wwPDB X-ray Structure Validation Summary Report (i)

Oct 9, 2023 – 04:18 PM EDT

PDB ID	:	8D5N
Title	:	Crystal structure of Ld-HF10
Authors	:	Wang, Y.; Dai, S.
Deposited on	:	2022-06-05
Resolution	:	1.80 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 (i)) were used in the production of this report:

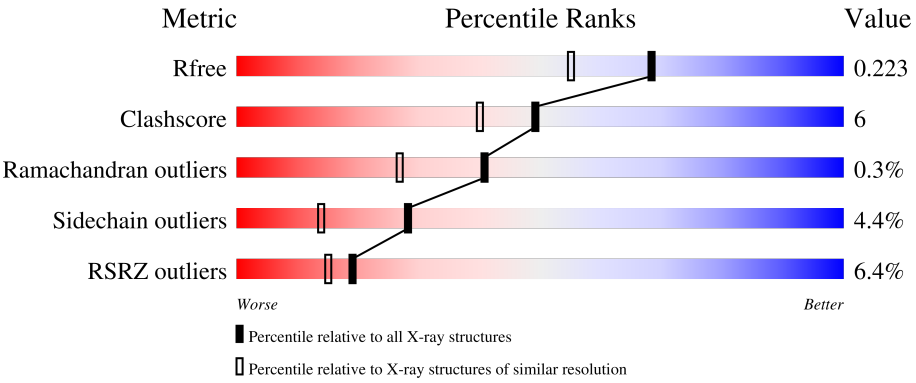
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.80 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.



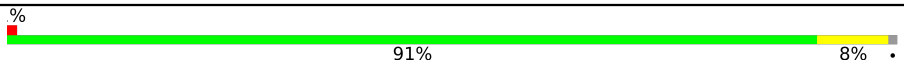
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	282	8% 83% 13%
1	C	282	9% 78% 12% 7%
2	B	16	69% 6% 25%
2	E	16	88% 6% 6%
3	F	104	84% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	104	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	EDO	F	101	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 6715 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called H-2 class I histocompatibility antigen, L-D alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	262	Total	C	N	O	S	0	2	0
			2149	1362	376	401	10			
1	A	275	Total	C	N	O	S	0	0	0
			2238	1417	390	421	10			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	84	CYS	TYR	conflict	UNP P01897
C	121	SER	CYS	conflict	UNP P01897
C	275	GLN	-	expression tag	UNP P01897
C	276	SER	-	expression tag	UNP P01897
C	277	THR	-	expression tag	UNP P01897
C	278	ARG	-	expression tag	UNP P01897
C	279	GLY	-	expression tag	UNP P01897
C	280	GLY	-	expression tag	UNP P01897
C	281	ALA	-	expression tag	UNP P01897
C	282	SER	-	expression tag	UNP P01897
A	84	CYS	TYR	conflict	UNP P01897
A	121	SER	CYS	conflict	UNP P01897
A	275	GLN	-	expression tag	UNP P01897
A	276	SER	-	expression tag	UNP P01897
A	277	THR	-	expression tag	UNP P01897
A	278	ARG	-	expression tag	UNP P01897
A	279	GLY	-	expression tag	UNP P01897
A	280	GLY	-	expression tag	UNP P01897
A	281	ALA	-	expression tag	UNP P01897
A	282	SER	-	expression tag	UNP P01897

- Molecule 2 is a protein called Dense granule protein 6, HF10 peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	15	Total	C	N	O	S	0	0	0
			105	64	18	22	1			
2	B	12	Total	C	N	O	S	0	0	0
			91	57	15	18	1			

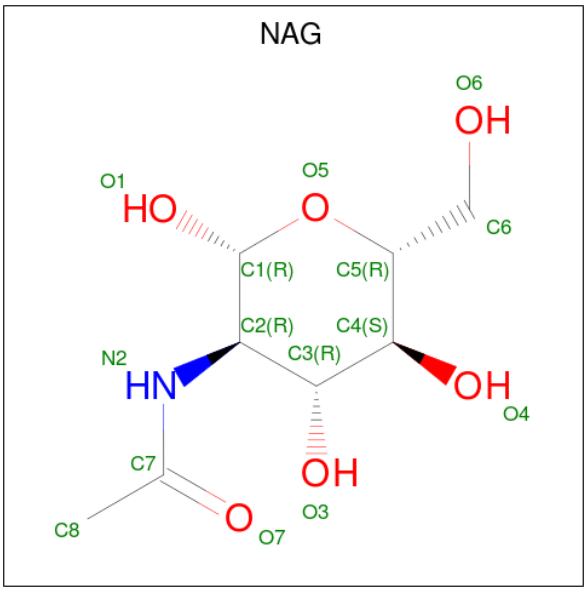
- Molecule 3 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	102	Total	C	N	O	S	0	0	0
			832	530	141	154	7			
3	H	103	Total	C	N	O	S	0	1	0
			845	537	144	157	7			

There are 10 discrepancies between the modelled and reference sequences:

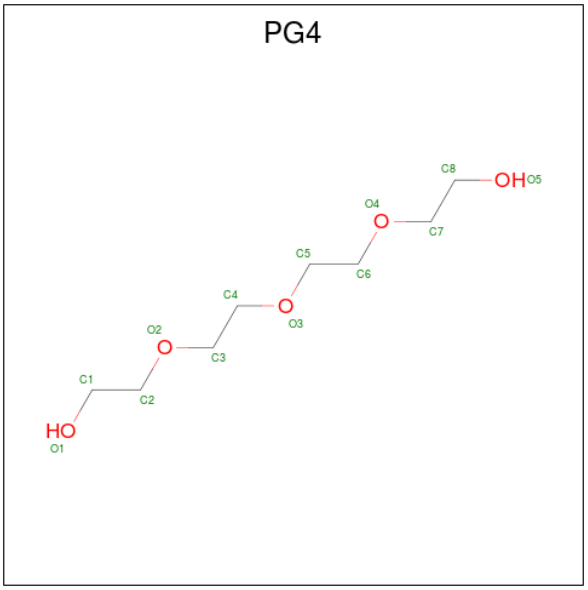
Chain	Residue	Modelled	Actual	Comment	Reference
F	-4	GLY	-	expression tag	UNP P01887
F	-3	GLY	-	expression tag	UNP P01887
F	-2	SER	-	expression tag	UNP P01887
F	-1	GLY	-	expression tag	UNP P01887
F	0	GLY	-	expression tag	UNP P01887
H	-4	GLY	-	expression tag	UNP P01887
H	-3	GLY	-	expression tag	UNP P01887
H	-2	SER	-	expression tag	UNP P01887
H	-1	GLY	-	expression tag	UNP P01887
H	0	GLY	-	expression tag	UNP P01887

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



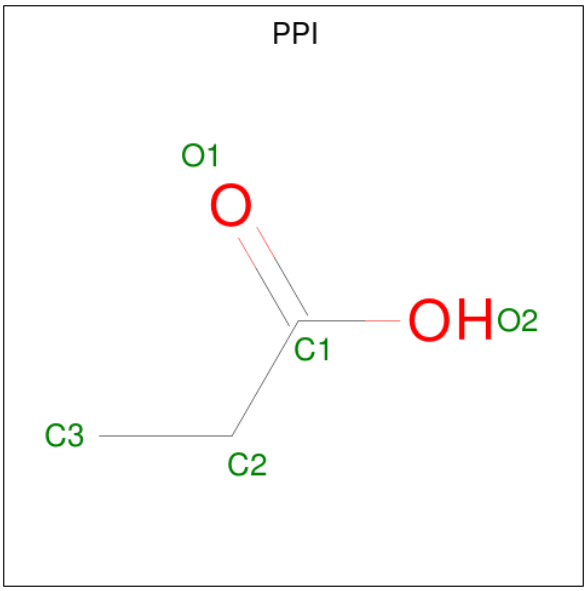
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



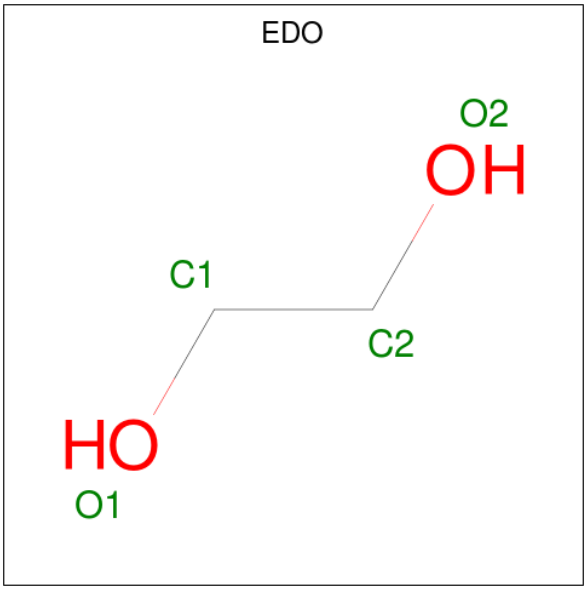
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			13	8	5		

- Molecule 6 is PROPANOIC ACID (three-letter code: PPI) (formula: $C_3H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	1	Total C O 5 3 2	0	0

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	1	Total C O 4 2 2	0	0
7	H	1	Total C O 4 2 2	0	0

- Molecule 8 is water.

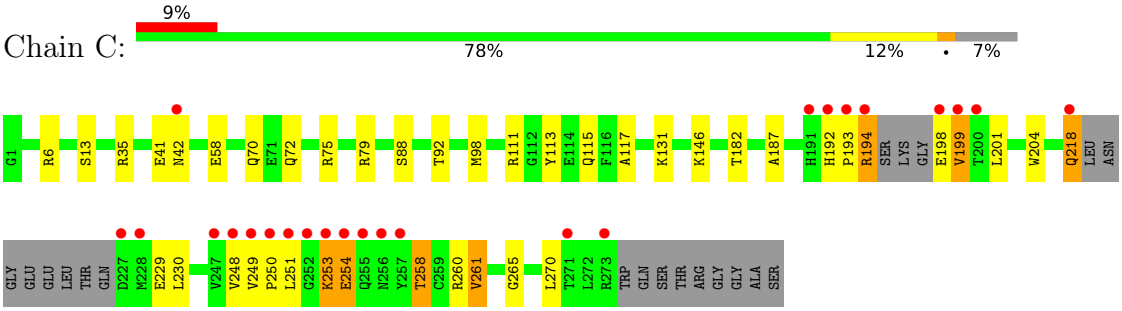
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	141	Total O 141 141	0	0
8	E	9	Total O 9 9	0	0
8	A	127	Total O 127 127	0	0
8	B	14	Total O 14 14	0	0
8	F	62	Total O 62 62	0	0
8	H	48	Total O 48 48	0	0

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.

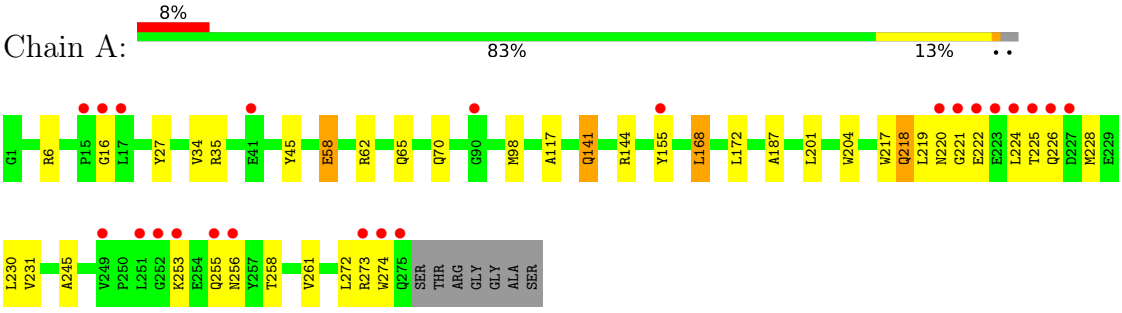
• Molecule 1: H-2 class I histocompatibility antigen, L-D alpha chain

Chain C:



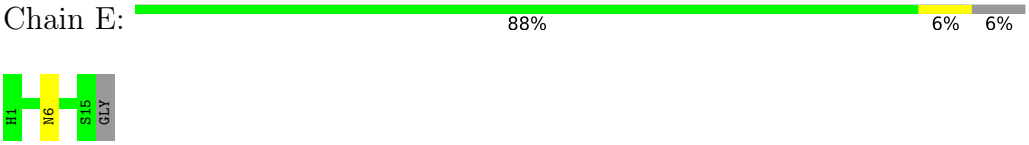
• Molecule 1: H-2 class I histocompatibility antigen, L-D alpha chain

Chain A:



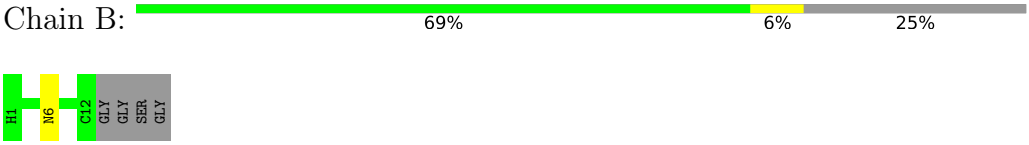
• Molecule 2: Dense granule protein 6, HF10 peptide

Chain E:



• Molecule 2: Dense granule protein 6, HF10 peptide

Chain B:

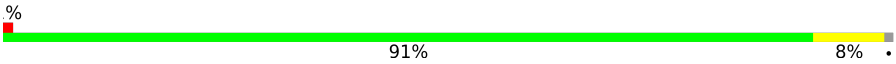


• Molecule 3: Beta-2-microglobulin

Chain F:

• Molecule 3: Beta-2-microglobulin

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	121.50Å 139.82Å 87.34Å 90.00° 130.82° 90.00°	Depositor
Resolution (Å)	50.01 – 1.80 48.03 – 1.78	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.01-1.80) 97.9 (48.03-1.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.178 , 0.219 0.186 , 0.223	Depositor DCC
R_{free} test set	5160 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	34.7	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 45.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6715	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% 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: PG4, EDO, NAG, PPI

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	0.92	1/2304 (0.0%)	1.08	3/3134 (0.1%)
1	C	0.96	2/2214 (0.1%)	1.11	4/3010 (0.1%)
2	B	0.98	0/94	0.97	0/126
2	E	0.98	0/108	1.01	0/144
3	F	1.02	3/858 (0.3%)	1.09	1/1162 (0.1%)
3	H	0.92	0/871	1.07	0/1179
All	All	0.95	6/6449 (0.1%)	1.09	8/8755 (0.1%)

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	C	0	1
All	All	0	2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	55	SER	CA-CB	-6.13	1.43	1.52
3	F	15	PRO	C-O	-5.71	1.11	1.23
3	F	74	GLU	CD-OE2	5.39	1.31	1.25
1	C	88	SER	CA-CB	-5.24	1.45	1.52
1	A	58	GLU	CD-OE1	5.23	1.31	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	35	ARG	NE-CZ-NH1	9.01	124.80	120.30
1	C	6	ARG	NE-CZ-NH1	-8.76	115.92	120.30
1	C	6	ARG	CG-CD-NE	-7.08	96.93	111.80
1	A	27	TYR	CB-CG-CD1	-6.85	116.89	121.00
1	A	6	ARG	CG-CD-NE	-5.50	100.24	111.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	221	GLY	Peptide
1	C	131	LYS	Mainchain

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	2238	0	2101	20	0
1	C	2149	0	2024	37	0
2	B	91	0	74	1	0
2	E	105	0	85	3	0
3	F	832	0	808	14	0
3	H	845	0	818	3	0
4	A	14	0	13	0	0
4	C	14	0	13	0	0
5	C	13	0	18	1	0
6	E	5	0	5	0	0
7	F	4	0	6	10	0
7	H	4	0	6	0	0
8	A	127	0	0	0	0
8	B	14	0	0	2	0
8	C	141	0	0	3	0
8	E	9	0	0	0	0
8	F	62	0	0	0	0
8	H	48	0	0	0	0
All	All	6715	0	5971	72	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 6.

The worst 5 of 72 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:F:54:MET:HA	7:F:101:EDO:H12	1.38	1.04
1:A:70:GLN:HE22	2:B:6:ASN:H	1.14	0.91
3:F:54:MET:HG3	7:F:101:EDO:C1	1.99	0.91
1:C:198:GLU:CA	1:C:251:LEU:HD11	2.02	0.89
3:F:54:MET:CA	7:F:101:EDO:H12	2.05	0.86

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	273/282 (97%)	263 (96%)	8 (3%)	2 (1%)	22 10
1	C	258/282 (92%)	251 (97%)	7 (3%)	0	100 100
2	B	10/16 (62%)	10 (100%)	0	0	100 100
2	E	13/16 (81%)	12 (92%)	1 (8%)	0	100 100
3	F	100/104 (96%)	96 (96%)	4 (4%)	0	100 100
3	H	102/104 (98%)	101 (99%)	1 (1%)	0	100 100
All	All	756/804 (94%)	733 (97%)	21 (3%)	2 (0%)	41 27

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	16	GLY
1	A	226	GLN

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	229/234 (98%)	219 (96%)	10 (4%)	28 14
1	C	221/234 (94%)	210 (95%)	11 (5%)	24 10
2	B	10/11 (91%)	10 (100%)	0	100 100
2	E	11/11 (100%)	11 (100%)	0	100 100
3	F	94/94 (100%)	89 (95%)	5 (5%)	22 9
3	H	95/94 (101%)	92 (97%)	3 (3%)	39 25
All	All	660/678 (97%)	631 (96%)	29 (4%)	28 14

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	201	LEU
3	H	70	PHE
1	A	225	THR
3	F	70	PHE
1	A	219	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	ASN
1	A	42	ASN
2	B	6	ASN
1	A	149	GLN
1	A	256	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)

6 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	$\#\|Z\| > 2$
4	NAG	A	301	1	14,14,15	0.82	0	17,19,21	1.59	2 (11%)
6	PPI	E	101	-	4,4,4	1.03	0	4,4,4	0.95	0
4	NAG	C	301	1	14,14,15	0.63	0	17,19,21	1.45	3 (17%)
7	EDO	F	101	-	3,3,3	0.68	0	2,2,2	0.98	0
5	PG4	C	302	-	12,12,12	0.24	0	11,11,11	0.43	0
7	EDO	H	101	-	3,3,3	0.12	0	2,2,2	0.17	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
4	NAG	A	301	1	-	0/6/23/26	0/1/1/1
6	PPI	E	101	-	-	1/2/2/2	-
4	NAG	C	301	1	-	2/6/23/26	0/1/1/1
7	EDO	F	101	-	-	0/1/1/1	-
5	PG4	C	302	-	-	6/10/10/10	-
7	EDO	H	101	-	-	0/1/1/1	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	301	NAG	C1-O5-C5	5.01	118.97	112.19
4	C	301	NAG	C2-N2-C7	-3.26	118.26	122.90
4	A	301	NAG	C4-C3-C2	-2.33	107.61	111.02
4	C	301	NAG	C3-C4-C5	-2.12	106.46	110.24
4	C	301	NAG	C1-O5-C5	2.02	114.93	112.19

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	302	PG4	O3-C5-C6-O4
5	C	302	PG4	C4-C3-O2-C2
5	C	302	PG4	C8-C7-O4-C6
5	C	302	PG4	C3-C4-O3-C5
5	C	302	PG4	C6-C5-O3-C4

There are no ring outliers.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	101	EDO	10	0
5	C	302	PG4	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

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	275/282 (97%)	0.05	23 (8%) 11 8	26, 41, 73, 99	0
1	C	262/282 (92%)	0.12	24 (9%) 9 6	23, 37, 83, 107	0
2	B	12/16 (75%)	-0.29	0 100 100	34, 37, 47, 56	0
2	E	15/16 (93%)	-0.29	0 100 100	27, 31, 49, 67	0
3	F	102/104 (98%)	-0.31	1 (0%) 82 80	26, 36, 52, 69	0
3	H	103/104 (99%)	-0.17	1 (0%) 82 80	25, 37, 58, 74	0
All	All	769/804 (95%)	-0.02	49 (6%) 19 15	23, 38, 71, 107	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	225	THR	7.1
1	A	17	LEU	6.0
1	C	253	LYS	5.5
1	C	248	VAL	5.3
1	C	251	LEU	5.1

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 (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
4	NAG	C	301	14/15	0.78	0.16	54,66,78,82	0
4	NAG	A	301	14/15	0.81	0.16	53,61,71,78	0
6	PPI	E	101	5/5	0.86	0.12	45,52,58,59	0
5	PG4	C	302	13/13	0.93	0.15	34,45,52,60	0
7	EDO	F	101	4/4	0.93	0.15	41,41,42,49	0
7	EDO	H	101	4/4	0.93	0.24	53,58,59,61	0

6.5 Other polymers (i)

There are no such residues in this entry.