



wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 10:00 pm BST

PDB ID : 3BXB
Title : Monomeric Far-red Fluorescent Protein mKate Crystallized at pH 7.0
Authors : Pletnev, S.; Pletneva, N.; Pletnev, V.
Deposited on : 2008-01-12
Resolution : 2.60 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

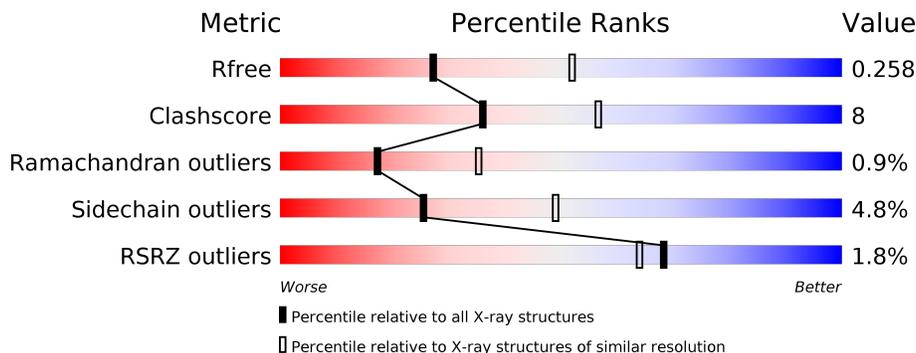
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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	A	243	
1	B	243	
1	C	243	
1	D	243	
1	E	243	
1	F	243	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	243	
1	H	243	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14615 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 Far-red fluorescent protein mKate.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	1776	1129	300	334	13	0	0	0
1	B	223	1776	1129	300	334	13	0	0	0
1	C	223	1787	1135	304	335	13	0	1	0
1	D	223	1799	1144	305	337	13	0	2	0
1	E	223	1787	1135	304	335	13	0	1	0
1	F	223	1787	1135	304	335	13	0	1	0
1	G	223	1776	1129	300	334	13	0	0	0
1	H	223	1776	1129	300	334	13	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	48	Total	O	0	0
			48	48		
2	B	51	Total	O	0	0
			51	51		
2	C	42	Total	O	0	0
			42	42		
2	D	60	Total	O	0	0
			60	60		
2	E	51	Total	O	0	0
			51	51		
2	F	45	Total	O	0	0
			45	45		

Continued on next page...

Continued from previous page...

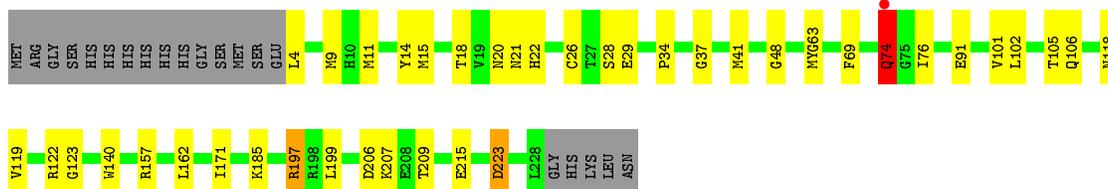
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	25	Total	O	0	0
			25	25		
2	H	29	Total	O	0	0
			29	29		

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 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.

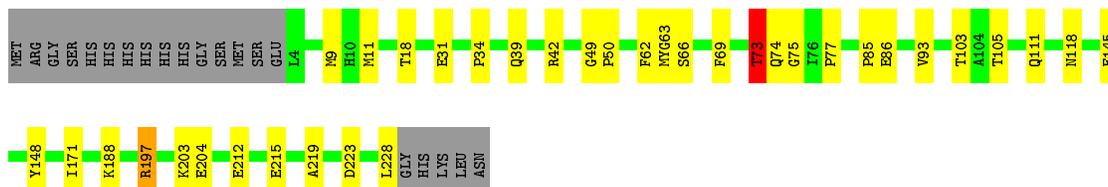
- Molecule 1: Far-red fluorescent protein mKate

Chain A: 



- Molecule 1: Far-red fluorescent protein mKate

Chain B: 



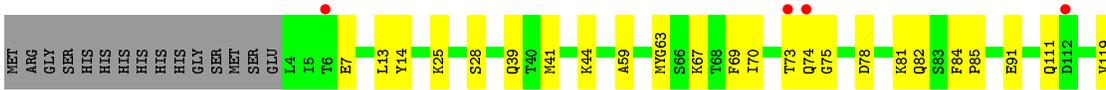
- Molecule 1: Far-red fluorescent protein mKate

Chain C: 

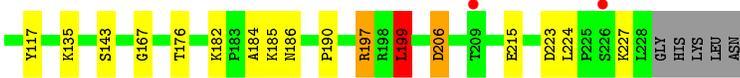
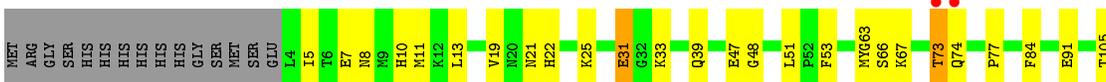


- Molecule 1: Far-red fluorescent protein mKate

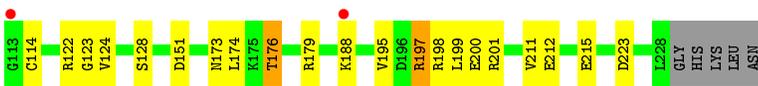
Chain D: 



• Molecule 1: Far-red fluorescent protein mKate



• Molecule 1: Far-red fluorescent protein mKate



• Molecule 1: Far-red fluorescent protein mKate



• Molecule 1: Far-red fluorescent protein mKate



E141	E145	Y148	P149	L162	R163	L164	V165	L170	R179	K185	V192	R197	R198	Y210	V216	C222	D223	L224	P225	S226	K227	L228	GLY	HIS	LYS	LEU	ASN
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	74.18Å 104.97Å 123.03Å 90.00° 105.81° 90.00°	Depositor
Resolution (Å)	29.84 – 2.60 29.84 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (29.84-2.60) 99.7 (29.84-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.92 (at 2.61Å)	Xtrriage
Refinement program	PHENIX, REFMAC	Depositor
R, R_{free}	0.170 , 0.263 0.165 , 0.258	Depositor DCC
R_{free} test set	2814 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	41.8	Xtrriage
Anisotropy	0.200	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 47.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.020 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14615	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.89% 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: NRQ

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	0.60	0/1792	0.67	0/2417
1	B	0.59	0/1792	0.70	1/2417 (0.0%)
1	C	0.62	0/1803	0.66	0/2431
1	D	0.57	0/1816	0.67	0/2449
1	E	0.56	0/1803	0.71	1/2431 (0.0%)
1	F	0.57	0/1803	0.66	0/2431
1	G	0.58	0/1792	0.66	0/2417
1	H	0.52	0/1792	0.64	0/2417
All	All	0.58	0/14393	0.67	2/19410 (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	C	0	1
1	F	0	1
1	G	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	199	LEU	CA-CB-CG	5.83	128.71	115.30
1	B	62	PHE	CA-C-O	-5.31	108.95	120.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	84	PHE	Peptide
1	F	74	GLN	Peptide
1	G	84	PHE	Peptide

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	A	1776	0	1745	31	0
1	B	1776	0	1745	19	0
1	C	1787	0	1757	38	0
1	D	1799	0	1765	24	0
1	E	1787	0	1757	28	0
1	F	1787	0	1757	39	0
1	G	1776	0	1745	41	0
1	H	1776	0	1745	27	0
2	A	48	0	0	1	0
2	B	51	0	0	0	0
2	C	42	0	0	7	0
2	D	60	0	0	1	0
2	E	51	0	0	1	0
2	F	45	0	0	3	0
2	G	25	0	0	5	0
2	H	29	0	0	1	0
All	All	14615	0	14016	227	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 8.

The worst 5 of 227 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:A:41:MET:HG3	1:A:63:NRQ:HE3	1.18	1.11
1:A:41:MET:HG3	1:A:63:NRQ:CE	1.85	1.06
1:C:73:THR:HA	2:C:266:HOH:O	1.55	1.05
1:E:67:LYS:NZ	1:E:197[B]:ARG:HH22	1.59	1.00
1:H:78:ASP:O	1:H:82:GLN:HB2	1.62	0.98

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	218/243 (90%)	211 (97%)	6 (3%)	1 (0%)	29	52
1	B	218/243 (90%)	209 (96%)	7 (3%)	2 (1%)	17	35
1	C	219/243 (90%)	209 (95%)	9 (4%)	1 (0%)	29	52
1	D	220/243 (90%)	211 (96%)	8 (4%)	1 (0%)	29	52
1	E	219/243 (90%)	204 (93%)	10 (5%)	5 (2%)	6	11
1	F	219/243 (90%)	211 (96%)	8 (4%)	0	100	100
1	G	218/243 (90%)	200 (92%)	14 (6%)	4 (2%)	8	16
1	H	218/243 (90%)	205 (94%)	11 (5%)	2 (1%)	17	35
All	All	1749/1944 (90%)	1660 (95%)	73 (4%)	16 (1%)	17	35

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	73	THR
1	E	73	THR
1	E	74	GLN
1	G	206	ASP
1	E	185	LYS

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	191/208 (92%)	183 (96%)	8 (4%)	30	55
1	B	191/208 (92%)	181 (95%)	10 (5%)	23	46
1	C	192/208 (92%)	182 (95%)	10 (5%)	23	46
1	D	193/208 (93%)	187 (97%)	6 (3%)	40	66
1	E	192/208 (92%)	180 (94%)	12 (6%)	18	36
1	F	192/208 (92%)	180 (94%)	12 (6%)	18	36
1	G	191/208 (92%)	179 (94%)	12 (6%)	18	36
1	H	191/208 (92%)	185 (97%)	6 (3%)	40	66
All	All	1533/1664 (92%)	1457 (95%)	76 (5%)	25	47

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	7	GLU
1	E	199	LEU
1	H	11	MET
1	E	11	MET
1	E	176	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	22	HIS
1	D	213	GLN
1	G	74	GLN
1	D	21	ASN
1	G	22	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

8 non-standard protein/DNA/RNA residues 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
1	NRQ	B	63	1	23,24,25	0.92	0	23,32,34	1.05	1 (4%)
1	NRQ	D	63	1	23,24,25	0.89	1 (4%)	23,32,34	0.92	1 (4%)
1	NRQ	F	63	1	23,24,25	0.90	1 (4%)	23,32,34	0.91	1 (4%)
1	NRQ	H	63	1	23,24,25	0.76	0	23,32,34	0.88	1 (4%)
1	NRQ	E	63	1	23,24,25	0.88	0	23,32,34	1.34	2 (8%)
1	NRQ	C	63	1	23,24,25	0.81	0	23,32,34	1.03	2 (8%)
1	NRQ	G	63	1	23,24,25	0.81	1 (4%)	23,32,34	0.97	1 (4%)
1	NRQ	A	63	1	23,24,25	0.95	1 (4%)	23,32,34	1.01	1 (4%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	NRQ	B	63	1	-	2/9/31/32	0/2/2/2
1	NRQ	D	63	1	-	1/9/31/32	0/2/2/2
1	NRQ	F	63	1	-	1/9/31/32	0/2/2/2
1	NRQ	H	63	1	-	1/9/31/32	0/2/2/2
1	NRQ	E	63	1	-	1/9/31/32	0/2/2/2
1	NRQ	C	63	1	-	1/9/31/32	0/2/2/2
1	NRQ	G	63	1	-	2/9/31/32	0/2/2/2
1	NRQ	A	63	1	-	2/9/31/32	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	63	NRQ	CA1-N1	3.65	1.36	1.27
1	F	63	NRQ	CA1-N1	3.24	1.35	1.27
1	D	63	NRQ	OH-CZ	-2.22	1.31	1.37
1	G	63	NRQ	OH-CZ	-2.03	1.32	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	NRQ	O3-C3-CA3	-2.82	117.87	126.39
1	E	63	NRQ	CD1-CG2-CD2	2.82	121.81	117.64
1	F	63	NRQ	O3-C3-CA3	-2.71	118.22	126.39
1	E	63	NRQ	O3-C3-CA3	-2.71	118.22	126.39
1	H	63	NRQ	O3-C3-CA3	-2.65	118.39	126.39

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	F	63	NRQ	C1-CA1-CB1-CG1
1	D	63	NRQ	CB1-CG1-SD-CE
1	C	63	NRQ	CB1-CG1-SD-CE
1	A	63	NRQ	CB1-CG1-SD-CE
1	G	63	NRQ	CA1-CB1-CG1-SD

There are no ring outliers.

8 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	63	NRQ	1	0
1	D	63	NRQ	3	0
1	F	63	NRQ	3	0
1	H	63	NRQ	1	0
1	E	63	NRQ	4	0
1	C	63	NRQ	7	0
1	G	63	NRQ	2	0
1	A	63	NRQ	7	0

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 [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	A	222/243 (91%)	-0.53	1 (0%) 91 89	23, 39, 62, 85	0
1	B	222/243 (91%)	-0.43	0 100 100	23, 36, 64, 88	0
1	C	222/243 (91%)	-0.39	1 (0%) 91 89	25, 42, 66, 82	0
1	D	222/243 (91%)	-0.30	6 (2%) 54 48	23, 39, 68, 94	0
1	E	222/243 (91%)	-0.38	4 (1%) 68 64	22, 39, 68, 88	0
1	F	222/243 (91%)	-0.39	3 (1%) 75 71	24, 40, 68, 91	0
1	G	222/243 (91%)	-0.04	14 (6%) 20 15	26, 49, 78, 90	0
1	H	222/243 (91%)	-0.26	3 (1%) 75 71	28, 52, 73, 97	0
All	All	1776/1944 (91%)	-0.34	32 (1%) 68 64	22, 42, 71, 97	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	73	THR	5.8
1	G	113	GLY	3.6
1	E	73	THR	3.6
1	G	112	ASP	3.6
1	D	112	ASP	3.4

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	NRQ	G	63	23/24	0.93	0.30	32,51,86,107	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
1	NRQ	H	63	23/24	0.94	0.27	30,62,100,111	0
1	NRQ	C	63	23/24	0.95	0.25	30,46,57,97	0
1	NRQ	B	63	23/24	0.95	0.31	17,40,61,73	0
1	NRQ	E	63	23/24	0.96	0.22	23,42,59,73	0
1	NRQ	F	63	23/24	0.96	0.23	20,33,69,71	0
1	NRQ	D	63	23/24	0.96	0.23	28,36,61,89	0
1	NRQ	A	63	23/24	0.98	0.23	25,36,64,79	0

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.