Supplementary Information

Supplementary Tables

Tables S1 lists the gold standard PPIs of the *E. coli* genome and is available to download at https://zhanglab.ccmb.med.umich.edu/Threpp/download/Table_S1.xlsx.

Tables S2 lists PPIs predicted by Threpp_threading on the *E. coli* and is available to download at https://zhanglab.ccmb.med.umich.edu/Threpp/download/Table_S2.xlsx.

Tables S3 lists PPIs detected by high-throughput methods on *E. coli* and is available to download at https://zhanglab.ccmb.med.umich.edu/Threpp/download/Table_S3.xlsx.

Tables S4 lists the final PPIs predicted by Threpp on the *E. coli* genome and is available to download at https://zhanglab.ccmb.med.umich.edu/Threpp/download/Table_S4.xlsx.

Tables S5 lists the betweenness centrality for all proteins in *E. coli* that have at least one PPI partner in the Threpp predicted networks and is available to download at https://zhanglab.ccmb.med.umich.edu/Threpp/download/Table_S5.xlsx.

Table S6. Summary of Threpp modeling parameters for the two PPI complexes shown in Figure 5.

Target ID	Template ID	Template function	Z-score[a]	S-score[b]	%ID[c]
Threading model for monomer chains					
DmsA	1EU1_A	Oxidoreductase	56		32%
DmsB	2VPZ_B	Oxidoreductase	32		34%
YagR	1RM6_A	CoA reductase (alpha-unit)	159		26%
YagS	1RM6_B	CoA reductase (beta-unit)	80		27%
YagT	3SR6_A	Oxidoreductase	59		25%
Threpp model for complex					
DmsAB	2IVF_AB	Ethylbenzene/dehydrogenase	31	52.3	33%
YagRS	1RM6_AB	Oxidoreductase	79	110.9	31%
YagRT	1FIQ_CA	Oxidoreductase	58	142.8	28%
YagST	3HRD_CD	Oxidoreductase	58	98.0	25%

[a]Z-score for monomer or Z_{com} of Threpp for dimer complex
[b]S-score is defined by Eq. (4)
[c]Sequence identity between target and template

Table S7. Summary of predicted models on 39 proteins solved after the Threpp modeling.

PDB ID of experimentally solved structures	Comparison of Threpp models to the PDB structures					
	TM-score	Template ID	S-score	Sequence identity		
				Monomer1	Monomer2	Complex
Homo-dimer						
4UHT_AB	0.45	1K66_AB	56.4	17%	17%	49%
4YZE_AC	0.81	3EUP_AB	115.4	21%	20%	49%
5AEE_AB	0.43	1WE5_AD	146.8	25%	25%	39%
5CB0_AB	0.96	4R2K_AB	104.8	94%	94%	50%
5CNX_AB	0.45	1CHM_AB	128.3	27%	27%	48%
5DUD_AC	0.46	3GI0_AB	242	15%	15%	46%
5DUD_BD	0.43	3GI0_AB	198.2	19%	19%	47%
5FSR_AB	0.54	2WUQ_AB	63.3	17%	16%	45%
5HW4_AB	0.94	3KWP_AB	118.1	40%	40%	49%
5IMJ_AB	0.91	2OEZ_AB	139.2	41%	42%	48%
5J43_AE	0.96	1V7C_AB	159.5	28%	28%	50%
5NJ9_BD	0.49	1VPB_AB	144	27%	27%	50%
5TPM_AB	0.71	2DI3_AB	111.2	23%	22%	45%
5VM2_AB	0.48	2DQ4_AB	106.6	27%	27%	49%
5WQL_AB	0.94	1XNF_AB	76.12	87%	88%	47%
5XU7_AB	0.5	2WDO_AB	62.3	39%	39%	33%
5Z03_AB	0.88	3T5M_AB	146.2	28%	29%	48%
5Z1Z_AB	0.87	2GO1_AB	158.5	26%	26%	50%
5ZE6_AB	0.48	4JYX_BD	140.4	60%	58%	50%
5ZXL_AB	0.96	1TA9_AB	118.7	43%	44%	49%
6AGL_AB	0.49	1KFL_AB	194.1	52%	52%	50%
6BPM_CA	0.67	4AIP_AC	75.9	20%	20%	45%
6E4B_AB	0.81	3LGA_AB	80.3	29%	29%	48%
6EI9_AB	0.41	1GWJ_AB	70.6	24%	24%	47%
6GAM_LM	0.95	4UE3_LM	124	43%	43%	49%
6GAM_ST	0.97	4GD3_ST	122.3	33%	33%	43%
6GFL_AB	0.49	3BQ9_AB	181.5	61%	61%	50%
6MTG_AB	0.88	2H2W_AB	113.7	48%	48%	52%
6MX1_AB	0.64	3C1N_AB	214.7	21%	21%	49%
6OFU_AB	0.95	1GVF_AC	136.9	37%	37%	50%
6OHB_AC	0.91	3LS9_AB	124.8	26%	26%	48%
Average	0.71±0.22		128.2±44.8	(35±19)%	(35±19)%	(47±4)%
Hetero-dimer						
5DUD_AB	0.55	3GI0_AB	251.5	15%	19%	54%
5G5G_AB	0.9	3HRD_CD	90	25%	25%	25%
5G5G_AC	0.94	1FIQ_CA	142.8	27%	30%	28%
5G5G_BC	1	1RM6_AB	110.9	30%	32%	31%

5JFF_AB	0.74	3ZGY_BA	42.3	29%	33%	77%
5NJ9_AB	0.75	1VPB_AB	139.2	22%	27%	47%
5T1O_AB	0.6	2XDF_CA	51.6	28%	27%	71%
6GAM_SL	0.96	4UE3_SL	243.9	30%	43%	50%
Average	0.81±0.16		134.0±73.9	(26±5)%	(30±7)%	(48±18)%
Average of all 39 targets	0.73±0.21		129.4±52.2	(33±18)%	(34±18)%	(48±9)%

Table S8. PPIs (given by Figure 4) described in IntAct database.

Gene name	Gene name	IntAct	Gene name	Gene name	IntAct
yqaB	dnaK	1	dnaK	murC	1
crp	rcsB	1	dnaK	gapA	1
ndk	dnaK	1	dnaK	ftsA	1
pgk	dnaK	1	dnaK	ftsZ	1
alaS	dnaK	1	dnaK	eno	1
thrS	dnaK	1	dnaK	gadA	1
mlaF	dnaK	1	dnaK	rne	1
yegD	dnaK		dnaK	uspG	1
tyrS	dnaK	1	dnaK	fis	1
holA	dnaK	1	dnaK	ravA	1
mog	dnaK	1	dnaK	rplF	1
yjjI	dnaK	1	dnaK	pepB	1
hfq	dnaK	1	dnaK	yihQ	
mutL	dnaK	1	dnaK	helD	1
ydcR	dnaK	1	dnaK	accD	1
accC	dnaK	1	dnaK	yidA	1
rcsD	rcsB	1	dnaK	ddlA	1
rcsB	yagA	1	dnaK	rsmD	1
rcsB	cusR		dnaK	ygcP	1
rcsB	ssuD	1	dnaK	nikD	1
rcsB	rplM	1	dnaK	srlR	1
rcsB	prmA	1	dnaK	mreB	1
rcsB	rlmN	1	dnaK	rplD	1
rcsB	rplI	1	dnaK	tnaA	1
rcsB	rpsC	1	dnaK	ppa	1
rcsB	rplV	1	dnaK	holC	1
rcsB	arcA		dnaK	fabF	1
rcsB	bglJ	1	dnaK	aspA	1
rcsB	groS		dnaK	rhlB	1
rcsB	basR		dnaK	trmB	1
rcsB	yjjQ		dnaK	mutY	1
rcsB	tilS	1	dnaK	gatY	1
rcsB	polA	1	dnaK	fbaB	1
rcsB	rpsU	1	dnaK	rpe	1
rcsB	yhjB		dnaK	rpoH	1
rcsB	torR		dnaK	nudB	1
rcsB	narP		dnaK	aidB	1
rcsB	qseB		dnaK	ruvB	1
rcsB	csgD		dnaK	ssb	1
rcsB	kdpE		dnaK	cbpA	1

rcsB	bcsQ	1	dnaK	kdsB	1
rcsB	rpsJ	1	dnaK	pyrH	1
rcsB	rplS	1	dnaK	ygdG	1
rcsB	yhjC		dnaK	ispA	1
rcsB	mgtA	1	dnaK	rng	1
rcsB	rplE	1	dnaK	iscS	1
rcsB	phoB		dnaK	proS	1
rcsB	ompR		dnaK	hupA	1
rcsB	dnaK	1	dnaK	clpX	1
rcsB	baeR		dnaK	entB	1
rcsB	rstA		dnaK	recJ	1
rcsB	rpmB	1	dnaK	dhaM	1
rcsB	rcsA	1	dnaK	serC	1
rcsB	ygfK	1	dnaK	hisC	1
rcsB	phoP		dnaK	pykA	1
rcsB	rplC	1	dnaK	clpP	1
rcsB	cspC	1	dnaK	metK	1
rcsB	ygeK		dnaK	prs	1
rcsB	yedW		dnaK	yceA	1
rcsB	cpxR		dnaK	capR	1
rcsB	creB		dnaK	dnaE	1
rcsB	yfaY	1	dnaK	accA	1
rcsB	rhmD	1	dnaK	ulaR	1
rcsB	rpsD	1	dnaK	gloA	1
rcsB	gfcD	1	dnaK	tufA	1
rcsB	evgA		dnaK	sapA	1
rcsB	etk	1	dnaK	potH	
rcsB	narL		dnaK	prmB	1
rcsB	fimZ		dnaK	hinT	1
rcsB	lldP	1	dnaK	yfjK	1
rcsB	dctR		dnaK	nrdB	1
rcsB	uhpA		dnaK	holE	1
rcsB	glnB	1	dnaK	ydhR	1
rcsB	rpoA	1	dnaK	luxS	1
rcsB	uvrY		dnaK	greA	1
rcsB	citB		dnaK	sspA	1
fbaA	dnaK	1	dnaK	iscU	1
yfiF	ftsA	1	dnaK	panC	1
rplK	ftsA	1	dnaK	rpmG	1
clpA	dnaK	1	dnaK	dcuD	1
nusG	ftsA	1	dnaK	groS	1
rplS	dnaK	1	dnaK	groL	1
hldD	dnaK	1	dnaK	hslU	1

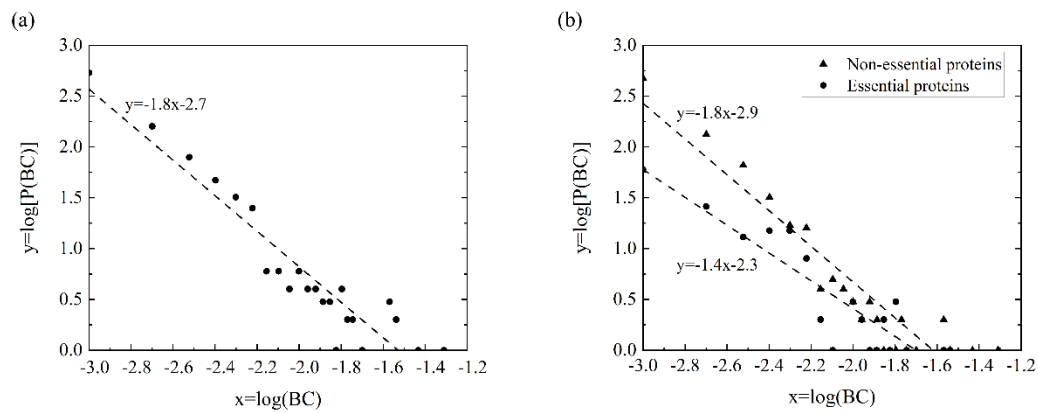|					
dnaK	yibA	1	dnaK	pepD	1
dnaK	dnaJ	1	dnaK	yghA	1
dnaK	hscC		dnaK	fabB	1
dnaK	tsaE	1	dnaK	rpoZ	1
dnaK	rplJ	1	dnaK	purC	1
dnaK	ydfJ	1	dnaK	queC	1
dnaK	rsgA	1	dnaK	yiiD	1
dnaK	trxC	1	dnaK	rph	1
dnaK	tdcG	1	dnaK	yicC	1
dnaK	eutJ		dnaK	tpiA	1
dnaK	rcsA	1	dnaJ	ftsA	1
dnaK	gyrA	1	rplL	ftsA	1
dnaK	grpE	1	tdcE	ftsA	
dnaK	qorA	1	eutJ	ftsA	
dnaK	rplI	1	rplT	ftsA	1
dnaK	aphA	1	rpsB	ftsA	1
dnaK	gadB	1	ftsA	cutA	1
dnaK	htpG	1	ftsA	yhhA	1
dnaK	miaB	1	ftsA	rplU	1
dnaK	sspB	1	ftsA	accD	1
dnaK	rpsB	1	ftsA	mreB	1
dnaK	infB	1	ftsA	slyD	1
dnaK	secA	1	ftsA	rho	1
dnaK	hscA	1	ftsA	aceE	1
dnaK	ropN		ftsA	tufA	1
dnaK	narQ	1			

Supplementary Figures

Figure S1. Betweenness centrality (BC) distributions of the *E. coli* genome. (a) Distribution of PPI node BC that is defined as the normalized number of shortest paths passing through the considered node in the network. (b) PPI BC distributions for essential (circles) and non-essential (triangles) proteins, respectively. Lines in (a) and (b) are power law fit to $P(BC) \propto BC^{-\gamma}$.
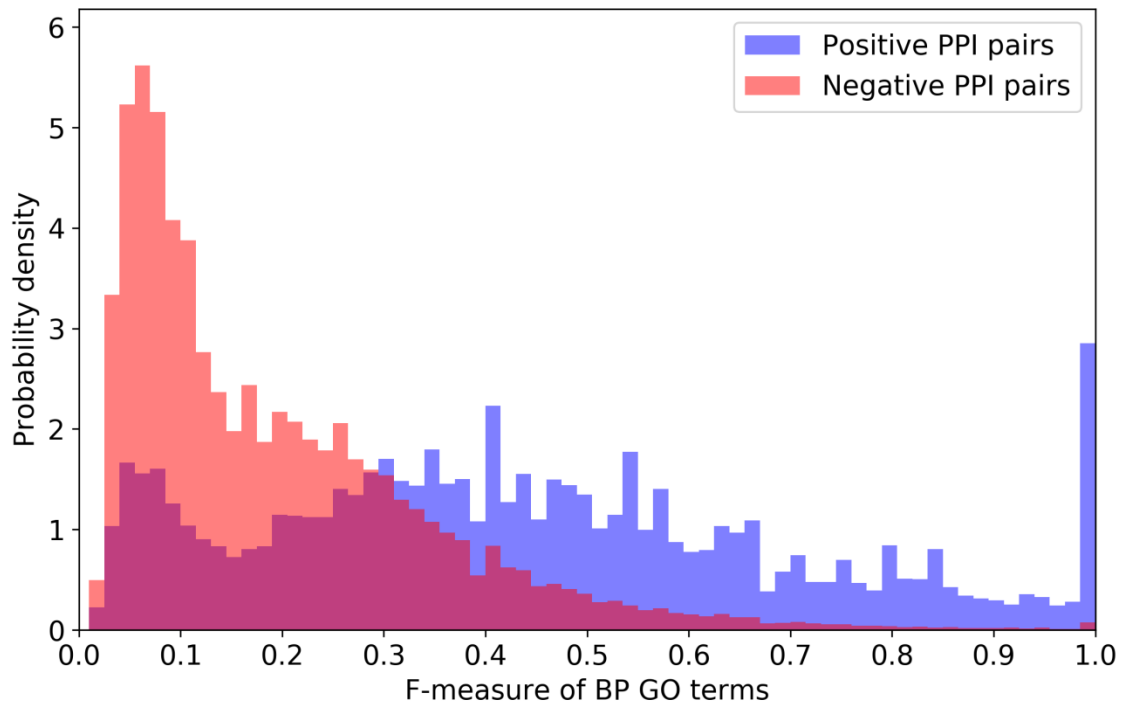
Figure S2. Distribution of Biological Process Gene Ontology (BP GO) annotation similarity among protein pairs predicted to have PPI (blue) and among predicted non-interacting pairs (red).
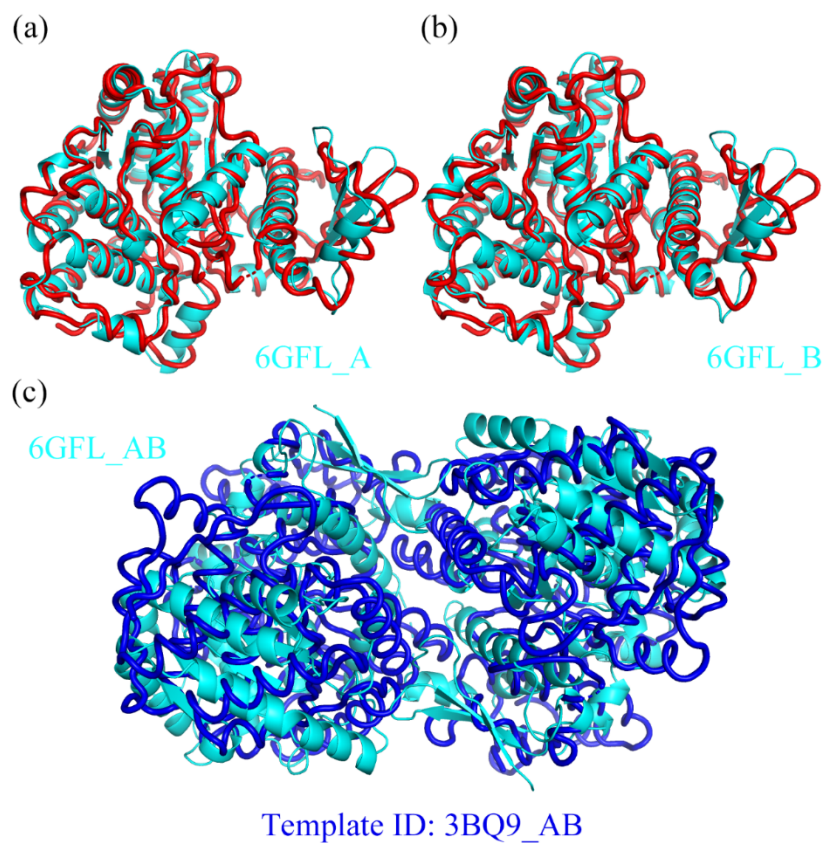
Figure S3. Illustrative example of complex structure modelling by Threpp for protein complex 6GFL_AB. (a) and (b) Superpositions of modeled monomers on the corresponding parts of the X-ray complex structure (6GFL_AB). (c) Superposition of the complex template (3BQ9_AB) on the X-ray complex structure (6GFL_AB).