



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 29, 2017 – 02:13 AM EDT

PDB ID : 4P2O
Title : Crystal structure of the 2B4 TCR in complex with 2A/I-Ek
Authors : Birnbaum, M.E.; Ozkan, E.; Garcia, K.C.
Deposited on : unknown
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

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-20030345
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-20030345

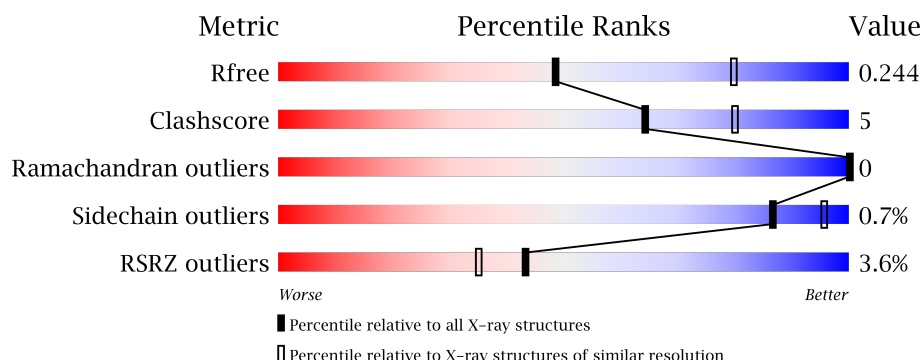
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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}	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (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 ≥ 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	204	<div> <div>0%</div> <div> <div></div> <div>76%</div> <div>10%</div> <div>14%</div> </div> </div>
2	B	236	<div> <div>6%</div> <div> <div></div> <div>63%</div> <div>11%</div> <div>25%</div> </div> </div>
3	C	220	<div> <div>2%</div> <div> <div></div> <div>83%</div> <div>8%</div> <div>9%</div> </div> </div>
4	D	255	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>13%</div> <div>•</div> </div> </div>
5	P	22	<div> <div>5%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>9%</div> </div> </div>

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	FUC	B	903	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 6701 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 H-2 class II histocompatibility antigen, E-K alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	175	Total	C	N	O	S	0	0	0
			1431	928	236	263	4			

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	ALA	-	expression tag	UNP P04224
A	-1	ASP	-	expression tag	UNP P04224
A	0	PRO	-	expression tag	UNP P04224
A	192	SER	-	expression tag	UNP P04224
A	193	ARG	-	expression tag	UNP P04224
A	194	GLY	-	expression tag	UNP P04224
A	195	GLY	-	expression tag	UNP P04224
A	196	LEU	-	expression tag	UNP P04224
A	197	GLU	-	expression tag	UNP P04224
A	198	VAL	-	expression tag	UNP P04224
A	199	LEU	-	expression tag	UNP P04224
A	200	PHE	-	expression tag	UNP P04224
A	201	GLN	-	expression tag	UNP P04224

- Molecule 2 is a protein called MHC class II E-beta-k.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1450	926	247	271	6			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-27	ALA	-	expression tag	UNP Q31163
B	-26	ASP	-	expression tag	UNP Q31163
B	-25	PRO	-	expression tag	UNP Q31163

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-24	ALA	-	expression tag	UNP Q31163
B	-23	ASP	-	expression tag	UNP Q31163
B	-22	PRO	-	expression tag	UNP Q31163
B	-21	LEU	-	expression tag	UNP Q31163
B	-20	ALA	-	expression tag	UNP Q31163
B	-19	PHE	-	expression tag	UNP Q31163
B	-18	PHE	-	expression tag	UNP Q31163
B	-17	SER	-	expression tag	UNP Q31163
B	-16	SER	-	expression tag	UNP Q31163
B	-15	ALA	-	expression tag	UNP Q31163
B	-14	ILE	-	expression tag	UNP Q31163
B	-13	LYS	-	expression tag	UNP Q31163
B	-12	GLY	-	expression tag	UNP Q31163
B	-11	GLY	-	expression tag	UNP Q31163
B	-10	GLY	-	expression tag	UNP Q31163
B	-9	GLY	-	expression tag	UNP Q31163
B	-8	SER	-	expression tag	UNP Q31163
B	-7	LEU	-	expression tag	UNP Q31163
B	-6	VAL	-	expression tag	UNP Q31163
B	-5	PRO	-	expression tag	UNP Q31163
B	-4	ARG	-	expression tag	UNP Q31163
B	-3	GLY	-	expression tag	UNP Q31163
B	-2	SER	-	expression tag	UNP Q31163
B	-1	GLY	-	expression tag	UNP Q31163
B	0	GLY	-	expression tag	UNP Q31163
B	1	GLY	-	expression tag	UNP Q31163
B	2	GLY	-	expression tag	UNP Q31163
B	199	SER	-	expression tag	UNP Q31163
B	200	ARG	-	expression tag	UNP Q31163
B	201	GLY	-	expression tag	UNP Q31163
B	202	GLY	-	expression tag	UNP Q31163
B	203	LEU	-	expression tag	UNP Q31163
B	204	GLU	-	expression tag	UNP Q31163
B	205	VAL	-	expression tag	UNP Q31163
B	206	LEU	-	expression tag	UNP Q31163
B	207	PHE	-	expression tag	UNP Q31163
B	208	GLN	-	expression tag	UNP Q31163

- Molecule 3 is a protein called 2B4 T-cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	200	Total	C	N	O	S	0	0	0
			1526	946	261	311	8			

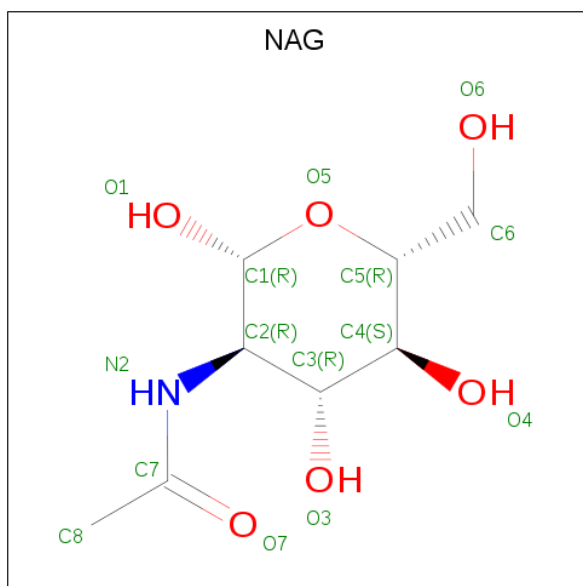
- Molecule 4 is a protein called 2B4 T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	244	Total	C	N	O	S	0	0	0
			1953	1234	339	371	9			

- Molecule 5 is a protein called 2A peptide.

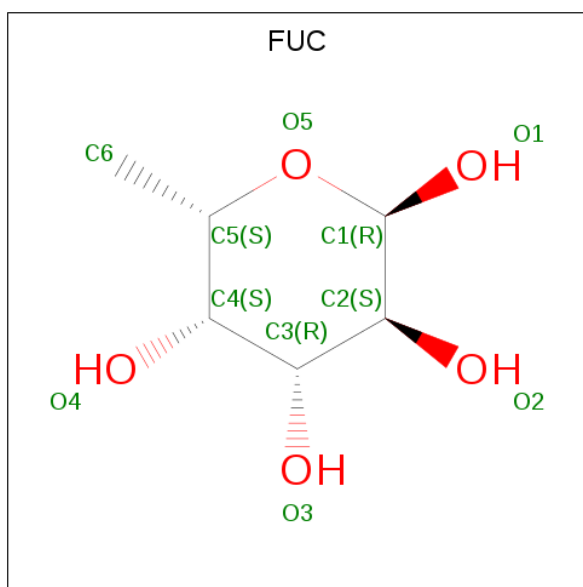
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	P	20	Total	C	N	O	0	0	0
			133	87	21	25			

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: $C_6H_{12}O_5$).

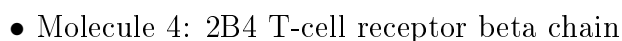


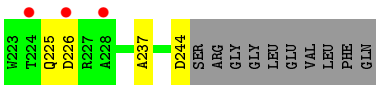
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			10	6	4		
7	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 8 is water.

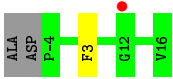
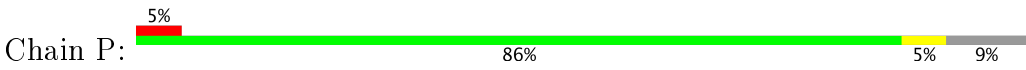
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	31	Total	O	0	0
			31	31		
8	B	28	Total	O	0	0
			28	28		
8	C	28	Total	O	0	0
			28	28		
8	D	30	Total	O	0	0
			30	30		
8	P	1	Total	O	0	0
			1	1		

- Molecule 1: H-2 class II histocompatibility antigen, E-K alpha chain





● Molecule 5: 2A peptide



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	239.94Å 60.18Å 78.36Å 90.00° 104.33° 90.00°	Depositor
Resolution (Å)	47.67 – 2.60 47.67 – 2.60	Depositor EDS
% Data completeness (in resolution range)	96.8 (47.67-2.60) 96.9 (47.67-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.29 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.189 , 0.245 0.190 , 0.244	Depositor DCC
R_{free} test set	1648 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	47.5	Xtriage
Anisotropy	0.677	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 44.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6701	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, FUC

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	0.22	0/1472	0.38	0/1999
2	B	0.23	0/1489	0.39	0/2029
3	C	0.23	0/1556	0.40	0/2112
4	D	0.23	0/2006	0.38	0/2730
5	P	0.24	0/136	0.34	0/182
All	All	0.23	0/6659	0.39	0/9052

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	1431	0	1367	12	0
2	B	1450	0	1359	18	0
3	C	1526	0	1424	11	0
4	D	1953	0	1861	17	0
5	P	133	0	135	1	0
6	A	14	0	13	0	0
6	B	28	0	23	3	0
6	C	28	0	26	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	20	0	20	4	0
8	A	31	0	0	2	0
8	B	28	0	0	5	0
8	C	28	0	0	2	0
8	D	30	0	0	0	0
8	P	1	0	0	0	0
All	All	6701	0	6228	61	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 5.

The worst 5 of 61 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)
2:B:21:THR:O	2:B:80:ARG:NH1	2.07	0.88
6:C:901:NAG:O4	8:C:1007:HOH:O	1.91	0.87
6:B:902:NAG:O4	8:B:1005:HOH:O	1.95	0.85
2:B:10:TYR:HB3	2:B:31:PHE:HB2	1.67	0.77
2:B:20:GLY:HA3	7:B:903:FUC:H4	1.69	0.74

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	171/204 (84%)	166 (97%)	5 (3%)	0	100	100
2	B	170/236 (72%)	163 (96%)	7 (4%)	0	100	100
3	C	198/220 (90%)	192 (97%)	6 (3%)	0	100	100
4	D	242/255 (95%)	232 (96%)	10 (4%)	0	100	100
5	P	18/22 (82%)	18 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	799/937 (85%)	771 (96%)	28 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	155/184 (84%)	155 (100%)	0	100	100
2	B	160/207 (77%)	159 (99%)	1 (1%)	89	96
3	C	170/191 (89%)	169 (99%)	1 (1%)	89	96
4	D	215/225 (96%)	212 (99%)	3 (1%)	71	89
5	P	13/14 (93%)	13 (100%)	0	100	100
All	All	713/821 (87%)	708 (99%)	5 (1%)	87	96

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	67	PHE
3	C	89	LEU
4	D	72	ASN
4	D	193	ARG
4	D	244	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

7 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$
6	NAG	A	901	1	14,14,15	0.19	0	15,19,21	0.45	0
6	NAG	B	901	2,7,6	14,14,15	0.70	1 (7%)	15,19,21	1.51	3 (20%)
6	NAG	B	902	6	14,14,15	0.36	0	15,19,21	0.78	1 (6%)
7	FUC	B	903	6	9,10,11	0.76	0	13,14,16	1.04	2 (15%)
7	FUC	B	904	6	9,10,11	0.61	0	13,14,16	1.24	2 (15%)
6	NAG	C	901	3	14,14,15	0.18	0	15,19,21	0.49	0
6	NAG	C	902	3	14,14,15	0.37	0	15,19,21	0.59	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	901	1	-	0/6/23/26	0/1/1/1
6	NAG	B	901	2,7,6	-	0/6/23/26	0/1/1/1
6	NAG	B	902	6	-	0/6/23/26	0/1/1/1
7	FUC	B	903	6	-	0/0/17/20	0/1/1/1
7	FUC	B	904	6	-	0/0/17/20	0/1/1/1
6	NAG	C	901	3	-	0/6/23/26	0/1/1/1
6	NAG	C	902	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	901	NAG	C1-C2	2.09	1.55	1.52

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	901	NAG	C4-C3-C2	-2.11	107.93	111.02
7	B	903	FUC	C1-O5-C5	2.11	117.06	112.39
7	B	903	FUC	O5-C5-C4	2.29	113.40	109.62
6	B	901	NAG	C1-O5-C5	2.41	115.49	112.17
7	B	904	FUC	O5-C5-C4	2.47	113.69	109.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	902	NAG	3	0
7	B	903	FUC	2	0
7	B	904	FUC	2	0
6	C	901	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, 95th 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 ²)	Q<0.9
1	A	175/204 (85%)	0.03	3 (1%) 70 65	37, 53, 80, 89	0
2	B	176/236 (74%)	0.21	15 (8%) 11 7	37, 54, 99, 117	0
3	C	200/220 (90%)	0.06	4 (2%) 65 59	38, 55, 94, 115	0
4	D	244/255 (95%)	-0.08	6 (2%) 58 50	33, 48, 80, 120	0
5	P	20/22 (90%)	0.09	1 (5%) 30 23	37, 50, 86, 101	0
All	All	815/937 (86%)	0.05	29 (3%) 43 35	33, 51, 91, 120	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	133	ARG	4.5
2	B	165	PRO	4.4
2	B	113	ASN	3.7
4	D	220	ASN	3.4
2	B	189	LYS	3.2

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, 95th 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(Å ²)	Q<0.9
7	FUC	B	903	10/11	0.91	0.31	5.62	77,87,94,100	0
6	NAG	C	902	14/15	0.84	0.16	-	82,89,98,99	0
6	NAG	A	901	14/15	0.81	0.29	-	83,96,103,105	0
7	FUC	B	904	10/11	0.82	0.33	-	100,104,107,110	0
6	NAG	B	901	14/15	0.87	0.15	-	82,93,102,104	0
6	NAG	C	901	14/15	0.81	0.18	-	63,79,91,93	0
6	NAG	B	902	14/15	0.84	0.22	-	93,97,101,106	0

6.5 Other polymers

There are no such residues in this entry.