WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 25, 2024 – 12:04 PM EST

PDB ID	:	9DH2
Title	:	Structure of Fab in complex with NKG2D extracellular domain
Authors	:	Fallon, D.; Huang, C.S.
Deposited on	:	2024-09-03
Resolution	:	2.98 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:

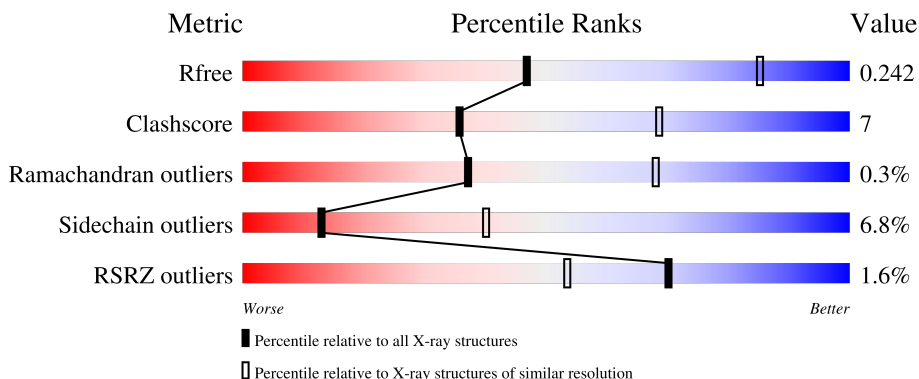|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.98 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		0.242
Clashscore		7
Ramachandran outliers		0.3%
Sidechain outliers		6.8%
RSRZ outliers		1.6%

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$_{free}$	164625	3360 (3.00-2.96)
Clashscore	180529	3751 (3.00-2.96)
Ramachandran outliers	177936	3628 (3.00-2.96)
Sidechain outliers	177891	3631 (3.00-2.96)
RSRZ outliers	164620	3372 (3.00-2.96)

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	220	% 81% 16% ··
1	G	220	2% 80% 16% ··
1	K	220	4% 80% 19% ·
1	P	220	3% 82% 15% ··
2	D	212	79% 18% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	212	83% 15%
2	L	212	83% 15%
2	Q	212	82% 17%
3	M	137	2% 66% 22% 10%
3	R	137	67% 22% 9%
3	S	137	6% 63% 25% 10%
3	T	137	% 70% 20% 9%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 17120 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 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1646	1043	270	326	7			
1	G	217	Total	C	N	O	S	0	0	0
			1646	1043	270	326	7			
1	K	217	Total	C	N	O	S	0	0	0
			1646	1043	270	326	7			
1	P	217	Total	C	N	O	S	0	0	0
			1646	1043	270	326	7			

- Molecule 2 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	211	Total	C	N	O	S	0	0	0
			1634	1029	271	329	5			
2	H	211	Total	C	N	O	S	0	0	0
			1634	1029	271	329	5			
2	L	211	Total	C	N	O	S	0	0	0
			1634	1029	271	329	5			
2	Q	211	Total	C	N	O	S	0	0	0
			1634	1029	271	329	5			

- Molecule 3 is a protein called NKG2-D type II integral membrane protein.

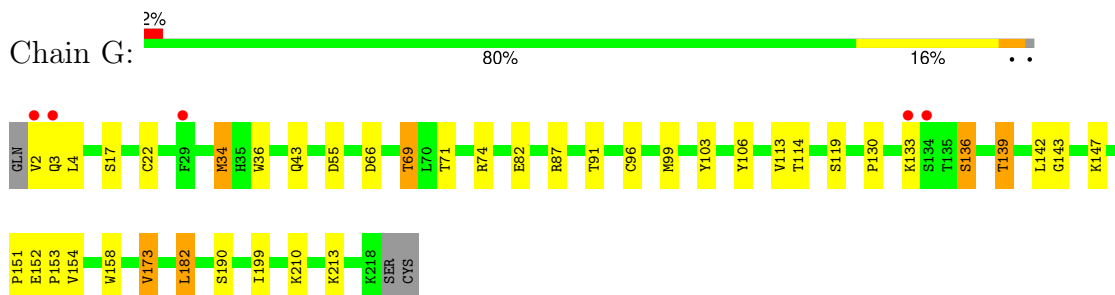| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	M	123	Total	C	N	O	S	0	0	0
			996	633	160	191	12			
3	R	124	Total	C	N	O	S	0	0	0
			1004	638	161	193	12			
3	S	123	Total	C	N	O	S	0	0	0
			996	633	160	191	12			
3	T	124	Total	C	N	O	S	0	0	0
			1004	638	161	193	12			

3 Residue-property plots ⓘ

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.
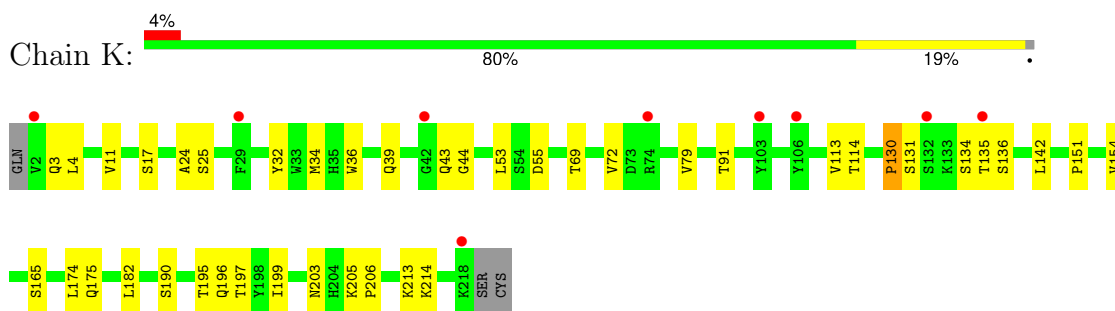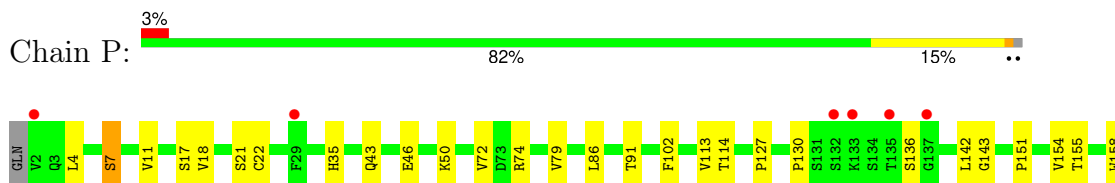
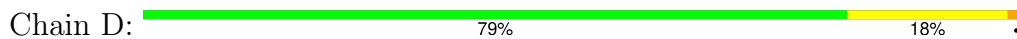- Molecule 1: Fab heavy chain

Chain A:

- Molecule 1: Fab heavy chain

Chain G:

- Molecule 1: Fab heavy chain

Chain K:

- Molecule 1: Fab heavy chain

Chain P:

L174 Y180 S181 L182 T187 S191 I199 K213 K218 SER CYS

• Molecule 2: Fab light chain

Chain D:

79% 18%

D1 I2 S7 P8 S9 I29 L33 Q38 I48 Y49 N50 A51 K52 T53 L54 S63 G66 S67 D70 Y71 W92 S93 E105 R108 A111 A112 P113 S114 V115 T129 L136 N137 N138 F139 R142 V150 Q155 S159 Q166 Y173

T178 E195 S202 K205 S206 G210 E211 CYS

• Molecule 2: Fab light chain

Chain H:

83% 15%

D1 S7 P8 S9 S10 T22 L33 K39 L46 D56 S63 G66 S67 D70 Y71 W92 S93 I94 P95 E105 R108 T109 Y140 P141 V150 S159 Q160 Q166 Y173 T178 E195 V196 T197 Q198 T201 S202 V203 T204 E211

CYS

• Molecule 2: Fab light chain

Chain L:

83% 15%

D1 S7 P8 T20 I21 T22 C23 R24 L33 Q38 I48 L54 G66 Y71 T74 H90 F91 W92 S93 R108 V115 F118 P119 S127 Y140 P141 R142 E143 Q155 N158 S159 Q166 Y173 T178 L179 D185 K188

S202 K205 G210 E211 CYS

• Molecule 2: Fab light chain

Chain Q:

82% 17%

D1 S9 S10 L11 S12 T22 Y30 N31 Y32 L33 Y49 N50 L54 A55 D56 A57 V58 P59 S60 R61 S65 G66 Y71 T72 S93 I94 P95 E105 R108 T109 V115 S121 S127 G128 T129 L136 R142 V150 S159 Q166 Y173

S174 L175 T178 S202 K205 E211 CYS

• Molecule 3: NKG2-D type II integral membrane protein

Chain M:

2%
66% 22% 10%

ASN SER LEU PHE ASN GLN GLU VAL GLN ILE PRO LEU T92 E93 S94 Y95 C96 Y106 K107 N108 S117 Q124 A125 S126 S133 V137 Y138 S139 L145 L146 K147 L148 V149 Y152 M155 I160 P161 T162 E169 L173 L174 L178 I181 M184 A193

S194 S195 F196 K197 G198 Y199 N207 R214 THR VAL

• Molecule 3: NKG2-D type II integral membrane protein

Chain R:

ASN SER LEU PHE ASN GLN GLU VAL GLN ILE PRO LEU THR E93 P98 Y106 Y111 Q112 F113 F114 D115 E116 S117 L135 K136 V137 Y138 S139 K140 Q143 D144 L145 V149 N163 G164 S165 I173 N177 L178 I181 K186 L191 S194 S195 E201 N202

C203 S204 M212 T215 V216

• Molecule 3: NKG2-D type II integral membrane protein

Chain S:

ASN SER LEU PHE ASN GLN GLU VAL GLN ILE PRO LEU T92 P98 C99 P100 K101 N102 Y106 D115 E116 S117 Q124 L135 K136 V137 K140 Q143 D144 L145 L146 K147 L148 V149 K150 S151 M155 I160 P161 T162 N163 G164 S165 D170 S175 L178 I181

I182 Q185 K186 Y192 A193 S194 S195 F196 K197 G198 Y199 I200 E201 Q213 R214 THR VAL

• Molecule 3: NKG2-D type II integral membrane protein

Chain T:

ASN SER LEU PHE ASN GLN GLU VAL GLN ILE PRO LEU THR E93 S94 Y95 P98 Y106 Q112 F113 V137 Y138 S139 K140 Q143 D144 L145 Y152 M155 I160 S165 W166 Q167 S172 I173 N177 L178 I181 A193 S194 S195 I200 T215 V216

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.06Å 224.43Å 136.30Å 90.00° 91.22° 90.00°	Depositor
Resolution (Å)	136.64 – 2.98 136.27 – 2.98	Depositor EDS
% Data completeness (in resolution range)	97.3 (136.64-2.98) 97.3 (136.27-2.98)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.35 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.197 , 0.243 0.198 , 0.242	Depositor DCC
R_{free} test set	2818 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	45.9	Xtriage
Anisotropy	0.456	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.5	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.033 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17120	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.36% 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	0.52	0/1688	0.90	1/2298 (0.0%)
1	G	0.47	0/1688	0.94	4/2298 (0.2%)
1	K	0.48	1/1688 (0.1%)	0.88	0/2298
1	P	0.48	0/1688	0.87	0/2298
2	D	0.58	2/1672 (0.1%)	0.89	0/2274
2	H	0.52	0/1672	0.92	1/2274 (0.0%)
2	L	0.48	0/1672	0.90	2/2274 (0.1%)
2	Q	0.48	0/1672	0.85	2/2274 (0.1%)
3	M	0.46	0/1024	0.85	0/1388
3	R	0.46	0/1032	0.85	1/1398 (0.1%)
3	S	0.45	0/1024	0.87	0/1388
3	T	0.46	0/1032	0.82	0/1398
All	All	0.49	3/17552 (0.0%)	0.88	11/23860 (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	2
1	G	0	2
1	K	0	1
1	P	0	1
3	M	0	2
3	T	0	1
All	All	0	9

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	142	ARG	CD-NE	-10.42	1.28	1.46
2	D	142	ARG	CG-CD	7.10	1.69	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	165	SER	CA-CB	-5.36	1.45	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	158	ASN	CB-CA-C	-10.20	90.01	110.40
2	L	108	ARG	CG-CD-NE	9.14	131.00	111.80
1	G	147	LYS	C-N-CA	-6.86	104.54	121.70
1	G	66	ASP	CB-CA-C	-6.81	96.78	110.40
1	A	62	GLN	CB-CG-CD	6.69	128.98	111.60

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	24	ALA	Peptide
1	A	83	LEU	Peptide
1	G	151	PRO	Peptide
1	G	43	GLN	Peptide
1	K	130	PRO	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	1646	0	1601	20	0
1	G	1646	0	1601	25	0
1	K	1646	0	1601	23	0
1	P	1646	0	1601	16	0
2	D	1634	0	1584	22	0
2	H	1634	0	1584	19	0
2	L	1634	0	1584	22	0
2	Q	1634	0	1584	16	0
3	M	996	0	933	19	0
3	R	1004	0	942	29	0
3	S	996	0	933	36	0
3	T	1004	0	942	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	17120	0	16490	243	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 7.

The worst 5 of 243 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)
3:S:155:MET:HE2	3:S:192:TYR:HB2	1.48	0.95
1:G:152:GLU:O	1:G:152:GLU:HG2	1.67	0.94
1:A:39:GLN:HE22	2:D:38:GLN:HE22	1.12	0.92
1:A:6:GLN:HE21	1:A:108:GLY:HA3	1.35	0.91
3:R:135:LEU:HD12	3:R:212:MET:HE3	1.54	0.90

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	215/220 (98%)	205 (95%)	10 (5%)	0	100	100
1	G	215/220 (98%)	205 (95%)	10 (5%)	0	100	100
1	K	215/220 (98%)	208 (97%)	7 (3%)	0	100	100
1	P	215/220 (98%)	206 (96%)	9 (4%)	0	100	100
2	D	209/212 (99%)	199 (95%)	8 (4%)	2 (1%)	13	43
2	H	209/212 (99%)	197 (94%)	12 (6%)	0	100	100
2	L	209/212 (99%)	201 (96%)	8 (4%)	0	100	100
2	Q	209/212 (99%)	200 (96%)	8 (4%)	1 (0%)	25	59
3	M	121/137 (88%)	113 (93%)	7 (6%)	1 (1%)	16	48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	R	122/137 (89%)	115 (94%)	6 (5%)	1 (1%)	16	48
3	S	121/137 (88%)	111 (92%)	9 (7%)	1 (1%)	16	48
3	T	122/137 (89%)	114 (93%)	8 (7%)	0	100	100
All	All	2182/2276 (96%)	2074 (95%)	102 (5%)	6 (0%)	37	68

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	R	195	SER
3	S	195	SER
2	Q	50	ASN
2	D	50	ASN
2	D	138	ASN

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	186/189 (98%)	172 (92%)	14 (8%)	11	36
1	G	186/189 (98%)	172 (92%)	14 (8%)	11	36
1	K	186/189 (98%)	173 (93%)	13 (7%)	12	39
1	P	186/189 (98%)	174 (94%)	12 (6%)	14	41
2	D	184/185 (100%)	170 (92%)	14 (8%)	11	35
2	H	184/185 (100%)	171 (93%)	13 (7%)	12	38
2	L	184/185 (100%)	175 (95%)	9 (5%)	21	52
2	Q	184/185 (100%)	170 (92%)	14 (8%)	11	35
3	M	113/127 (89%)	105 (93%)	8 (7%)	12	38
3	R	114/127 (90%)	108 (95%)	6 (5%)	19	49
3	S	113/127 (89%)	104 (92%)	9 (8%)	10	33
3	T	114/127 (90%)	109 (96%)	5 (4%)	24	56
All	All	1934/2004 (96%)	1803 (93%)	131 (7%)	13	40

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	R	186	LYS
3	S	117	SER
3	T	173	ILE
2	H	63	SER
2	H	56	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
2	Q	31	ASN
3	T	153	HIS
3	R	109	ASN
3	T	185	GLN
3	S	102	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 (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, 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	217/220 (98%)	-0.58	2 (0%)	81	67	29, 41, 61, 99	0
1	G	217/220 (98%)	-0.45	5 (2%)	61	44	30, 44, 73, 94	0
1	K	217/220 (98%)	-0.30	9 (4%)	42	28	28, 46, 80, 144	0
1	P	217/220 (98%)	-0.37	6 (2%)	55	38	29, 46, 81, 140	0
2	D	211/212 (99%)	-0.51	0	100	100	31, 47, 65, 80	0
2	H	211/212 (99%)	-0.67	0	100	100	26, 38, 53, 69	0
2	L	211/212 (99%)	-0.56	0	100	100	30, 46, 65, 79	0
2	Q	211/212 (99%)	-0.43	0	100	100	29, 51, 75, 90	0
3	M	123/137 (89%)	0.12	3 (2%)	59	42	46, 65, 95, 130	0
3	R	124/137 (90%)	-0.41	0	100	100	36, 51, 79, 104	0
3	S	123/137 (89%)	0.26	8 (6%)	26	17	47, 69, 96, 111	0
3	T	124/137 (90%)	-0.22	2 (1%)	70	53	39, 55, 87, 104	0
All	All	2206/2276 (96%)	-0.39	35 (1%)	70	53	26, 47, 80, 144	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	2	VAL	5.6
1	P	135	THR	5.3
3	M	92	THR	4.1
1	G	2	VAL	3.9
3	S	92	THR	3.4

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.