STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 77 (2021)

Supporting information for article:

Using yeast surface display to engineer a soluble and crystallizable construct of HPK1

Wai L. Lau, Bradley Pearce, Heather Malakian, Iyoncy Rodrigo, Dianlin Xie, Mian Gao, Frank Marsilio, Chiehying Chang, Max Ruzanov, Jodi K. Muckelbauer, John A. Newitt, Daša Lipovšek and Steven Sheriff

S1. Construction of HPK1 library

Full-length double-stranded HPK1 with 5' and 3' overhang complementary to pDV-154 cloning sites sequences was assembled by overlap extension of eight pairs of DNA oligos which encode different overlapping fragments of HPK1 (Figure S2).

Step 1: resuspend all oligos to 100nM concentration in deionized water.
For each of R4, R6, R8, F9, F11, and F13, mix oligos that encode the given overlapping fragments with wildtype amino acids and designed mutations (E, Q, T, D, N) into one tube per fragment. Next, combined oligos to obtain the following eight mixes of primers:
Mix 1: 2 vol. oligo F1 and 1 vol. oligo R2
Mix 2: 1 vol. oligo F3 and 2 vol. oligo R4
Mix 3: 2 vol. oligo F5 and 1 vol. oligo R6
Mix 4: 1 vol. oligo F7 and 2 vol. oligo R8
Mix 5: 2 vol. oligo F9 and 1 vol. oligo R10
Mix 6: 1 vol. oligo F11 and 2 vol. oligo R12
Mix 7: 2 vol. oligo F13 and 1 vol. oligo R14
Mix 8: 1 vol. oligo F15 and 2 vol. oligo R16

Assembled PCR reaction mixes for each of the eight oligo mixes are above. The composition of the reaction mix is 1x KOD HotStart buffer, 0.2mM dNTP mix, 1.5mM $MgSO_4$, 1M betaine, 3% (v/v) DMSO, 3% (v/v) oligo mix*, and 2U KOD HotStart polymerase in water. Ran PCR thermal cycles (of the second extension): 368 K for 2 m, then 18 cycles of (368 K for 20 s, 328 K for 10 s, 343 K for 10 s).

After reaction was completed, prepared four reaction mixes, designated "Frag 1" to "Frag 4", by mixing PCR products from Step 1 in the following manner:
Frag 1: 0.05ml of each reaction from Mix 1 and Mix 2
Frag 2: 0.05ml of each reaction from Mix 3 and Mix 4
Frag 3: 0.05ml of each reaction from Mix 5 and Mix 6
Frag 4: 0.05ml of each reaction from Mix 7 and Mix 8
Ran PCR thermal cycles using the same settings as before.

Next, combined the reaction products from the second PCR extension to obtain mixes "Frag N" and "Frag C" as follows:
Frag N: 0.02ml of each of Frag 1 and Frag 2

Frag C: 0.02ml of each of Frag 3 and Frag 4

Assembled two PCR reaction mixes. The composition of the reaction mix is 1x KOD HotStart buffer, 0.2mM dNTP mix, 1.5mM MgSO4, 1M betaine, 3% v/v DMSO, 3% v/v oligo mix, “Frag N” or “Frag C”, and 0.02U/µl KOD HotStart polymerase in water. In the mix, primers F1 and R8 were used in reaction for “Frag N” whereas primers F9 and R16 were used in reaction for “Frag C”. Ran PCR thermal cycles (of the third extension): 368 K for 2 m, then 18 cycles of (368 K for 20 s, 328 K for 10 s, 343 K for 10 s).

After the third extension PCR was completed, combined 0.2ml of each product from PCR using “Frag N” and “Frag C”, and ran the same PCR thermal cycles again. The PCR products were purified by agarose gel electrophoresis and extraction from gel slices using GE Healthcare Illustra Gel Band Purification Kit (Marlborough, MA).

S1.1. Amplification and purification of HPK1 library

Full-length HPK1 with 5’ and 3’ overhang complementary to pDV-154 was amplified by polymerase chain reaction (PCR). A typical 0.1ml volume PCR reaction was composed of 0.01ml DNA template, 0.2 µM forward and reverse primers, 1x buffer, 1mM $MgSO_4$, 3% DMSO, 1M betaine, 0.2 mM of dNTP mix, and 2U *KOD Hot Start Polymerase*. KOD HotStart Polymerase and components were purchased from EMD Millipore (Bellerica, MA) and betaine and DMSO were purchased from Sigma-Aldrich (St. Louis, MO). To amplify PROfusion population of different designs, different pairs of forward and reverse primers are used in PCR amplification. Nucleotide sequences of the forward and reverse primers are 5’-TTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGACGTCGTGGACCCTGACATTTTC-3’ and 5’-GTTCAGATCCTCTTCTGAGATGAGTTTTTGTTCTAAGCTTCCGCTTGCGGCCGCTCCTCGGAGCTTCCTGAACTCCATGTGCCGCCGACA-3’, respectively.

PCR reactions were conducted on a BioRad MJ Mini Gradient Thermal Cycler (BioRad, Hercules, CA) running the following conditions: one cycle at 300 K for 2 m, followed by 30 cycles at 300 K for 20 s, 328 K for 10 s, and 343 K for 15 s. The amplified DNAs were purified by agarose gel electrophoresis and extracted from gel slices using Qiagen Gel Extraction Kit. After purification and gel extraction, a second PCR was performed to extend 5’ and 3’ overhang sequence for homologous recombination using the purified DNA as template. PCR products were amplified and purified by applying the similar set of conditions, except annealing time was increased to 15 s.

S2. Yeast display plasmid pDV-154

Yeast display plasmid pDV-154 was derived from pDV-23 (Lipovšek *et al.*, 2018) reported earlier. The yeast 2µ origin of replication (Zakian *et al.*, 1979) in pDV-23 was replaced by a CEN (Grugge *et al.*, 2017) origin of replication to become pDV-154. Complete nucleotide sequence of pDV-154 is shown below:

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTC
ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGG
TGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGT
GCACCATATCGACTACGTCGTTAAGGCCGTTTCTGACAGAGTAAAATTCTTGAGGGAACT
TTCACCATTATGGGAAATGGTTCAAGAAGGTATTGACTTAAACTCCATCAAATGGTCAGG
TCATTGAGTGTTTTTTATTTGTTGTATTTTTTTTTTTTTAGAGAAAATCCTCCAATATATAA
ATTAGGAATCATAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTGCCCTCCTC
CTTGTCAATATTAATGTTAAAGTGCAATTCTTTTTCCTTATCACGTTGAGCCATTAGTATC
AATTTGCTTACCTGTATTCCTTTACATCCTCCTTTTTCTCCTTCTTGATAAATGTATGTAGA
TTGCGTATATAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATTTCGTGTCGTTTCT
ATTATGAATTTCATTTATAAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTTA
AGCAAGGATTTTCTTAACTTCTTCGGCGACAGCATCACCGACTTCGGTGGTACTGTTGGA
ACCACCTAAATCACCAGTTCTGATACCTGCATCCAAAACCTTTTTAACTGCATCTTCAATG
GCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCATTGCAGCAGACAAGATA
GTGGCGATAGGGTTGACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGGCATGGTTCG
TACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACAA
ACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGC
TGGTGATTATAATACCATTTAGGTGGGTTGGGTTCTTAACTAGGATCATGGCGGCAGAAT
CAATCAATTGATGTTGAACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGT
TTTTCTCCATAATCTTGAAGAGGCCAAAACATTAGCTTTATCCAAGGACCAAATAGGCAA
TGGTGGCTCATGTTGTAGGGCCATGAAAGCGGCCATTCTTGTGATTCTTTGCACTTCTGG
AACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCATCGTCTTCCTTTCTCTTACCA
AAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTTTAGCAAATTGT
GGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAGTT
GGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTCAGGTCTAACACTACC
GGTACCCCATTTAGGACCACCCACAGCACCTAACAAAACGGCATCAGCCTTCTTGGAGG
CTTCCAGCGCCTCATCTGGAAGTGGAACACCTGTAGCATCGATAGCAGCACCACCAATTA
AATGATTTTCGAAATCGAACTTGACATTGGAACGAACATCAGAAATAGCTTTAAGAACCT
TAATGGCTTCGGCTGTGATTTCTTGACCAACGTGGTCACCTGGCAAAACGACGATCTTCT
TAGGGGCAGACATTAGAATGGTATATCCTTGAAATATATATATATATATTGCTGAAATGT
AAAAGGTAAGAAAAGTTAGAAAGTAAGACGATTGCTAACCACCTATTGGAAAAAACAAT

AGGTCCTTAAATAATATTGTCAACTTCAAGTATTGTGATGCAAGCATTTAGTCATGAACG
CTTCTCTATTCTATATGAAAAGCCGGTTCCGGCGCTCTCACCTTTCCTTTTTCTCCCAATTT
TTCAGTTGAAAAAGGTATATGCGTCAGGCGACCTCTGAAATTAACAAAAAATTTCCAGTC
ATCGAATTTGATTCTGTGCGATAGCGCCCCTGTGTGTTCTCGTTATGTTGAGGAAAAAAA
TAATGGTTGCTAAGAGATTCGAACTCTTGCATCTTACGATACCTGAGTATTCCCACAGTT
AACTGCGGTCAAGATATTTCTTGAATCAGGCGCCTTAGACCGCTCGGCCAAACAACCAAT
TACTTGTTGAGAAATAGAGTATAATTATCCTATAAATATAACGTTTTTGAACACACATGA
ACAAGGAAGTACAGGACAATTGATTTTGAAGAGAATGTGGATTTTGATGTAATTGTTGG
GATTCCATTTTTAATAAGGCAATAATATTAGGTATGTAGATATACTAGAAGTTCTCCTCG
ACCGGTCGATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCA
GGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCA
TTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC
CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCAC
CCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGG
AGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGA
AGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGT
AACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCATTCGCCATTCAGG
CTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCG
AAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACG
ACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGC
TTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCCGCCCGGGGGATCTAGCTA
TACTTCGGAGCACTGTTGAGCGAAGGCTCATTAGATATATTTTCTGTCATTTTCCTTAACC
CAAAAATAAGGGAGAGGATCCAAAAAGCGCTCGGACAACTGTTGCCCGTGATCCGAAGG
ACTGGCTTATACAGTGTTCACAAAATAGCCAAGCTGAAAATAATGTGTAGCCTTTAGCTA
TGTTCAGTTAGTTTGGCTAGCAAAGATATAAAAGCAGGTCGGAAATATTTATGGGCATTA
TTATGCAGAGCATCAACATGATAAAAAAAAACAGTTGAATATTCCGAGCTCATCACACAA
ACAAACAAAACAAAATGATGAGATTTCCTTCAATTTTTACTGCCGTTTTATTCGCAGCAT
CCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCG
GCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCA
TTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCT
GCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGCTAGCGTTTCTGATGTGCCGCGCGA
CCTGGAAGTGGTTGCTGCCACCCCCACCAGCCTGCTGATCAGCTGGCCATGGATCTGCTC
TGCTTCCTATCCATGGGCCGGGTCATCGTGGATTCATATGTCTTCCGCTGCGACGCATATG
CCAATTTCCATTAATTACCGCACAGAAATTGACAAACCATCCCAGGCGGCCGCAAGCGG
AAGCTTAGAACAAAAACTCATCTCAGAAGAGGATCTGAACGCTAGCGCCAAAAGCTCTT
TTATCTCAACCACTACTACTGATTTAACAAGTATAAACACTAGTGCGTATTCCACTGGAT

CCATTTCCACAGTAGAAACAGGCAATCGAACTACATCAGAAGTGATCAGTCATGTGGTG
ACTACCAGCACAAAACTGTCTCCAACTGCTACTACCAGCCTGACAATTGCACAAACCAGT
ATCTATTCTACTGACTCAAATATCACAGTAGGAACAGATATTCACACCACATCAGAAGTG
ATTAGTGATGTGGAAACCATTAGCAGAGAAACAGCTTCGACCGTTGTAGCCGCTCCAAC
CTCAACAACTGGATGGACAGGCGCTATGAATACTTACATCCCGCAATTTACATCCTCTTC
TTTCGCAACAATCAACAGCACACCAATAATCTCTTCATCAGCAGTATTTGAAACCTCAGA
TGCTTCAATTGTCAATGTGCACACTGAAAATATCACGAATACTGCTGCTGTTCCATCTGA
AGAACCCACTTTTGTAAATGCCACGAGAAACTCCTTAAATTCCTTTTGCAGCAGCAAACA
GCCATCCAGTCCCTCATCTTATACGTCTTCCCCACTCGTATCGTCCCTCTCCGTAAGCAAA
ACATTACTAAGCACCAGTTTTACGCCTTCTGTGCCAACATCTAATACATATATCAAAACG
GAAAATACGGGTTACTTTGAGCACACGGCTTTGACAACATCTTCAGTTGGCCTTAATTCT
TTTAGTGAAACAGCACTCTCATCTCAGGGAACGAAAATTGACACCTTTTTAGTGTCATCC
TTGATCGCATATCCTTCTTCTGCATCAGGAAGCCAATTGTCCGGTATCCAACAGAATTTC
ACATCAACTTCTCTCATGATTTCAACCTATGAAGGTAAAGCGTCTATATTTTTCTCAGCTG
AGCTGGGTTCGATCATTTTTCTGCTTTTGTCGTACCTGCTATTCTAAGATCTGAATAAACG
CGTGAAGCTTATCGATAGATCAATTTTTTTCTTTTCTCTTTCCCCATCCTTTACGCTAAAAT
AATAGTTTATTTTATTTTTTGAATATTTTTTATTTATATACGTATATATAGACTATTATTTA
TCTTTTAATGATTATTAAGATTTTTATTAAAAAAAAATTCGCTCCTCTTTTAATGCCTTTAT
GCAGTTTTTTTTTCCCATTCGATATTTCTATGTTCGGGTTCAGCGTATTTTAAGTTTAATAA
CTCGAAAATTCTGCGTTCGTTAAAGCTAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTG
TGTGAAATTGTTATCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTT
AATTTCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTC
ACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATG
AGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCT
GTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT
TGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC
TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC
CTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT
TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG
CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC

TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTA
AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG
CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGA
CTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCA
ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGC
CGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAA
TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
CATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCC
TTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATG
GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGG
CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGA
AAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
TAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAAT
GTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCT
CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
CATTTCCCCGAAAAGTGCCACCTGGGTCCTTTTCATCACGTGCTATAAAAATAATTATAA
TTTAAATTTTTTAATATAAATATATAAATTAAAAATAGAAAGTAAAAAAAGAAATTAAA
GAAAAAATAGTTTTTGTTTTCCGAAGATGTAAAAGACTCTAGGGGGATCGCCAACAAAT
ACTACCTTTTATCTTGCTCTTCCTGCTCTCAGGTATTAATGCCGAATTGTTTCATCTTGTCT
GTGTAGAAGACCACACACGAAAATCCTGTGATTTTACATTTTACTTATCGTTAATCGAAT
GTATATCTATTTAATCTGCTTTTCTTGTCTAATAAATATATATGTAAAGTACGCTTTTTGTT
GAAATTTTTTAAACCTTTGTTTATTTTTTTTTCTTCATTCCGTAACTCTTCTACCTTCTTTAT
TTACTTTCTAAAATCCAAATACAAAACATAAAAATAAATAAACACAGAGTAAATTCCCA
AATTATTCCATCATTAAAAGATACGAGGCGCGTGTAAGTTACAGGCAAGCGATCCGTCCT
AAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTC
GTC

S2.1. Preparation of linerarized pDV-154 by restriction enzyme digestion

Yeast display plasmid pDV-154 was linearized by restriction enzyme digestion using XhoI and NotI-HF. In a typical reaction, uncut pDV-154 was mixed with NotI-HF and XhoI to a final concentration of 1µg/µl, ~1U/µl, and ~1U/µl in 1x CutSmart buffer. The digestion reaction was incubated at 310 K for two hours. To ensure complete removal of spacer sequence, NdeI was mixed into the double-digested reaction to a final concentration of ~1U/µl for digestion at 310 K for two hours. All restriction enzymes used were purchased from New England Biolabs (Ipswich, MA). After reaction was completed, the triple-digested pDV-154 was purified by agarose gel electrophoresis and extracted from gel slices by applying QIAquick Gel Extraction Kit purchased from Qiagen Inc.

S3. Growth and induction of yeast cultures

Yeast cells grown in selective media were harvested by centrifugation, and resuspended in induction media, as described previously (Lipovšek *et al*., 2018). However, in addition to a selection that used induction in YPG (10 g/l extract, 20 g/l bacto peptone, 20 g/l galactose), a parallel selection was performed using induction in YDOG media (6.7 g/l yeast nitrogen base with ammonium sulfate, 1.6 g/l yeast synthetic drop out media supplements without leucine (Sigma Y1376), and 20 g/l galactose).

Table S1 Percent yeast cells captured by fluorescence activated cell sorting during yeast display selection in Round 1-4.

	R1	R2	R3	R4
Selection using YPG induction	3.7%	10.9%	30.6%	5.6%
Selection using YDOG induction	0.1%	0.9%	14.9%	5.3%

Table S2 Enriched amino acid residue in aggregation prone sites detected by DNA sequencing of post round-3 and -4 yeast display populations.

	Pos. 64	Pos. 80	Pos. 112	Pos. 170	Pos. 188	Pos. 221	Pos. 225	Pos. 285
mutation	E, Q	Q	E	L*	D, T	D, E	E	D, Q

*L170 is the residue found in wild-type HPK1; the remaining residues found in enriched variants are mutations from the wild type.

Table S3 Amino acid residue in selected positions of wild-type HPK1 and variants expressed/characterized. Yield is listed as the amount of protein, in milligrams, purified per liter culture of baculovirus-infected insect cell expressions in shake flasks.

	Expression vector	Pos. 64	Pos. 80	Pos. 112	Pos. 170	Pos. 188	Pos. 221	Pos. 225	Pos. 285	Yield (mg/L)
WT	TVMV-HPK1(1-346)-pFB	L	L	L	L	L	L	F	L	N.D.*
M1	TVMV-HPK1(1-346)-M1-pFB	Q	Q	E	L	D	D	E	Q	9.4
M2	TVMV-HPK1(1-346)-M2-pFB	Q	Q	E	L	D	D	E	D	9.1
M3	TVMV-HPK1(1-346)-M3-pFB	Q	Q	E	L	D	E	E	Q	9.7
M4	TVMV-HPK1(1-346)-M4-pFB	Q	Q	E	L	D	E	E	D	N.D.
M5	TVMV-HPK1(1-346)-M5-pFB	Q	Q	E	L	T	D	E	Q	11.9
M6	TVMV-HPK1(1-346)-M6-pFB	Q	Q	E	L	T	D	E	D	8.9
M7	TVMV-HPK1(1-346)-M7-pFB	Q	Q	E	L	T	E	E	Q	N.D.
M8	TVMV-HPK1(1-346)-M8-pFB	Q	Q	E	L	T	E	E	D	9.5
M9	TVMV-HPK1(1-346)-M9-pFB	E	Q	E	L	D	D	E	Q	8.4
M10	TVMV-HPK1(1-346)-M10-pFB	E	Q	E	L	D	D	E	D	N.D.
M11	TVMV-HPK1(1-346)-M12-pFB	E	Q	E	L	D	E	E	Q	N.D.
M12	TVMV-HPK1(1-346)-M12-pFB	E	Q	E	L	D	E	E	D	6.9
M13	TVMV-HPK1(1-346)-M13-pFB	E	Q	E	L	T	D	E	Q	3.4
M14	TVMV-HPK1(1-346)-M14-pFB	E	Q	E	L	T	D	E	D	10.3
M15	TVMV-HPK1(1-346)-M15-pFB	E	Q	E	L	T	E	E	Q	6.9
M16	TVMV-HPK1(1-346)-M16-pFB	E	Q	E	L	T	E	E	D	N.D.
M17	TVMV-HPK1(1-346)-M17-pFB	L	Q	E	L	T	D	E	Q	N.D.
M18	TVMV-HPK1(1-346)-M18-pFB	Q	L	E	L	T	D	E	Q	N.D.
M19	TVMV-HPK1(1-346)-M19-pFB	Q	Q	L	L	T	D	E	Q	N.D.
M20	TVMV-HPK1(1-346)-M20-pFB	Q	Q	E	L	L	D	E	Q	N.D.
M21	TVMV-HPK1(1-346)-M21-pFB	Q	Q	E	L	T	L	E	Q	2.4
M22	TVMV-HPK1(1-346)-M22-pFB	Q	Q	E	L	T	D	F	Q	2.8
M23	TVMV-HPK1(1-346)-M23-pFB	Q	Q	E	L	T	D	E	L	N.D.
M24	TVMV-HPK1(1-346)-M24-pFB	Q	Q	E	L	T	L	F	Q	N.D.
M25	TVMV-HPK1(1-346)-M25-pFB	L	L	L	L	L	D	E	L	4.4
Final construct	TVMV-HPK1(1-319)-M25-pFB	L	L	L	L	L	D	E	L	4.3

*N.D. Not determined due low yield.

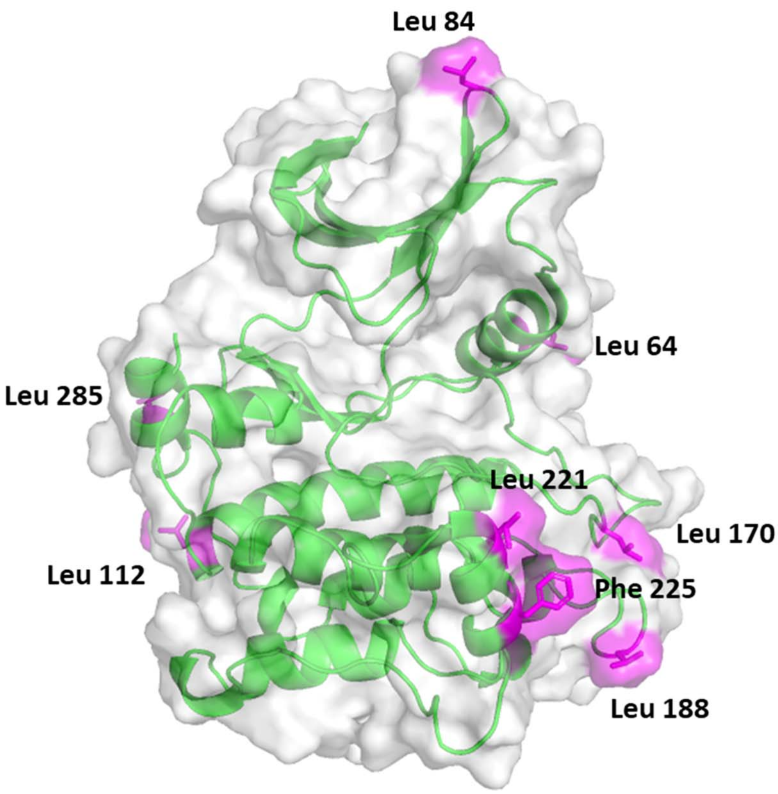


Figure S1 Homology model of HPK1 based on the crystal structure of MST1 (3COM). The backbone of the protein is shown as green cartoon and the eight residues to interrogate for YSD selections are shown in magenta.

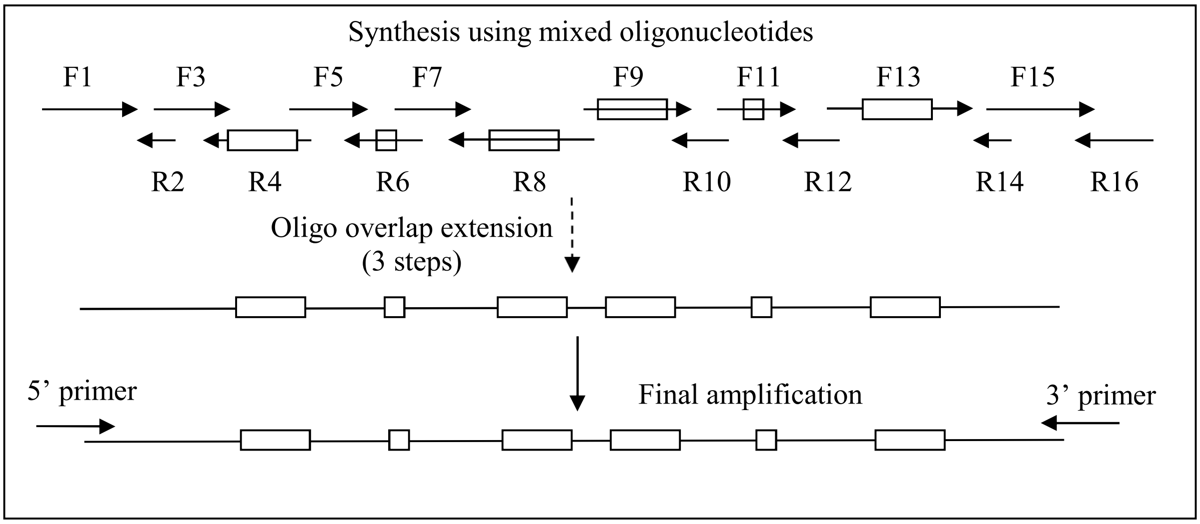
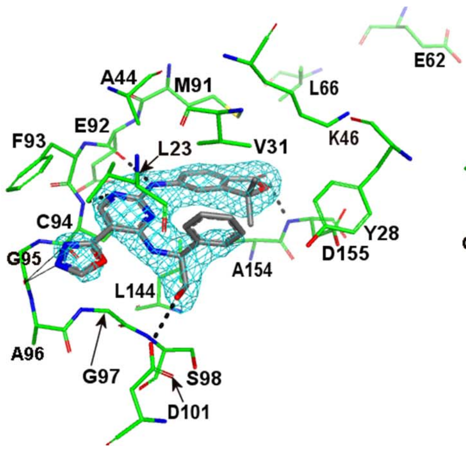
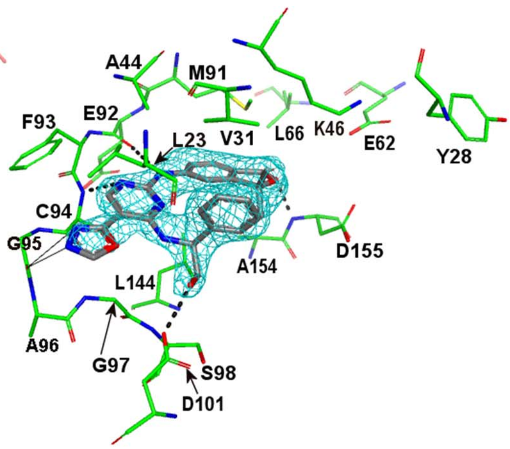


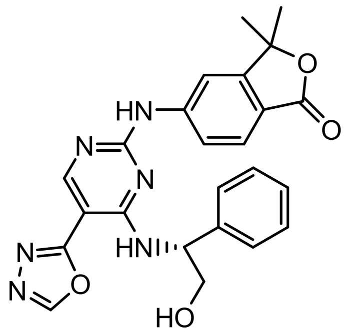
Figure S2 Schematics illustrating library construction by extension PCR using overlapping oligos encoding HPK1 fragments. In the diagram, R4, R6, R8, F9, F11, and F13 are represent mix of oligos encoding diversification in selected positions.



(a)



(b)



(c)

Figure S3 Omit Fo-Fc electron density contoured at 3 r.m.s.d. (cyan) with the final model. At the same contour level, the electron density is poorer for the ligand in chain A than chain B. The ligand (Compound K, You *et al*., 2020) forms direct hydrogen bonds from one of the pyrimidinyl nitrogen atoms (N3) to Cys 94 N, from the amino (N9) to Glu 92 O, from the hydroxyl (O13) to Asp 101 OD2, and from the carbonyl oxygen (O30) to Asp 155 N (of the DGF motif). The two nitrogen atoms (N33, N34) of the oxadiazolyl are close enough to Gly 95 O to form hydrogen bonds, but none of the atoms has a hydrogen atom attached, so those interactions are shown with thin lines. Atom colors: carbon (protein - green; ligand - gray), nitrogen - blue; oxygen -red; sulfur - yellow). Hydrogen bonds are shown as a series of small prolate ellipsoids (black). (a) chain A; (b) chain B; (c) Two-dimensional representation of the Compound K.

References

Grugge, R., Rudolf, F. (2017). *Yeast* **34**, 205-221.

Lipovšek, D., Carvajal, I., Allentoff, A. J., Barros Jr, A., Brailsford, J., Cong, Q., Cotter, P., Gangwar, S., Hollander, C., Lafont, V., Lau, W. L., Li, W., Moreta, M., O'Neil, S., Pinckney, J., Smith, M. J., Su, J., Terragni, C., Wallace, M. A., Wang, L., Wright, D. M., Marsh, H. N., Bryson, J. W. (2018). *Protein Eng. Des. Sel.* **31**,159-171.

Zakian, V. A., Brewer, B. J., Fangman, W. L. (1979). *Cell* **17**, 923-934.