

Supplementary Information

Control
hPCIF1 WW domain
PDB; 2JX8

Sequence
IPEELVHAGWEKCWSRR-
ENRPYYFNRFTNQSLWE-
MPVLG

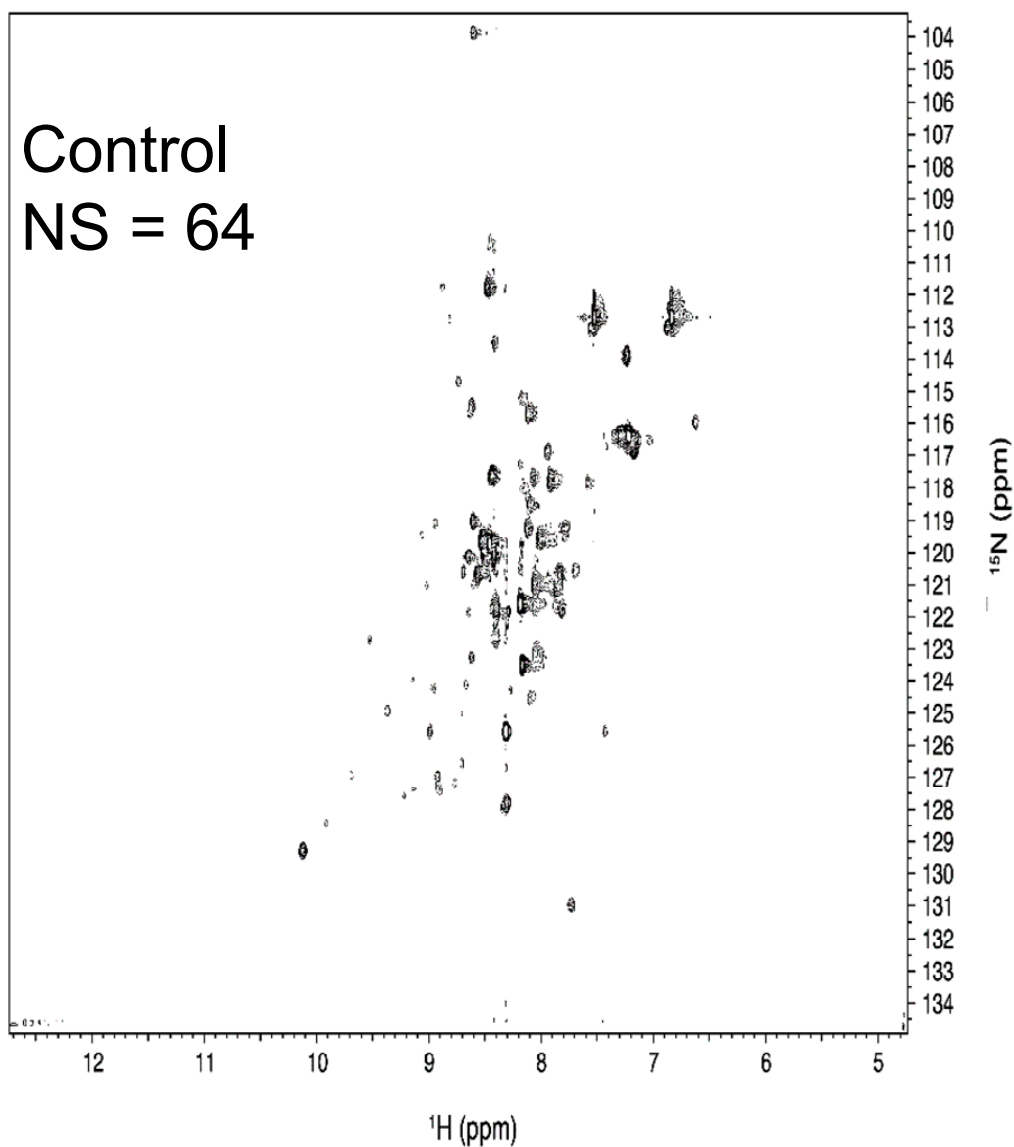


Figure S1. Examples of ^1H - ^{15}N HSQC spectrum of PH0471 chimera with a control folded protein domain. (**Top panel**) Amino acid sequence of hPCIF1 WW domain with its ribbon diagram; (**Bottom panel**) ^1H - ^{15}N HSQC spectrum of PH0471-hPCIF1-WW chimera (scan number = 64).

Table S1. The list of the predicted IDP genes from human genome whose genes were cloned. Gene accession numbers with front/rear PCR primer sequences are summarized.

Sample ID	Accession	Template	Front Primer	Rear Primer
A1	AAH22328.1	synthetic	aggtgttatgaatgtggagaaaagggc cactatgcatatgattgtcaccgagcca	ggctcggtgacaatcatatgcatagtggcc ctttctccacattcataacaccta
A2	NP_005924.2	synthetic	aggaagccacgaggacacctgaagcggtagtgatcgggcca	ggccccgatactaccgcttcgaggtcgtcctctggcttcta
A3	NP_038286.2	synthetic	aagcgcccaccgctccatatccaggacctggttaaggcca	ggccttaccaggtcctggataggaggcggggcgctta
A4	NP_003922.1	synthetic	gcacgaagcgtattactcgaggcaatccg aaagggaatccagttgcgaaagggtggcca	ggccaccttcgcaactggattcccttctg gattgcctcgagtaatacgtctctgca
A5	NP_060834.2	synthetic	aagcacaaggacataaggagcgtcataagccatcatcaccgagcca	ggctcggtgatgatggcttctatgacgctccttatgtcccttgctta
A6	NP_001196.1	synthetic	aaggcagagttattgcaggagggtgattacttcgacaacgagcca	ggctcgtgtcgaagtaaatcacctccctgcaataactctgcctta
B1	AAH07821.1	HeLa	atgatcatggctcactgcag	ggcgctcctgagtagctg
B2	NP_995312.1	placenta	atgatgaggctgccagt	ggtgggtttgttctcagc
B3	NP_066926.1	HeLa	atggcagacaaaccagacatg	ggaatttactccgcttctc
B4	NP_001028678.1	HeLa	atgtcaggagacggagc	ggcaacctgtgcctctcgaatg
C1	NP_570859.1	mouse brain	atggcagctctgcgctat	ggctttggcgtagtagaccag
C2	NP_898870.1	mouse spleen	atgtctgacaaaccgatatgg	ggattcgccagcttgcct
C3	NP_066927.1	mouse brain	atgagagcgaagtggcg	ggccttggacctctgcctcat
C4	NP_002036.1	mouse brain	aaggctcataaggctgcg	ggctttgagcttttctgttatgtgtcc
C5	NP_958782.1	mouse brain	agcctgatcttctgagggcc	ggcgtgcagctcaggacc
C6	NP_003160.2	mouse brain	aatgaactcataggccagactg	ggactctgtggcatcc
C7	NP_005977.2	mouse brain	agctactcgtggccggc	ggccacacctggccacg
C8	NP_055792.1	mouse brain	agccttcaggtctcgggg	ggctagagcccctttagccg
C9	NP_006697.2	mouse brain	agccccactgaggagatatgg	ggctccatctggcttggctcat
C10	NP_055725.2	mouse brain	gtagctctcccagagg	ggcaaggacatactgatgctg
D1	NP_705612.1	HeLa	atgtccaggaaaccacgag	gggagagcctttgttct
D2	AAH14919.1	HeLa	atgactccccgcgtc	ggctctgcctcgcgcac
D3	NP_060710.2	HeLa	gacgtgacctgaggaa	ggctgctgtggagacctg
D4	NP_005115.2	HeLa	gcaggcacttgggattgtgatac	ggccggtgttccacaggctac

Table S1. *Cont.*

Sample ID	Accession	Template	Front Primer	Rear Primer
D5	NP_001012768.1	HeLa	gaaaacagtggtagcagtagt	ggcattgtgccatactgggaac
D6	NP_005091.2	HeLa	agagaaggtgtcactccctg	ggcctcatctctttatcactttccgaag
D7	NP_998734.1	HeLa	gcatttactcccgtcgtggacg	ggcgtgagggtggaccctttata
D8	NP_006281.1	HeLa	gatgtaaatgtgaaacgcccaactg	ggcctctgagcactcatggcataaag
D9	NP_001844.3	HeLa	gacggccaggacggt	ggcccaatggcacct
D10	NP_002537.3	HeLa	gcatttctccaaagtaccttcattatgac	ggcgcacacggcttccactttg
E1	NP_005222.2	HeLa	gcaggattcggcgggaagtatg	ggcttcgacagggtactgtctctg
E2	NP_001964.2	HeLa	agccgcaatgactacatacac	ggcgttcaaagagttgagagtaaatgaag
E3	NP_004703.1	HeLa	gcacaggaggaggaggagct	ggccttctctccgcctctg
E4	NP_002077.1	HeLa	gcatggtttttggcaaatccccaga	ggagaagtccccaggag
E5	NP_998759.1	mouse testis	agtgacatgtacccccgc	ggcaaagtgtgcatgtgctcactc
E6	NP_751896.1	HeLa	gcagagaccactaccag	ggctctgttctcaatgacttcaatagattg
E7	NP_742068.1	HeLa	gcatgtcgagatatgaagcagagag	ggcgtgtgcaaatgtacagctgg
E8	NP_055894.3	HeLa	actgtgtttataccaatgaagagtg	ggctcctcaggggtcagctt
E9	NP_006416.3	HeLa	aaagtgatttggcaaacgaacattg	ggcctgtccactaatttctgagg

Table S2. Summary of the putative IDPs examined in this study.

Status ¹	Sample ID	Accession	Location ²	Sequence	Length
NMR	A2	NP_005924.2	173 to 185	RKPRGRPRSGSDR	13
	A3	NP_038286.2	878 to 889	KRPPPPYPGPGK	12
	A4	NP_003922.1	497 to 514	ARSVLLLEAIRKGIQLRKV	18
	A5	NP_060834.2	184 to 198	KHKGHKERHKKHHHR	15
	A6	NP_001196.1	111 to 124	KAELLQGGDLLRQR	14
	B2	NP_995312.1	full	MMRLPVPSSEGPLRGRDGRGTEDPRADYACIAENKPT	39
	B3	NP_066926.1	full	MADKPDMEIASFDKAKLKKTTTQEKNTLPTKETIEQEKRESEIS	44
	B4	NP_001028678.1	full	MSGDGATEQAAEYVPEKVKKAEEKLEENPYDLDAWSILIREAQV	44
	C1	NP_570859.1	full	MAALRYAGLDDTDEDELPPGWERTTKDGVVYAK	36
	C2	NP_898870.1	full	MSDKPDMAEIEKFDKPKLKKTTTQEKNTLPSKETIEQEKQAGES	44
	C3	NP_066927.1	full	MRAKWRKKRMRLLKRKRKRMRQRSK	25
	C4	NP_002036.1	30 to 52	KAHKAAATKIQASFRKHITRKKLK	23
	C6	NP_003160.2	703 to 730	NELIGQTVRISQGPYKGYIGVVKDATES	28
	C8	NP_055792.1	72 to 106	LQALRVTDLKAALRGLAKSGQKSAVVKRLKDAL	35
	C9	NP_006697.2	132 to 164	PTEEIWVENKTPDGKVVYNNARTRESAWTKPDG	33
	D1	NP_705612.1	full	MSRKPRASSPFSNNHPSTPKRFRQPKREKGPVKEVPGTKGSP	43
	D2	AAH14919.1	full	MTPPRPPEEARRRPASLGHFRQGRSRPVVRRER	34
	D3	NP_060710.2	366 to 400	DVTLRKLDIKVTENVREPLTMHSDREDDTASVSTA	35
	D4	NP_005115.2	724 to 748	GTWDCDTCLVQNKPEAIKCVACETP	25
	D6	NP_005091.2	604 to 634	REGVTPWASFKKMVTPKKRVRPSESDEKED	31
D7	NP_998734.1	134 to 164	FTPVVDDPVTERRKNTQVSDAAYKGVHPH	31	
D8	NP_006281.1	384 to 409	DVKCETPNCPFFMSVNTQPLCHECSE	26	
D10	NP_002537.3	24 to 62	FPPKYLHYDEETSHQLLCKKCPGTYLKQHTAKWKTVC	39	
E1	NP_005222.2	305 to 341	GFGGKYGQKDRMDKNASTFEDVTQVSSAYQKTPVE	37	
E4	NP_002077.1	60 to 96	WFFGKIPRAKAEEMLSKQRHDGAFLIRESESAPGDFS	37	
E5	NP_998759.1	699 to 719	SDMYTPLTTRRNSEYEHMQHF	21	
exp. ³	E7	NP_742068.1	419 to 438	CRDMKQRGGCPRGASCTFAH	20
	E9	NP_006416.3	187 to 212	KVDLAKRTLKAQKLQEELASGKLVEQ	26
not expressed	A1	AAH22328.1	full	RCYECGKGYAYDCHR	17
	C5	NP_958782.1	2864 to 2901	LILLEAQAASGFLDVPVRNRRLLTVNEAVKEGVVGPGLH	38
	D5	NP_001012768.1	250 to 287	ENSGSSSIGIPIAVPTSPPTIGPAAPGSAPGSQYGTM	38
	D9	NP_001844.3	632 to 655	DGQDQAPGEPGPPGDPGLPGAIGA	24
	E2	NP_001964.2	137 to 156	SRNDYIHSGLYSSFTLNSLN	20
	E3	NP_004703.1	258 to 277	QEEEEELQALALSQSEAEK	20
	E6	NP_751896.1	586 to 619	AETHYQLGLAYGNSQYDEAVAQFSKSIEVIENR	34
only PCR	B1	AAH07821.1	full	MIMAHCSLLGSGDPPASTSYSGG	24
	C10	NP_055725.2	565 to 587	VALPEGRQELSDGQVKTGISMSL	39
	C7	NP_005977.2	135 to 157	YSLAGLLAAGAGGGGAAMGV	23
	E8	NP_055894.3	120 to 154	TVFYTNEEWELLDPTPKDLEESIVQEEKKLTPEG	35
		NP_976056.1	full	MSRKPRASSPLSNNHPPTPKRRGSGRFRQPGREKGPKEVPGTKGSP	48
PCR primer not designed		AAH04343.1	full	MQQAAGFSHCLVENKPVFGASQLTSEPFPTSEGEAPKATVDG	43
		NP_068832.1	full	MSDKPDLSEVEKFDKSLKKTNTTEKNTLPSKETIQEKECVQTS	45
		NP_001178.1	full	MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF	43
		NP_001009956.1	full	MEEIPAQEAAGSPRVQFQSLQSECLSPPEQFVQDQDMEQGLTGG	46
		AAM53524.1	full	MGPPGDEEPLGPELHVLMINAPSVLAGFSNAS	33
		NP_067005.1	full	MEEQPECREEKRGSLHVWKSSELVEEDDYYLRHSSSLTYRL	41
		NP_001005414.1	full	MEAAETEAEEAALEVLAEVAGILEPVGLQEEAELPAKILVEFVVVCTR	48
		NP_851813.1	full	MDPEHAKPESSEAPSGNLKQPETAALR	28
		NP_853651.1	full	MSFDNNYHGGQGYAKGGLGCSYGCGLSGYGYACYCPWCYERSWFSGCF	48
		NP_001008505.2	385 to 404	TNHQCLPIPSLSCWALEQGC	20
		NP_892018.1	153 to 191	RCPGNSFSCGNSQCVTQVNPEDDQEDCSDGSDAEHCEC	39
		NP_055379.1	368 to 402	GNGPSSDLSTGSSGGYPDFPASPASWLDVDEHQAQF	35
		NP_443723.2	1077 to 1104	FEDIKILKEKNAELQMTLKLKEESLTR	28

¹ Experimental status of each genes (see Figure 3); ² For intrinsically disorder regions, there amino acid position relative to the full length proteins are indicated; ³ Expression level was too low.