



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 14, 2024 – 07:13 am GMT

PDB ID : 6QLB  
Title : Calpain small subunit 1, RNA-binding protein Hfq  
Authors : Rizkallah, P.J.; Cresser-Brown, J.O.  
Deposited on : 2019-01-31  
Resolution : 2.32 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
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

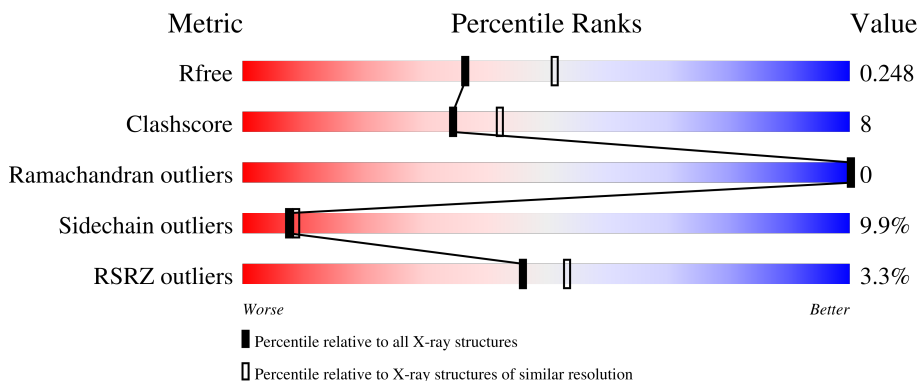
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.32 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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  $\geq 3$ , 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	173	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 82%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 14%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">3%      82%      14%      . .</p>
1	B	173	<div style="display: flex; align-items: center;"> <div style="width: 0%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 80%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">%      80%      17%      .</p>
1	C	173	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 79%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">3%      79%      17%      .</p>
1	D	173	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 83%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 15%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">4%      83%      15%      .</p>
2	E	63	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 79%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 19%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">5%      79%      19%      .</p>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	63	<p>3% 71% 25%</p>
2	G	63	<p>2% 86% 10% 5%</p>
2	H	63	<p>6% 78% 13% 10%</p>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GUN	H	101	-	-	X	-
5	EDO	A	207	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7880 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 Calpain small subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	172	1394	880	241	263	10	0	0	0
1	B	173	1413	891	245	267	10	0	1	0
1	C	173	1413	891	245	267	10	0	1	0
1	D	173	1403	885	242	266	10	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	LYS	ARG	conflict	UNP P04632
A	11	VAL	ALA	conflict	UNP P04632
A	78	LYS	ARG	conflict	UNP P04632
A	81	GLY	ALA	conflict	UNP P04632
A	85	ARG	GLN	conflict	UNP P04632
A	95	GLY	CYS	conflict	UNP P04632
A	97	ASN	SER	conflict	UNP P04632
A	112	GLN	GLU	conflict	UNP P04632
A	114	ILE	LEU	conflict	UNP P04632
A	116	SER	ASN	conflict	UNP P04632
A	126	THR	SER	conflict	UNP P04632
A	149	ARG	LYS	conflict	UNP P04632
A	154	ASN	ASP	conflict	UNP P04632
B	8	LYS	ARG	conflict	UNP P04632
B	11	VAL	ALA	conflict	UNP P04632
B	78	LYS	ARG	conflict	UNP P04632
B	81	GLY	ALA	conflict	UNP P04632
B	85	ARG	GLN	conflict	UNP P04632
B	95	GLY	CYS	conflict	UNP P04632
B	97	ASN	SER	conflict	UNP P04632
B	112	GLN	GLU	conflict	UNP P04632

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	114	ILE	LEU	conflict	UNP P04632
B	116	SER	ASN	conflict	UNP P04632
B	126	THR	SER	conflict	UNP P04632
B	149	ARG	LYS	conflict	UNP P04632
B	154	ASN	ASP	conflict	UNP P04632
C	8	LYS	ARG	conflict	UNP P04632
C	11	VAL	ALA	conflict	UNP P04632
C	78	LYS	ARG	conflict	UNP P04632
C	81	GLY	ALA	conflict	UNP P04632
C	85	ARG	GLN	conflict	UNP P04632
C	95	GLY	CYS	conflict	UNP P04632
C	97	ASN	SER	conflict	UNP P04632
C	112	GLN	GLU	conflict	UNP P04632
C	114	ILE	LEU	conflict	UNP P04632
C	116	SER	ASN	conflict	UNP P04632
C	126	THR	SER	conflict	UNP P04632
C	149	ARG	LYS	conflict	UNP P04632
C	154	ASN	ASP	conflict	UNP P04632
D	8	LYS	ARG	conflict	UNP P04632
D	11	VAL	ALA	conflict	UNP P04632
D	78	LYS	ARG	conflict	UNP P04632
D	81	GLY	ALA	conflict	UNP P04632
D	85	ARG	GLN	conflict	UNP P04632
D	95	GLY	CYS	conflict	UNP P04632
D	97	ASN	SER	conflict	UNP P04632
D	112	GLN	GLU	conflict	UNP P04632
D	114	ILE	LEU	conflict	UNP P04632
D	116	SER	ASN	conflict	UNP P04632
D	126	THR	SER	conflict	UNP P04632
D	149	ARG	LYS	conflict	UNP P04632
D	154	ASN	ASP	conflict	UNP P04632

- Molecule 2 is a protein called RNA-binding protein Hfq.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	63	Total	C	N	O	S	0	0	0
			497	320	86	90	1			
2	F	63	Total	C	N	O	S	0	1	0
			505	324	88	92	1			
2	G	63	Total	C	N	O	S	0	0	0
			497	320	86	90	1			
2	H	63	Total	C	N	O	S	0	2	0
			512	328	89	94	1			

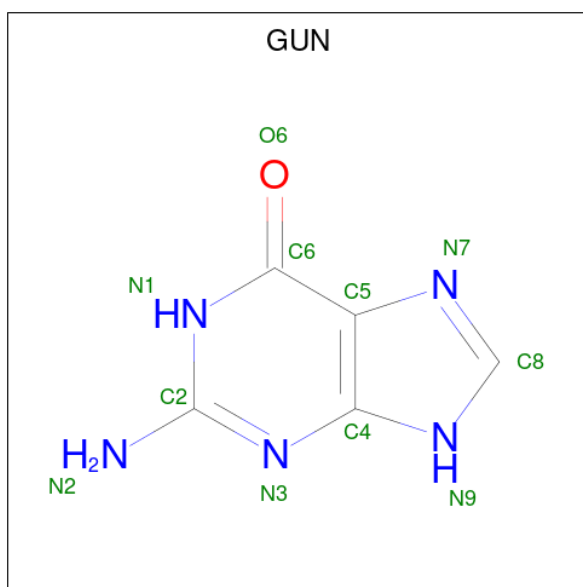
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	63	ALA	-	expression tag	UNP A0A376T1I0
F	63	ALA	-	expression tag	UNP A0A376T1I0
G	63	ALA	-	expression tag	UNP A0A376T1I0
H	63	ALA	-	expression tag	UNP A0A376T1I0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total Ca 5 5	0	0
3	B	5	Total Ca 5 5	0	0
3	C	5	Total Ca 5 5	0	0
3	D	5	Total Ca 5 5	0	0

- Molecule 4 is GUANINE (three-letter code: GUN) (formula: C<sub>5</sub>H<sub>5</sub>N<sub>5</sub>O).



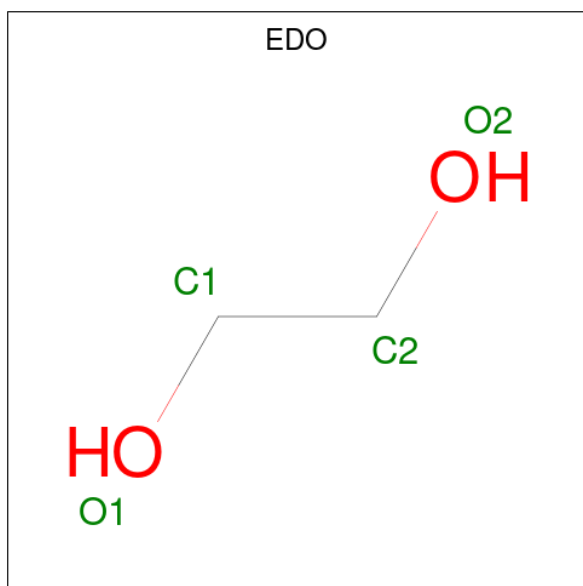
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 11 5 5 1	0	0
4	E	1	Total C N O 11 5 5 1	0	0
4	F	1	Total C N O 11 5 5 1	0	0

*Continued on next page...*

Continued from previous page...

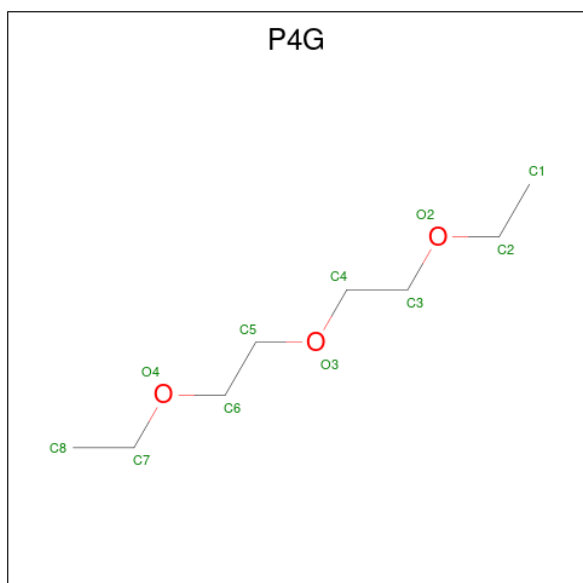
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	H	1	11	5	5	1	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	4	2	2	0	0

- Molecule 6 is 1-ETHOXY-2-(2-ETHOXYETHOXY)ETHANE (three-letter code: P4G) (formula:  $C_8H_{18}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			11	8	3		
6	C	1	Total	C	O	0	0
			11	8	3		

- Molecule 7 is water.

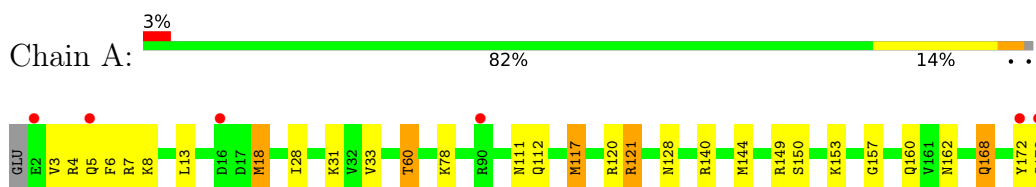
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	34	Total	O	0	0
			34	34		
7	B	40	Total	O	0	0
			40	40		
7	C	26	Total	O	0	0
			26	26		
7	D	23	Total	O	0	0
			23	23		
7	E	8	Total	O	0	0
			8	8		
7	F	8	Total	O	0	0
			8	8		
7	G	8	Total	O	0	0
			8	8		
7	H	9	Total	O	0	0
			9	9		



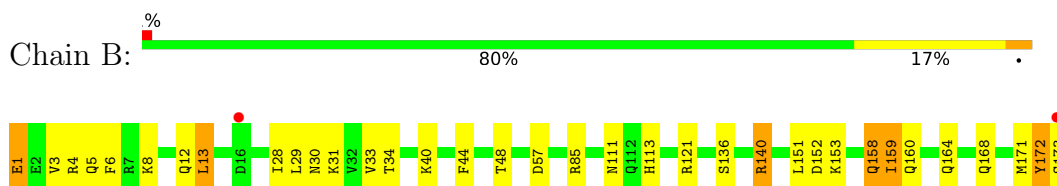
### 3 Residue-property plots [i](#)

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.

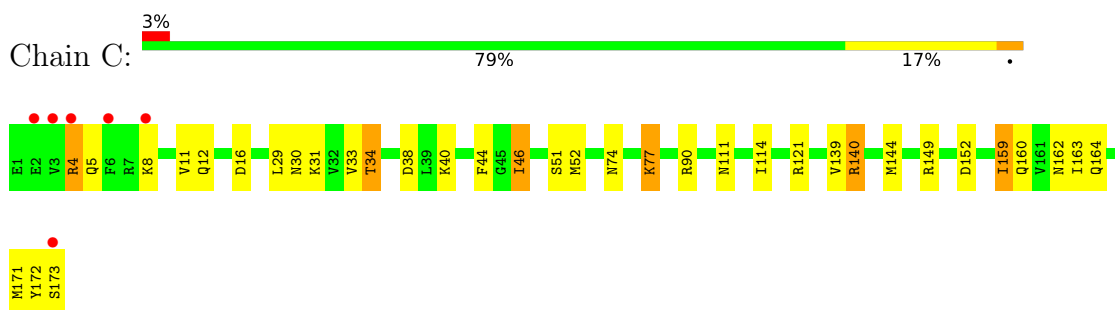
- Molecule 1: Calpain small subunit 1



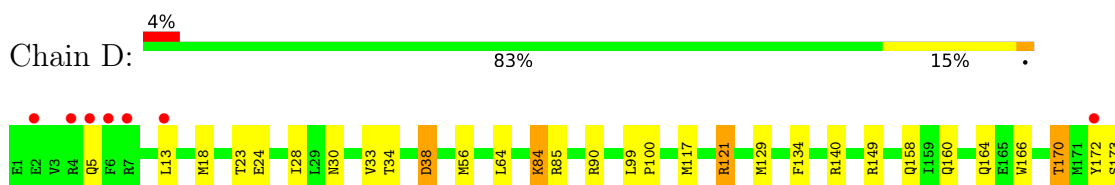
- Molecule 1: Calpain small subunit 1



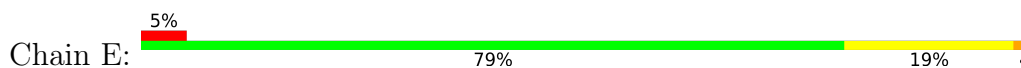
- Molecule 1: Calpain small subunit 1

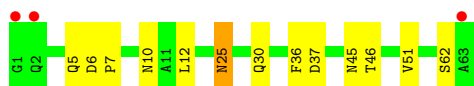


- Molecule 1: Calpain small subunit 1

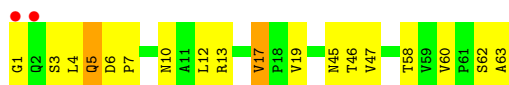
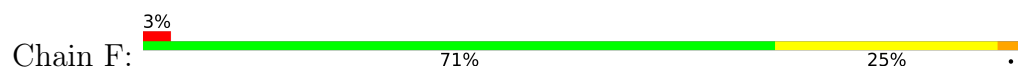


- Molecule 2: RNA-binding protein Hfq

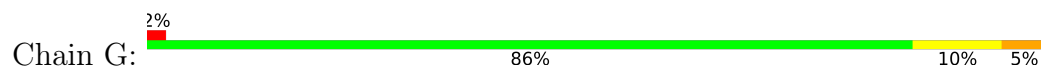




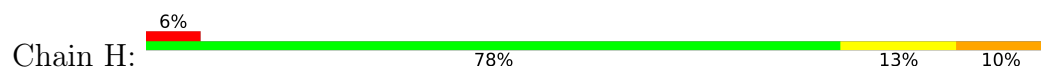
- Molecule 2: RNA-binding protein Hfq



- Molecule 2: RNA-binding protein Hfq



- Molecule 2: RNA-binding protein Hfq



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	147.61Å 147.61Å 147.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	66.01 – 2.32 66.01 – 2.32	Depositor EDS
% Data completeness (in resolution range)	99.0 (66.01-2.32) 99.0 (66.01-2.32)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.72 (at 2.32Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.201 , 0.247 0.205 , 0.248	Depositor DCC
$R_{free}$ test set	2337 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.1	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.033 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7880	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: GUN, P4G, CA, EDO

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	0.80	0/1421	1.02	4/1910 (0.2%)
1	B	0.82	0/1441	1.00	3/1937 (0.2%)
1	C	0.78	0/1441	0.98	3/1937 (0.2%)
1	D	0.84	2/1430 (0.1%)	0.99	1/1922 (0.1%)
2	E	0.81	0/505	0.92	0/684
2	F	0.80	0/513	0.99	0/695
2	G	0.77	0/505	0.93	0/684
2	H	0.78	0/520	0.96	0/705
All	All	0.80	2/7776 (0.0%)	0.99	11/10474 (0.1%)

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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	172	TYR	C-O	6.50	1.35	1.23
1	D	24	GLU	CD-OE2	-6.36	1.18	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	140	ARG	NE-CZ-NH2	-8.96	115.82	120.30
1	A	121	ARG	CB-CG-CD	7.92	132.19	111.60
1	A	121	ARG	NE-CZ-NH2	-7.00	116.80	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	140	ARG	NE-CZ-NH1	6.72	123.66	120.30
1	B	85	ARG	NE-CZ-NH2	6.50	123.55	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	172	TYR	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	1394	0	1352	36	0
1	B	1413	0	1367	36	0
1	C	1413	0	1368	29	0
1	D	1403	0	1361	21	0
2	E	497	0	520	9	0
2	F	505	0	525	11	0
2	G	497	0	520	6	0
2	H	512	0	531	15	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	D	5	0	0	0	0
4	A	11	0	5	0	0
4	E	11	0	5	0	0
4	F	11	0	5	2	0
4	H	11	0	5	4	0
5	A	4	0	6	4	0
6	B	11	0	18	0	0
6	C	11	0	18	0	0
7	A	34	0	0	1	0
7	B	40	0	0	0	0
7	C	26	0	0	0	0
7	D	23	0	0	2	0
7	E	8	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	F	8	0	0	2	0
7	G	8	0	0	2	0
7	H	9	0	0	0	0
All	All	7880	0	7606	128	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 128 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)
4:F:101:GUN:H8	7:F:201:HOH:O	1.36	1.22
7:G:104:HOH:O	4:H:101:GUN:H8	1.50	1.09
1:A:162:ASN:HD22	2:E:30:GLN:NE2	1.57	1.02
1:C:30:ASN:HD21	1:C:44:PHE:H	1.11	0.97
1:A:162:ASN:HD22	2:E:30:GLN:HE22	1.00	0.97

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	170/173 (98%)	169 (99%)	1 (1%)	0	100	100
1	B	172/173 (99%)	172 (100%)	0	0	100	100
1	C	172/173 (99%)	172 (100%)	0	0	100	100
1	D	171/173 (99%)	169 (99%)	2 (1%)	0	100	100
2	E	61/63 (97%)	57 (93%)	4 (7%)	0	100	100
2	F	62/63 (98%)	54 (87%)	8 (13%)	0	100	100
2	G	61/63 (97%)	55 (90%)	6 (10%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	63/63 (100%)	55 (87%)	8 (13%)	0	100	100
All	All	932/944 (99%)	903 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	151/152 (99%)	135 (89%)	16 (11%)	6	7
1	B	153/152 (101%)	138 (90%)	15 (10%)	8	9
1	C	153/152 (101%)	140 (92%)	13 (8%)	10	12
1	D	152/152 (100%)	140 (92%)	12 (8%)	12	15
2	E	57/57 (100%)	51 (90%)	6 (10%)	7	7
2	F	58/57 (102%)	52 (90%)	6 (10%)	7	8
2	G	57/57 (100%)	50 (88%)	7 (12%)	4	4
2	H	59/57 (104%)	49 (83%)	10 (17%)	2	2
All	All	840/836 (100%)	755 (90%)	85 (10%)	8	8

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	10	ASN
2	G	15	GLU
2	E	25	ASN
2	F	45[B]	ASN
2	H	4	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	168	GLN
2	H	25	ASN
2	E	25	ASN
2	G	30	GLN
2	E	10	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 20 are monoatomic - leaving 7 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	P4G	C	201	-	10,10,10	0.70	0	9,9,9	0.69	0
6	P4G	B	206	-	10,10,10	0.64	0	9,9,9	0.42	0
5	EDO	A	207	-	3,3,3	0.32	0	2,2,2	0.12	0
4	GUN	A	206	-	7,12,12	0.94	1 (14%)	8,17,17	1.24	1 (12%)
4	GUN	F	101	-	7,12,12	1.09	1 (14%)	8,17,17	0.89	0
4	GUN	E	101	-	7,12,12	1.12	1 (14%)	8,17,17	0.90	0
4	GUN	H	101	-	7,12,12	1.32	1 (14%)	8,17,17	0.81	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral



centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	P4G	C	201	-	-	5/8/8/8	-
6	P4G	B	206	-	-	6/8/8/8	-
5	EDO	A	207	-	-	1/1/1/1	-
4	GUN	A	206	-	-	-	0/2/2/2
4	GUN	F	101	-	-	-	0/2/2/2
4	GUN	E	101	-	-	-	0/2/2/2
4	GUN	H	101	-	-	-	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	101	GUN	C5-C6	-2.89	1.41	1.47
4	E	101	GUN	C5-C6	-2.36	1.42	1.47
4	F	101	GUN	C5-C6	-2.33	1.42	1.47
4	A	206	GUN	C5-C6	-2.10	1.43	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	206	GUN	O6-C6-C5	2.62	129.48	124.37

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	206	P4G	O3-C5-C6-O4
6	B	206	P4G	O2-C3-C4-O3
6	C	201	P4G	O2-C3-C4-O3
5	A	207	EDO	O1-C1-C2-O2
6	B	206	P4G	C8-C7-O4-C6

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	207	EDO	4	0
4	F	101	GUN	2	0
4	H	101	GUN	4	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	172/173 (99%)	0.04	6 (3%) 44 51	35, 53, 92, 135	0
1	B	173/173 (100%)	-0.15	2 (1%) 79 83	30, 44, 70, 83	0
1	C	173/173 (100%)	-0.03	6 (3%) 44 51	36, 53, 78, 102	0
1	D	173/173 (100%)	0.05	7 (4%) 38 45	37, 58, 91, 107	0
2	E	63/63 (100%)	0.19	3 (4%) 30 38	29, 51, 84, 106	0
2	F	63/63 (100%)	0.11	2 (3%) 47 55	31, 54, 84, 117	0
2	G	63/63 (100%)	0.01	1 (1%) 72 78	30, 53, 85, 123	0
2	H	63/63 (100%)	0.12	4 (6%) 20 26	33, 53, 85, 99	0
All	All	943/944 (99%)	0.01	31 (3%) 46 53	29, 52, 85, 135	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	1	GLY	5.5
2	F	2	GLN	4.6
1	B	173	SER	4.3
1	C	173	SER	3.8
2	E	2	GLN	3.6

### 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	CA	D	205	1/1	0.29	0.18	96,96,96,96	0
3	CA	C	205	1/1	0.37	0.20	63,63,63,63	0
3	CA	B	205	1/1	0.45	0.30	100,100,100,100	0
3	CA	D	201	1/1	0.55	0.16	77,77,77,77	0
3	CA	A	205	1/1	0.68	0.24	86,86,86,86	0
3	CA	C	206	1/1	0.68	0.18	105,105,105,105	0
3	CA	D	204	1/1	0.74	0.13	71,71,71,71	0
6	P4G	B	206	11/11	0.82	0.13	77,85,89,93	0
6	P4G	C	201	11/11	0.82	0.14	72,78,87,89	0
3	CA	C	202	1/1	0.92	0.08	53,53,53,53	0
5	EDO	A	207	4/4	0.94	0.20	50,54,58,63	0
4	GUN	A	206	11/11	0.94	0.12	37,41,45,45	0
4	GUN	F	101	11/11	0.94	0.13	38,41,43,43	0
4	GUN	E	101	11/11	0.96	0.13	38,41,45,50	0
3	CA	A	204	1/1	0.96	0.05	70,70,70,70	0
4	GUN	H	101	11/11	0.96	0.10	36,39,48,48	0
3	CA	B	202	1/1	0.98	0.13	51,51,51,51	0
3	CA	D	202	1/1	0.98	0.04	51,51,51,51	0
3	CA	D	203	1/1	0.98	0.13	59,59,59,59	0
3	CA	C	203	1/1	0.99	0.05	45,45,45,45	0
3	CA	C	204	1/1	0.99	0.09	54,54,54,54	0
3	CA	B	201	1/1	0.99	0.11	45,45,45,45	0
3	CA	A	203	1/1	0.99	0.06	43,43,43,43	0
3	CA	B	203	1/1	0.99	0.09	37,37,37,37	0
3	CA	B	204	1/1	0.99	0.08	37,37,37,37	0
3	CA	A	201	1/1	0.99	0.11	49,49,49,49	0
3	CA	A	202	1/1	0.99	0.05	57,57,57,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.