STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 71 (2015)

Supporting information for article:

Expression, purification, crystallization and X-ray diffraction analysis of ChiL, a chitinase from *Chitiniphilus shinanonensis*

Miruku Ueda, Makoto Shimosaka and Ryoichi Arai

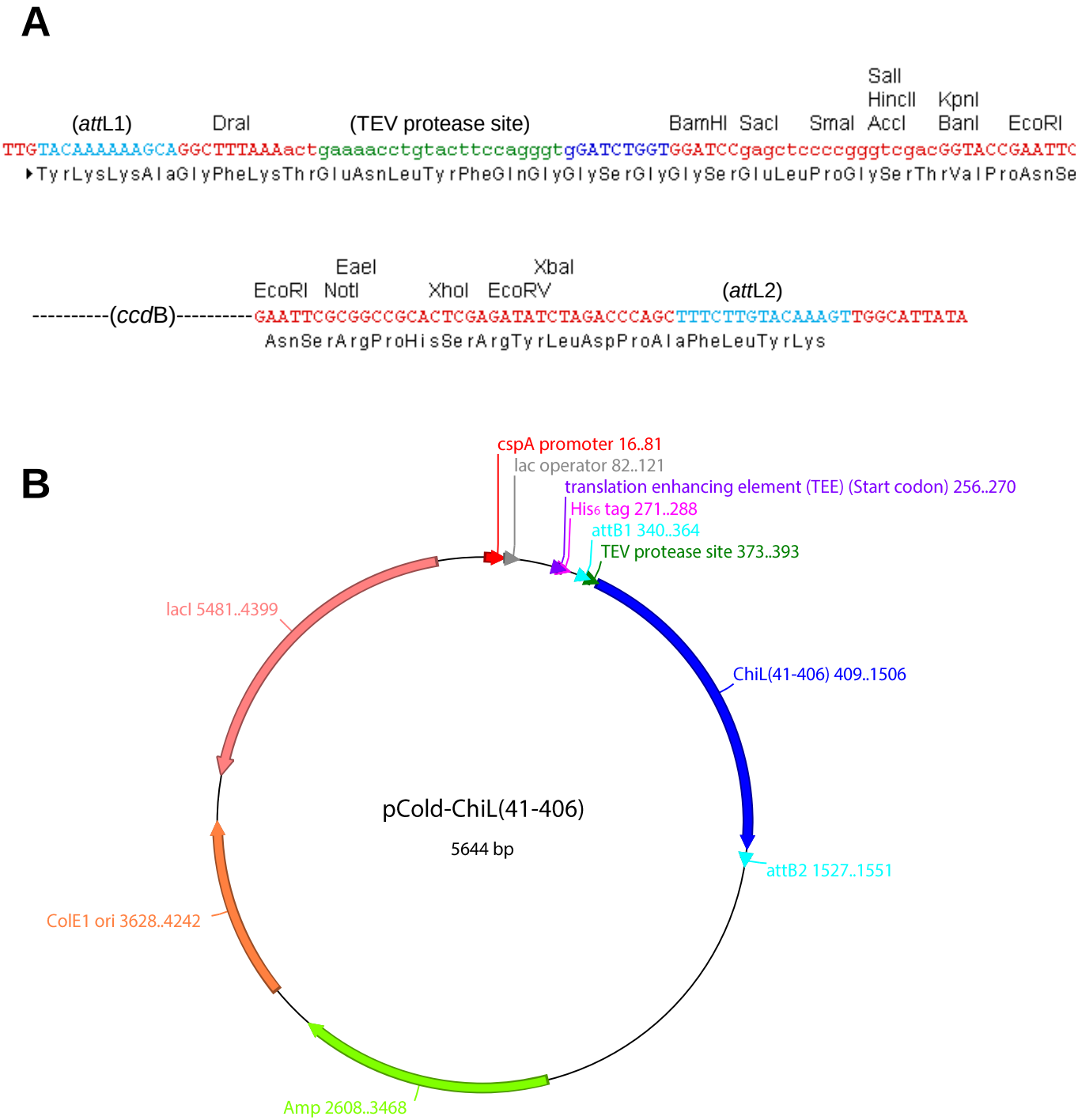


Figure S1 (A) Multi-cloning site of pENTR-TEV-*ccdB*, a modified Gateway entry vector with a tobacco etch virus (TEV) protease cleavage site (ENLYFQ|G), derived from pENTR1A (Invitrogen, Thermo Fisher Scientific). (B) Map of the protein expression plasmid vector, pCold-ChiL(41-406), which was constructed using LR reaction of Gateway technology (Invitrogen) with pENTR-TEV-L-ChiL(41-406) and pCold I-DEST, a modified pCold I DNA (Takara Bio) expression vector bearing a cold shock protein *cspA* promoter, an N-terminal His_6-tag, and a Gateway reading frame cassette (Invitrogen). The vector map was created using *ApE* (A plasmid Editor by M. Wayne Davis) (http://biologylabs.utah.edu/jorgensen/wayned/ape/).

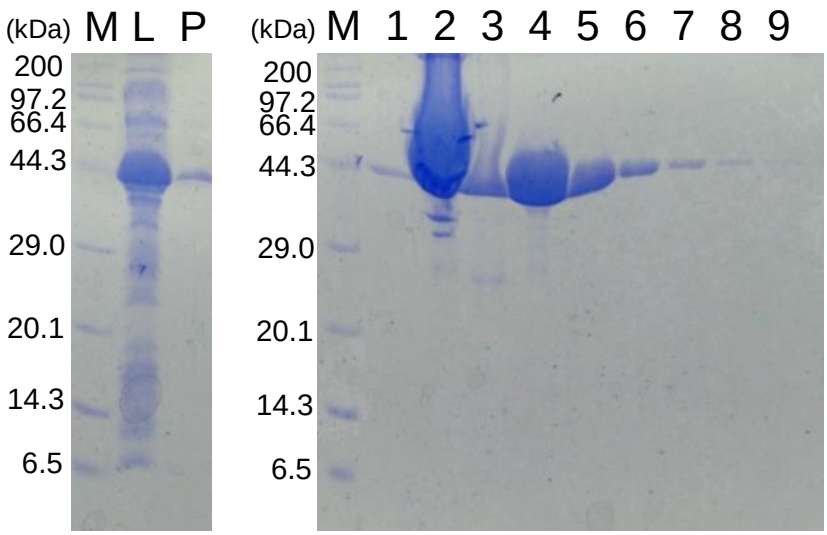


Figure S2 SDS−PAGE analysis of ChiL(41-406) purification by immobilized-metal affinity chromatography (IMAC). Lanes M, molecular mass standards (protein molecular weight marker (broad), Takara Bio); lane L, lysate of cells; lane P, pellet of cell extract, lanes 1–9, eluted fractions of ChiL(41-406) by IMAC. Proteins were detected by staining with Coomassie Brilliant Blue (CBB).

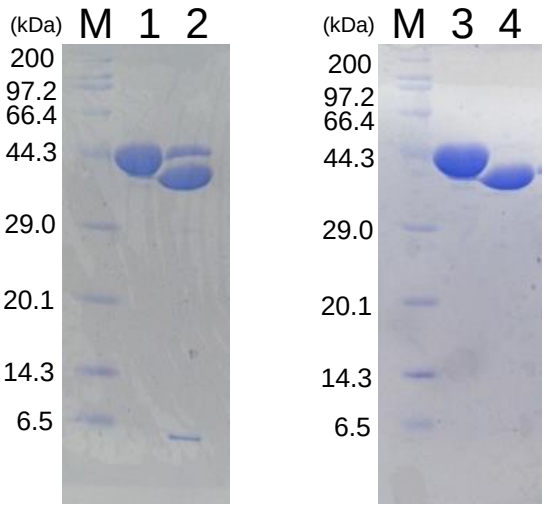
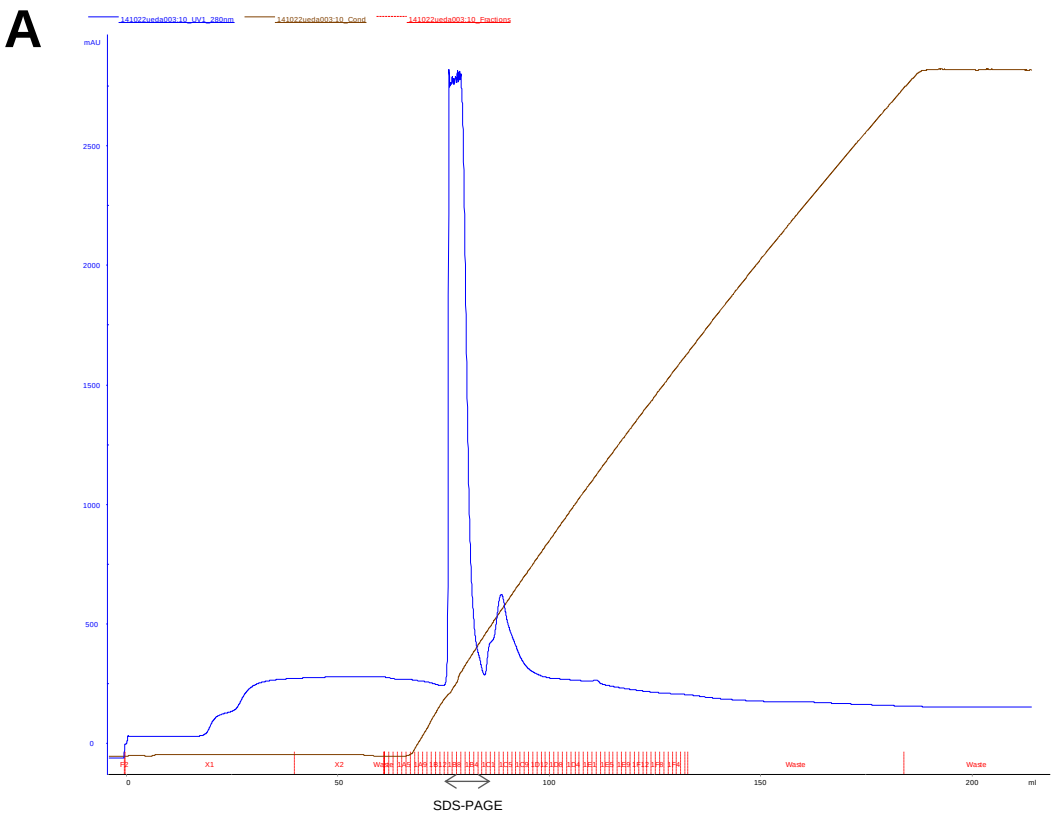


Figure S3 SDS–PAGE analysis of ChiL(41-406) purification process: cleavage of a His_6 tag by TEV protease. Lanes M, molecular mass standards; lanes 1 and 3, the sample before TEV protease treatment; lane 2, the sample after TEV protease treatment; lane 4, the TEV protease-treated sample after removal of the His_6 tag by IMAC. Proteins were detected by staining with CBB.



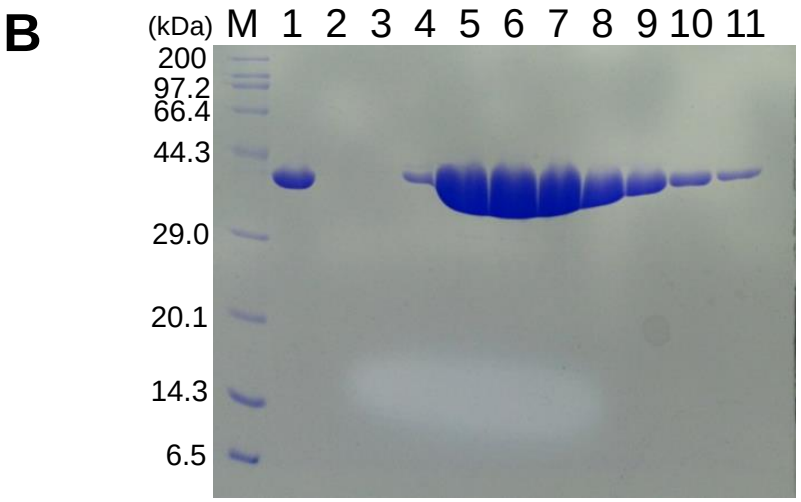


Figure S4 ChiL(41-406) purification by anion exchange chromatography (AEC). (A) Chromatogram of ChiL(41-406) purification by AEC using an AKTAexplorer HPLC system with a RESOURCE Q (6 ml) column (GE healthcare). (B) SDS–PAGE analysis of ChiL(41-406) purification by AEC. Lanes M, molecular mass standards; lane 1, ChiL(41-406) protein sample before AEC; lanes 2–11, eluted peak fractions of ChiL(41-406) by AEC. Proteins were detected by staining with CBB.

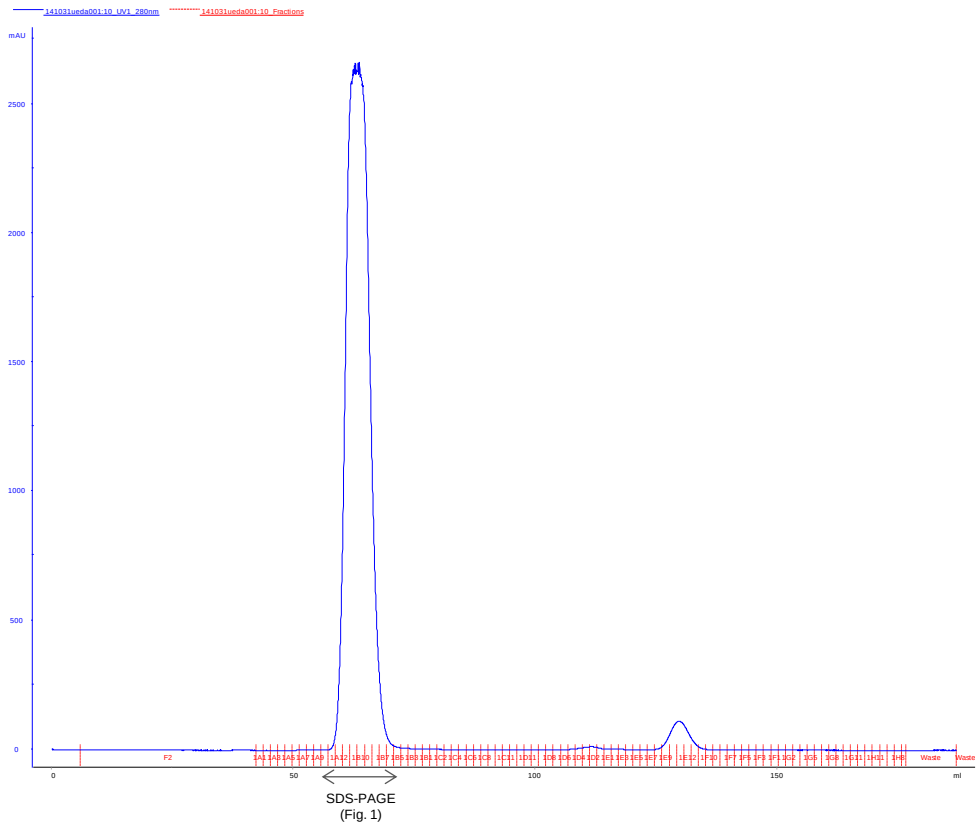


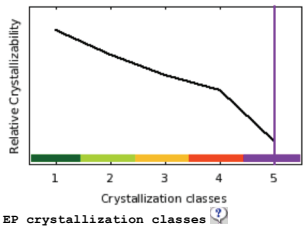
Figure S5 Chromatogram of ChiL(41-406) purification by size exclusion chromatography (SEC) using an AKTAexplorer 10S HPLC system with a HiLoad 16/600 Superdex 75 pg column (GE Healthcare). Proteins were detected by staining with CBB.

XtalPred-RF

Target: ChiL

JCMM | JCSG
Godzik Lab
SBMRI

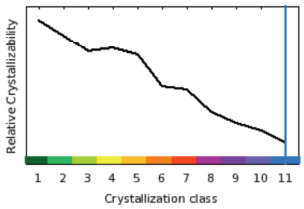
Expert pool crystallizability class: 5



EP crystallization classes

↓Most promising
1 2 3 4 5
↑Least promising

Random forest crystallizability class:11



Homologs in microbial genomes

EP Class	1	2	3	4	5	total
Homologs	4	6	15	11	99	135

RF Class	1	2	3	4	5	6	7	8	9	10	11	total
Homologs	1	3	10	12	12	3	8	5	32	7	42	135

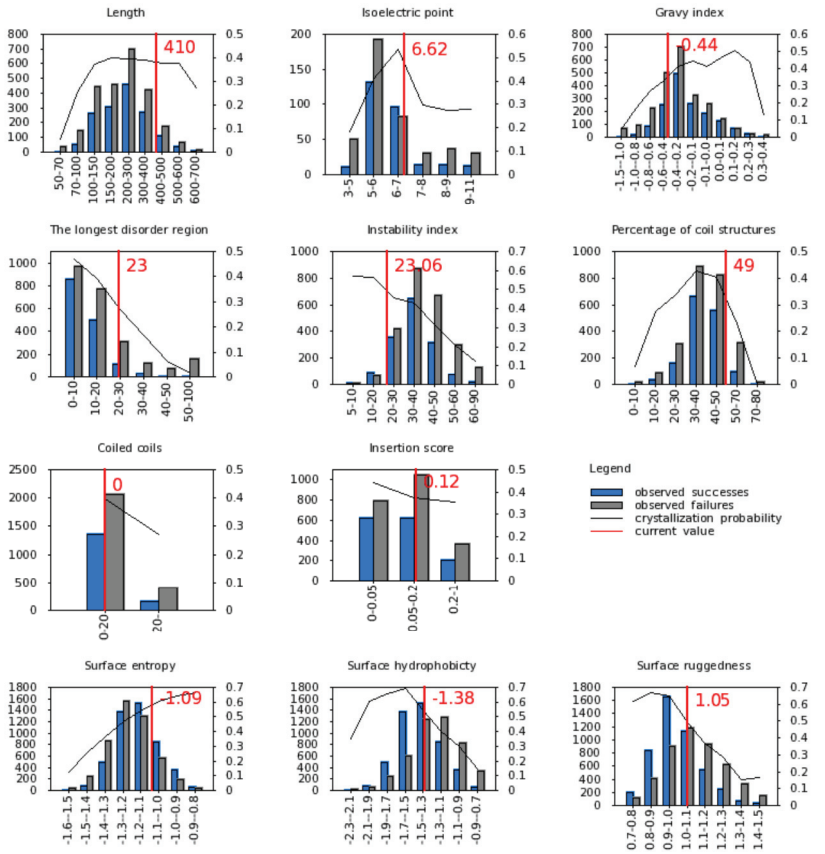
RF crystallization classes

↓Most promising
1 2 3 4 5 6 7 8 9 10 11
Least promising↑

Combined class (score): 5 (range 1..5)

Homologs (by PSI-BLAST)	
All known proteins (NR)	2046
Non-redundant NR database (NR60)	825
Solved strucutres (PDB)	168

Comparison of target features with distributions of crystallization probabilities obtained from TargetDB



Protein features	
Length	410
Molecular weight	45310
Gravy index	-0.44
Isoelectic point	6.62
Instability index	23.06
Predictions	
Transmembrane helices (number)	No
Signal peptides (length)	21
Longest disorder reg.	23
Longest low complexity reg.	36
Coiled coils	0
% disorder residues	7
% coil residues	49
% helix residues	32
% strand residues	18
Predicted surface features	
Surface entropy	-1.09
Surface hydrophobicity	-1.38
Surface rugedness	1.05
Other	
Number of Cys residues	2
Number of Met residues	5
Number of Trp residues	10
Number of Tyr residues	24
Number of Phe residues	17
Epsilon 280	90760
Insertions score	0.12

```
1...*...10....*...20....*...30....*...40....*...50....*...60....*...70....*...80....*...90....*..100
MRQHLLALLLAAALSFLAAQAADTPAAKPAAVAVPEASKPYRVVAYYISWGAYGRSYFPSDIDYSKVTHINYAFANIKDGEVVVGDPGVDDGGKNNFTAL

....*..110....*..120....*..130....*..140....*..150....*..160....*..170....*..180....*..190....*..200
RKAKKAHPHLRNLISVGGWSWSSGFSDAAATPEARKRFADSAVAFIRKYGFDGVDIDWEYPVEGGAENMKHRPEDKQNYTLLTRSLREALDTAGKADGKY

....*..210....*..220....*..230....*..240....*..250....*..260....*..270....*..280....*..290....*..300
YELTTAVWGNDKFIANTEMDKVSRDFDFINVMSYDFNGTWNKFSGHNAPFVNDPAYDKPGIGKTFNVVSAVEAYLKAGVPADKLVVGVPLYGYSWKGCAA

....*..310....*..320....*..330....*..340....*..350....*..360....*..370....*..380....*..390....*..400
GERNGEYQDCNGKGRGTWEDGNLDFTDIEKNLLNKKGFKRYWNDTAKAAYLYNAETGEFVTYEDPQALKIKLDYIKSKGLGGAMYWEITADRKQTLVNLI

....*..410
ADELLTKSVK
```

Legend

LOOP	loop secondary structure predicted by PSIPRED
HELIX	helix secondary structure predicted by PSIPRED
STRAND	strand secondary structure predicted by PSIPRED
DISORDER	disordered region predicted by DISOPRED2
LOW COMPLEXITY	low complexity region predicted by SEG
COILS	coiled coils region predicted by COILS
TRANSMEMBRANE HELICES	transmembrane helices predicted by TMHMM
SIGNAL PEPTIDES	signal peptites predicted by RPSP

Download

FASTA sequence	text file with protein sequence in FASTA format
Summary	text file with a summary of protein features

This server is supported by the NIH Protein Structure Initiative grant U54 GM074898 (JCSG)
and NIGMS R01 grant GM095847 'Engineering of Proteins for Crystallography'

Last update May 12, 2013

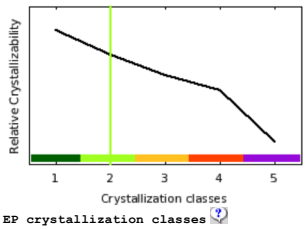
Figure S6 Crystallizability prediction results of the full-length ChiL(1-410) by *XtalPred* web server (Slabinski *et al.*, 2007).

XtalPred-RF

Target: ChiL_41-406

JCMM | JCSG
Godzik Lab
SBMRI

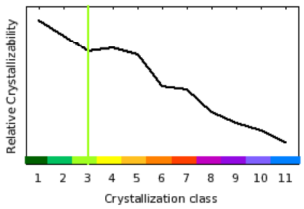
Expert pool crystallizability class: 2



EP crystallization classes

↓Most promising
1 2 3 4 5
↑Least promising

Random forest crystallizability class:3



Homologs in microbial genomes

EP Class	1	2	3	4	5	total
Homologs	4	6	15	11	101	137

RF Class	1	2	3	4	5	6	7	8	9	10	11	total
Homologs	1	3	10	12	12	2	8	5	32	9	43	137

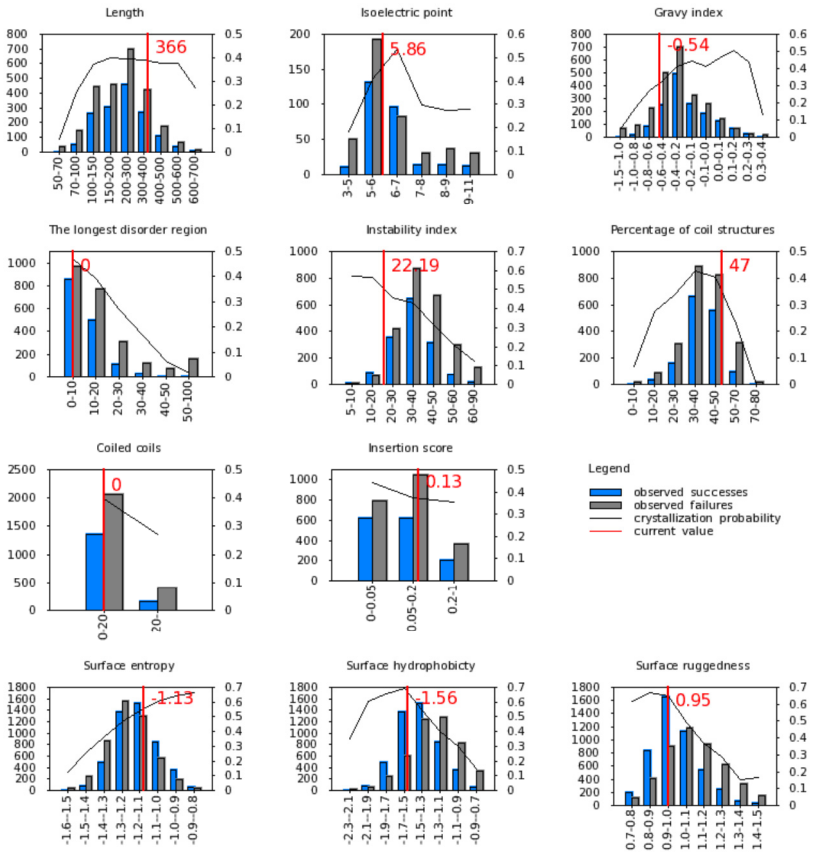
RF crystallization classes

↓Most promising
1 2 3 4 5 6 7 8 9 10 11
Least promising↑

Combined class (score): 2 (range 1..5)

Homologs (by PSI-BLAST)	
All known proteins (NR)	2048
Non-redundant NR database (NR60)	823
Solved strucutres (PDB)	166

Comparison of target features with distributions of crystallization probabilities obtained from TargetDB



Protein features	
Length	366
Molecular weight	40939
Gravy index	-0.54
Isoelectic point	5.86
Instability index	22.19
Predictions	
Transmembrane helices (number)	No
Signal peptides (length)	No
Longest disorder reg.	0
Longest low complexity reg.	15
Coiled coils	0
% disorder residues	0
% coil residues	47
% helix residues	32
% strand residues	20
Predicted surface features	
Surface entropy	-1.13
Surface hydrophobicity	-1.56
Surface rugedness	0.95
Other	
Number of Cys residues	2
Number of Met residues	4
Number of Trp residues	10
Number of Tyr residues	24
Number of Phe residues	17
Epsilon 280	90760
Insertions score	0.13

```
1...*...10....*...20....*...30....*...40....*...50....*...60....*...70....*...80....*...90....*..100
YRVVAYYISWGAYGRSYFPSDIDYSKVTHINYAFANIKDGEVVVGDPGVDDGGKNNFTALRKAKKAHPHLRNLISVGGWSWSSGFSDAAATPEARKRFAD

....*...110....*..120....*..130....*..140....*..150....*..160....*..170....*..180....*..190....*..200
SAVAFIRKYGFDGVDIDWEYPVEGGAENMKHRPEDKQNYTLLTRSLREALDTAGKADGKYYELTTAVWGNDKFIANTEMDKVSRDFDFINVMSYDFNGTW

....*..210....*..220....*..230....*..240....*..250....*..260....*..270....*..280....*..290....*..300
NKFSGHNAPFVNDPAYDKPGIGKTFNVVSAVEAYLKAGVPADKLVVGVPLYGYSWKGCAAGERNGEYQDCNGKGRGTWEDGNLDFTDIEKNLLNKKGFKR

....*..310....*..320....*..330....*..340....*..350....*..360....*.
YWNDTAKAAYLYNAETGEFVTYEDPQALKIKLDYIKSKGLGGAMYWEITADRKQTLVNLIADELLT
```

Legend

LOOP	loop secondary structure predicted by PSIPRED
HELIX	helix secondary structure predicted by PSIPRED
STRAND	strand secondary structure predicted by PSIPRED
DISORDER	disordered region predicted by DISOPRED2
LOW COMPLEXITY	low complexity region predicted by SEG
COILS	coiled coils region predicted by COILS
TRANSMEMBRANE HELICES	transmembrane helices predicted by TMHMM
SIGNAL PEPTIDES	signal peptites predicted by RPSP

Download

FASTA sequence	text file with protein sequence in FASTA format
Summary	text file with a summary of protein features

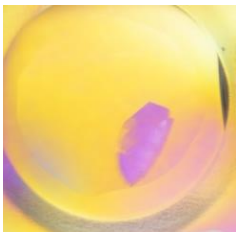
This server is supported by the NIH Protein Structure Initiative grant U54 GM074898 (JCSG)
and NIGMS R01 grant GM095847 'Engineering of Proteins for Crystallography'

Last update May 12, 2013

Figure S7 Crystallizability prediction results of the truncated ChiL(41-406) (i.e. the catalytic domain) by *XtalPred* web server (Slabinski *et al.*, 2007).

Index

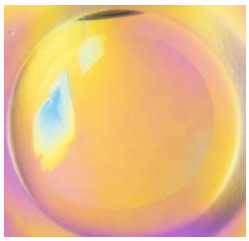
D7 (25% PEG 3350, 0.1 M BIS-TRIS pH 6.5)



D8 (25% PEG 3350, 0.1 M HEPES pH 7.5)

F2 (20% PEG 2000, 0.1 M Tris pH 8.5, 0.2M Trimethylamine N-oxide)

F6 (25% PEG 3350, 0.1 M BIS-TRIS pH 5.5, 0.2M Ammonium sulfate)



F9 (25% PEG 3350, 0.1 M Tris pH 8.5, 0.2M Ammonium sulfate)

F11 (25% PEG 3350, 0.1 M BIS-TRIS pH 6.5, 0.2M Sodium chloride)

G4 (25% PEG 3350, 0.1 M HEPES pH 7.5, 0.2M Lithium sulfate)

H3 (20% PEG 3350, 0.2 M Sodium malonate pH 7.0)

PEGRx

G4 (20% PEG 4000, 0.1 M Sodium citrate pH 5.5, 18% 2-propanol)

G9 (20% PEG 5000, 0.1 M BIS-TRIS pH 9.0, 2-propanol)

Wizard

A6 (20% PEG 3000, 0.1 M Sodium citrate pH 5.5)

G4 (20% PEG 8000, 0.1 M MES pH 6.0, 0.2 M Calcium acetate)

Figure S8 Crystals obtained by the sitting drop vapor diffusion method with 96-well protein crystallization plates using crystallization screens, Index HT (Hampton Research), PEGRx HT (Hampton Research), and Wizard Classic 1 and 2 block (Rigaku Reagents).