



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 9, 2025 – 03:17 PM JST

PDB ID : 9U9D / pdb\_00009u9d  
Title : Bipartite Genetically Encoded Biosensor sG-GECO1  
Authors : Wen, Y.; Campbell, R.E.  
Deposited on : 2025-03-27  
Resolution : 1.80 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](mailto: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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.45.1

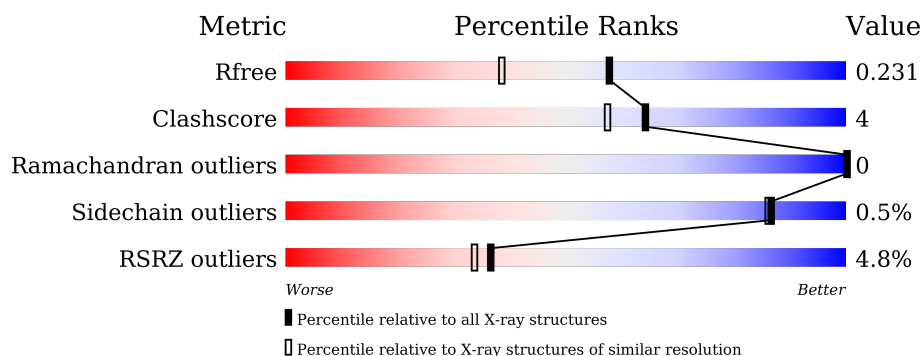
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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  $\geq 3$ , 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	234	<div> <div style="width: 10%; background-color: red;"></div> <div style="width: 10%; background-color: orange;"></div> <div style="width: 10%; background-color: yellow;"></div> <div style="width: 10%; background-color: green;"></div> <div style="width: 50%; background-color: grey;"></div> </div> <div> <div style="width: 10%; background-color: red;"></div> <div style="width: 10%; background-color: orange;"></div> <div style="width: 10%; background-color: yellow;"></div> <div style="width: 10%; background-color: green;"></div> <div style="width: 50%; background-color: grey;"></div> </div>
2	B	232	<div> <div style="width: 3%; background-color: red;"></div> <div style="width: 76%; background-color: green;"></div> <div style="width: 13%; background-color: grey;"></div> </div> <div> <div style="width: 3%; background-color: red;"></div> <div style="width: 76%; background-color: green;"></div> <div style="width: 13%; background-color: grey;"></div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2033 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 Myosin light chain kinase, smooth muscle, deglutamylated form, Green fluorescent protein, Calmodulin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	30	Total	C	N	O	S	0	1	0
			239	148	46	42	3			

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	MET	LYS	conflict	UNP P11799
A	38	VAL	LYS	conflict	UNP P11799
A	39	ASP	TYR	conflict	UNP P11799
A	40	SER	MET	conflict	UNP P11799
A	41	SER	ALA	conflict	UNP P11799
A	46	TYR	GLN	conflict	UNP P11799
A	54	VAL	ALA	conflict	UNP P11799
A	55	LEU	ILE	conflict	UNP P11799
A	56	ARG	GLY	conflict	UNP P11799
A	70	ARG	LYS	conflict	UNP P42212
A	75	ALA	VAL	conflict	UNP P42212
A	78	GLN	LYS	conflict	UNP P42212
A	80	CYS	ARG	conflict	UNP P42212
A	83	LEU	ILE	conflict	UNP P42212
A	84	VAL	GLU	conflict	UNP P42212
A	85	SER	-	linker	UNP P42212
A	86	LEU	-	linker	UNP P42212
A	87	ARG	-	linker	UNP P42212
A	98	ILE	PHE	conflict	UNP P0DP29
A	101	LEU	ALA	conflict	UNP P0DP29
A	116	ARG	LYS	conflict	UNP P0DP29
A	146	ASP	ASN	conflict	UNP P0DP29
A	156	ALA	THR	conflict	UNP P0DP29
A	164	TYR	ASP	conflict	UNP P0DP29
A	165	ARG	THR	conflict	UNP P0DP29
A	167	THR	SER	conflict	UNP P0DP29

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	175	TYR	PHE	conflict	UNP P0DP29
A	176	GLY	ARG	conflict	UNP P0DP29
A	211	LEU	ILE	conflict	UNP P0DP29

- Molecule 2 is a protein called Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	201	Total	C	N	O	S	0	2	0
			1620	1042	265	308	5			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P42212
B	2	ALA	-	expression tag	UNP P42212
B	3	SER	-	expression tag	UNP P42212
B	4	TRP	-	expression tag	UNP P42212
B	5	SER	-	expression tag	UNP P42212
B	6	HIS	-	expression tag	UNP P42212
B	7	PRO	-	expression tag	UNP P42212
B	8	GLN	-	expression tag	UNP P42212
B	9	PHE	-	expression tag	UNP P42212
B	10	GLU	-	expression tag	UNP P42212
B	11	LYS	-	expression tag	UNP P42212
B	12	VAL	-	expression tag	UNP P42212
B	14	SER	ILE	conflict	UNP P42212
B	15	MET	GLU	conflict	UNP P42212
B	18	GLY	SER	conflict	UNP P42212
B	23	TYR	ASP	conflict	UNP P42212
B	48	THR	SER	conflict	UNP P42212
B	49	LYS	ALA	conflict	UNP P42212
B	74	ARG	HIS	conflict	UNP P42212
B	79	PRO	LEU	conflict	UNP P42212
B	82	GLY	-	linker	UNP P42212
B	83	GLY	-	linker	UNP P42212
B	84	THR	-	linker	UNP P42212
B	85	GLY	-	linker	UNP P42212
B	86	GLY	-	linker	UNP P42212
B	87	SER	-	linker	UNP P42212
B	88	MET	-	linker	UNP P42212
B	89	VAL	-	linker	UNP P42212
B	152	LEU	PHE	conflict	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	154	CRO	SER	chromophore	UNP P42212
B	154	CRO	TYR	chromophore	UNP P42212
B	154	CRO	GLY	chromophore	UNP P42212
B	181	ILE	VAL	conflict	UNP P42212
B	233	PHE	TYR	conflict	UNP P42212

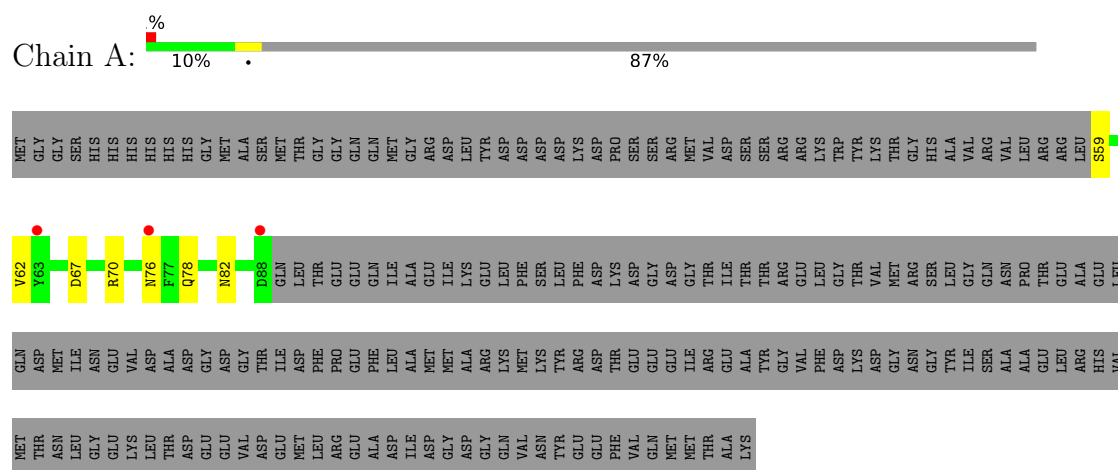
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	8	Total O 8 8	0	0
3	B	166	Total O 166 166	0	0

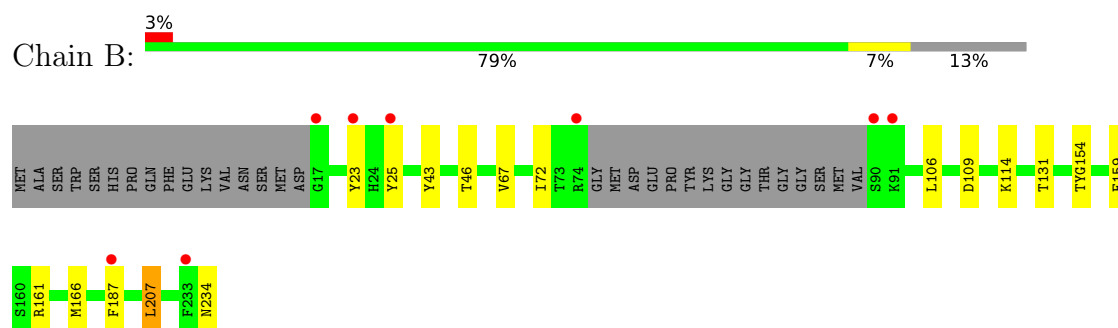
### 3 Residue-property plots

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: Myosin light chain kinase, smooth muscle, deglutamylated form, Green fluorescent protein, Calmodulin-1



- Molecule 2: Green fluorescent protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.10Å 51.10Å 201.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.14 – 1.80 36.14 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (36.14-1.80) 99.9 (36.14-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.01 (at 1.81Å)	Xtriage
Refinement program	PHENIX (1.19_4092: ???)	Depositor
R, $R_{free}$	0.183 , 0.227 0.184 , 0.231	Depositor DCC
$R_{free}$ test set	2000 reflections (7.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.6	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 37.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2033	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: CRO

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	0.37	0/244	0.72	0/326
2	B	0.33	0/1642	0.60	1/2218 (0.0%)
All	All	0.33	0/1886	0.61	1/2544 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	207	LEU	CA-CB-CG	-5.11	98.42	116.30

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	A	239	0	242	5	0
2	B	1620	0	1569	13	0
3	A	8	0	0	0	0
3	B	166	0	0	2	0
All	All	2033	0	1811	14	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.



All (14) 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:B:187[B]:PHE:H	2:B:187[B]:PHE:HD1	1.45	0.65
2:B:109:ASP:OD2	2:B:114:LYS:HE3	2.00	0.62
2:B:46:THR:HG23	2:B:67:VAL:HG22	1.81	0.62
1:A:78:GLN:OE1	2:B:23[B]:TYR:OH	2.04	0.60
1:A:76:ASN:OD1	2:B:25:TYR:HD1	1.84	0.59
2:B:131:THR:HG23	3:B:422:HOH:O	2.07	0.55
2:B:72:ILE:HD12	2:B:166:MET:HE3	1.90	0.54
2:B:159:PHE:HE2	2:B:207:LEU:HD22	1.81	0.46
1:A:67:ASP:CG	1:A:70:ARG:HG2	2.42	0.45
1:A:82:ASN:OD1	2:B:234:ASN:ND2	2.52	0.42
1:A:62:VAL:O	2:B:43:TYR:HA	2.20	0.42
2:B:161:ARG:NH1	3:B:301:HOH:O	2.25	0.42
2:B:106:LEU:HD23	2:B:106:LEU:C	2.46	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	29/234 (12%)	29 (100%)	0	0	100	100
2	B	196/232 (84%)	194 (99%)	2 (1%)	0	100	100
All	All	225/466 (48%)	223 (99%)	2 (1%)	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	27/200 (14%)	26 (96%)	1 (4%)	29	17
2	B	174/199 (87%)	174 (100%)	0	100	100
All	All	201/399 (50%)	200 (100%)	1 (0%)	86	86

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	ASN
2	B	193	ASN
2	B	223	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
2	CRO	B	154	2	23,23,24	2.43	7 (30%)	30,32,34	2.40	7 (23%)

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
2	CRO	B	154	2	-	0/12/31/32	0/2/2/2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	154	CRO	CA2-C2	-8.92	1.39	1.48
2	B	154	CRO	CA1-C1	3.72	1.56	1.51
2	B	154	CRO	C1-N2	2.97	1.36	1.32
2	B	154	CRO	CG2-CB2	2.75	1.52	1.46
2	B	154	CRO	C2-N3	-2.33	1.34	1.39
2	B	154	CRO	CA2-N2	-2.11	1.34	1.38
2	B	154	CRO	CG1-CB1	2.04	1.57	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	154	CRO	CA2-C2-N3	9.23	107.74	103.37
2	B	154	CRO	O2-C2-CA2	-6.30	127.42	130.96
2	B	154	CRO	OG1-CB1-CA1	2.81	115.05	109.04
2	B	154	CRO	C1-CA1-N1	-2.69	105.60	109.96
2	B	154	CRO	N3-C1-N2	-2.58	109.67	111.45
2	B	154	CRO	O3-C3-CA3	-2.21	119.72	126.39
2	B	154	CRO	CG1-CB1-CA1	-2.02	107.40	112.16

There are no chirality outliers.

There are no torsion outliers.

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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	30/234 (12%)	0.69	3 (10%) 14 11	32, 50, 76, 99	1 (3%)
2	B	200/232 (86%)	-0.03	8 (4%) 43 40	26, 41, 59, 88	2 (1%)
All	All	230/466 (49%)	0.06	11 (4%) 36 34	26, 42, 64, 99	3 (1%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	90	SER	4.5
2	B	23[A]	TYR	3.3
2	B	25	TYR	3.2
2	B	187[A]	PHE	2.9
2	B	91	LYS	2.8
1	A	88	ASP	2.8
2	B	17	GLY	2.7
1	A	76	ASN	2.6
2	B	74	ARG	2.1
2	B	233	PHE	2.1
1	A	63	TYR	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	CRO	B	154	22/23	0.96	0.06	28,31,37,41	0

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides 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.