Full wwPDB X-ray Structure Validation Report ⓘ

Mar 20, 2025 – 10:18 AM EDT

PDB ID	:	9EJH
Title	:	Peptide-independent T cell receptor recognition of HLA-DQ2
Authors	:	Lim, J.J.; Loh, T.J.; Reid, H.H.; Rossjohn, J.
Deposited on	:	2024-11-27
Resolution	:	2.45 Å(reported)

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

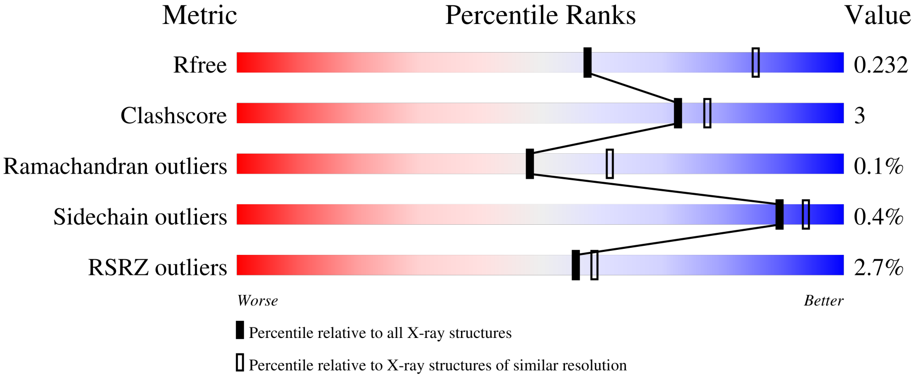
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

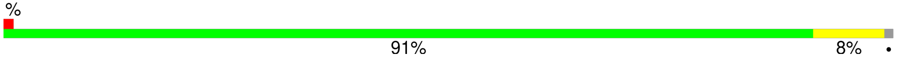
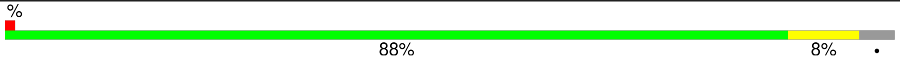
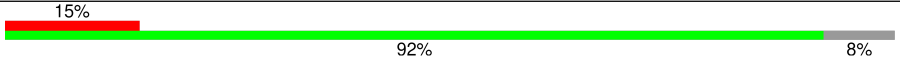
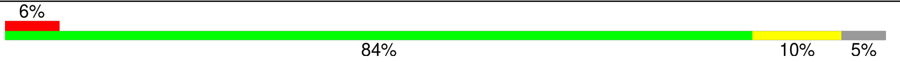
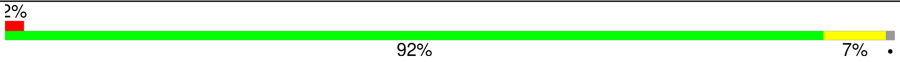
The reported resolution of this entry is 2.45 Å.

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}	164625	1096 (2.46-2.46)
Clashscore	180529	1178 (2.46-2.46)
Ramachandran outliers	177936	1170 (2.46-2.46)
Sidechain outliers	177891	1170 (2.46-2.46)
RSRZ outliers	164620	1096 (2.46-2.46)

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	183	 % 91% 8% •
2	B	194	 % 88% 8% •
3	C	13	 15% 92% 8%
4	D	204	 6% 84% 10% 5%
5	E	242	 2% 92% 7% •

2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 6638 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, DQ alpha 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	182	1399	906	225	266	2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	SER	CYS	conflict	UNP P01909

- Molecule 2 is a protein called MHC class II HLA-DQ-beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	187	1515	954	272	282	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	193	THR	-	expression tag	UNP O19712
B	194	GLY	-	expression tag	UNP O19712

- Molecule 3 is a protein called HLA class II histocompatibility antigen gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	12	84	55	13	14	2	0	0	0

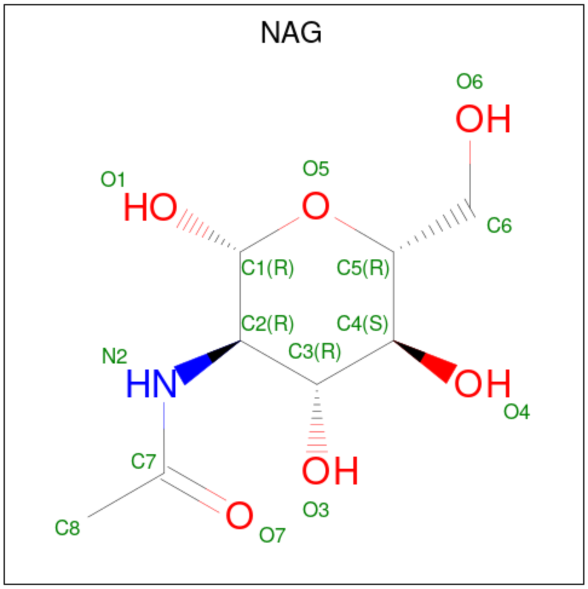
- Molecule 4 is a protein called G9 T cell receptor alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	194	1434	911	236	277	10	0	0	0

- Molecule 5 is a protein called G9 T cell receptor beta chain.

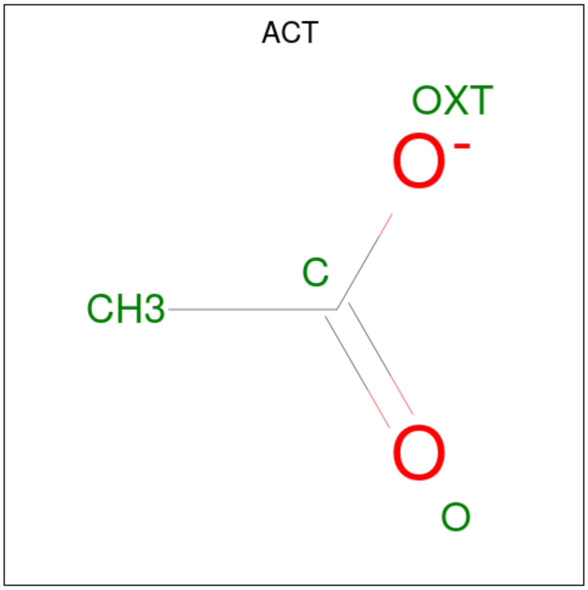
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	240	1859	1174	327	352	6	0	0	0

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



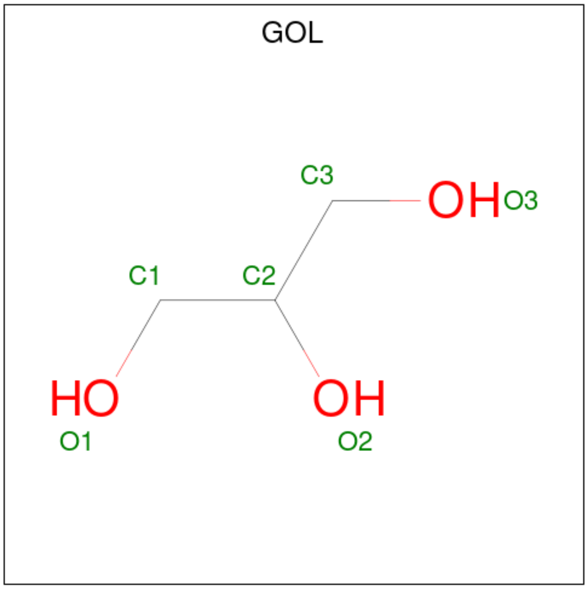
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	A	1	14	8	1	5	0	0
6	B	1	14	8	1	5	0	0

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	E	1	Total C O 4 2 2	0	0
7	E	1	Total C O 4 2 2	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

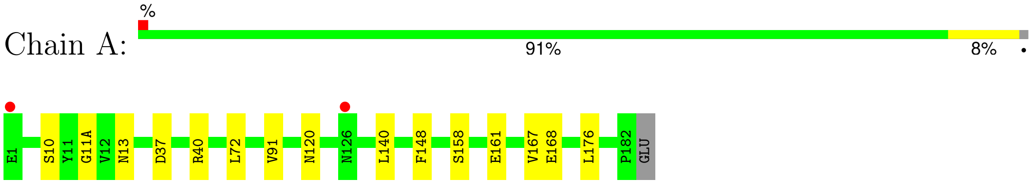
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	63	Total O 63 63	0	0
9	B	83	Total O 83 83	0	0
9	C	3	Total O 3 3	0	0
9	D	59	Total O 59 59	0	0
9	E	89	Total O 89 89	0	0

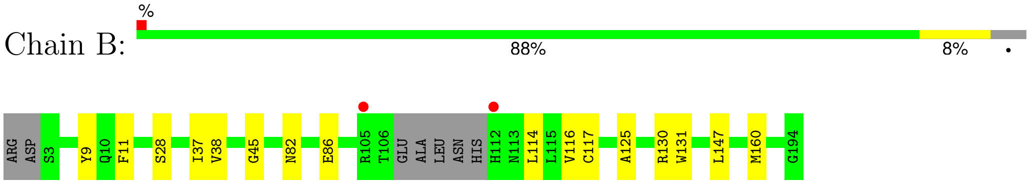
3 Residue-property plots ⓘ

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.

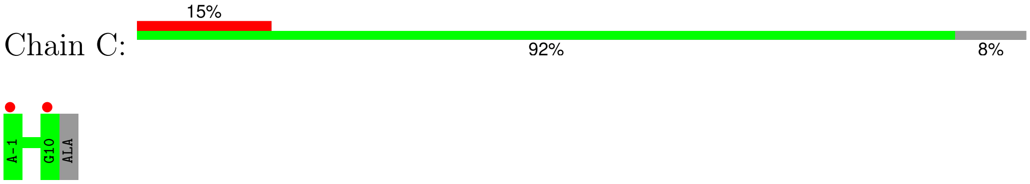
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain



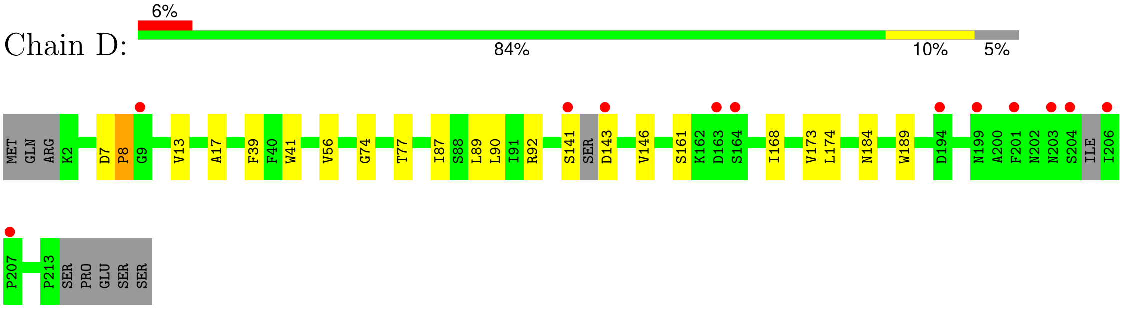
- Molecule 2: MHC class II HLA-DQ-beta-1



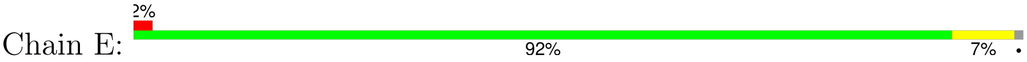
- Molecule 3: HLA class II histocompatibility antigen gamma chain

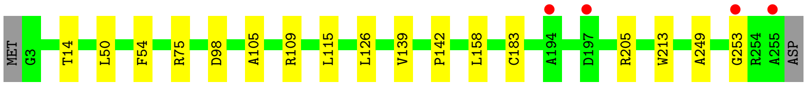


- Molecule 4: G9 T cell receptor alpha chain



- Molecule 5: G9 T cell receptor beta chain




MET
G3
T14
L50
F54
R75
D98
A105
R109
L115
L126
V139
P142
L158
C183
A194
D197
R205
W213
A249
G253
R254
A255
ASP

4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	92.97Å 43.27Å 257.96Å 90.00° 90.31° 90.00°	Depositor
Resolution (Å)	45.22 – 2.45 45.22 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.22-2.45) 99.9 (45.22-2.45)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.45Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.176 , 0.232 0.176 , 0.232	Depositor DCC
R_{free} test set	1954 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	37.3	Xtriage
Anisotropy	0.478	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 58.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.008 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6638	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.13% 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: ACT, GOL, 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	0.28	0/1441	0.46	0/1978
2	B	0.29	0/1549	0.56	0/2107
3	C	0.28	0/85	0.50	0/115
4	D	0.28	0/1464	0.50	0/1991
5	E	0.30	0/1909	0.52	0/2598
All	All	0.29	0/6448	0.51	0/8789

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 ⓘ

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	1399	0	1312	9	0
2	B	1515	0	1464	10	0
3	C	84	0	95	0	0
4	D	1434	0	1292	12	0
5	E	1859	0	1740	11	0
6	A	14	0	13	0	0
6	B	14	0	13	0	0
7	B	8	0	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	E	8	0	6	0	0
8	B	6	0	8	0	0
9	A	63	0	0	0	0
9	B	83	0	0	1	0
9	C	3	0	0	0	0
9	D	59	0	0	0	0
9	E	89	0	0	0	0
All	All	6638	0	5949	38	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 3.

All (38) 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:91:VAL:HG21	1:A:167:VAL:HG21	1.60	0.82
5:E:75:ARG:NH1	5:E:98:ASP:OD2	2.17	0.76
1:A:37:ASP:OD2	1:A:40:ARG:NH1	2.18	0.75
2:B:37:ILE:HG13	2:B:38:VAL:HG23	1.72	0.71
2:B:116:VAL:HG22	2:B:160:MET:HG3	1.77	0.67
2:B:125:ALA:HB1	2:B:147:LEU:HD21	1.78	0.66
4:D:74:GLY:O	4:D:92:ARG:NH2	2.30	0.64
2:B:130:ARG:NH1	9:B:305:HOH:O	2.37	0.58
1:A:120:ASN:HB2	1:A:168:GLU:HB2	1.89	0.54
2:B:82:ASN:O	2:B:86:GLU:HG2	2.09	0.53
2:B:114:LEU:HD11	2:B:160:MET:HB3	1.91	0.53
4:D:56:VAL:HG11	4:D:87:ILE:HG23	1.91	0.52
5:E:139:VAL:HG23	5:E:249:ALA:HB3	1.92	0.52
5:E:105:ALA:HB1	5:E:115:LEU:HB3	1.92	0.51
4:D:173:VAL:HG22	4:D:184:ASN:OD1	2.13	0.48
4:D:161:SER:HB2	4:D:168:ILE:HD12	1.96	0.48
4:D:141:SER:O	4:D:143:ASP:N	2.47	0.48
4:D:146:VAL:HG12	4:D:189:TRP:HB3	1.96	0.47
4:D:174:LEU:HB3	5:E:183:CYS:HB2	1.96	0.47
1:A:10:SER:C	1:A:11(A):GLY:HA2	2.35	0.47
4:D:77:THR:HB	4:D:90:LEU:HB2	1.97	0.46
5:E:213:TRP:CZ2	5:E:253:GLY:HA2	2.50	0.46
1:A:13:ASN:HB2	2:B:11:PHE:HB3	1.97	0.46
1:A:158:SER:HB2	1:A:161:GLU:HG3	1.97	0.46
2:B:45:GLY:HA3	5:E:109:ARG:HD3	1.98	0.45
1:A:140:LEU:HD12	1:A:148:PHE:CE2	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:117:CYS:HB2	2:B:131:TRP:CZ2	2.53	0.44
1:A:167:VAL:CG2	1:A:176:LEU:HB3	2.49	0.43
1:A:72:LEU:HD13	2:B:9:TYR:HB2	2.01	0.43
5:E:50:LEU:HD23	5:E:50:LEU:HA	1.87	0.42
5:E:142:PRO:HD2	5:E:213:TRP:CZ2	2.55	0.42
4:D:7:ASP:HA	4:D:8:PRO:HD3	1.93	0.41
5:E:14:THR:HA	5:E:126:LEU:O	2.21	0.41
4:D:39:PHE:HB2	4:D:56:VAL:HB	2.02	0.41
5:E:158:LEU:HD23	5:E:158:LEU:HA	1.85	0.41
4:D:41:TRP:CE2	4:D:89:LEU:HB2	2.56	0.40
4:D:13:VAL:HG13	4:D:17:ALA:HB3	2.04	0.40
5:E:75:ARG:HH12	5:E:98:ASP:CG	2.18	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	180/183 (98%)	176 (98%)	4 (2%)	0	100	100
2	B	183/194 (94%)	178 (97%)	5 (3%)	0	100	100
3	C	10/13 (77%)	10 (100%)	0	0	100	100
4	D	188/204 (92%)	181 (96%)	6 (3%)	1 (0%)	25	32
5	E	238/242 (98%)	234 (98%)	4 (2%)	0	100	100
All	All	799/836 (96%)	779 (98%)	19 (2%)	1 (0%)	48	61

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	8	PRO

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	153/167 (92%)	153 (100%)	0	100	100
2	B	167/177 (94%)	166 (99%)	1 (1%)	84	91
3	C	9/9 (100%)	9 (100%)	0	100	100
4	D	146/185 (79%)	146 (100%)	0	100	100
5	E	197/211 (93%)	195 (99%)	2 (1%)	73	83
All	All	672/749 (90%)	669 (100%)	3 (0%)	89	94

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	28	SER
5	E	54	PHE
5	E	205	ARG

Sometimes 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 oligosaccharides 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	$\#\|Z\| > 2$
7	ACT	E	302	-	3,3,3	1.27	0	3,3,3	1.42	0
7	ACT	E	301	-	3,3,3	1.32	0	3,3,3	1.42	0
7	ACT	B	201	-	3,3,3	1.46	1 (33%)	3,3,3	1.34	0
8	GOL	B	204	-	5,5,5	0.82	0	5,5,5	1.31	1 (20%)
6	NAG	A	201	1	14,14,15	0.23	0	17,19,21	0.48	0
6	NAG	B	202	2	14,14,15	0.20	0	17,19,21	0.55	0
7	ACT	B	203	-	3,3,3	1.54	1 (33%)	3,3,3	1.27	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	B	202	2	-	0/6/23/26	0/1/1/1
8	GOL	B	204	-	-	4/4/4/4	-
6	NAG	A	201	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	203	ACT	CH3-C	2.27	1.58	1.49
7	B	201	ACT	CH3-C	2.17	1.57	1.49

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	B	204	GOL	C3-C2-C1	-2.27	103.49	111.80

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	204	GOL	O1-C1-C2-C3
8	B	204	GOL	C1-C2-C3-O3
8	B	204	GOL	O1-C1-C2-O2
8	B	204	GOL	O2-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	182/183 (99%)	-0.30	2 (1%)	77	79	22, 42, 62, 83	0
2	B	187/194 (96%)	-0.47	2 (1%)	77	79	22, 33, 53, 87	0
3	C	12/13 (92%)	0.13	2 (16%)	5	5	31, 35, 53, 78	0
4	D	194/204 (95%)	0.15	12 (6%)	28	28	22, 46, 87, 103	0
5	E	240/242 (99%)	-0.28	4 (1%)	69	70	21, 37, 69, 87	0
All	All	815/836 (97%)	-0.22	22 (2%)	56	59	21, 39, 73, 103	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	255	ALA	4.9
4	D	164	SER	3.5
4	D	203	ASN	3.5
4	D	207	PRO	3.4
4	D	206	ILE	3.4
4	D	141	SER	3.2
4	D	204	SER	2.9
4	D	143	ASP	2.9
5	E	197	ASP	2.9
1	A	1	GLU	2.9
2	B	105	ARG	2.8
4	D	199	ASN	2.7
3	C	10	GLY	2.7
1	A	126	ASN	2.6
4	D	201	PHE	2.6
5	E	253	GLY	2.5
2	B	112	HIS	2.3
4	D	163	ASP	2.2
3	C	-1	ALA	2.2
5	E	194	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	194	ASP	2.1
4	D	9	GLY	2.0

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

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. 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	B-factors(Å^2)	Q<0.9
6	NAG	B	202	14/15	0.67	0.18	61,75,86,95	0
6	NAG	A	201	14/15	0.73	0.15	48,75,91,92	0
7	ACT	B	203	4/4	0.73	0.20	56,56,56,56	0
8	GOL	B	204	6/6	0.82	0.19	48,48,48,48	0
7	ACT	B	201	4/4	0.83	0.15	37,51,57,59	0
7	ACT	E	302	4/4	0.87	0.15	47,47,60,60	0
7	ACT	E	301	4/4	0.90	0.12	45,47,55,62	0

6.5 Other polymers ⓘ

There are no such residues in this entry.