wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 02:21 AM EDT

PDB ID	:	6GCM
Title	:	Escherichia coli DPS
Authors	:	Kovalenko, V.V.; Loiko, N.G.; Tereshkin, E.V.; Tereshkina, K.B.; Chulichkov, A.L.; Popov, A.N.; Krupyanskii, Y.F.
Deposited on	:	2018-04-18
Resolution	:	2.45 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 (1)) were used in the production of this report:

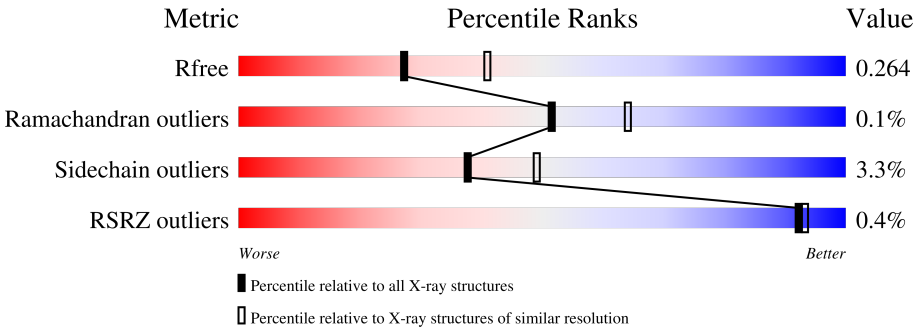
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.45 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	1544 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	159	96%
1	b	159	3% 96%
2	B	154	98%
2	C	154	96%
2	E	154	2% 97%
2	F	154	% 97%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	154	99%
2	H	154	3% 96%
2	I	154	97%
2	J	154	99%
2	K	154	% 94% 6%
2	L	154	97%
2	c	154	97%
2	d	154	98%
2	g	154	97%
2	h	154	96%
2	i	154	97%
2	j	154	97%
2	k	154	97%
2	l	154	97%
2	m	154	97%
3	D	156	% 97%
3	e	156	% 96%
4	f	151	97%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 30236 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 DNA protection during starvation protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	155	Total	C	N	O	S	0	0	0
			1226	771	214	238	3			
1	b	159	Total	C	N	O	S	0	0	0
			1252	785	219	245	3			

- Molecule 2 is a protein called DNA protection during starvation protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	154	Total	C	N	O	S	0	0	0
			1214	763	211	237	3			
2	C	154	Total	C	N	O	S	0	0	0
			1218	766	212	237	3			
2	E	154	Total	C	N	O	S	0	0	0
			1219	767	212	237	3			
2	F	154	Total	C	N	O	S	0	0	0
			1219	767	212	237	3			
2	G	154	Total	C	N	O	S	0	0	0
			1217	765	212	237	3			
2	H	154	Total	C	N	O	S	0	0	0
			1219	767	212	237	3			
2	I	154	Total	C	N	O	S	0	0	0
			1216	765	211	237	3			
2	J	154	Total	C	N	O	S	0	0	0
			1220	768	212	237	3			
2	K	154	Total	C	N	O	S	0	0	0
			1221	768	213	237	3			
2	L	154	Total	C	N	O	S	0	0	0
			1217	766	211	237	3			
2	c	154	Total	C	N	O	S	0	0	0
			1214	763	211	237	3			
2	d	154	Total	C	N	O	S	0	0	0
			1218	766	212	237	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	g	154	Total	C	N	O	S	0	0	0
			1219	767	212	237	3			
2	h	154	Total	C	N	O	S	0	0	0
			1217	765	212	237	3			
2	i	154	Total	C	N	O	S	0	0	0
			1219	767	212	237	3			
2	j	154	Total	C	N	O	S	0	0	0
			1216	765	211	237	3			
2	k	154	Total	C	N	O	S	0	0	0
			1220	768	212	237	3			
2	l	154	Total	C	N	O	S	0	0	0
			1221	768	213	237	3			
2	m	154	Total	C	N	O	S	0	0	0
			1217	766	211	237	3			

- Molecule 3 is a protein called DNA protection during starvation protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	156	Total	C	N	O	S	0	0	0
			1236	776	216	241	3			
3	e	156	Total	C	N	O	S	0	0	0
			1236	776	216	241	3			

- Molecule 4 is a protein called DNA protection during starvation protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	f	151	Total	C	N	O	S	0	0	0
			1191	746	209	233	3			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	40	Total	O	0	0
			40	40		
5	B	37	Total	O	0	0
			37	37		
5	C	36	Total	O	0	0
			36	36		
5	D	41	Total	O	0	0
			41	41		
5	E	47	Total	O	0	0
			47	47		

Continued on next page...

Continued from previous page...

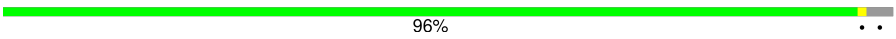
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	F	26	Total O 26 26	0	0
5	G	43	Total O 43 43	0	0
5	H	40	Total O 40 40	0	0
5	I	34	Total O 34 34	0	0
5	J	42	Total O 42 42	0	0
5	K	34	Total O 34 34	0	0
5	L	38	Total O 38 38	0	0
5	b	52	Total O 52 52	0	0
5	c	39	Total O 39 39	0	0
5	d	40	Total O 40 40	0	0
5	e	24	Total O 24 24	0	0
5	f	42	Total O 42 42	0	0
5	g	28	Total O 28 28	0	0
5	h	42	Total O 42 42	0	0
5	i	45	Total O 45 45	0	0
5	j	48	Total O 48 48	0	0
5	k	44	Total O 44 44	0	0
5	l	42	Total O 42 42	0	0
5	m	50	Total O 50 50	0	0

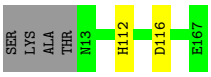
3 Residue-property plots (i)

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: DNA protection during starvation protein

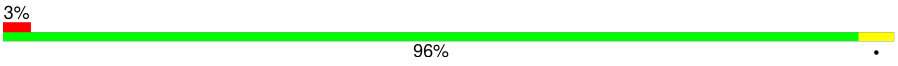
Chain A:

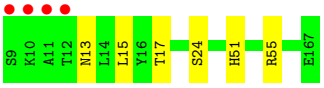




- Molecule 1: DNA protection during starvation protein

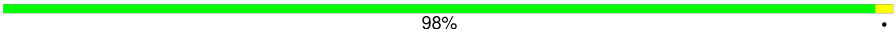
Chain b:

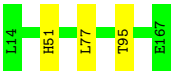




- Molecule 2: DNA protection during starvation protein

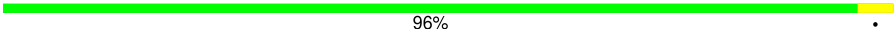
Chain B:





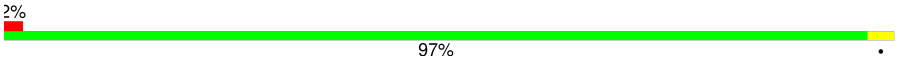
- Molecule 2: DNA protection during starvation protein

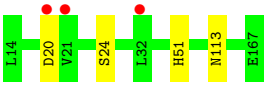
Chain C:



- Molecule 2: DNA protection during starvation protein

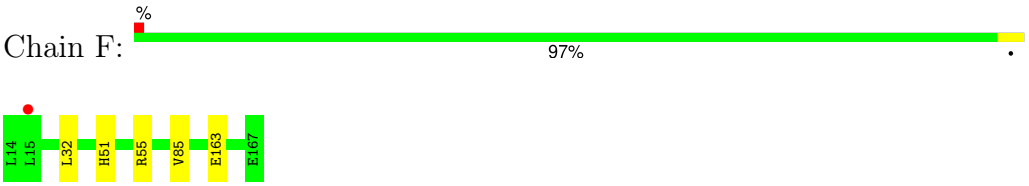
Chain E:





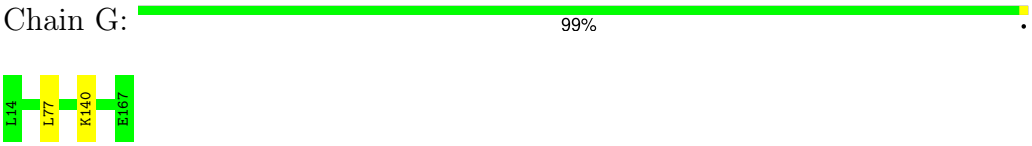
- Molecule 2: DNA protection during starvation protein

Chain F:



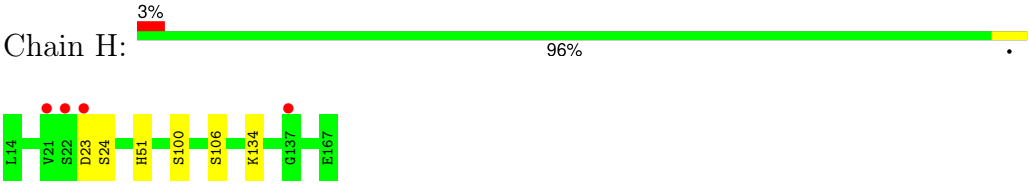
• Molecule 2: DNA protection during starvation protein

Chain G:



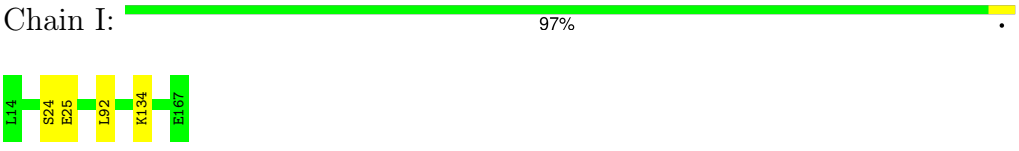
• Molecule 2: DNA protection during starvation protein

Chain H:



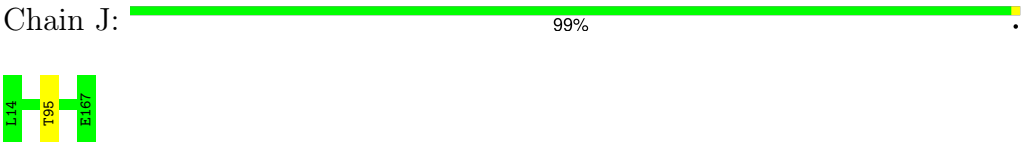
• Molecule 2: DNA protection during starvation protein

Chain I:



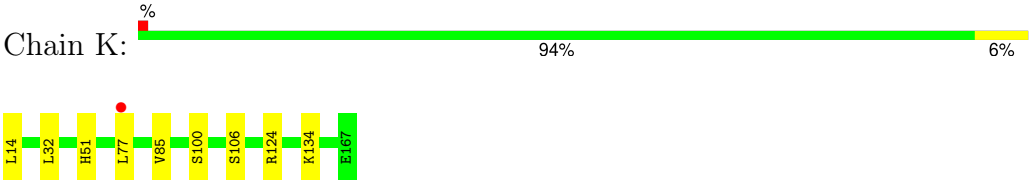
• Molecule 2: DNA protection during starvation protein

Chain J:



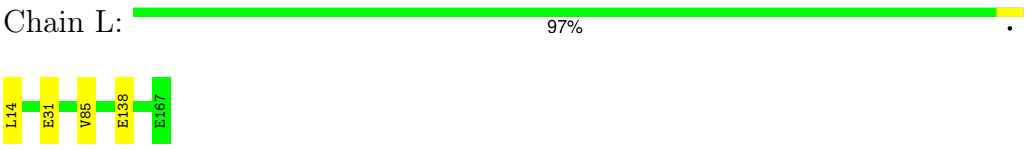
• Molecule 2: DNA protection during starvation protein

Chain K:



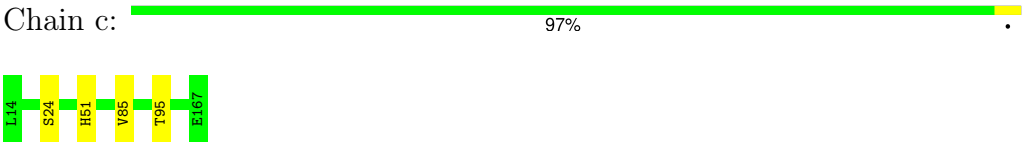
• Molecule 2: DNA protection during starvation protein

Chain L:



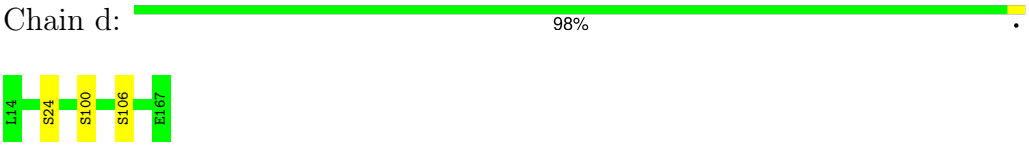
• Molecule 2: DNA protection during starvation protein

Chain c:



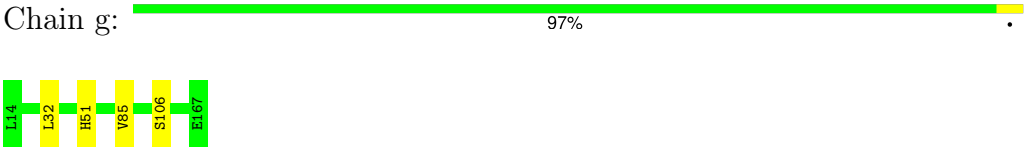
• Molecule 2: DNA protection during starvation protein

Chain d:



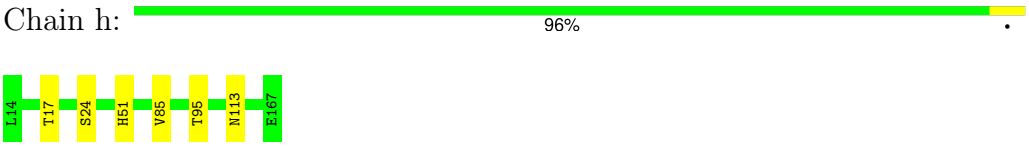
• Molecule 2: DNA protection during starvation protein

Chain g:



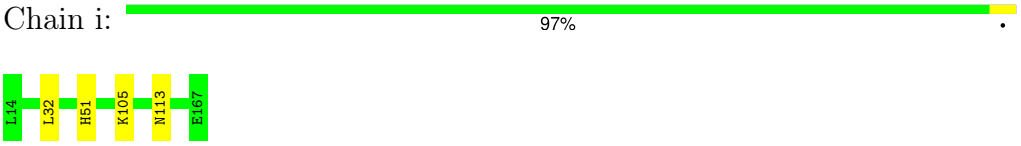
• Molecule 2: DNA protection during starvation protein

Chain h:



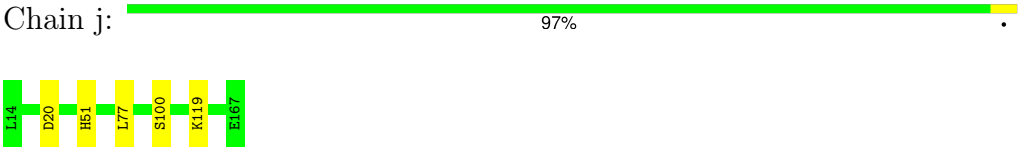
• Molecule 2: DNA protection during starvation protein

Chain i:



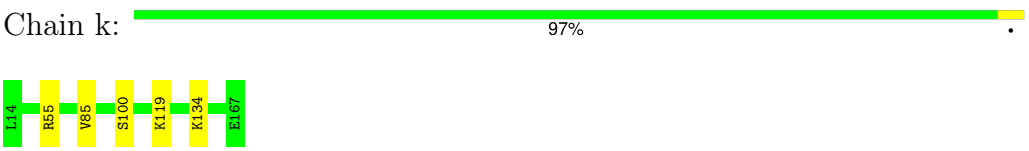
• Molecule 2: DNA protection during starvation protein

Chain j:



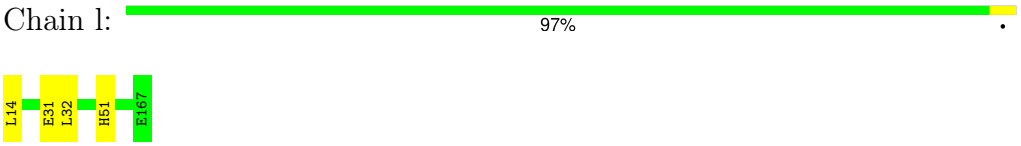
• Molecule 2: DNA protection during starvation protein

Chain k:



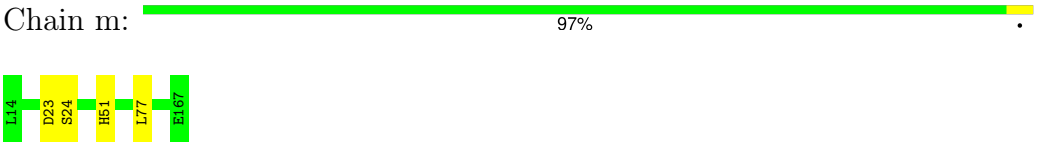
• Molecule 2: DNA protection during starvation protein

Chain l:



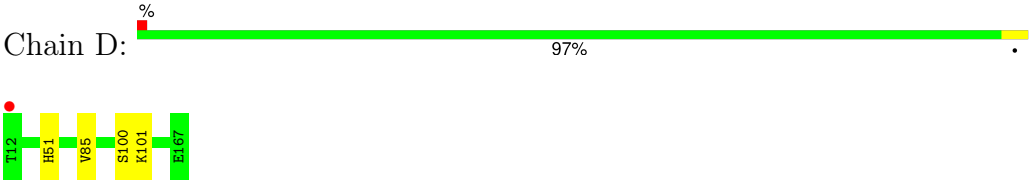
• Molecule 2: DNA protection during starvation protein

Chain m:



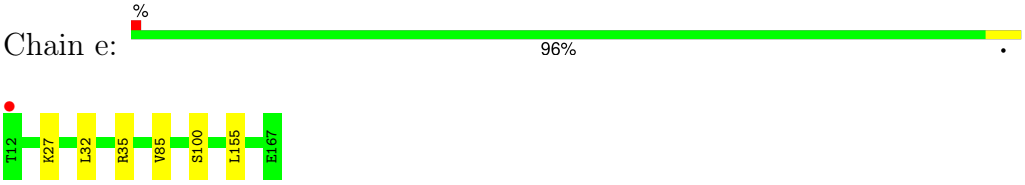
• Molecule 3: DNA protection during starvation protein

Chain D:



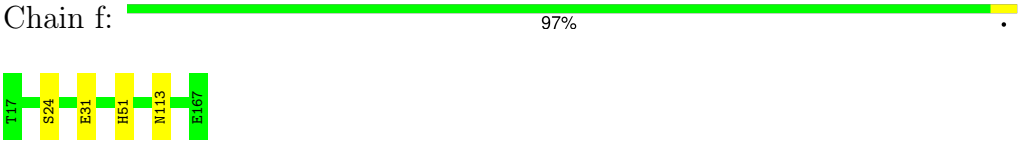
• Molecule 3: DNA protection during starvation protein

Chain e:



• Molecule 4: DNA protection during starvation protein

Chain f:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	150.15Å 89.42Å 150.65Å 90.00° 90.24° 90.00°	Depositor
Resolution (Å)	100.02 – 2.45 47.68 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.8 (100.02-2.45) 99.8 (47.68-2.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.203 , 0.258 0.210 , 0.264	Depositor DCC
R_{free} test set	7407 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	30.4	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 19.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.014 for l,k,-h 0.027 for h,-k,-l 0.022 for l,-k,h	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	30236	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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	0.73	0/1244	0.80	0/1684
1	b	0.73	0/1270	0.84	0/1720
2	B	0.70	0/1232	0.82	0/1670
2	C	0.70	0/1236	0.83	0/1674
2	E	0.71	0/1237	0.80	0/1675
2	F	0.65	0/1237	0.78	0/1675
2	G	0.68	0/1234	0.79	0/1672
2	H	0.63	0/1237	0.81	0/1675
2	I	0.66	0/1234	0.78	0/1672
2	J	0.66	0/1238	0.77	0/1676
2	K	0.65	0/1239	0.81	0/1677
2	L	0.64	0/1235	0.80	0/1673
2	c	0.68	0/1232	0.77	0/1670
2	d	0.68	0/1236	0.84	0/1674
2	g	0.65	0/1237	0.78	0/1675
2	h	0.69	0/1234	0.81	0/1672
2	i	0.67	0/1237	0.85	0/1675
2	j	0.74	0/1234	0.83	0/1672
2	k	0.69	0/1238	0.83	0/1676
2	l	0.71	0/1239	0.82	0/1677
2	m	0.67	0/1235	0.79	0/1673
3	D	0.66	0/1254	0.78	0/1698
3	e	0.70	0/1254	0.82	0/1698
4	f	0.72	0/1208	0.80	0/1635
All	All	0.68	0/29711	0.81	0/40238

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

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	153/159 (96%)	149 (97%)	4 (3%)	0	100	100
1	b	157/159 (99%)	153 (98%)	4 (2%)	0	100	100
2	B	152/154 (99%)	148 (97%)	4 (3%)	0	100	100
2	C	152/154 (99%)	148 (97%)	4 (3%)	0	100	100
2	E	152/154 (99%)	146 (96%)	6 (4%)	0	100	100
2	F	152/154 (99%)	145 (95%)	7 (5%)	0	100	100
2	G	152/154 (99%)	151 (99%)	1 (1%)	0	100	100
2	H	152/154 (99%)	147 (97%)	3 (2%)	2 (1%)	12	11
2	I	152/154 (99%)	145 (95%)	7 (5%)	0	100	100
2	J	152/154 (99%)	146 (96%)	6 (4%)	0	100	100
2	K	152/154 (99%)	146 (96%)	6 (4%)	0	100	100
2	L	152/154 (99%)	152 (100%)	0	0	100	100
2	c	152/154 (99%)	148 (97%)	4 (3%)	0	100	100
2	d	152/154 (99%)	148 (97%)	4 (3%)	0	100	100
2	g	152/154 (99%)	147 (97%)	5 (3%)	0	100	100
2	h	152/154 (99%)	151 (99%)	1 (1%)	0	100	100
2	i	152/154 (99%)	150 (99%)	2 (1%)	0	100	100
2	j	152/154 (99%)	148 (97%)	4 (3%)	0	100	100
2	k	152/154 (99%)	150 (99%)	2 (1%)	0	100	100
2	l	152/154 (99%)	146 (96%)	6 (4%)	0	100	100
2	m	152/154 (99%)	148 (97%)	4 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	154/156 (99%)	150 (97%)	4 (3%)	0	100	100
3	e	154/156 (99%)	152 (99%)	2 (1%)	0	100	100
4	f	149/151 (99%)	144 (97%)	5 (3%)	0	100	100
All	All	3655/3707 (99%)	3558 (97%)	95 (3%)	2 (0%)	51	64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	23	ASP
2	H	24	SER

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	132/136 (97%)	130 (98%)	2 (2%)	65	76
1	b	135/136 (99%)	129 (96%)	6 (4%)	28	37
2	B	130/132 (98%)	127 (98%)	3 (2%)	50	63
2	C	131/132 (99%)	125 (95%)	6 (5%)	27	35
2	E	131/132 (99%)	127 (97%)	4 (3%)	40	52
2	F	131/132 (99%)	126 (96%)	5 (4%)	33	43
2	G	130/132 (98%)	128 (98%)	2 (2%)	65	76
2	H	131/132 (99%)	127 (97%)	4 (3%)	40	52
2	I	130/132 (98%)	126 (97%)	4 (3%)	40	52
2	J	131/132 (99%)	130 (99%)	1 (1%)	81	88
2	K	132/132 (100%)	123 (93%)	9 (7%)	16	19
2	L	130/132 (98%)	126 (97%)	4 (3%)	40	52
2	c	130/132 (98%)	126 (97%)	4 (3%)	40	52
2	d	131/132 (99%)	128 (98%)	3 (2%)	50	63
2	g	131/132 (99%)	127 (97%)	4 (3%)	40	52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	h	130/132 (98%)	124 (95%)	6 (5%)	27	35
2	i	131/132 (99%)	127 (97%)	4 (3%)	40	52
2	j	130/132 (98%)	125 (96%)	5 (4%)	33	43
2	k	131/132 (99%)	126 (96%)	5 (4%)	33	43
2	l	132/132 (100%)	128 (97%)	4 (3%)	41	52
2	m	130/132 (98%)	126 (97%)	4 (3%)	40	52
3	D	134/134 (100%)	130 (97%)	4 (3%)	41	52
3	e	134/134 (100%)	128 (96%)	6 (4%)	27	36
4	f	128/129 (99%)	124 (97%)	4 (3%)	40	52
All	All	3146/3177 (99%)	3043 (97%)	103 (3%)	38	49

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	c	95	THR
2	g	32	LEU
2	m	23	ASP
2	d	100	SER
3	e	100	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 84 such sidechains are listed below:

Mol	Chain	Res	Type
3	e	115	GLN
2	i	115	GLN
4	f	99	ASN
2	g	115	GLN
2	j	117	HIS

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	155/159 (97%)	-0.42	0 100 100	23, 34, 53, 66	0
1	b	159/159 (100%)	-0.28	4 (2%) 57 53	24, 32, 55, 162	0
2	B	154/154 (100%)	-0.46	0 100 100	27, 35, 49, 66	0
2	C	154/154 (100%)	-0.42	0 100 100	23, 35, 52, 68	0
2	E	154/154 (100%)	-0.43	3 (1%) 66 64	23, 30, 45, 63	0
2	F	154/154 (100%)	-0.31	1 (0%) 89 89	30, 39, 53, 69	0
2	G	154/154 (100%)	-0.52	0 100 100	21, 32, 46, 72	0
2	H	154/154 (100%)	-0.33	4 (2%) 56 52	29, 39, 54, 67	0
2	I	154/154 (100%)	-0.43	0 100 100	24, 34, 51, 65	0
2	J	154/154 (100%)	-0.40	0 100 100	30, 38, 53, 65	0
2	K	154/154 (100%)	-0.41	1 (0%) 89 89	27, 38, 54, 69	0
2	L	154/154 (100%)	-0.44	0 100 100	26, 35, 48, 68	0
2	c	154/154 (100%)	-0.44	0 100 100	27, 36, 50, 70	0
2	d	154/154 (100%)	-0.53	0 100 100	24, 32, 44, 56	0
2	g	154/154 (100%)	-0.44	0 100 100	28, 36, 51, 66	0
2	h	154/154 (100%)	-0.49	0 100 100	25, 34, 48, 63	0
2	i	154/154 (100%)	-0.48	0 100 100	24, 33, 46, 57	0
2	j	154/154 (100%)	-0.45	0 100 100	24, 31, 43, 57	0
2	k	154/154 (100%)	-0.51	0 100 100	25, 33, 46, 57	0
2	l	154/154 (100%)	-0.51	0 100 100	24, 32, 50, 65	0
2	m	154/154 (100%)	-0.53	0 100 100	23, 33, 47, 66	0
3	D	156/156 (100%)	-0.47	1 (0%) 89 89	25, 36, 54, 102	0
3	e	156/156 (100%)	-0.42	1 (0%) 89 89	26, 37, 56, 109	0
4	f	151/151 (100%)	-0.53	0 100 100	24, 33, 47, 56	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	3703/3707 (99%)	-0.44	15 (0%)	92	93	21, 34, 51, 162	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	b	11	ALA	10.9
3	e	12	THR	10.3
1	b	12	THR	8.6
2	H	22	SER	7.0
1	b	9	SER	6.7

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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.