



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 18, 2016 – 09:04 PM EDT

PDB ID : 5HJP  
Title : Identification of LXRBeta selective agonists for the treatment of Alzheimer's Disease  
Authors : Parthasarathy, G.; Klein, D.  
Deposited on : 2016-01-13  
Resolution : 2.60 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20027257

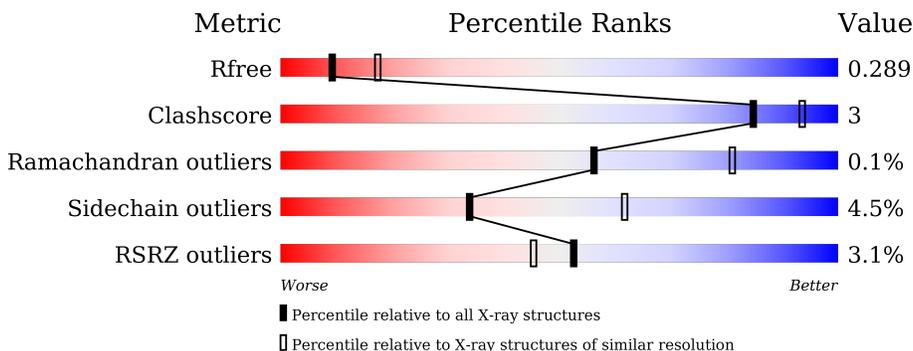
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	254	 6% 74% 10% 15%
1	C	254	 4% 78% 6% • 15%
2	B	264	 81% 16% ••
2	D	264	 2% 87% 8% 5%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7679 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 Retinoic acid receptor RXR-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	215	1715	1101	300	304	10	0	0	0
1	C	216	1722	1106	301	305	10	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	534	GLY	-	expression tag	UNP P28702
A	535	SER	-	expression tag	UNP P28702
A	536	GLY	-	expression tag	UNP P28702
A	537	SER	-	expression tag	UNP P28702
A	538	GLY	-	expression tag	UNP P28702
A	539	SER	-	expression tag	UNP P28702
A	540	HIS	-	expression tag	UNP P28702
A	541	LYS	-	expression tag	UNP P28702
A	542	ILE	-	expression tag	UNP P28702
A	543	LEU	-	expression tag	UNP P28702
A	544	HIS	-	expression tag	UNP P28702
A	545	ARG	-	expression tag	UNP P28702
A	546	LEU	-	expression tag	UNP P28702
A	547	LEU	-	expression tag	UNP P28702
A	548	GLN	-	expression tag	UNP P28702
A	549	ASP	-	expression tag	UNP P28702
A	550	SER	-	expression tag	UNP P28702
A	551	SER	-	expression tag	UNP P28702
A	552	SER	-	expression tag	UNP P28702
C	534	GLY	-	expression tag	UNP P28702
C	535	SER	-	expression tag	UNP P28702
C	536	GLY	-	expression tag	UNP P28702
C	537	SER	-	expression tag	UNP P28702
C	538	GLY	-	expression tag	UNP P28702
C	539	SER	-	expression tag	UNP P28702

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	540	HIS	-	expression tag	UNP P28702
C	541	LYS	-	expression tag	UNP P28702
C	542	ILE	-	expression tag	UNP P28702
C	543	LEU	-	expression tag	UNP P28702
C	544	HIS	-	expression tag	UNP P28702
C	545	ARG	-	expression tag	UNP P28702
C	546	LEU	-	expression tag	UNP P28702
C	547	LEU	-	expression tag	UNP P28702
C	548	GLN	-	expression tag	UNP P28702
C	549	ASP	-	expression tag	UNP P28702
C	550	SER	-	expression tag	UNP P28702
C	551	SER	-	expression tag	UNP P28702
C	552	SER	-	expression tag	UNP P28702

- Molecule 2 is a protein called Oxysterols receptor LXR-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	259	Total 2079	C 1323	N 367	O 382	S 7	0	0	0
2	D	251	Total 2033	C 1296	N 359	O 371	S 7	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

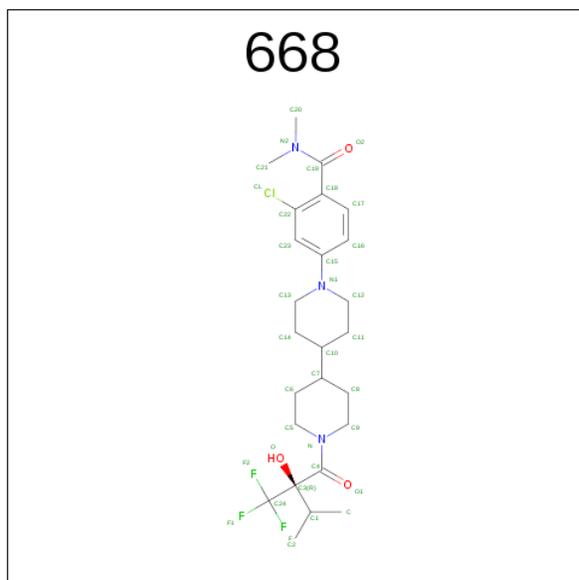
Chain	Residue	Modelled	Actual	Comment	Reference
B	259	ALA	GLN	engineered mutation	UNP P55055
B	261	GLY	ARG	engineered mutation	UNP P55055
B	262	SER	ASP	engineered mutation	UNP P55055
B	264	SER	ARG	engineered mutation	UNP P55055
B	462	GLY	-	expression tag	UNP P55055
B	463	SER	-	expression tag	UNP P55055
B	464	GLY	-	expression tag	UNP P55055
B	465	SER	-	expression tag	UNP P55055
B	466	GLY	-	expression tag	UNP P55055
B	467	SER	-	expression tag	UNP P55055
B	468	HIS	-	expression tag	UNP P55055
B	469	LYS	-	expression tag	UNP P55055
B	470	ILE	-	expression tag	UNP P55055
B	471	LEU	-	expression tag	UNP P55055
B	472	HIS	-	expression tag	UNP P55055
B	473	ARG	-	expression tag	UNP P55055
B	474	LEU	-	expression tag	UNP P55055

*Continued on next page...*

*Continued from previous page...*

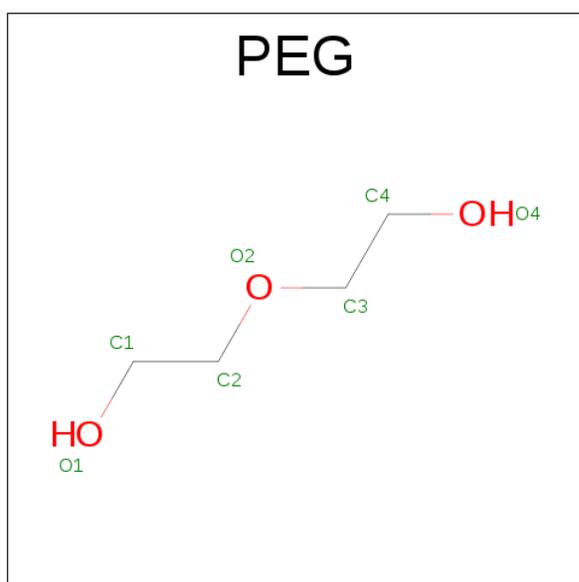
Chain	Residue	Modelled	Actual	Comment	Reference
B	475	LEU	-	expression tag	UNP P55055
B	476	GLN	-	expression tag	UNP P55055
B	477	ASP	-	expression tag	UNP P55055
B	478	SER	-	expression tag	UNP P55055
B	479	SER	-	expression tag	UNP P55055
B	480	SER	-	expression tag	UNP P55055
D	259	ALA	GLN	engineered mutation	UNP P55055
D	261	GLY	ARG	engineered mutation	UNP P55055
D	262	SER	ASP	engineered mutation	UNP P55055
D	264	SER	ARG	engineered mutation	UNP P55055
D	462	GLY	-	expression tag	UNP P55055
D	463	SER	-	expression tag	UNP P55055
D	464	GLY	-	expression tag	UNP P55055
D	465	SER	-	expression tag	UNP P55055
D	466	GLY	-	expression tag	UNP P55055
D	467	SER	-	expression tag	UNP P55055
D	468	HIS	-	expression tag	UNP P55055
D	469	LYS	-	expression tag	UNP P55055
D	470	ILE	-	expression tag	UNP P55055
D	471	LEU	-	expression tag	UNP P55055
D	472	HIS	-	expression tag	UNP P55055
D	473	ARG	-	expression tag	UNP P55055
D	474	LEU	-	expression tag	UNP P55055
D	475	LEU	-	expression tag	UNP P55055
D	476	GLN	-	expression tag	UNP P55055
D	477	ASP	-	expression tag	UNP P55055
D	478	SER	-	expression tag	UNP P55055
D	479	SER	-	expression tag	UNP P55055
D	480	SER	-	expression tag	UNP P55055

- Molecule 3 is 2-chloro-4-{1'-[(2R)-2-hydroxy-3-methyl-2-(trifluoromethyl)butanoyl]-4,4'-biperidin-1-yl}-N,N-dimethylbenzamide (three-letter code: 668) (formula: C<sub>25</sub>H<sub>35</sub>ClF<sub>3</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	Cl	F	N			O
3	B	1	Total	35	25	1	3	3	0	0
3	D	1	Total	35	25	1	3	3	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	O			
4	B	1	Total	7	4	3	0	0
4	C	1	Total	7	4	3	0	0

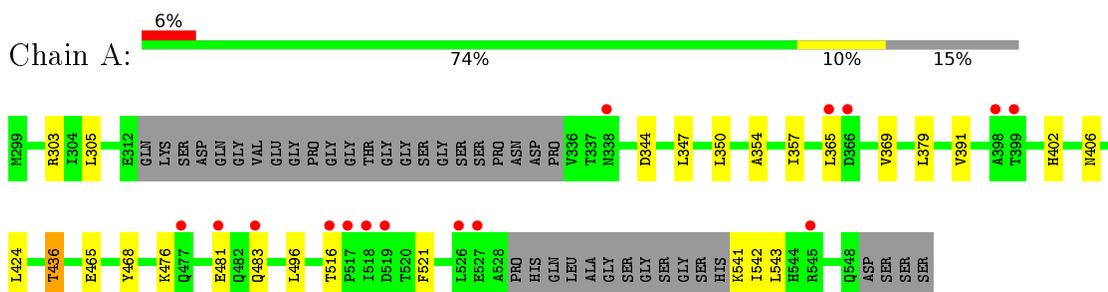
- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
5	A	4	Total O 4 4	0	0
5	B	15	Total O 15 15	0	0
5	C	11	Total O 11 11	0	0
5	D	16	Total O 16 16	0	0

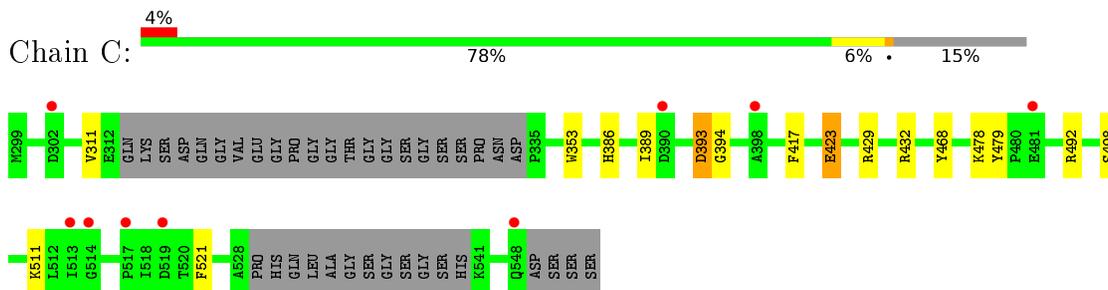
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

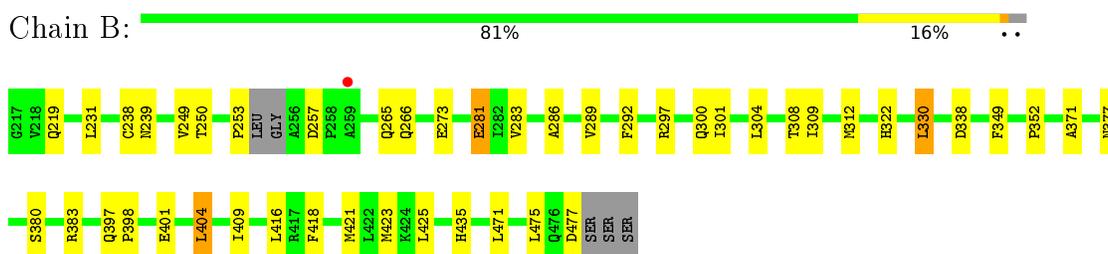
- Molecule 1: Retinoic acid receptor RXR-beta



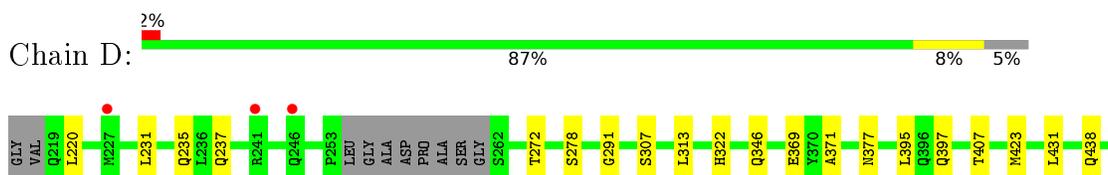
- Molecule 1: Retinoic acid receptor RXR-beta

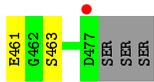


- Molecule 2: Oxysterols receptor LXR-beta



- Molecule 2: Oxysterols receptor LXR-beta





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.77Å 106.03Å 139.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.70 – 2.60 35.70 – 2.60	Depositor EDS
% Data completeness (in resolution range)	96.4 (35.70-2.60) 96.4 (35.70-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.43 (at 2.61Å)	Xtrriage
Refinement program	BUSTER-TNT	Depositor
R, $R_{free}$	0.236 , 0.283 0.236 , 0.289	Depositor DCC
$R_{free}$ test set	1525 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.7	Xtrriage
Anisotropy	0.080	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 33.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7679	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PEG, 668

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	0.45	0/1744	0.64	0/2351
1	C	0.49	0/1752	0.66	0/2362
2	B	0.49	0/2120	0.68	0/2866
2	D	0.49	0/2073	0.65	0/2801
All	All	0.48	0/7689	0.66	0/10380

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 [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	1715	0	1775	11	0
1	C	1722	0	1783	8	0
2	B	2079	0	2095	21	0
2	D	2033	0	2054	7	0
3	B	35	0	0	1	0
3	D	35	0	0	0	0
4	B	7	0	10	0	0
4	C	7	0	10	0	0
5	A	4	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	15	0	0	0	0
5	C	11	0	0	0	0
5	D	16	0	0	0	0
All	All	7679	0	7727	44	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 3.

The worst 5 of 44 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:476:LYS:HE2	1:A:483:GLN:HE22	1.57	0.69
1:C:423:GLU:HG3	1:C:492:ARG:HH12	1.59	0.67
2:B:250:THR:HG23	2:B:273:GLU:OE1	1.95	0.66
2:B:292:PHE:CZ	2:B:300:GLN:HB2	2.33	0.63
1:A:468:TYR:OH	2:B:423:MET:HG3	2.00	0.61

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	209/254 (82%)	204 (98%)	5 (2%)	0	100	100
1	C	210/254 (83%)	202 (96%)	7 (3%)	1 (0%)	34	60
2	B	255/264 (97%)	246 (96%)	9 (4%)	0	100	100
2	D	247/264 (94%)	243 (98%)	4 (2%)	0	100	100
All	All	921/1036 (89%)	895 (97%)	25 (3%)	1 (0%)	56	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	393	ASP

### 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	188/216 (87%)	179 (95%)	9 (5%)	31	58
1	C	189/216 (88%)	183 (97%)	6 (3%)	46	74
2	B	227/231 (98%)	214 (94%)	13 (6%)	25	49
2	D	223/231 (96%)	214 (96%)	9 (4%)	38	67
All	All	827/894 (92%)	790 (96%)	37 (4%)	34	62

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	338	ASP
2	B	421	MET
2	D	397	GLN
2	B	397	GLN
2	B	404	LEU

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	483	GLN
2	B	239	ASN
2	B	265	GLN
2	B	377	ASN
2	B	397	GLN

### 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](#)

4 ligands 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
3	668	B	501	-	36,37,37	1.12	3 (8%)	49,56,56	1.25	5 (10%)
4	PEG	B	502	-	6,6,6	0.29	0	5,5,5	0.24	0
4	PEG	C	601	-	6,6,6	0.33	0	5,5,5	0.20	0
3	668	D	501	-	36,37,37	1.06	3 (8%)	49,56,56	1.00	4 (8%)

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	668	B	501	-	-	0/40/61/61	0/3/3/3
4	PEG	B	502	-	-	0/4/4/4	0/0/0/0
4	PEG	C	601	-	-	0/4/4/4	0/0/0/0
3	668	D	501	-	-	0/40/61/61	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	501	668	C12-N1	3.23	1.51	1.46
3	D	501	668	C13-N1	3.36	1.51	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	501	668	C15-N1	3.51	1.48	1.38
3	B	501	668	C15-N1	3.51	1.48	1.38
3	B	501	668	C12-N1	3.76	1.52	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	501	668	O1-C4-C3	-2.82	110.45	118.50
3	B	501	668	O1-C4-C3	-2.57	111.17	118.50
3	B	501	668	C23-C15-N1	-2.13	118.88	121.35
3	D	501	668	C23-C15-N1	-2.11	118.90	121.35
3	D	501	668	C11-C12-N1	-2.05	107.04	110.88

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
3	B	501	668	1	0

## 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	215/254 (84%)	0.53	15 (6%) 19 13	35, 61, 86, 92	0
1	C	216/254 (85%)	0.20	9 (4%) 40 32	25, 44, 65, 90	0
2	B	259/264 (98%)	-0.03	1 (0%) 93 91	27, 39, 56, 68	0
2	D	251/264 (95%)	-0.03	4 (1%) 74 69	22, 37, 55, 72	0
All	All	941/1036 (90%)	0.15	29 (3%) 52 45	22, 42, 72, 92	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	514	GLY	7.8
1	A	399	THR	4.3
2	D	477	ASP	4.2
1	C	390	ASP	3.7
1	A	477	GLN	3.7

### 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](#)

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	668	B	501	35/35	0.82	0.24	1.70	38,47,60,63	0
3	668	D	501	35/35	0.90	0.18	0.82	26,42,54,62	0
4	PEG	C	601	7/7	0.91	0.16	-0.17	37,39,43,45	0
4	PEG	B	502	7/7	0.84	0.17	-	49,53,55,55	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.