



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 29, 2017 – 08:34 AM EDT

PDB ID : 5IZ8  
Title : Protein-protein interaction  
Authors : Zhao, Y.; Jiang, H.; Yang, X.; Jiang, F.; Song, K.; Zhang, J.  
Deposited on : 2016-03-25  
Resolution : 3.06 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

<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-20029077  
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-20029077

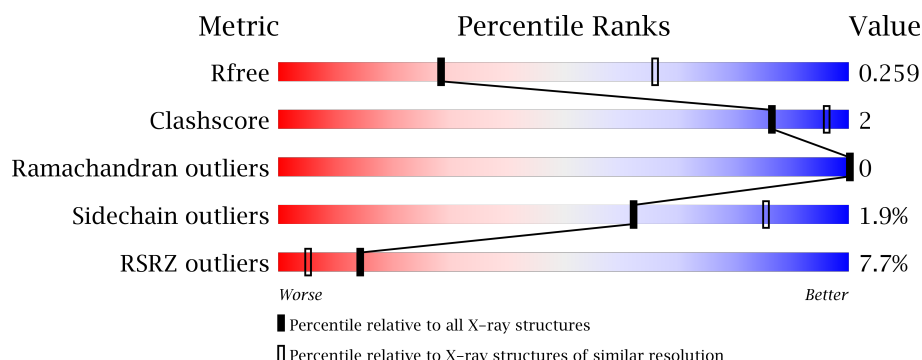
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.06 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	1348 (3.10-3.02)
Clashscore	112137	1462 (3.10-3.02)
Ramachandran outliers	110173	1410 (3.10-3.02)
Sidechain outliers	110143	1410 (3.10-3.02)
RSRZ outliers	101464	1355 (3.10-3.02)

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	354	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="width: 91%; height: 10px; background-color: green;"></div> <div style="width: 6%; height: 10px; background-color: yellow;"></div> <div style="width: 3%; height: 10px; background-color: grey;"></div> <div style="position: absolute; left: 0; top: -10px;">%</div> <div style="position: absolute; left: 45%; top: -10px;">91%</div> <div style="position: absolute; left: 95%; top: -10px;">6%</div> <div style="position: absolute; left: 98%; top: -10px;">.</div> </div>
1	B	354	<div> <div style="width: 14%; height: 10px; background-color: red;"></div> <div style="width: 86%; height: 10px; background-color: green;"></div> <div style="width: 9%; height: 10px; background-color: yellow;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> <div style="position: absolute; left: 14%; top: -10px;">14%</div> <div style="position: absolute; left: 55%; top: -10px;">86%</div> <div style="position: absolute; left: 95%; top: -10px;">9%</div> <div style="position: absolute; left: 98%; top: -10px;">.</div> </div>
2	C	9	<div> <div style="width: 100%; height: 10px; background-color: green;"></div> <div style="position: absolute; left: 50%; top: -10px;">100%</div> </div>
2	D	9	<div> <div style="width: 67%; height: 10px; background-color: green;"></div> <div style="width: 33%; height: 10px; background-color: yellow;"></div> <div style="position: absolute; left: 33%; top: -10px;">67%</div> <div style="position: absolute; left: 80%; top: -10px;">33%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5377 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 Adenomatous polyposis coli protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	343	Total	C	N	O	S	0	0	0
			2661	1654	482	497	28			
1	B	322	Total	C	N	O	S	0	0	0
			2504	1556	452	470	26			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	398	MET	-	expression tag	UNP P25054
A	399	GLY	-	expression tag	UNP P25054
A	400	HIS	-	expression tag	UNP P25054
A	401	HIS	-	expression tag	UNP P25054
A	402	HIS	-	expression tag	UNP P25054
A	403	HIS	-	expression tag	UNP P25054
A	404	HIS	-	expression tag	UNP P25054
A	405	HIS	-	expression tag	UNP P25054
A	406	MET	-	expression tag	UNP P25054
B	398	MET	-	expression tag	UNP P25054
B	399	GLY	-	expression tag	UNP P25054
B	400	HIS	-	expression tag	UNP P25054
B	401	HIS	-	expression tag	UNP P25054
B	402	HIS	-	expression tag	UNP P25054
B	403	HIS	-	expression tag	UNP P25054
B	404	HIS	-	expression tag	UNP P25054
B	405	HIS	-	expression tag	UNP P25054
B	406	MET	-	expression tag	UNP P25054

- Molecule 2 is a protein called ACE-ALA-GLY-GLU-ALA-LEU-ALA-ASP-NH2.

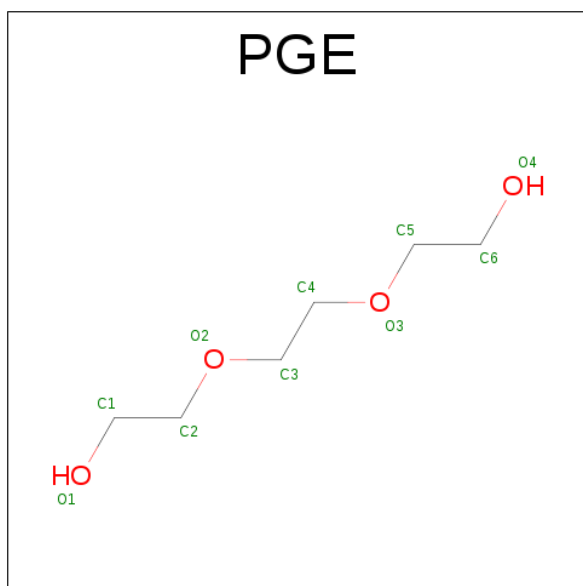
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	9	Total	C	N	O	0	0	1
			48	28	8	12			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	9	Total	C	N	O	0	0	1
			48	28	8	12			

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is water.

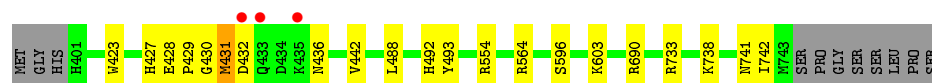
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	71	Total	O	0	0
			71	71		
4	B	33	Total	O	0	0
			33	33		
4	D	2	Total	O	0	0
			2	2		

### 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.

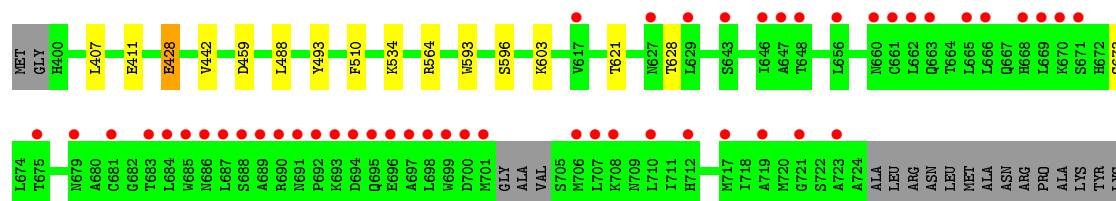
- Molecule 1: Adenomatous polyposis coli protein

Chain A: 



- Molecule 1: Adenomatous polyposis coli protein

Chain B: 



- Molecule 2: ACE-ALA-GLY-GLU-ALA-LEU-ALA-ASP-NH2

Chain C: 

There are no outlier residues recorded for this chain.

- Molecule 2: ACE-ALA-GLY-GLU-ALA-LEU-ALA-ASP-NH2

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.07Å 114.07Å 308.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	98.79 – 3.06 49.87 – 3.06	Depositor EDS
% Data completeness (in resolution range)	99.9 (98.79-3.06) 99.9 (49.87-3.06)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.55 (at 3.07Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.213 , 0.253 0.218 , 0.259	Depositor DCC
$R_{free}$ test set	1121 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	46.7	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 51.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	5377	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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: PGE, ACE, NH2

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	0.51	0/2705	0.74	4/3656 (0.1%)
1	B	0.50	0/2546	0.70	3/3441 (0.1%)
2	C	0.59	0/44	0.64	0/59
2	D	0.52	0/44	0.73	0/59
All	All	0.50	0/5339	0.72	7/7215 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	430	GLY	N-CA-C	-7.94	93.24	113.10
1	A	564	ARG	NE-CZ-NH1	7.51	124.06	120.30
1	B	564	ARG	NE-CZ-NH2	6.72	123.66	120.30
1	A	431	MET	N-CA-C	5.92	127.00	111.00
1	B	428	GLU	N-CA-C	-5.90	95.07	111.00
1	A	564	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	B	564	ARG	NE-CZ-NH1	-5.34	117.63	120.30

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	2661	0	2669	21	0
1	B	2504	0	2497	6	0
2	C	48	0	42	0	0
2	D	48	0	42	2	0
3	A	10	0	14	0	0
4	A	71	0	0	1	0
4	B	33	0	0	0	0
4	D	2	0	0	0	0
All	All	5377	0	5264	26	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 2.

All (26) 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:423:TRP:NE1	1:A:427:HIS:HD2	1.42	1.17
1:A:428:GLU:HG3	1:A:429:PRO:HD2	1.09	1.05
1:A:423:TRP:NE1	1:A:427:HIS:CD2	2.24	1.04
1:A:428:GLU:CG	1:A:429:PRO:HD2	1.92	0.99
1:A:428:GLU:HG3	1:A:429:PRO:CD	2.01	0.88
1:A:423:TRP:CE2	1:A:427:HIS:HD2	2.00	0.78
1:A:423:TRP:HE1	1:A:427:HIS:HD2	1.37	0.71
1:A:423:TRP:HE1	1:A:427:HIS:CD2	2.11	0.66
1:A:423:TRP:CD1	1:A:427:HIS:CD2	2.84	0.65
1:A:431:MET:HG2	1:A:432:ASP:N	2.15	0.62
1:A:428:GLU:CG	1:A:429:PRO:CD	2.70	0.60
1:A:428:GLU:O	1:A:431:MET:HB2	2.02	0.59
1:A:554:ARG:NH2	1:B:534:LYS:O	2.36	0.58
1:A:431:MET:HG2	1:A:432:ASP:O	2.06	0.56
1:B:510:PHE:CG	2:D:5:LEU:HD23	2.45	0.52
1:A:423:TRP:CE2	1:A:427:HIS:CD2	2.86	0.51
1:A:427:HIS:NE2	1:A:436:ASN:O	2.36	0.51
1:A:492:HIS:HB2	4:A:968:HOH:O	2.13	0.49
1:B:442:VAL:HA	1:B:493:TYR:CZ	2.47	0.48
1:B:593:TRP:CE3	2:D:1:ALA:HB2	2.50	0.45
1:A:431:MET:CG	1:A:432:ASP:N	2.79	0.45
1:A:596:SER:O	1:A:603:LYS:HE3	2.17	0.45
1:B:407:LEU:O	1:B:411:GLU:HG2	2.17	0.45
1:A:442:VAL:HA	1:A:493:TYR:CZ	2.53	0.43
1:A:733:ARG:HD2	1:A:738:LYS:HA	2.02	0.42
1:B:596:SER:O	1:B:603:LYS:HE3	2.20	0.42



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	341/354 (96%)	326 (96%)	15 (4%)	0	100	100
1	B	318/354 (90%)	305 (96%)	13 (4%)	0	100	100
2	C	7/9 (78%)	7 (100%)	0	0	100	100
2	D	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
All	All	673/726 (93%)	643 (96%)	30 (4%)	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	291/300 (97%)	287 (99%)	4 (1%)	71	89
1	B	276/300 (92%)	270 (98%)	6 (2%)	57	83
2	C	3/3 (100%)	3 (100%)	0	100	100
2	D	3/3 (100%)	2 (67%)	1 (33%)	0	0
All	All	573/606 (95%)	562 (98%)	11 (2%)	62	85

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

Mol	Chain	Res	Type
1	A	488	LEU
1	A	690	ARG
1	A	741	ASN
1	A	742	ILE
1	B	428	GLU
1	B	459	ASP
1	B	488	LEU
1	B	621	THR
1	B	628	THR
1	B	673	SER
2	D	7	ASP

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

Mol	Chain	Res	Type
1	A	427	HIS
1	A	659	ASN
1	A	741	ASN
1	B	433	GLN
1	B	659	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](#)

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](#)

1 ligand 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	$\# Z  > 2$
3	PGE	A	801	-	9,9,9	0.45	0	8,8,8	0.67	0

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
3	PGE	A	801	-	-	0/7/7/7	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

### 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<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	343/354 (96%)	-0.20	3 (0%) 84 66	16, 41, 83, 128	0
1	B	322/354 (90%)	0.54	49 (15%) 2 1	16, 67, 148, 168	0
2	C	7/9 (77%)	-0.42	0 100 100	21, 21, 26, 44	0
2	D	7/9 (77%)	-0.19	0 100 100	48, 59, 67, 78	0
All	All	679/726 (93%)	0.15	52 (7%) 14 5	16, 48, 136, 168	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	688	SER	6.3
1	B	698	LEU	5.7
1	B	687	LEU	5.5
1	B	707	LEU	5.2
1	B	694	ASP	5.2
1	B	666	LEU	5.1
1	B	662	LEU	5.0
1	B	701	MET	4.9
1	B	668	HIS	4.6
1	B	661	CYS	4.5
1	B	721	GLY	4.5
1	B	683	THR	4.4
1	B	665	LEU	4.2
1	B	691	ASN	4.1
1	B	643	SER	4.1
1	B	699	TRP	4.0
1	B	700	ASP	3.9
1	B	692	PRO	3.9
1	B	686	ASN	3.9
1	B	717	MET	3.6
1	B	710	LEU	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	617	VAL	3.3
1	B	629	LEU	3.1
1	B	684	LEU	3.1
1	B	663	GLN	3.1
1	B	647	ALA	3.0
1	B	708	LYS	2.9
1	B	690	ARG	2.9
1	B	681	CYS	2.8
1	B	695	GLN	2.7
1	B	646	ILE	2.7
1	A	433	GLN	2.7
1	B	679	ASN	2.7
1	B	685	TRP	2.6
1	B	689	ALA	2.6
1	B	671	SER	2.5
1	B	669	LEU	2.5
1	B	693	LYS	2.5
1	B	696	GLU	2.4
1	B	627	ASN	2.4
1	B	648	THR	2.4
1	A	435	LYS	2.3
1	B	656	LEU	2.3
1	B	697	ALA	2.3
1	B	719	ALA	2.2
1	B	723	ALA	2.2
1	B	675	THR	2.2
1	A	432	ASP	2.2
1	B	706	MET	2.1
1	B	670	LYS	2.1
1	B	712	HIS	2.1
1	B	660	ASN	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands [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( $\text{\AA}^2$ )	Q<0.9
3	PGE	A	801	10/10	0.89	0.22	1.67	63,65,67,67	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.