Full wwPDB X-ray Structure Validation Report ⓘ

Jun 4, 2025 – 01:05 pm BST

PDB ID	:	9FGI / pdb_00009fgi
Title	:	LSSmOrange (P1) - Directionality of Optical Properties of Fluorescent Proteins
Authors	:	Myskova, J.; Brynda, J.; Lazar, J.
Deposited on	:	2024-05-24
Resolution	:	1.77 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:

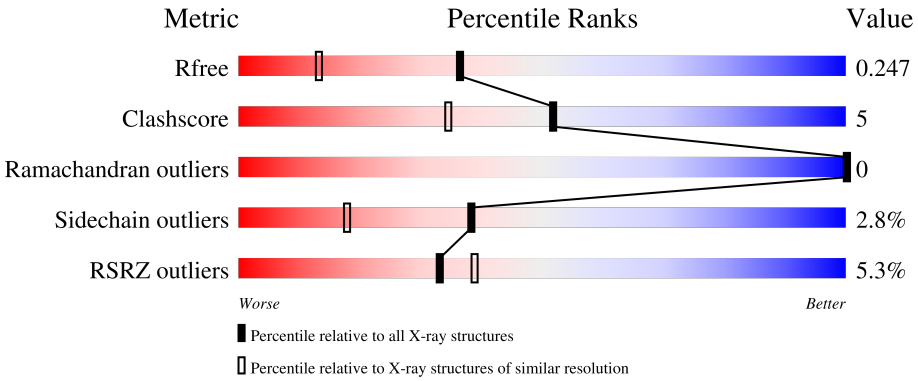
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 1.77 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	1191 (1.78-1.78)
Clashscore	180529	1282 (1.78-1.78)
Ramachandran outliers	177936	1270 (1.78-1.78)
Sidechain outliers	177891	1270 (1.78-1.78)
RSRZ outliers	164620	1191 (1.78-1.78)

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	243	4% 81% 12% 6%
1	B	243	6% 81% 13% • 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3807 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 nowGFP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	228	Total	C	N	O	S	0	1	0
			1812	1152	306	347	7			
1	B	230	Total	C	N	O	S	0	0	0
			1824	1160	308	348	8			

- Molecule 2 is water.

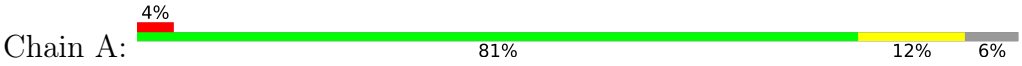
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	99	Total	O	0	0
			99	99		
2	B	72	Total	O	0	0
			72	72		

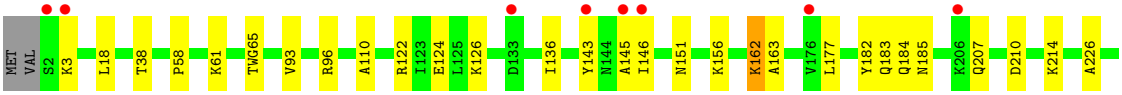
3 Residue-property plots (i)

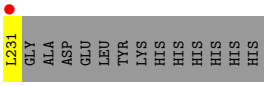
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: nowGFP

Chain A:

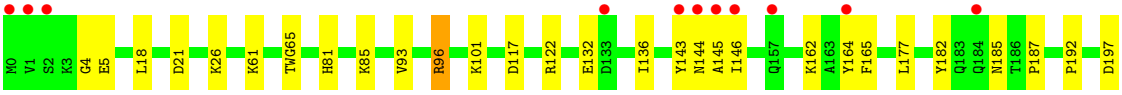


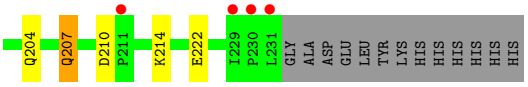




• Molecule 1: nowGFP

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.26Å 51.39Å 195.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.59 – 1.77 49.59 – 1.77	Depositor EDS
% Data completeness (in resolution range)	99.1 (49.59-1.77) 99.2 (49.59-1.77)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 1.77Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.199 , 0.240 0.211 , 0.247	Depositor DCC
R_{free} test set	2098 reflections (4.11%)	wwPDB-VP
Wilson B-factor (Å^2)	32.9	Xtriage
Anisotropy	0.102	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 35.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.021 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3807	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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

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

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	1.20	3/1830 (0.2%)	1.33	2/2469 (0.1%)
1	B	1.21	1/1839 (0.1%)	1.37	4/2481 (0.2%)
All	All	1.20	4/3669 (0.1%)	1.35	6/4950 (0.1%)

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	1
1	B	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	187	PRO	C-O	-6.39	1.16	1.23
1	A	226	ALA	C-O	6.04	1.31	1.23
1	A	38	THR	C-O	5.29	1.30	1.24
1	A	146	ILE	C-O	5.23	1.29	1.24

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	231	LEU	CA-C-O	-12.57	99.43	120.80
1	B	207	GLN	CB-CG-CD	-6.05	102.31	112.60
1	B	132	GLU	CB-CG-CD	5.76	122.39	112.60
1	A	210	ASP	CA-CB-CG	5.44	118.04	112.60
1	B	192	PRO	CA-C-O	-5.33	115.36	121.86

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	210	ASP	CB-CA-C	5.16	118.34	109.82

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	96	ARG	Peptide
1	B	96	ARG	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	1812	0	1772	18	0
1	B	1824	0	1787	19	0
2	A	99	0	0	3	0
2	B	72	0	0	3	0
All	All	3807	0	3559	36	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 5.

All (36) 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:61:LYS:HE2	1:A:207:GLN:OE1	1.40	1.19
1:B:61:LYS:HE2	1:B:207:GLN:OE1	1.54	1.05
1:B:117:ASP:HB2	2:B:358:HOH:O	1.56	1.05
1:A:61:LYS:CE	1:A:207:GLN:OE1	2.09	1.00
1:B:61:LYS:CE	1:B:207:GLN:OE1	2.15	0.95
1:B:145:ALA:HA	2:B:327:HOH:O	1.73	0.89
1:B:61:LYS:NZ	1:B:207:GLN:OE1	2.15	0.79
1:A:61:LYS:NZ	1:A:207:GLN:OE1	2.26	0.67
1:B:61:LYS:HE2	1:B:207:GLN:CD	2.20	0.67
1:B:81:HIS:HD2	1:B:197:ASP:H	1.41	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:81:HIS:CD2	1:B:197:ASP:H	2.12	0.66
1:B:18:LEU:C	1:B:18:LEU:HD23	2.26	0.61
1:B:143:TYR:CE1	1:B:207:GLN:NE2	2.68	0.58
1:A:124:GLU:HG2	2:A:390:HOH:O	2.07	0.55
1:A:145:ALA:HA	2:A:305:HOH:O	2.06	0.55
1:B:164:TYR:O	1:B:165:PHE:HB3	2.09	0.52
1:B:21:ASP:OD2	1:B:26:LYS:HE2	2.12	0.50
1:A:151:ASN:O	1:A:163:ALA:HA	2.12	0.50
1:A:58:PRO:HA	1:A:61:LYS:HE3	1.93	0.49
1:A:18:LEU:C	1:A:18:LEU:HD23	2.36	0.49
1:A:93:VAL:O	1:A:185:ASN:HA	2.12	0.49
1:A:143:TYR:CE1	1:A:207:GLN:NE2	2.80	0.49
1:A:143:TYR:CD1	1:A:207:GLN:NE2	2.81	0.48
1:A:162:LYS:HG2	1:A:184:GLN:OE1	2.13	0.48
1:A:182:TYR:CD2	1:B:101:LYS:HE2	2.50	0.46
1:A:143:TYR:CE2	1:A:207:GLN:HG2	2.50	0.46
1:A:136:ILE:HD11	1:A:177:LEU:HD21	1.98	0.45
1:A:163:ALA:HB3	1:A:183:GLN:HB3	1.98	0.44
1:B:136:ILE:HD11	1:B:177:LEU:HD21	1.98	0.44
1:B:96:ARG:HA	1:B:182:TYR:O	2.19	0.43
1:B:61:LYS:CE	1:B:207:GLN:CD	2.86	0.42
1:B:144:ASN:OD1	2:B:301:HOH:O	2.22	0.42
1:A:110:ALA:HA	1:A:122:ARG:O	2.21	0.41
1:B:4:GLY:HA3	1:B:85:LYS:O	2.20	0.41
1:A:126:LYS:HE2	2:A:337:HOH:O	2.21	0.40
1:B:93:VAL:O	1:B:185:ASN:HA	2.21	0.40

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	224/243 (92%)	220 (98%)	4 (2%)	0	100	100
1	B	225/243 (93%)	218 (97%)	7 (3%)	0	100	100
All	All	449/486 (92%)	438 (98%)	11 (2%)	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	196/208 (94%)	192 (98%)	4 (2%)	50	32
1	B	197/208 (95%)	190 (96%)	7 (4%)	30	10
All	All	393/416 (94%)	382 (97%)	11 (3%)	38	18

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	LYS
1	A	156	LYS
1	A	162	LYS
1	A	214	LYS
1	B	5	GLU
1	B	122	ARG
1	B	146	ILE
1	B	162	LYS
1	B	204	GLN
1	B	214	LYS
1	B	222	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	81	HIS
1	A	151	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	204	GLN
1	B	81	HIS
1	B	212	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	CRF	B	65	1	25,26,27	3.77	10 (40%)	32,37,39	4.01	18 (56%)
1	CRF	A	65	1	25,26,27	4.97	8 (32%)	32,37,39	2.85	12 (37%)

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	CRF	B	65	1	-	1/10/31/32	0/3/3/3
1	CRF	A	65	1	-	0/10/31/32	0/3/3/3

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	65	CRF	CB2-CA2	22.34	1.53	1.35
1	B	65	CRF	CB2-CA2	16.54	1.48	1.35
1	A	65	CRF	CA2-C2	-6.65	1.42	1.48
1	A	65	CRF	O2-C2	4.89	1.33	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	65	CRF	CE3-CD2	-3.86	1.34	1.42
1	A	65	CRF	C1-N2	3.18	1.36	1.32
1	A	65	CRF	CZ3-CE3	3.10	1.43	1.36
1	B	65	CRF	CA1-C1	-2.91	1.47	1.51
1	A	65	CRF	CH2-CZ2	2.86	1.43	1.36
1	B	65	CRF	CA2-C2	-2.86	1.45	1.48
1	B	65	CRF	CG2-CD2	2.83	1.47	1.41
1	B	65	CRF	C2-N3	-2.52	1.33	1.39
1	A	65	CRF	CA1-C1	-2.49	1.47	1.51
1	B	65	CRF	CA1-N1	-2.41	1.40	1.47
1	A	65	CRF	CG2-CD2	2.32	1.46	1.41
1	B	65	CRF	CD2-CE2	2.21	1.48	1.42
1	B	65	CRF	CZ3-CE3	2.14	1.41	1.36
1	B	65	CRF	CH2-CZ2	2.06	1.41	1.36

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	65	CRF	O2-C2-CA2	-13.03	123.64	130.96
1	B	65	CRF	CA2-C2-N3	11.08	108.61	103.37
1	A	65	CRF	CE3-CD2-CE2	6.77	127.14	118.17
1	A	65	CRF	CA2-C2-N3	6.43	106.41	103.37
1	B	65	CRF	C1-CA1-N1	-5.64	100.81	109.96
1	A	65	CRF	CB2-CA2-C2	5.49	128.83	122.28
1	A	65	CRF	CZ3-CE3-CD2	-5.46	113.32	120.89
1	A	65	CRF	CB2-CA2-N2	-5.01	121.88	128.83
1	A	65	CRF	O3-C3-CA3	-4.70	112.20	126.39
1	B	65	CRF	C2-N3-C1	-4.66	105.61	107.97
1	B	65	CRF	CG2-CB2-CA2	-4.49	122.09	130.81
1	B	65	CRF	CB2-CA2-C2	4.32	127.43	122.28
1	B	65	CRF	OG1-CB1-CG1	-4.13	97.52	109.74
1	B	65	CRF	CA1-C1-N3	-3.67	120.34	124.75
1	B	65	CRF	C2-CA2-N2	-3.57	106.43	108.93
1	B	65	CRF	CE3-CD2-CE2	3.47	122.77	118.17
1	A	65	CRF	CE3-CD2-CG2	-3.28	129.87	135.45
1	B	65	CRF	O3-C3-CA3	-3.22	116.65	126.39
1	A	65	CRF	OG1-CB1-CG1	-3.17	100.36	109.74
1	B	65	CRF	CA3-N3-C2	3.05	130.80	123.80
1	B	65	CRF	CA3-N3-C1	-3.02	123.54	127.16
1	B	65	CRF	CG1-CB1-CA1	-2.78	105.62	112.16
1	A	65	CRF	CG2-CB2-CA2	-2.68	125.61	130.81
1	B	65	CRF	OG1-CB1-CA1	2.48	114.35	109.04

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	65	CRF	CA3-N3-C1	-2.31	124.39	127.16
1	A	65	CRF	C1-CA1-N1	-2.19	106.41	109.96
1	B	65	CRF	CA1-C1-N2	2.16	126.92	123.89
1	A	65	CRF	O2-C2-CA2	2.11	132.15	130.96
1	B	65	CRF	CZ2-CE2-CD2	-2.08	116.96	120.76
1	B	65	CRF	CZ3-CE3-CD2	-2.06	118.04	120.89

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	65	CRF	C3-CA3-N3-C2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides 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 (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	64:LEU	C	65:CRF	N1	1.66

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	227/243 (93%)	0.36	9 (3%)	43	50	21, 38, 63, 82	1 (0%)
1	B	229/243 (94%)	0.59	15 (6%)	26	31	31, 41, 67, 99	0
All	All	456/486 (93%)	0.48	24 (5%)	33	39	21, 40, 67, 99	1 (0%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	231	LEU	5.5
1	A	146	ILE	4.5
1	B	164	TYR	4.4
1	B	145	ALA	4.2
1	B	143	TYR	3.6
1	A	145	ALA	3.1
1	B	230	PRO	3.0
1	A	143	TYR	3.0
1	A	2	SER	2.6
1	B	133	ASP	2.6
1	B	1	VAL	2.5
1	B	146	ILE	2.4
1	A	231	LEU	2.3
1	B	0	MET	2.3
1	A	133	ASP	2.3
1	A	176	VAL	2.2
1	B	229	ILE	2.2
1	B	211	PRO	2.2
1	B	144	ASN	2.2
1	A	3	LYS	2.1
1	B	157	GLN	2.1
1	B	184	GLN	2.1
1	A	206	LYS	2.0
1	B	2	SER	2.0

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, 95^{th} 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	CRF	A	65	24/25	0.93	0.09	28,34,43,47	0
1	CRF	B	65	24/25	0.95	0.08	28,36,43,45	0

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.