wwPDB X-ray Structure Validation Summary Report (i)

Jul 6, 2022 – 01:09 pm BST

PDB ID	:	6ZC5
Title	:	Human Adenovirus serotype D10 FiberKnob protein
Authors	:	Baker, A.T.; Mundy, R.M.; Rizkallah, P.J.; Parker, A.L.
Deposited on	:	2020-06-09
Resolution	:	2.50 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:

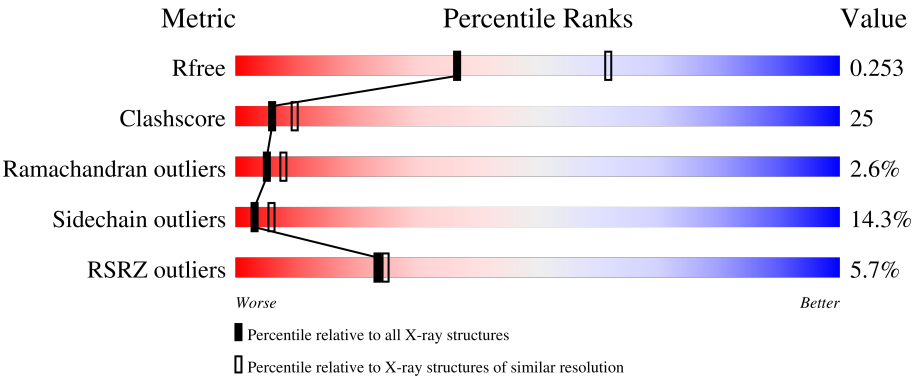
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	189	4% 56% 36% 5%
1	B	189	7% 61% 30%
1	C	189	12% 54% 38%
1	D	189	% 45% 40% 12%
1	E	189	4% 44% 40% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	189	7% 52% 41% 5%
1	G	189	3% 56% 34% 10%
1	H	189	2% 44% 40% 12%
1	I	189	5% 47% 40% 8%
1	J	189	10% 56% 37%
1	K	189	5% 50% 34% 11%
1	L	189	8% 56% 32% 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 17536 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 Fiber.

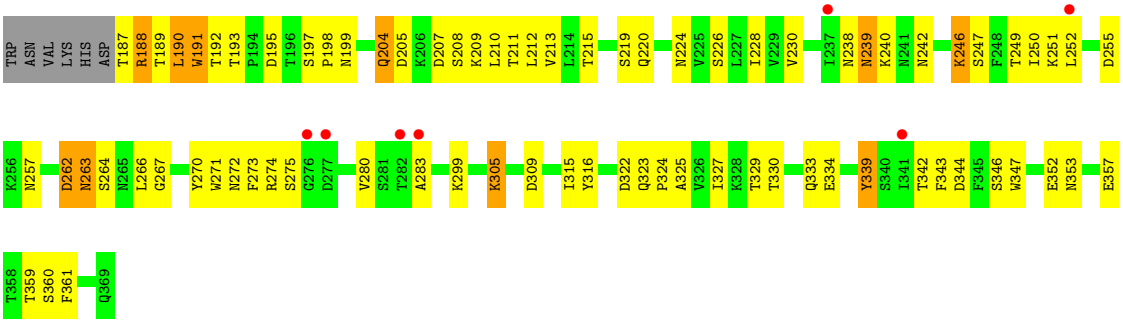
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	B	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	C	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	D	184	Total	C	N	O	S	0	0	0
			1458	933	235	286	4			
1	E	186	Total	C	N	O	S	0	0	0
			1477	945	240	288	4			
1	F	186	Total	C	N	O	S	0	0	0
			1477	945	240	288	4			
1	G	189	Total	C	N	O	S	0	0	0
			1506	965	245	292	4			
1	H	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	I	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	J	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			
1	K	185	Total	C	N	O	S	0	0	0
			1468	939	238	287	4			
1	L	183	Total	C	N	O	S	0	0	0
			1450	929	234	283	4			

3 Residue-property plots (i)

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.

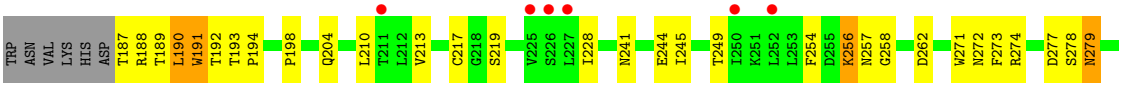
• Molecule 1: Fiber

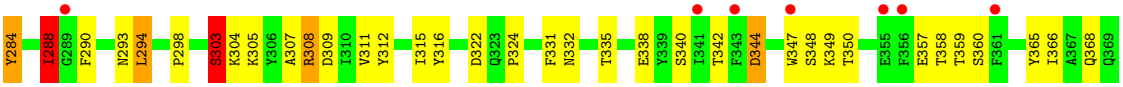
Chain A:



• Molecule 1: Fiber

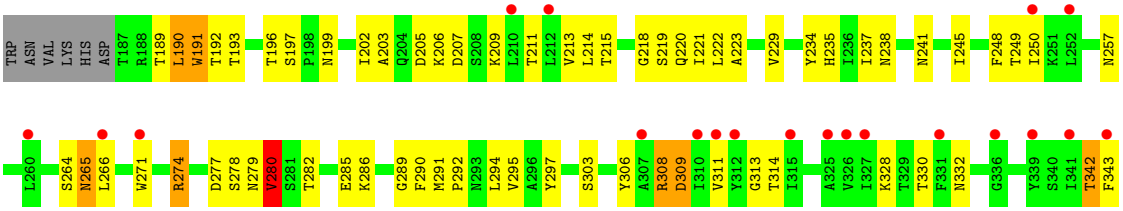
Chain B:





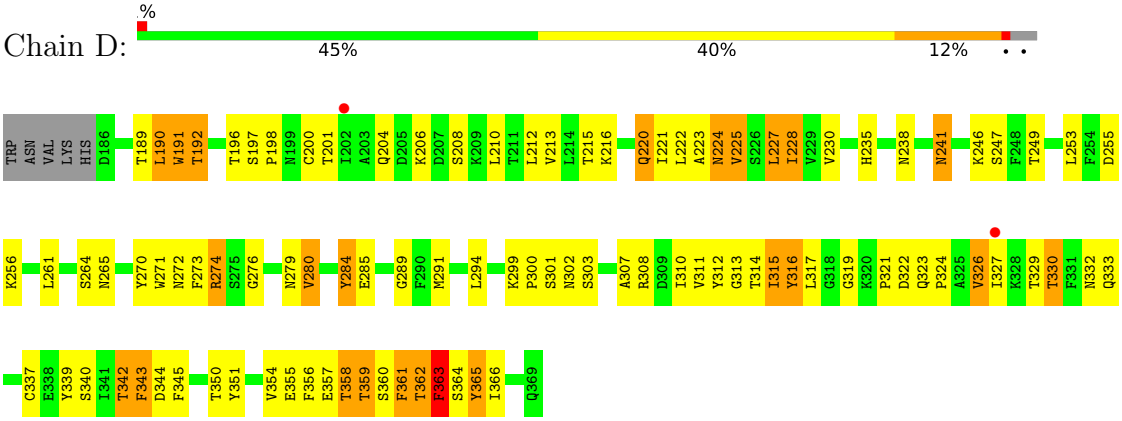
• Molecule 1: Fiber

Chain C:



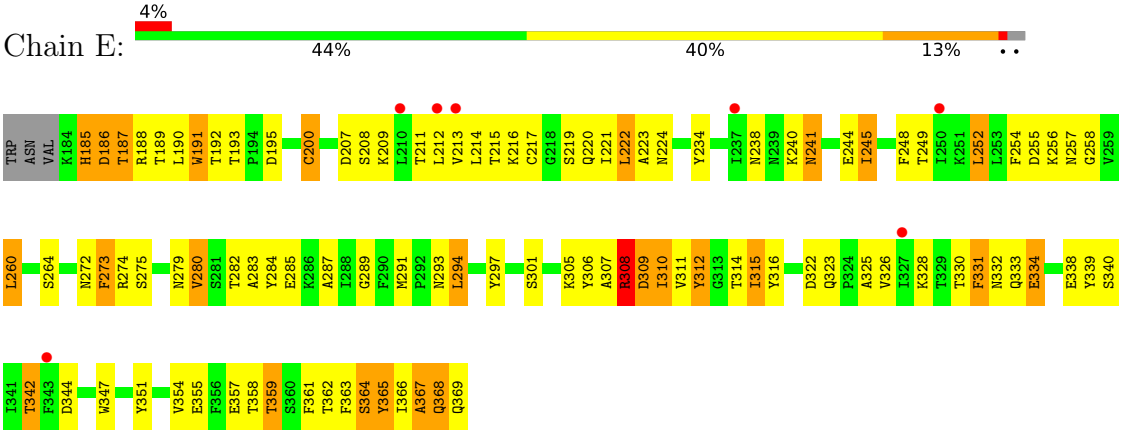
• Molecule 1: Fiber

Chain D:



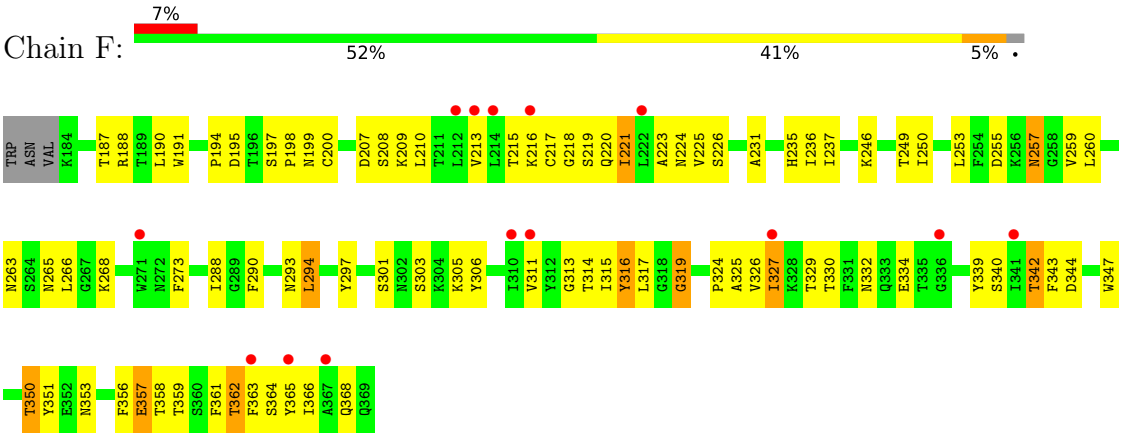
• Molecule 1: Fiber

Chain E:



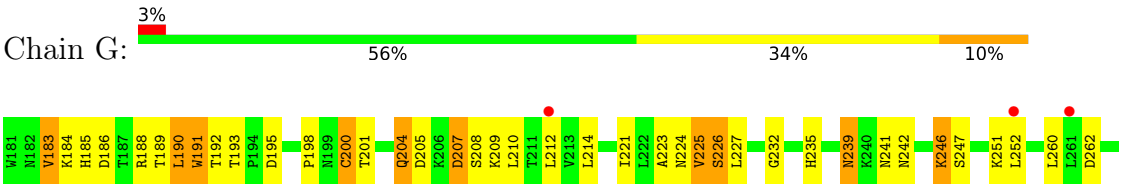
• Molecule 1: Fiber

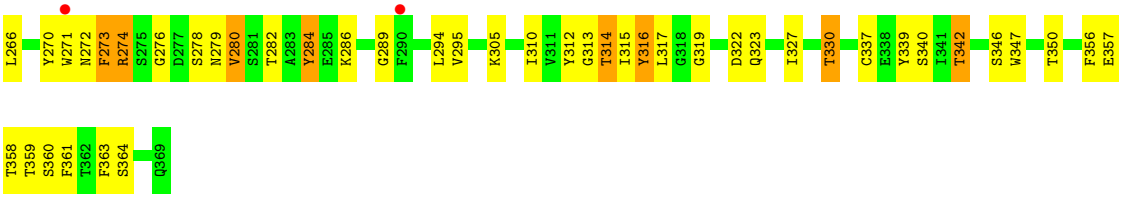
Chain F:



• Molecule 1: Fiber

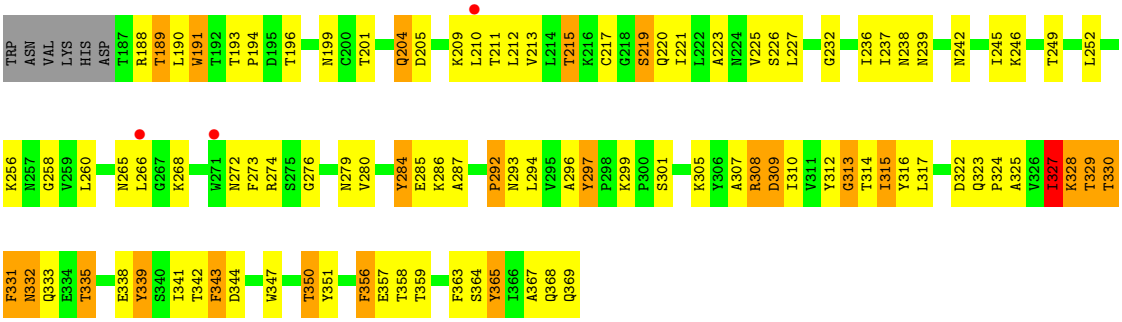
Chain G:





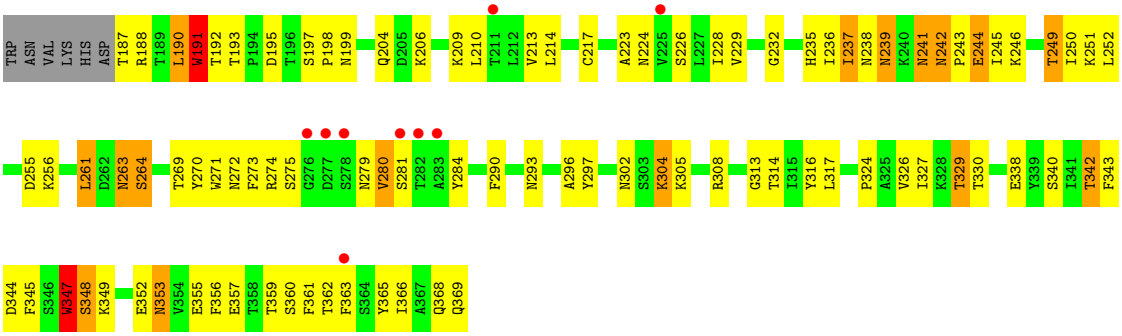
• Molecule 1: Fiber

Chain H:



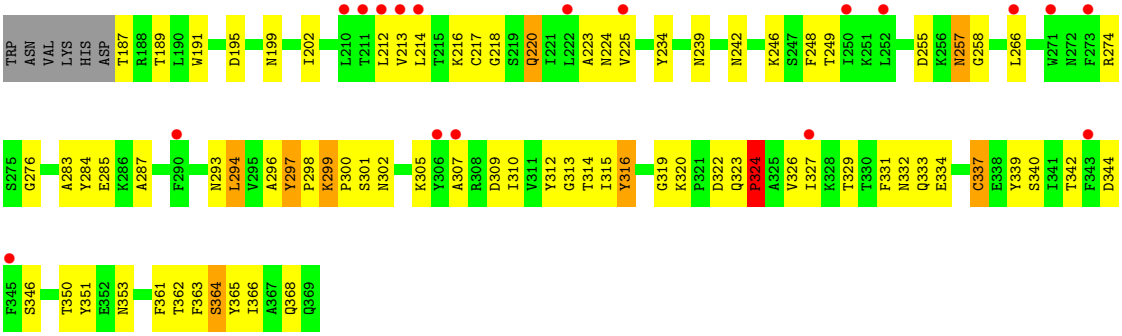
• Molecule 1: Fiber

Chain I:



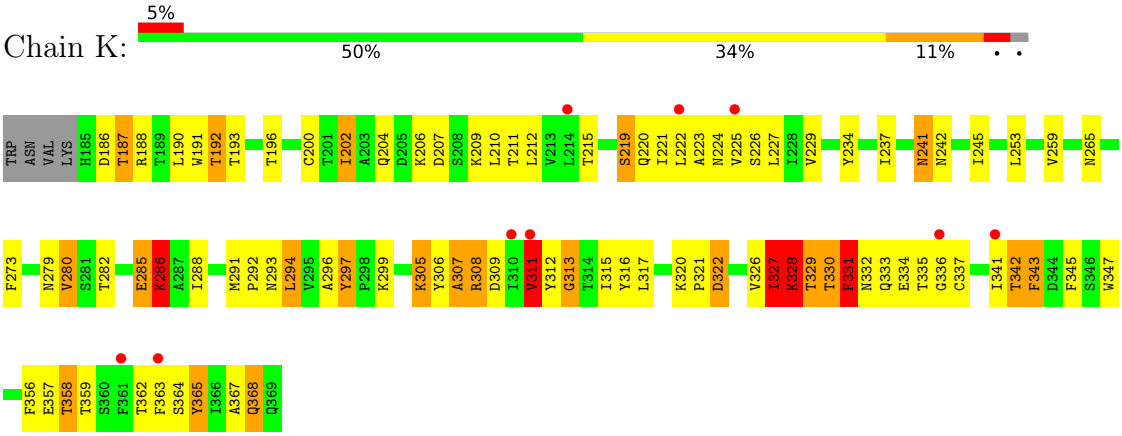
• Molecule 1: Fiber

Chain J:



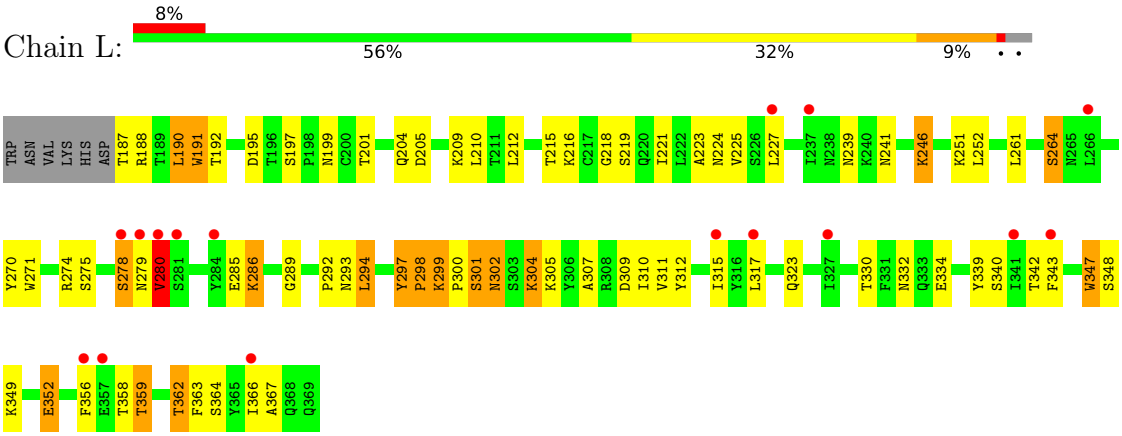
• Molecule 1: Fiber

Chain K:



• Molecule 1: Fiber

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	101.36Å 101.36Å 326.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	108.91 – 2.50 108.91 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.6 (108.91-2.50) 98.6 (108.91-2.50)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.64 (at 2.52Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.219 , 0.251 0.220 , 0.253	Depositor DCC
R_{free} test set	6341 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	54.6	Xtriage
Anisotropy	0.859	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.467 for -h,-k,l 0.477 for h,-h-k,-l 0.467 for -k,-h,-l	Xtriage
Reported twinning fraction	0.253 for H, K, L 0.254 for K, H, -L 0.245 for -h,-k,l 0.248 for -K, -H, -L	Depositor
Outliers	0 of 128203 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17536	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.91% 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	0.85	0/1484	0.86	0/2016
1	B	0.73	0/1484	0.78	0/2016
1	C	0.71	0/1484	0.75	0/2016
1	D	1.17	10/1492 (0.7%)	1.04	5/2027 (0.2%)
1	E	1.02	8/1512 (0.5%)	0.94	2/2053 (0.1%)
1	F	0.85	1/1512 (0.1%)	0.81	0/2053
1	G	0.86	3/1543 (0.2%)	0.84	0/2097
1	H	1.21	10/1484 (0.7%)	0.98	1/2016 (0.0%)
1	I	0.86	0/1484	0.85	0/2016
1	J	0.81	1/1484 (0.1%)	0.82	0/2016
1	K	1.13	7/1503 (0.5%)	0.92	0/2042
1	L	0.78	1/1484 (0.1%)	0.83	0/2016
All	All	0.93	41/17950 (0.2%)	0.87	8/24384 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	D	0	2
1	E	0	4
1	G	0	2
1	I	0	2
1	J	0	3
1	K	0	3
All	All	0	17

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	361	PHE	CG-CD1	-12.38	1.20	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	292	PRO	N-CD	-8.92	1.35	1.47
1	D	363	PHE	CG-CD1	-8.43	1.26	1.38
1	E	331	PHE	CG-CD1	-8.20	1.26	1.38
1	D	365	TYR	CE1-CZ	-7.97	1.28	1.38

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	308	ARG	NE-CZ-NH1	-7.53	116.53	120.30
1	D	316	TYR	O-C-N	5.76	131.91	122.70
1	D	358	THR	C-N-CA	5.39	135.18	121.70
1	D	359	THR	CA-CB-OG1	-5.35	97.76	109.00
1	E	308	ARG	CB-CA-C	5.35	121.11	110.40

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	283	ALA	Peptide
1	D	190	LEU	Peptide
1	D	363	PHE	Mainchain
1	E	185	HIS	Peptide
1	E	186	ASP	Peptide

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	1450	0	1430	67	0
1	B	1450	0	1430	70	0
1	C	1450	0	1430	53	1
1	D	1458	0	1434	105	0
1	E	1477	0	1454	107	0
1	F	1477	0	1454	83	0
1	G	1506	0	1478	77	0
1	H	1450	0	1429	129	0
1	I	1450	0	1430	75	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	1450	0	1430	46	0
1	K	1468	0	1441	95	1
1	L	1450	0	1430	69	0
All	All	17536	0	17270	861	1

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 25.

The worst 5 of 861 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:H:315:ILE:HD13	1:H:358:THR:OG1	1.15	1.27
1:E:222:LEU:O	1:F:220:GLN:NE2	1.66	1.26
1:D:364:SER:OG	1:F:362:THR:HG21	1.33	1.23
1:H:313:GLY:O	1:H:327:ILE:N	1.68	1.22
1:K:309:ASP:OD2	1:K:333:GLN:NE2	1.74	1.20

All (1) 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:C:202:ILE:O	1:K:336:GLY:O[3_554]	2.11	0.09

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	181/189 (96%)	158 (87%)	19 (10%)	4 (2%)	6 10
1	B	181/189 (96%)	153 (84%)	25 (14%)	3 (2%)	9 16
1	C	181/189 (96%)	155 (86%)	20 (11%)	6 (3%)	4 5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	182/189 (96%)	150 (82%)	27 (15%)	5 (3%)	5 7
1	E	184/189 (97%)	150 (82%)	27 (15%)	7 (4%)	3 4
1	F	184/189 (97%)	154 (84%)	27 (15%)	3 (2%)	9 17
1	G	187/189 (99%)	153 (82%)	27 (14%)	7 (4%)	3 4
1	H	181/189 (96%)	156 (86%)	24 (13%)	1 (1%)	25 43
1	I	181/189 (96%)	150 (83%)	26 (14%)	5 (3%)	5 7
1	J	181/189 (96%)	152 (84%)	25 (14%)	4 (2%)	6 10
1	K	183/189 (97%)	152 (83%)	25 (14%)	6 (3%)	4 5
1	L	181/189 (96%)	152 (84%)	24 (13%)	5 (3%)	5 7
All	All	2187/2268 (96%)	1835 (84%)	296 (14%)	56 (3%)	5 8

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	280	VAL
1	D	280	VAL
1	E	280	VAL
1	E	284	TYR
1	K	280	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	A	164/170 (96%)	149 (91%)	15 (9%)	9 18
1	B	164/170 (96%)	138 (84%)	26 (16%)	2 4
1	C	164/170 (96%)	141 (86%)	23 (14%)	3 6
1	D	165/170 (97%)	146 (88%)	19 (12%)	5 11
1	E	167/170 (98%)	143 (86%)	24 (14%)	3 6
1	F	167/170 (98%)	148 (89%)	19 (11%)	5 11
1	G	170/170 (100%)	142 (84%)	28 (16%)	2 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	H	164/170 (96%)	136 (83%)	28 (17%)	2 3
1	I	164/170 (96%)	137 (84%)	27 (16%)	2 4
1	J	164/170 (96%)	143 (87%)	21 (13%)	4 8
1	K	166/170 (98%)	134 (81%)	32 (19%)	1 2
1	L	164/170 (96%)	143 (87%)	21 (13%)	4 8
All	All	1983/2040 (97%)	1700 (86%)	283 (14%)	3 6

5 of 283 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	187	THR
1	K	219	SER
1	K	365	TYR
1	E	312	TYR
1	E	301	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	353	ASN
1	L	241	ASN
1	H	235	HIS
1	L	239	ASN
1	J	333	GLN

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 monosaccharides 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, 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	183/189 (96%)	0.16	7 (3%) 40 43	30, 74, 122, 175	0
1	B	183/189 (96%)	0.14	13 (7%) 16 16	44, 108, 144, 201	0
1	C	183/189 (96%)	0.47	22 (12%) 4 4	67, 100, 140, 174	0
1	D	184/189 (97%)	0.03	2 (1%) 80 82	8, 70, 105, 147	0
1	E	186/189 (98%)	-0.16	7 (3%) 40 43	20, 87, 131, 150	0
1	F	186/189 (98%)	0.25	14 (7%) 14 14	30, 86, 114, 131	0
1	G	189/189 (100%)	-0.08	5 (2%) 56 59	30, 100, 139, 172	0
1	H	183/189 (96%)	-0.10	3 (1%) 72 74	9, 73, 112, 134	0
1	I	183/189 (96%)	0.21	9 (4%) 29 31	37, 78, 121, 179	0
1	J	183/189 (96%)	0.11	18 (9%) 7 7	30, 99, 139, 169	0
1	K	185/189 (97%)	-0.03	9 (4%) 29 31	8, 82, 113, 136	0
1	L	183/189 (96%)	0.38	16 (8%) 10 10	37, 83, 124, 156	0
All	All	2211/2268 (97%)	0.11	125 (5%) 23 25	8, 87, 134, 201	0

The worst 5 of 125 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	325	ALA	9.3
1	A	282	THR	8.2
1	B	361	PHE	7.9
1	J	213	VAL	6.4
1	L	356	PHE	6.1

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 monosaccharides 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.