



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:02 am GMT

PDB ID : 5A7D  
Title : Tetrameric assembly of LGN with Inscuteable  
Authors : Culurgioni, S.; Mari, S.; Bonetto, G.; Gallini, S.; Brennich, M.; Round, A.;  
Mapelli, M.  
Deposited on : 2015-07-07  
Resolution : 3.40 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

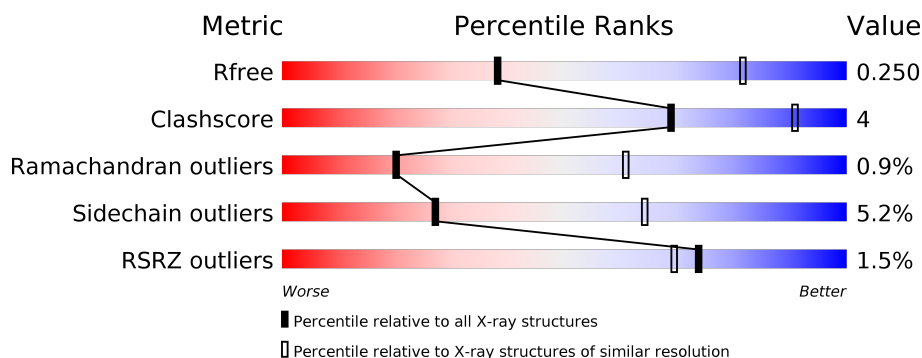
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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))
$R_{free}$	100719	1679 (3.50-3.30)
Clashscore	112137	1832 (3.50-3.30)
Ramachandran outliers	110173	1789 (3.50-3.30)
Sidechain outliers	110143	1789 (3.50-3.30)
RSRZ outliers	101464	1709 (3.50-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	B	382	<div> <div>77%</div> <div>13% • 7%</div> </div>
1	C	382	<div> <div>2%</div> <div>80%</div> <div>10% • 9%</div> </div>
1	D	382	<div> <div>2%</div> <div>80%</div> <div>10% 10%</div> </div>
1	E	382	<div> <div>2%</div> <div>77%</div> <div>14% 8%</div> </div>
1	F	382	<div> <div>2%</div> <div>82%</div> <div>9% 9%</div> </div>
1	G	382	<div> <div>2%</div> <div>78%</div> <div>12% • 9%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	H	382	<p>78% 12% 9%</p>
1	I	382	<p>80% 10% 9%</p>
2	L	341	<p>70% 13% 17%</p>
2	M	341	<p>64% 11% 24%</p>
2	N	341	<p>72% 11% 17%</p>
2	O	341	<p>68% 15% 16%</p>
2	P	341	<p>71% 12% 16%</p>
2	Q	341	<p>66% 16% 18%</p>
2	R	341	<p>78% 10% 9%</p>
2	S	341	<p>70% 13% 16%</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 38980 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 PINS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	354	Total	C	N	O	S	0	0	0
			2742	1695	513	521	13			
1	C	348	Total	C	N	O	S	0	0	0
			2685	1655	504	513	13			
1	D	345	Total	C	N	O	S	0	0	0
			2643	1635	489	506	13			
1	E	352	Total	C	N	O	S	0	0	0
			2616	1618	489	497	12			
1	F	346	Total	C	N	O	S	0	0	0
			2676	1652	502	509	13			
1	G	346	Total	C	N	O	S	0	0	0
			2679	1655	504	507	13			
1	H	347	Total	C	N	O	S	0	0	0
			2668	1650	499	506	13			
1	I	346	Total	C	N	O	S	0	0	0
			2608	1615	481	499	13			

- Molecule 2 is a protein called INSCUTEABLE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	284	Total	C	N	O	S	0	0	0
			2127	1329	386	402	10			
2	M	259	Total	C	N	O	S	0	0	0
			1977	1237	361	369	10			
2	N	283	Total	C	N	O	S	0	0	0
			2228	1395	407	416	10			
2	O	288	Total	C	N	O	S	0	0	0
			2250	1410	409	421	10			
2	P	288	Total	C	N	O	S	0	0	0
			2233	1396	406	421	10			
2	Q	281	Total	C	N	O	S	0	0	0
			2215	1388	401	416	10			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	R	311	Total	C	N	O	S	0	0	0
			2386	1490	439	447	10			
2	S	287	Total	C	N	O	S	0	0	0
			2243	1405	408	420	10			

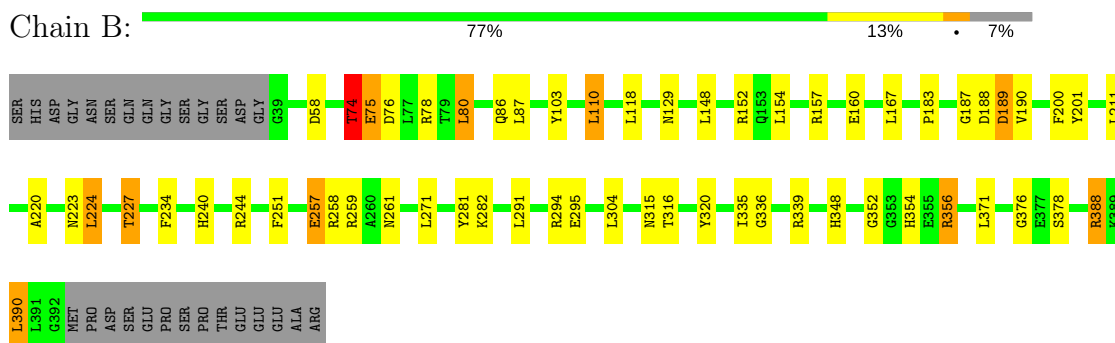
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	O	0	0
			2	2		
3	H	2	Total	O	0	0
			2	2		

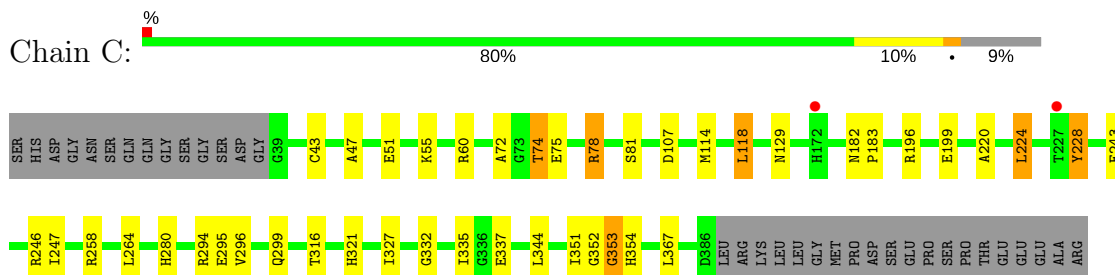
### 3 Residue-property plots

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.

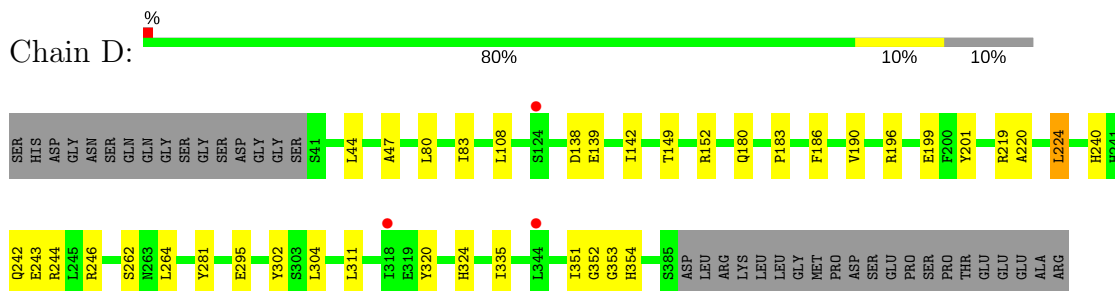
- Molecule 1: PINS



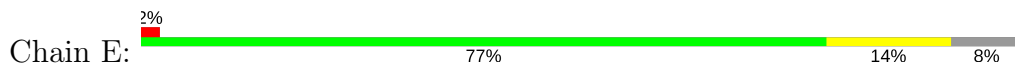
- Molecule 1: PINS

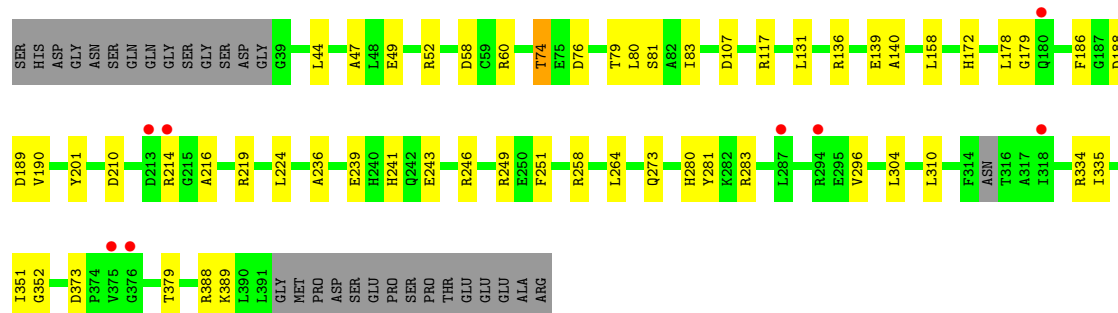


- Molecule 1: PINS

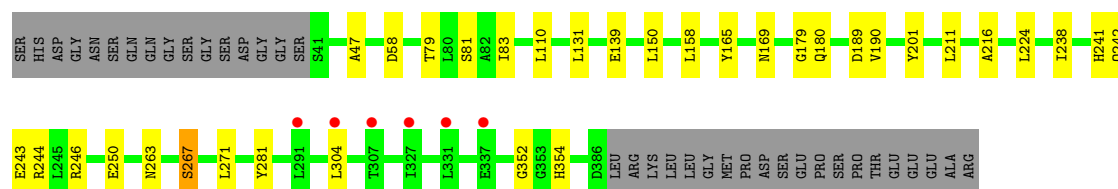
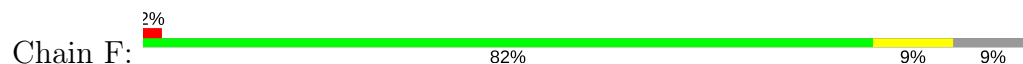


- Molecule 1: PINS

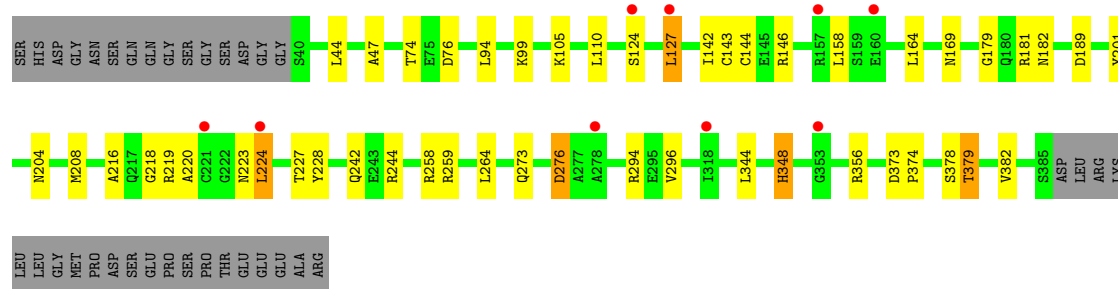
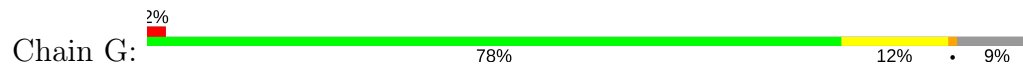




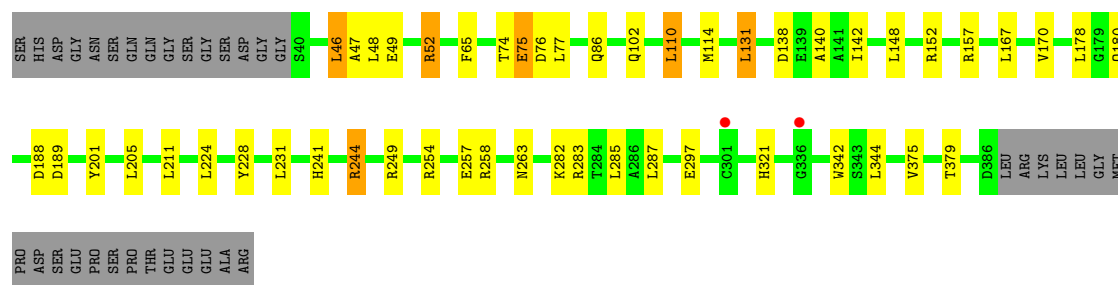
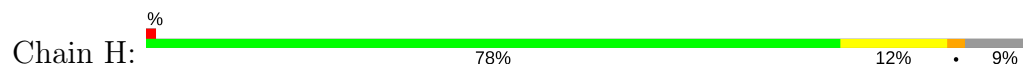
- Molecule 1: PINS



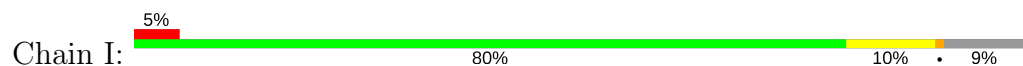
- Molecule 1: PINS



- Molecule 1: PINS



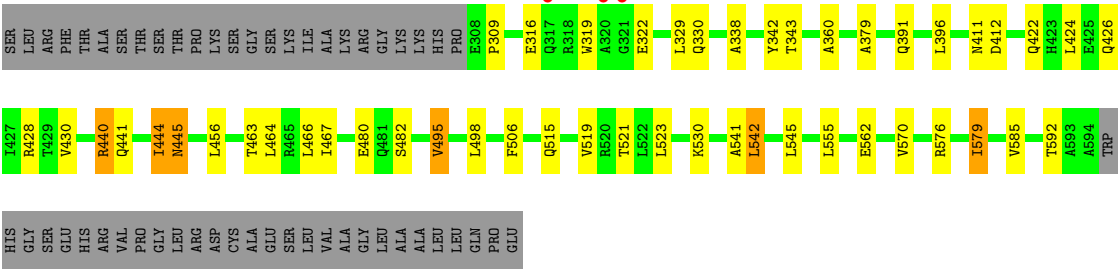
- Molecule 1: PINS











## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.19Å 212.58Å 280.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	75.58 – 3.40 86.47 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (75.58-3.40) 99.9 (86.47-3.40)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 3.41Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.209 , 0.250 0.210 , 0.250	Depositor DCC
$R_{free}$ test set	5218 reflections (4.93%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.1	Xtriage
Anisotropy	0.393	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 66.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	38980	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.06% 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

### 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	B	0.22	0/2787	0.35	0/3748
1	C	0.22	0/2729	0.35	0/3673
1	D	0.21	0/2688	0.34	0/3623
1	E	0.21	0/2657	0.33	0/3583
1	F	0.21	0/2721	0.34	0/3662
1	G	0.21	0/2724	0.34	0/3665
1	H	0.22	0/2713	0.35	0/3654
1	I	0.21	0/2650	0.35	0/3576
2	L	0.22	0/2149	0.39	1/2914 (0.0%)
2	M	0.23	0/1999	0.39	0/2710
2	N	0.23	0/2254	0.39	0/3049
2	O	0.22	0/2278	0.38	0/3085
2	P	0.22	0/2260	0.39	0/3063
2	Q	0.22	0/2242	0.37	0/3034
2	R	0.23	0/2416	0.40	1/3273 (0.0%)
2	S	0.22	0/2270	0.39	0/3074
All	All	0.22	0/39537	0.36	2/53386 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	603	PRO	N-CA-CB	6.14	110.67	103.30
2	R	603	PRO	N-CA-CB	6.02	110.52	103.30

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	B	2742	0	2647	32	0
1	C	2685	0	2569	23	0
1	D	2643	0	2503	19	0
1	E	2616	0	2438	31	0
1	F	2676	0	2560	18	0
1	G	2679	0	2570	27	0
1	H	2668	0	2545	28	0
1	I	2608	0	2457	21	0
2	L	2127	0	2088	18	0
2	M	1977	0	1958	19	0
2	N	2228	0	2268	19	0
2	O	2250	0	2285	29	0
2	P	2233	0	2253	25	0
2	Q	2215	0	2251	25	0
2	R	2386	0	2372	22	0
2	S	2243	0	2277	17	0
3	B	2	0	0	0	0
3	H	2	0	0	0	0
All	All	38980	0	38041	330	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 4.

The worst 5 of 330 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)
2:M:567:ILE:HG23	2:M:587:LEU:HD21	1.63	0.80
1:B:74:THR:OG1	1:B:75:GLU:N	2.16	0.76
2:O:523:LEU:HD22	2:O:563:ILE:HD11	1.70	0.74
2:P:342:TYR:N	2:P:343:THR:HA	2.04	0.73
1:E:47:ALA:HB1	2:O:334:ILE:HD13	1.71	0.72

There are no symmetry-related clashes.

## 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	B	352/382 (92%)	329 (94%)	17 (5%)	6 (2%)	11	47
1	C	346/382 (91%)	334 (96%)	11 (3%)	1 (0%)	44	79
1	D	343/382 (90%)	332 (97%)	10 (3%)	1 (0%)	44	79
1	E	348/382 (91%)	330 (95%)	16 (5%)	2 (1%)	28	68
1	F	344/382 (90%)	334 (97%)	9 (3%)	1 (0%)	44	79
1	G	344/382 (90%)	331 (96%)	13 (4%)	0	100	100
1	H	345/382 (90%)	337 (98%)	7 (2%)	1 (0%)	44	79
1	I	344/382 (90%)	330 (96%)	9 (3%)	5 (2%)	12	49
2	L	278/341 (82%)	255 (92%)	20 (7%)	3 (1%)	17	57
2	M	251/341 (74%)	236 (94%)	13 (5%)	2 (1%)	22	62
2	N	279/341 (82%)	263 (94%)	16 (6%)	0	100	100
2	O	286/341 (84%)	270 (94%)	12 (4%)	4 (1%)	13	51
2	P	286/341 (84%)	273 (96%)	12 (4%)	1 (0%)	44	79
2	Q	277/341 (81%)	269 (97%)	6 (2%)	2 (1%)	25	65
2	R	307/341 (90%)	283 (92%)	17 (6%)	7 (2%)	7	41
2	S	285/341 (84%)	266 (93%)	10 (4%)	9 (3%)	5	34
All	All	5015/5784 (87%)	4772 (95%)	198 (4%)	45 (1%)	20	61

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	183	PRO
1	E	351	ILE
2	M	408	VAL
2	S	343	THR
1	B	75	GLU

### 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	B	271/298 (91%)	252 (93%)	19 (7%)	18	55
1	C	263/298 (88%)	254 (97%)	9 (3%)	42	75
1	D	256/298 (86%)	247 (96%)	9 (4%)	41	74
1	E	241/298 (81%)	232 (96%)	9 (4%)	39	73
1	F	262/298 (88%)	256 (98%)	6 (2%)	56	82
1	G	262/298 (88%)	248 (95%)	14 (5%)	26	63
1	H	259/298 (87%)	247 (95%)	12 (5%)	31	68
1	I	247/298 (83%)	238 (96%)	9 (4%)	40	74
2	L	216/292 (74%)	203 (94%)	13 (6%)	22	60
2	M	206/292 (70%)	195 (95%)	11 (5%)	26	63
2	N	242/292 (83%)	231 (96%)	11 (4%)	32	69
2	O	243/292 (83%)	225 (93%)	18 (7%)	16	51
2	P	241/292 (82%)	227 (94%)	14 (6%)	23	61
2	Q	242/292 (83%)	229 (95%)	13 (5%)	26	63
2	R	248/292 (85%)	231 (93%)	17 (7%)	18	55
2	S	242/292 (83%)	223 (92%)	19 (8%)	14	48
All	All	3941/4720 (84%)	3738 (95%)	203 (5%)	27	64

5 of 203 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	410	PHE
2	N	406	ARG
2	S	412	ASP
2	L	495	VAL
2	M	366	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	202	GLN
1	D	240	HIS
2	L	481	GLN
2	M	366	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 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, 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	B	354/382 (92%)	0.07	0 100 100	35, 56, 99, 127	0
1	C	348/382 (91%)	0.22	2 (0%) 89 86	46, 79, 126, 152	0
1	D	345/382 (90%)	0.30	3 (0%) 84 80	51, 89, 121, 167	0
1	E	352/382 (92%)	0.21	8 (2%) 61 56	46, 104, 147, 191	0
1	F	346/382 (90%)	0.28	6 (1%) 70 66	31, 70, 132, 150	0
1	G	346/382 (90%)	0.48	9 (2%) 56 52	50, 84, 113, 154	0
1	H	347/382 (90%)	0.10	2 (0%) 89 86	31, 64, 119, 132	0
1	I	346/382 (90%)	0.55	20 (5%) 24 23	50, 103, 147, 161	0
2	L	284/341 (83%)	0.17	9 (3%) 48 45	43, 75, 125, 150	0
2	M	259/341 (75%)	0.20	3 (1%) 79 75	45, 82, 115, 132	0
2	N	283/341 (82%)	0.20	2 (0%) 87 85	42, 69, 109, 135	0
2	O	288/341 (84%)	0.10	0 100 100	33, 62, 124, 161	0
2	P	288/341 (84%)	0.18	3 (1%) 82 78	28, 55, 115, 176	0
2	Q	281/341 (82%)	0.31	2 (0%) 87 85	43, 68, 99, 117	0
2	R	311/341 (91%)	0.16	3 (0%) 82 78	28, 52, 134, 153	0
2	S	287/341 (84%)	0.09	3 (1%) 82 78	36, 67, 128, 153	0
All	All	5065/5784 (87%)	0.23	75 (1%) 74 69	28, 74, 130, 191	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	376	GLY	7.6
2	R	340	LYS	5.1
1	I	307	THR	4.4
1	I	348	HIS	4.2
2	R	338	ALA	4.2

## 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 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.