Full wwPDB X-ray Structure Validation Report ⓘ

Jun 16, 2024 – 08:33 PM EDT

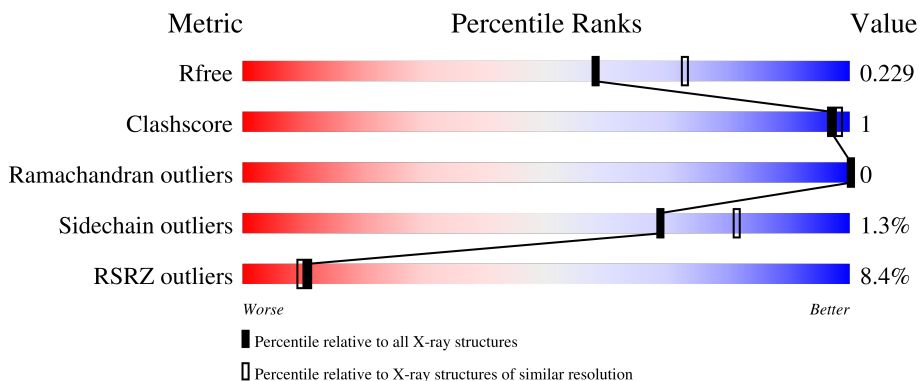|		
PDB ID	:	5NUS
Title	:	Structure of a minimal complex between p44 and p34 from Chaetomium thermophilum
Authors	:	Koelmel, W.; Schoenwetter, E.; Kuper, J.; Schmitt, D.R.; Kisker, C.
Deposited on	:	2017-05-02
Resolution	:	2.20 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.229
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		1.3%
RSRZ outliers		8.4%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	303	4% 68% 30%
2	B	113	11% 62% 35%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4457 atoms, of which 2228 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called p34.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	211	3348	1052	1704	291	298	3	0	7	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-25	MET	-	initiating methionine	UNP G0RXV8
A	-24	LYS	-	expression tag	UNP G0RXV8
A	-23	HIS	-	expression tag	UNP G0RXV8
A	-22	HIS	-	expression tag	UNP G0RXV8
A	-21	HIS	-	expression tag	UNP G0RXV8
A	-20	HIS	-	expression tag	UNP G0RXV8
A	-19	HIS	-	expression tag	UNP G0RXV8
A	-18	HIS	-	expression tag	UNP G0RXV8
A	-17	PRO	-	expression tag	UNP G0RXV8
A	-16	MET	-	expression tag	UNP G0RXV8
A	-15	SER	-	expression tag	UNP G0RXV8
A	-14	ASP	-	expression tag	UNP G0RXV8
A	-13	TYR	-	expression tag	UNP G0RXV8
A	-12	ASP	-	expression tag	UNP G0RXV8
A	-11	ILE	-	expression tag	UNP G0RXV8
A	-10	PRO	-	expression tag	UNP G0RXV8
A	-9	THR	-	expression tag	UNP G0RXV8
A	-8	THR	-	expression tag	UNP G0RXV8
A	-7	GLU	-	expression tag	UNP G0RXV8
A	-6	ASN	-	expression tag	UNP G0RXV8
A	-5	LEU	-	expression tag	UNP G0RXV8
A	-4	TYR	-	expression tag	UNP G0RXV8
A	-3	PHE	-	expression tag	UNP G0RXV8
A	-2	GLN	-	expression tag	UNP G0RXV8
A	-1	GLY	-	expression tag	UNP G0RXV8
A	0	ALA	-	expression tag	UNP G0RXV8

- Molecule 2 is a protein called p44.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	74	1089	355	524	101	100	9	0	0	0

There are 59 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	LEU	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	GLY	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6
B	?	-	GLU	deletion	UNP G0RZE6
B	?	-	LYS	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	GLY	deletion	UNP G0RZE6
B	?	-	LYS	deletion	UNP G0RZE6
B	?	-	GLU	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	LYS	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	GLY	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	GLU	deletion	UNP G0RZE6
B	?	-	ARG	deletion	UNP G0RZE6
B	?	-	GLN	deletion	UNP G0RZE6
B	?	-	GLY	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6
B	?	-	ASN	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASN	deletion	UNP G0RZE6
B	?	-	ASN	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	?	-	LYS	deletion	UNP G0RZE6
B	?	-	LYS	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	GLY	deletion	UNP G0RZE6
B	?	-	ILE	deletion	UNP G0RZE6
B	?	-	SER	deletion	UNP G0RZE6
B	?	-	LEU	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	?	-	THR	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	?	-	LEU	deletion	UNP G0RZE6
B	?	-	PRO	deletion	UNP G0RZE6
B	?	-	GLU	deletion	UNP G0RZE6
B	?	-	ALA	deletion	UNP G0RZE6
B	410	SER	ARG	conflict	UNP G0RZE6
B	411	ASN	ALA	conflict	UNP G0RZE6
B	412	GLY	VAL	conflict	UNP G0RZE6
B	413	ASN	GLY	conflict	UNP G0RZE6
B	414	GLY	VAL	conflict	UNP G0RZE6

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	Zn 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	14	Total 14	O 14	0	0
4	B	4	Total 4	O 4	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: p34

Chain A:

• Molecule 2: p44

Chain B:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	147.06Å 147.06Å 87.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.42 – 2.20 42.19 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.7 (41.42-2.20) 99.8 (42.19-2.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.30 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.210 , 0.227 0.213 , 0.229	Depositor DCC
R_{free} test set	1423 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	48.8	Xtriage
Anisotropy	0.359	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 57.9	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4457	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality (i)

5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.23	0/1696	0.41	0/2312
2	B	0.24	0/583	0.37	0/788
All	All	0.23	0/2279	0.40	0/3100

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1644	1704	1704	3	0
2	B	565	524	523	1	0
3	B	2	0	0	0	0
4	A	14	0	0	0	0
4	B	4	0	0	0	0
All	All	2229	2228	2227	4	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (4) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:241:GLY:O	1:A:259[B]:ARG:NH2	2.39	0.54
2:B:380:PHE:N	2:B:381:PRO:CD	2.76	0.48
1:A:24:LEU:HD13	1:A:67:GLN:HG3	2.02	0.42
1:A:241:GLY:O	1:A:259[B]:ARG:NH1	2.52	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	212/303 (70%)	208 (98%)	4 (2%)	0	100	100
2	B	72/113 (64%)	67 (93%)	5 (7%)	0	100	100
All	All	284/416 (68%)	275 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	176/235 (75%)	174 (99%)	2 (1%)	73	85
2	B	61/93 (66%)	59 (97%)	2 (3%)	38	49
All	All	237/328 (72%)	233 (98%)	4 (2%)	69	74

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25[B]	ARG
1	A	25[C]	ARG
2	B	388	VAL
2	B	505	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	211/303 (69%)	0.33	12 (5%)	23	22	42, 62, 113, 166	0
2	B	74/113 (65%)	0.69	12 (16%)	1	1	57, 85, 157, 179	0
All	All	285/416 (68%)	0.43	24 (8%)	11	9	42, 67, 120, 179	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	413	ASN	8.8
2	B	411	ASN	5.8
2	B	412	GLY	5.5
1	A	198	ALA	5.2
1	A	106	ILE	4.8
1	A	88	PRO	4.8
2	B	394	ARG	4.3
1	A	20	ASP	4.0
1	A	166	ALA	3.9
1	A	18	THR	3.7
1	A	87	GLU	3.3
2	B	414	GLY	3.2
2	B	410	SER	3.2
1	A	104	ALA	3.1
2	B	395	LYS	3.1
2	B	390	TRP	2.8
1	A	86	PRO	2.7
1	A	160	LEU	2.7
1	A	105	THR	2.5
2	B	472	GLY	2.4
1	A	242	SER	2.4
2	B	388	VAL	2.4
2	B	469	SER	2.2
2	B	409	PRO	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	ZN	B	601	1/1	0.97	0.17	62,62,62,62	0
3	ZN	B	602	1/1	0.99	0.12	67,67,67,67	0

6.5 Other polymers ⓘ

There are no such residues in this entry.