wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 04:35 am BST

PDB ID	:	6SB4
Title	:	Crystal structure of murine perforin-2 P2 domain crystal form 2
Authors	:	Ni, T.; Yu, X.; Ginger, L.; Gilbert, R.J.C.
Deposited on	:	2019-07-18
Resolution	:	3.17 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 following versions of software and data (see references (1)) were used in the production of this report:

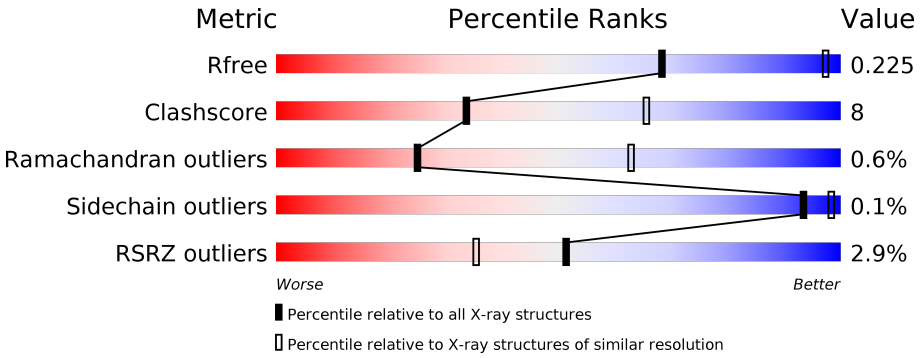
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.17 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.



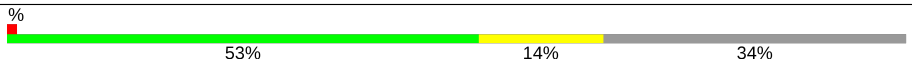
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1467 (3.20-3.16)
Clashscore	141614	1599 (3.20-3.16)
Ramachandran outliers	138981	1574 (3.20-3.16)
Sidechain outliers	138945	1573 (3.20-3.16)
RSRZ outliers	127900	1423 (3.20-3.16)

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 <=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	295	2% 53% 14% 34%
1	B	295	5% 55% 12% 33%
1	C	295	% 56% 11% 33%
1	D	295	4% 50% 16% 34%
1	E	295	57% 10% 33%
1	F	295	% 51% 15% 34%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	295	
1	H	295	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 11992 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 Macrophage-expressed gene 1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	196	Total	C	N	O	S	0	0	0
			1493	958	241	278	16			
1	B	198	Total	C	N	O	S	0	0	0
			1511	972	243	280	16			
1	C	197	Total	C	N	O	S	0	0	0
			1504	968	242	278	16			
1	D	195	Total	C	N	O	S	0	0	0
			1492	960	240	276	16			
1	E	197	Total	C	N	O	S	0	0	0
			1504	968	242	278	16			
1	F	196	Total	C	N	O	S	0	0	0
			1497	965	240	276	16			
1	G	196	Total	C	N	O	S	0	0	0
			1495	960	241	278	16			
1	H	196	Total	C	N	O	S	0	0	0
			1496	961	241	278	16			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	346	GLU	-	expression tag	UNP A1L314
A	347	THR	-	expression tag	UNP A1L314
A	348	GLY	-	expression tag	UNP A1L314
A	632	GLY	-	expression tag	UNP A1L314
A	633	THR	-	expression tag	UNP A1L314
A	634	LYS	-	expression tag	UNP A1L314
A	635	HIS	-	expression tag	UNP A1L314
A	636	HIS	-	expression tag	UNP A1L314
A	637	HIS	-	expression tag	UNP A1L314
A	638	HIS	-	expression tag	UNP A1L314
A	639	HIS	-	expression tag	UNP A1L314
A	640	HIS	-	expression tag	UNP A1L314
B	346	GLU	-	expression tag	UNP A1L314

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	347	THR	-	expression tag	UNP A1L314
B	348	GLY	-	expression tag	UNP A1L314
B	632	GLY	-	expression tag	UNP A1L314
B	633	THR	-	expression tag	UNP A1L314
B	634	LYS	-	expression tag	UNP A1L314
B	635	HIS	-	expression tag	UNP A1L314
B	636	HIS	-	expression tag	UNP A1L314
B	637	HIS	-	expression tag	UNP A1L314
B	638	HIS	-	expression tag	UNP A1L314
B	639	HIS	-	expression tag	UNP A1L314
B	640	HIS	-	expression tag	UNP A1L314
C	346	GLU	-	expression tag	UNP A1L314
C	347	THR	-	expression tag	UNP A1L314
C	348	GLY	-	expression tag	UNP A1L314
C	632	GLY	-	expression tag	UNP A1L314
C	633	THR	-	expression tag	UNP A1L314
C	634	LYS	-	expression tag	UNP A1L314
C	635	HIS	-	expression tag	UNP A1L314
C	636	HIS	-	expression tag	UNP A1L314
C	637	HIS	-	expression tag	UNP A1L314
C	638	HIS	-	expression tag	UNP A1L314
C	639	HIS	-	expression tag	UNP A1L314
C	640	HIS	-	expression tag	UNP A1L314
D	346	GLU	-	expression tag	UNP A1L314
D	347	THR	-	expression tag	UNP A1L314
D	348	GLY	-	expression tag	UNP A1L314
D	632	GLY	-	expression tag	UNP A1L314
D	633	THR	-	expression tag	UNP A1L314
D	634	LYS	-	expression tag	UNP A1L314
D	635	HIS	-	expression tag	UNP A1L314
D	636	HIS	-	expression tag	UNP A1L314
D	637	HIS	-	expression tag	UNP A1L314
D	638	HIS	-	expression tag	UNP A1L314
D	639	HIS	-	expression tag	UNP A1L314
D	640	HIS	-	expression tag	UNP A1L314
E	346	GLU	-	expression tag	UNP A1L314
E	347	THR	-	expression tag	UNP A1L314
E	348	GLY	-	expression tag	UNP A1L314
E	632	GLY	-	expression tag	UNP A1L314
E	633	THR	-	expression tag	UNP A1L314
E	634	LYS	-	expression tag	UNP A1L314
E	635	HIS	-	expression tag	UNP A1L314

Continued on next page...

Continued from previous page...

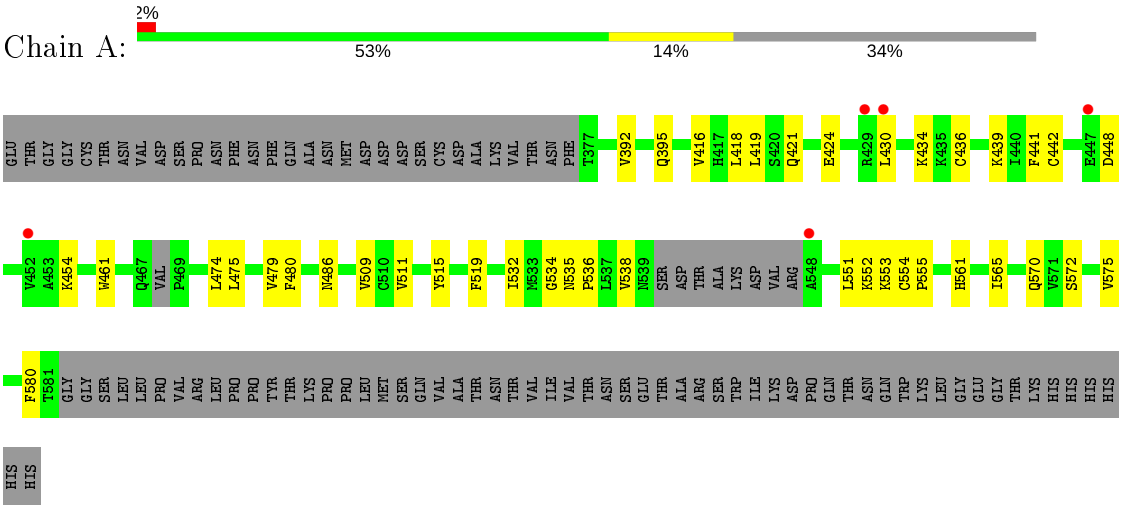
Chain	Residue	Modelled	Actual	Comment	Reference
E	636	HIS	-	expression tag	UNP A1L314
E	637	HIS	-	expression tag	UNP A1L314
E	638	HIS	-	expression tag	UNP A1L314
E	639	HIS	-	expression tag	UNP A1L314
E	640	HIS	-	expression tag	UNP A1L314
F	346	GLU	-	expression tag	UNP A1L314
F	347	THR	-	expression tag	UNP A1L314
F	348	GLY	-	expression tag	UNP A1L314
F	632	GLY	-	expression tag	UNP A1L314
F	633	THR	-	expression tag	UNP A1L314
F	634	LYS	-	expression tag	UNP A1L314
F	635	HIS	-	expression tag	UNP A1L314
F	636	HIS	-	expression tag	UNP A1L314
F	637	HIS	-	expression tag	UNP A1L314
F	638	HIS	-	expression tag	UNP A1L314
F	639	HIS	-	expression tag	UNP A1L314
F	640	HIS	-	expression tag	UNP A1L314
G	346	GLU	-	expression tag	UNP A1L314
G	347	THR	-	expression tag	UNP A1L314
G	348	GLY	-	expression tag	UNP A1L314
G	632	GLY	-	expression tag	UNP A1L314
G	633	THR	-	expression tag	UNP A1L314
G	634	LYS	-	expression tag	UNP A1L314
G	635	HIS	-	expression tag	UNP A1L314
G	636	HIS	-	expression tag	UNP A1L314
G	637	HIS	-	expression tag	UNP A1L314
G	638	HIS	-	expression tag	UNP A1L314
G	639	HIS	-	expression tag	UNP A1L314
G	640	HIS	-	expression tag	UNP A1L314
H	346	GLU	-	expression tag	UNP A1L314
H	347	THR	-	expression tag	UNP A1L314
H	348	GLY	-	expression tag	UNP A1L314
H	632	GLY	-	expression tag	UNP A1L314
H	633	THR	-	expression tag	UNP A1L314
H	634	LYS	-	expression tag	UNP A1L314
H	635	HIS	-	expression tag	UNP A1L314
H	636	HIS	-	expression tag	UNP A1L314
H	637	HIS	-	expression tag	UNP A1L314
H	638	HIS	-	expression tag	UNP A1L314
H	639	HIS	-	expression tag	UNP A1L314
H	640	HIS	-	expression tag	UNP A1L314

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 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.

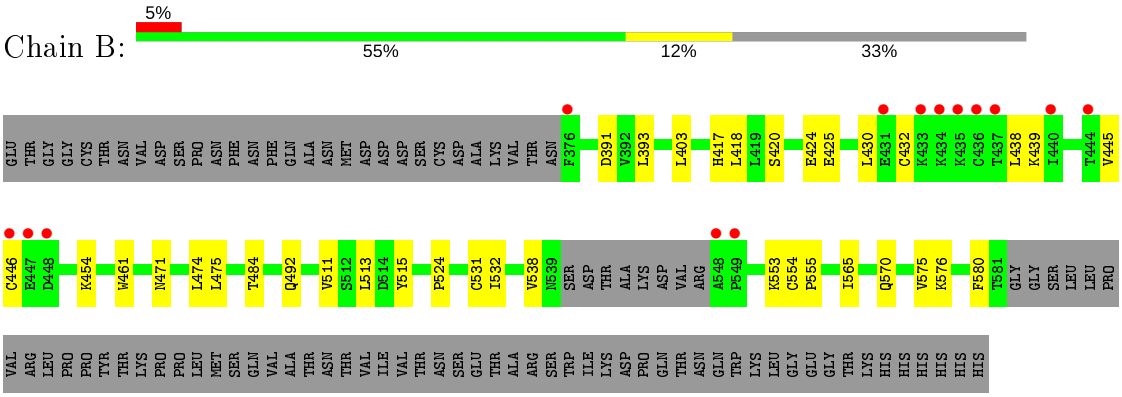
• Molecule 1: Macrophage-expressed gene 1 protein

Chain A:



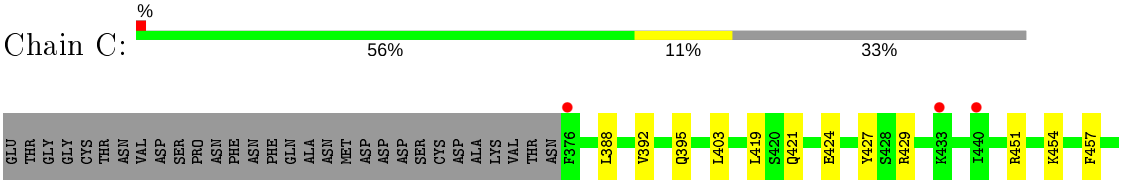
• Molecule 1: Macrophage-expressed gene 1 protein

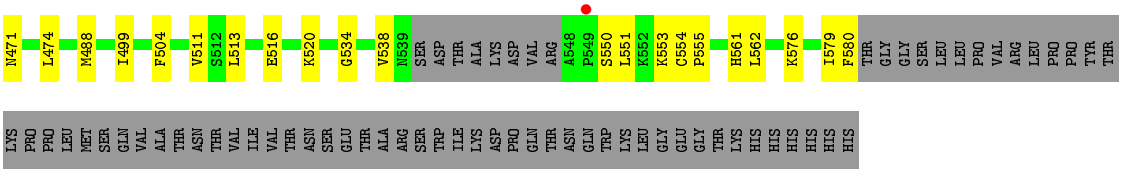
Chain B:



• Molecule 1: Macrophage-expressed gene 1 protein

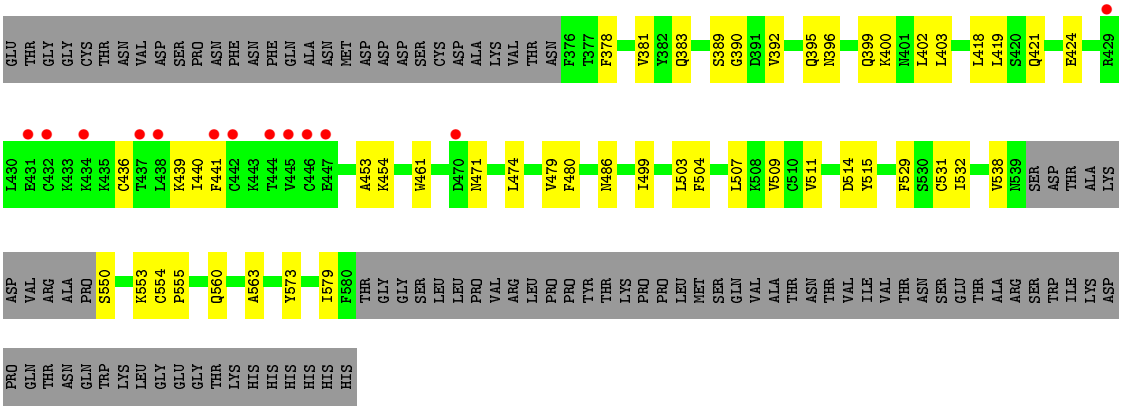
Chain C:





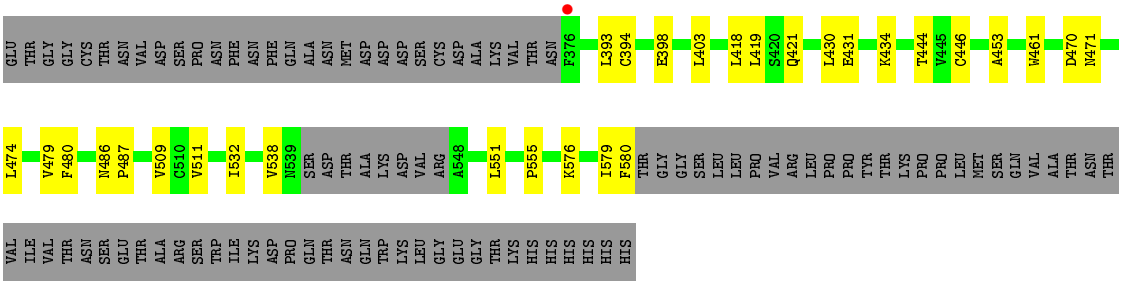
• Molecule 1: Macrophage-expressed gene 1 protein

Chain D:



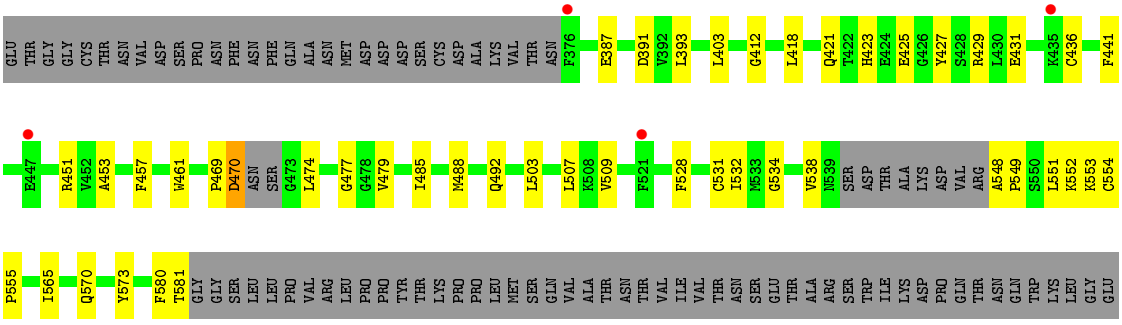
• Molecule 1: Macrophage-expressed gene 1 protein

Chain E:



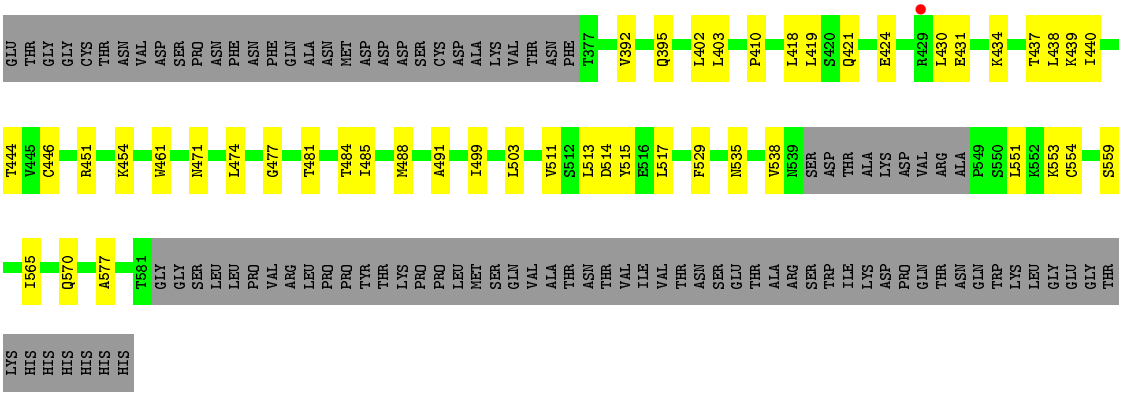
• Molecule 1: Macrophage-expressed gene 1 protein

Chain F:



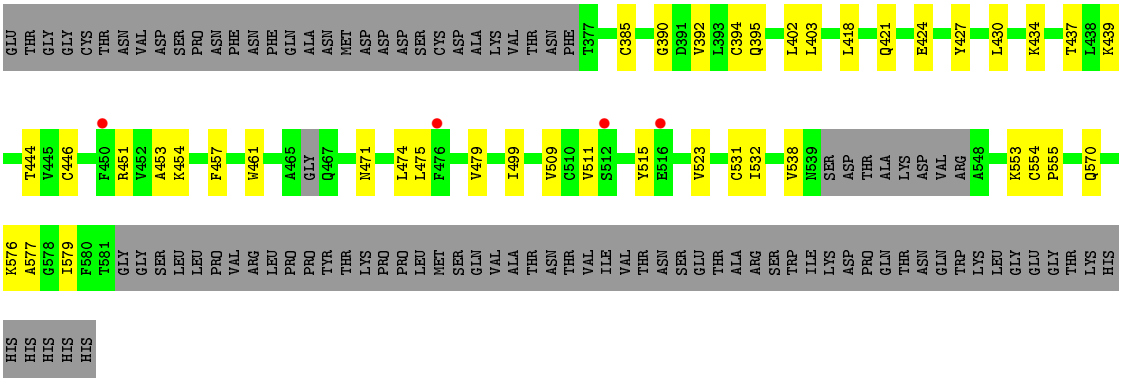
• Molecule 1: Macrophage-expressed gene 1 protein

Chain G:



• Molecule 1: Macrophage-expressed gene 1 protein

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	70.23Å 70.05Å 128.14Å 78.50° 79.05° 85.01°	Depositor
Resolution (Å)	56.09 – 3.17 65.36 – 3.17	Depositor EDS
% Data completeness (in resolution range)	95.3 (56.09-3.17) 95.4 (65.36-3.17)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.50 (at 3.19Å)	Xtriage
Refinement program	PHENIX dev_3488, PHENIX dev_3488	Depositor
R, R_{free}	0.189 , 0.225 0.188 , 0.225	Depositor DCC
R_{free} test set	1976 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å^2)	72.0	Xtriage
Anisotropy	0.664	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 57.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.016 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11992	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.01% 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

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	0.28	0/1528	0.54	0/2064
1	B	0.27	0/1548	0.52	0/2094
1	C	0.28	0/1541	0.52	0/2084
1	D	0.27	0/1528	0.53	0/2065
1	E	0.28	0/1541	0.54	0/2084
1	F	0.28	0/1534	0.52	0/2075
1	G	0.28	0/1531	0.53	0/2070
1	H	0.27	0/1531	0.52	0/2070
All	All	0.28	0/12282	0.53	0/16606

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

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	1493	0	1450	27	0
1	B	1511	0	1468	28	0
1	C	1504	0	1461	24	0
1	D	1492	0	1449	31	0
1	E	1504	0	1465	20	0
1	F	1497	0	1455	29	0
1	G	1495	0	1453	30	0
1	H	1496	0	1453	28	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11992	0	11654	201	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 8.

The worst 5 of 201 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:B:430:LEU:HD21	1:B:446:CYS:HB3	1.52	0.91
1:G:485:ILE:HD11	1:G:491:ALA:HA	1.54	0.88
1:C:516:GLU:O	1:C:520:LYS:NZ	2.15	0.80
1:H:427:TYR:OH	1:H:451:ARG:NH1	2.16	0.79
1:B:524:PRO:HG2	1:B:576:LYS:HD2	1.64	0.78

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	190/295 (64%)	182 (96%)	7 (4%)	1 (0%)	29 66
1	B	194/295 (66%)	189 (97%)	4 (2%)	1 (0%)	29 66
1	C	193/295 (65%)	186 (96%)	7 (4%)	0	100 100
1	D	191/295 (65%)	181 (95%)	7 (4%)	3 (2%)	9 41
1	E	193/295 (65%)	185 (96%)	8 (4%)	0	100 100
1	F	192/295 (65%)	184 (96%)	7 (4%)	1 (0%)	29 66
1	G	192/295 (65%)	182 (95%)	9 (5%)	1 (0%)	29 66
1	H	190/295 (64%)	182 (96%)	6 (3%)	2 (1%)	14 50
All	All	1535/2360 (65%)	1471 (96%)	55 (4%)	9 (1%)	25 63

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	579	ILE
1	H	390	GLY
1	G	439	LYS
1	H	577	ALA
1	A	439	LYS

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	A	167/255 (66%)	167 (100%)	0	100	100
1	B	169/255 (66%)	169 (100%)	0	100	100
1	C	168/255 (66%)	168 (100%)	0	100	100
1	D	167/255 (66%)	167 (100%)	0	100	100
1	E	168/255 (66%)	168 (100%)	0	100	100
1	F	167/255 (66%)	167 (100%)	0	100	100
1	G	168/255 (66%)	168 (100%)	0	100	100
1	H	168/255 (66%)	166 (99%)	2 (1%)	71	87
All	All	1342/2040 (66%)	1340 (100%)	2 (0%)	93	98

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	385	CYS
1	H	394	CYS

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

Mol	Chain	Res	Type
1	F	423	HIS
1	H	395	GLN

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	A	196/295 (66%)	-0.09	5 (2%) 56 40	41, 72, 144, 178	0
1	B	198/295 (67%)	0.08	14 (7%) 16 9	38, 73, 167, 205	0
1	C	197/295 (66%)	-0.06	4 (2%) 65 50	44, 74, 133, 196	0
1	D	195/295 (66%)	0.11	13 (6%) 17 10	46, 88, 183, 214	0
1	E	197/295 (66%)	-0.12	1 (0%) 91 86	43, 75, 145, 192	0
1	F	196/295 (66%)	0.00	4 (2%) 65 50	48, 90, 138, 166	0
1	G	196/295 (66%)	-0.16	1 (0%) 91 86	47, 85, 141, 152	0
1	H	196/295 (66%)	0.00	4 (2%) 65 50	53, 100, 149, 180	0
All	All	1571/2360 (66%)	-0.03	46 (2%) 51 35	38, 83, 150, 214	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	549	PRO	5.5
1	E	376	PHE	4.2
1	D	438	LEU	4.1
1	D	445	VAL	4.1
1	D	446	CYS	3.9

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.