Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 10:06 PM GMT

PDB ID	:	1S4U
Title	:	Crystal Structure analysis of the beta-propeller protein Ski8p
Authors	:	Cheng, Z.; Song, H.
Deposited on	:	2004-01-18
Resolution	:	2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

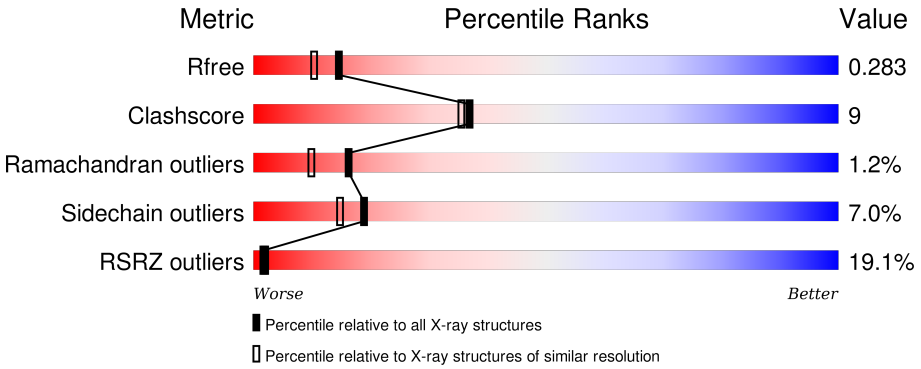
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.10 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	X	407	17% 67% 19% • 13%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2974 atoms, of which 0 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 Antiviral protein SKI8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	356	Total	C	N	O	S	0	0	0
			2816	1790	474	539	13			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	-9	GLY	-	CLONING ARTIFACT	UNP Q02793
X	-8	PRO	-	CLONING ARTIFACT	UNP Q02793
X	-7	LEU	-	CLONING ARTIFACT	UNP Q02793
X	-6	GLY	-	CLONING ARTIFACT	UNP Q02793
X	-5	SER	-	CLONING ARTIFACT	UNP Q02793
X	-4	PRO	-	CLONING ARTIFACT	UNP Q02793
X	-3	GLU	-	CLONING ARTIFACT	UNP Q02793
X	-2	PHE	-	CLONING ARTIFACT	UNP Q02793
X	-1	PRO	-	CLONING ARTIFACT	UNP Q02793
X	0	GLY	-	CLONING ARTIFACT	UNP Q02793

- Molecule 2 is water.

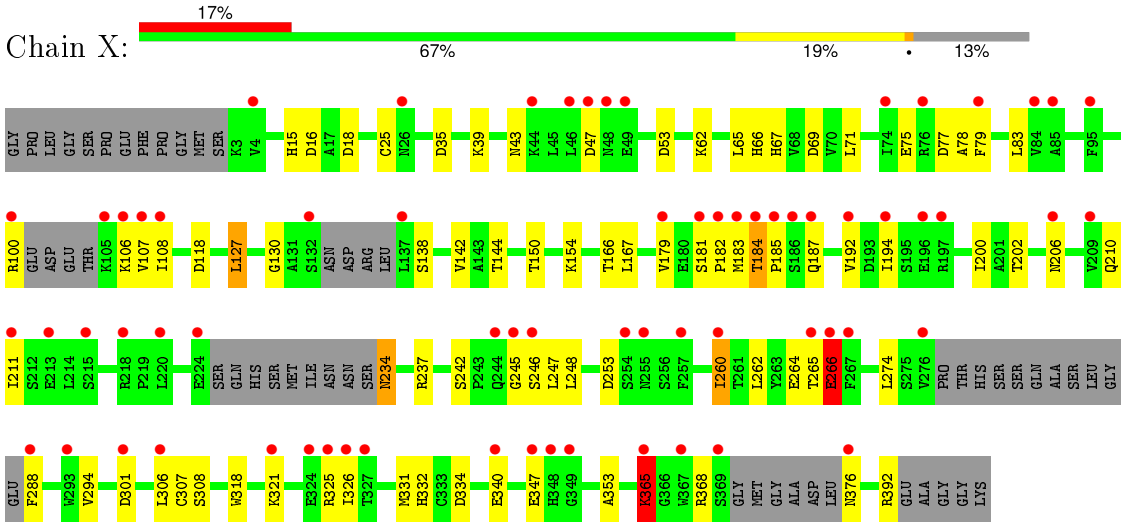
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	X	158	Total	O	0	0
			158	158		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Antiviral protein SKI8

Chain X:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.56Å 66.79Å 81.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 2.10 19.96 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.96-2.10) 99.9 (19.96-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.44 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.251 , 0.283 0.255 , 0.283	Depositor DCC
R_{free} test set	1108 reflections (5.45%)	DCC
Wilson B-factor (Å^2)	27.8	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.8	EDS
Estimated twinning fraction	0.034 for k,h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Outliers	0 of 21439 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2974	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.00% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	X	0.36	0/2879	0.73	9/3891 (0.2%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	X	47	ASP	CB-CG-OD2	5.83	123.54	118.30
1	X	118	ASP	CB-CG-OD2	5.68	123.41	118.30
1	X	301	ASP	CB-CG-OD2	5.57	123.31	118.30
1	X	334	ASP	CB-CG-OD2	5.33	123.10	118.30
1	X	16	ASP	CB-CG-OD2	5.28	123.05	118.30
1	X	69	ASP	CB-CG-OD2	5.27	123.04	118.30
1	X	35	ASP	CB-CG-OD2	5.17	122.95	118.30
1	X	18	ASP	CB-CG-OD2	5.16	122.94	118.30
1	X	53	ASP	CB-CG-OD2	5.02	122.82	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	X	2816	0	2723	52	0
2	X	158	0	0	13	0
All	All	2974	0	2723	52	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 9.

All (52) 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:X:245:GLY:O	1:X:265:THR:HG21	1.74	0.86
1:X:265:THR:HG23	1:X:265:THR:O	1.87	0.75
1:X:260:ILE:HG21	1:X:274:LEU:HD12	1.72	0.71
1:X:184:THR:HB	1:X:185:PRO:CD	2.21	0.70
1:X:66:HIS:HD2	1:X:67:HIS:ND1	1.90	0.70
1:X:150:THR:HG21	1:X:202:THR:OG1	1.93	0.68
1:X:294:VAL:HG13	2:X:416:HOH:O	1.92	0.68
1:X:211:ILE:HD11	1:X:248:LEU:HD22	1.75	0.68
1:X:307:CYS:O	2:X:534:HOH:O	2.13	0.66
1:X:308:SER:CB	2:X:416:HOH:O	2.45	0.65
1:X:260:ILE:HD11	2:X:416:HOH:O	1.95	0.65
1:X:144:THR:HG23	1:X:192:VAL:CG1	2.28	0.64
1:X:294:VAL:CG1	2:X:416:HOH:O	2.45	0.64
1:X:194:ILE:HG22	1:X:200:ILE:HG12	1.82	0.62
1:X:184:THR:HB	1:X:185:PRO:HD3	1.83	0.60
1:X:15:HIS:CE1	1:X:39:LYS:HD3	2.38	0.59
1:X:144:THR:HG23	1:X:192:VAL:HG11	1.85	0.58
1:X:130:GLY:CA	1:X:194:ILE:HD11	2.33	0.58
1:X:237:ARG:HD2	1:X:253:ASP:OD2	2.04	0.57
1:X:184:THR:CB	1:X:185:PRO:CD	2.83	0.56
1:X:211:ILE:HD11	1:X:248:LEU:CD2	2.37	0.55
1:X:166:THR:O	1:X:166:THR:HG22	2.06	0.55
1:X:67:HIS:HB3	1:X:127:LEU:HD22	1.88	0.54
1:X:182:PRO:HG3	1:X:210:GLN:NE2	2.22	0.54
1:X:264:GLU:OE2	1:X:266:GLU:HB3	2.08	0.54
1:X:142:VAL:HG21	1:X:194:ILE:HG21	1.89	0.53
1:X:75:GLU:HB3	1:X:79:PHE:HB3	1.90	0.53
1:X:130:GLY:N	1:X:194:ILE:HD11	2.25	0.51
1:X:265:THR:CG2	1:X:265:THR:O	2.55	0.50
1:X:308:SER:HB3	2:X:416:HOH:O	2.10	0.50
1:X:192:VAL:HB	2:X:436:HOH:O	2.11	0.50
1:X:67:HIS:HB3	1:X:127:LEU:CD2	2.41	0.49
1:X:144:THR:CG2	1:X:192:VAL:CG1	2.90	0.49
1:X:150:THR:HG23	2:X:399:HOH:O	2.11	0.49
1:X:144:THR:HG22	1:X:150:THR:HG22	1.96	0.48
1:X:71:LEU:HD23	1:X:83:LEU:HD12	1.96	0.48
1:X:166:THR:O	1:X:166:THR:CG2	2.63	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:X:234:ASN:HB3	1:X:253:ASP:O	2.13	0.47
1:X:138:SER:HB2	1:X:154:LYS:HG2	1.97	0.45
1:X:247:LEU:HD11	1:X:262:LEU:HB3	1.97	0.45
1:X:288:PHE:N	2:X:455:HOH:O	2.50	0.44
1:X:78:ALA:HB1	1:X:167:LEU:HD21	1.99	0.44
1:X:106:LYS:O	1:X:108:ILE:HG13	2.18	0.43
1:X:144:THR:HG23	1:X:192:VAL:HG13	2.00	0.43
1:X:187:GLN:HB3	2:X:505:HOH:O	2.18	0.43
1:X:25:CYS:O	1:X:43:ASN:ND2	2.51	0.43
1:X:246:SER:HA	1:X:265:THR:HG22	2.02	0.42
1:X:365:LYS:NZ	2:X:518:HOH:O	2.52	0.42
1:X:332:HIS:HA	1:X:353:ALA:O	2.21	0.41
1:X:274:LEU:HD13	1:X:318:TRP:CG	2.55	0.41
1:X:100:ARG:CZ	2:X:543:HOH:O	2.69	0.41
1:X:206:ASN:ND2	2:X:505:HOH:O	2.55	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	X	344/407 (84%)	322 (94%)	18 (5%)	4 (1%)	16 10

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	184	THR
1	X	266	GLU
1	X	365	LYS
1	X	107	VAL

5.3.2 Protein sidechains (i)

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	X	314/354 (89%)	292 (93%)	22 (7%)	19 15

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	X	62	LYS
1	X	65	LEU
1	X	77	ASP
1	X	127	LEU
1	X	179	VAL
1	X	181	SER
1	X	183	MET
1	X	234	ASN
1	X	242	SER
1	X	260	ILE
1	X	266	GLU
1	X	306	LEU
1	X	321	LYS
1	X	325	ARG
1	X	326	ILE
1	X	331	MET
1	X	340	GLU
1	X	347	GLU
1	X	365	LYS
1	X	368	ARG
1	X	376	ASN
1	X	392	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	X	66	HIS
1	X	210	GLN
1	X	234	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	X	376	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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, 95^{th} 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	X	356/407 (87%)	1.14	68 (19%) 2 2	20, 34, 49, 54	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	X	184	THR	8.1
1	X	215	SER	6.2
1	X	255	ASN	6.0
1	X	47	ASP	5.8
1	X	288	PHE	5.2
1	X	48	ASN	5.1
1	X	185	PRO	4.9
1	X	46	LEU	4.8
1	X	267	PHE	4.7
1	X	321	LYS	4.7
1	X	105	LYS	4.4
1	X	348	HIS	4.4
1	X	376	ASN	4.3
1	X	76	ARG	4.2
1	X	257	PHE	4.1
1	X	244	GLN	4.1
1	X	49	GLU	4.0
1	X	197	ARG	3.9
1	X	196	GLU	3.8
1	X	218	ARG	3.7
1	X	365	LYS	3.6
1	X	369	SER	3.5
1	X	347	GLU	3.5
1	X	182	PRO	3.3
1	X	132	SER	3.3
1	X	266	GLU	3.3
1	X	209	VAL	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	X	324	GLU	3.2
1	X	245	GLY	3.1
1	X	301	ASP	3.1
1	X	79	PHE	3.0
1	X	260	ILE	3.0
1	X	100	ARG	3.0
1	X	276	VAL	3.0
1	X	187	GLN	2.9
1	X	106	LYS	2.9
1	X	265	THR	2.9
1	X	220	LEU	2.9
1	X	44	LYS	2.8
1	X	326	ILE	2.8
1	X	340	GLU	2.8
1	X	137	LEU	2.8
1	X	108	ILE	2.7
1	X	186	SER	2.5
1	X	95	PHE	2.5
1	X	107	VAL	2.5
1	X	183	MET	2.5
1	X	367	TRP	2.4
1	X	254	SER	2.3
1	X	74	ILE	2.2
1	X	325	ARG	2.2
1	X	85	ALA	2.2
1	X	213	GLU	2.2
1	X	211	ILE	2.2
1	X	306	LEU	2.1
1	X	181	SER	2.1
1	X	224	GLU	2.1
1	X	349	GLY	2.1
1	X	192	VAL	2.1
1	X	84	VAL	2.1
1	X	179	VAL	2.1
1	X	327	THR	2.1
1	X	206	ASN	2.1
1	X	4	VAL	2.1
1	X	293	TRP	2.0
1	X	194	ILE	2.0
1	X	26	ASN	2.0
1	X	246	SER	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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.