



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2024 – 03:58 pm BST

PDB ID : 8QFY
Title : Crystal structure of high affinity TCR in complex with pHLA harbouring bacterial peptide
Authors : Pengelly, R.J.; Robinson, R.A.
Deposited on : 2023-09-05
Resolution : 2.33 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

<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
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36.2
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) : 2.36.2

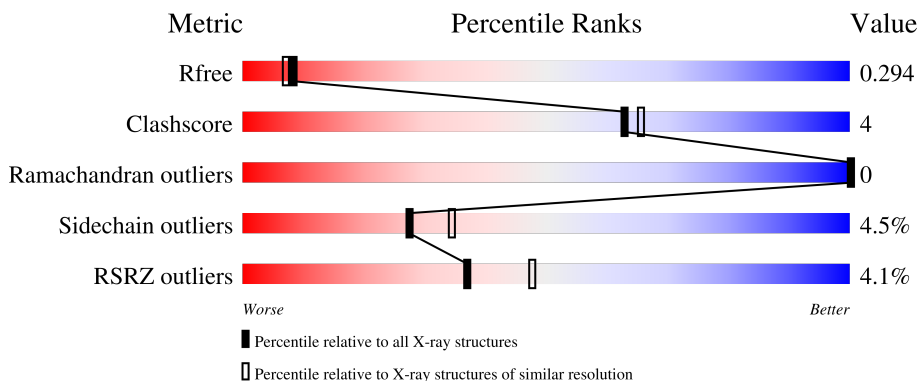
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.33 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.







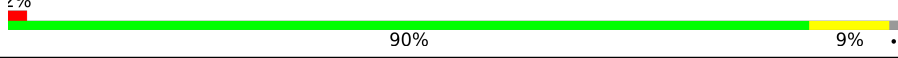
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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 $\leq 5\%$. 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	AAA	276	 11% 82% 8% 9%
1	FFF	276	 0% 83% 12% 3%
2	BBB	100	 7% 91% 9%
2	GGG	100	 3% 88% 12%
3	CCC	9	 78% 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	HHH	9	 78% 22%
4	DDD	197	 3% 78% 16% • 5%
4	III	197	 5% 80% 14% • 5%
5	EEE	243	 % 86% 13% ••
5	JJJ	243	 2% 90% 9% •

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13151 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, alpha chain E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	250	Total 2038	C 1278	N 363	O 390	S 7	0	1	0
1	FFF	267	Total 2185	C 1371	N 390	O 417	S 7	0	1	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	100	Total 837	C 533	N 141	O 159	S 4	0	0	0
2	GGG	100	Total 837	C 533	N 141	O 159	S 4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	0	MET	-	initiating methionine	UNP P61769
GGG	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Peptide from inhA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	CCC	9	Total 69	C 46	N 13	O 10	0	0	0
3	HHH	9	Total 69	C 46	N 13	O 10	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	DDD	187	Total 1484	C 917	N 254	O 304	S 9	0	2	0

Continued on next page...

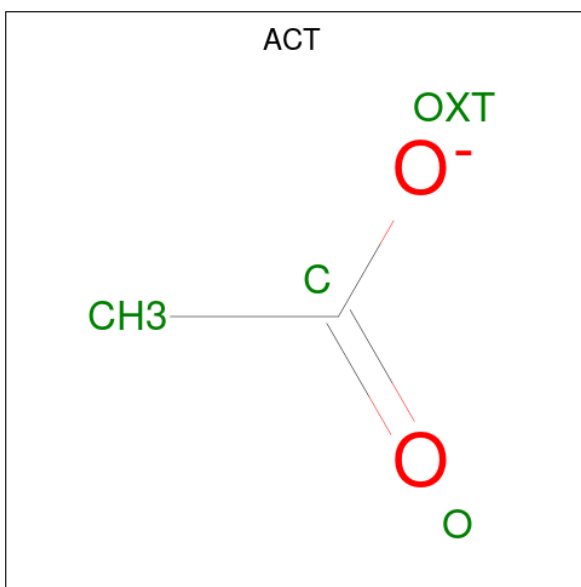
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	III	187	Total	C	N	O	S	0	0	0
			1471	909	251	302	9			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	EEE	241	Total	C	N	O	S	0	0	0
			1936	1213	345	372	6			
5	JJJ	241	Total	C	N	O	S	0	0	0
			1936	1213	345	372	6			

- Molecule 6 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	FFF	1	Total	C	O	0	0
			4	2	2		
6	JJJ	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	43	Total	O	0	0
			43	43		
7	BBB	10	Total	O	0	0
			10	10		

Continued on next page...

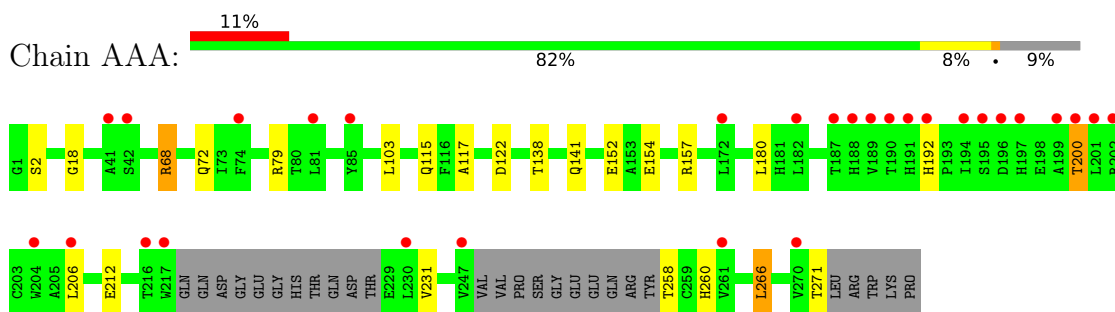
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	CCC	2	Total O 2 2	0	0
7	DDD	38	Total O 38 38	0	0
7	EEE	51	Total O 51 51	0	0
7	FFF	63	Total O 63 63	0	0
7	GGG	11	Total O 11 11	0	0
7	HHH	6	Total O 6 6	0	0
7	III	33	Total O 33 33	0	0
7	JJJ	24	Total O 24 24	0	0

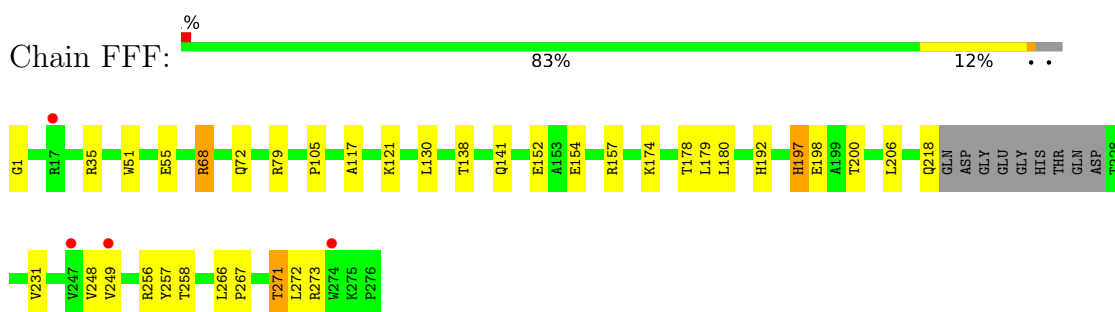
3 Residue-property plots [i](#)

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.

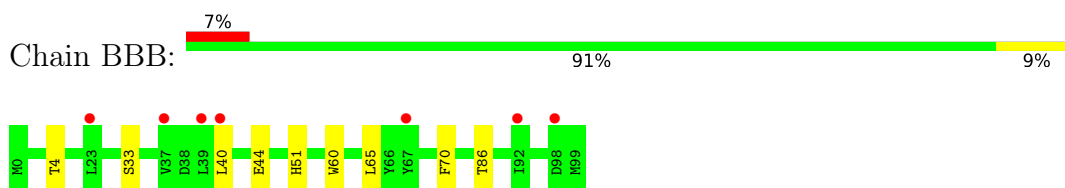
- Molecule 1: HLA class I histocompatibility antigen, alpha chain E



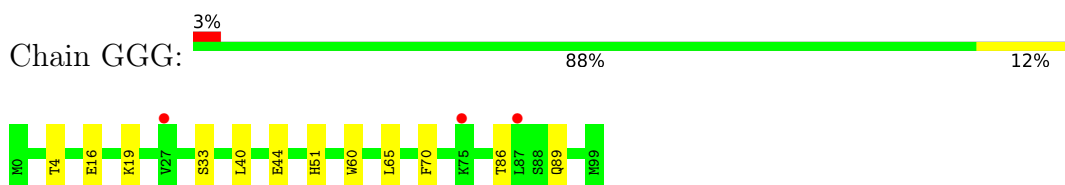
- Molecule 1: HLA class I histocompatibility antigen, alpha chain E




- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin




- Molecule 3: Peptide from inhA

Chain CCC:  78% 22%




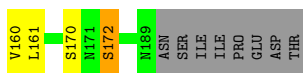
- Molecule 3: Peptide from inhA

Chain HHH:  78% 22%




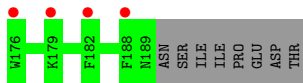
- Molecule 4: T-cell receptor alpha chain

Chain DDD:  78% 16% 3% 5%




- Molecule 4: T-cell receptor alpha chain

Chain III:  80% 14% 5% 5%



- Molecule 5: T-cell receptor beta chain

Chain EEE:  86% 13% 1% 2%



- Molecule 5: T-cell receptor beta chain

Chain JJJ:  90% 9% 2%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	326.83Å 79.88Å 104.07Å 90.00° 96.89° 90.00°	Depositor
Resolution (Å)	82.90 – 2.33 82.76 – 2.33	Depositor EDS
% Data completeness (in resolution range)	99.9 (82.90-2.33) 98.9 (82.76-2.33)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.00 (at 2.32Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.242 , 0.289 0.250 , 0.294	Depositor DCC
R_{free} test set	5632 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	59.8	Xtrriage
Anisotropy	0.483	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 46.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	13151	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ACT

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	AAA	0.66	0/2097	0.78	0/2849
1	FFF	0.67	0/2250	0.82	0/3058
2	BBB	0.64	0/860	0.73	0/1162
2	GGG	0.65	0/860	0.76	0/1162
3	CCC	0.64	0/70	0.82	0/93
3	HHH	0.80	0/70	0.99	0/93
4	DDD	0.74	0/1518	0.85	0/2050
4	III	0.73	0/1499	0.83	0/2025
5	EEE	0.65	0/1987	0.82	0/2699
5	JJJ	0.63	0/1987	0.79	0/2699
All	All	0.67	0/13198	0.81	0/17890

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 [i](#)

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	AAA	2038	0	1896	17	0
1	FFF	2185	0	2041	30	0
2	BBB	837	0	803	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	GGG	837	0	803	6	0
3	CCC	69	0	85	2	0
3	HHH	69	0	85	4	0
4	DDD	1484	0	1410	19	0
4	III	1471	0	1393	18	0
5	EEE	1936	0	1843	17	0
5	JJJ	1936	0	1843	16	0
6	FFF	4	0	3	0	0
6	JJJ	4	0	3	0	0
7	AAA	43	0	0	2	0
7	BBB	10	0	0	0	0
7	CCC	2	0	0	0	0
7	DDD	38	0	0	0	0
7	EEE	51	0	0	1	0
7	FFF	63	0	0	9	0
7	GGG	11	0	0	0	0
7	HHH	6	0	0	0	0
7	III	33	0	0	1	0
7	JJJ	24	0	0	0	0
All	All	13151	0	12208	107	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 107 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:FFF:68:ARG:HH11	1:FFF:68:ARG:HG2	1.54	0.72
1:AAA:72:GLN:HG2	5:EEE:31:ASN:HB3	1.73	0.69
1:FFF:272:LEU:HD12	7:FFF:404:HOH:O	1.96	0.66
4:DDD:33:SER:HB3	4:DDD:52:TYR:CD1	2.31	0.65
4:III:161:LEU:HB3	5:JJJ:170:CYS:HB2	1.79	0.65

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	AAA	245/276 (89%)	236 (96%)	9 (4%)	0	100	100
1	FFF	264/276 (96%)	252 (96%)	12 (4%)	0	100	100
2	BBB	98/100 (98%)	91 (93%)	7 (7%)	0	100	100
2	GGG	98/100 (98%)	92 (94%)	6 (6%)	0	100	100
3	CCC	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	HHH	7/9 (78%)	7 (100%)	0	0	100	100
4	DDD	187/197 (95%)	175 (94%)	12 (6%)	0	100	100
4	III	185/197 (94%)	170 (92%)	15 (8%)	0	100	100
5	EEE	239/243 (98%)	225 (94%)	14 (6%)	0	100	100
5	JJJ	239/243 (98%)	228 (95%)	11 (5%)	0	100	100
All	All	1569/1650 (95%)	1482 (94%)	87 (6%)	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	AAA	215/237 (91%)	207 (96%)	8 (4%)	34	43
1	FFF	231/237 (98%)	222 (96%)	9 (4%)	32	41
2	BBB	95/95 (100%)	93 (98%)	2 (2%)	53	65
2	GGG	95/95 (100%)	93 (98%)	2 (2%)	53	65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	CCC	7/7 (100%)	7 (100%)	0	100	100
3	HHH	7/7 (100%)	7 (100%)	0	100	100
4	DDD	170/177 (96%)	159 (94%)	11 (6%)	17	19
4	III	168/177 (95%)	158 (94%)	10 (6%)	19	22
5	EEE	212/214 (99%)	197 (93%)	15 (7%)	14	15
5	JJJ	212/214 (99%)	205 (97%)	7 (3%)	38	46
All	All	1412/1460 (97%)	1348 (96%)	64 (4%)	27	34

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	JJJ	28	SER
5	JJJ	52	SER
5	EEE	37	ARG
5	EEE	29	ASP
5	JJJ	90	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
6	ACT	FFF	301	-	3,3,3	1.16	0	3,3,3	0.87	0
6	ACT	JJJ	301	-	3,3,3	1.07	0	3,3,3	0.70	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	AAA	250/276 (90%)	0.73	29 (11%) 4 8	49, 83, 126, 166	0
1	FFF	267/276 (96%)	0.36	4 (1%) 73 81	34, 71, 153, 198	0
2	BBB	100/100 (100%)	0.44	7 (7%) 16 24	68, 106, 138, 157	0
2	GGG	100/100 (100%)	0.47	3 (3%) 50 60	52, 75, 109, 129	0
3	CCC	9/9 (100%)	0.64	0 100 100	46, 60, 76, 77	0
3	HHH	9/9 (100%)	0.43	0 100 100	43, 47, 57, 61	0
4	DDD	187/197 (94%)	0.47	6 (3%) 47 58	41, 69, 132, 156	1 (0%)
4	III	187/197 (94%)	0.47	9 (4%) 30 41	39, 66, 141, 162	1 (0%)
5	EEE	241/243 (99%)	0.48	2 (0%) 86 90	39, 61, 95, 135	0
5	JJJ	241/243 (99%)	0.45	6 (2%) 57 66	35, 83, 123, 150	0
All	All	1591/1650 (96%)	0.49	66 (4%) 37 48	34, 75, 133, 198	2 (0%)

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	JJJ	60	ARG	7.7
1	AAA	194	ILE	6.2
1	AAA	247	VAL	6.1
4	III	179	LYS	5.4
4	DDD	149	LYS	5.3

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	ACT	FFF	301	4/4	0.86	0.13	70,73,73,84	0
6	ACT	JJJ	301	4/4	0.87	0.20	94,95,99,101	0

6.5 Other polymers [i](#)

There are no such residues in this entry.