wwPDB X-ray Structure Validation Summary Report (i)

Dec 19, 2023 – 02:42 AM EST

PDB ID	:	2P5E
Title	:	Crystal Structures of High Affinity Human T-Cell Receptors Bound to pMHC Reveal Native Diagonal Binding Geometry
Authors	:	Sami, M.; Rizkallah, P.J.; Dunn, S.; Li, Y.; Moysey, R.; Vuidepot, A.; Baston, E.; Todorov, P.; Molloy, P.; Gao, F.; Boulter, J.M.; Jakobsen, B.K.
Deposited on	:	2007-03-15
Resolution	:	1.89 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
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 (1)) were used in the production of this report:

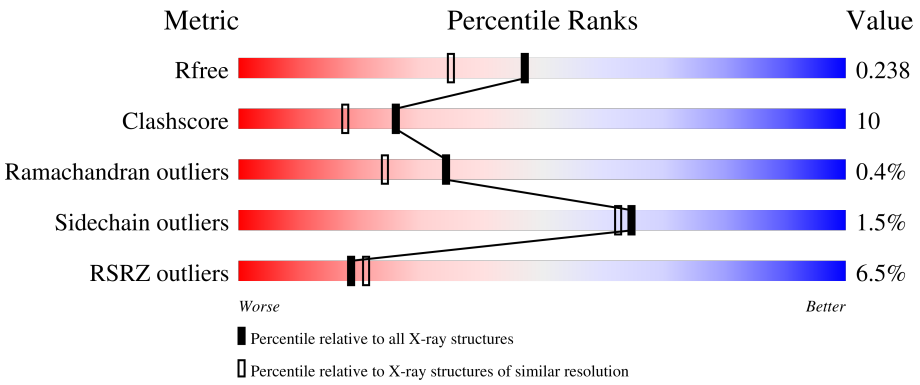
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.89 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	276	5% 82% 18%
2	B	100	3% 87% 12% •
3	C	9	67% 33%
4	D	195	16% 75% 23% • •
5	E	242	2% 80% 20%

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
9	IPA	C	909	-	-	X	-

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 7474 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 I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	276	2271	1423	412	427	9	0	2	0

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	100	853	542	145	161	5	0	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	insertion	UNP P61769
B	91	CYS	LYS	conflict	UNP P61769

- Molecule 3 is a protein called Cancer/testis antigen 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	9	85	59	11	13	2	0	2	0

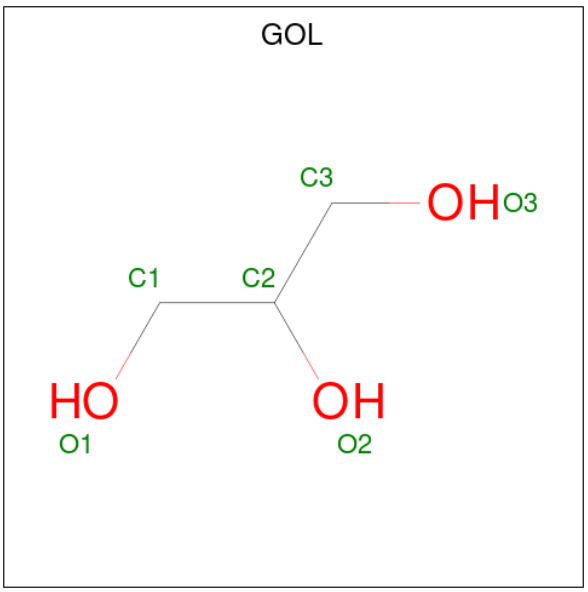
- Molecule 4 is a protein called T-Cell Receptor, Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	194	1498	936	251	302	9	0	1	0

- Molecule 5 is a protein called Hypothetical protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	241	1935	1222	330	372	11	0	4	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

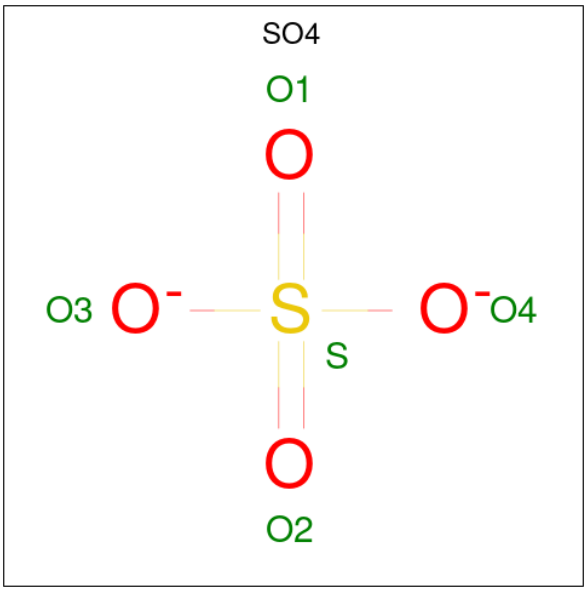


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	B	1	6	3	3	0	0
			Total	C	O		
6	E	1	6	3	3	0	0
			Total	C	O		
6	E	1	6	3	3	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

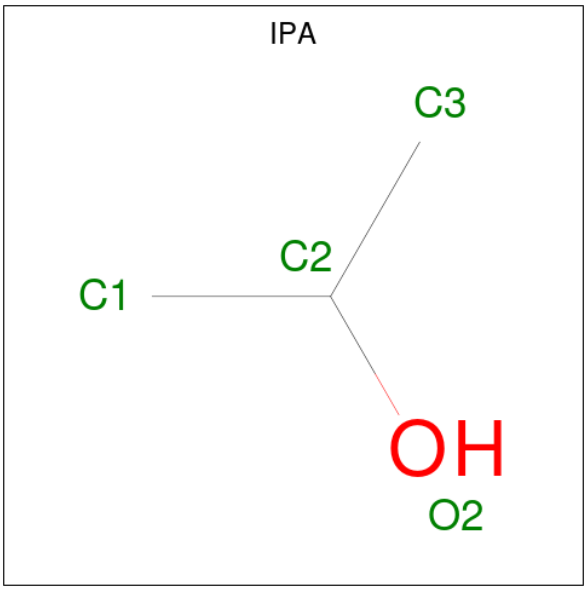
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
7	B	1	1	1	0	0

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



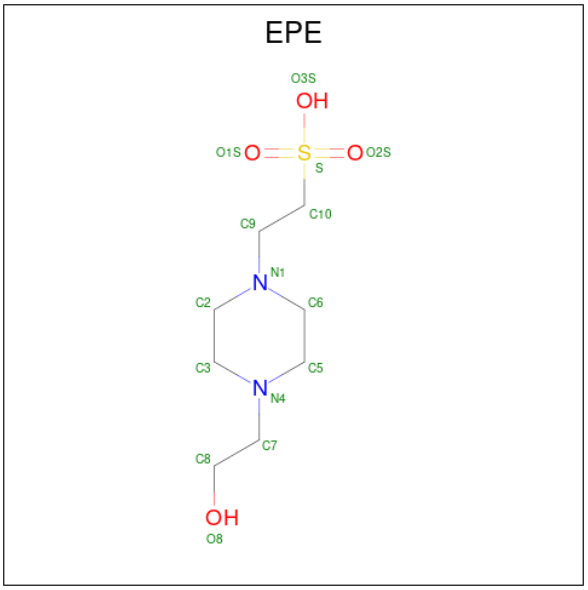
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total O S 5 4 1	0	0
8	D	1	Total O S 5 4 1	0	0
8	D	1	Total O S 5 4 1	0	0
8	E	1	Total O S 5 4 1	0	0
8	E	1	Total O S 5 4 1	0	0

- Molecule 9 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	C	1	4	3	1	0	0

- Molecule 10 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
10	E	1	15	8	2	4	1	0	0

- Molecule 11 is water.

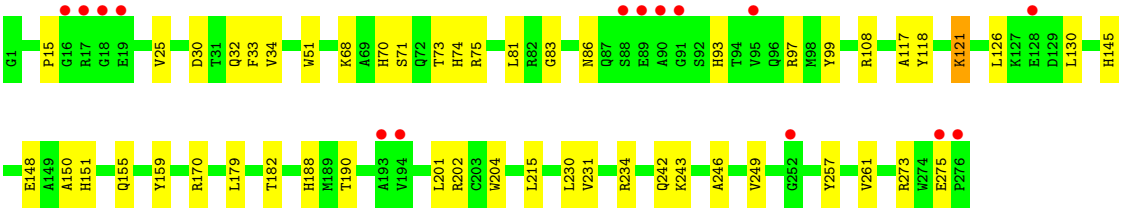
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
11	A	237	237	237	0	0
			Total	O		
11	B	120	120	120	0	0
			Total	O		
11	C	8	8	8	0	0
			Total	O		
11	D	121	121	121	0	0
			Total	O		
11	E	253	253	253	0	0

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.

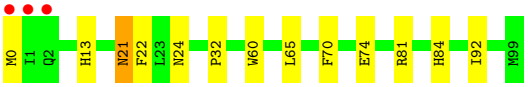
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain

Chain A:



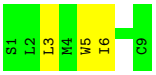
- Molecule 2: Beta-2-microglobulin

Chain B:



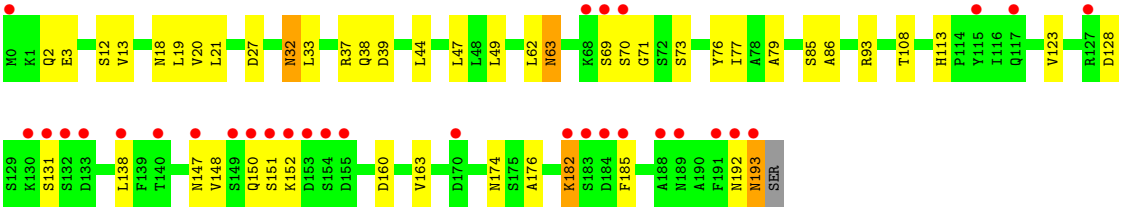
- Molecule 3: Cancer/testis antigen 1B

Chain C:



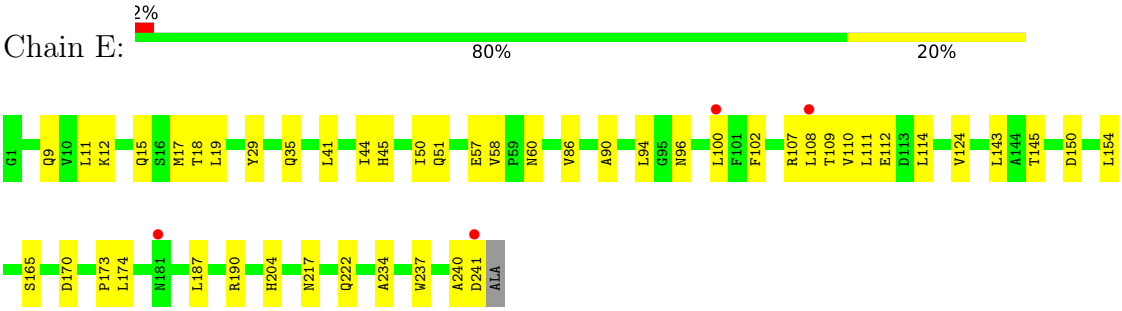
- Molecule 4: T-Cell Receptor, Alpha Chain

Chain D:



- Molecule 5: Hypothetical protein

Chain E:


2%
80%
20%
G1 Q9 V10 L11 K12 Q15 S16 M17 T18 L19 Y29 Q35 L41 I44 H45 I50 Q51 E57 V58 P59 N60 V86 A90 L94 G95 N96 L100 F101 F102 R107 L108 T109 V110 L111 E112 D113 L114 V124 L143 A144 T145 D150 L154
S165 D170 P173 L174 N181 L187 R190 H204 N217 Q222 A234 W237 A240 D241 ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.00Å 51.77Å 118.27Å 90.00° 98.24° 90.00°	Depositor
Resolution (Å)	117.04 – 1.89 23.41 – 1.89	Depositor EDS
% Data completeness (in resolution range)	97.4 (117.04-1.89) 97.4 (23.41-1.89)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.178 , 0.236 0.181 , 0.238	Depositor DCC
R_{free} test set	3552 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	22.9	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 46.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7474	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: EPE, IPA, MG, GOL, SO4

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	0.86	1/2342 (0.0%)	0.56	0/3182
2	B	0.79	0/876	0.58	0/1184
3	C	0.79	0/92	0.70	0/123
4	D	0.76	0/1532	0.60	0/2084
5	E	0.89	1/1992 (0.1%)	0.58	0/2713
All	All	0.84	2/6834 (0.0%)	0.58	0/9286

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	57	GLU	CB-CG	-5.94	1.40	1.52
1	A	204	TRP	CB-CG	-5.56	1.40	1.50

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	2271	0	2121	44	0
2	B	853	0	817	14	0
3	C	85	0	101	12	0
4	D	1498	0	1448	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	1935	0	1847	41	1
6	A	30	0	40	4	0
6	B	6	0	8	0	0
6	E	12	0	16	0	0
7	B	1	0	0	0	0
8	B	5	0	0	0	0
8	D	10	0	0	0	0
8	E	10	0	0	0	0
9	C	4	0	8	7	0
10	E	15	0	17	0	0
11	A	237	0	0	3	0
11	B	120	0	0	1	0
11	C	8	0	0	0	0
11	D	121	0	0	1	0
11	E	253	0	0	4	1
All	All	7474	0	6423	138	1

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 10.

The worst 5 of 138 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)
3:C:3[B]:LEU:HD11	9:C:909:IPA:H2	1.38	1.04
3:C:3[B]:LEU:CD1	9:C:909:IPA:H2	1.99	0.93
4:D:2:GLN:HE22	4:D:27:ASP:H	1.17	0.92
2:B:21:ASN:HD22	2:B:22:PHE:H	1.16	0.91
5:E:109:THR:HG22	5:E:111:LEU:HD11	1.52	0.89

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:217:ASN:OD1	11:E:960:HOH:O[2_545]	2.16	0.04

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	276/276 (100%)	271 (98%)	4 (1%)	1 (0%)	34	24
2	B	100/100 (100%)	97 (97%)	3 (3%)	0	100	100
3	C	9/9 (100%)	9 (100%)	0	0	100	100
4	D	193/195 (99%)	180 (93%)	11 (6%)	2 (1%)	15	6
5	E	243/242 (100%)	238 (98%)	5 (2%)	0	100	100
All	All	821/822 (100%)	795 (97%)	23 (3%)	3 (0%)	34	24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	70	SER
4	D	151	SER
1	A	15	PRO

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	234/232 (101%)	233 (100%)	1 (0%)	91	91
2	B	97/95 (102%)	95 (98%)	2 (2%)	53	48
3	C	11/9 (122%)	11 (100%)	0	100	100
4	D	172/172 (100%)	166 (96%)	6 (4%)	36	27
5	E	214/210 (102%)	212 (99%)	2 (1%)	78	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	728/718 (101%)	717 (98%)	11 (2%)	65 62

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	182	LYS
4	D	193	ASN
5	E	190	ARG
5	E	96	ASN
4	D	63	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
5	E	35	GLN
5	E	222	GLN
5	E	55	GLN
5	E	136	GLN
2	B	24	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 1 is monoatomic - leaving 15 for Mogul analysis.

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	GOL	A	902	-	5,5,5	0.45	0	5,5,5	0.28	0
6	GOL	A	903	-	5,5,5	0.37	0	5,5,5	0.39	0
8	SO4	D	802	-	4,4,4	0.18	0	6,6,6	0.16	0
6	GOL	B	901	-	5,5,5	0.41	0	5,5,5	0.30	0
8	SO4	E	803	-	4,4,4	0.07	0	6,6,6	0.11	0
6	GOL	A	906	-	5,5,5	0.31	0	5,5,5	0.23	0
8	SO4	B	805	-	4,4,4	0.20	0	6,6,6	0.05	0
6	GOL	A	904	-	5,5,5	0.56	0	5,5,5	0.29	0
8	SO4	E	804	-	4,4,4	0.15	0	6,6,6	0.06	0
9	IPA	C	909	-	3,3,3	0.50	0	3,3,3	0.22	0
10	EPE	E	806	-	15,15,15	0.74	1 (6%)	18,20,20	1.19	2 (11%)
6	GOL	A	907	-	5,5,5	0.43	0	5,5,5	0.33	0
6	GOL	E	908	-	5,5,5	0.40	0	5,5,5	0.35	0
8	SO4	D	801	-	4,4,4	0.14	0	6,6,6	0.11	0
6	GOL	E	905	-	5,5,5	0.40	0	5,5,5	0.26	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	GOL	A	902	-	-	2/4/4/4	-
6	GOL	A	903	-	-	2/4/4/4	-
6	GOL	B	901	-	-	2/4/4/4	-
6	GOL	A	906	-	-	2/4/4/4	-
10	EPE	E	806	-	-	2/9/19/19	0/1/1/1
6	GOL	A	904	-	-	2/4/4/4	-
6	GOL	A	907	-	-	3/4/4/4	-
6	GOL	E	908	-	-	2/4/4/4	-
6	GOL	E	905	-	-	3/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	E	806	EPE	C10-S	2.56	1.81	1.77

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	E	806	EPE	O2S-S-C10	2.38	109.78	106.92
10	E	806	EPE	O1S-S-C10	2.12	109.47	106.92

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	902	GOL	O1-C1-C2-O2
6	A	902	GOL	O1-C1-C2-C3
6	A	903	GOL	C1-C2-C3-O3
6	A	907	GOL	C1-C2-C3-O3
6	E	905	GOL	O1-C1-C2-C3

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	903	GOL	1	0
6	A	904	GOL	2	0
9	C	909	IPA	7	0
6	A	907	GOL	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	276/276 (100%)	0.24	15 (5%) 25 29	5, 13, 28, 44	0
2	B	100/100 (100%)	0.17	3 (3%) 50 53	5, 14, 27, 55	0
3	C	9/9 (100%)	0.55	0 100 100	6, 7, 9, 12	0
4	D	194/195 (99%)	0.74	31 (15%) 1 2	4, 13, 41, 64	0
5	E	241/242 (99%)	0.05	4 (1%) 70 72	5, 12, 23, 46	0
All	All	820/822 (99%)	0.30	53 (6%) 18 21	4, 13, 29, 64	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	ILE	7.8
4	D	0	MET	7.5
4	D	193	ASN	7.0
4	D	70	SER	6.0
2	B	0	MET	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)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 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	GOL	A	902	6/6	0.80	0.24	23,31,33,34	6
8	SO4	E	804	5/5	0.81	0.31	75,78,84,85	0
9	IPA	C	909	4/4	0.81	0.18	2,2,12,15	4
6	GOL	A	906	6/6	0.84	0.18	24,29,34,37	0
6	GOL	A	907	6/6	0.88	0.14	35,39,40,45	0
8	SO4	B	805	5/5	0.88	0.27	70,71,75,75	0
6	GOL	A	904	6/6	0.89	0.15	24,27,31,33	0
6	GOL	E	908	6/6	0.91	0.28	19,30,33,44	0
6	GOL	B	901	6/6	0.92	0.15	14,29,30,36	0
6	GOL	E	905	6/6	0.92	0.12	17,31,34,36	0
6	GOL	A	903	6/6	0.92	0.09	19,30,31,32	0
8	SO4	D	802	5/5	0.94	0.21	2,8,24,31	5
8	SO4	E	803	5/5	0.95	0.30	43,45,50,54	0
8	SO4	D	801	5/5	0.96	0.19	37,46,48,60	0
7	MG	B	800	1/1	0.98	0.15	23,23,23,23	0
10	EPE	E	806	15/15	0.98	0.21	24,38,46,49	0

6.5 Other polymers (i)

There are no such residues in this entry.