



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 11, 2018 – 02:40 PM EST

PDB ID : 6B7T  
Title : Truncated strand 10-less green fluorescent protein  
Authors : Deng, A.; Boxer, S.G.  
Deposited on : 2017-10-05  
Resolution : 1.91 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

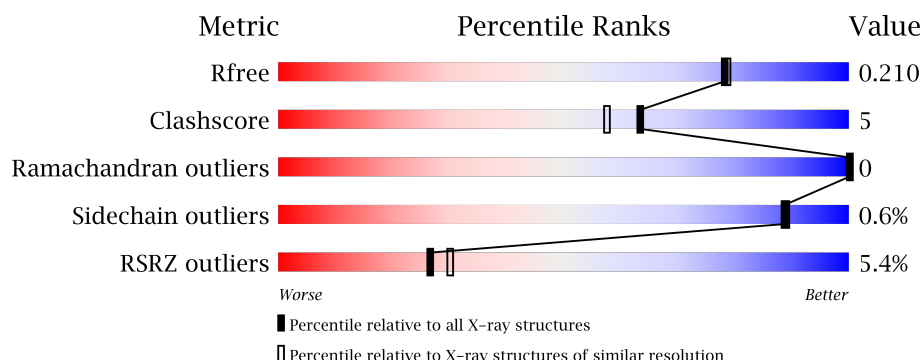
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.91 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))
$R_{free}$	100719	6276 (1.94-1.90)
Clashscore	112137	7025 (1.94-1.90)
Ramachandran outliers	110173	6947 (1.94-1.90)
Sidechain outliers	110143	6948 (1.94-1.90)
RSRZ outliers	101464	6332 (1.94-1.90)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	246	<div> <div>2%</div> <div>77%</div> <div>8%</div> <div>15%</div> </div>
1	B	246	<div> <div>7%</div> <div>74%</div> <div>9%</div> <div>17%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3282 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 Green fluorescent protein, Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	0	1	0
			1634	1041	288	302	3			
1	B	204	Total	C	N	O	S	0	0	0
			1555	994	269	289	3			

There are 110 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	expression tag	UNP P42212
A	-22	GLY	-	expression tag	UNP P42212
A	-21	SER	-	expression tag	UNP P42212
A	-20	SER	-	expression tag	UNP P42212
A	-19	HIS	-	expression tag	UNP P42212
A	-18	HIS	-	expression tag	UNP P42212
A	-17	HIS	-	expression tag	UNP P42212
A	-16	HIS	-	expression tag	UNP P42212
A	-15	HIS	-	expression tag	UNP P42212
A	-14	HIS	-	expression tag	UNP P42212
A	-13	SER	-	expression tag	UNP P42212
A	-12	SER	-	expression tag	UNP P42212
A	-11	GLY	-	expression tag	UNP P42212
A	-10	LEU	-	expression tag	UNP P42212
A	-9	VAL	-	expression tag	UNP P42212
A	-8	PRO	-	expression tag	UNP P42212
A	-4	GLY	-	expression tag	UNP P42212
A	-3	GLY	-	expression tag	UNP P42212
A	-2	SER	-	expression tag	UNP P42212
A	-1	HIS	-	expression tag	UNP P42212
A	0	MET	-	expression tag	UNP P42212
A	8	HIS	LEU	engineered mutation	UNP P42212
A	10	TYR	PHE	engineered mutation	UNP P42212
A	12	ASN	THR	engineered mutation	UNP P42212
A	25	GLY	-	linker	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	GLY	-	linker	UNP P42212
A	27	THR	-	linker	UNP P42212
A	28	GLY	-	linker	UNP P42212
A	29	GLY	-	linker	UNP P42212
A	30	SER	-	linker	UNP P42212
A	31	ALA	-	linker	UNP P42212
A	32	SER	-	linker	UNP P42212
A	33	GLN	LYS	engineered mutation	UNP P42212
A	60	ARG	SER	engineered mutation	UNP P42212
A	69	ILE	TYR	engineered mutation	UNP P42212
A	78	SER	CYS	engineered mutation	UNP P42212
A	94	LEU	PHE	engineered mutation	UNP P42212
A	96	GYS	SER	chromophore	UNP P42212
A	96	GYS	TYR	chromophore	UNP P42212
A	96	GYS	GLY	chromophore	UNP P42212
A	100	ALA	CYS	engineered mutation	UNP P42212
A	110	ARG	GLN	engineered mutation	UNP P42212
A	129	SER	PHE	engineered mutation	UNP P42212
A	135	LYS	ASN	engineered mutation	UNP P42212
A	141	VAL	GLU	engineered mutation	UNP P42212
A	158	THR	ILE	engineered mutation	UNP P42212
A	175	PHE	TYR	engineered mutation	UNP P42212
A	183	THR	MET	engineered mutation	UNP P42212
A	193	ALA	VAL	engineered mutation	UNP P42212
A	196	THR	LYS	engineered mutation	UNP P42212
A	197	VAL	ILE	engineered mutation	UNP P42212
A	201	VAL	ILE	engineered mutation	UNP P42212
A	225	GLY	-	expression tag	UNP P42212
A	226	THR	-	expression tag	UNP P42212
A	227	ARG	-	expression tag	UNP P42212
B	-23	MET	-	expression tag	UNP P42212
B	-22	GLY	-	expression tag	UNP P42212
B	-21	SER	-	expression tag	UNP P42212
B	-20	SER	-	expression tag	UNP P42212
B	-19	HIS	-	expression tag	UNP P42212
B	-18	HIS	-	expression tag	UNP P42212
B	-17	HIS	-	expression tag	UNP P42212
B	-16	HIS	-	expression tag	UNP P42212
B	-15	HIS	-	expression tag	UNP P42212
B	-14	HIS	-	expression tag	UNP P42212
B	-13	SER	-	expression tag	UNP P42212
B	-12	SER	-	expression tag	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-11	GLY	-	expression tag	UNP P42212
B	-10	LEU	-	expression tag	UNP P42212
B	-9	VAL	-	expression tag	UNP P42212
B	-8	PRO	-	expression tag	UNP P42212
B	-4	GLY	-	expression tag	UNP P42212
B	-3	GLY	-	expression tag	UNP P42212
B	-2	SER	-	expression tag	UNP P42212
B	-1	HIS	-	expression tag	UNP P42212
B	0	MET	-	expression tag	UNP P42212
B	8	HIS	LEU	engineered mutation	UNP P42212
B	10	TYR	PHE	engineered mutation	UNP P42212
B	12	ASN	THR	engineered mutation	UNP P42212
B	25	GLY	-	linker	UNP P42212
B	26	GLY	-	linker	UNP P42212
B	27	THR	-	linker	UNP P42212
B	28	GLY	-	linker	UNP P42212
B	29	GLY	-	linker	UNP P42212
B	30	SER	-	linker	UNP P42212
B	31	ALA	-	linker	UNP P42212
B	32	SER	-	linker	UNP P42212
B	33	GLN	LYS	engineered mutation	UNP P42212
B	60	ARG	SER	engineered mutation	UNP P42212
B	69	ILE	TYR	engineered mutation	UNP P42212
B	78	SER	CYS	engineered mutation	UNP P42212
B	94	LEU	PHE	engineered mutation	UNP P42212
B	96	GYS	SER	chromophore	UNP P42212
B	96	GYS	TYR	chromophore	UNP P42212
B	96	GYS	GLY	chromophore	UNP P42212
B	100	ALA	CYS	engineered mutation	UNP P42212
B	110	ARG	GLN	engineered mutation	UNP P42212
B	129	SER	PHE	engineered mutation	UNP P42212
B	135	LYS	ASN	engineered mutation	UNP P42212
B	141	VAL	GLU	engineered mutation	UNP P42212
B	158	THR	ILE	engineered mutation	UNP P42212
B	175	PHE	TYR	engineered mutation	UNP P42212
B	183	THR	MET	engineered mutation	UNP P42212
B	193	ALA	VAL	engineered mutation	UNP P42212
B	196	THR	LYS	engineered mutation	UNP P42212
B	197	VAL	ILE	engineered mutation	UNP P42212
B	201	VAL	ILE	engineered mutation	UNP P42212
B	225	GLY	-	expression tag	UNP P42212
B	226	THR	-	expression tag	UNP P42212

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	227	ARG	-	expression tag	UNP P42212

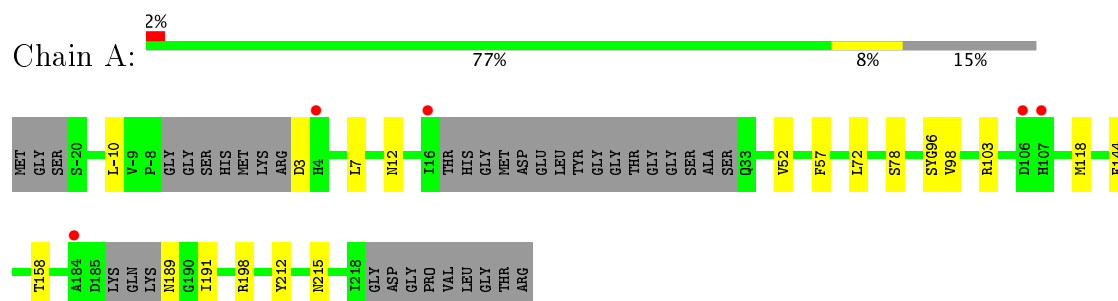
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	55	Total O 55 55	0	0
2	B	38	Total O 38 38	0	0

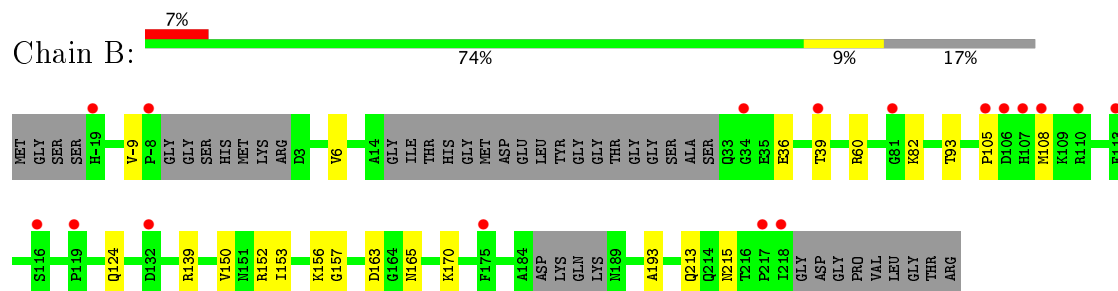
### 3 Residue-property plots [i](#)

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: Green fluorescent protein, Green fluorescent protein



- Molecule 1: Green fluorescent protein, Green fluorescent protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.37Å 67.21Å 59.99Å 90.00° 98.81° 90.00°	Depositor
Resolution (Å)	35.12 – 1.91 35.12 – 1.91	Depositor EDS
% Data completeness (in resolution range)	98.2 (35.12-1.91) 98.2 (35.12-1.91)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 1.91Å)	Xtriage
Refinement program	PHENIX 1.10.1 _2155	Depositor
R, $R_{free}$	0.179 , 0.209 0.178 , 0.210	Depositor DCC
$R_{free}$ test set	1482 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.1	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 69.6	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	3282	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.81% 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

### 5.1 Standard geometry

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

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	0.56	0/1652	0.71	0/2237
1	B	0.48	0/1569	0.67	0/2130
All	All	0.52	0/3221	0.69	0/4367

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

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	1634	0	1530	15	1
1	B	1555	0	1437	15	1
2	A	55	0	0	4	0
2	B	38	0	0	2	0
All	All	3282	0	2967	30	1

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.

The worst 5 of 30 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:3:ASP:N	2:A:301:HOH:O	1.91	1.01
1:A:118:MET:HE2	1:A:144:PHE:HD1	1.33	0.92
1:B:124:GLN:HE21	1:B:215:ASN:HD21	1.29	0.80
1:B:-9:VAL:HG23	1:B:6:VAL:HB	1.69	0.74
1:B:105:PRO:HG2	1:B:108:MET:HG2	1.71	0.71

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:158:THR:OG1	1:B:152:ARG:NH2[1_554]	2.17	0.03

## 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	200/246 (81%)	196 (98%)	4 (2%)	0	100	100
1	B	195/246 (79%)	191 (98%)	4 (2%)	0	100	100
All	All	395/492 (80%)	387 (98%)	8 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	167/205 (82%)	166 (99%)	1 (1%)	89	89
1	B	154/205 (75%)	153 (99%)	1 (1%)	89	89
All	All	321/410 (78%)	319 (99%)	2 (1%)	89	89

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	198	ARG
1	B	139	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	124	GLN
1	B	179	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	GYS	A	96	1	22,22,23	1.88	9 (40%)	27,30,32	1.85	5 (18%)
1	GYS	B	96	1	22,22,23	2.00	4 (18%)	27,30,32	1.70	4 (14%)

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	GYS	A	96	1	-	0/9/29/30	0/2/2/2
1	GYS	B	96	1	-	0/9/29/30	0/2/2/2

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	96	GYS	CA1-C1	-6.47	1.42	1.51
1	A	96	GYS	CB2-CA2	-2.64	1.32	1.35
1	A	96	GYS	OG1-CB1	-2.10	1.33	1.42
1	A	96	GYS	CA2-N2	2.20	1.43	1.38
1	A	96	GYS	CG2-CB2	2.30	1.51	1.46

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	96	GYS	O-C-CA3	-6.04	106.40	126.38
1	A	96	GYS	O-C-CA3	-6.02	106.47	126.38
1	A	96	GYS	CA1-C1-N3	-3.43	120.63	124.75
1	B	96	GYS	CA1-C1-N3	-3.27	120.83	124.75
1	A	96	GYS	CD1-CE1-CZ	-2.54	117.00	119.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	96	GYS	1	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 [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	207/246 (84%)	0.09	5 (2%) 59 63	19, 38, 66, 80	0
1	B	203/246 (82%)	0.49	17 (8%) 12 14	20, 47, 77, 97	0
All	All	410/492 (83%)	0.29	22 (5%) 26 30	19, 42, 72, 97	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	106	ASP	7.1
1	B	105	PRO	5.5
1	B	107	HIS	4.9
1	B	108	MET	4.5
1	A	16	ILE	4.2

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	GYS	A	96	21/22	0.97	0.13	-	18,21,26,29	0
1	GYS	B	96	21/22	0.98	0.12	-	19,24,26,29	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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.