



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 08:47 am BST

PDB ID : 5XDA
Title : Structural basis for Ufm1 recognition by UfSP
Authors : Kim, K.H.; Ha, B.H.; Kim, E.E.
Deposited on : 2017-03-28
Resolution : 3.29 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

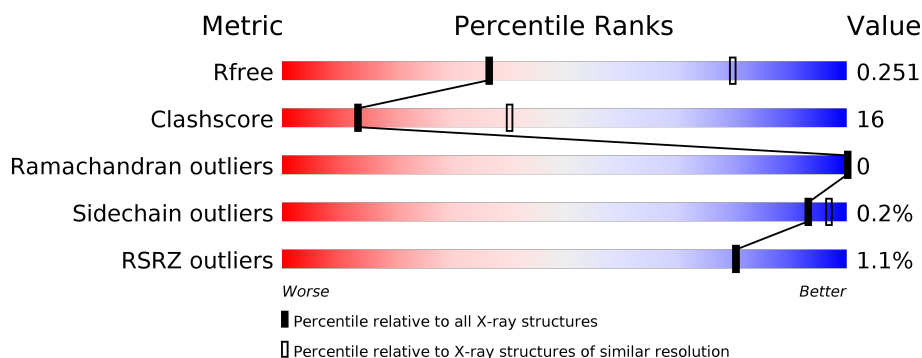
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 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}	130704	1177 (3.32-3.24)
Clashscore	141614	1044 (3.30-3.26)
Ramachandran outliers	138981	1026 (3.30-3.26)
Sidechain outliers	138945	1025 (3.30-3.26)
RSRZ outliers	127900	1141 (3.32-3.24)

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 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 $\leq 5\%$. 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	565	<div> <div>2%</div> <div>73% 20% • 5%</div> </div>
1	B	565	<div> <div>76% 19% • •</div> </div>
1	C	565	<div> <div>74% 20% • •</div> </div>
1	D	565	<div> <div>2%</div> <div>69% 25% • 5%</div> </div>
1	E	565	<div> <div>72% 21% • 5%</div> </div>
1	F	565	<div> <div>2%</div> <div>70% 23% • 5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	84	<div><div></div><div>77%23%</div></div>
2	H	84	<div><div></div><div>%89%11%</div></div>
2	I	84	<div><div></div><div>%81%17%•</div></div>
2	J	84	<div><div></div><div>2%73%24%•</div></div>
2	K	84	<div><div></div><div>98%•</div></div>
2	L	84	<div><div></div><div>4%86%13%•</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 29730 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 Ufm1-specific protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	537	Total	C	N	O	S	0	0	0
			4306	2733	755	802	16			
1	B	542	Total	C	N	O	S	0	0	0
			4342	2754	762	810	16			
1	C	541	Total	C	N	O	S	0	0	0
			4332	2748	759	809	16			
1	D	538	Total	C	N	O	S	0	0	0
			4312	2736	756	804	16			
1	E	535	Total	C	N	O	S	0	0	0
			4289	2723	751	799	16			
1	F	535	Total	C	N	O	S	0	0	0
			4289	2723	752	798	16			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	407	SER	CYS	engineered mutation	UNP Q94218
B	407	SER	CYS	engineered mutation	UNP Q94218
C	407	SER	CYS	engineered mutation	UNP Q94218
D	407	SER	CYS	engineered mutation	UNP Q94218
E	407	SER	CYS	engineered mutation	UNP Q94218
F	407	SER	CYS	engineered mutation	UNP Q94218

- Molecule 2 is a protein called Ubiquitin-fold modifier 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	84	Total	C	N	O	0	0	0
			627	405	106	116			
2	H	84	Total	C	N	O	0	0	0
			627	405	106	116			
2	I	84	Total	C	N	O	0	0	0
			627	405	106	116			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	J	84	Total 627	C 405	N 106	O 116	0	0	0
2	K	84	Total 627	C 405	N 106	O 116	0	0	0
2	L	84	Total 627	C 405	N 106	O 116	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	94	GLY	HIS	engineered mutation	UNP P34661
H	94	GLY	HIS	engineered mutation	UNP P34661
I	94	GLY	HIS	engineered mutation	UNP P34661
J	94	GLY	HIS	engineered mutation	UNP P34661
K	94	GLY	HIS	engineered mutation	UNP P34661
L	94	GLY	HIS	engineered mutation	UNP P34661

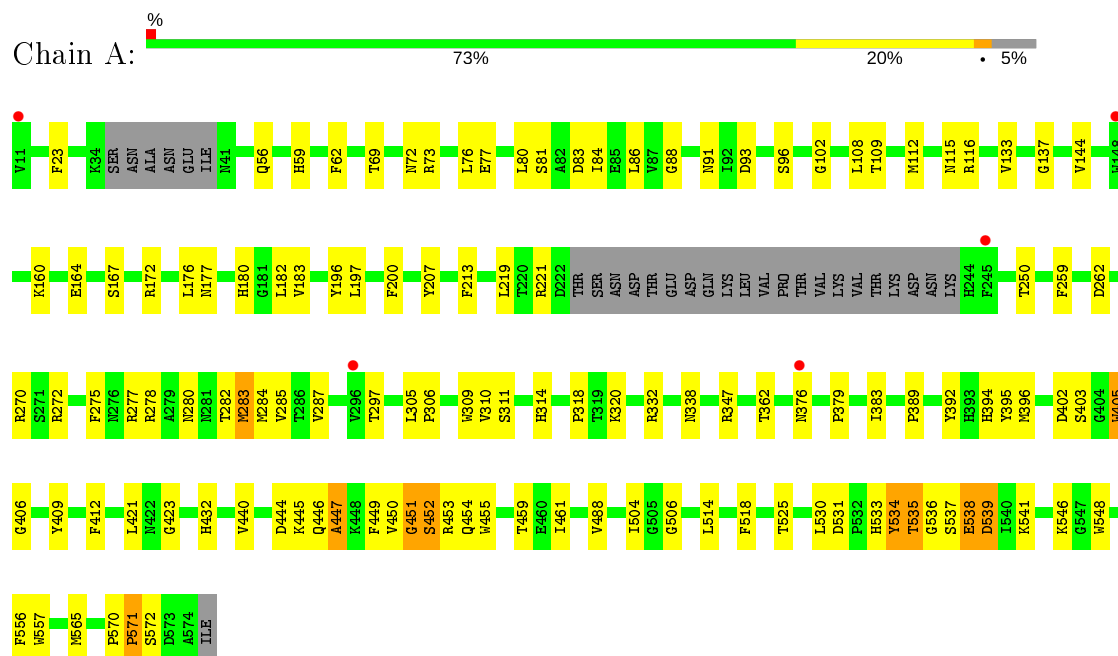
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total 18	O 18	0	0
3	B	19	Total 19	O 19	0	0
3	C	27	Total 27	O 27	0	0
3	D	6	Total 6	O 6	0	0
3	E	11	Total 11	O 11	0	0
3	F	5	Total 5	O 5	0	0
3	G	4	Total 4	O 4	0	0
3	H	2	Total 2	O 2	0	0
3	I	4	Total 4	O 4	0	0
3	L	2	Total 2	O 2	0	0

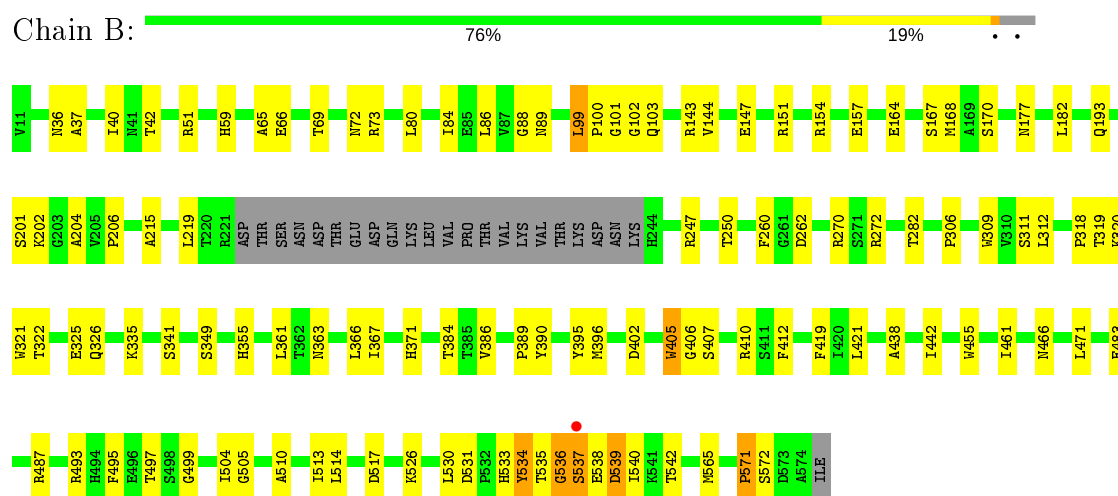
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Ufm1-specific protease

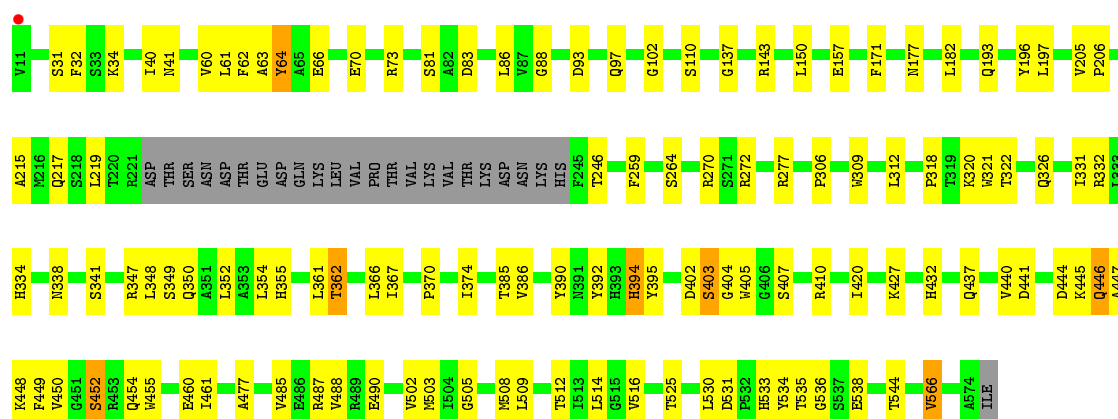


- Molecule 1: Ufm1-specific protease



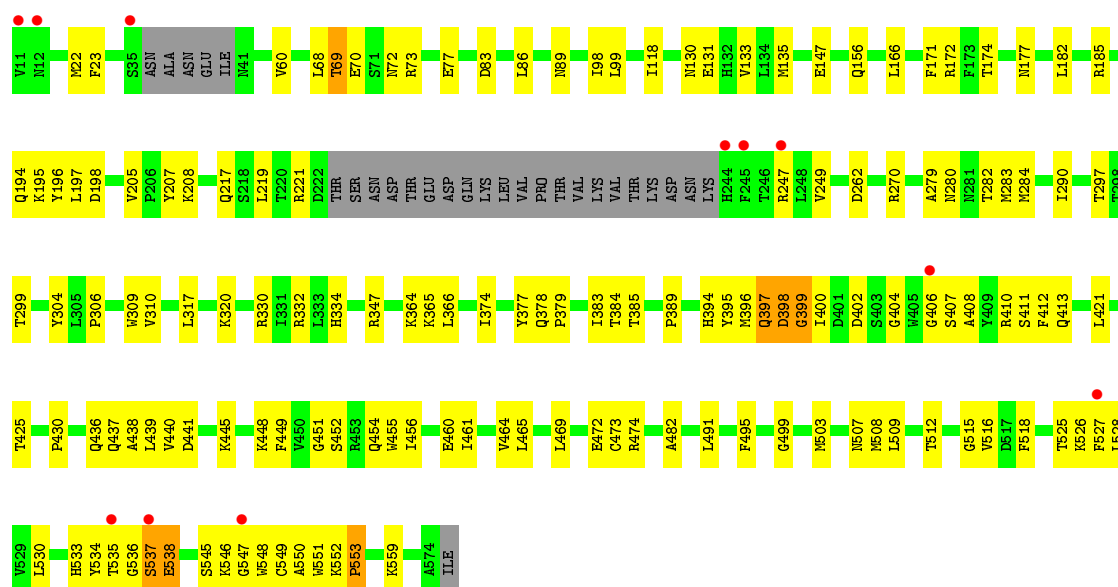
- Molecule 1: Ufm1-specific protease

Chain C: 



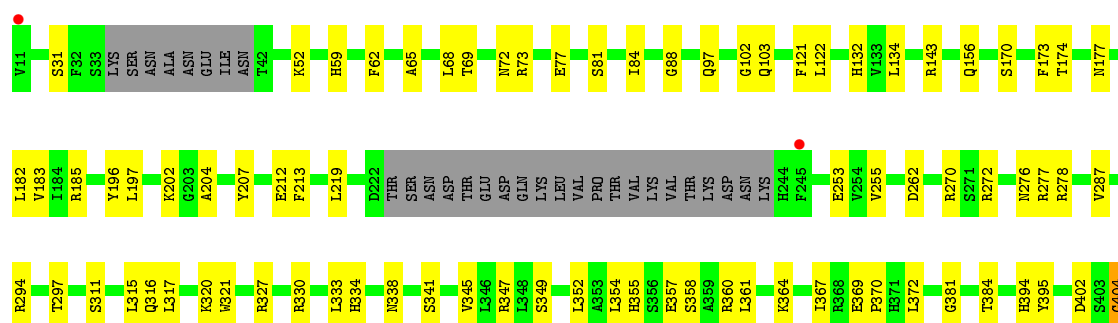
• Molecule 1: Ufm1-specific protease

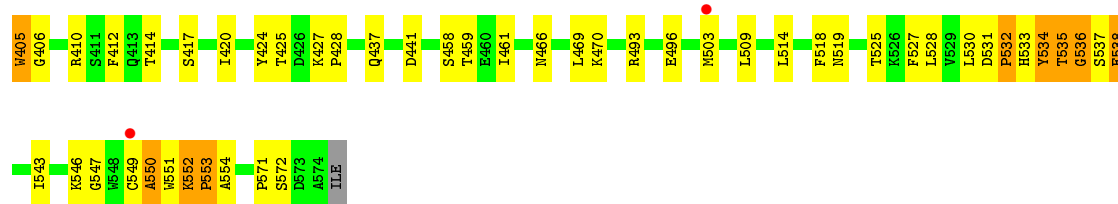
Chain D: 



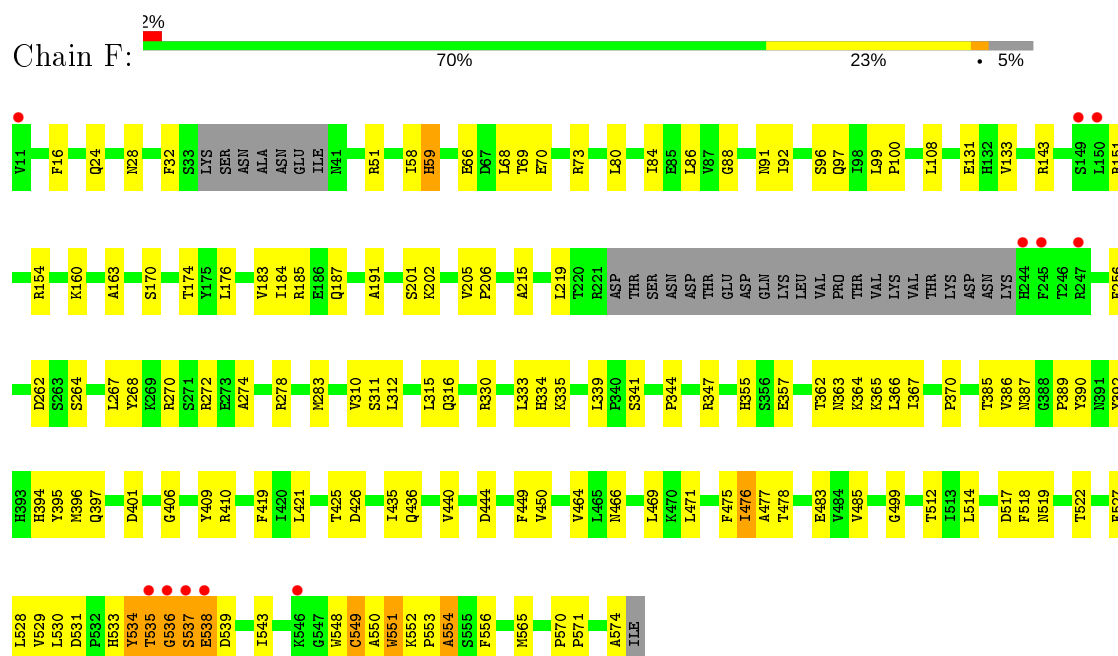
• Molecule 1: Ufm1-specific protease

Chain E: 

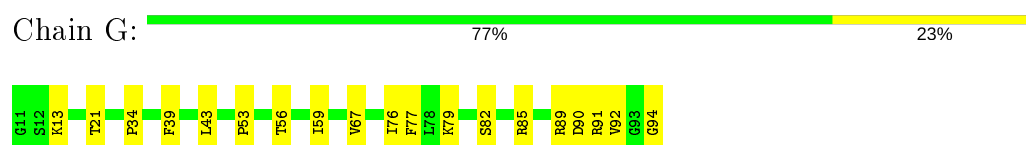




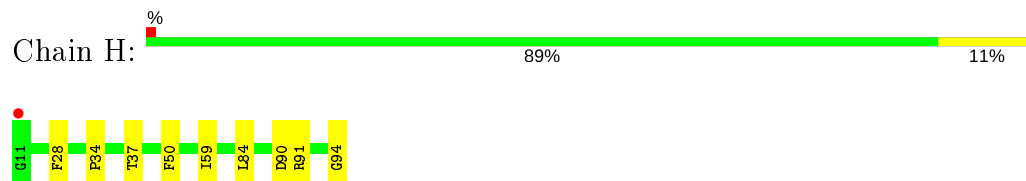
- Molecule 1: Ufm1-specific protease



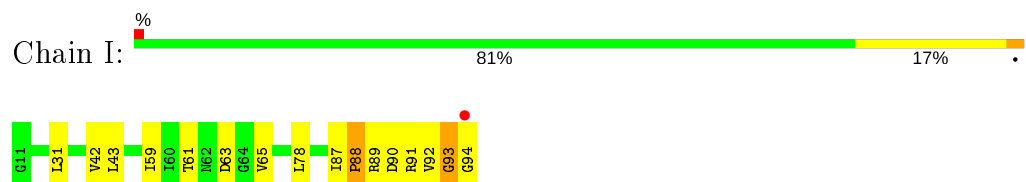
- Molecule 2: Ubiquitin-fold modifier 1



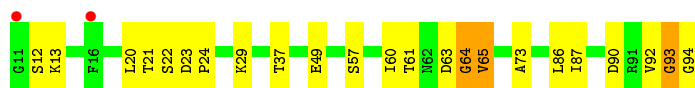
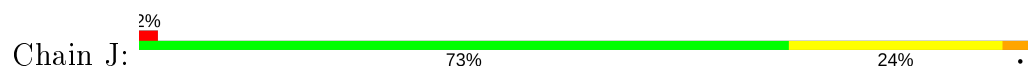
- Molecule 2: Ubiquitin-fold modifier 1



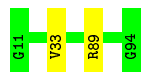
- Molecule 2: Ubiquitin-fold modifier 1



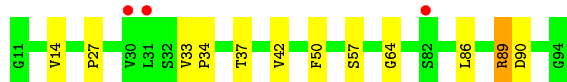
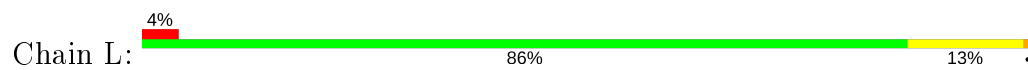
- Molecule 2: Ubiquitin-fold modifier 1



- Molecule 2: Ubiquitin-fold modifier 1



- Molecule 2: Ubiquitin-fold modifier 1



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	267.45Å 455.33Å 195.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.16 – 3.29 46.17 – 3.29	Depositor EDS
% Data completeness (in resolution range)	96.0 (46.16-3.29) 96.0 (46.17-3.29)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.26 (at 3.25Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.207 , 0.243 0.218 , 0.251	Depositor DCC
R_{free} test set	2018 reflections (1.16%)	wwPDB-VP
Wilson B-factor (Å ²)	71.1	Xtriage
Anisotropy	0.452	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 32.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.045 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.037 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	29730	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.81% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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

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	0.31	0/4407	0.57	12/5968 (0.2%)
1	B	0.32	0/4444	0.74	16/6020 (0.3%)
1	C	0.44	3/4433 (0.1%)	0.54	9/6005 (0.1%)
1	D	0.34	0/4413	0.61	12/5976 (0.2%)
1	E	0.33	1/4390 (0.0%)	0.69	25/5946 (0.4%)
1	F	0.29	0/4390	0.80	17/5946 (0.3%)
2	G	0.46	0/641	0.49	0/870
2	H	0.24	0/641	0.43	0/870
2	I	0.43	1/641 (0.2%)	0.73	3/870 (0.3%)
2	J	0.23	0/641	0.60	4/870 (0.5%)
2	K	0.23	0/641	0.44	0/870
2	L	0.22	0/641	0.48	1/870 (0.1%)
All	All	0.34	5/30323 (0.0%)	0.65	99/41081 (0.2%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	60	VAL	CB-CG1	8.76	1.71	1.52
2	I	88	PRO	N-CD	-8.45	1.36	1.47
1	C	566	VAL	C-O	8.09	1.38	1.23
1	C	331	ILE	CB-CG1	5.88	1.70	1.54
1	E	405	TRP	C-N	-5.72	1.22	1.33

The worst 5 of 99 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	550	ALA	N-CA-CB	-23.78	76.81	110.10
1	E	550	ALA	N-CA-CB	-19.38	82.97	110.10
1	F	550	ALA	CB-CA-C	-19.07	81.49	110.10
1	F	537	SER	CB-CA-C	17.32	143.01	110.10
1	B	538	GLU	N-CA-CB	-16.86	80.25	110.60

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	4306	0	4218	149	0
1	B	4342	0	4254	120	0
1	C	4332	0	4247	153	0
1	D	4312	0	4223	209	0
1	E	4289	0	4199	126	0
1	F	4289	0	4201	150	0
2	G	627	0	652	16	0
2	H	627	0	652	7	0
2	I	627	0	652	40	0
2	J	627	0	652	31	0
2	K	627	0	652	2	0
2	L	627	0	652	8	0
3	A	18	0	0	0	0
3	B	19	0	0	0	0
3	C	27	0	0	2	0
3	D	6	0	0	1	0
3	E	11	0	0	1	0
3	F	5	0	0	0	0
3	G	4	0	0	0	0
3	H	2	0	0	0	0
3	I	4	0	0	0	0
3	L	2	0	0	0	0
All	All	29730	0	29254	950	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 16.

The worst 5 of 950 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:411:SER:HB3	1:D:503:MET:CE	1.32	1.59
1:C:321:TRP:HE1	1:C:326:GLN:CB	1.15	1.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:394:HIS:CE1	1:A:535:THR:HB	1.45	1.50
1:C:321:TRP:NE1	1:C:326:GLN:HB2	1.21	1.44
1:D:411:SER:CB	1:D:503:MET:HE2	1.48	1.40

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	A	531/565 (94%)	490 (92%)	41 (8%)	0	100	100
1	B	538/565 (95%)	509 (95%)	29 (5%)	0	100	100
1	C	537/565 (95%)	508 (95%)	29 (5%)	0	100	100
1	D	532/565 (94%)	495 (93%)	37 (7%)	0	100	100
1	E	529/565 (94%)	497 (94%)	32 (6%)	0	100	100
1	F	529/565 (94%)	471 (89%)	58 (11%)	0	100	100
2	G	82/84 (98%)	80 (98%)	2 (2%)	0	100	100
2	H	82/84 (98%)	81 (99%)	1 (1%)	0	100	100
2	I	82/84 (98%)	81 (99%)	1 (1%)	0	100	100
2	J	82/84 (98%)	77 (94%)	5 (6%)	0	100	100
2	K	82/84 (98%)	78 (95%)	4 (5%)	0	100	100
2	L	82/84 (98%)	75 (92%)	7 (8%)	0	100	100
All	All	3688/3894 (95%)	3442 (93%)	246 (7%)	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	468/495 (94%)	466 (100%)	2 (0%)	91	95
1	B	472/495 (95%)	472 (100%)	0	100	100
1	C	471/495 (95%)	470 (100%)	1 (0%)	93	97
1	D	469/495 (95%)	469 (100%)	0	100	100
1	E	466/495 (94%)	466 (100%)	0	100	100
1	F	466/495 (94%)	464 (100%)	2 (0%)	91	95
2	G	69/69 (100%)	69 (100%)	0	100	100
2	H	69/69 (100%)	69 (100%)	0	100	100
2	I	69/69 (100%)	69 (100%)	0	100	100
2	J	69/69 (100%)	69 (100%)	0	100	100
2	K	69/69 (100%)	69 (100%)	0	100	100
2	L	69/69 (100%)	69 (100%)	0	100	100
All	All	3226/3384 (95%)	3221 (100%)	5 (0%)	93	97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	283	MET
1	A	332	ARG
1	C	394	HIS
1	F	59	HIS
1	F	518	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	72	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	537/565 (95%)	-0.22	5 (0%) 84 84	6, 35, 74, 110	0
1	B	542/565 (95%)	-0.28	1 (0%) 95 96	7, 27, 76, 117	0
1	C	541/565 (95%)	-0.25	1 (0%) 95 96	6, 31, 77, 120	0
1	D	538/565 (95%)	0.01	11 (2%) 65 63	19, 58, 105, 131	0
1	E	535/565 (94%)	-0.12	4 (0%) 87 88	16, 53, 94, 128	0
1	F	535/565 (94%)	0.05	11 (2%) 63 61	31, 72, 111, 134	0
2	G	84/84 (100%)	-0.19	0 100 100	10, 23, 50, 87	0
2	H	84/84 (100%)	-0.06	1 (1%) 79 78	15, 28, 49, 83	0
2	I	84/84 (100%)	-0.11	1 (1%) 79 78	5, 24, 48, 84	0
2	J	84/84 (100%)	0.29	2 (2%) 59 55	33, 62, 88, 131	0
2	K	84/84 (100%)	-0.19	0 100 100	9, 27, 53, 91	0
2	L	84/84 (100%)	0.20	3 (3%) 42 40	24, 70, 90, 101	0
All	All	3732/3894 (95%)	-0.12	40 (1%) 80 80	5, 45, 95, 134	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	94	GLY	4.2
1	D	11	VAL	4.2
1	C	11	VAL	4.1
1	F	536	GLY	3.8
1	D	537	SER	3.6

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 [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.