



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 12, 2021 – 03:24 PM EDT

PDB ID : 1YMM  
Title : TCR/HLA-DR2b/MBP-peptide complex  
Authors : Hahn, M.; Nicholson, M.J.; Pyrdol, J.; Wucherpennig, K.W.  
Deposited on : 2005-01-21  
Resolution : 3.50 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
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) : 2.23.2

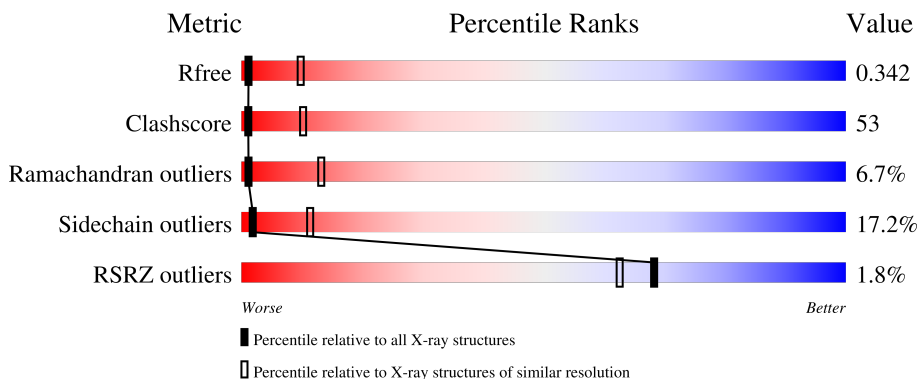
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 of 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	191	
2	B	198	
3	C	23	
4	D	207	
5	E	249	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
6	NAG	A	192	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 5789 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	182	1494	968	243	278	5	79	0	0

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	178	1466	930	257	273	6	68	0	0

- Molecule 3 is a protein called MBP peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	14	116	78	19	19	0	0	0

- Molecule 4 is a protein called T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	96	750	461	137	149	3	133	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	92	ASP	ALA	see remark 999	GB 29293741
D	93	THR	ASN	see remark 999	GB 29293741
D	94	THR	ALA	see remark 999	GB 29293741
D	95	SER	GLY	see remark 999	GB 29293741
D	?	-	SER	see remark 999	GB 29293741
D	?	-	GLY	see remark 999	GB 29293741
D	100	TYR	LEU	see remark 999	GB 29293741
D	101	ILE	THR	see remark 999	GB 29293741

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	104	THR	GLN	see remark 999	GB 29293741
D	107	ARG	ILE	see remark 999	GB 29293741
D	109	LYS	THR	see remark 999	GB 29293741
D	111	LEU	HIS	see remark 999	GB 29293741
D	112	ALA	PRO	see remark 999	GB 29293741

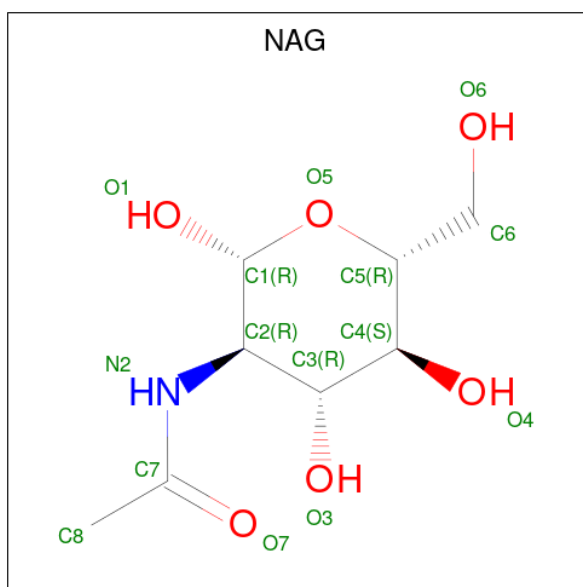
- Molecule 5 is a protein called T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	246	1948	1229	335	377	7	395	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	13	SER	CYS	engineered mutation	UNP P01850
E	98	ASP	GLU	see remark 999	UNP P01850
E	99	LEU	SER	see remark 999	UNP P01850
E	102	GLY	ASP	see remark 999	UNP P01850
E	103	ALA	PRO	see remark 999	UNP P01850
E	104	ASN	LYS	see remark 999	UNP P01850
E	193	SER	CYS	engineered mutation	UNP P01850

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



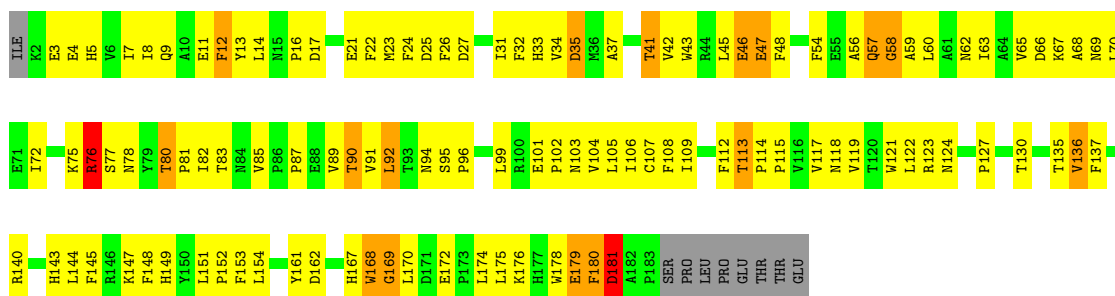
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	A	1	15	8	1	6	1	0

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

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

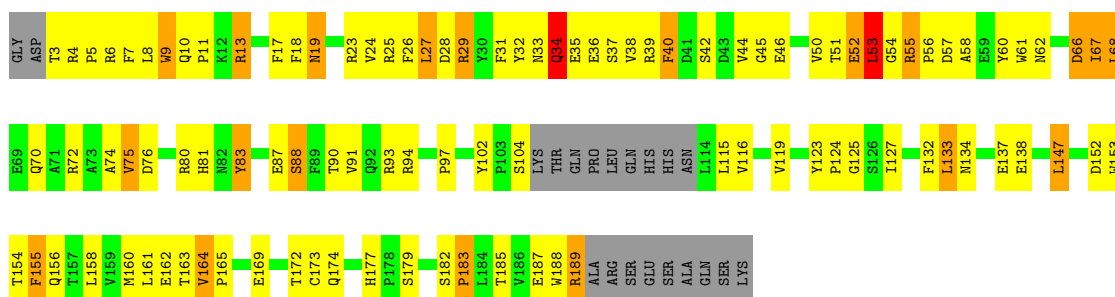
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain A: 

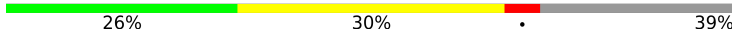


- Molecule 2: HLA class II histocompatibility antigen, DR beta chain

Chain B: 




- Molecule 3: MBP peptide

Chain C: 



- Molecule 4: T cell receptor alpha chain

Chain D: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	137.34Å 212.62Å 278.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.72 – 3.50 19.72 – 3.20	Depositor EDS
% Data completeness (in resolution range)	96.4 (19.72-3.50) 95.9 (19.72-3.20)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 3.22Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.274 , 0.318 0.292 , 0.342	Depositor DCC
$R_{free}$ test set	1634 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.5	Xtrriage
Anisotropy	0.186	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 89.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	5789	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0/1540	0.76	0/2100
2	B	0.49	0/1505	0.78	2/2044 (0.1%)
3	C	0.94	1/120 (0.8%)	0.85	0/164
4	D	0.95	6/760 (0.8%)	1.25	10/1024 (1.0%)
5	E	0.76	7/2000 (0.3%)	1.03	21/2722 (0.8%)
All	All	0.67	14/5925 (0.2%)	0.94	33/8054 (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
3	C	0	1
4	D	0	1
5	E	0	1
All	All	0	3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	86	SER	CA-CB	-9.63	1.38	1.52
4	D	59	GLY	CA-C	-9.43	1.36	1.51
4	D	62	ARG	CA-CB	9.08	1.74	1.53
5	E	55	LYS	CA-C	7.15	1.71	1.52
5	E	202	ALA	CA-CB	6.97	1.67	1.52

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	74	LEU	CA-C-N	-11.35	92.23	117.20
5	E	74	LEU	O-C-N	11.17	140.57	122.70
5	E	74	LEU	C-N-CA	10.26	147.34	121.70
5	E	54	SER	O-C-N	-9.54	107.44	122.70
4	D	61	LEU	O-C-N	-9.53	107.46	122.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	86	ASN	Mainchain
4	D	87	TYR	Sidechain
5	E	74	LEU	Mainchain

## 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	1494	0	1430	142	0
2	B	1466	0	1394	137	1
3	C	116	0	114	21	0
4	D	750	0	739	82	3
5	E	1948	0	1866	191	11
6	A	15	0	14	1	0
All	All	5789	0	5557	513	15

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 53.

The worst 5 of 513 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)
4:D:13:ILE:HB	4:D:79:SER:HB2	1.16	1.16
5:E:31:THR:HA	5:E:74:LEU:HD12	1.26	1.13
5:E:16:GLY:H	5:E:84:ALA:HB3	1.19	1.07
5:E:25:SER:HB2	5:E:28:PHE:CE1	1.89	1.06
4:D:56:LYS:HB3	4:D:63:VAL:H	1.19	1.03

The worst 5 of 15 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)
5:E:168:LYS:C	5:E:169:GLU:OE1[4_555]	0.53	1.67
5:E:168:LYS:CA	5:E:169:GLU:OE1[4_555]	1.16	1.04
5:E:169:GLU:N	5:E:169:GLU:OE1[4_555]	1.37	0.83
4:D:15:GLU:OE1	4:D:77:THR:O[3_455]	1.38	0.82
5:E:223:GLU:OE1	5:E:232:ALA:CB[4_555]	1.42	0.78

## 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	180/191 (94%)	147 (82%)	23 (13%)	10 (6%)	<b>2</b> <b>17</b>
2	B	174/198 (88%)	143 (82%)	24 (14%)	7 (4%)	<b>3</b> <b>24</b>
3	C	12/23 (52%)	11 (92%)	0	1 (8%)	<b>1</b> <b>9</b>
4	D	94/207 (45%)	67 (71%)	17 (18%)	10 (11%)	<b>0</b> <b>7</b>
5	E	244/249 (98%)	171 (70%)	54 (22%)	19 (8%)	<b>1</b> <b>10</b>
All	All	704/868 (81%)	539 (77%)	118 (17%)	47 (7%)	<b>1</b> <b>13</b>

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	180	PHE
4	D	44	VAL
4	D	46	LEU
5	E	7	HIS
5	E	54	SER

### 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	166/175 (95%)	148 (89%)	18 (11%)	6	30
2	B	161/178 (90%)	138 (86%)	23 (14%)	3	19
3	C	14/16 (88%)	14 (100%)	0	100	100
4	D	83/185 (45%)	58 (70%)	25 (30%)	0	2
5	E	217/219 (99%)	173 (80%)	44 (20%)	1	6
All	All	641/773 (83%)	531 (83%)	110 (17%)	2	11

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	60	ARG
5	E	28	PHE
5	E	237	GLN
5	E	205	TRP
4	D	70	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	52	ASN
5	E	38	GLN
5	E	215	GLN
5	E	211	HIS
2	B	134	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 monosaccharides 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
6	NAG	A	192	1	15,15,15	0.48	0	21,21,21	0.81	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
6	NAG	A	192	1	-	2/6/26/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	192	NAG	C4-C5-C6-O6
6	A	192	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	192	NAG	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	182/191 (95%)	-0.81	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	12, 44, 82, 164	22 (12%)
2	B	178/198 (89%)	-0.77	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	8, 46, 84, 104	19 (10%)
3	C	14/23 (60%)	-0.64	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	18, 63, 93, 102	0
4	D	96/207 (46%)	-0.29	3 (3%) <span style="border: 1px solid gray; padding: 2px;">49</span> <span style="border: 1px solid gray; padding: 2px;">43</span>	36, 112, 159, 174	44 (45%)
5	E	246/249 (98%)	0.05	10 (4%) <span style="border: 1px solid gray; padding: 2px;">37</span> <span style="border: 1px solid gray; padding: 2px;">33</span>	30, 117, 155, 180	120 (48%)
All	All	716/868 (82%)	-0.43	13 (1%) <span style="border: 1px solid blue; padding: 2px;">68</span> <span style="border: 1px solid blue; padding: 2px;">62</span>	8, 70, 150, 180	205 (28%)

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	9	GLN	3.7
5	E	156	PRO	3.1
5	E	168	LYS	3.1
4	D	18	ASN	3.0
5	E	167	GLY	2.9

### 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 monosaccharides 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	A	192	15/15	0.73	0.43	1,154,168,176	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.