



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 11:47 pm BST

PDB ID : 4HJ0  
Title : Crystal structure of the human GIPr ECD in complex with Gipg013 Fab at 3-Å resolution  
Authors : Madhurantakam, C.; Ravn, P.; Gruetter, M.G.; Jackson, R.H.  
Deposited on : 2012-10-12  
Resolution : 3.00 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

<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
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

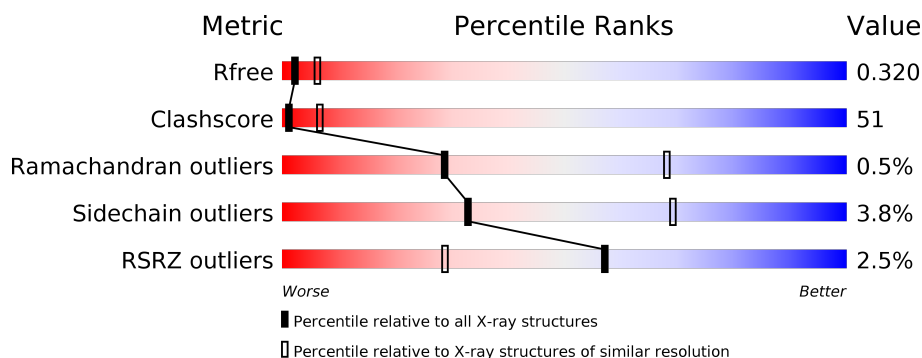
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	136	<div> <div>35%</div> <div>30%</div> <div>••</div> <div>32%</div> </div>
1	B	136	<div> <div>8%</div> <div>20%</div> <div>41%</div> <div>••</div> <div>33%</div> </div>
2	C	227	<div> <div>3%</div> <div>39%</div> <div>47%</div> <div>•</div> <div>10%</div> </div>
2	P	227	<div> <div>51%</div> <div>37%</div> <div>•</div> <div>9%</div> </div>
3	D	215	<div> <div>3%</div> <div>35%</div> <div>57%</div> <div>6%</div> <div>•</div> </div>
3	Q	215	<div> <div>44%</div> <div>49%</div> <div>••</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7635 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 Gastric inhibitory polypeptide receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	92	Total	C	N	O	S	2	0	0
			742	465	135	135	7			
1	B	91	Total	C	N	O	S	2	0	0
			727	457	131	132	7			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	MET	-	EXPRESSION TAG	UNP P48546
A	4	GLY	-	EXPRESSION TAG	UNP P48546
A	5	SER	-	EXPRESSION TAG	UNP P48546
A	6	SER	-	EXPRESSION TAG	UNP P48546
A	7	HIS	-	EXPRESSION TAG	UNP P48546
A	8	HIS	-	EXPRESSION TAG	UNP P48546
A	9	HIS	-	EXPRESSION TAG	UNP P48546
A	10	HIS	-	EXPRESSION TAG	UNP P48546
A	11	HIS	-	EXPRESSION TAG	UNP P48546
A	12	HIS	-	EXPRESSION TAG	UNP P48546
A	13	SER	-	EXPRESSION TAG	UNP P48546
A	14	ASP	-	EXPRESSION TAG	UNP P48546
A	15	TYR	-	EXPRESSION TAG	UNP P48546
A	16	LYS	-	EXPRESSION TAG	UNP P48546
A	17	ASP	-	EXPRESSION TAG	UNP P48546
A	18	ASP	-	EXPRESSION TAG	UNP P48546
A	19	ASP	-	EXPRESSION TAG	UNP P48546
A	20	ASP	-	EXPRESSION TAG	UNP P48546
A	21	LYS	-	EXPRESSION TAG	UNP P48546
A	22	HIS	-	EXPRESSION TAG	UNP P48546
A	23	MET	-	EXPRESSION TAG	UNP P48546
B	3	MET	-	EXPRESSION TAG	UNP P48546
B	4	GLY	-	EXPRESSION TAG	UNP P48546
B	5	SER	-	EXPRESSION TAG	UNP P48546
B	6	SER	-	EXPRESSION TAG	UNP P48546

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	7	HIS	-	EXPRESSION TAG	UNP P48546
B	8	HIS	-	EXPRESSION TAG	UNP P48546
B	9	HIS	-	EXPRESSION TAG	UNP P48546
B	10	HIS	-	EXPRESSION TAG	UNP P48546
B	11	HIS	-	EXPRESSION TAG	UNP P48546
B	12	HIS	-	EXPRESSION TAG	UNP P48546
B	13	SER	-	EXPRESSION TAG	UNP P48546
B	14	ASP	-	EXPRESSION TAG	UNP P48546
B	15	TYR	-	EXPRESSION TAG	UNP P48546
B	16	LYS	-	EXPRESSION TAG	UNP P48546
B	17	ASP	-	EXPRESSION TAG	UNP P48546
B	18	ASP	-	EXPRESSION TAG	UNP P48546
B	19	ASP	-	EXPRESSION TAG	UNP P48546
B	20	ASP	-	EXPRESSION TAG	UNP P48546
B	21	LYS	-	EXPRESSION TAG	UNP P48546
B	22	HIS	-	EXPRESSION TAG	UNP P48546
B	23	MET	-	EXPRESSION TAG	UNP P48546

- Molecule 2 is a protein called Gipg013 Fab, Antagonizing antibody to the GIP Receptor, Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	206	Total	C	N	O	S	0	0	0
			1518	959	252	301	6			
2	C	205	Total	C	N	O	S	0	0	0
			1512	956	251	299	6			

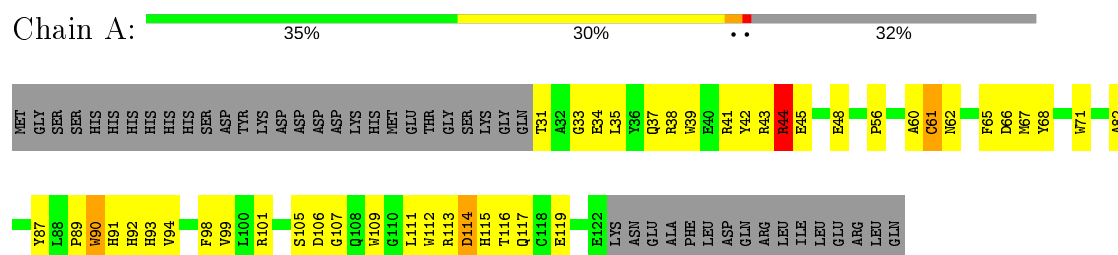
- Molecule 3 is a protein called Gipg013 Fab, Antagonizing antibody to the GIP Receptor, Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	Q	210	Total	C	N	O	S	0	0	0
			1566	976	266	320	4			
3	D	211	Total	C	N	O	S	0	0	0
			1570	978	267	321	4			

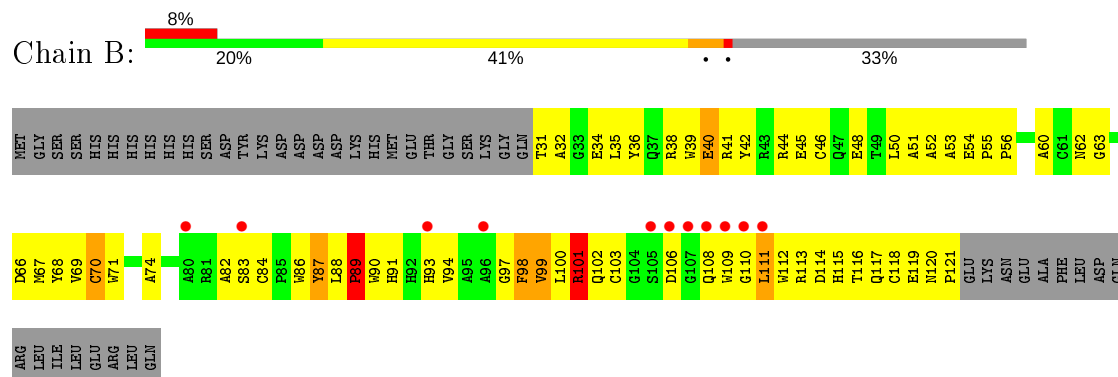
### 3 Residue-property plots [i](#)

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.

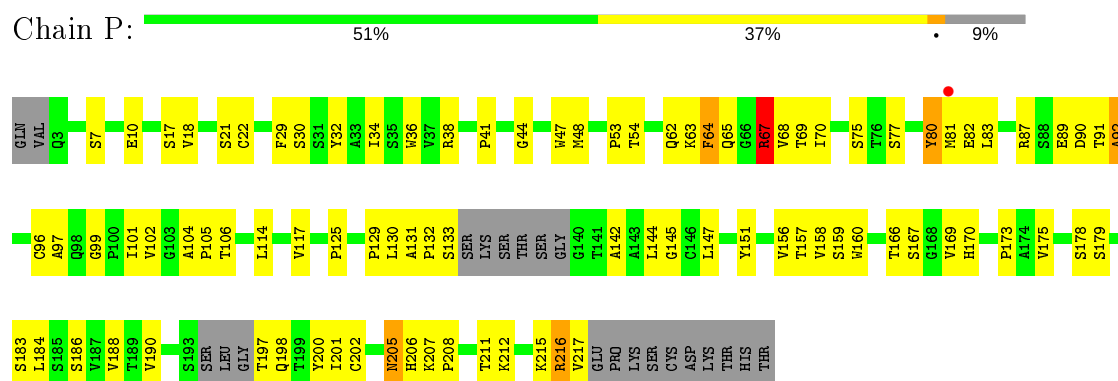
- Molecule 1: Gastric inhibitory polypeptide receptor



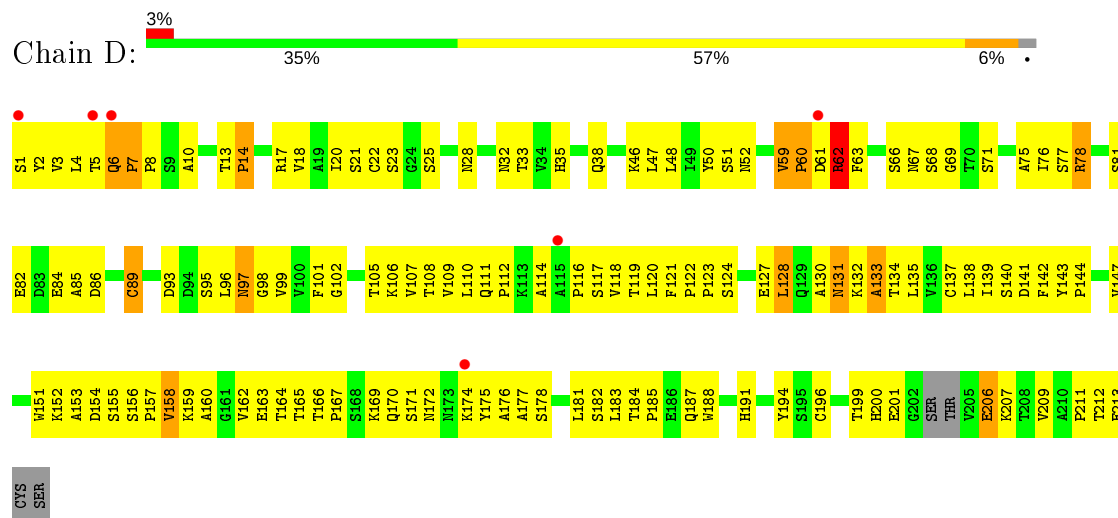
- Molecule 1: Gastric inhibitory polypeptide receptor



- Molecule 2: Gipg013 Fab, Antagonizing antibody to the GIP Receptor, Heavy chain



- Molecule 2: Gipg013 Fab, Antagonizing antibody to the GIP Receptor, Heavy chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.27Å 109.85Å 105.94Å 90.00° 97.76° 90.00°	Depositor
Resolution (Å)	48.66 – 3.00 48.67 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.66-3.00) 100.0 (48.67-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.255 , 0.311 0.272 , 0.320	Depositor DCC
$R_{free}$ test set	1999 reflections (9.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.9	Xtriage
Anisotropy	0.501	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	7635	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% 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

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	0.53	1/770 (0.1%)	0.86	4/1053 (0.4%)
1	B	1.14	7/755 (0.9%)	1.10	10/1034 (1.0%)
2	C	0.33	0/1547	0.71	6/2109 (0.3%)
2	P	0.33	1/1553 (0.1%)	0.96	4/2117 (0.2%)
3	D	0.55	3/1608 (0.2%)	0.86	14/2198 (0.6%)
3	Q	0.59	3/1604 (0.2%)	0.68	5/2193 (0.2%)
All	All	0.57	15/7837 (0.2%)	0.85	43/10704 (0.4%)

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	B	0	1
2	C	0	1
2	P	0	3
3	D	0	2
All	All	0	7

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	40	GLU	CD-OE2	18.46	1.46	1.25
3	D	60	PRO	N-CD	-15.34	1.26	1.47
3	Q	51	SER	C-N	12.99	1.64	1.34
3	Q	196	CYS	CB-SG	12.84	2.04	1.82
1	B	70	CYS	CB-SG	12.23	2.03	1.82

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	67	ARG	NE-CZ-NH2	25.64	133.12	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	67	ARG	NE-CZ-NH1	-24.38	108.11	120.30
1	B	98	PHE	O-C-N	13.37	144.08	122.70
1	A	61	CYS	CA-CB-SG	12.05	135.69	114.00
1	B	98	PHE	CA-C-N	-10.61	93.86	117.20

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	101	ARG	Sidechain
3	D	62	ARG	Peptide
2	P	178	SER	Peptide
2	P	64	PHE	Peptide
2	P	67	ARG	Sidechain

## 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	742	0	656	64	0
1	B	727	0	640	148	0
2	C	1512	0	1488	169	0
2	P	1518	0	1491	95	0
3	D	1570	0	1513	189	0
3	Q	1566	0	1508	149	0
All	All	7635	0	7296	752	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 51.

The worst 5 of 752 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:Q:196:CYS:SG	3:Q:196:CYS:CB	2.04	1.46
1:B:70:CYS:CB	1:B:70:CYS:SG	2.03	1.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:Q:173:ASN:O	3:Q:174:LYS:HG3	1.28	1.29
3:Q:207:LYS:NZ	3:Q:207:LYS:HB3	1.43	1.19
3:D:1:SER:HB2	3:D:2:TYR:CA	1.72	1.18

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	90/136 (66%)	82 (91%)	8 (9%)	0	100	100
1	B	89/136 (65%)	76 (85%)	12 (14%)	1 (1%)	14	50
2	C	199/227 (88%)	175 (88%)	23 (12%)	1 (0%)	29	68
2	P	200/227 (88%)	187 (94%)	13 (6%)	0	100	100
3	D	207/215 (96%)	168 (81%)	38 (18%)	1 (0%)	29	68
3	Q	206/215 (96%)	185 (90%)	19 (9%)	2 (1%)	15	53
All	All	991/1156 (86%)	873 (88%)	113 (11%)	5 (0%)	29	68

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	Q	156	SER
3	D	7	PRO
1	B	89	PRO
2	C	152	PHE
3	Q	7	PRO

### 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	73/113 (65%)	71 (97%)	2 (3%)	44	77
1	B	71/113 (63%)	68 (96%)	3 (4%)	30	66
2	C	168/188 (89%)	162 (96%)	6 (4%)	35	70
2	P	169/188 (90%)	162 (96%)	7 (4%)	30	67
3	D	176/180 (98%)	169 (96%)	7 (4%)	31	68
3	Q	176/180 (98%)	169 (96%)	7 (4%)	31	68
All	All	833/962 (87%)	801 (96%)	32 (4%)	33	69

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Q	77	SER
3	Q	207	LYS
2	C	116	THR
3	Q	206	GLU
3	Q	208	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
3	Q	54	GLN
3	Q	173	ASN
3	D	172	ASN
2	P	98	GLN
3	D	170	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	Q	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	Q	51:SER	C	52:ASN	N	1.63

## 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	92/136 (67%)	0.00	0	100	100	29, 53, 71, 87	1 (1%)
1	B	91/136 (66%)	0.86	11 (12%)	4	1	23, 100, 129, 141	1 (1%)
2	C	205/227 (90%)	0.15	6 (2%)	51	23	33, 53, 77, 91	0
2	P	206/227 (90%)	-0.26	1 (0%)	91	75	19, 30, 56, 76	0
3	D	211/215 (98%)	0.25	6 (2%)	53	25	22, 53, 67, 84	0
3	Q	210/215 (97%)	-0.16	1 (0%)	91	75	16, 27, 40, 48	0
All	All	1015/1156 (87%)	0.07	25 (2%)	57	29	16, 45, 98, 141	2 (0%)

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	105	SER	6.0
1	B	93	HIS	5.8
1	B	110	GLY	4.2
1	B	107	GLY	3.6
3	D	61	ASP	3.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 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.