Full wwPDB X-ray Structure Validation Report ⓘ

Jan 23, 2025 – 04:23 PM JST

PDB ID	:	8K4T
Title	:	Crystal structure of HLA-A*11:01 in complex with KRAS G12C peptide (VVVGACGVGK)
Authors	:	Xu, Y.
Deposited on	:	2023-07-20
Resolution	:	2.30 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:

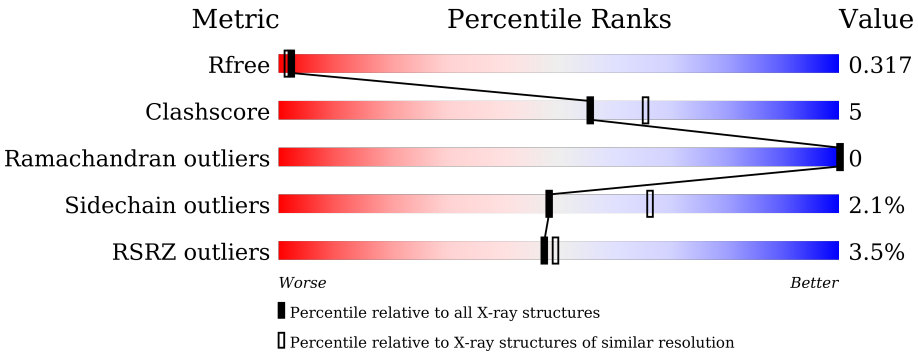
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.30 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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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 >=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	3% 91% 8%
1	D	276	4% 85% 14%
2	B	122	2% 70% 10% 20%
2	E	122	2% 62% 16% 21%
3	C	10	30% 60% 40%
3	F	10	20% 80% 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6213 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 antigen, A alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total	C	N	O	S	0	0	0
			2238	1391	405	433	9			
1	D	275	Total	C	N	O	S	0	0	0
			2244	1394	408	433	9			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MET	-	initiating methionine	UNP P04439
A	33	TYR	PHE	variant	UNP P04439
A	114	ASP	ALA	variant	UNP P04439
A	129	PRO	SER	variant	UNP P04439
D	24	MET	-	initiating methionine	UNP P04439
D	33	TYR	PHE	variant	UNP P04439
D	114	ASP	ALA	variant	UNP P04439
D	129	PRO	SER	variant	UNP P04439

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	97	Total	C	N	O	S	0	0	0
			812	519	138	153	2			
2	E	96	Total	C	N	O	S	0	0	0
			797	510	133	152	2			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
B	100	GLY	-	expression tag	UNP P61769
B	101	SER	-	expression tag	UNP P61769
B	102	GLY	-	expression tag	UNP P61769

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	103	GLY	-	expression tag	UNP P61769
B	104	SER	-	expression tag	UNP P61769
B	105	GLY	-	expression tag	UNP P61769
B	106	ALA	-	expression tag	UNP P61769
B	107	GLY	-	expression tag	UNP P61769
B	108	LEU	-	expression tag	UNP P61769
B	109	ASN	-	expression tag	UNP P61769
B	110	ASP	-	expression tag	UNP P61769
B	111	ILE	-	expression tag	UNP P61769
B	112	PHE	-	expression tag	UNP P61769
B	113	GLU	-	expression tag	UNP P61769
B	114	ALA	-	expression tag	UNP P61769
B	115	GLN	-	expression tag	UNP P61769
B	116	LYS	-	expression tag	UNP P61769
B	117	ILE	-	expression tag	UNP P61769
B	118	GLU	-	expression tag	UNP P61769
B	119	TRP	-	expression tag	UNP P61769
B	120	HIS	-	expression tag	UNP P61769
B	121	GLU	-	expression tag	UNP P61769
E	0	MET	-	initiating methionine	UNP P61769
E	100	GLY	-	expression tag	UNP P61769
E	101	SER	-	expression tag	UNP P61769
E	102	GLY	-	expression tag	UNP P61769
E	103	GLY	-	expression tag	UNP P61769
E	104	SER	-	expression tag	UNP P61769
E	105	GLY	-	expression tag	UNP P61769
E	106	ALA	-	expression tag	UNP P61769
E	107	GLY	-	expression tag	UNP P61769
E	108	LEU	-	expression tag	UNP P61769
E	109	ASN	-	expression tag	UNP P61769
E	110	ASP	-	expression tag	UNP P61769
E	111	ILE	-	expression tag	UNP P61769
E	112	PHE	-	expression tag	UNP P61769
E	113	GLU	-	expression tag	UNP P61769
E	114	ALA	-	expression tag	UNP P61769
E	115	GLN	-	expression tag	UNP P61769
E	116	LYS	-	expression tag	UNP P61769
E	117	ILE	-	expression tag	UNP P61769
E	118	GLU	-	expression tag	UNP P61769
E	119	TRP	-	expression tag	UNP P61769
E	120	HIS	-	expression tag	UNP P61769
E	121	GLU	-	expression tag	UNP P61769

- Molecule 3 is a protein called KRAS G12C peptide (VVVGACGVGK).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	10	Total	C	N	O	S	0	0	0
			61	38	11	11	1			
3	F	10	Total	C	N	O	S	0	0	0
			61	38	11	11	1			

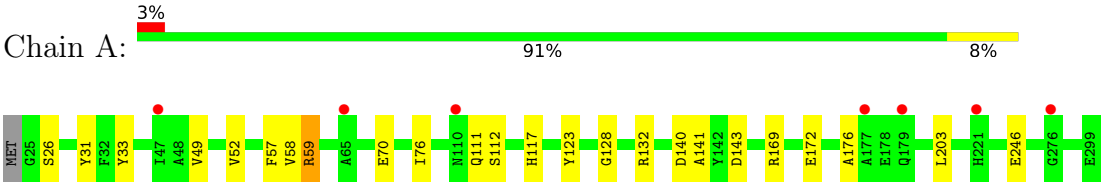
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	6	CYS	GLY	engineered mutation	UNP P01116
F	6	CYS	GLY	engineered mutation	UNP P01116

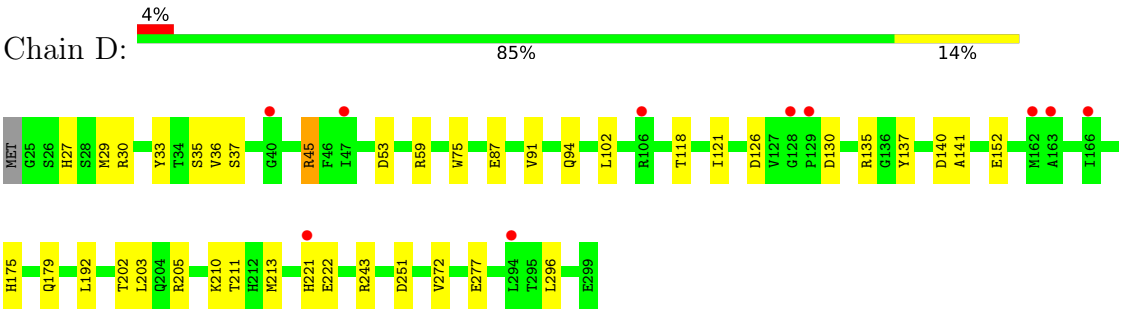
3 Residue-property plots (i)

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.

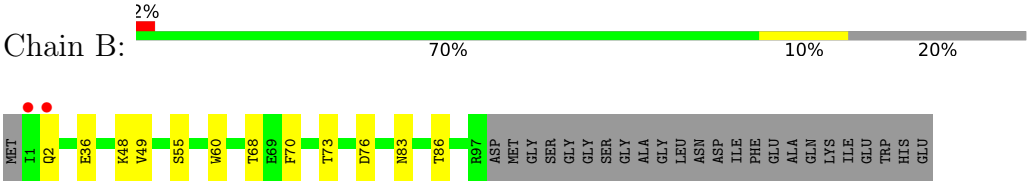
• Molecule 1: HLA class I histocompatibility antigen, A alpha chain



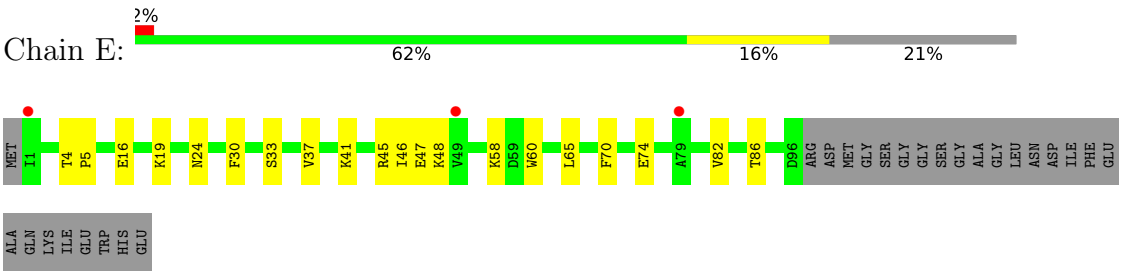
• Molecule 1: HLA class I histocompatibility antigen, A alpha chain



• Molecule 2: Beta-2-microglobulin

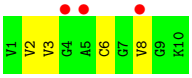


• Molecule 2: Beta-2-microglobulin



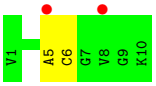
• Molecule 3: KRAS G12C peptide (VVVGACGVGK)

Chain C:



• Molecule 3: KRAS G12C peptide (VVVGACGVGK)

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.71Å 63.58Å 75.51Å 105.86° 92.66° 97.70°	Depositor
Resolution (Å)	28.78 – 2.30 28.78 – 2.30	Depositor EDS
% Data completeness (in resolution range)	94.6 (28.78-2.30) 87.5 (28.78-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.16 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.258 , 0.316 0.260 , 0.317	Depositor DCC
R_{free} test set	35359 reflections (5.41%)	wwPDB-VP
Wilson B-factor (Å^2)	45.1	Xtriage
Anisotropy	0.473	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 35.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6213	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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	0/2299	0.42	0/3122
1	D	0.24	0/2305	0.43	0/3129
2	B	0.24	0/835	0.43	0/1131
2	E	0.29	0/820	0.47	0/1113
3	C	0.29	0/60	0.48	0/78
3	F	0.34	0/60	0.42	0/78
All	All	0.25	0/6379	0.43	0/8651

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	2238	0	2077	15	0
1	D	2244	0	2088	21	0
2	B	812	0	781	7	0
2	E	797	0	757	16	0
3	C	61	0	70	3	1
3	F	61	0	70	1	1
All	All	6213	0	5843	55	1

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 5.

All (55) 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:E:16:GLU:HB2	2:E:19:LYS:HD2	1.63	0.78
2:E:58:LYS:H	2:E:58:LYS:HE2	1.55	0.72
1:D:211:THR:HB	1:D:296:LEU:HD21	1.77	0.66
1:D:33:TYR:HB2	1:D:121:ILE:HB	1.80	0.63
1:D:179:GLN:HB3	3:F:5:ALA:HB1	1.83	0.60
2:E:48:LYS:HG2	2:E:48:LYS:O	2.03	0.59
1:D:202:THR:HA	1:D:205:ARG:HD3	1.87	0.57
1:A:52:VAL:HG11	1:A:203:LEU:HD13	1.86	0.57
2:B:36:GLU:HG3	2:B:83:ASN:HB3	1.86	0.57
1:D:29:MET:HB2	1:D:192:LEU:HD13	1.88	0.56
1:A:49:VAL:HG21	2:B:55:SER:HB3	1.88	0.54
2:B:2:GLN:HB2	2:B:86:THR:HG22	1.88	0.54
2:E:37:VAL:HG22	2:E:82:VAL:HG22	1.91	0.53
2:B:49:VAL:HG22	2:B:68:THR:HB	1.91	0.52
1:D:130:ASP:OD1	1:D:130:ASP:N	2.39	0.52
2:E:41:LYS:HB2	2:E:46:ILE:HD11	1.92	0.52
1:D:221:HIS:CD2	1:D:222:GLU:HG3	2.45	0.51
2:E:47:GLU:HG3	2:E:48:LYS:H	1.74	0.51
2:E:46:ILE:HG22	2:E:47:GLU:O	2.10	0.51
1:A:31:TYR:CE1	3:C:2:VAL:HG12	2.47	0.50
1:A:111:GLN:OE1	1:A:117:HIS:NE2	2.45	0.50
1:D:87:GLU:O	1:D:91:VAL:HG22	2.12	0.50
1:D:251:ASP:HB3	1:D:272:VAL:HG12	1.93	0.50
1:A:246:GLU:HB3	1:D:175:HIS:CE1	2.48	0.49
2:E:4:THR:HG22	2:E:86:THR:HB	1.94	0.49
2:B:48:LYS:HD2	2:B:48:LYS:N	2.28	0.48
1:D:27:HIS:ND1	1:D:53:ASP:OD2	2.45	0.47
2:E:47:GLU:O	2:E:48:LYS:HB3	2.14	0.47
1:D:243:ARG:NH2	1:D:277:GLU:OE1	2.48	0.47
2:E:45:ARG:HG2	2:E:45:ARG:HH11	1.80	0.46
1:A:176:ALA:HB2	3:C:8:VAL:HG11	1.97	0.46
1:D:37:SER:HB3	1:D:102:LEU:HD13	1.97	0.46
1:D:75:TRP:CZ2	1:D:203:LEU:HD11	2.51	0.46
1:A:169:ARG:HA	1:A:172:GLU:HB2	1.97	0.45
1:D:33:TYR:CE2	1:D:94:GLN:HG2	2.51	0.45
1:D:251:ASP:HB3	1:D:272:VAL:CG1	2.47	0.45
2:E:24:ASN:HB3	2:E:65:LEU:HD11	1.98	0.45
1:A:57:PHE:CD2	1:A:58:VAL:HG13	2.52	0.44
2:E:45:ARG:HH12	2:E:47:GLU:CA	2.31	0.44
1:D:141:ALA:HB2	2:E:60:TRP:CE2	2.53	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:73:THR:OG1	2:B:76:ASP:OD2	2.31	0.44
1:D:135:ARG:HD2	1:D:137:TYR:OH	2.17	0.44
1:A:26:SER:OG	1:A:128:GLY:HA2	2.18	0.43
1:D:36:VAL:HG22	1:D:118:THR:HG23	2.00	0.43
2:E:47:GLU:CG	2:E:48:LYS:H	2.31	0.43
1:D:30:ARG:NH2	1:D:126:ASP:OD1	2.40	0.43
1:A:141:ALA:HB2	2:B:60:TRP:CE2	2.54	0.43
2:E:45:ARG:HG2	2:E:45:ARG:NH1	2.34	0.42
1:A:59:ARG:NH2	1:A:70:GLU:OE1	2.53	0.42
2:E:5:PRO:HB3	2:E:30:PHE:HB3	2.02	0.42
1:A:76:ILE:HG13	1:A:76:ILE:O	2.21	0.41
1:A:117:HIS:HD2	1:A:143:ASP:OD2	2.04	0.41
1:A:123:TYR:CZ	3:C:3:VAL:HG12	2.56	0.41
1:A:33:TYR:HH	1:A:123:TYR:HH	1.64	0.40
1:D:35:SER:HA	1:D:45:ARG:O	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:6:CYS:CB	3:F:6:CYS:SG[1_565]	1.84	0.36

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	273/276 (99%)	266 (97%)	7 (3%)	0	100	100
1	D	273/276 (99%)	265 (97%)	8 (3%)	0	100	100
2	B	95/122 (78%)	93 (98%)	2 (2%)	0	100	100
2	E	94/122 (77%)	92 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	C	8/10 (80%)	7 (88%)	1 (12%)	0	100 100
3	F	8/10 (80%)	7 (88%)	1 (12%)	0	100 100
All	All	751/816 (92%)	730 (97%)	21 (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	231/233 (99%)	227 (98%)	4 (2%)	56 72
1	D	232/233 (100%)	226 (97%)	6 (3%)	41 58
2	B	92/110 (84%)	91 (99%)	1 (1%)	70 83
2	E	90/110 (82%)	87 (97%)	3 (3%)	33 48
3	C	6/6 (100%)	6 (100%)	0	100 100
3	F	6/6 (100%)	6 (100%)	0	100 100
All	All	657/698 (94%)	643 (98%)	14 (2%)	48 66

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	ARG
1	A	112	SER
1	A	132	ARG
1	A	140	ASP
2	B	70	PHE
1	D	45	ARG
1	D	59	ARG
1	D	140	ASP
1	D	152	GLU
1	D	210	LYS
1	D	213	MET
2	E	33	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	70	PHE
2	E	74	GLU

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

Mol	Chain	Res	Type
2	B	83	ASN

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 oligosaccharides 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, 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	275/276 (99%)	0.59	7 (2%) 58 59	48, 54, 61, 66	0
1	D	275/276 (99%)	0.74	10 (3%) 46 48	54, 59, 65, 68	0
2	B	97/122 (79%)	0.65	2 (2%) 63 64	47, 53, 65, 67	0
2	E	96/122 (78%)	0.75	3 (3%) 51 53	55, 63, 68, 69	0
3	C	10/10 (100%)	1.47	3 (30%) 1 1	53, 57, 60, 62	0
3	F	10/10 (100%)	1.41	2 (20%) 3 4	57, 60, 61, 64	0
All	All	763/816 (93%)	0.69	27 (3%) 47 49	47, 58, 66, 69	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	8	VAL	3.7
1	D	221	HIS	3.2
1	D	166	ILE	3.2
3	C	4	GLY	3.1
2	B	1	ILE	2.8
1	D	163	ALA	2.8
1	D	162	MET	2.8
3	C	5	ALA	2.7
3	F	5	ALA	2.7
1	A	65	ALA	2.6
1	D	40	GLY	2.5
1	D	47	ILE	2.5
1	A	276	GLY	2.4
1	A	179	GLN	2.3
1	A	110	ASN	2.3
2	B	2	GLN	2.3
1	D	129	PRO	2.2
1	D	106	ARG	2.2
1	A	221	HIS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	8	VAL	2.2
1	D	294	LEU	2.2
2	E	79	ALA	2.2
1	A	177	ALA	2.2
2	E	1	ILE	2.1
1	D	128	GLY	2.1
1	A	47	ILE	2.1
2	E	49	VAL	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.