



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:10 PM GMT

PDB ID : 4R09  
Title : Crystal structure of human TLR8 in complex with ORN06S  
Authors : Tanji, H.; Ohto, U.; Shimizu, T.  
Deposited on : 2014-07-30  
Resolution : 2.62 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

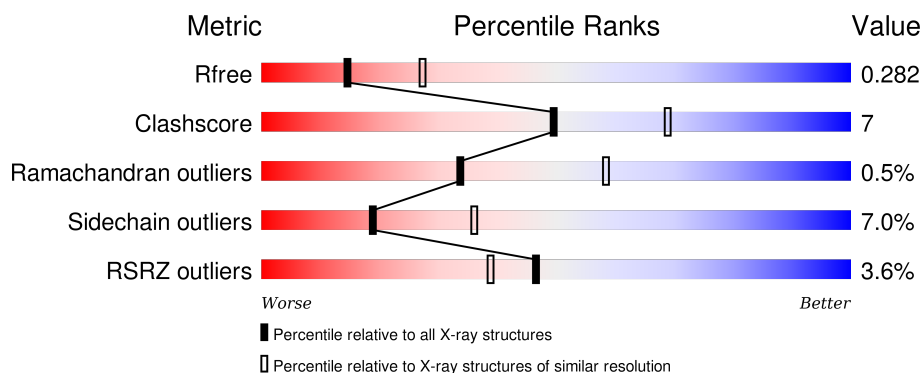
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.62 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	2700 (2.64-2.60)
Clashscore	102246	3065 (2.64-2.60)
Ramachandran outliers	100387	3015 (2.64-2.60)
Sidechain outliers	100360	3015 (2.64-2.60)
RSRZ outliers	91569	2706 (2.64-2.60)

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	811	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 15%, green 76%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>76%</span> <span>15%</span> <span>• 8%</span> </div> </div>
1	B	811	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 2%, yellow 17%, green 74%, grey 7%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>74%</span> <span>17%</span> <span>• 7%</span> </div> </div>
1	C	811	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 4%, orange 4%, yellow 19%, green 70%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>70%</span> <span>19%</span> <span>• 8%</span> </div> </div>
1	D	811	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 6%, orange 6%, yellow 19%, green 72%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>72%</span> <span>19%</span> <span>• 8%</span> </div> </div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 24895 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 Toll-like receptor 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	749	Total	C	N	O	S	0	0	0
			6033	3862	1022	1130	19			
1	B	752	Total	C	N	O	S	0	0	0
			6056	3872	1027	1138	19			
1	C	746	Total	C	N	O	S	0	0	0
			6008	3845	1019	1125	19			
1	D	746	Total	C	N	O	S	0	0	0
			6009	3844	1020	1126	19			

There are 40 discrepancies between the modelled and reference sequences:

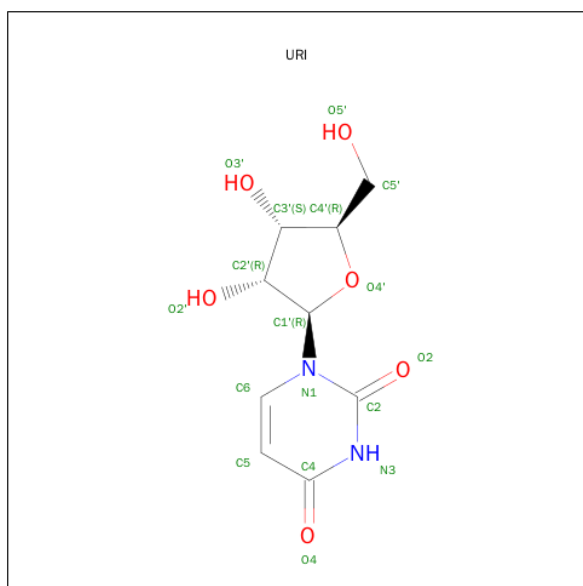
Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ARG	-	EXPRESSION TAG	UNP Q9NR97
A	23	SER	-	EXPRESSION TAG	UNP Q9NR97
A	24	PRO	-	EXPRESSION TAG	UNP Q9NR97
A	25	TRP	-	EXPRESSION TAG	UNP Q9NR97
A	828	GLU	-	EXPRESSION TAG	UNP Q9NR97
A	829	PHE	-	EXPRESSION TAG	UNP Q9NR97
A	830	LEU	-	EXPRESSION TAG	UNP Q9NR97
A	831	VAL	-	EXPRESSION TAG	UNP Q9NR97
A	832	PRO	-	EXPRESSION TAG	UNP Q9NR97
A	833	ARG	-	EXPRESSION TAG	UNP Q9NR97
B	22	ARG	-	EXPRESSION TAG	UNP Q9NR97
B	23	SER	-	EXPRESSION TAG	UNP Q9NR97
B	24	PRO	-	EXPRESSION TAG	UNP Q9NR97
B	25	TRP	-	EXPRESSION TAG	UNP Q9NR97
B	828	GLU	-	EXPRESSION TAG	UNP Q9NR97
B	829	PHE	-	EXPRESSION TAG	UNP Q9NR97
B	830	LEU	-	EXPRESSION TAG	UNP Q9NR97
B	831	VAL	-	EXPRESSION TAG	UNP Q9NR97
B	832	PRO	-	EXPRESSION TAG	UNP Q9NR97
B	833	ARG	-	EXPRESSION TAG	UNP Q9NR97
C	22	ARG	-	EXPRESSION TAG	UNP Q9NR97

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	23	SER	-	EXPRESSION TAG	UNP Q9NR97
C	24	PRO	-	EXPRESSION TAG	UNP Q9NR97
C	25	TRP	-	EXPRESSION TAG	UNP Q9NR97
C	828	GLU	-	EXPRESSION TAG	UNP Q9NR97
C	829	PHE	-	EXPRESSION TAG	UNP Q9NR97
C	830	LEU	-	EXPRESSION TAG	UNP Q9NR97
C	831	VAL	-	EXPRESSION TAG	UNP Q9NR97
C	832	PRO	-	EXPRESSION TAG	UNP Q9NR97
C	833	ARG	-	EXPRESSION TAG	UNP Q9NR97
D	22	ARG	-	EXPRESSION TAG	UNP Q9NR97
D	23	SER	-	EXPRESSION TAG	UNP Q9NR97
D	24	PRO	-	EXPRESSION TAG	UNP Q9NR97
D	25	TRP	-	EXPRESSION TAG	UNP Q9NR97
D	828	GLU	-	EXPRESSION TAG	UNP Q9NR97
D	829	PHE	-	EXPRESSION TAG	UNP Q9NR97
D	830	LEU	-	EXPRESSION TAG	UNP Q9NR97
D	831	VAL	-	EXPRESSION TAG	UNP Q9NR97
D	832	PRO	-	EXPRESSION TAG	UNP Q9NR97
D	833	ARG	-	EXPRESSION TAG	UNP Q9NR97

- Molecule 2 is URIDINE (three-letter code: URI) (formula: C<sub>9</sub>H<sub>12</sub>N<sub>2</sub>O<sub>6</sub>).



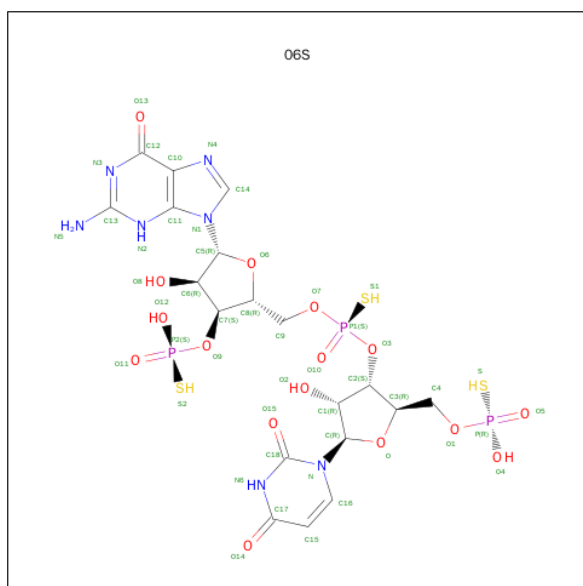
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			17	9	2	6		
2	B	1	Total	C	N	O	0	0
			17	9	2	6		

Continued on next page...

Continued from previous page...

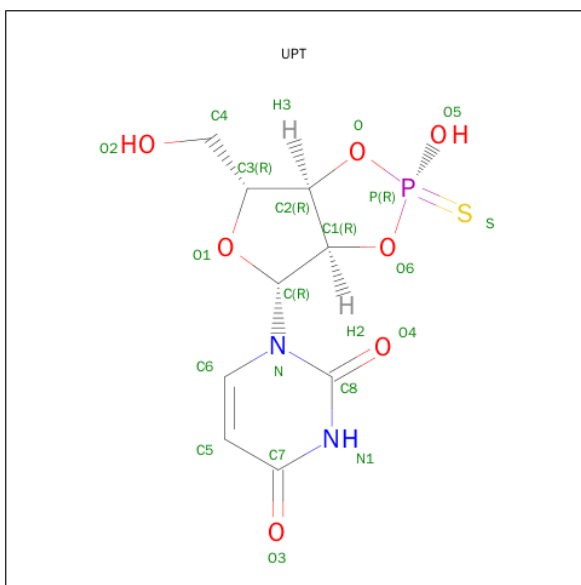
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			17	9	2	6		
2	D	1	Total	C	N	O	0	0
			17	9	2	6		

- Molecule 3 is O-[(2R,3S,4R,5R)-5-(2-AMINO-6-OXO-3,6-DIHYDRO-9H-PURIN-9-YL)-2-({[(S)-{(2R,3S,4R,5R)-5-(2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)-4-HYDROXY-2-[(THIOPHOSPHONOOXY)METHYL]TETRAHYDROFURAN-3-YL}OXY}(SULFANYL)PHOSPHORYL)OXY}METHYL)-4-HYDROXYTETRAHYDROFURAN-3-YL] DIHYDROGEN (S)-PHOSPHOROTHIOATE (three-letter code: 06S) (formula: C<sub>19</sub>H<sub>26</sub>N<sub>7</sub>O<sub>16</sub>P<sub>3</sub>S<sub>3</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	S	0	0
			48	19	7	16	3	3		
3	B	1	Total	C	N	O	P	S	0	0
			48	19	7	16	3	3		
3	C	1	Total	C	N	O	P	S	0	0
			48	19	7	16	3	3		
3	D	1	Total	C	N	O	P	S	0	0
			48	19	7	16	3	3		

- Molecule 4 is 1-[(2R,3AR,4R,6R,6AR)-2-HYDROXY-6-(HYDROXYMETHYL)-2-SULFIDOTETRAHYDROFURO[3,4-D][1,3,2]DIOXAPHOSPHOL-4-YL]PYRIMIDINE-2,4(1H,3H)-DIONE (three-letter code: UPT) (formula: C<sub>9</sub>H<sub>11</sub>N<sub>2</sub>O<sub>7</sub>PS).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	S	0	0
			20	9	2	7	1	1		
4	B	1	Total	C	N	O	P	S	0	0
			20	9	2	7	1	1		
4	C	1	Total	C	N	O	P	S	0	0
			20	9	2	7	1	1		
4	D	1	Total	C	N	O	P	S	0	0
			20	9	2	7	1	1		

- Molecule 5 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	5	Total	C	N	O	0	0
			61	34	2	25		
5	C	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

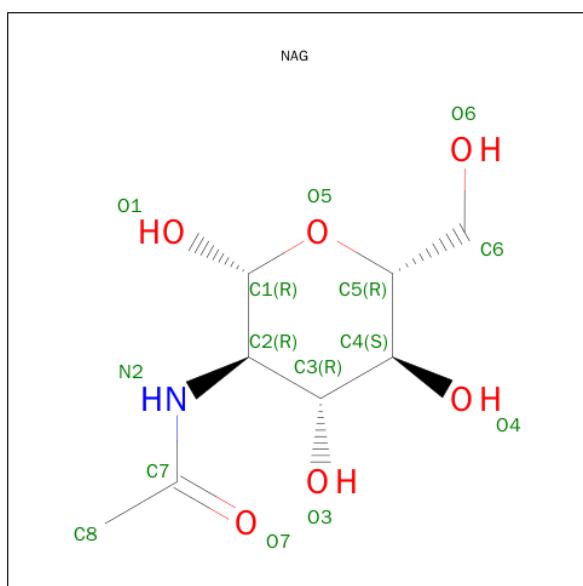
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	2	Total	C	N	O	0	0
			28	16	2	10		
6	A	2	Total	C	N	O	0	0
			28	16	2	10		
6	B	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	2	Total	C	N	O	0	0
			28	16	2	10		
6	C	2	Total	C	N	O	0	0
			28	16	2	10		
6	D	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 7 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

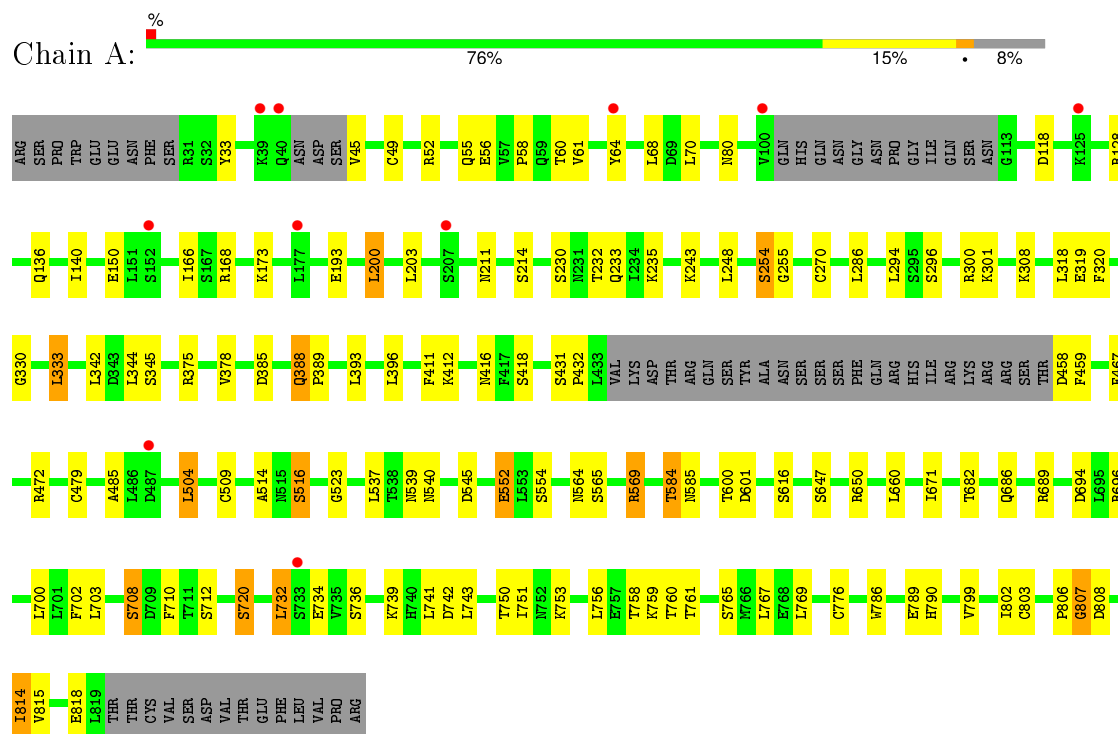
- Molecule 8 is a polymer of unknown type called SUGAR (3-MER).

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

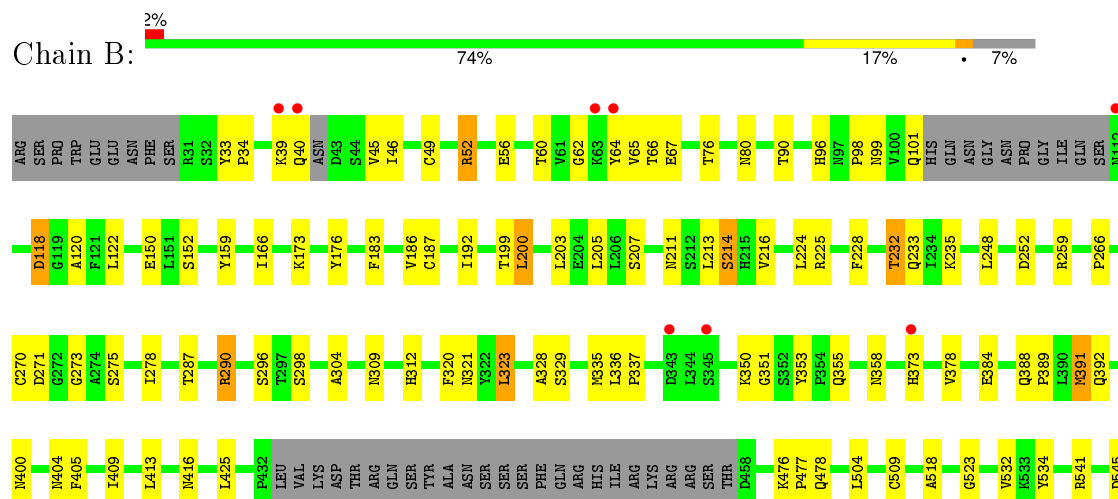
### 3 Residue-property plots [i](#)

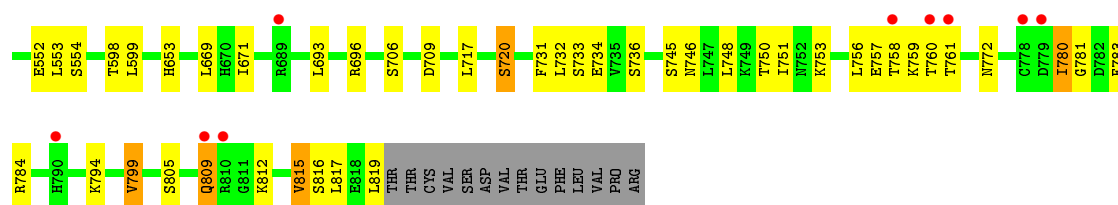
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Toll-like receptor 8

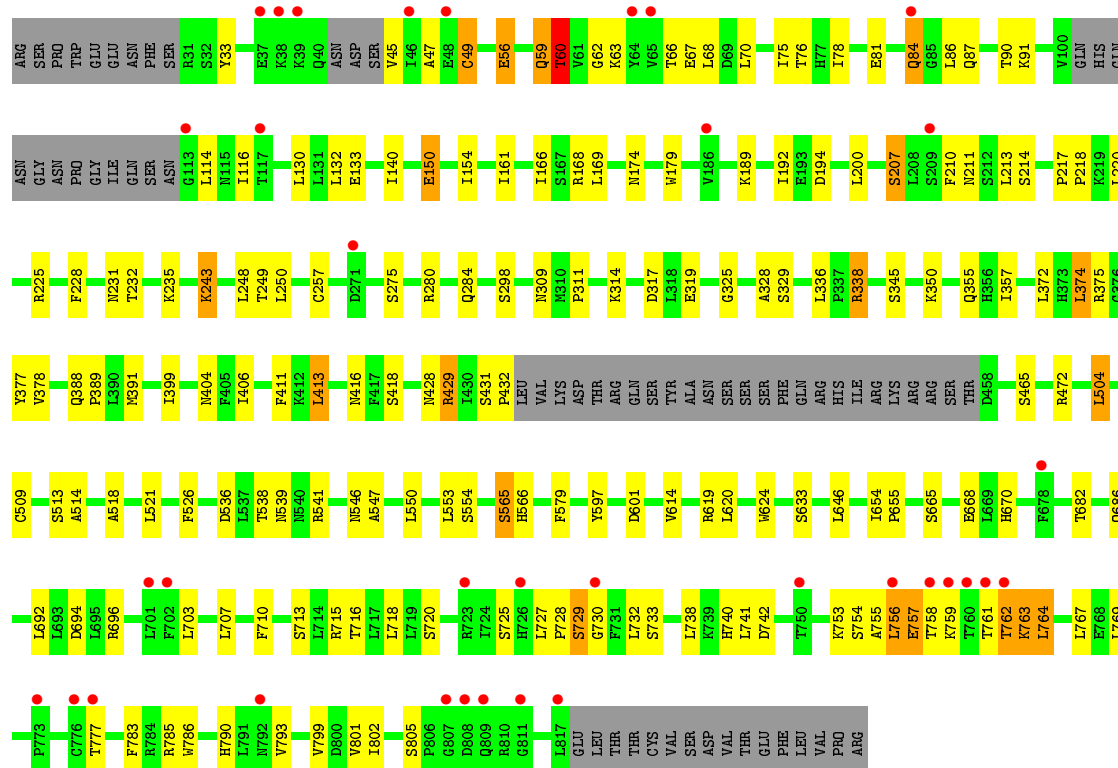


#### • Molecule 1: Toll-like receptor 8

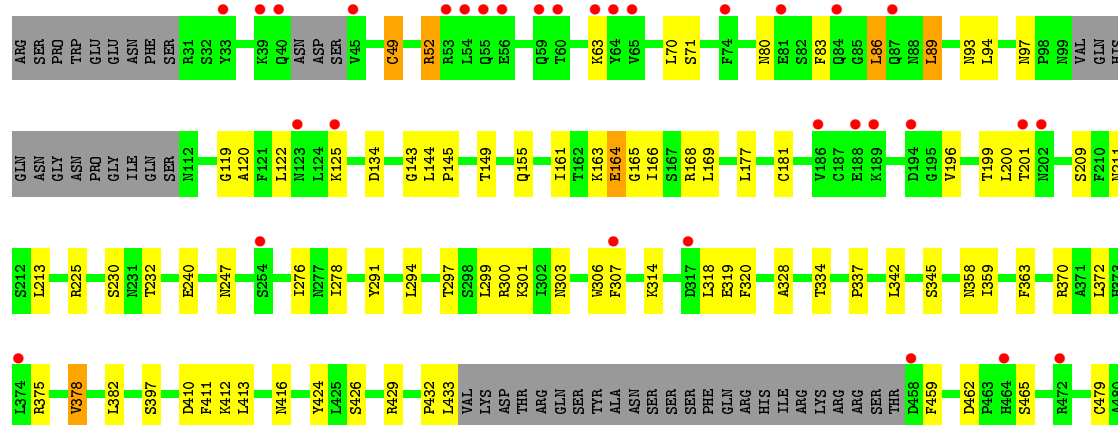


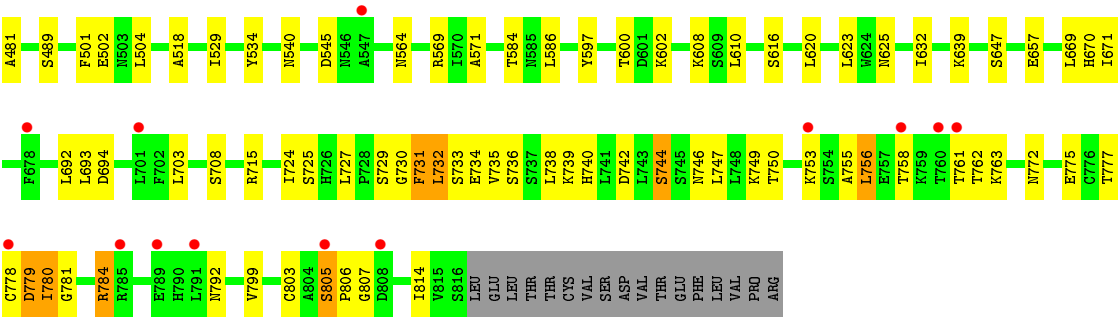


• Molecule 1: Toll-like receptor 8



• Molecule 1: Toll-like receptor 8





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.68Å 141.34Å 170.04Å 90.00° 89.46° 90.00°	Depositor
Resolution (Å)	31.56 – 2.62 31.56 – 2.62	Depositor EDS
% Data completeness (in resolution range)	98.9 (31.56-2.62) 99.0 (31.56-2.62)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.89 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.216 , 0.285 0.217 , 0.282	Depositor DCC
$R_{free}$ test set	6112 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.7	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 31.5	EDS
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 121802 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	24895	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 06S, BMA, NAG, URI, UPT, MAN

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	0.62	0/6157	0.64	0/8350
1	B	0.60	1/6180 (0.0%)	0.62	1/8381 (0.0%)
1	C	0.54	0/6132	0.60	0/8316
1	D	0.52	0/6133	0.58	0/8317
All	All	0.57	1/24602 (0.0%)	0.61	1/33364 (0.0%)

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
1	A	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	477	PRO	N-CD	5.19	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	476	LYS	C-N-CD	5.35	139.63	128.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	254	SER	Peptide
1	B	756	LEU	Peptide

## 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	6033	0	6023	60	0
1	B	6056	0	6035	78	0
1	C	6008	0	5996	101	0
1	D	6009	0	5993	109	0
2	A	17	0	12	0	0
2	B	17	0	12	0	0
2	D	34	0	24	3	0
3	A	48	0	24	0	0
3	B	48	0	24	0	0
3	C	48	0	24	0	0
3	D	48	0	24	1	0
4	A	20	0	11	0	0
4	B	20	0	11	0	0
4	C	20	0	11	1	0
4	D	20	0	11	4	0
5	A	61	0	52	2	0
5	C	61	0	52	0	0
6	A	56	0	50	0	0
6	B	56	0	50	2	0
6	C	28	0	25	0	0
6	D	28	0	25	0	0
7	A	14	0	13	0	0
7	B	14	0	13	0	0
7	D	14	0	13	0	0
8	B	39	0	34	1	0
8	C	39	0	34	0	0
8	D	39	0	34	0	0
All	All	24895	0	24630	349	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 7.

The worst 5 of 349 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:D:731:PHE:HA	1:D:733:SER:H	1.25	1.00
1:C:429:ARG:NH1	2:D:901:URI:O4	1.95	0.99
1:D:432:PRO:HA	1:D:433:LEU:HB2	1.46	0.95
1:C:764:LEU:HD11	1:C:793:VAL:CG1	1.97	0.94
1:D:742:ASP:OD1	1:D:744:SER:OG	1.86	0.92

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	741/811 (91%)	672 (91%)	66 (9%)	3 (0%)	39	63
1	B	744/811 (92%)	671 (90%)	70 (9%)	3 (0%)	39	63
1	C	738/811 (91%)	668 (90%)	65 (9%)	5 (1%)	26	49
1	D	738/811 (91%)	640 (87%)	94 (13%)	4 (0%)	34	58
All	All	2961/3244 (91%)	2651 (90%)	295 (10%)	15 (0%)	34	58

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	780	ILE
1	C	60	THR
1	C	757	GLU
1	D	792	ASN
1	A	807	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	695/755 (92%)	644 (93%)	51 (7%)	17	34
1	B	698/755 (92%)	655 (94%)	43 (6%)	23	44
1	C	692/755 (92%)	637 (92%)	55 (8%)	15	29
1	D	692/755 (92%)	646 (93%)	46 (7%)	21	39
All	All	2777/3020 (92%)	2582 (93%)	195 (7%)	19	36

5 of 195 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	799	VAL
1	C	207	SER
1	D	639	LYS
1	B	809	GLN
1	C	66	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	416	ASN
1	C	59	GLN
1	D	358	ASN
1	B	809	GLN
1	C	202	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 ⓘ

31 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$
5	NAG	A	904	1,5	14,14,15	0.91	0	15,19,21	1.46	4 (26%)
5	NAG	A	905	5	14,14,15	0.68	0	15,19,21	1.52	4 (26%)
5	BMA	A	906	5	11,11,12	1.29	1 (9%)	14,15,17	3.04	4 (28%)
5	MAN	A	907	5	11,11,12	1.08	1 (9%)	14,15,17	2.23	3 (21%)
5	MAN	A	908	5	11,11,12	0.85	1 (9%)	14,15,17	2.29	3 (21%)
6	NAG	A	909	1,6	14,14,15	0.71	0	15,19,21	1.41	5 (33%)
6	NAG	A	910	6	14,14,15	0.62	0	15,19,21	1.42	2 (13%)
6	NAG	A	911	1,6	14,14,15	0.81	0	15,19,21	2.65	4 (26%)
6	NAG	A	912	6	14,14,15	0.89	1 (7%)	15,19,21	2.02	5 (33%)
8	NAG	B	904	1,8	14,14,15	0.97	0	15,19,21	1.41	2 (13%)
8	NAG	B	905	8	14,14,15	0.55	0	15,19,21	1.09	1 (6%)
8	BMA	B	906	8	11,11,12	0.70	0	14,15,17	1.07	1 (7%)
6	NAG	B	907	1,6	14,14,15	0.48	0	15,19,21	1.03	0
6	NAG	B	908	6	14,14,15	0.67	0	15,19,21	1.43	3 (20%)
6	NAG	B	909	1,6	14,14,15	0.58	0	15,19,21	1.00	1 (6%)
6	NAG	B	910	6	14,14,15	1.55	2 (14%)	15,19,21	2.40	7 (46%)
5	NAG	C	903	1,5	14,14,15	0.42	0	15,19,21	1.33	2 (13%)
5	NAG	C	904	5	14,14,15	0.65	0	15,19,21	1.07	1 (6%)
5	BMA	C	905	5	11,11,12	0.67	0	14,15,17	2.71	3 (21%)
5	MAN	C	906	5	11,11,12	1.04	0	14,15,17	1.90	1 (7%)
5	MAN	C	907	5	11,11,12	1.01	1 (9%)	14,15,17	1.89	4 (28%)
8	NAG	C	908	1,8	14,14,15	0.66	0	15,19,21	1.20	2 (13%)
8	NAG	C	909	8	14,14,15	0.96	1 (7%)	15,19,21	1.09	1 (6%)
8	BMA	C	910	8	11,11,12	0.54	0	14,15,17	1.29	2 (14%)
6	NAG	C	911	1,6	14,14,15	0.53	0	15,19,21	1.86	3 (20%)
6	NAG	C	912	6	14,14,15	0.69	0	15,19,21	1.57	3 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	NAG	D	905	1,8	14,14,15	0.62	0	15,19,21	1.08	1 (6%)
8	NAG	D	906	8	14,14,15	0.68	0	15,19,21	1.53	2 (13%)
8	BMA	D	907	8	11,11,12	0.87	0	14,15,17	1.57	4 (28%)
6	NAG	D	909	1,6	14,14,15	0.64	0	15,19,21	1.27	2 (13%)
6	NAG	D	910	6	14,14,15	0.97	1 (7%)	15,19,21	1.46	1 (6%)

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
5	NAG	A	904	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	905	5	-	0/6/23/26	0/1/1/1
5	BMA	A	906	5	-	0/2/19/22	0/1/1/1
5	MAN	A	907	5	-	0/2/19/22	0/1/1/1
5	MAN	A	908	5	-	0/2/19/22	0/1/1/1
6	NAG	A	909	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	910	6	-	0/6/23/26	0/1/1/1
6	NAG	A	911	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	912	6	-	0/6/23/26	0/1/1/1
8	NAG	B	904	1,8	-	0/6/23/26	0/1/1/1
8	NAG	B	905	8	-	0/6/23/26	0/1/1/1
8	BMA	B	906	8	-	0/2/19/22	0/1/1/1
6	NAG	B	907	1,6	-	0/6/23/26	0/1/1/1
6	NAG	B	908	6	-	0/6/23/26	0/1/1/1
6	NAG	B	909	1,6	-	0/6/23/26	0/1/1/1
6	NAG	B	910	6	-	0/6/23/26	0/1/1/1
5	NAG	C	903	1,5	-	0/6/23/26	0/1/1/1
5	NAG	C	904	5	-	0/6/23/26	0/1/1/1
5	BMA	C	905	5	-	0/2/19/22	0/1/1/1
5	MAN	C	906	5	-	0/2/19/22	0/1/1/1
5	MAN	C	907	5	-	0/2/19/22	0/1/1/1
8	NAG	C	908	1,8	-	0/6/23/26	0/1/1/1
8	NAG	C	909	8	-	0/6/23/26	0/1/1/1
8	BMA	C	910	8	-	0/2/19/22	0/1/1/1
6	NAG	C	911	1,6	-	0/6/23/26	0/1/1/1
6	NAG	C	912	6	-	0/6/23/26	0/1/1/1
8	NAG	D	905	1,8	-	0/6/23/26	0/1/1/1
8	NAG	D	906	8	-	0/6/23/26	0/1/1/1
8	BMA	D	907	8	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	D	909	1,6	-	0/6/23/26	0/1/1/1
6	NAG	D	910	6	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	910	NAG	C1-C2	-3.69	1.47	1.52
6	B	910	NAG	O5-C1	-3.46	1.37	1.43
8	C	909	NAG	O5-C1	-3.12	1.38	1.43
6	D	910	NAG	O5-C1	-2.83	1.39	1.43
6	A	912	NAG	O5-C1	-2.41	1.39	1.43

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	908	MAN	O5-C1-C2	-7.12	99.31	110.86
8	D	906	NAG	O4-C4-C3	-3.66	102.09	110.34
6	A	911	NAG	O7-C7-C8	-3.61	115.44	122.06
6	B	910	NAG	O3-C3-C2	-3.36	102.47	109.11
8	B	904	NAG	C3-C4-C5	-3.27	104.50	110.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	904	NAG	2	0
5	A	905	NAG	1	0
8	B	905	NAG	1	0
6	B	910	NAG	2	0

## 5.6 Ligand geometry ⓘ

15 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$
2	URI	A	901	-	12,18,18	1.19	1 (8%)	15,26,26	3.63	2 (13%)
3	06S	A	902	-	35,52,52	1.16	5 (14%)	44,81,81	2.01	11 (25%)
4	UPT	A	903	-	14,22,22	1.28	2 (14%)	17,34,34	1.59	1 (5%)
7	NAG	A	913	1	14,14,15	0.69	0	15,19,21	0.68	0
2	URI	B	901	-	12,18,18	1.18	1 (8%)	15,26,26	3.64	2 (13%)
3	06S	B	902	-	35,52,52	1.09	3 (8%)	44,81,81	1.91	7 (15%)
4	UPT	B	903	-	14,22,22	0.73	0	17,34,34	1.70	1 (5%)
7	NAG	B	911	1	14,14,15	0.64	0	15,19,21	1.21	1 (6%)
3	06S	C	901	-	35,52,52	1.19	3 (8%)	44,81,81	1.98	9 (20%)
4	UPT	C	902	-	14,22,22	0.89	0	17,34,34	1.77	1 (5%)
2	URI	D	901	-	12,18,18	1.11	1 (8%)	15,26,26	3.66	3 (20%)
2	URI	D	902	-	12,18,18	1.11	1 (8%)	15,26,26	3.59	2 (13%)
3	06S	D	903	-	35,52,52	1.29	5 (14%)	44,81,81	2.15	13 (29%)
4	UPT	D	904	-	14,22,22	1.13	1 (7%)	17,34,34	1.73	1 (5%)
7	NAG	D	908	1	14,14,15	0.84	0	15,19,21	1.63	3 (20%)

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
2	URI	A	901	-	-	0/2/22/22	0/2/2/2
3	06S	A	902	-	-	0/17/62/62	0/5/5/5
4	UPT	A	903	-	-	0/2/32/32	0/3/3/3
7	NAG	A	913	1	-	0/6/23/26	0/1/1/1
2	URI	B	901	-	-	0/2/22/22	0/2/2/2
3	06S	B	902	-	-	0/17/62/62	0/5/5/5
4	UPT	B	903	-	-	0/2/32/32	0/3/3/3
7	NAG	B	911	1	-	0/6/23/26	0/1/1/1
3	06S	C	901	-	-	0/17/62/62	0/5/5/5
4	UPT	C	902	-	-	0/2/32/32	0/3/3/3
2	URI	D	901	-	-	0/2/22/22	0/2/2/2
2	URI	D	902	-	-	0/2/22/22	0/2/2/2
3	06S	D	903	-	-	0/17/62/62	0/5/5/5
4	UPT	D	904	-	-	0/2/32/32	0/3/3/3
7	NAG	D	908	1	-	0/6/23/26	0/1/1/1

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	902	06S	C12-C10	-2.99	1.35	1.41
4	A	903	UPT	C6-N	-2.18	1.32	1.35
3	B	902	06S	P2-O12	-2.12	1.51	1.56
3	A	902	06S	P1-O7	2.04	1.61	1.58
3	A	902	06S	P-O1	2.27	1.61	1.58

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	06S	N2-C13-N3	-4.80	120.13	127.44
3	B	902	06S	N2-C13-N3	-4.38	120.77	127.44
3	C	901	06S	N2-C13-N3	-4.31	120.88	127.44
3	D	903	06S	N2-C13-N3	-4.18	121.08	127.44
3	D	903	06S	C10-C12-N3	-3.78	118.42	123.59

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	902	UPT	1	0
2	D	901	URI	1	0
2	D	902	URI	2	0
3	D	903	06S	1	0
4	D	904	UPT	4	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	749/811 (92%)	-0.31	10 (1%) 79 75	29, 48, 83, 114	0
1	B	752/811 (92%)	-0.12	17 (2%) 64 57	29, 55, 99, 138	0
1	C	746/811 (91%)	0.06	35 (4%) 35 28	36, 65, 119, 170	0
1	D	746/811 (91%)	0.13	45 (6%) 25 19	37, 73, 118, 184	0
All	All	2993/3244 (92%)	-0.06	107 (3%) 46 39	29, 59, 110, 184	0

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	758	THR	8.6
1	D	758	THR	7.4
1	C	38	LYS	7.1
1	C	762	THR	6.8
1	D	63	LYS	5.9

### 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](#)

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
5	NAG	A	904	14/15	0.98	0.15	-0.25	31,32,37,38	0
8	NAG	D	905	14/15	0.95	0.21	-0.29	53,58,62,64	0
5	NAG	C	903	14/15	0.96	0.17	-0.33	40,43,46,46	0
8	NAG	C	908	14/15	0.95	0.15	-0.75	43,45,54,65	0
8	NAG	B	904	14/15	0.95	0.16	-0.81	38,41,44,46	0
5	NAG	C	904	14/15	0.97	0.10	-1.12	42,49,54,56	0
6	NAG	C	911	14/15	0.94	0.10	-1.26	49,55,61,66	0
6	NAG	A	909	14/15	0.97	0.12	-1.27	34,40,45,50	0
6	NAG	D	909	14/15	0.97	0.10	-1.32	39,44,49,54	0
6	NAG	A	911	14/15	0.96	0.12	-1.51	30,36,41,42	0
8	NAG	D	906	14/15	0.94	0.11	-1.61	53,57,63,65	0
6	NAG	B	907	14/15	0.97	0.09	-1.62	33,36,39,39	0
8	NAG	B	905	14/15	0.96	0.09	-1.62	42,45,52,52	0
5	NAG	A	905	14/15	0.98	0.09	-1.97	31,36,38,44	0
6	NAG	B	909	14/15	0.96	0.09	-2.06	36,40,41,42	0
5	MAN	C	907	11/12	0.86	0.28	-	61,67,75,87	0
6	NAG	A	912	14/15	0.91	0.16	-	34,37,40,49	0
5	BMA	C	905	11/12	0.95	0.16	-	52,59,66,70	0
5	MAN	A	908	11/12	0.78	0.22	-	54,70,80,83	0
8	BMA	D	907	11/12	0.89	0.16	-	61,70,73,77	0
5	BMA	A	906	11/12	0.95	0.11	-	46,56,64,64	0
8	BMA	B	906	11/12	0.92	0.14	-	48,54,57,60	0
6	NAG	D	910	14/15	0.93	0.14	-	46,50,55,56	0
6	NAG	B	910	14/15	0.93	0.14	-	38,41,49,54	0
6	NAG	B	908	14/15	0.91	0.16	-	48,55,66,67	0
8	NAG	C	909	14/15	0.92	0.16	-	55,64,74,82	0
5	MAN	C	906	11/12	0.92	0.15	-	70,72,77,78	0
6	NAG	C	912	14/15	0.90	0.19	-	42,49,57,59	0
6	NAG	A	910	14/15	0.92	0.22	-	46,53,62,62	0
8	BMA	C	910	11/12	0.85	0.34	-	73,80,84,89	0
5	MAN	A	907	11/12	0.91	0.17	-	59,67,73,75	0

## 6.4 Ligands ⓘ

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
7	NAG	A	913	14/15	0.93	0.19	0.02	56,61,67,70	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	UPT	A	903	20/20	0.97	0.14	-0.09	36,40,44,45	0
3	06S	B	902	48/48	0.87	0.18	-0.55	45,59,82,90	0
2	URI	D	901	17/17	0.96	0.12	-0.66	35,40,45,45	0
4	UPT	B	903	20/20	0.97	0.14	-0.71	37,43,50,52	0
7	NAG	D	908	14/15	0.93	0.13	-0.76	42,44,49,54	0
4	UPT	C	902	20/20	0.97	0.12	-0.77	45,49,58,60	0
3	06S	A	902	48/48	0.93	0.12	-0.81	36,52,75,87	0
2	URI	A	901	17/17	0.95	0.12	-0.94	36,40,43,48	0
2	URI	D	902	17/17	0.93	0.11	-1.00	58,64,69,72	0
2	URI	B	901	17/17	0.96	0.10	-1.16	34,38,47,53	0
3	06S	C	901	48/48	0.93	0.13	-1.40	47,61,77,87	0
4	UPT	D	904	20/20	0.97	0.12	-1.61	57,65,75,76	0
3	06S	D	903	48/48	0.91	0.12	-1.77	57,70,87,90	0
7	NAG	B	911	14/15	0.93	0.17	-	56,61,64,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.