



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:03 PM GMT

PDB ID : 4QT8  
Title : Crystal Structure of RON Sema-PSI-IPT1 extracellular domains in complex with MSP beta-chain  
Authors : Herzberg, O.; Chao, K.L.  
Deposited on : 2014-07-07  
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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

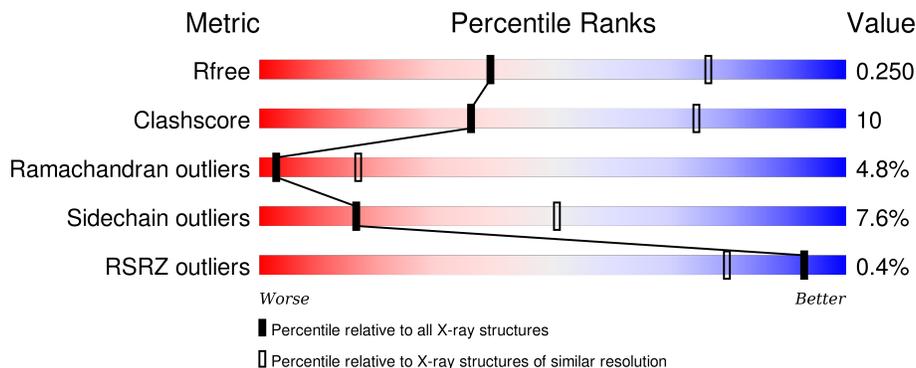
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (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. 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	669	
1	B	669	
2	C	253	
2	D	253	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12617 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 Macrophage-stimulating protein receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	629	4544	2871	794	847	32	0	0	0
1	A	628	4565	2886	798	849	32	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	ARG	-	EXPRESSION TAG	UNP Q04912
A	24	SER	-	EXPRESSION TAG	UNP Q04912
A	209	GLY	ALA	CONFLICT	UNP Q04912
A	306	LEU	ARG	ENGINEERED MUTATION	UNP Q04912
A	307	VAL	ARG	ENGINEERED MUTATION	UNP Q04912
A	308	PRO	ARG	ENGINEERED MUTATION	UNP Q04912
A	311	SER	ALA	ENGINEERED MUTATION	UNP Q04912
A	322	GLN	ARG	ENGINEERED MUTATION	UNP Q04912
A	684	THR	-	EXPRESSION TAG	UNP Q04912
A	685	GLY	-	EXPRESSION TAG	UNP Q04912
A	686	HIS	-	EXPRESSION TAG	UNP Q04912
A	687	HIS	-	EXPRESSION TAG	UNP Q04912
A	688	HIS	-	EXPRESSION TAG	UNP Q04912
A	689	HIS	-	EXPRESSION TAG	UNP Q04912
A	690	HIS	-	EXPRESSION TAG	UNP Q04912
A	691	HIS	-	EXPRESSION TAG	UNP Q04912
B	23	ARG	-	EXPRESSION TAG	UNP Q04912
B	24	SER	-	EXPRESSION TAG	UNP Q04912
B	209	GLY	ALA	CONFLICT	UNP Q04912
B	306	LEU	ARG	ENGINEERED MUTATION	UNP Q04912
B	307	VAL	ARG	ENGINEERED MUTATION	UNP Q04912
B	308	PRO	ARG	ENGINEERED MUTATION	UNP Q04912
B	311	SER	ALA	ENGINEERED MUTATION	UNP Q04912
B	322	GLN	ARG	ENGINEERED MUTATION	UNP Q04912
B	684	THR	-	EXPRESSION TAG	UNP Q04912

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	685	GLY	-	EXPRESSION TAG	UNP Q04912
B	686	HIS	-	EXPRESSION TAG	UNP Q04912
B	687	HIS	-	EXPRESSION TAG	UNP Q04912
B	688	HIS	-	EXPRESSION TAG	UNP Q04912
B	689	HIS	-	EXPRESSION TAG	UNP Q04912
B	690	HIS	-	EXPRESSION TAG	UNP Q04912
B	691	HIS	-	EXPRESSION TAG	UNP Q04912

- Molecule 2 is a protein called Hepatocyte growth factor-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	231	Total	C	N	O	S	0	0	0
			1725	1094	313	302	16			
2	D	231	Total	C	N	O	S	0	0	0
			1716	1089	311	300	16			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	672	SER	CYS	ENGINEERED MUTATION	UNP P26927
D	712	HIS	-	EXPRESSION TAG	UNP P26927
D	713	HIS	-	EXPRESSION TAG	UNP P26927
D	714	HIS	-	EXPRESSION TAG	UNP P26927
D	715	HIS	-	EXPRESSION TAG	UNP P26927
D	716	HIS	-	EXPRESSION TAG	UNP P26927
D	717	HIS	-	EXPRESSION TAG	UNP P26927
C	672	SER	CYS	ENGINEERED MUTATION	UNP P26927
C	712	HIS	-	EXPRESSION TAG	UNP P26927
C	713	HIS	-	EXPRESSION TAG	UNP P26927
C	714	HIS	-	EXPRESSION TAG	UNP P26927
C	715	HIS	-	EXPRESSION TAG	UNP P26927
C	716	HIS	-	EXPRESSION TAG	UNP P26927
C	717	HIS	-	EXPRESSION TAG	UNP P26927

- Molecule 3 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	2	Total	C	N	O	0	0
			28	16	2	10		

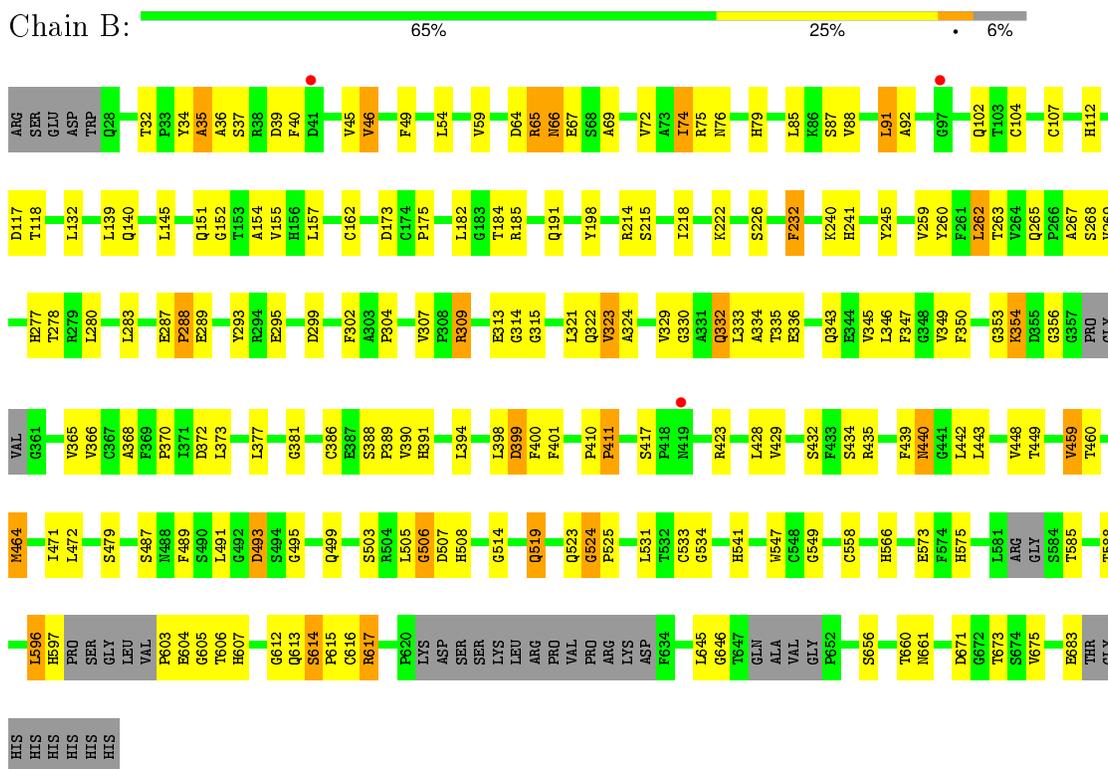
- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	B	3	39	22	2	15	0	0

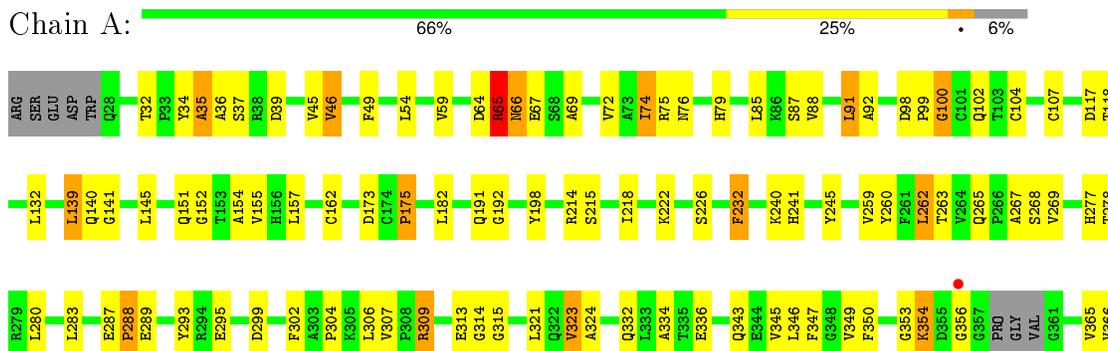
### 3 Residue-property plots [i](#)

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: Macrophage-stimulating protein receptor



- Molecule 1: Macrophage-stimulating protein receptor





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.58Å 63.77Å 146.05Å 90.00° 90.08° 90.00°	Depositor
Resolution (Å)	19.98 – 3.00 19.98 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.4 (19.98-3.00) 98.2 (19.98-3.00)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 2.98Å)	Xtriage
Refinement program	REFMAC 5.8.0071	Depositor
R, $R_{free}$	0.234 , 0.291 0.210 , 0.250	Depositor DCC
$R_{free}$ test set	2006 reflections (5.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.8	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.22 , -10.3	EDS
Estimated twinning fraction	0.577 for H, K, L 0.423 for -H, -K, L 0.209 for h,-k,-l	Xtriage
Reported twinning fraction	0.577 for H, K, L 0.423 for -H, -K, L	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 39246 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12617	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% 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: BMA, NAG

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	0.49	0/4685	0.68	0/6404
1	B	0.45	0/4664	0.68	0/6380
2	C	0.44	0/1766	0.65	0/2408
2	D	0.43	0/1757	0.63	0/2397
All	All	0.46	0/12872	0.67	0/17589

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	4565	0	4291	88	0
1	B	4544	0	4230	98	0
2	C	1725	0	1661	29	0
2	D	1716	0	1643	25	0
3	A	28	0	25	0	0
4	B	39	0	34	2	0
All	All	12617	0	11884	236	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 10.

The worst 5 of 236 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:B:645:LEU:HG	1:B:646:GLY:N	2.01	0.75
1:B:269:VAL:HG21	1:B:399:ASP:HB2	1.69	0.74
1:A:269:VAL:HG21	1:A:399:ASP:HB2	1.71	0.73
1:A:198:TYR:CD1	1:A:259:VAL:HG21	2.29	0.67
2:C:483:ARG:NH1	2:C:485:VAL:O	2.28	0.67

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	616/669 (92%)	487 (79%)	97 (16%)	32 (5%)	2	15
1	B	617/669 (92%)	488 (79%)	99 (16%)	30 (5%)	3	16
2	C	225/253 (89%)	182 (81%)	34 (15%)	9 (4%)	4	21
2	D	225/253 (89%)	184 (82%)	32 (14%)	9 (4%)	4	21
All	All	1683/1844 (91%)	1341 (80%)	262 (16%)	80 (5%)	3	17

5 of 80 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	35	ALA
1	B	46	VAL
1	B	66	ASN
1	B	232	PHE
1	B	356	GLY

### 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	479/560 (86%)	439 (92%)	40 (8%)	14	45
1	B	471/560 (84%)	433 (92%)	38 (8%)	15	47
2	C	181/218 (83%)	170 (94%)	11 (6%)	23	61
2	D	178/218 (82%)	168 (94%)	10 (6%)	26	65
All	All	1309/1556 (84%)	1210 (92%)	99 (8%)	16	51

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	GLN
1	A	309	ARG
2	D	507	CYS
1	A	117	ASP
1	A	182	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	325	HIS
1	A	332	GLN
1	A	597	HIS
1	A	254	HIS
1	A	564	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

5 carbohydrates 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	NAG	A	701	1,3	14,14,15	0.61	0	15,19,21	1.55	5 (33%)
3	NAG	A	702	3	14,14,15	1.01	1 (7%)	15,19,21	2.01	5 (33%)
4	NAG	B	701	1,4	14,14,15	0.44	0	15,19,21	3.39	5 (33%)
4	NAG	B	702	4	14,14,15	0.60	0	15,19,21	1.04	1 (6%)
4	BMA	B	703	4	11,11,12	0.59	0	14,15,17	1.62	2 (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
3	NAG	A	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	702	3	-	0/6/23/26	0/1/1/1
4	NAG	B	701	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	702	4	-	0/6/23/26	0/1/1/1
4	BMA	B	703	4	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	702	NAG	C1-C2	3.34	1.57	1.52

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	702	NAG	C3-C4-C5	-4.20	102.87	110.20
4	B	703	BMA	O5-C1-C2	-3.50	105.17	110.86
3	A	702	NAG	C4-C3-C2	-2.87	106.77	111.23
4	B	701	NAG	O3-C3-C2	-2.82	103.52	109.11

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	701	NAG	O7-C7-C8	-2.49	117.49	122.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	701	NAG	2	0

## 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	628/669 (93%)	-0.71	1 (0%) 95 87	8, 46, 89, 126	0
1	B	629/669 (94%)	-0.65	3 (0%) 91 76	22, 53, 92, 117	0
2	C	231/253 (91%)	-0.67	2 (0%) 85 64	24, 53, 95, 134	3 (1%)
2	D	231/253 (91%)	-0.59	1 (0%) 93 80	26, 54, 102, 134	3 (1%)
All	All	1719/1844 (93%)	-0.67	7 (0%) 93 80	8, 51, 93, 134	6 (0%)

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	41	ASP	2.6
2	C	473	ASP	2.5
2	C	582	GLN	2.4
1	A	356	GLY	2.3
2	D	540	THR	2.1

### 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](#)

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
4	NAG	B	701	14/15	0.95	0.14	-0.59	38,42,46,53	0
3	NAG	A	701	14/15	0.96	0.10	-1.08	23,29,33,33	0
3	NAG	A	702	14/15	0.95	0.20	-	34,38,41,42	0
4	BMA	B	703	11/12	0.93	0.13	-	53,59,61,61	0
4	NAG	B	702	14/15	0.94	0.17	-	58,64,78,83	0

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.