



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:28 PM GMT

PDB ID : 4OV5
Title : Structure of HLA-DR1 with a bound peptide with non-optimal alanine in the P1 pocket
Authors : Trenh, P.; Yin, L.; Stern, L.J.
Deposited on : 2014-02-20
Resolution : 2.20 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
<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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

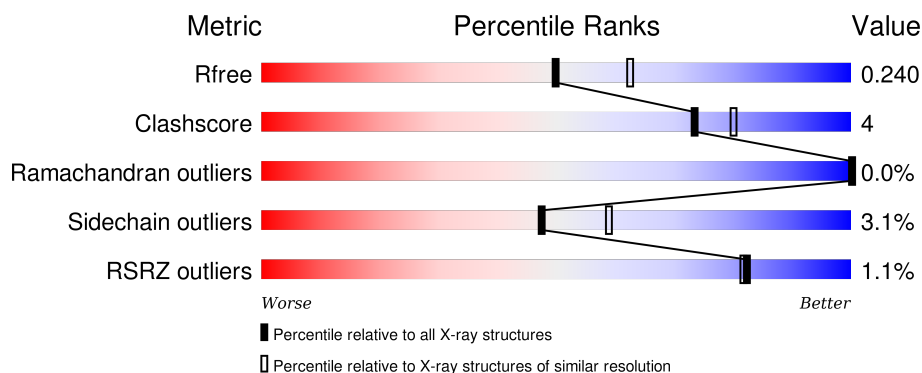
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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 ≥ 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	A	182	<div> <div></div> <div>87% 12% ..</div> </div>
1	D	182	<div> <div></div> <div>85% 13% ..</div> </div>
1	G	182	<div> <div></div> <div>92% 5% ..</div> </div>
1	J	182	<div> <div></div> <div>86% 10% ..</div> </div>
1	M	182	<div> <div></div> <div>86% 13% .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	182	
2	B	190	
2	E	190	
2	H	190	
2	K	190	
2	N	190	
2	Q	190	
3	C	14	
3	F	14	
3	I	14	
3	L	14	
3	O	14	
3	R	14	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 19066 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1478	957	240	276	5			
1	D	179	Total	C	N	O	S	0	0	0
			1461	947	237	272	5			
1	G	179	Total	C	N	O	S	0	0	0
			1444	938	233	268	5			
1	J	177	Total	C	N	O	S	0	0	0
			1431	932	231	263	5			
1	M	179	Total	C	N	O	S	0	0	0
			1434	932	234	263	5			
1	P	179	Total	C	N	O	S	0	0	0
			1436	933	234	264	5			

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	189	Total	C	N	O	S	0	0	0
			1523	960	267	290	6			
2	E	186	Total	C	N	O	S	0	0	0
			1505	947	266	286	6			
2	H	189	Total	C	N	O	S	0	0	0
			1528	963	270	289	6			
2	K	187	Total	C	N	O	S	0	0	0
			1496	944	261	285	6			
2	N	185	Total	C	N	O	S	0	0	0
			1502	946	267	283	6			
2	Q	189	Total	C	N	O	S	0	0	0
			1530	963	272	289	6			

- Molecule 3 is a protein called HLA class I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	14	Total	C	N	O	0	0	0
			115	74	22	19			
3	F	14	Total	C	N	O	0	0	0
			115	74	22	19			
3	I	11	Total	C	N	O	0	0	0
			94	60	19	15			
3	L	14	Total	C	N	O	0	0	0
			108	68	22	18			
3	O	14	Total	C	N	O	0	0	0
			115	74	22	19			
3	R	14	Total	C	N	O	0	0	0
			112	73	22	17			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
C	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892
F	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
F	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892
I	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
I	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892
L	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
L	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892
O	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
O	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892
R	5	ALA	TRP	ENGINEERED MUTATION	UNP P01892
R	13	LEU	GLN	ENGINEERED MUTATION	UNP P01892

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	53	Total	O	0	0
			53	53		
4	B	64	Total	O	0	0
			64	64		
4	D	68	Total	O	0	0
			68	68		
4	E	59	Total	O	0	0
			59	59		
4	G	48	Total	O	0	0
			48	48		
4	H	53	Total	O	0	0
			53	53		

Continued on next page...


Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	J	53	Total 53	O 53	0	0
4	K	35	Total 35	O 35	0	0
4	M	63	Total 63	O 63	0	0
4	N	43	Total 43	O 43	0	0
4	P	42	Total 42	O 42	0	0
4	Q	38	Total 38	O 38	0	0
4	C	2	Total 2	O 2	0	0
4	F	5	Total 5	O 5	0	0
4	I	3	Total 3	O 3	0	0
4	L	1	Total 1	O 1	0	0
4	O	4	Total 4	O 4	0	0
4	R	5	Total 5	O 5	0	0

3 Residue-property plots [i](#)


These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: HLA class II histocompatibility antigen, DR alpha chain

Chain A: 



- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain D: 




- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain G: 




- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain J: 




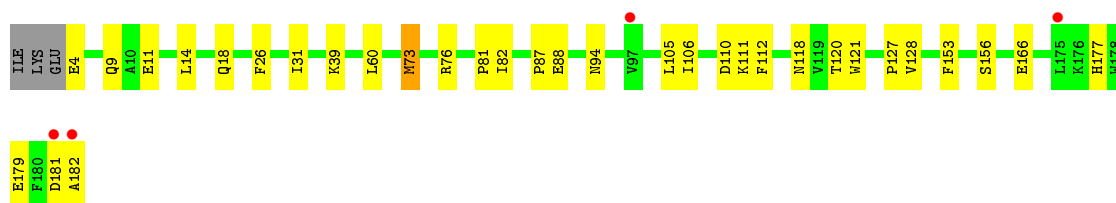
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain M: 



- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain P: 



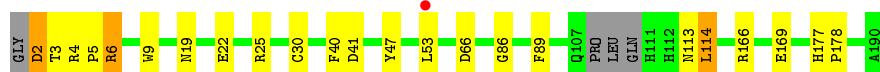
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain B: 85% 11% ..



- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain E: 86% 11% ..



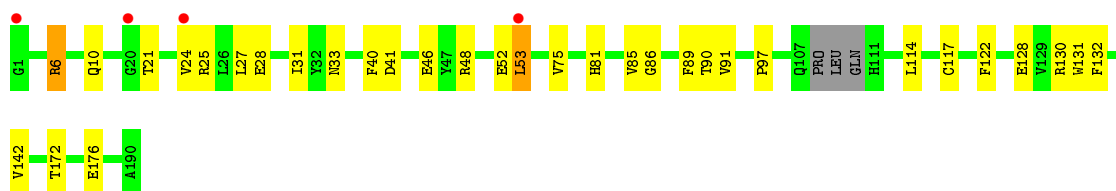
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain H: 85% 14% ..



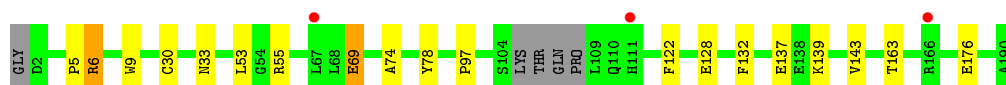
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain K: 81% 16% ..



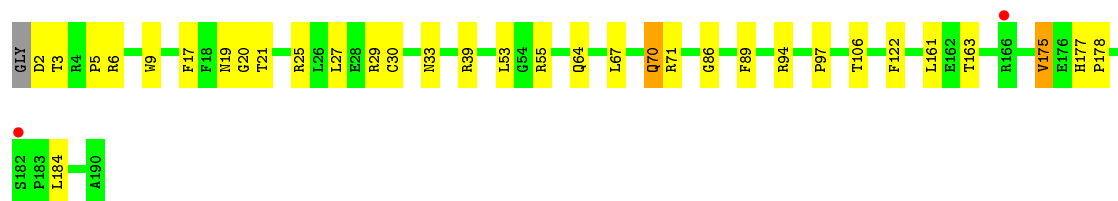
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain N: 87% 9% ..



- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain Q: 82% 16% ..



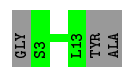
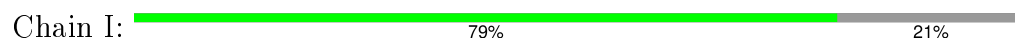
- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



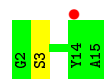
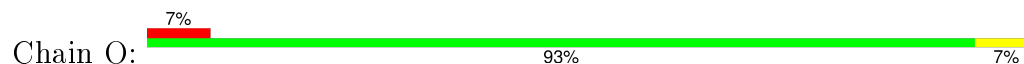
- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



- Molecule 3: HLA class I histocompatibility antigen, A-2 alpha chain



There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	95.95Å 173.19Å 96.48Å 90.00° 109.72° 90.00°	Depositor
Resolution (Å)	41.51 – 2.20 41.51 – 2.20	Depositor EDS
% Data completeness (in resolution range)	93.6 (41.51-2.20) 92.1 (41.51-2.20)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.205 , 0.239 0.204 , 0.240	Depositor DCC
R_{free} test set	7083 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	40.4	Xtriage
Anisotropy	0.253	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 7.3	EDS
Estimated twinning fraction	0.110 for l,-k,h	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 140586 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	19066	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0/1523	0.44	0/2077
1	D	0.23	0/1506	0.43	0/2056
1	G	0.23	0/1489	0.43	0/2036
1	J	0.23	0/1476	0.42	0/2017
1	M	0.24	0/1479	0.44	0/2024
1	P	0.23	0/1481	0.43	0/2025
2	B	0.23	0/1563	0.41	0/2129
2	E	0.22	0/1543	0.41	0/2098
2	H	0.23	0/1568	0.41	0/2136
2	K	0.24	0/1534	0.43	0/2088
2	N	0.24	0/1540	0.47	0/2095
2	Q	0.23	0/1570	0.42	0/2139
3	C	0.23	0/118	0.32	0/157
3	F	0.25	0/118	0.32	0/157
3	I	0.22	0/96	0.31	0/127
3	L	0.23	0/110	0.32	0/146
3	O	0.22	0/118	0.30	0/157
3	R	0.23	0/115	0.33	0/153
All	All	0.23	0/18947	0.42	0/25817

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 ⓘ

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	1478	0	1412	15	0
1	D	1461	0	1386	16	0
1	G	1444	0	1358	6	0
1	J	1431	0	1357	12	0
1	M	1434	0	1347	12	0
1	P	1436	0	1353	19	0
2	B	1523	0	1423	19	0
2	E	1505	0	1407	16	0
2	H	1528	0	1434	18	0
2	K	1496	0	1396	20	0
2	N	1502	0	1409	13	0
2	Q	1530	0	1434	21	0
3	C	115	0	106	0	0
3	F	115	0	106	0	0
3	I	94	0	89	0	0
3	L	108	0	99	0	0
3	O	115	0	106	0	0
3	R	112	0	104	0	0
4	A	53	0	0	0	0
4	B	64	0	0	0	0
4	C	2	0	0	0	0
4	D	68	0	0	0	0
4	E	59	0	0	0	0
4	F	5	0	0	0	0
4	G	48	0	0	0	0
4	H	53	0	0	0	0
4	I	3	0	0	0	0
4	J	53	0	0	0	0
4	K	35	0	0	0	0
4	L	1	0	0	0	0
4	M	63	0	0	0	0
4	N	43	0	0	0	0
4	O	4	0	0	0	0
4	P	42	0	0	0	0
4	Q	38	0	0	0	0
4	R	5	0	0	0	0
All	All	19066	0	17326	148	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 148 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:G:111:LYS:NZ	1:J:88:GLU:OE2	2.24	0.71
2:K:130:ARG:NH1	2:K:176:GLU:OE1	2.24	0.71
1:M:111:LYS:NZ	1:P:88:GLU:OE2	2.26	0.69
2:N:55:ARG:HH12	2:Q:55:ARG:CZ	2.06	0.69
2:Q:175:VAL:HG13	2:Q:184:LEU:HB2	1.76	0.67

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	178/182 (98%)	175 (98%)	3 (2%)	0	100	100
1	D	177/182 (97%)	176 (99%)	1 (1%)	0	100	100
1	G	177/182 (97%)	175 (99%)	2 (1%)	0	100	100
1	J	175/182 (96%)	174 (99%)	1 (1%)	0	100	100
1	M	177/182 (97%)	175 (99%)	2 (1%)	0	100	100
1	P	177/182 (97%)	176 (99%)	1 (1%)	0	100	100
2	B	187/190 (98%)	182 (97%)	5 (3%)	0	100	100
2	E	182/190 (96%)	177 (97%)	5 (3%)	0	100	100
2	H	187/190 (98%)	181 (97%)	6 (3%)	0	100	100
2	K	183/190 (96%)	178 (97%)	5 (3%)	0	100	100
2	N	181/190 (95%)	178 (98%)	3 (2%)	0	100	100
2	Q	187/190 (98%)	182 (97%)	5 (3%)	0	100	100
3	C	12/14 (86%)	10 (83%)	1 (8%)	1 (8%)	1	0
3	F	12/14 (86%)	12 (100%)	0	0	100	100
3	I	9/14 (64%)	9 (100%)	0	0	100	100
3	L	12/14 (86%)	11 (92%)	1 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	O	12/14 (86%)	12 (100%)	0	0	100	100
3	R	12/14 (86%)	12 (100%)	0	0	100	100
All	All	2237/2316 (97%)	2195 (98%)	41 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	14	TYR

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	A	164/166 (99%)	159 (97%)	5 (3%)	48	60
1	D	161/166 (97%)	157 (98%)	4 (2%)	55	67
1	G	157/166 (95%)	151 (96%)	6 (4%)	40	49
1	J	156/166 (94%)	151 (97%)	5 (3%)	46	57
1	M	154/166 (93%)	148 (96%)	6 (4%)	39	48
1	P	155/166 (93%)	150 (97%)	5 (3%)	46	57
2	B	165/171 (96%)	157 (95%)	8 (5%)	31	37
2	E	163/171 (95%)	158 (97%)	5 (3%)	47	59
2	H	166/171 (97%)	163 (98%)	3 (2%)	66	79
2	K	161/171 (94%)	157 (98%)	4 (2%)	55	67
2	N	163/171 (95%)	158 (97%)	5 (3%)	47	59
2	Q	166/171 (97%)	161 (97%)	5 (3%)	48	60
3	C	10/10 (100%)	10 (100%)	0	100	100
3	F	10/10 (100%)	10 (100%)	0	100	100
3	I	9/10 (90%)	9 (100%)	0	100	100
3	L	9/10 (90%)	9 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	O	10/10 (100%)	9 (90%)	1 (10%)	9	8
3	R	9/10 (90%)	9 (100%)	0	100	100
All	All	1988/2082 (96%)	1926 (97%)	62 (3%)	47	59

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	114	LEU
1	J	128	VAL
2	Q	70	GLN
1	J	46	GLU
2	K	6	ARG

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

Mol	Chain	Res	Type
2	H	19	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 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 ⓘ

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, 95th 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	180/182 (98%)	0.16	2 (1%) 82 82	17, 32, 50, 61	0
1	D	179/182 (98%)	0.03	0 100 100	16, 26, 42, 49	0
1	G	179/182 (98%)	0.05	2 (1%) 82 82	17, 30, 48, 57	0
1	J	177/182 (97%)	-0.06	0 100 100	22, 32, 46, 55	0
1	M	179/182 (98%)	-0.03	1 (0%) 90 90	14, 27, 45, 55	0
1	P	179/182 (98%)	0.11	4 (2%) 65 64	23, 35, 50, 75	0
2	B	189/190 (99%)	0.12	0 100 100	17, 29, 48, 59	0
2	E	186/190 (97%)	0.08	1 (0%) 91 91	14, 28, 48, 54	0
2	H	189/190 (99%)	0.10	3 (1%) 74 73	16, 30, 56, 71	0
2	K	187/190 (98%)	0.24	4 (2%) 67 65	25, 40, 61, 73	0
2	N	185/190 (97%)	0.12	3 (1%) 74 73	13, 30, 51, 72	0
2	Q	189/190 (99%)	0.12	2 (1%) 82 82	23, 36, 52, 58	0
3	C	14/14 (100%)	0.38	1 (7%) 19 18	23, 34, 55, 61	0
3	F	14/14 (100%)	0.18	1 (7%) 19 18	23, 33, 51, 54	0
3	I	11/14 (78%)	-0.08	0 100 100	25, 33, 45, 46	0
3	L	14/14 (100%)	0.55	1 (7%) 19 18	33, 40, 51, 52	0
3	O	14/14 (100%)	0.61	1 (7%) 19 18	26, 34, 47, 53	0
3	R	14/14 (100%)	0.13	0 100 100	32, 38, 57, 59	0
All	All	2279/2316 (98%)	0.10	26 (1%) 82 82	13, 32, 52, 75	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	2	GLY	5.3
1	P	182	ALA	4.7
1	P	97	VAL	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	G	51	PHE	3.5
1	M	182	ALA	3.2

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.