Full wwPDB X-ray Structure Validation Report (i)

Mar 1, 2026 – 06:38 AM UTC

PDB ID	:	9NIH / pdb_00009nih
Title	:	Crystal structure of HLA-DR4 presenting citrullinated Tenascin C peptide
Authors	:	Dao, H.T.; Loh, T.J.; Lim, J.J.; Rossjohn, J.
Deposited on	:	2025-02-26
Resolution	:	2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

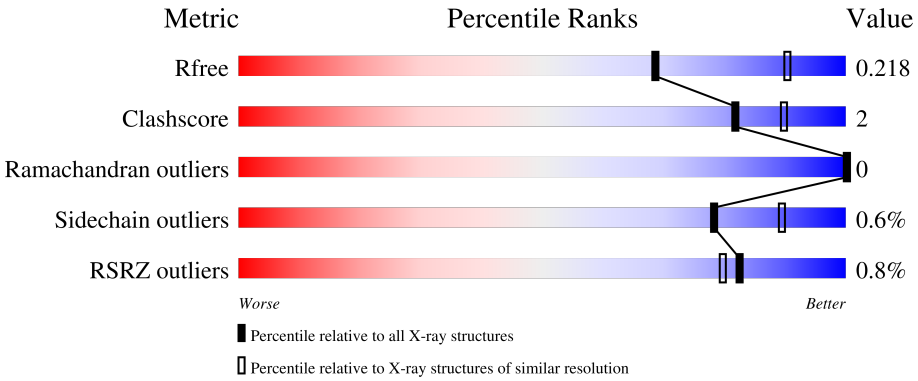
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.40 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}	180053	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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 >=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 <=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	189	89% 7% .
2	B	198	2% 86% 9% 5%
3	C	14	57% 14% 7% 21%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3232 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
1	A	182	Total	C	N	O	S	0	0	0
			1458	945	239	269	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	ILE	-	expression tag	UNP P01903
A	2	LYS	-	expression tag	UNP P01903
A	3	GLU	-	expression tag	UNP P01903
A	4	GLU	-	expression tag	UNP P01903
A	182	THR	-	expression tag	UNP P01903
A	183	SER	-	expression tag	UNP P01903
A	184	GLY	-	expression tag	UNP P01903
A	185	ASP	-	expression tag	UNP P01903
A	186	ASP	-	expression tag	UNP P01903
A	187	ASP	-	expression tag	UNP P01903
A	188	ASP	-	expression tag	UNP P01903
A	189	LYS	-	expression tag	UNP P01903

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	188	Total	C	N	O	S	0	0	0
			1520	962	270	283	5			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	191	THR	-	expression tag	UNP A0A1V1IGJ9
B	192	GLY	-	expression tag	UNP A0A1V1IGJ9
B	193	GLY	-	expression tag	UNP A0A1V1IGJ9
B	194	ASP	-	expression tag	UNP A0A1V1IGJ9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	ASP	-	expression tag	UNP A0A1V1IGJ9
B	196	ASP	-	expression tag	UNP A0A1V1IGJ9
B	197	ASP	-	expression tag	UNP A0A1V1IGJ9
B	198	LYS	-	expression tag	UNP A0A1V1IGJ9

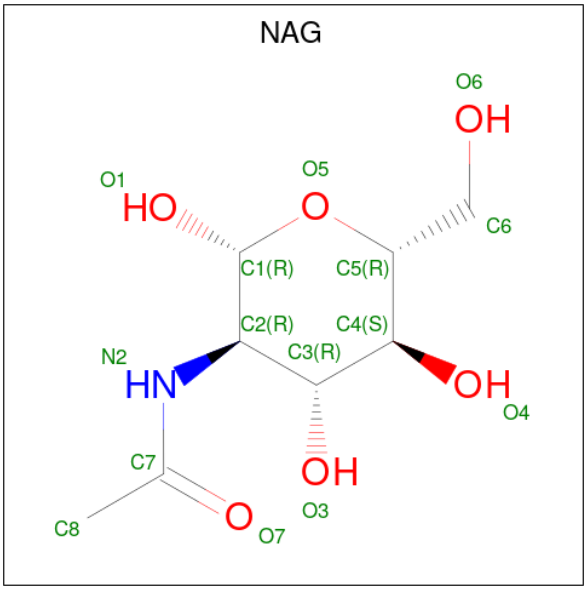
- Molecule 3 is a protein called Tenascin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	11	Total	C	N	O	0	0	0
			98	62	16	20			

There are 2 discrepancies between the modelled and reference sequences:

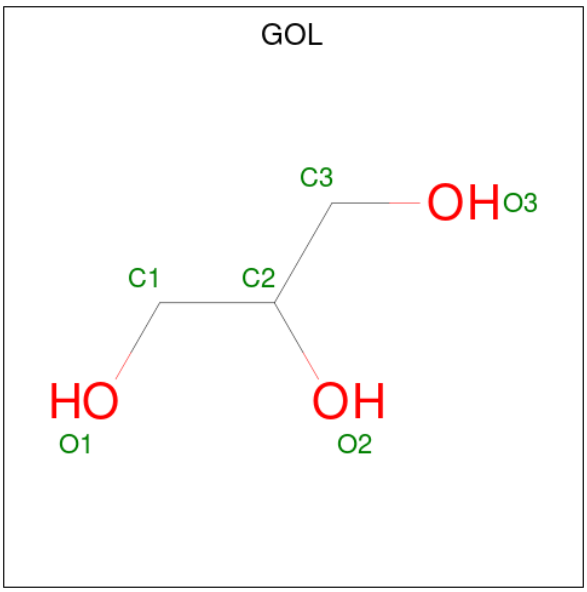
Chain	Residue	Modelled	Actual	Comment	Reference
C	1025	LYS	-	insertion	UNP P24821
C	1026	LYS	-	insertion	UNP P24821

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 5 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0

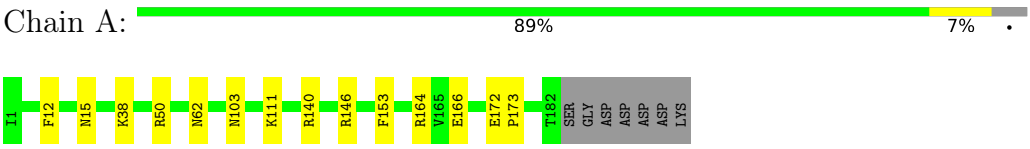
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	54	Total O 54 54	0	0
6	B	49	Total O 49 49	0	0
6	C	7	Total O 7 7	0	0

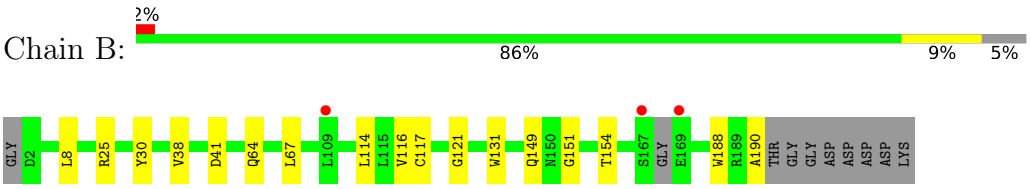
3 Residue-property plots (i)

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



• Molecule 2: HLA class II histocompatibility antigen DR beta chain



• Molecule 3: Tenascin

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	119.30Å 119.30Å 73.27Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.53 – 2.40 34.53 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.6 (34.53-2.40) 99.6 (34.53-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.72 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.184 , 0.219 0.184 , 0.218	Depositor DCC
R_{free} test set	786 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 40.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.010 for -1/3*h+1/3*k+4/3*l,-k,2/3*h+1/3*k+1/3*l 0.013 for -2/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+4/3*l,-1/3*h+1/3*k+1/3*l 0.013 for -h,1/3*h-1/3*k-4/3*l,-1/3*h-2/3*k+1/3*l 0.021 for -1/3*h-2/3*k+4/3*l,-2/3*h-1/3*k-4/3*l,1/3*h-1/3*k-1/3*l 0.012 for -h,2/3*h+1/3*k+4/3*l,1/3*h+2/3*k-1/3*l 0.015 for 1/3*h+2/3*k-4/3*l,-k,-2/3*h-1/3*k-1/3*l 0.023 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3232	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: NAG, GOL, CIR

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	0.19	0/1502	0.38	0/2053
2	B	0.18	0/1562	0.40	0/2130
3	C	0.17	0/76	0.32	0/101
All	All	0.19	0/3140	0.39	0/4284

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	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	1013	ASP	Mainchain
3	C	1016	CIR	Mainchain

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	1458	0	1382	7	1
2	B	1520	0	1397	10	0
3	C	98	0	82	0	0
4	A	14	0	13	0	0
4	B	14	0	13	0	0
5	A	6	0	8	0	0
5	B	12	0	16	0	0
6	A	54	0	0	0	1
6	B	49	0	0	0	1
6	C	7	0	0	0	0
All	All	3232	0	2911	15	2

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 2.

All (15) 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:164:ARG:NH1	1:A:166:GLU:OE1	2.33	0.61
2:B:25:ARG:NH1	2:B:41:ASP:OD2	2.34	0.60
1:A:111:LYS:HG2	1:A:140:ARG:CZ	2.35	0.57
2:B:30:TYR:HB2	2:B:38:VAL:HG12	1.89	0.53
2:B:188:TRP:CZ3	2:B:190:ALA:HB2	2.47	0.50
1:A:172:GLU:HG2	1:A:173:PRO:HD2	1.95	0.47
2:B:114:LEU:HA	2:B:114:LEU:HD23	1.85	0.45
1:A:12:PHE:HB2	2:B:8:LEU:HD11	1.98	0.45
2:B:64:GLN:HB2	2:B:67:LEU:HB2	2.00	0.43
1:A:146:ARG:NH1	2:B:149:GLN:OE1	2.51	0.43
1:A:103:ASN:HB3	1:A:153:PHE:CE1	2.54	0.42
2:B:121:GLY:HA2	2:B:154:THR:HB	2.03	0.41
2:B:149:GLN:HG2	2:B:151:GLY:H	1.85	0.41
1:A:38:LYS:HA	1:A:38:LYS:HD3	1.86	0.41
2:B:117:CYS:HB2	2:B:131:TRP:CZ2	2.57	0.40

All (2) 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)
1:A:50:ARG:NH2	1:A:62:ASN:OD1[9_544]	2.03	0.17
6:A:306:HOH:O	6:B:343:HOH:O[5_445]	2.12	0.08

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	180/189 (95%)	178 (99%)	2 (1%)	0	100	100
2	B	184/198 (93%)	176 (96%)	8 (4%)	0	100	100
3	C	7/14 (50%)	7 (100%)	0	0	100	100
All	All	371/401 (92%)	361 (97%)	10 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	158/173 (91%)	157 (99%)	1 (1%)	78	89
2	B	160/177 (90%)	159 (99%)	1 (1%)	78	89
3	C	9/11 (82%)	9 (100%)	0	100	100
All	All	327/361 (91%)	325 (99%)	2 (1%)	78	89

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
2	B	116	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	15	ASN
1	A	149	HIS
2	B	113	ASN
2	B	156	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)

2 non-standard protein/DNA/RNA residues 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	CIR	C	1016	3	9,10,11	7.24	2 (22%)	6,11,13	1.67	1 (16%)
3	CIR	C	1014	3	9,10,11	4.70	4 (44%)	6,11,13	1.51	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
3	CIR	C	1016	3	-	3/8/9/11	-
3	CIR	C	1014	3	-	2/8/9/11	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1016	CIR	C7-N6	20.85	1.65	1.34
3	C	1014	CIR	O7-C7	9.92	1.40	1.24
3	C	1014	CIR	C7-N6	8.08	1.46	1.34
3	C	1016	CIR	O-C	-5.68	0.98	1.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1014	CIR	O-C	5.31	1.40	1.20
3	C	1014	CIR	C7-N8	2.29	1.39	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	1016	CIR	C5-N6-C7	-2.54	119.33	122.64

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1016	CIR	C3-C4-C5-N6
3	C	1014	CIR	C3-C4-C5-N6
3	C	1014	CIR	CA-C3-C4-C5
3	C	1016	CIR	O7-C7-N6-C5
3	C	1016	CIR	N8-C7-N6-C5

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

5 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	$\#\|Z\| > 2$
5	GOL	A	202	-	5,5,5	1.02	0	5,5,5	0.97	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	A	201	1	14,14,15	0.20	0	17,19,21	0.54	0
4	NAG	B	201	2	14,14,15	0.53	0	17,19,21	0.43	0
5	GOL	B	203	-	5,5,5	0.82	0	5,5,5	1.10	0
5	GOL	B	202	-	5,5,5	0.92	0	5,5,5	1.07	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
5	GOL	A	202	-	-	0/4/4/4	-
4	NAG	A	201	1	-	0/6/23/26	0/1/1/1
4	NAG	B	201	2	-	0/6/23/26	0/1/1/1
5	GOL	B	203	-	-	2/4/4/4	-
5	GOL	B	202	-	-	0/4/4/4	-

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
5	B	203	GOL	O1-C1-C2-C3
5	B	203	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

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^{th} 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	182/189 (96%)	-0.19	0 100 100	27, 40, 67, 82	0
2	B	188/198 (94%)	-0.08	3 (1%) 70 66	27, 41, 74, 118	0
3	C	9/14 (64%)	-0.09	0 100 100	31, 35, 59, 62	0
All	All	379/401 (94%)	-0.13	3 (0%) 82 80	27, 40, 71, 118	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	167	SER	3.4
2	B	169	GLU	2.8
2	B	109	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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^{th} 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^2)	Q<0.9
3	CIR	C	1014	11/12	0.91	0.13	34,45,88,94	0
3	CIR	C	1016	11/12	0.91	0.12	30,40,64,68	0

6.3 Carbohydrates

There are no oligosaccharides 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^{th} 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^2)	Q<0.9
5	GOL	A	202	6/6	0.63	0.14	69,72,75,78	0
4	NAG	A	201	14/15	0.64	0.15	86,97,101,106	0
4	NAG	B	201	14/15	0.71	0.14	62,71,74,79	0
5	GOL	B	202	6/6	0.80	0.10	67,72,77,77	0
5	GOL	B	203	6/6	0.90	0.12	56,64,67,75	0

6.5 Other polymers (i)

There are no such residues in this entry.