Full wwPDB X-ray Structure Validation Report ⓘ

Dec 1, 2022 – 12:38 pm GMT

PDB ID	:	7Q72
Title	:	Structure of Pla1 in complex with Red1
Authors	:	Soni, K.; Wild, K.; Sinning, I.
Deposited on	:	2021-11-09
Resolution	:	2.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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:

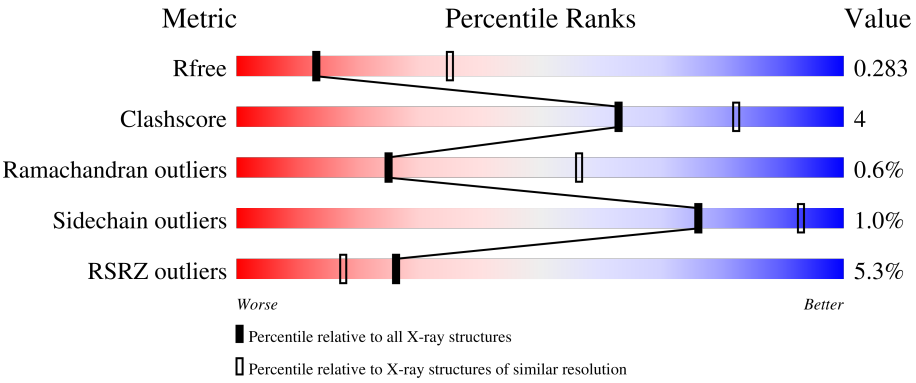
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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

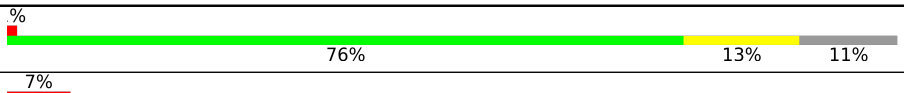
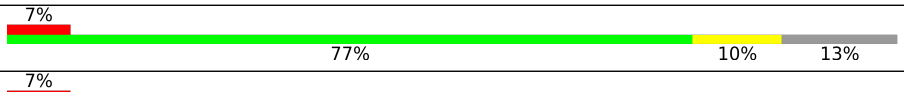
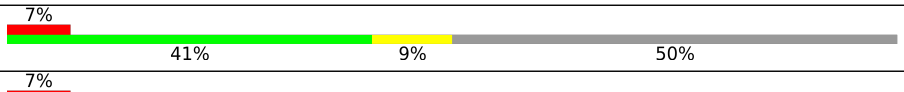
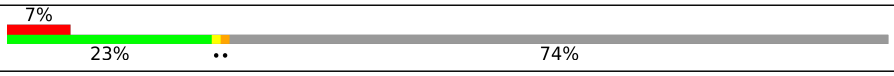
The reported resolution of this entry is 2.80 Å.

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	575	
1	B	575	
2	C	70	
2	D	70	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8531 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 Poly(A) polymerase pla1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	511	Total	C	N	O	S	0	0	0
			4097	2638	699	744	16			
1	B	502	Total	C	N	O	S	0	0	0
			4027	2593	686	732	16			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q10295
A	2	LYS	-	expression tag	UNP Q10295
A	3	HIS	-	expression tag	UNP Q10295
A	4	HIS	-	expression tag	UNP Q10295
A	5	HIS	-	expression tag	UNP Q10295
A	6	HIS	-	expression tag	UNP Q10295
A	7	HIS	-	expression tag	UNP Q10295
A	8	HIS	-	expression tag	UNP Q10295
A	9	PRO	-	expression tag	UNP Q10295
B	1	MET	-	initiating methionine	UNP Q10295
B	2	LYS	-	expression tag	UNP Q10295
B	3	HIS	-	expression tag	UNP Q10295
B	4	HIS	-	expression tag	UNP Q10295
B	5	HIS	-	expression tag	UNP Q10295
B	6	HIS	-	expression tag	UNP Q10295
B	7	HIS	-	expression tag	UNP Q10295
B	8	HIS	-	expression tag	UNP Q10295
B	9	PRO	-	expression tag	UNP Q10295

- Molecule 2 is a protein called NURS complex subunit red1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	35	Total	C	N	O	0	0	0
			268	169	40	59			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	18	Total	C	N	O	0	0	0
			136	88	21	27			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	284	GLY	-	expression tag	UNP Q9UTR8
C	285	ALA	-	expression tag	UNP Q9UTR8
C	286	MET	-	expression tag	UNP Q9UTR8
C	287	GLY	-	expression tag	UNP Q9UTR8
C	346	GLY	-	expression tag	UNP Q9UTR8
C	347	SER	-	expression tag	UNP Q9UTR8
C	348	HIS	-	expression tag	UNP Q9UTR8
C	349	HIS	-	expression tag	UNP Q9UTR8
C	350	HIS	-	expression tag	UNP Q9UTR8
C	351	HIS	-	expression tag	UNP Q9UTR8
C	352	HIS	-	expression tag	UNP Q9UTR8
C	353	HIS	-	expression tag	UNP Q9UTR8
D	284	GLY	-	expression tag	UNP Q9UTR8
D	285	ALA	-	expression tag	UNP Q9UTR8
D	286	MET	-	expression tag	UNP Q9UTR8
D	287	GLY	-	expression tag	UNP Q9UTR8
D	346	GLY	-	expression tag	UNP Q9UTR8
D	347	SER	-	expression tag	UNP Q9UTR8
D	348	HIS	-	expression tag	UNP Q9UTR8
D	349	HIS	-	expression tag	UNP Q9UTR8
D	350	HIS	-	expression tag	UNP Q9UTR8
D	351	HIS	-	expression tag	UNP Q9UTR8
D	352	HIS	-	expression tag	UNP Q9UTR8
D	353	HIS	-	expression tag	UNP Q9UTR8

- Molecule 3 is water.

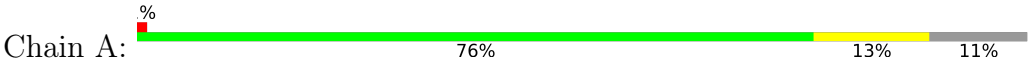
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O	0	0
			1	1		
3	B	2	Total	O	0	0
			2	2		

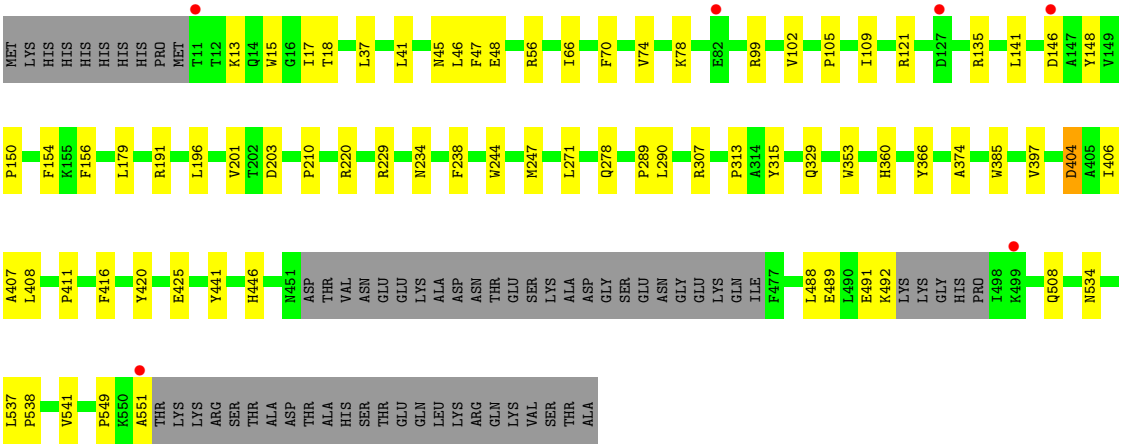
3 Residue-property plots (i)

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: Poly(A) polymerase pla1

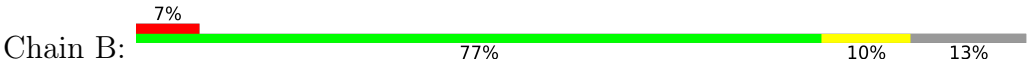
Chain A:

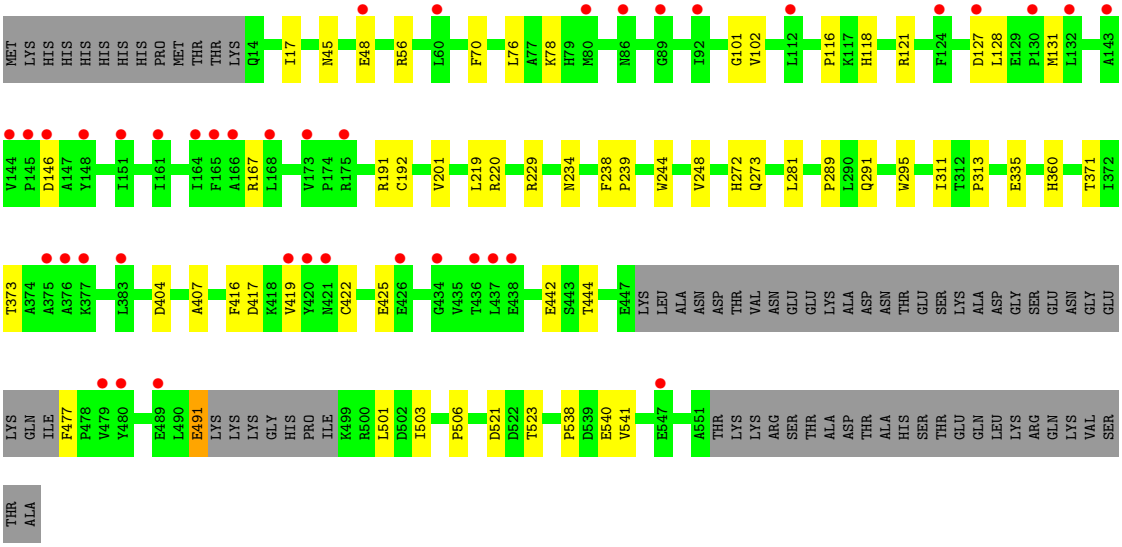




• Molecule 1: Poly(A) polymerase pla1

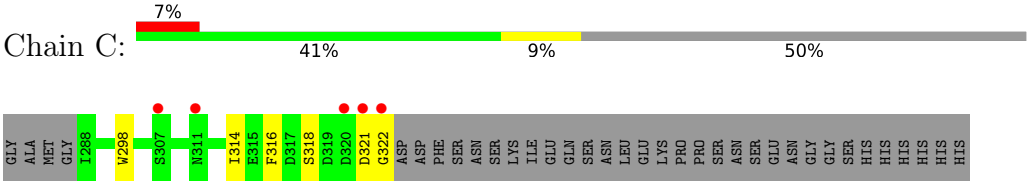
Chain B:





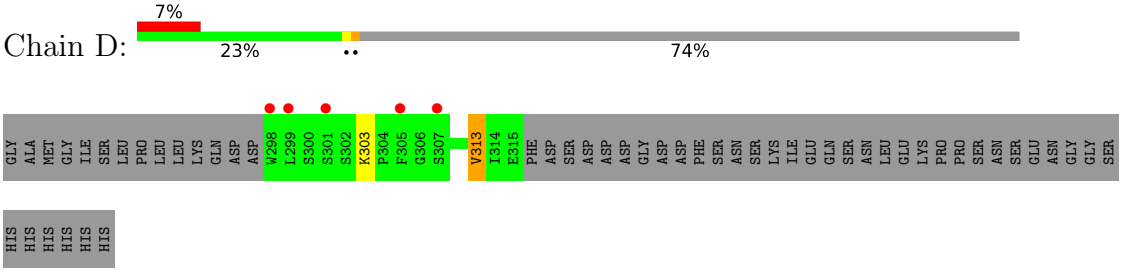
• Molecule 2: NURS complex subunit red1

Chain C:



• Molecule 2: NURS complex subunit red1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	128.64Å 151.34Å 65.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	98.02 – 2.80 98.02 – 2.81	Depositor EDS
% Data completeness (in resolution range)	64.0 (98.02-2.80) 64.0 (98.02-2.81)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 2.82Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158+SVN	Depositor
R, R_{free}	0.241 , 0.283 0.240 , 0.283	Depositor DCC
R_{free} test set	1999 reflections (9.74%)	wwPDB-VP
Wilson B-factor (Å^2)	51.1	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	8531	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% 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	0.24	0/4194	0.45	0/5697
1	B	0.24	0/4124	0.45	0/5604
2	C	0.24	0/274	0.39	0/373
2	D	0.25	0/140	0.47	0/191
All	All	0.24	0/8732	0.45	0/11865

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	B	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	404	ASP	Peptide
1	B	404	ASP	Peptide

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	4097	0	4139	40	0
1	B	4027	0	4053	33	0
2	C	268	0	250	3	0
2	D	136	0	132	2	0
3	A	1	0	0	1	0
3	B	2	0	0	0	0
All	All	8531	0	8574	73	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.

All (73) 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:78:LYS:NZ	1:B:118:HIS:O	2.22	0.71
1:A:407:ALA:HB2	1:A:491:GLU:HG2	1.74	0.68
1:A:397:VAL:HG11	1:A:411:PRO:HG3	1.77	0.67
1:B:501:LEU:HB3	2:D:313:VAL:HG22	1.76	0.67
1:B:102:VAL:HG13	1:B:220:ARG:HG2	1.75	0.67
1:A:121:ARG:HH12	1:A:191:ARG:HH11	1.44	0.65
1:A:102:VAL:HG13	1:A:220:ARG:HG2	1.78	0.65
1:B:422:CYS:HB2	1:B:477:PHE:HB2	1.78	0.63
1:B:229:ARG:NE	1:B:335:GLU:OE1	2.27	0.62
1:A:48:GLU:HA	1:A:105:PRO:HG3	1.81	0.62
1:B:425:GLU:OE2	2:D:303:LYS:NZ	2.33	0.61
1:B:17:ILE:HD13	1:B:281:LEU:HB2	1.82	0.61
1:A:46:LEU:O	1:A:220:ARG:NH2	2.35	0.59
1:A:508:GLN:NE2	3:A:601:HOH:O	2.36	0.57
1:A:56:ARG:HD3	1:A:109:ILE:HD11	1.87	0.56
1:B:121:ARG:HH22	1:B:191:ARG:HD3	1.71	0.56
1:B:521:ASP:OD2	1:B:523:THR:OG1	2.27	0.53
1:A:290:LEU:O	1:A:315:TYR:OH	2.26	0.53
1:A:37:LEU:HB2	1:A:353:TRP:CE2	2.45	0.52
1:A:154:PHE:HE2	1:A:156:PHE:HB2	1.76	0.51
1:A:229:ARG:HA	1:A:360:HIS:HB3	1.92	0.51
1:A:201:VAL:HG13	1:A:313:PRO:HD2	1.92	0.51
1:B:229:ARG:HA	1:B:360:HIS:HB3	1.91	0.51
1:A:408:LEU:HB3	1:A:489:GLU:HB2	1.93	0.50
1:B:407:ALA:HB2	1:B:491:GLU:HG2	1.94	0.50
1:A:534:ASN:HB3	1:A:549:PRO:HB2	1.93	0.49
1:B:419:VAL:HB	1:B:442:GLU:HB2	1.94	0.49
1:B:121:ARG:HH12	1:B:191:ARG:HH11	1.61	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:141:LEU:HD13	1:A:154:PHE:HB3	1.95	0.49
1:B:101:GLY:O	1:B:220:ARG:NH1	2.46	0.48
1:B:76:LEU:HA	1:B:76:LEU:HD23	1.60	0.47
1:A:244:TRP:HA	1:A:247:MET:HE3	1.97	0.47
1:B:48:GLU:OE2	1:B:56:ARG:NH2	2.48	0.46
1:A:146:ASP:N	1:A:146:ASP:OD1	2.49	0.46
1:A:366:TYR:HD1	1:A:489:GLU:HG2	1.81	0.46
1:B:538:PRO:HB2	1:B:540:GLU:OE1	2.16	0.46
1:A:374:ALA:HB1	1:A:385:TRP:CD1	2.51	0.45
1:A:416:PHE:CE2	1:A:541:VAL:HG12	2.52	0.45
1:B:416:PHE:CE2	1:B:541:VAL:HG12	2.52	0.45
1:B:127:ASP:O	1:B:131:MET:HG3	2.17	0.45
2:C:316:PHE:CE1	2:C:318:SER:HB2	2.52	0.45
1:B:234:ASN:HA	1:B:238:PHE:O	2.16	0.45
1:A:551:ALA:HB1	2:C:322:GLY:HA2	1.98	0.45
1:B:371:THR:HG22	1:B:373:THR:HG23	1.99	0.45
1:A:13:LYS:HD3	1:A:15:TRP:HE1	1.81	0.45
1:B:201:VAL:HG13	1:B:313:PRO:HD2	1.98	0.45
1:A:45:ASN:HB2	1:B:45:ASN:HB2	1.99	0.44
1:A:538:PRO:HG2	1:A:541:VAL:HG13	2.00	0.44
1:A:41:LEU:HB3	1:A:47:PHE:HE2	1.83	0.43
1:A:406:ILE:HG21	1:A:488:LEU:HD22	2.00	0.43
1:B:167:ARG:NE	1:B:192:CYS:SG	2.91	0.43
1:B:272:HIS:CE1	1:B:273:GLN:HG3	2.53	0.43
1:A:135:ARG:HA	1:A:135:ARG:HD2	1.87	0.43
1:A:491:GLU:O	1:A:492:LYS:HG2	2.19	0.43
1:A:179:LEU:HD13	1:A:196:LEU:HD22	2.01	0.43
1:A:537:LEU:HD23	1:A:541:VAL:HG21	2.01	0.42
1:A:307:ARG:HB3	1:A:329:GLN:OE1	2.19	0.42
1:A:17:ILE:HG13	1:A:18:THR:HG23	2.01	0.42
1:B:239:PRO:HG2	1:B:244:TRP:CE2	2.55	0.42
1:B:503:ILE:C	1:B:506:PRO:HD2	2.40	0.41
1:A:74:VAL:O	1:A:78:LYS:HG2	2.19	0.41
1:B:70:PHE:CE2	1:B:128:LEU:HB2	2.56	0.41
1:A:271:LEU:HA	1:A:271:LEU:HD23	1.74	0.41
1:B:219:LEU:HD13	1:B:248:VAL:HG21	2.03	0.41
1:A:66:ILE:O	1:A:70:PHE:N	2.53	0.41
1:A:234:ASN:HA	1:A:238:PHE:O	2.21	0.41
1:A:420:TYR:CE1	1:A:441:TYR:HB3	2.56	0.41
1:A:425:GLU:HG3	2:C:298:TRP:CZ3	2.56	0.41
1:A:99:ARG:HD3	1:A:203:ASP:OD1	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:70:PHE:HB2	1:B:128:LEU:HD13	2.03	0.40
1:B:295:TRP:CE2	1:B:311:ILE:HD11	2.56	0.40
1:B:102:VAL:HG21	1:B:219:LEU:HB3	2.03	0.40
1:B:417:ASP:HB2	1:B:444:THR:O	2.21	0.40

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	505/575 (88%)	480 (95%)	22 (4%)	3 (1%)	25	56
1	B	496/575 (86%)	461 (93%)	32 (6%)	3 (1%)	25	56
2	C	33/70 (47%)	30 (91%)	3 (9%)	0	100	100
2	D	16/70 (23%)	13 (81%)	3 (19%)	0	100	100
All	All	1050/1290 (81%)	984 (94%)	60 (6%)	6 (1%)	25	56

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	289	PRO
1	A	210	PRO
1	B	146	ASP
1	B	116	PRO
1	A	150	PRO
1	A	289	PRO

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	446/502 (89%)	442 (99%)	4 (1%)	78 94
1	B	438/502 (87%)	436 (100%)	2 (0%)	88 96
2	C	33/63 (52%)	31 (94%)	2 (6%)	18 48
2	D	17/63 (27%)	16 (94%)	1 (6%)	19 49
All	All	934/1130 (83%)	925 (99%)	9 (1%)	76 93

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	148	TYR
1	A	278	GLN
1	A	404	ASP
1	A	446	HIS
2	C	314	ILE
2	C	321	ASP
1	B	291	GLN
1	B	491	GLU
2	D	313	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	272	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	511/575 (88%)	0.10	6 (1%) 79 73	7, 34, 72, 89	0
1	B	502/575 (87%)	0.49	40 (7%) 12 6	16, 52, 100, 134	0
2	C	35/70 (50%)	0.71	5 (14%) 2 1	23, 44, 88, 89	0
2	D	18/70 (25%)	1.10	5 (27%) 0 0	52, 87, 111, 114	0
All	All	1066/1290 (82%)	0.32	56 (5%) 26 17	7, 43, 95, 134	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	143	ALA	4.9
2	D	298	TRP	4.5
1	B	132	LEU	4.3
1	B	145	PRO	4.0
2	C	320	ASP	4.0
1	B	480	TYR	3.9
1	B	438	GLU	3.9
1	B	165	PHE	3.8
1	B	127	ASP	3.8
1	B	421	ASN	3.8
1	B	146	ASP	3.5
1	A	499	LYS	3.4
1	B	479	VAL	3.3
1	A	146	ASP	3.2
1	B	89	GLY	3.2
1	B	173	VAL	3.1
2	D	307	SER	3.1
1	B	80	MET	3.1
1	B	168	LEU	3.0
2	C	322	GLY	3.0
2	D	305	PHE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	437	LEU	3.0
1	B	112	LEU	2.9
1	B	124	PHE	2.8
1	B	130	PRO	2.8
1	B	166	ALA	2.8
2	D	301	SER	2.7
1	B	148	TYR	2.7
1	A	11	THR	2.7
1	B	436	THR	2.7
1	A	82	GLU	2.6
1	B	375	ALA	2.6
1	B	377	LYS	2.6
1	B	164	ILE	2.6
1	B	60	LEU	2.6
1	B	426	GLU	2.5
1	B	547	GLU	2.5
1	A	127	ASP	2.5
1	B	175	ARG	2.5
1	B	489	GLU	2.5
1	B	434	GLY	2.5
1	B	383	LEU	2.4
1	B	144	VAL	2.4
1	B	420	TYR	2.3
2	C	307	SER	2.3
1	B	376	ALA	2.3
1	B	48	GLU	2.3
1	B	151	ILE	2.3
2	C	311	ASN	2.2
2	D	299	LEU	2.2
2	C	321	ASP	2.2
1	B	92	ILE	2.2
1	B	161	ILE	2.1
1	B	86	ASN	2.0
1	B	419	VAL	2.0
1	A	551	ALA	2.0

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.