WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 10, 2021 – 06:33 PM EDT

PDB ID	:	3DXA
Title	:	Crystal Structure of the DM1 TCR in complex with HLA-B*4405 and decamer EBV antigen
Authors	:	Archbold, J.K.; Macdonald, W.A.; Gras, S.; Rossjohn, J.
Deposited on	:	2008-07-23
Resolution	:	3.50 Å(reported)

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

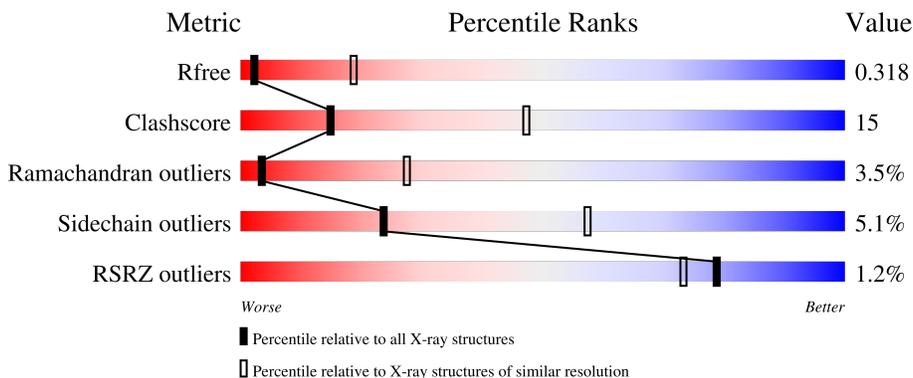|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.50 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	Percentile Ranks	Value
Rfree		0.318
Clashscore		15
Ramachandran outliers		3.5%
Sidechain outliers		5.1%
RSRZ outliers		1.2%

Worse — *Better*

▌ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	276	65% 29% 5% •
1	F	276	67% 27% 6%
1	K	276	5% 69% 25% 5% •
2	B	99	69% 26% 5%
2	G	99	69% 26% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	99	3% / 68% / 27% / 5%
3	C	10	30% / 40% / 30%
3	H	10	50% / 40% / 10%
3	M	10	50% / 20% / 20% / 10%
4	D	199	56% / 31% / 10% / •
4	I	199	2% / 63% / 27% / 8% / •
4	N	199	3% / 65% / 26% / 8% / •
5	E	244	77% / 21% / •
5	J	244	76% / 22% / •
5	O	244	% / 77% / 21% / •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 20171 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 I histocompatibility complex HLA-B*4402.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			
1	F	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			
1	K	276	Total	C	N	O	S	0	0	0
			2258	1408	407	436	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	116	TYR	ASP	engineered mutation	UNP P30481
F	116	TYR	ASP	engineered mutation	UNP P30481
K	116	TYR	ASP	engineered mutation	UNP P30481

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			
2	G	99	Total	C	N	O	S	0	1	0
			832	530	140	159	3			
2	L	99	Total	C	N	O	S	0	0	0
			829	528	140	158	3			

- Molecule 3 is a protein called EBV decapeptide epitope.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	0	0	0
			91	59	14	18			
3	H	10	Total	C	N	O	0	0	0
			90	59	14	17			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	10	Total C N O 91 59 14 18				0	0	0

• Molecule 4 is a protein called DM1 T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	199	Total C N O S 1560 975 262 313 10					0	0	0
4	I	199	Total C N O S 1560 975 262 313 10					0	0	0
4	N	199	Total C N O S 1560 975 262 313 10					0	0	0

• Molecule 5 is a protein called DM1 T cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	244	Total C N O S 1985 1244 356 379 6					0	0	0
5	J	244	Total C N O S 1985 1244 356 379 6					0	0	0
5	O	244	Total C N O S 1985 1244 356 379 6					0	0	0

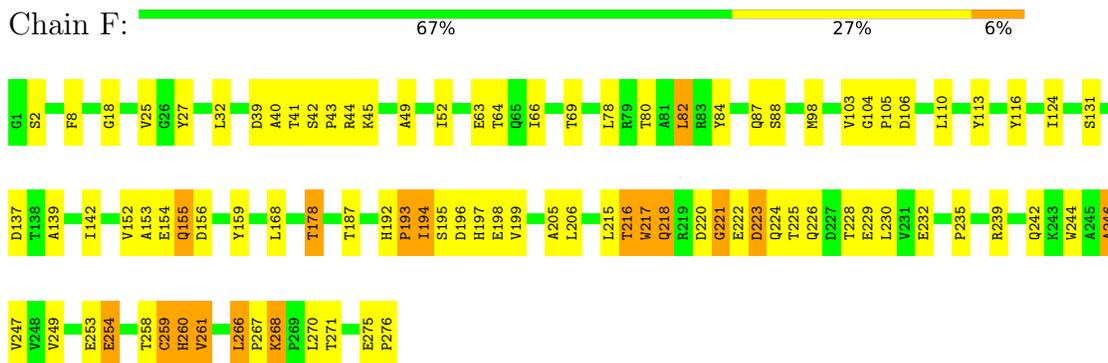# 3 Residue-property plots (i)

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.
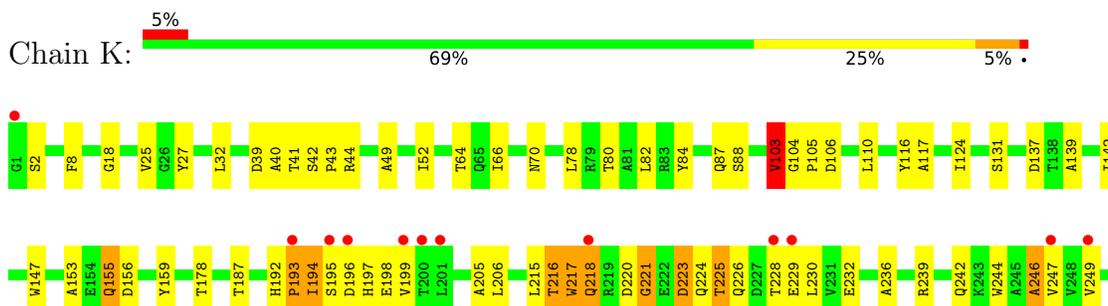
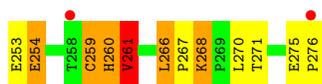• Molecule 1: HLA class I histocompatibility complex HLA-B*4402
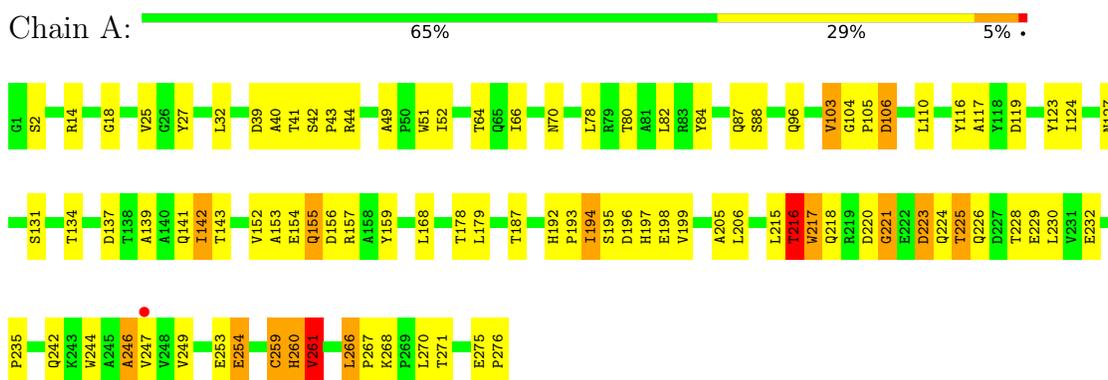
Chain A:

• Molecule 1: HLA class I histocompatibility complex HLA-B*4402
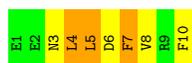
Chain F:

• Molecule 1: HLA class I histocompatibility complex HLA-B*4402

Chain K:

E253 E254 T258 C259 H260 V261 L266 P267 K268 P269 L270 T271 E275 P276

• Molecule 2: Beta-2-microglobulin

Chain B:

| 69% | 26% | 5% |

I1 Q2 R3 T4 A15 E16 N17 G18 K19 S20 L23 N24 H31 V37 D38 L39 L40 K41 R45 I46 V49 L54 K58 D59 W60 L64 L65 T68 E69 F70 T71 P73 E74 Y78 A79 V82 M99

• Molecule 2: Beta-2-microglobulin

Chain G:

| 69% | 26% | 5% |

I1 Q2 R3 T4 A15 E16 N17 G18 K19 S20 L23 N24 V37 D38 L39 L40 K41 R45 I46 V49 L54 K58 L64 L65 T68 E69 F70 T71 P72 T73 E74 Y78 A79 C80 R81 V82 M99

• Molecule 2: Beta-2-microglobulin

Chain L:

3%

| 68% | 27% | 5% |

I1 Q2 R3 T4 A15 E16 N17 G18 K19 S20 L23 N24 C25 V37 D38 L39 L40 K41 R45 I46 V49 L54 K58 D59 W60 L64 T68 E69 F70 T71 P72 T73 E74 Y78 A79 C80 R81 V82 D98 M99

• Molecule 3: EBV decapeptide epitope

Chain C:

| 30% | 40% | 30% |

E1 E2 N3 L4 L5 D6 F7 V8 R9 F10

• Molecule 3: EBV decapeptide epitope

Chain H:

| 50% | 40% | 10% |

E1 L4 L5 D6 F7 V8 R9 F10

• Molecule 3: EBV decapeptide epitope

Chain M:

| 50% | 20% | 20% | 10% |

E1 L4 L5 D6 F7 V8 R9 F10

• Molecule 4: DM1 T cell receptor alpha chain

Chain D:

| 56% | 31% | 10% | • |

• Molecule 4: DM1 T cell receptor alpha chain

Chain I:

2% 63% 27% 8%

• Molecule 4: DM1 T cell receptor alpha chain

Chain N:

3% 65% 26% 8%

• Molecule 5: DM1 T cell receptor beta chain

Chain E:

77% 21%

• Molecule 5: DM1 T cell receptor beta chain

Chain J:

76% 22%
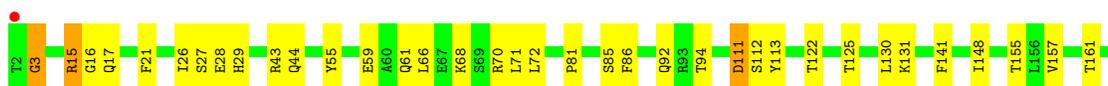
• Molecule 5: DM1 T cell receptor beta chain

Chain O:

4　Data and refinement statistics　ⓘ

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	121.95Å　121.95Å　695.91Å 90.00°　　90.00°　　120.00°	Depositor
Resolution (Å)	100.00　–　3.50 49.37　–　3.50	Depositor EDS
% Data completeness (in resolution range)	97.4 (100.00-3.50) 97.5 (49.37-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.21 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.286　,　0.330 0.283　,　0.318	Depositor DCC
R_{free} test set	1970 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	73.8	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 59.3	EDS
L-test for twinning [2]	$<|L|>=0.42$, $<L^2>=0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	20171	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.62% 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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	0.42	0/2319	0.75	2/3154 (0.1%)
1	F	0.43	0/2319	0.65	0/3154
1	K	0.39	0/2319	0.65	1/3154 (0.0%)
2	B	0.41	0/852	0.66	0/1152
2	G	0.40	0/858	0.61	0/1160
2	L	0.36	0/852	0.58	0/1152
3	C	0.55	0/92	0.91	0/121
3	H	0.49	0/91	0.97	0/121
3	M	0.46	0/92	1.00	1/121 (0.8%)
4	D	0.50	0/1596	0.93	4/2171 (0.2%)
4	I	0.44	0/1596	0.66	0/2171
4	N	0.41	0/1596	0.68	1/2171 (0.0%)
5	E	0.43	0/2038	0.57	1/2767 (0.0%)
5	J	0.42	0/2038	0.56	0/2767
5	O	0.40	0/2038	0.55	0/2767
All	All	0.42	0/20696	0.67	10/28103 (0.0%)

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	20
1	F	0	18
1	K	0	18
2	B	0	9
2	G	0	9
2	L	0	9
4	D	0	32
4	I	0	21
4	N	0	22
All	All	0	158

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	173	VAL	CG1-CB-CG2	7.37	122.69	110.90
4	N	166	VAL	CG1-CB-CG2	7.15	122.34	110.90
5	E	244	VAL	CG1-CB-CG2	6.49	121.28	110.90
1	K	103	VAL	CG1-CB-CG2	6.48	121.27	110.90
1	A	261	VAL	CG1-CB-CG2	6.00	120.49	110.90

There are no chirality outliers.

5 of 158 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	PRO	Peptide
1	A	106	ASP	Peptide
1	A	14	ARG	Peptide
1	A	18	GLY	Peptide
1	A	39	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	2258	0	2128	85	0
1	F	2258	0	2128	82	0
1	K	2258	0	2128	73	0
2	B	829	0	794	22	0
2	G	832	0	799	21	0
2	L	829	0	794	22	0
3	C	91	0	86	15	0
3	H	90	0	86	17	0
3	M	91	0	86	10	0
4	D	1560	0	1485	63	0
4	I	1560	0	1485	56	0
4	N	1560	0	1485	52	0
5	E	1985	0	1879	46	0
5	J	1985	0	1879	49	0
5	O	1985	0	1879	44	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	20171	0	19121	596	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 15.

The worst 5 of 596 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:N:18:ALA:CB	4:N:19:ALA:HA	1.83	1.08
4:I:18:ALA:CB	4:I:19:ALA:HA	1.84	1.06
1:A:246:ALA:HB3	1:A:247:VAL:HA	1.42	1.02
1:F:246:ALA:HB3	1:F:247:VAL:HA	1.40	1.01
1:K:246:ALA:HB3	1:K:247:VAL:HA	1.40	1.01

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	A	274/276 (99%)	229 (84%)	35 (13%)	10 (4%)	3	26
1	F	274/276 (99%)	234 (85%)	33 (12%)	7 (3%)	5	33
1	K	274/276 (99%)	232 (85%)	34 (12%)	8 (3%)	4	31
2	B	97/99 (98%)	83 (86%)	13 (13%)	1 (1%)	15	54
2	G	98/99 (99%)	84 (86%)	13 (13%)	1 (1%)	15	54
2	L	97/99 (98%)	84 (87%)	12 (12%)	1 (1%)	15	54
3	C	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	5
3	H	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	5
3	M	8/10 (80%)	6 (75%)	1 (12%)	1 (12%)	0	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	197/199 (99%)	149 (76%)	33 (17%)	15 (8%)	1	10
4	I	197/199 (99%)	150 (76%)	34 (17%)	13 (7%)	1	13
4	N	197/199 (99%)	151 (77%)	33 (17%)	13 (7%)	1	13
5	E	242/244 (99%)	217 (90%)	20 (8%)	5 (2%)	7	38
5	J	242/244 (99%)	218 (90%)	19 (8%)	5 (2%)	7	38
5	O	242/244 (99%)	216 (89%)	23 (10%)	3 (1%)	13	50
All	All	2455/2484 (99%)	2065 (84%)	305 (12%)	85 (4%)	3	27

5 of 85 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	198	GLU
1	A	246	ALA
4	D	11	MET
4	D	17	ARG
4	D	18	ALA

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	237/237 (100%)	229 (97%)	8 (3%)	37	68
1	F	237/237 (100%)	229 (97%)	8 (3%)	37	68
1	K	237/237 (100%)	228 (96%)	9 (4%)	33	65
2	B	94/94 (100%)	89 (95%)	5 (5%)	22	55
2	G	95/94 (101%)	89 (94%)	6 (6%)	18	51
2	L	94/94 (100%)	88 (94%)	6 (6%)	17	50
3	C	10/10 (100%)	8 (80%)	2 (20%)	1	7
3	H	10/10 (100%)	9 (90%)	1 (10%)	7	32
3	M	10/10 (100%)	8 (80%)	2 (20%)	1	7
4	D	178/178 (100%)	166 (93%)	12 (7%)	16	48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	I	178/178 (100%)	166 (93%)	12 (7%)	16	48
4	N	178/178 (100%)	165 (93%)	13 (7%)	14	45
5	E	216/216 (100%)	206 (95%)	10 (5%)	27	61
5	J	216/216 (100%)	207 (96%)	9 (4%)	30	63
5	O	216/216 (100%)	207 (96%)	9 (4%)	30	63
All	All	2206/2205 (100%)	2094 (95%)	112 (5%)	24	57

5 of 112 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	84	THR
5	O	188	GLN
5	J	188	GLN
5	O	125	THR
4	N	126	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	96	GLN
5	O	61	GLN
1	K	144	GLN
4	N	20	ASN
5	O	167	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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(Å²)	Q<0.9
1	A	276/276 (100%)	-0.13	1 (0%)	92	90	64, 64, 64, 64	13 (4%)
1	F	276/276 (100%)	-0.16	0	100	100	64, 64, 64, 64	12 (4%)
1	K	276/276 (100%)	0.14	14 (5%)	28	25	64, 64, 64, 64	16 (5%)
2	B	99/99 (100%)	0.05	0	100	100	64, 64, 64, 64	6 (6%)
2	G	99/99 (100%)	-0.11	0	100	100	64, 64, 64, 64	5 (5%)
2	L	99/99 (100%)	0.31	3 (3%)	50	44	64, 64, 64, 64	9 (9%)
3	C	10/10 (100%)	-0.09	0	100	100	64, 64, 64, 64	0
3	H	10/10 (100%)	0.08	0	100	100	64, 64, 64, 64	3 (30%)
3	M	10/10 (100%)	-0.13	0	100	100	64, 64, 64, 64	0
4	D	199/199 (100%)	-0.22	0	100	100	64, 64, 64, 65	13 (6%)
4	I	199/199 (100%)	0.03	4 (2%)	65	60	64, 64, 64, 65	17 (8%)
4	N	199/199 (100%)	0.12	5 (2%)	57	51	64, 64, 64, 65	21 (10%)
5	E	244/244 (100%)	-0.25	1 (0%)	92	90	25, 64, 64, 65	21 (8%)
5	J	244/244 (100%)	-0.21	1 (0%)	92	90	25, 64, 64, 65	20 (8%)
5	O	244/244 (100%)	-0.02	2 (0%)	86	81	25, 64, 64, 65	20 (8%)
All	All	2484/2484 (100%)	-0.06	31 (1%)	79	73	25, 64, 64, 65	176 (7%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	O	209	VAL	4.2
1	K	228	THR	3.9
5	E	2	THR	3.8
1	K	201	LEU	3.7
4	I	215	PRO	3.7

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.