSP
SBF2	SET binding factor 2	Myotubularins	DUSP
MTMR14	myotubularin related protein 14	Myotubularins	DUSP
MTMR15	myotubularin related protein 15	Myotubularins	DUSP
ACPI	acid phosphatase 1	LMW	LMW
CDC25A	cell division cycle 25 homolog A	cdc25	cdc25
CDC25B	cell division cycle 25 homolog B	cdc25	cdc25
CDC25C	cell division cycle 25 homolog C	cdc25	cdc25

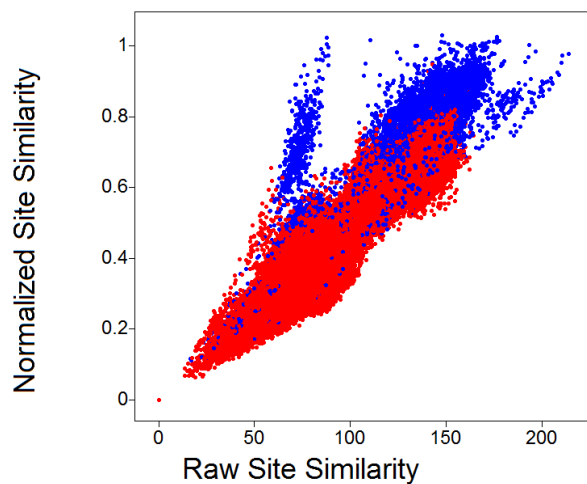
**Table S2.** Modeled sequence to template identities for 455 PTP domain structure models

Gene	Template	Identity	Gene	Template	Identity	Gene	Template	Identity
PTPRA	2ooqA	0.51	PTPN6	2h02A	0.30	DUSP21	2g6zA	0.32
PTPRA	2h02A	0.44	PTPN6	2f71A	0.45	DUSP21	1d5rA	0.14
PTPRA	2g59A	0.40	PTPN7	2ooqA	0.40	DUSP22	1wrmA	1.00
PTPRA	2f71A	0.36	PTPN7	2g59A	0.36	DUSP22	1yz4A	0.56
PTPRA_2	2ooqA	0.44	PTPN7	2h02A	0.44	DUSP22	1zzwA	0.37
PTPRA_2	2i1yA	0.38	PTPN7	2f71A	0.42	DUSP22	2g6zA	0.35
PTPRA_2	2g59A	0.38	PTPN9	2f71A	0.41	DUSP22	1d5rA	0.12
PTPRA_2	2h02A	0.37	PTPN9	2ooqA	0.36	PTPMT1	2imgA	0.22
PTPRA_2	2f71A	0.34	PTPN9	2h02A	0.39	PTPMT1	1wrmA	0.21
PTPRB	2h02A	1.00	PTPN9	2g59A	0.39	PTPMT1	1zzwA	0.21
PTPRB	2g59A	0.52	PTPN11	2ooqA	0.38	PTPMT1	2g6zA	0.20
PTPRB	2ooqA	0.49	PTPN11	2h02A	0.36	PTPMT1	1vhrA	0.19
PTPRB	2f71A	0.40	PTPN11	2f71A	0.42	PTPMT1	1d5rA	0.14
PTPRC	1ygrA	0.96	PTPN11	2i1yA	0.38	DUSP26	2e0tA	1.00
PTPRC	2ooqA	0.48	PTPN12	2ooqA	0.36	DUSP26	2pq5A	0.54
PTPRC	2h02A	0.45	PTPN12	2g59A	0.36	DUSP26	1vhrA	0.44
PTPRC	2g59A	0.42	PTPN12	2f71A	0.42	DUSP26	2g6zA	0.36
PTPRC	2f71A	0.34	PTPN12	2i1yA	0.38	DUSP26	1zzwA	0.33
PTPRC_2	1ygrA	0.97	PTPN12	2h02A	0.37	DUSP26	1d5rA	0.13
PTPRC_2	2ooqA	0.34	PTPN13	1wchA	0.36	DUSP23	2imgA	0.99
PTPRC_2	2g59A	0.31	PTPN13	2h02A	0.35	DUSP23	1oheA	0.28
PTPRC_2	2h02A	0.28	PTPN13	2ooqA	0.99	DUSP23	1d5rA	0.20
PTPRC_2	2f71A	0.26	PTPN13	2g59A	0.39	DUSP23	2e0tA	0.19
PTPRD	2ooqA	0.59	PTPN13	2f71A	0.38	DUSP23	1zzwA	0.18
PTPRD	2g59A	0.45	PTPN14	2f71A	0.34	DUSP23	2g6zA	0.18
PTPRD	2h02A	0.43	PTPN14	2h02A	0.33	DUSP28	2g6zA	0.35
PTPRD	2f71A	0.40	PTPN14	2ooqA	0.33	DUSP28	2nt2A	0.35
PTPRD_2	2ooqA	0.46	PTPN14	2g59A	0.31	DUSP28	1zzwA	0.30
PTPRD_2	2h02A	0.41	PTPN18	2h02A	0.36	DUSP28	2r0bA	0.29
PTPRD_2	2i1yA	0.40	PTPN18	2ooqA	0.36	DUSP28	1d5rA	0.14
PTPRD_2	2f71A	0.35	PTPN18	1wchA	0.36	DUPD1	2e0tA	0.55
PTPRE	2ooqA	0.53	PTPN18	2f71A	0.35	DUPD1	2pq5A	0.55
PTPRE	2h02A	0.44	PTPN18	2i1yA	0.35	DUPD1	1vhrA	0.47
PTPRE	2f71A	0.36	PTPN20A	1wchA	0.49	DUPD1	1zzwA	0.31
PTPRE_2	2ooqA	0.41	PTPN20A	2ooqA	0.39	DUPD1	2g6zA	0.30
PTPRE_2	2h02A	0.39	PTPN20A	2h02A	0.38	DUPD1	1d5rA	0.11
PTPRE_2	2i1yA	0.35	PTPN20A	2f71A	0.36	EPM2A	2g6zA	0.26
PTPRE_2	2f71A	0.29	PTPN21	2f71A	0.34	RNGTT	2imgA	0.22
PTPRF	2ooqA	0.59	PTPN21	2ooqA	0.34	RNGTT	1xm2A	0.21
PTPRF	2g59A	0.47	PTPN21	1ygrA	0.33	RNGTT	2pq5A	0.14
PTPRF	2h02A	0.46	PTPN21	1wchA	0.33	RNGTT	1d5rA	0.13
PTPRF	2f71A	0.41	PTPN21	2h02A	0.32	RNGTT	1zzwA	0.12
PTPRF_2	2ooqA	0.45	PTPN22	2ooqA	0.41	RNGTT	2g6zA	0.11
PTPRF_2	2h02A	0.41	PTPN22	2g59A	0.39	RNGTT	2esbA	0.11
PTPRF_2	2i1yA	0.38	PTPN22	2i1yA	0.38	STYX	2r0bA	1.00
PTPRF_2	2f71A	0.33	PTPN22	2f71A	0.37	STYX	1zzwA	0.35

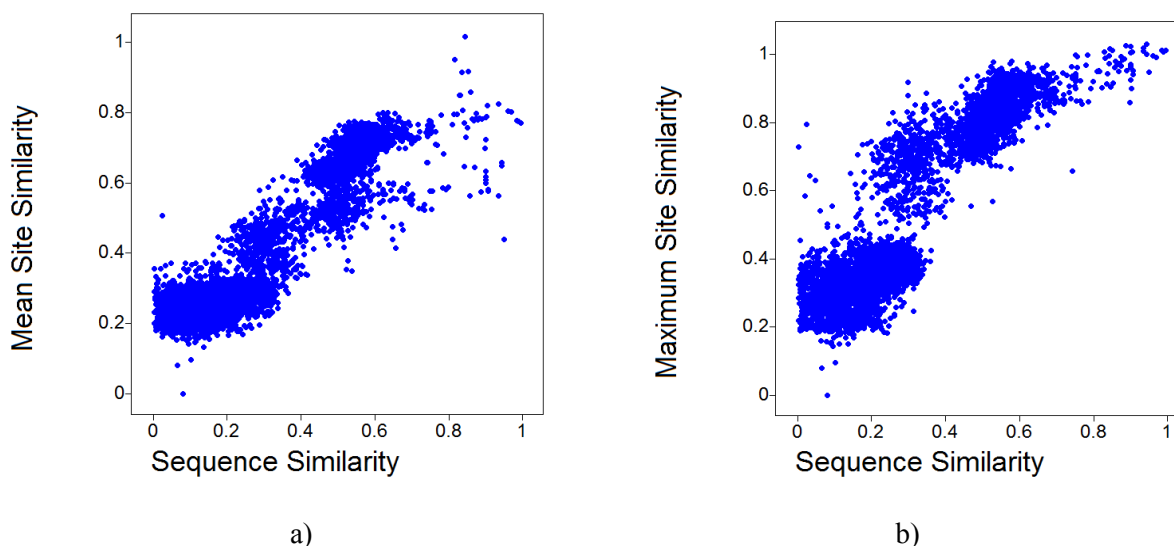
PTPRG	2ooqA	0.49	PTPN22	2h02A	0.36	STYX	2g6zA	0.34
PTPRG	2h02A	0.43	PTPN23	1wchA	0.34	STYX	2nt2A	0.33
PTPRG	2g59A	0.43	PTPN23	2f71A	0.33	STYX	1d5rA	0.13
PTPRG	2f71A	0.37	PTPN23	2h02A	0.30	SSH1	2nt2A	0.74
PTPRG 2	2ooqA	0.31	PTPN23	2ooqA	0.30	SSH1	2g6zA	0.40
PTPRG 2	2g59A	0.29	PTPN23	2g59A	0.30	SSH1	1zzwA	0.38
PTPRG 2	2f71A	0.28	DUSP1	2g6zA	0.65	SSH1	2r0bA	0.34
PTPRG 2	2h02A	0.26	DUSP1	1zzwA	0.43	SSH1	1d5rA	0.12
PTPRH	2h02A	0.49	DUSP1	2r0bA	0.36	SSH2	2nt2A	0.99
PTPRH	2g59A	0.49	DUSP1	2nt2A	0.35	SSH2	1zzwA	0.36
PTPRH	2ooqA	0.44	DUSP1	1d5rA	0.15	SSH2	2g6zA	0.35
PTPRH	2f71A	0.38	DUSP2	2g6zA	0.58	SSH2	2r0bA	0.33
PTPRJ	2h02A	0.56	DUSP2	1zzwA	0.44	SSH2	1d5rA	0.15
PTPRJ	2ooqA	0.50	DUSP2	2nt2A	0.39	SSH3	2nt2A	0.55
PTPRJ	2g59A	0.48	DUSP2	2r0bA	0.38	SSH3	2e0tA	0.36
PTPRJ	2f71A	0.36	DUSP2	1d5rA	0.15	SSH3	1zzwA	0.33
PTPRK	2ooqA	0.77	DUSP4	2g6zA	0.65	SSH3	2g6zA	0.33
PTPRK	2h02A	0.46	DUSP4	1zzwA	0.43	SSH3	1d5rA	0.15
PTPRK	2g59A	0.46	DUSP4	2r0bA	0.33	PTP4A1	1xm2A	0.98
PTPRK	2ilyA	0.39	DUSP4	2nt2A	0.32	PTP4A1	1d5rA	0.19
PTPRK	2f71A	0.35	DUSP4	1d5rA	0.15	PTP4A1	2g6zA	0.18
PTPRK 2	2ooqA	0.32	DUSP5	2g6zA	0.99	PTP4A1	2r0bA	0.17
PTPRK 2	2h02A	0.30	DUSP5	1zzwA	0.36	PTP4A1	1zzwA	0.15
PTPRK 2	2ilyA	0.30	DUSP5	1d5rA	0.15	PTP4A1	2nt2A	0.14
PTPRK 2	2f71A	0.27	DUSP6	1zzwA	0.49	PTP4A2	1xm2A	0.87
PTPRM	2ooqA	0.81	DUSP6	2g6zA	0.45	PTP4A2	1d5rA	0.19
PTPRM	2h02A	0.47	DUSP6	2r0bA	0.33	PTP4A2	2g6zA	0.19
PTPRM	2g59A	0.46	DUSP6	2nt2A	0.32	PTP4A2	2r0bA	0.18
PTPRM	2f71A	0.36	DUSP6	1d5rA	0.14	PTP4A2	1zzwA	0.14
PTPRM 2	2ilyA	0.33	DUSP7	1zzwA	0.48	PTP4A2	2nt2A	0.13
PTPRM 2	2ooqA	0.32	DUSP7	2g6zA	0.45	PTP4A3	1xm2A	0.79
PTPRM 2	2h02A	0.30	DUSP7	2nt2A	0.34	PTP4A3	1d5rA	0.18
PTPRM 2	2f71A	0.29	DUSP7	2r0bA	0.34	PTP4A3	2g6zA	0.17
PTPRN	2ilyA	0.99	DUSP7	1d5rA	0.15	PTP4A3	2r0bA	0.17
PTPRN	2ooqA	0.40	DUSP8	1zzwA	0.43	PTP4A3	1zzwA	0.15
PTPRN	2f71A	0.38	DUSP8	2g6zA	0.36	PTP4A3	2nt2A	0.14
PTPRN	2h02A	0.36	DUSP8	2r0bA	0.36	CDC14A	1oheA	0.70
PTPRN2	2ilyA	0.80	DUSP8	2e0tA	0.33	CDC14A	2imgA	0.23
PTPRN2	2ooqA	0.40	DUSP8	1d5rA	0.14	CDC14A	2r0bA	0.17
PTPRN2	2f71A	0.39	DUSP9	1zzwA	0.49	CDC14B	1oheA	0.99
PTPRN2	2h02A	0.36	DUSP9	2g6zA	0.45	CDC14B	2imgA	0.24
PTPRO	2g59A	1.00	DUSP9	2nt2A	0.37	CDC14B	1zzwA	0.17
PTPRO	2h02A	0.53	DUSP9	2r0bA	0.34	CDKN3	1fpzA	0.99
PTPRO	2ooqA	0.47	DUSP9	1d5rA	0.18	CDKN3	1xm2A	0.24
PTPRO	2f71A	0.38	DUSP10	1zzwA	1.00	CDKN3	1zzwA	0.16
PTPRQ	2g59A	0.51	DUSP10	2g6zA	0.36	CDKN3	2g59A	0.16
PTPRQ	2h02A	0.50	DUSP10	2nt2A	0.35	CDKN3	1d5rA	0.15
PTPRQ	2ooqA	0.42	DUSP10	2r0bA	0.34	PTPDC1	1oheA	0.25
PTPRQ	2f71A	0.34	DUSP10	1d5rA	0.15	PTPDC1	2imgA	0.25

PTPRR	2ooqA	0.44	DUSP16	1zzwA	0.42	PTPDC1	1d5rA	0.21
PTPRR	2g59A	0.42	DUSP16	2g6zA	0.40	PTPDC1	1zzwA	0.18
PTPRR	2h02A	0.40	DUSP16	2e0tA	0.35	PTPDC1	2e0tA	0.17
PTPRR	2f71A	0.34	DUSP16	2nt2A	0.34	PTPDC1	1vhrA	0.16
PTPRR	2i1yA	0.33	DUSP16	1d5rA	0.15	PTEN	1d5rA	1.00
PTPRS	2ooqA	0.59	STYXL1	2g6zA	0.29	PTEN	1zzwA	0.18
PTPRS	2g59A	0.46	STYXL1	1zzwA	0.27	PTEN	2imgA	0.15
PTPRS	2h02A	0.44	STYXL1	2r0bA	0.26	PTEN	2g6zA	0.13
PTPRS	2f71A	0.40	STYXL1	2nt2A	0.26	TPTE2	1d5rA	0.49
PTPRS_2	2ooqA	0.46	STYXL1	1wrmA	0.25	TPTE2	2imgA	0.20
PTPRS_2	2h02A	0.41	STYXL1	2e0tA	0.25	TPTE2	1vhrA	0.14
PTPRS_2	2i1yA	0.39	STYXL1	1d5rA	0.14	TPTE2	1zzwA	0.13
PTPRS_2	2f71A	0.34	DUSP3	1vhrA	1.00	TPTE2	2g6zA	0.12
PTPRT	2ooqA	1.00	DUSP3	2e0tA	0.44	TPTE	1d5rA	0.47
PTPRT	2h02A	0.49	DUSP3	2g6zA	0.33	TPTE	1xm2A	0.14
PTPRT	2g59A	0.46	DUSP3	1zzwA	0.32	TPTE	2g6zA	0.13
PTPRT	2f71A	0.37	DUSP3	1d5rA	0.13	TPTE	1zzwA	0.12
PTPRT_2	2h02A	0.31	DUSP11	2e0tA	0.17	TNS1	1wvhA	0.90
PTPRT_2	2ooqA	0.31	DUSP11	1d5rA	0.16	TENC1	1d5rA	0.32
PTPRT_2	2f71A	0.28	DUSP11	2imgA	0.15	TENC1	1xm2A	0.12
PTPRU	2ooqA	0.62	DUSP11	1zzwA	0.13	TENC1	1fpzA	0.11
PTPRU	2h02A	0.39	DUSP11	1vhrA	0.11	TENC1	1zzwA	0.08
PTPRU	2g59A	0.38	DUSP12	2g6zA	0.32	TENC1	2g6zA	0.08
PTPRU	2f71A	0.29	DUSP12	1wrmA	0.31	TNS3	1wvhA	0.75
PTPRU_2	2ooqA	0.32	DUSP12	2e0tA	0.31	MTM1	1zsqA	0.73
PTPRU_2	2i1yA	0.30	DUSP12	1zzwA	0.27	MTM1	1lw3A	0.73
PTPRU_2	2h02A	0.29	DUSP12	2r0bA	0.27	MTMR1	1zsqA	0.77
PTPRU_2	2f71A	0.27	DUSP12	2nt2A	0.26	MTMR1	1lw3A	0.77
PTPRZ1	2ooqA	0.50	DUSP12	1d5rA	0.13	MTMR2	1zsqA	1.00
PTPRZ1	2h02A	0.45	DUSP13	2pq5A	1.00	MTMR2	1lw3A	1.00
PTPRZ1	2i1yA	0.38	DUSP13	2e0tA	0.54	MTMR3	1zsqA	0.43
PTPRZ1	2f71A	0.36	DUSP13	1vhrA	0.44	MTMR3	1lw3A	0.43
PTPRZ1_2	2ooqA	0.31	DUSP13	1zzwA	0.31	MTMR4	1zsqA	0.41
PTPRZ1_2	2g59A	0.29	DUSP13	2g6zA	0.30	MTMR4	1lw3A	0.41
PTPRZ1_2	2f71A	0.27	DUSP13	1d5rA	0.12	SBF1	1lw3A	0.28
PTPRZ1_2	2h02A	0.26	DUSP14	2esbA	0.51	MTMR6	1zsqA	0.44
PTPN1	2f71A	1.00	DUSP14	2e0tA	0.39	MTMR6	1lw3A	0.44
PTPN1	2h02A	0.39	DUSP14	2g6zA	0.36	MTMR7	1zsqA	0.45
PTPN1	2ooqA	0.37	DUSP14	1zzwA	0.34	MTMR7	1lw3A	0.45
PTPN1	2g59A	0.36	DUSP14	1d5rA	0.13	MTMR8	1lw3A	0.46
PTPN2	2f71A	0.72	DUSP15	1yz4A	0.97	MTMR8	1zsqA	0.45
PTPN2	2h02A	0.38	DUSP15	1wrmA	0.57	MTMR9	1zsqA	0.41
PTPN2	2ooqA	0.37	DUSP15	1zzwA	0.36	MTMR9	1lw3A	0.41
PTPN2	2g59A	0.37	DUSP15	2g6zA	0.36	MTMR10	1zsqA	0.27
PTPN3	2f71A	0.37	DUSP15	1d5rA	0.13	MTMR10	1lw3A	0.27
PTPN3	2ooqA	0.37	DUSP18	2esbA	1.00	MTMR11	1zsqA	0.25
PTPN3	2g59A	0.37	DUSP18	2e0tA	0.37	MTMR11	1lw3A	0.25
PTPN3	2h02A	0.35	DUSP18	1zzwA	0.34	MTMR12	1lw3A	0.28
PTPN4	2ooqA	0.41	DUSP18	2g6zA	0.33	MTMR12	1zsqA	0.27

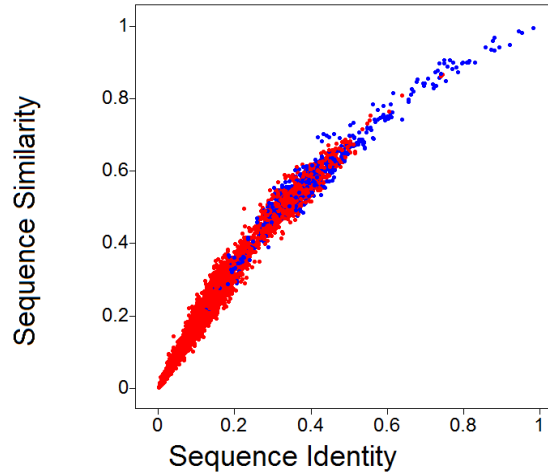
PTPN4	2f71A	0.39	DUSP18	1d5rA	0.15	SBF2	1lw3A	0.28
PTPN4	2g59A	0.39	DUSP19	1zzwA	0.40	MTMR14	1vhrA	0.21
PTPN4	2h02A	0.37	DUSP19	2r0bA	0.37	MTMR14	1zzwA	0.16
PTPN5	2ooqA	0.44	DUSP19	2g6zA	0.36	ACP1	5pntA	1.00
PTPN5	2g59A	0.39	DUSP19	1d5rA	0.13	CDC25A	1c25A	1.00
PTPN5	2h02A	0.36	DUSP21	2esbA	0.75	CDC25B	1c25A	0.69
PTPN5	2f71A	0.30	DUSP21	2e0tA	0.35	CDC25C	1c25A	0.66
PTPN6	2ooqA	0.45	DUSP21	1zzwA	0.32			



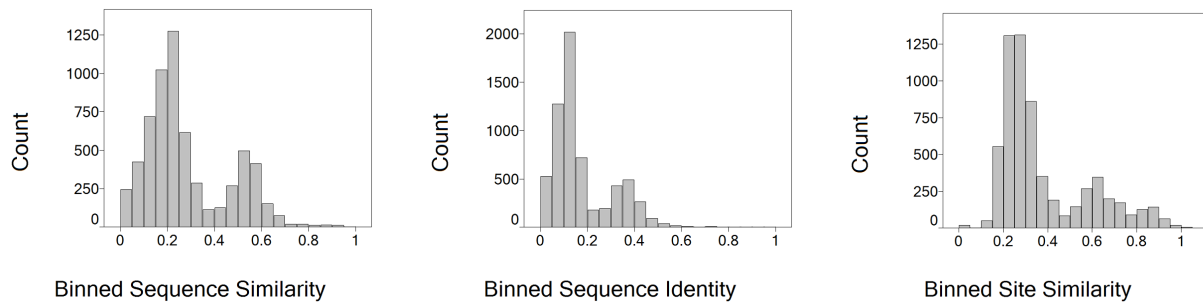
**Figure S1.** Correlation between normalized and raw site similarity values for all PTP combinations (455 models). Blue sites are modeled by same template and red sites are modeled by different templates.



**Figure S2.** Sequence similarity versus: a) mean site similarity between gene combinations and b) maximum site similarity between gene combinations (each gene is presented by ensemble of binding sites)

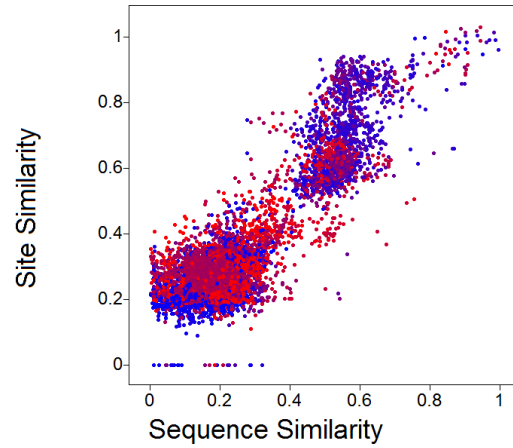


**Figure S3.** Sequence similarity versus sequence identity for PTP combinations (blue data represent site pairs where the corresponding models were generated with the same template while for red data points the templates were different).

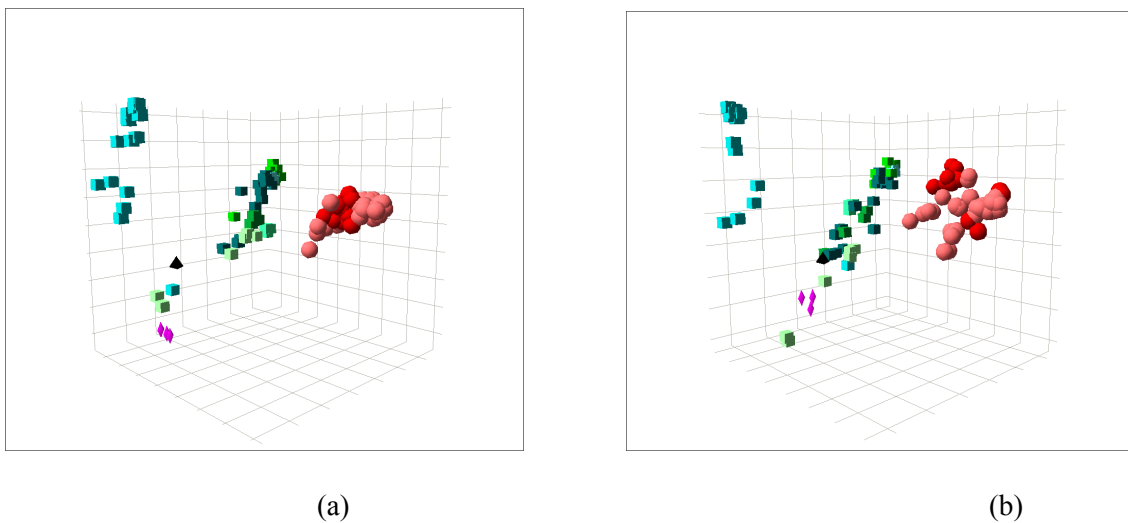


**Figure S4.** Histograms of sequence similarity, sequence identity, and site similarity pairs, respectively. Relative to the modes of the respective populations, the site similarity distribution contains a sub-population of highly similar pairs, which correspond to average lower sequence similarities (or identities).

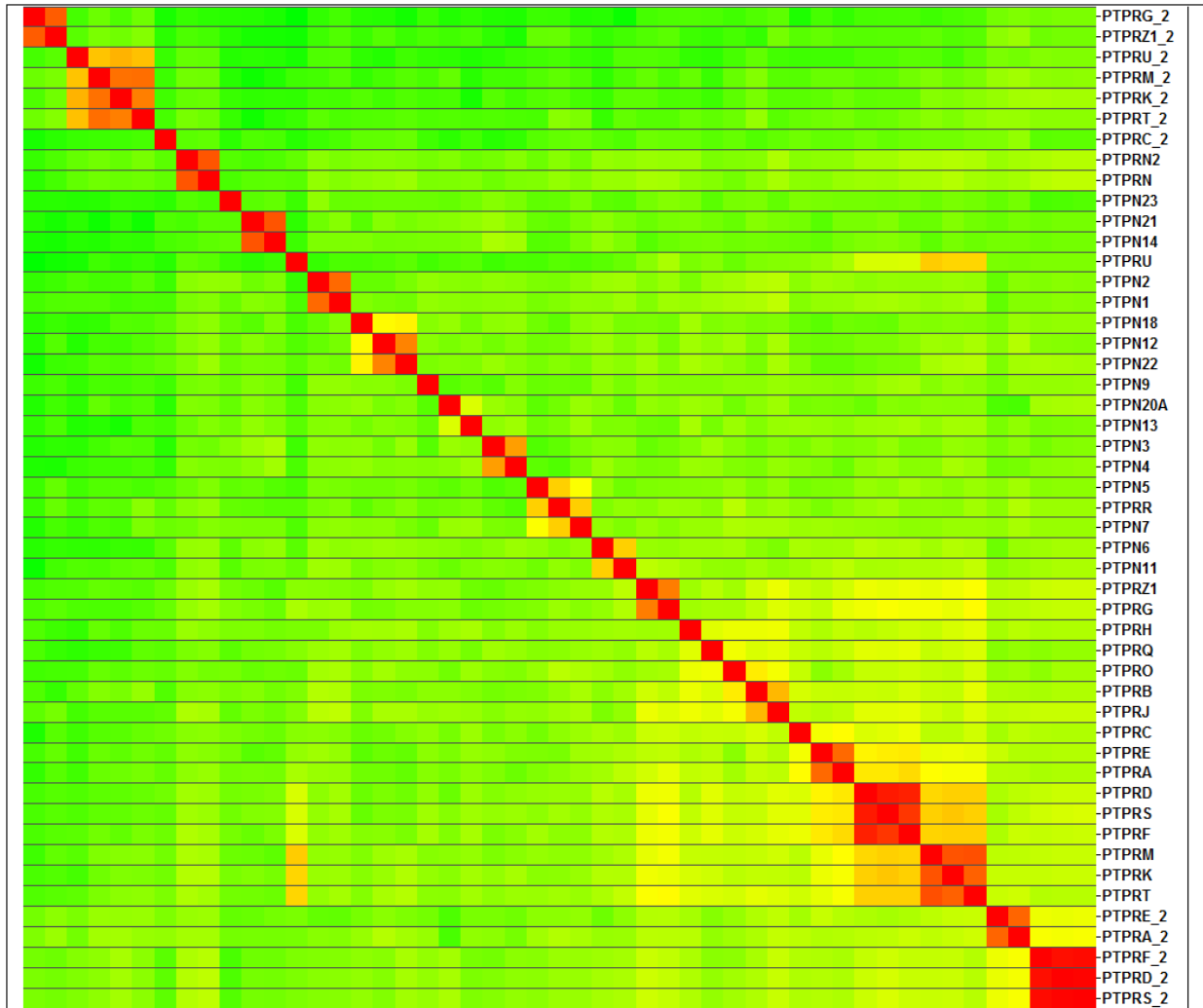




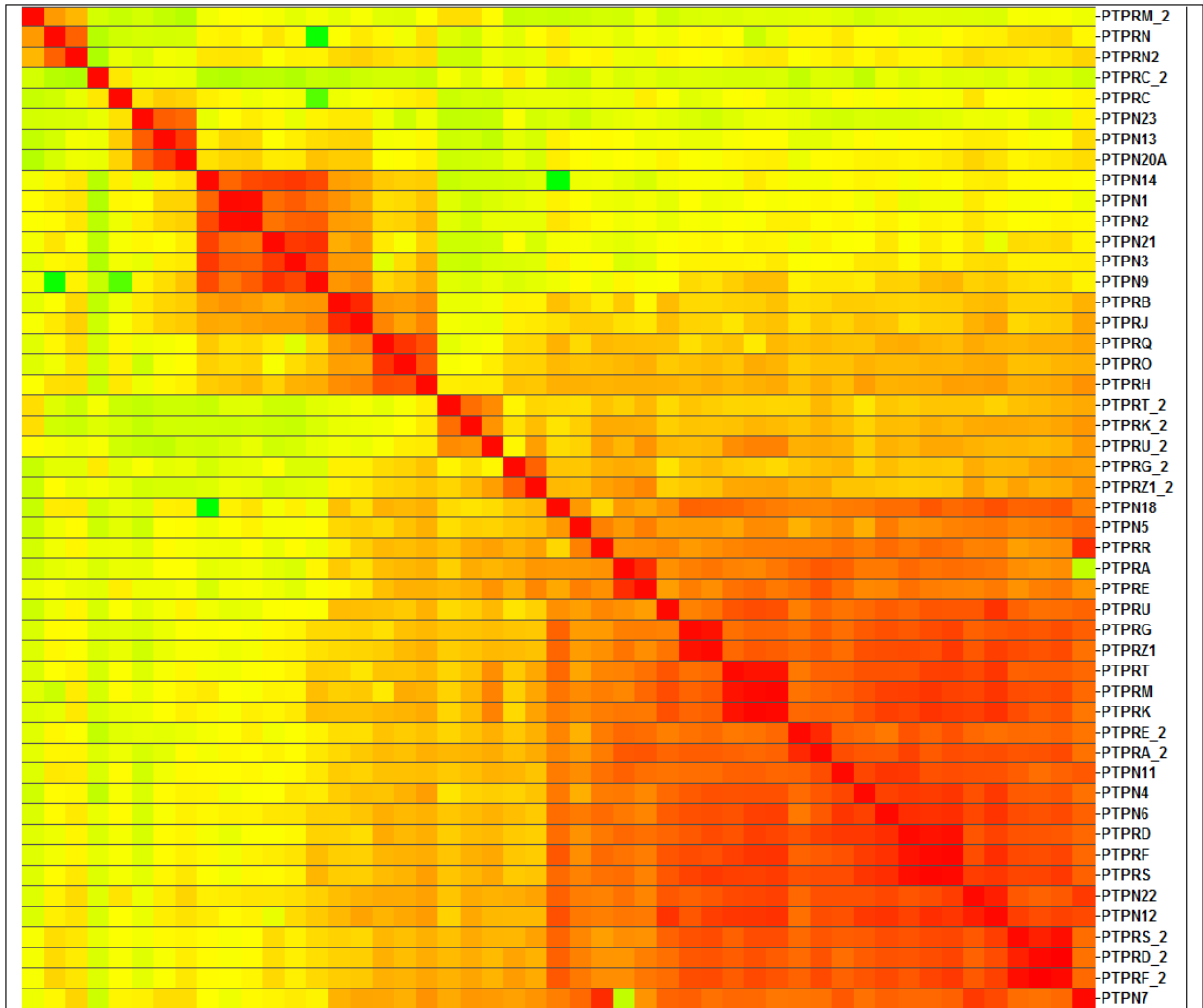
**Figure S5.** Correlation between (best model) site similarity and sequence similarity for all human PTP combinations and their dependence on average PDB template identity (higher average identity of the PTP domain to the PDB template is represented by increasing intensity of blue, lower average template identity as red).



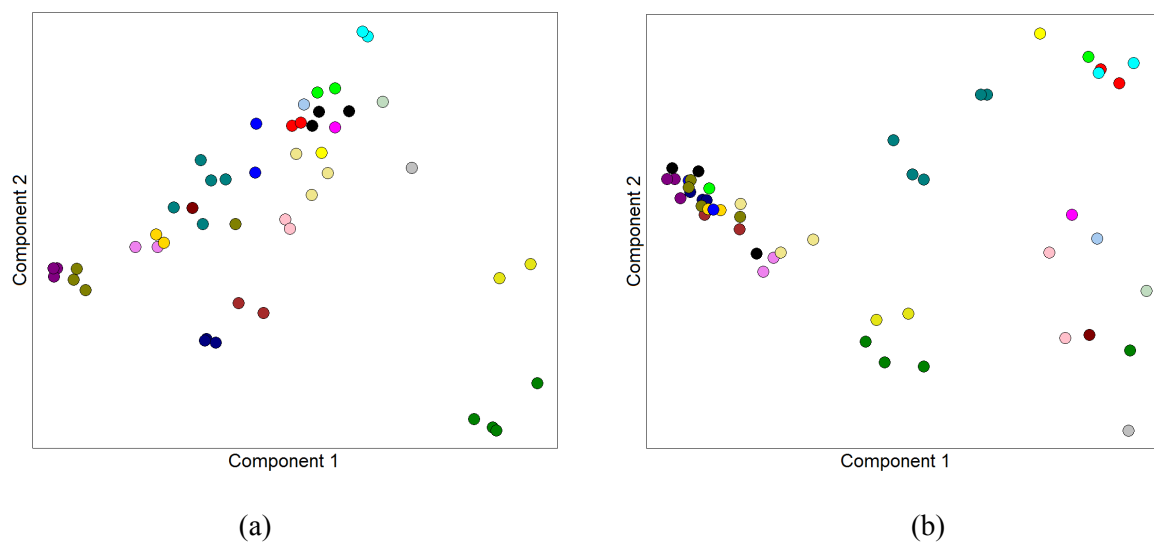
**Figure S6.** Three component PCA analysis of (a) sequence and (b) site similarities. 81.8% and 86.7% of variability is preserved for sequence and site similarity, respectively. Colors correspond to different PTP subtypes.



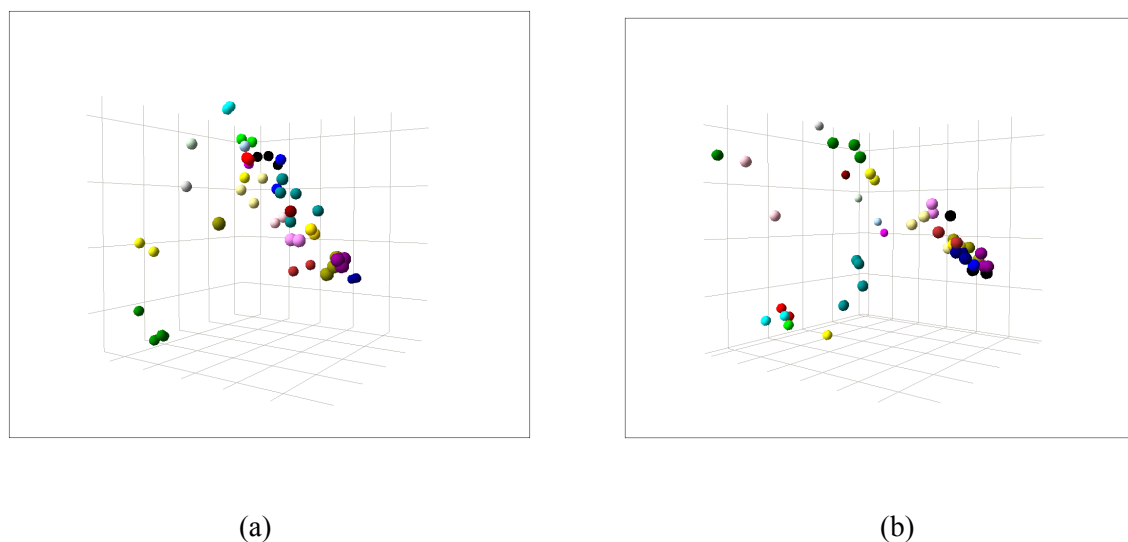
**Figure S7.** Hierarchical clustering of sequence similarities of 49 classical PTP domains (PTP domains annotated by genes).



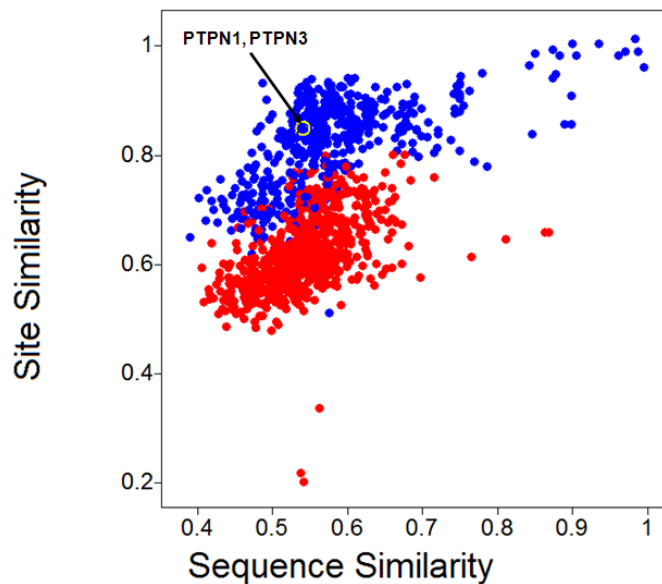
**Figure S8.** Hierarchical clustering of binding site similarities of 49 classical PTP domains (PTP domains annotated by genes).



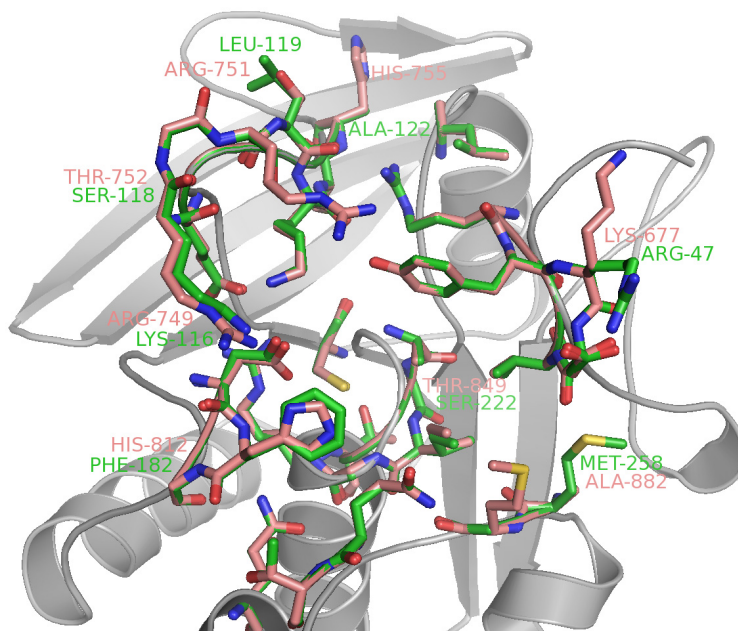
**Figure S9.** Two components PCA analysis of (a) sequence and (b) site similarities. 42.9% and 77.9% of variability is preserved for sequence and site similarity of classical PTPs, respectively. Colors correspond to different classical PTP subtypes.



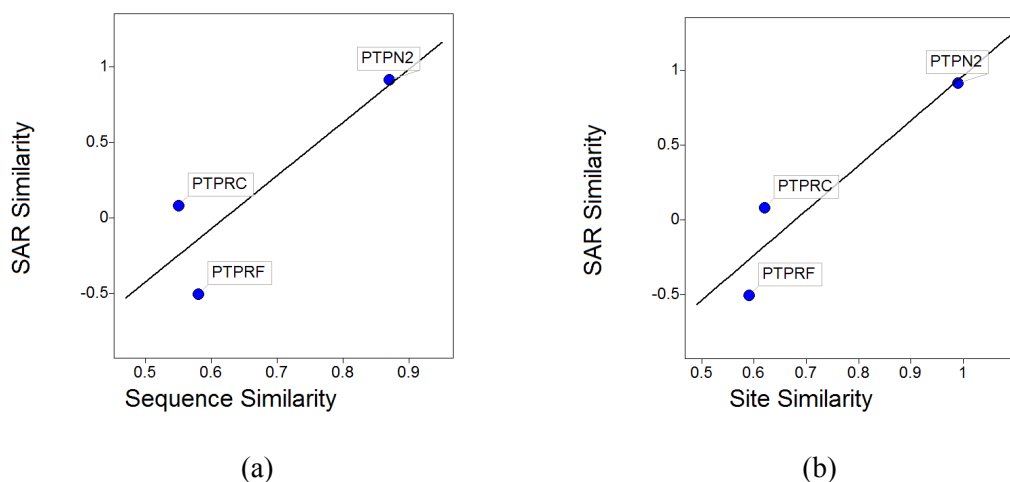
**Figure S10.** Three components PCA analysis of (a) sequence and (b) site similarities. 51.2% and 82.5% of variability is preserved for sequence and site similarity of classical PTPs, respectively. Colors correspond to different classical PTP subtypes.



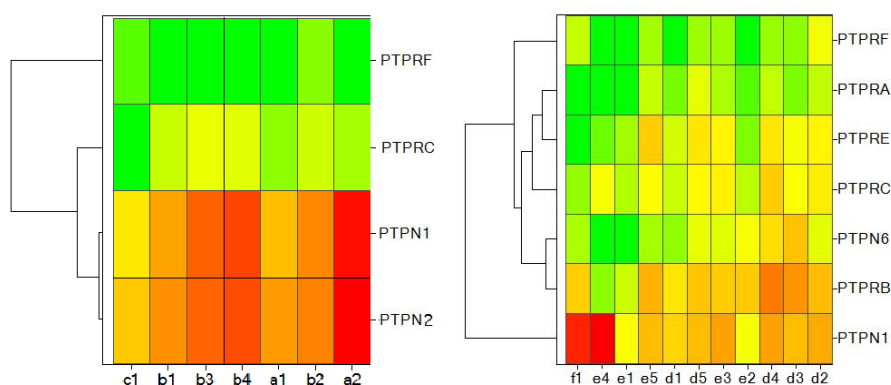
**Figure S11.** Correlation between site and sequence similarities for classical PTP combinations. Blue data present site pairs where the corresponding models were generated with the same template while for red data points the templates were different.



**Figure S12.** Despite relatively lower sequence similarity of 0.56, the catalytic binding sites of PTPN1 (green residues) and PTPN3 (pink residues) are much more conserved (site similarity 0.87). The majority of the residues (20 out of 29) are identical while most of the different residues exhibit similar physicochemical properties (such as Lys and Arg or Thr and Ser, which are aligned in the overlaid sites respectively).



**Figure S13.** Correlation of calculated small molecule activity-based similarity versus (a) sequence and (b) site similarity; relative to PTPN1 as reference.



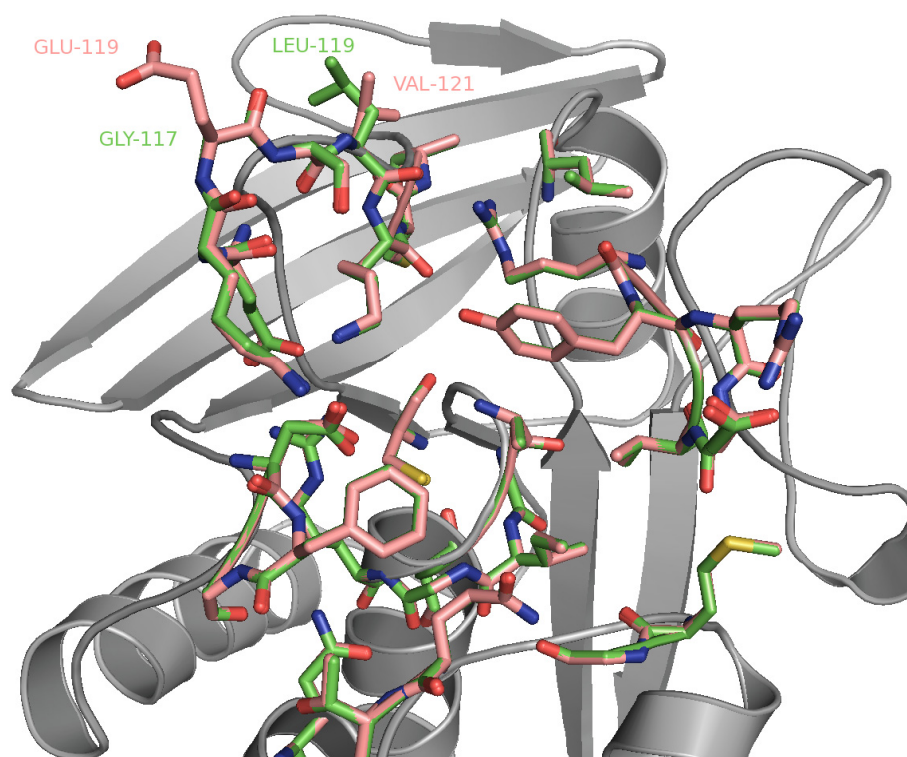
**Figure S14.** Hierarchical clustering of experimental activities of PTP1B inhibitors against a panel of classical PTPs.

**Table S3.** Proximity of PTPs to PTPN1 based on similarity by small molecule SAR, sequences, and sites.

Order derived from SAR data	Order by sequence similarity		Order by site similarity	
PTPN2	PTPN2	correct	PTPN2	correct
PTPRB	PTPRB	correct	PTPRB	correct
PTPN6	PTPRF	incorrect	PTPN6	correct
PTPRC	PTPRA	incorrect	PTPRC	correct
PTPRE	PTPRE	correct	PTPRE	correct
PTPRA	PTPN6	incorrect	PTPRF	incorrect
PTPRF	PTPRC	incorrect	PTPRA	incorrect

**Table S4.** Sequence and Site Similarities of classical PTPs to PTPN1 (ordered high to low similarity)

Gene	Sequence Similarity	Gene	Site Similarity
PTPN2	0.87	PTPN2	0.99
PTPRJ	0.62	PTPN3	0.87
PTPRB	0.6	PTPN14	0.85
PTPRO	0.59	PTPN21	0.84
PTPRS	0.59	PTPN9	0.82
PTPRT	0.59	PTPRB	0.78
PTPRD	0.58	PTPRJ	0.74
PTPRQ	0.58	PTPRH	0.7
PTPRF	0.58	PTPN20A	0.68
PTPRG	0.58	PTPN13	0.67
PTPRK	0.58	PTPRO	0.67
PTPN11	0.58	PTPRQ	0.66
PTPRH	0.57	PTPRN2	0.65
PTPRA	0.57	PTPN22	0.64
PTPRM	0.57	PTPRN	0.63
PTPRE	0.57	PTPN6	0.63
PTPN9	0.57	PTPN18	0.63
PTPN7	0.56	PTPRC	0.62
PTPN6	0.56	PTPRA 2	0.62
PTPN4	0.56	PTPN12	0.61
PTPN3	0.56	PTPN11	0.61
PTPN20A	0.56	PTPN23	0.61
PTPRD 2	0.56	PTPRF 2	0.61
PTPN13	0.55	PTPRM 2	0.61
PTPRS 2	0.55	PTPRG	0.61
PTPN21	0.55	PTPRM	0.61
PTPRA 2	0.55	PTPRE 2	0.61
PTPRC	0.55	PTPN5	0.6
PTPRF 2	0.55	PTPRD 2	0.6
PTPRN	0.55	PTPRK	0.6
PTPRZ1	0.55	PTPRS 2	0.6
PTPN18	0.55	PTPRE	0.59
PTPRR	0.55	PTPRF	0.59
PTPN14	0.54	PTPN4	0.59
PTPRN2	0.54	PTPRT	0.59
PTPN5	0.53	PTPRS	0.59
PTPN12	0.53	PTPRR	0.58
PTPN22	0.52	PTPN7	0.58
PTPN23	0.51	PTPRZ1	0.58
PTPRE 2	0.51	PTPRZ1 2	0.58
PTPRM 2	0.49	PTPRA	0.58
PTPRU 2	0.49	PTPRD	0.57
PTPRZ1 2	0.49	PTPRG 2	0.57
PTPRT 2	0.48	PTPRU	0.57
PTPRU	0.48	PTPRT 2	0.56
PTPRC 2	0.48	PTPRU 2	0.55
PTPRK 2	0.47	PTPRK 2	0.53
PTPRG 2	0.47	PTPRC 2	0.48



**Figure S15.** Binding sites of PTPN1 (green residues) and PTPN2 (pink residues) are extremely similar and the development of selective inhibitors is consequently very difficult. With the binding sites defined by 29 residues only two pairs of corresponding residues are different (Gly117 and Leu119 in PTPN1 vs. Glu119 and Val121 in PTPN2).