Supplemental Information:

Table S1: Primers used in this work. Relates to Experimental Procedures:

Primer Name	Sequence
MpPksCT-g5	ATGATTGACTCAACTTCGCACTCAAATCTGAG
MpPksCT-g3	TTAATCTAGAAATCCCATGGTCTTCCATGC
MpPksCT-ex1-5	GCATATTAATGATTGACTCAACTTCGCACTC
MpPksCT-ex1-3.2	GTCAACGATGACAGACACATATGCATCAGGGAATGTCTCGAGAACG
MpPksCT-ex2-5.2	CTCGAGACATTCCCTGATGCATATGTGTCTGTCATCGTTGAC
MpPksCT-ex2-3	GCATGCGGCCGCATCTAGAAATCCCATGGTCTTCCATG
MpPksCT-SAT-3	GCATGCGGCCGCTCCACCGAGAAGCTGCCC
MpPksCT-SKM3-3	GCATGCGGCCGCAGAGCGCTGGGATGCGTCTTG
MpPksCT-PT-5	GCATATTAATGACAGCAATTATCGAGGCACC
MpPksCT-PT-3	GCATGCGGCCGCGCCGCTTACCGCAGTAGACG
MpPksCT-ACP-5	GCATATTAATGACACCTCCGACAAAGAAAGCGC
MpPksCT-ACP-3	GCATGCGGCCGCACCGTTACCATTAACGACACCATTAGTGC
MpPksCT-CMeT-5	GCATATTAATGGTCTTGTTTCCTGAGCTAGGGGG
MpPksCT-CMeT-3	GCATGCGGCCGCGTCTGTCAATGGGACCTGCACC
MpPksCT-CMeTAseI-5	CGAGGAGGCCAGACTGATTGATATCAACG
MpPksCT-CMeTAseI-3	CGTTGATATCAATCAGTCTGGCCTCCTCG
MpAct-5	GGAATTCTGCAGATTCTACAACGAACTCCG
MpAct-3	GGAATTCTGCAGTCAGGGAGTTCATAGGAC
MpPksCT-CMeT-Y1955F-5	GTCACTGATGTCTTTGCCAAATCTCC
MpPksCT-CMeT-Y1955F-3	GGAGATTTGGCAAAGACATCAGTGAC
MpPksCT-CMeT-Y1955A-5	GTCACTGATGTCGCGGCCAAATCTCC
MpPksCT-CMeT-Y1955A-3	GGAGATTTGGCCGCGACATCAGTGAC
MpPksCT-CMeT-H2067A-5	CAAATTGTGTGGCGGCGACCCGG
MpPksCT-CMeT-H2067A-3	CCGGGTCGCCGCCACACAATTTG
MpPksCT-CMeT-H2067Q-5	CAAATTGTGTGCAGGCGACCCGG
MpPksCT-CMeT-H2067Q-3	CCGGGTCGCCTGCACACAATTTG

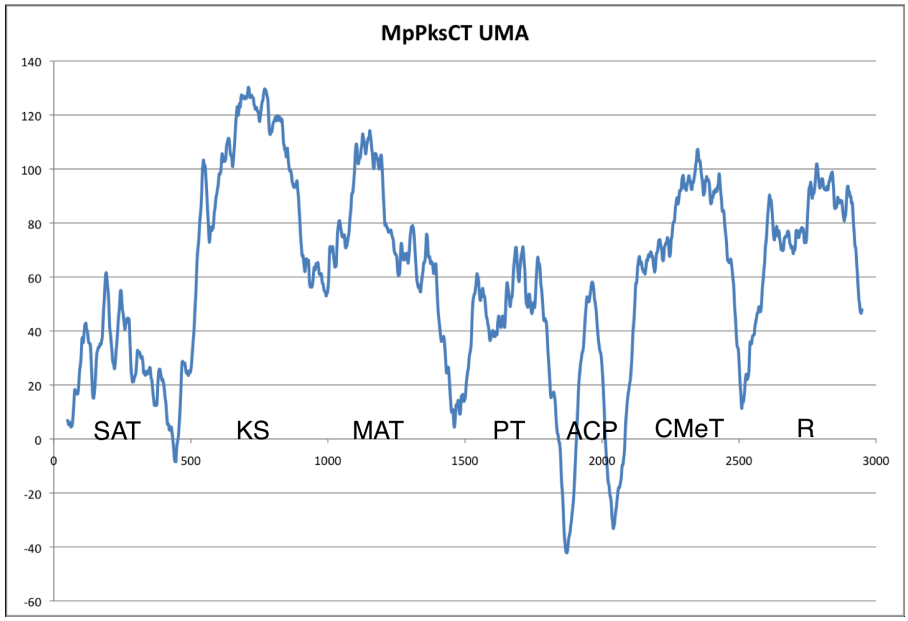
Figure S1, relates to Figures 1 and 2: Domain deconstruction, exon revision, and starter unit verification for PksCT.

A) Sequence conservation, predicted secondary structure, and local hydrophobicity are combined to generate an UMA score for each position. Low scores indicate likely linker regions and were used to guide the domain boundaries for mono- or multi-domain expression constructs. For details on input sequences and algorithm settings, please see Supplemental Experimental Procedures.

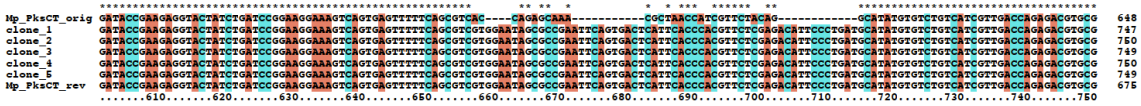
B) Five independent clones amplified from a cDNA library generated from total RNA under citrinin producing conditions. Each clone exactly matched the FGENESH prediction for the exon 1:exon 2 boundary.

C) Radiolabel assay of acetyl starter unit transfer in PksCT. Lane 1: PksCT SAT, Lane 2: PksCT SAT + ACP without Sfp activation, Lane 3: PksCT SAT + ACP with Sfp activation. NR-PKS ACP domains are often partially phosphopantetheinylated during expression, accounting for the low ACP signal in lane 2.

A)



B)



C)

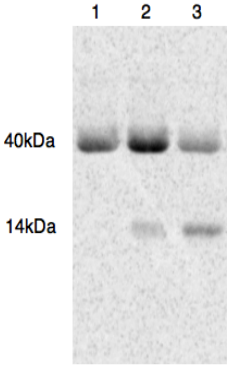


Table S2, relates to Figure 2: Plasmids used in this work:

PksCT domain(s)	Sequence	Plasmid name	Parent vector	Tag	MW (kDa)
SAT	M1-G380	pEMpPksCT-SAT	pET-24a	C-His6	41.8
SAT-KS-MAT	M1-I1288	pEMpPksCT-SKM3	pET-24a	C-His6	144.0
PT	T1285-G1651	pEMpPksCT-PT	pET-24a	C-His6	41.5
ACP	T1652-G1779	pEMpPksCT-ACP	pET-24a	C-His6	15.0
CMeT	V1780-D2182	pEMpPksCT-CMeT	pET-24a	C-His6	46.6
CMeT-Y1955A	V1780-D2182	pEMpPksCT-CMeT-Y1955A	pET-24a	C-His6	46.6
CMeT-Y1955F	V1780-D2182	pEMpPksCT-CMeT-Y1955F	pET-24a	C-His6	46.6
CMeT-H2067A	V1780-D2182	pEMpPksCT-CMeT-H2067A	pET-24a	C-His6	46.6
CMeT-H2067Q	V1780-D2182	pEMpPksCT-CMeT-H2067Q	pET-24a	C-His6	46.6
CMeT-R	V1780-D2593	pEMpPksCT-CMeT-R	pET-24a	C-His6	92.7

Table S3, relates to Figure 3: Crystallographic data collection and refinement statistics.

	CMeT native		CMeT SeMet
Data collection			
Space group	H3	H3†	H3†
Cell dimensions			
a, b, c (Å)	98.37, 98.37, 133.43	98.52, 98.52, 133.20	98.13, 98.13, 124.45
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Wavelength (Å)	0.99998	1.90747	0.97929
Swipes	1	2	2
Resolution (Å)	71.80 – 1.65‡	71.85 - 2.05	50.21 – 1.85
R_{merge} (%)*	8.0 (159.6)	8.4 (102.9)	13.5 (274.2)
$I/\sigma I$*	10.61 (1.37)	15.62 (1.38)	13.75 (1.12)
$CC_{1/2}$ (%)*	99.9 (76.0)	99.8 (71.1)	99.9 (57.1)
Completeness (%)*	98.9 (94.3)	97.9 (81.7)	99.8 (99.7)
Redundancy*	6.6 (6.6)	8.5 (2.6)	10.4 (10.2)
Unique reflections*	57,453 (4,045)	59,253 (3,685)	76,184 (5,646)
Refinement			
Resolution (Å)	71.80 – 1.65		
R_{work}/ R_{free}	0.19 / 0.22		
Ramachandran favored / outlier (%)	99.48 / 0.00		
No. atoms	6,512		
Protein	6,173		
Ligand	45		
Solvent	294		
B-factors	36.22		
Protein (Å^2)	36.29		
Ligand (Å^2)	26.84		
Solvent (Å^2)	36.11		
R.m.s deviations			
Bond lengths (Å)	0.009		
Bond angles (°)	0.975		

The resolution cutoff was determined by $CC_{1/2}$ criterion (Karplus and Diederichs, 2012). *, Highest resolution shell is shown in parenthesis. Datasets of native CMeT were acquired from the same crystal. †Anomalous data processing was applied. ‡A deprecated conservative resolution cutoff at $I/\sigma I$=2.91 and Rmerge= 69.8% is reached at a resolution of 1.84 Å.

Figure S2, relates to Figures 2 and 5A: UV-Vis spectra for compounds 2-8.

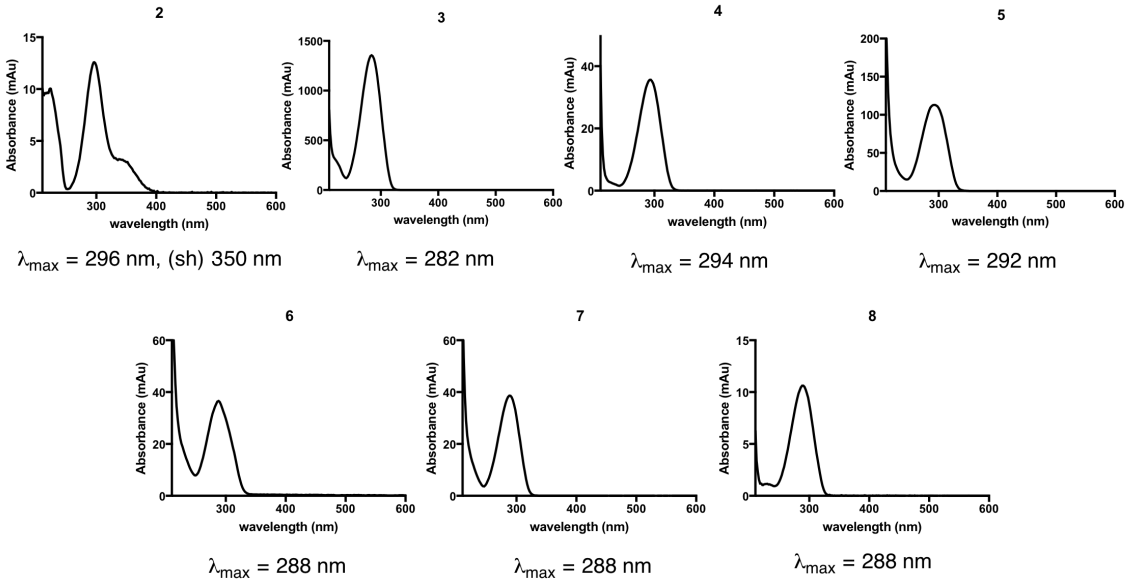


Table S4, relates to Figures 2 and 5A: Detected masses for compounds 2-8.

#	Calculated [MH+] or [MNa+] m/z	Detected m/z (ppm)	Formula
2	237.1127	237.1121(2.53)	$C_{13}H_{17}O_4$
3	127.0395	127.0391 (3.15)	$C_6H_7O_3$
4	155.0708	155.0704 (2.58)	$C_8H_{11}O_3$
5	211.0970	211.0970 (0.00)	$C_{11}H_{15}O_4$
	233.0787	233.0790 (1.29)	$C_{11}H_{14}O_4Na$
	169.0865	169.0859 (2.37)	$C_9H_{13}O_3[-C_2H_2O]$
6	253.1075	253.1076 (0.40)	$C_{13}H_{17}O_5$
	275.0895	275.0893 (0.73)	$C_{13}H_{16}O_5Na$
	211.0970	211.0965 (2.73)	$C_{11}H_{15}O_4[-C_2H_2O]$
	155.0708	155.0701 (4.51)	$C_8H_{11}O_3[-C_5H_6O_2]$
7	141.0552	141.0553 (0.71)	$C_7H_9O_3$
8	141.0552	141.0552 (0)	$C_7H_9O_3$


active
inactive
trans-AT PKS
cis-AT PKS
NR-PKS
HR-PKS
FAS
Sequence Logo
secondary structure:
active site residues:
PksCT (Mp)
linker
N-terminal subdomain
C-terminal subdomain
Onnl (Ts*) Q5MP00
PedF (Pf*) Q6VT95
BryC (Ces) AOM526
PksM (Bs) P40872
DifF (Ba) Q1RS54
MmpD (Pf) Q8RL72 1st
MmpD (Pf) Q8RL72 2nd
RhiD (Phr) E5ARL9
Xab8 (Xa) Q9AIT3
AlbI (Xa) Q70C47
OzmH (Sa) B2VW42
OzmK (Sa) B2VW44
OzmN (Sa) B2VW47
ChiC (Sc) A9EW96
Dsz/DisA (Sc) Q4U447
KirAII (Sc) S5VE98
KirAIV (Sc) S5V790
VirA (Sv) A4PHN0
LkcC (Sr) Q83X98
Curj (Mp) Q6DNE3
JamJ (Mp) Q6E7K0
CrpA (Nos) A4D933
HctD (Mp) Q1ED82
MpaC (Pb) F1D8A9
DtbA (An) G3Y802
CurS1 (At) L7X8J4
HmpB (Hs) B3FWS4
Rds5 (Mchh) B2FWU0
BfA (Pb) A0A068ABB7
FAS (Tc) D6W6D2
FAS (D) Q9VQL7
FAS (Bm) Q01678
FAS (Am) A0A088AP68
FAS (Dp) Q6CIR2
FAS (H) P49327
FAS (P) A5YV76
FAS (Xt) F6W5Y2
FAS (Ce) P91871

Figure S3, relates to Figure 4. Alignment of 51 CMeT domains from PKSs and FASs. The alignment reveals a high sequence conservation of active CMeT domains, whereas inactive ΨCMeT domains of FAS and some HR-PKS reveal deletions predominantly in the *N*-terminal subdomain and a low sequence conservation of active site residues. Sequence numbers, secondary structure elements, and active site residues correspond to PksCT. All sequences are labeled as protein name (organism abbreviation) Uniprot number and sorted according to their phylogenetic distance depicted in Figure 4. The sequence of PksCT corresponds to Supplemental Item 1. Multienzyme family classifications are indicated in colored groups. (°, endosymbiont of this organism; ‡, diketide synthase)

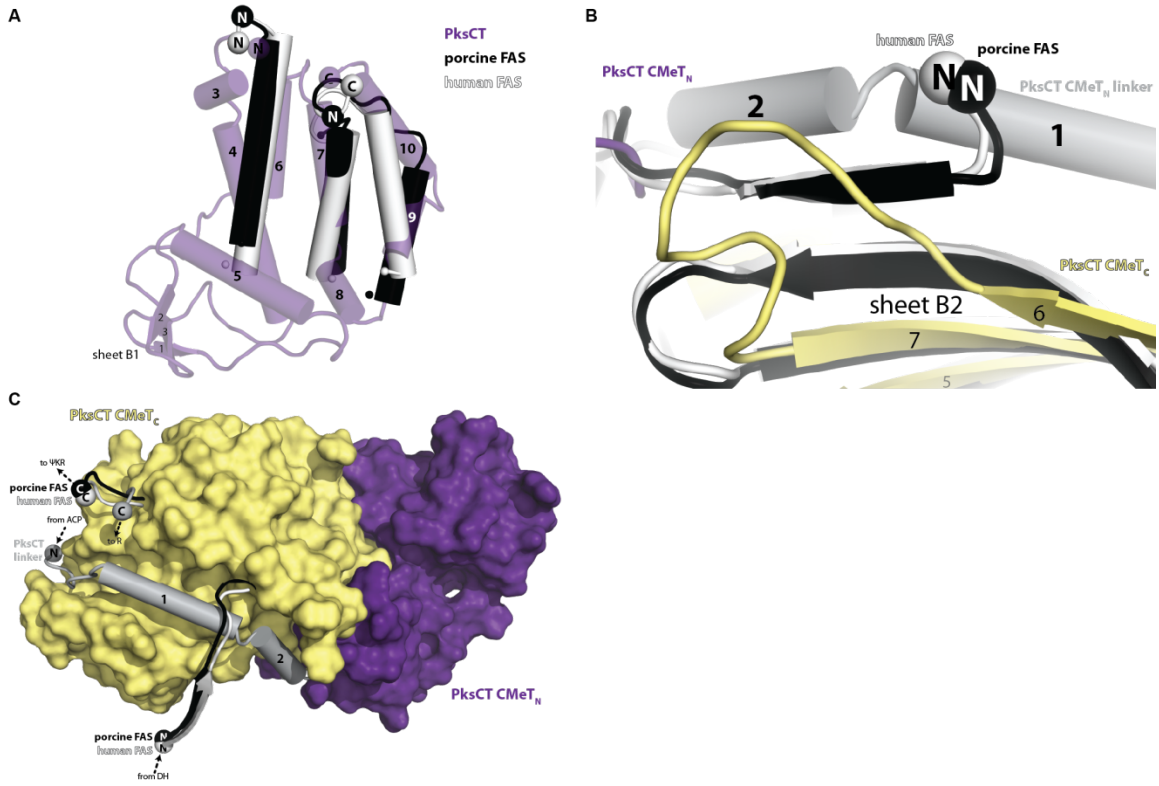


Figure S4, relates to Figure 3. Structural comparison of PksCT CMeT with mammalian FAS ΨCMeT domains.

(A) A superposition of the truncated and largely disordered *N*-terminal subdomain of the animal FAS ΨCMeT (porcine: black; human: white) with PksCT CMeT indicates a structural conservation of three helices, overlapping with PksCT CMeT helices 4, 7 and 9/10. Ends of disordered regions are indicated by small spheres. Secondary structures are assigned according to Figure 3B.
(B) In FAS, the linker connecting to the *N*-terminal subdomain contains a β-strand that extents the central β-sheet of the *C*-terminal subdomain. In PksCT CMeT, the position of this β-strand is occupied by helix 2 of the *N*-terminal linker leading into CMeT.
(C) In PksCT the *N*-terminal linker wraps in a surface groove around the *C*-terminal subdomain and positions the *N*- and *C*-terminal linker termini in close proximity to each other, whereas in FASs the *N*-terminal linker contacts this domain in a different location distant to the *C*-terminal linker terminus.

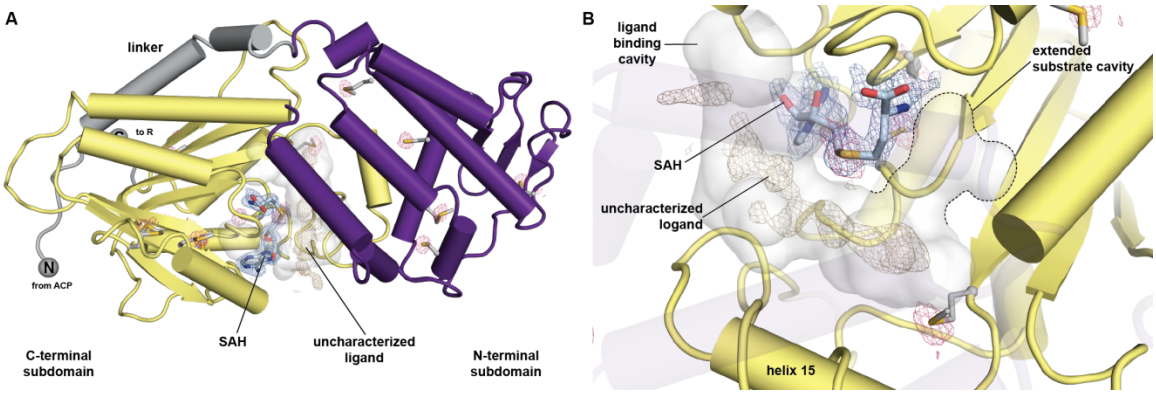


Figure S5, relates to Figure 3A and 3C. Uncharacterized ligand and anomalous difference density. (A) The active site tunnel reveals significant F_O-F_C omit difference density at 2.5 σ (brown) with a featured shape, which does not match components of the crystallization condition. Anomalous difference density at 3.7 σ (pink), indicates the presence of sulfur in methionine and cysteine residues as well as in SAH, but no significant difference density as part of the ligand.
(B) The close up view on the active site reveals the expansion of the ligand's difference density along the central tunnel with its end in the junction to the extended substrate cavity. The same contour level of the SAH difference density (blue) suggests either a partial occupancy or various conformations of the less defined ligand. The extent of difference density corresponds approximately to an atomistic binding interface of 500 $Å^2$.

Item S1, relates to Figure 1: Revised *M. purpureus* PksCT sequence. Residues encoded by alternate exon boundary are underlined:

MIDSTSHSNLRSKAFIFGPQDLSFDVRSFNKLHSQLQNHQWVLDALASLPKLWDNFAASDQKVQQ
SNTGKLLENLNAWISSGVAPEEAFPLPNVLLSPLVVIGQLVEYMTFLKAAFPDLGKKHDLPISIKED
TETFGLCTGTLCAFAVACSSNIADIQHYGAVAARLAMLVGAIVDTEEVLSDPEGKSVSFSASWNSA
EFSDSFTHVLETFPDAYVSVIVDQRRATLTASKKTAPAIIERLKQEGAHVTSIALSGRFHWKKHQD
AVSSLIQFCGLDPDLQLADATKMLLPSRSSSDGQYITTGKLHELALRAILLEQSEWYKTCRISYLSK
FIMDDAAVICFGPERCMPPTLARKLGPRLTYVSEIDISSSRVPGQLLGGTQKLNLTDLPDERIAVIG
MACRLPGAEDHEGFWEILKTGQSQHREVPEDRFGMATAWREADKRKWYGNFIDNYDTFDHKFF
KKSPREMASTDPQHRLMLQVAYQAVEQSGYFRNNGTNRRIGCFMGVGNVDYEDNIACYPANAY
SATGNLKSFLAGKISHHFGWTGPSLTLDTACSSSSVAIHQACRSILSGECNGALAGGVNVITSPNW
YHNLAGASFLSPTGQCKPFDAKGDGYCRGEGVGAVFLKRLSSAIADGDQVFGVIASTKVYQNQN
CTAITVPNAISLSELFTDVVRQARLEPKDITLVEAHGTGTAVGDPAEYDGIRAVFGGPIRSDVLSLG
SVKGLVGHTECASGVVSLIKTLLMIQQGFIPPQASFSSINPSLNAKAEEKIEISTRLKPWDAPFRAALI
NNYGASGSNASMVVTQPPNLTETPSTPLPGKSYPFWISAFDQQSLQSYVRRLRQFLEKHAADKNLS
VANLSFQVACQSNWSLPQALVFSASTKEELNRALASFEKGSTDFPSVQLPDPKPVILCFGGQVSTY
VGLDQEVYNSTAILRHYLDQCDAMCLSLGLQSIYPAIFQRSPIEDIVQLQTALFAMQYSCAKAWID
SGLKVASVVGHSFGELIALCVSNAVSLKDAVKMISGRARLIKERWGADKGSMIAVEADLSDVEAL
LAKVKSQMGSETGLAIACYNASKSFTLAGPTKDVDHAENLLKNDPDFSGIRYKRLNVTNAFHSVL
VDALIDDLESLGQGIRFKEPTIKLERATEQESTSTLNANYVATHMRKPVFFAQAVKRLSDKFPVAI
WLEAGSNSTITAMASRALGTSNSSFQAVNITSEGAFRFLCDTTVKLWKEGQKVSFWAHHRLQTPM
YTPVLLPPYQFEKSRHWMDLKVPPKPEASVQVAEQTAIIEAPKGLTTFVGYQDASQRSVRFRVNV
TTEKFNRLLSGHIMANAAAVCPGMFQVEIALDALTSLRPEFQARSFIPELHDLRHYQPLVRDESRA
VWIEAHCPNAEGLVWNWKLTASDDKGSGSVTHTSGTITFQAADSVQVKSEFEKLRRLIGRKRCLQ
LLDSNVADDILQGRNIYRAFTEVIDYKEIYRHVTKIAGRDNESAGRVIKTYDGETWLDTVLTDCFC
QVAGIFVNLMTTKIDLSERGIFICDGIDRWLRAPNAGSNNTPSQVYEVFALHHCESDSKYLSDVFAF
DAREGSLVEVALGISYQKVSISGIRRVLSKGMPAGLQPQVPTSPAAVAAIKTVSPPPVADSPLVDGS
STAVSGTPPTKKAPKAPSVDITGKMREIICNLSGLEPDEVKDDSDLVELGIDSLMSMELAREVDLAF
KTTIDVTQLIDVTDFRSLVECMQRILGIDNQEDNTYLAEGLNGHEGVVTNGNAYHVNGTNGVVN
GNGVLFPELGGSILPKSAILDAFRIAKEATDDFILNGQLGTYYNEVMPRSTELCVAHIVNAFEQLGC
PIRSAAAYQRLERVPYLPKHERFMNLIYGLLEEARLIDINGSEITRTSVPVSTKSVETMLEELLHDEP
LHAAEHKLTSLTGSKFADCITGKEDGLQLIFGSPEGREIVTDVYAKSPINAVWIQQAEFFLEQLVKR
LPNTGEPLRILEMGAGTGGTTVKMLPLLERLGVPVEYTMTDLSSSLIAAARKRFKKYPFMKFKVV
NIESPPDPQLVHSQHIILATNCVHATRNLEISTRNIHRILRPDGFLLLLEMTEQVPWVDFIFGLLEGW
WLFEDGRRHALQPATHWKKILTSVGYGHVDWTEGTRPEANIQRLIIALASEPRYDHTPQSLQPPVQ
VPLTDIAGRQEIIDTYIREYTEDFRALPIPGIQQAVMPAPTGHCVLVTGATGSLGSHVVGYLSRLPN
VHTVVCLNRRSTVPATIRQEEALKVRGISLDDNSRSKLVVLEVETAKPLLGLPVETYQKLVNTATH
IVHSAWPMSLTRPIRGYESQFKVMQNLINLAREVAAWRPVPFKFSFQFISSIGVVGYYPLRYGEIIAP
EETMTADSVLPVGYAEAKLVCERMLDETLHQYPDRFRPMAVRIAQIAGSTSNGHWNPVEHFAFLI
KSSQTLKALPDFDGSLSWCPVDDVSATLGELLISNTTPYSIYHIENPSRQQWRKMVKTLAQSLDIPR
DGIIPFDQWIERVRNSSASINDNPARQLLEFFDQHFIRMSCGNLILDTTKTREHSATLRERGPVGPGL
VEKYISAWKTMGFLD

Supplemental Procedures

UMA-guided domain dissection

UMAv2.18 was used with the BLOUSUM62 matrix. Secondary structure predictions were generated with ProteinPredict ((Yachdav et al., 2014), https://www.predictprotein.org/). Comparative sequences were curated from NCBI using the Conserved Domain Architecture Retrieval Tool (Geer et al., 2002) with PksCT as the query and included the following entries: XP_001217248, ANID_07903, ANID_00523, CAN87161, XP_001273475, XP_002381902, EHA28237, XP_001818926, XP_001393501, CAK40124, XP_660990.1, XP_658638.1, AAR90253.1, BAE66025.1, XP_001243185.1, CAK48487.1, XP_001559289.1, XP_002149737.1, XP_002567553.1, XP_002487778.1, XP_002340070.1, XP_002384396.1, XP_002848394.1, XP_003070229.1, XP_003010590.1, XP_001395291.2, EHA25844.1, EHA28237.1, and GAA92425.1.

Supplemental References

Geer, L.Y., Domrachev, M., Lipman, D.J., and Bryant, S.H. (2002). CDART: protein homology by domain architecture. Genome Res *12*, 1619-1623.

Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Honigschmid, P., Schafferhans, A., Roos, M., Bernhofer, M.*, et al.* (2014). PredictProtein-an open resource for online prediction of protein structural and functional features. Nucleic Acids Research *42*, W337-W343.