WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2024 – 02:09 PM EST

PDB ID	:	8FM1
Title	:	Structure of CBASS Cap5 from Pseudomonas syringae in the absence of a ligand (apo form dimer)
Authors	:	Rechkoblit, O.; Kreitler, D.F.; Aggarwal, A.K.
Deposited on	:	2022-12-22
Resolution	:	3.16 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

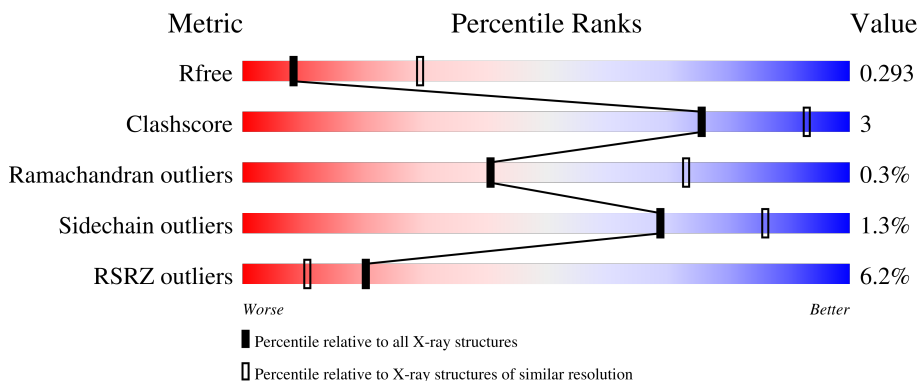|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.16 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		0.293
Clashscore		3
Ramachandran outliers		0.3%
Sidechain outliers		1.3%
RSRZ outliers		6.2%

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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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	388	2% · 84% 7% 9%
1	B	388	·% · 82% 10% 8%
1	C	388	3% · 82% 9% 9%
1	D	388	10% · 85% 6% 9%
1	E	388	6% · 84% 7% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	388	2% 85% 6% · 9%
1	G	388	5% 83% 8% 9%
1	H	388	5% 83% 8% 9%
1	I	388	6% 81% 9% 10%
1	J	388	12% 83% 6% · 11%
1	K	388	8% 86% 6% · 7%
1	L	388	9% 85% 7% · 8%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 65500 atoms, of which 32538 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 SAVED domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	353	Total	C	H	N	O	S	0	0	0
			5450	1735	2707	485	513	10			
1	B	358	Total	C	H	N	O	S	0	0	0
			5524	1754	2744	494	522	10			
1	C	352	Total	C	H	N	O	S	0	0	0
			5419	1723	2693	483	510	10			
1	D	352	Total	C	H	N	O	S	0	0	0
			5441	1731	2702	485	513	10			
1	E	352	Total	C	H	N	O	S	0	0	0
			5442	1731	2703	485	513	10			
1	F	352	Total	C	H	N	O	S	0	0	0
			5432	1725	2699	484	514	10			
1	G	354	Total	C	H	N	O	S	0	0	0
			5469	1739	2717	487	516	10			
1	H	354	Total	C	H	N	O	S	0	0	0
			5461	1734	2715	487	515	10			
1	I	350	Total	C	H	N	O	S	0	0	0
			5418	1724	2695	482	507	10			
1	J	347	Total	C	H	N	O	S	0	0	0
			5376	1708	2672	479	507	10			
1	K	360	Total	C	H	N	O	S	0	0	0
			5549	1766	2748	497	528	10			
1	L	357	Total	C	H	N	O	S	0	0	0
			5507	1747	2743	492	515	10			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0
2	E	1	Total 1	Zn 1	0	0
2	F	1	Total 1	Zn 1	0	0
2	G	1	Total 1	Zn 1	0	0
2	H	1	Total 1	Zn 1	0	0
2	I	1	Total 1	Zn 1	0	0
2	J	1	Total 1	Zn 1	0	0
2	K	1	Total 1	Zn 1	0	0
2	L	1	Total 1	Zn 1	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.
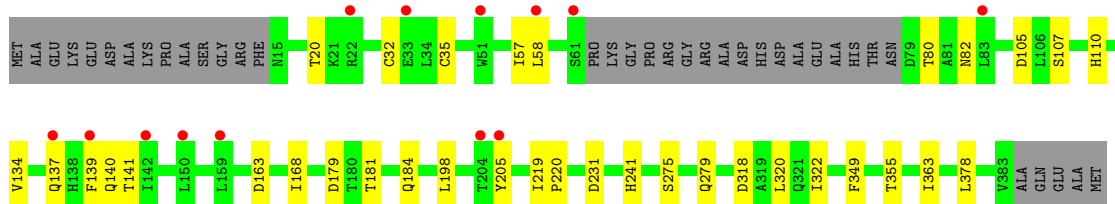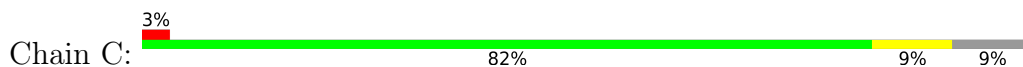
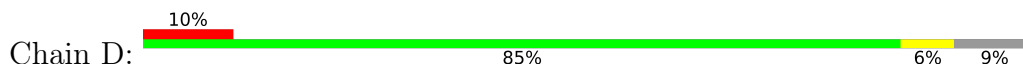• Molecule 1: SAVED domain-containing protein
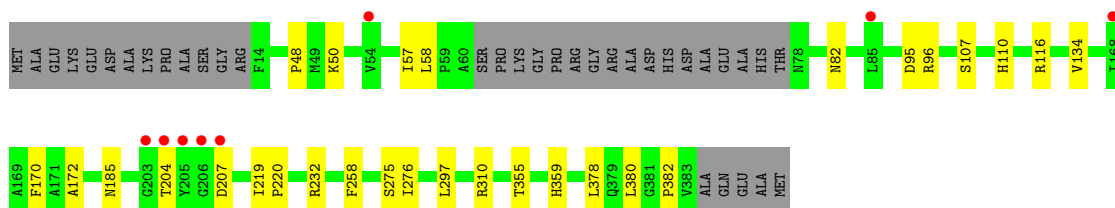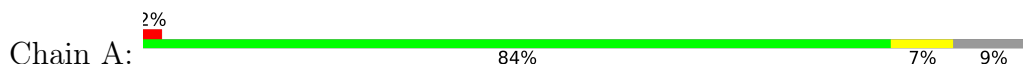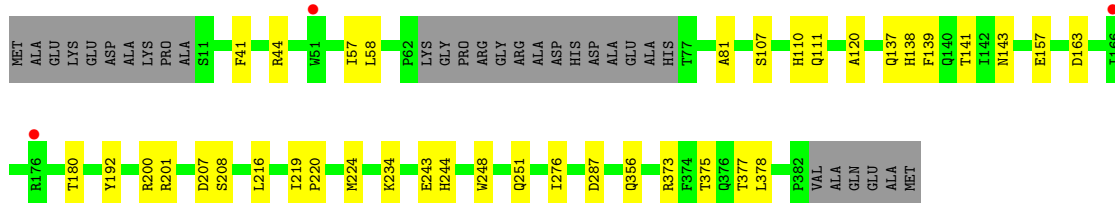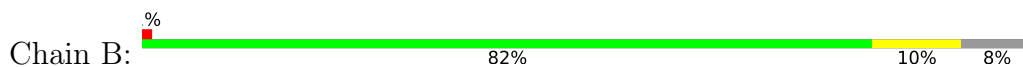
Chain A:

• Molecule 1: SAVED domain-containing protein

Chain B:

• Molecule 1: SAVED domain-containing protein

Chain C:

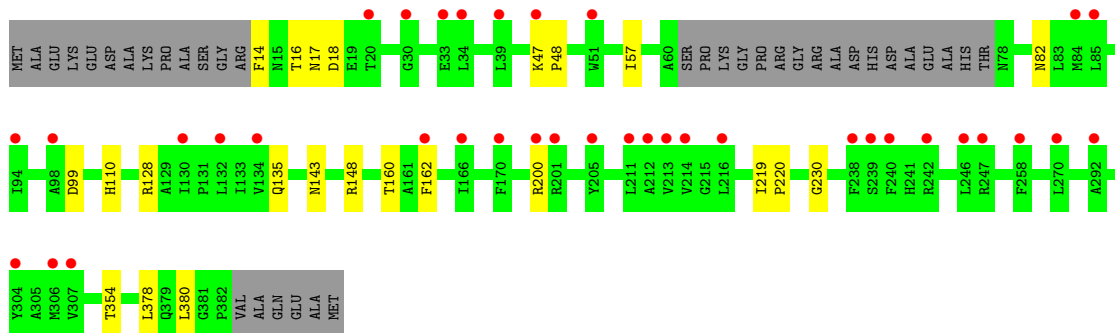• Molecule 1: SAVED domain-containing protein

Chain D:

• Molecule 1: SAVED domain-containing protein

Chain E:

• Molecule 1: SAVED domain-containing protein
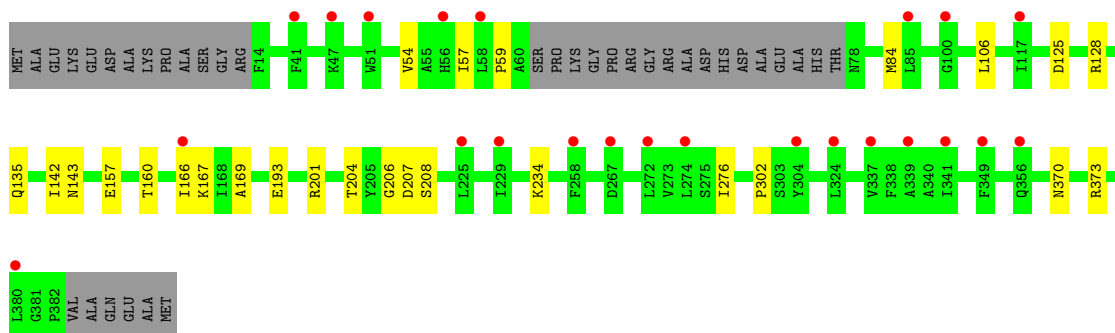
Chain F:

• Molecule 1: SAVED domain-containing protein

Chain G:

• Molecule 1: SAVED domain-containing protein

Chain H:

• Molecule 1: SAVED domain-containing protein

Chain I:

• Molecule 1: SAVED domain-containing protein
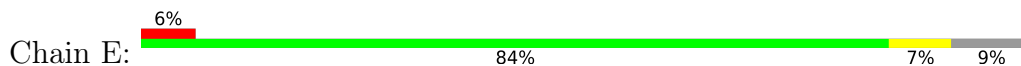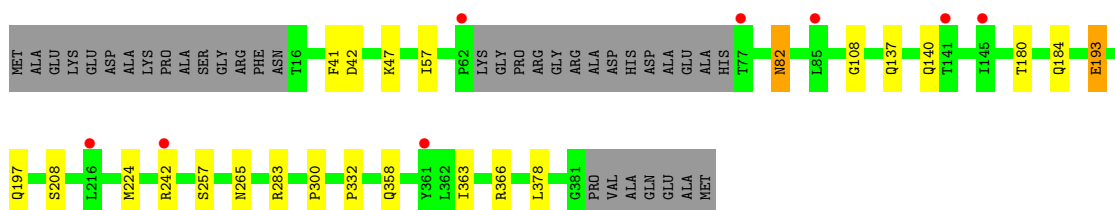
Chain J:
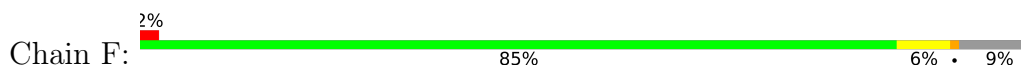

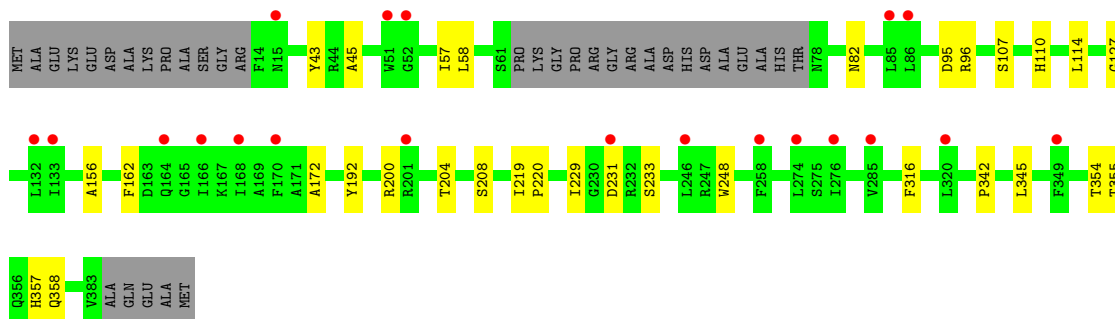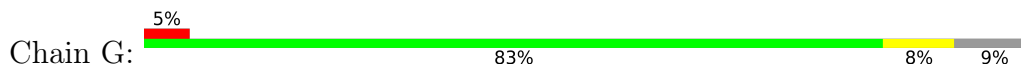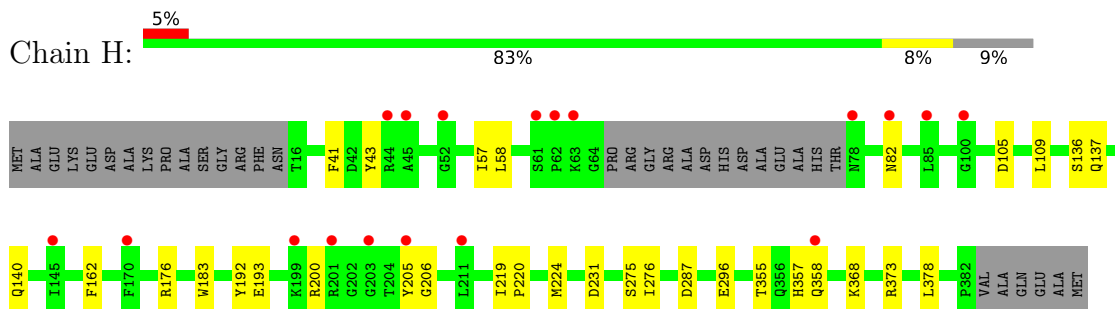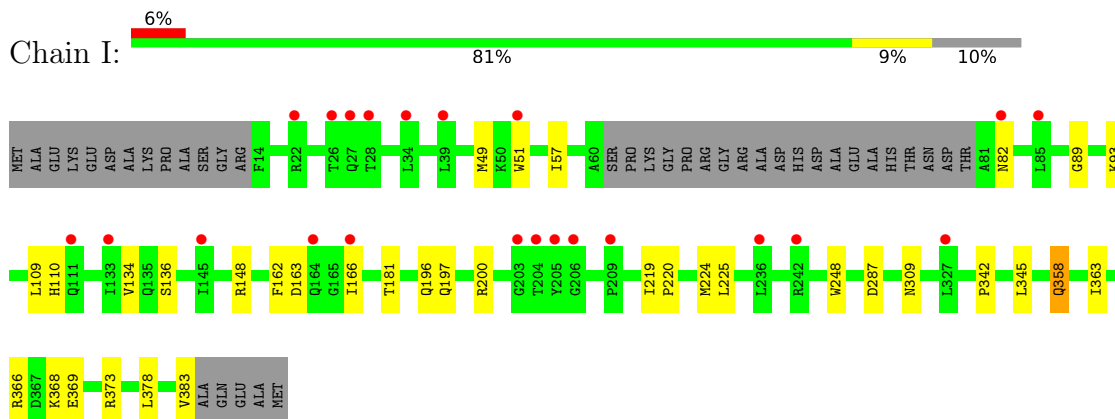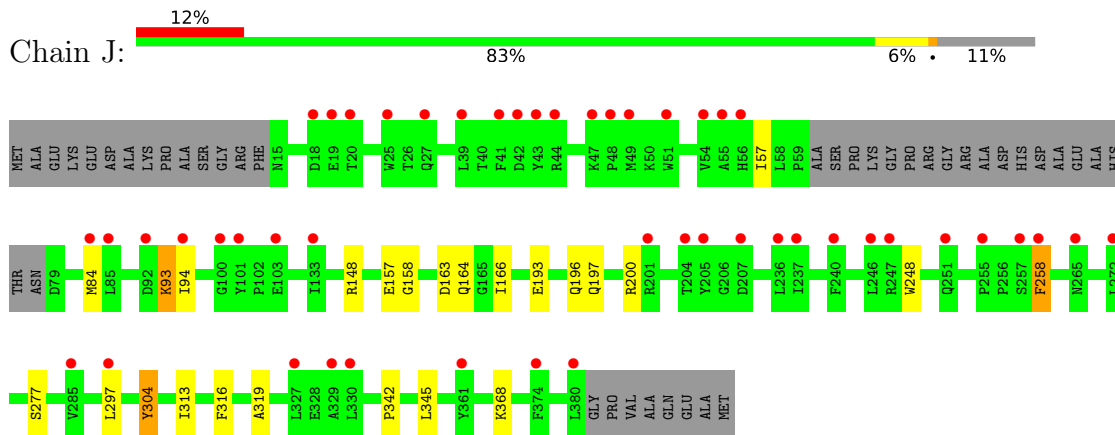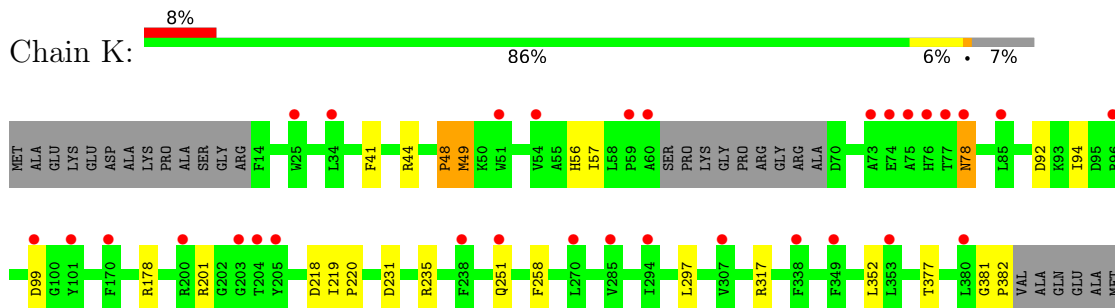• Molecule 1: SAVED domain-containing protein

Chain K:

• Molecule 1: SAVED domain-containing protein

Chain L:
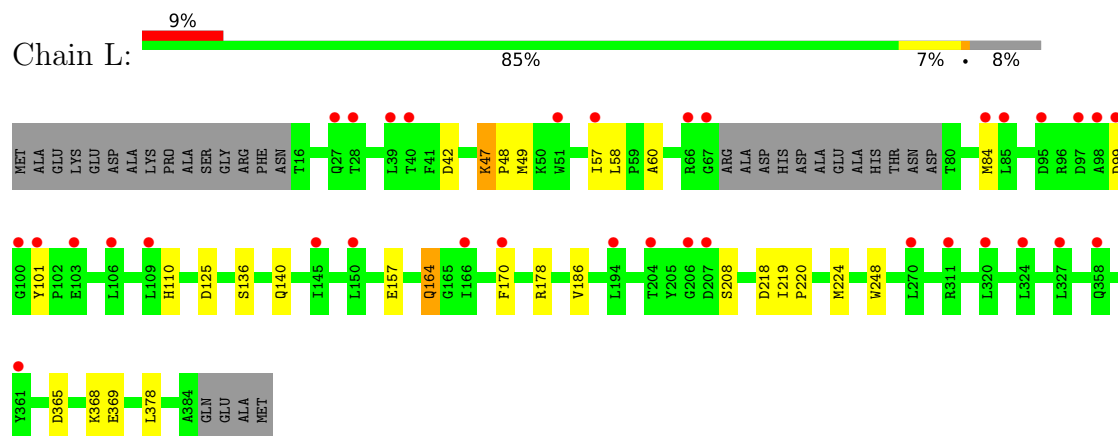
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	159.43Å 159.43Å 433.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	112.73 – 3.16 112.73 – 3.16	Depositor EDS
% Data completeness (in resolution range)	89.0 (112.73-3.16) 89.0 (112.73-3.16)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.18rc7_3834, PHENIX 1.18rc7_3834	Depositor
R, R_{free}	0.240 , 0.293 0.239 , 0.293	Depositor DCC
R_{free} test set	2000 reflections (2.33%)	wwPDB-VP
Wilson B-factor (Å^2)	129.5	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 80.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	65500	wwPDB-VP
Average B, all atoms (Å^2)	133.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.61% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.41	0/2807	0.65	0/3824
1	B	0.33	0/2845	0.59	0/3875
1	C	0.34	0/2789	0.62	0/3800
1	D	0.33	0/2803	0.61	0/3818
1	E	0.32	0/2803	0.58	0/3818
1	F	0.35	0/2796	0.62	1/3809 (0.0%)
1	G	0.41	2/2816 (0.1%)	0.62	0/3836
1	H	0.37	0/2810	0.63	0/3827
1	I	0.34	0/2787	0.60	0/3796
1	J	0.34	0/2766	0.57	0/3767
1	K	0.37	0/2867	0.67	1/3906 (0.0%)
1	L	0.45	0/2829	0.66	1/3853 (0.0%)
All	All	0.37	2/33718 (0.0%)	0.62	3/45929 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	248	TRP	CZ3-CH2	-5.41	1.31	1.40
1	G	248	TRP	CD1-NE1	-5.30	1.28	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	49	MET	CG-SD-CE	5.56	109.09	100.20
1	K	41	PHE	CB-CG-CD2	-5.50	116.95	120.80
1	F	41	PHE	CB-CG-CD2	-5.45	116.99	120.80

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	A	2743	2707	2701	16	0
1	B	2780	2744	2736	31	2
1	C	2726	2693	2687	23	0
1	D	2739	2702	2696	19	0
1	E	2739	2703	2696	18	2
1	F	2733	2699	2693	16	5
1	G	2752	2717	2710	17	0
1	H	2746	2715	2709	23	2
1	I	2723	2695	2688	15	9
1	J	2704	2672	2666	19	2
1	K	2801	2748	2741	14	0
1	L	2764	2743	2736	15	6
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
All	All	32962	32538	32459	202	14

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 3.

The worst 5 of 202 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:D:57:ILE:HD11	1:D:110:HIS:CD2	2.02	0.94
1:C:231:ASP:OD2	1:C:355:THR:OG1	1.86	0.93
1:H:200:ARG:O	1:H:200:ARG:HG3	1.75	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:231:ASP:OD2	1:H:355:THR:CG2	2.29	0.81
1:E:201:ARG:NH2	1:E:207:ASP:O	2.14	0.80

The worst 5 of 14 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:193:GLU:OE2	1:I:181:THR:OG1[6_444]	1.78	0.42
1:I:136:SER:OG	1:L:369:GLU:OE1[3_455]	1.99	0.21
1:B:180:THR:OG1	1:I:200:ARG:NH1[6_444]	2.00	0.20
1:F:193:GLU:OE2	1:I:181:THR:HG1[6_444]	1.40	0.20
1:I:136:SER:HG	1:L:369:GLU:OE1[3_455]	1.42	0.18

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	349/388 (90%)	333 (95%)	15 (4%)	1 (0%)	41	73
1	B	354/388 (91%)	342 (97%)	11 (3%)	1 (0%)	41	73
1	C	348/388 (90%)	329 (94%)	17 (5%)	2 (1%)	25	62
1	D	348/388 (90%)	329 (94%)	19 (6%)	0	100	100
1	E	348/388 (90%)	336 (97%)	12 (3%)	0	100	100
1	F	348/388 (90%)	329 (94%)	19 (6%)	0	100	100
1	G	350/388 (90%)	335 (96%)	15 (4%)	0	100	100
1	H	350/388 (90%)	327 (93%)	23 (7%)	0	100	100
1	I	346/388 (89%)	326 (94%)	19 (6%)	1 (0%)	41	73
1	J	343/388 (88%)	325 (95%)	18 (5%)	0	100	100
1	K	356/388 (92%)	333 (94%)	20 (6%)	3 (1%)	19	55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	353/388 (91%)	327 (93%)	23 (6%)	3 (1%)	19	55
All	All	4193/4656 (90%)	3971 (95%)	211 (5%)	11 (0%)	41	73

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	163	ASP
1	L	99	ASP
1	B	138	HIS
1	C	140	GLN
1	I	162	PHE

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	291/317 (92%)	286 (98%)	5 (2%)	60	82
1	B	296/317 (93%)	292 (99%)	4 (1%)	67	85
1	C	289/317 (91%)	286 (99%)	3 (1%)	76	89
1	D	291/317 (92%)	289 (99%)	2 (1%)	84	93
1	E	291/317 (92%)	290 (100%)	1 (0%)	92	97
1	F	291/317 (92%)	285 (98%)	6 (2%)	53	78
1	G	293/317 (92%)	289 (99%)	4 (1%)	67	85
1	H	292/317 (92%)	289 (99%)	3 (1%)	76	89
1	I	289/317 (91%)	283 (98%)	6 (2%)	53	78
1	J	288/317 (91%)	285 (99%)	3 (1%)	76	89
1	K	297/317 (94%)	293 (99%)	4 (1%)	69	86
1	L	293/317 (92%)	288 (98%)	5 (2%)	60	82
All	All	3501/3804 (92%)	3455 (99%)	46 (1%)	69	86

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	49	MET
1	J	258	PHE
1	I	51	TRP
1	I	196	GLN
1	K	48	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	82	ASN
1	G	357	HIS

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 12 ligands modelled in this entry, 12 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(Å²)	Q<0.9
1	A	353/388 (90%)	0.54	8 (2%)	60	46	79, 104, 134, 142	0
1	B	358/388 (92%)	0.42	3 (0%)	86	78	86, 110, 131, 139	0
1	C	352/388 (90%)	0.57	13 (3%)	41	25	80, 109, 142, 152	0
1	D	352/388 (90%)	0.71	37 (10%)	6	3	98, 121, 145, 163	0
1	E	352/388 (90%)	0.60	23 (6%)	18	10	101, 127, 152, 158	0
1	F	352/388 (90%)	0.43	8 (2%)	60	46	92, 113, 140, 151	0
1	G	354/388 (91%)	0.49	20 (5%)	24	12	105, 129, 150, 166	0
1	H	354/388 (91%)	0.56	18 (5%)	28	15	94, 113, 151, 168	0
1	I	350/388 (90%)	0.49	22 (6%)	20	10	91, 112, 158, 170	0
1	J	347/388 (89%)	0.82	48 (13%)	2	1	101, 138, 161, 179	0
1	K	360/388 (92%)	0.67	31 (8%)	10	5	112, 135, 166, 184	0
1	L	357/388 (92%)	0.61	34 (9%)	8	4	95, 119, 163, 180	0
All	All	4241/4656 (91%)	0.57	265 (6%)	20	11	79, 120, 152, 184	0

The worst 5 of 265 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	74	GLU	9.8
1	L	67	GLY	9.3
1	J	49	MET	8.1
1	K	77	THR	7.5
1	K	51	TRP	7.2

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
2	ZN	E	401	1/1	0.87	0.21	110,110,110,110	0
2	ZN	G	401	1/1	0.95	0.19	110,110,110,110	0
2	ZN	K	401	1/1	0.96	0.16	110,110,110,110	0
2	ZN	D	401	1/1	0.97	0.15	110,110,110,110	0
2	ZN	I	401	1/1	0.97	0.17	110,110,110,110	0
2	ZN	J	401	1/1	0.97	0.10	110,110,110,110	0
2	ZN	F	401	1/1	0.97	0.30	110,110,110,110	0
2	ZN	L	401	1/1	0.97	0.14	110,110,110,110	0
2	ZN	C	401	1/1	0.98	0.24	110,110,110,110	0
2	ZN	H	401	1/1	0.98	0.20	110,110,110,110	0
2	ZN	A	401	1/1	0.99	0.35	110,110,110,110	0
2	ZN	B	401	1/1	0.99	0.33	110,110,110,110	0

6.5 Other polymers ⓘ

There are no such residues in this entry.