Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 11:42 PM EDT

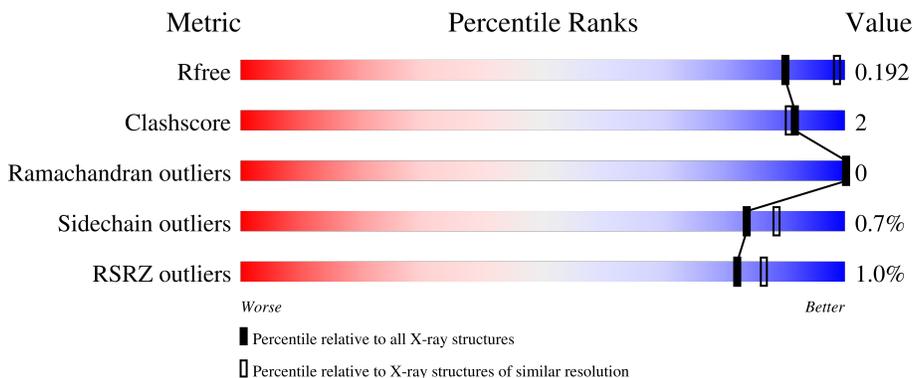|		
PDB ID	:	7MJ5
Title	:	complex of human thrombin with XC-43
Authors	:	Lu, S.; Tirloni, L.; Andersen, J.F.
Deposited on	:	2021-04-19
Resolution	:	2.15 Å(reported)

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 2.15 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.
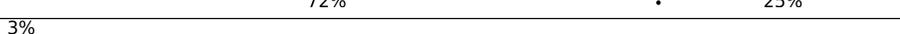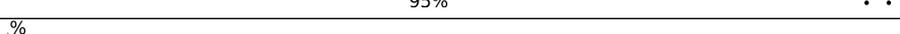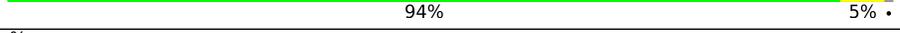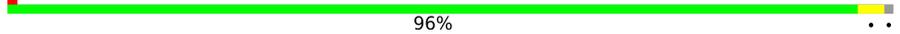
Metric	Percentile Ranks	Value
Rfree		0.192
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.7%
RSRZ outliers		1.0%

Worse ——— *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	36	86% 14%
1	N	36	83% 17%
1	O	36	81% • 17%
1	P	36	81% 19%
1	Q	36	78% 6% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	R	36	81% · 17%
2	B	36	8% 86% 14%
2	D	36	3% 69% 6% 25%
2	F	36	72% · 25%
2	I	36	3% 81% 19%
2	K	36	6% 75% 25%
2	L	36	3% 78% 6% 17%
3	C	259	93% 6% ·
3	E	259	% 95% · ·
3	G	259	% 94% 5% ·
3	H	259	% 96% · ·
3	J	259	95% · ·
3	M	259	93% 6%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 16237 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 Putative secreted salivary protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	31	Total	C	N	O	S	0	0	0
			233	143	40	49	1			
1	N	30	Total	C	N	O	S	0	0	0
			228	140	39	48	1			
1	O	30	Total	C	N	O	S	0	0	0
			228	140	39	48	1			
1	P	29	Total	C	N	O	S	0	0	0
			218	134	37	46	1			
1	Q	30	Total	C	N	O	S	0	0	0
			224	138	39	46	1			
1	R	30	Total	C	N	O	S	0	0	0
			228	140	39	48	1			

- Molecule 2 is a protein called Thrombin light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	30	Total	C	N	O	S	0	0	0
			235	147	35	52	1			
2	B	31	Total	C	N	O	S	0	0	0
			241	150	37	53	1			
2	D	27	Total	C	N	O	S	0	0	0
			208	131	31	45	1			
2	F	27	Total	C	N	O	S	0	0	0
			220	139	35	45	1			
2	I	29	Total	C	N	O	S	0	0	0
			231	145	38	47	1			
2	K	27	Total	C	N	O	S	0	0	0
			214	137	33	43	1			

- Molecule 3 is a protein called Thrombin heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	257	Total 2048	C 1305	N 363	O 366	S 14	0	0	0
3	C	257	Total 2056	C 1310	N 365	O 367	S 14	0	0	0
3	E	256	Total 2030	C 1291	N 358	O 367	S 14	0	0	0
3	G	257	Total 2029	C 1297	N 350	O 368	S 14	0	0	0
3	J	257	Total 2049	C 1305	N 362	O 368	S 14	0	0	0
3	M	258	Total 2064	C 1314	N 366	O 370	S 14	0	0	0

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

NAG

O6 OH
O1 HO
O5 O
C6
C1(R) C5(R)
C2(R) C4(S)
C3(R)
N2 HN
OH
O4
C7 OH
O3
C8 O
O7
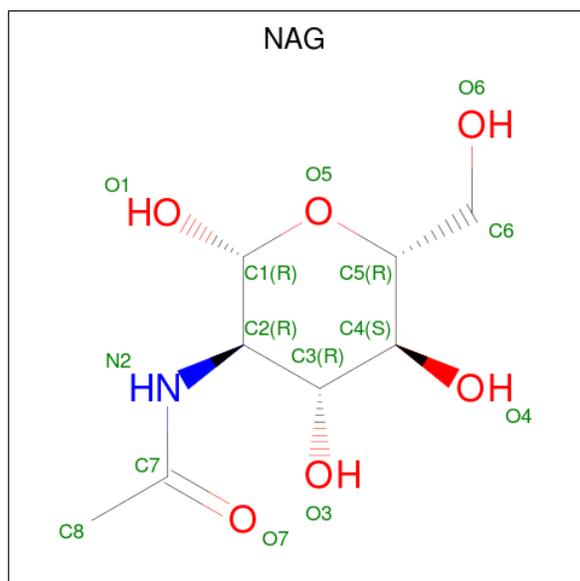
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	H	1	Total 14	C 8	N 1	O 5	0	0
4	C	1	Total 14	C 8	N 1	O 5	0	0
4	E	1	Total 14	C 8	N 1	O 5	0	0
4	G	1	Total 14	C 8	N 1	O 5	0	0
4	J	1	Total 14	C 8	N 1	O 5	0	0

Continued on next page...

W O R L D W I D E
PDB
PROTEIN DATA BANK

Continued from previous page...

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

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	1	Total 1	Na 1	0	0
5	C	1	Total 1	Na 1	0	0
5	E	1	Total 1	Na 1	0	0
5	G	1	Total 1	Na 1	0	0
5	J	1	Total 1	Na 1	0	0
5	M	1	Total 1	Na 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	26	Total 26	O 26	0	0
6	L	21	Total 21	O 21	0	0
6	H	166	Total 166	O 166	0	0
6	B	13	Total 13	O 13	0	0
6	C	193	Total 193	O 193	0	0
6	D	9	Total 9	O 9	0	0
6	E	151	Total 151	O 151	0	0
6	F	11	Total 11	O 11	0	0
6	G	144	Total 144	O 144	0	0
6	I	7	Total 7	O 7	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	154	Total 154	O 154	0	0
6	K	11	Total 11	O 11	0	0
6	M	149	Total 149	O 149	0	0
6	N	21	Total 21	O 21	0	0
6	O	21	Total 21	O 21	0	0
6	P	25	Total 25	O 25	0	0
6	Q	20	Total 20	O 20	0	0
6	R	21	Total 21	O 21	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: Putative secreted salivary protein

Chain A:

86%   14%

LYS PRO VAL GLU A5 N35 ALA

• Molecule 1: Putative secreted salivary protein

Chain N:

83%   17%

LYS PRO VAL GLU ALA E6 N35 ALA

• Molecule 1: Putative secreted salivary protein

Chain O:

81%   ·   17%

LYS PRO VAL GLU ALA E6 L34 N35 ALA

• Molecule 1: Putative secreted salivary protein

Chain P:

81%   19%

LYS PRO VAL GLU ALA E6 L34 ASN ALA

• Molecule 1: Putative secreted salivary protein

Chain Q:

78%   6%   17%

LYS PRO VAL GLU ALA E6 P24 I25 N35 ALA

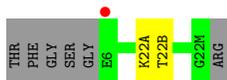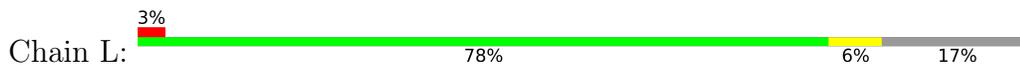• Molecule 1: Putative secreted salivary protein
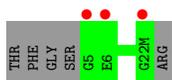
Chain R:

81%   ·   17%

• Molecule 2: Thrombin light chain
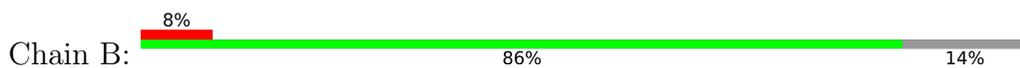
Chain L:

• Molecule 2: Thrombin light chain

Chain B:

• Molecule 2: Thrombin light chain

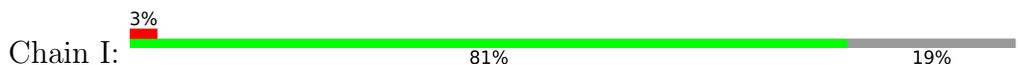Chain D:

• Molecule 2: Thrombin light chain

Chain F:

• Molecule 2: Thrombin light chain

Chain I:

• Molecule 2: Thrombin light chain

Chain K:

• Molecule 3: Thrombin heavy chain

Chain H:

%

96%

I16 E61 K87 T139 K145 A149A N149B V157 C182 G211 T229 F245 GLY GLU

• Molecule 3: Thrombin heavy chain

Chain C:

93% 6%

I16 W29 L60 E61 K87 V121 T139 K145 A149A V157 C182 K185 P186 D186A M201 M210 G211 T230 F246 GLY GLU

• Molecule 3: Thrombin heavy chain

Chain E:

%

95%

I16 E61 K81 K87 L105 I118 T149 A149A N149B V154 C182 D194 I239 Q245 PHE GLY GLU

• Molecule 3: Thrombin heavy chain

Chain G:

%

94% 5%

I16 W29 K36 L53 L65 I103 H119 P120 V121 L130 K145 A149A I162 C182 G211 T230 R234 V242 F246 GLY GLU

• Molecule 3: Thrombin heavy chain

Chain J:

95%

I16 Q38 L53 R93 R101 D102 I103 A104 L105 K109 K145 N149B C182 F204A I239 F246 GLY GLU

• Molecule 3: Thrombin heavy chain

Chain M:

93% 6%

I16 V31 K36 L53 V66 M84 K87 I88 Y89 I103 P124 V157 C182 F204A G211 K225 T230 F233 R234 G247 GLU
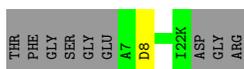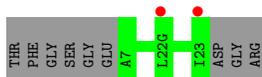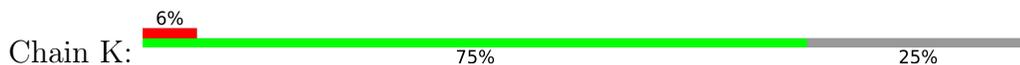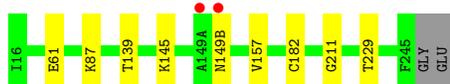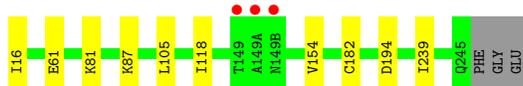
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	113.35Å 136.32Å 193.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.24 – 2.15 58.41 – 2.15	Depositor EDS
% Data completeness (in resolution range)	97.3 (44.24-2.15) 97.3 (58.41-2.15)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.30 (at 2.14Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.165 , 0.193 0.164 , 0.192	Depositor DCC
R_{free} test set	8016 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	35.6	Xtriage
Anisotropy	0.568	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.1	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	0.96	EDS
Total number of atoms	16237	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.16% 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: NA, 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	0/236	0.48	0/319
1	N	0.29	0/231	0.51	0/312
1	O	0.28	0/231	0.52	0/312
1	P	0.28	0/221	0.50	0/298
1	Q	0.30	0/227	0.49	0/307
1	R	0.29	0/231	0.50	0/312
2	B	0.30	0/243	0.50	0/324
2	D	0.25	0/210	0.51	0/282
2	F	0.27	0/222	0.54	0/296
2	I	0.27	0/233	0.54	0/310
2	K	0.29	0/216	0.53	0/288
2	L	0.29	0/237	0.52	0/317
3	C	0.30	0/2108	0.56	0/2851
3	E	0.30	0/2081	0.56	0/2819
3	G	0.28	0/2081	0.54	0/2820
3	H	0.30	0/2100	0.57	0/2841
3	J	0.30	0/2101	0.55	0/2844
3	M	0.30	0/2117	0.56	0/2864
All	All	0.29	0/15326	0.55	0/20716

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	233	0	221	0	0
1	N	228	0	216	0	0
1	O	228	0	216	1	0
1	P	218	0	203	0	0
1	Q	224	0	212	2	0
1	R	228	0	216	0	0
2	B	241	0	227	0	0
2	D	208	0	186	1	0
2	F	220	0	215	1	0
2	I	231	0	227	0	0
2	K	214	0	211	0	0
2	L	235	0	217	1	0
3	C	2056	0	2026	8	0
3	E	2030	0	1974	4	0
3	G	2029	0	1968	8	0
3	H	2048	0	2004	4	0
3	J	2049	0	2008	11	0
3	M	2064	0	2022	8	0
4	C	14	0	13	1	0
4	E	14	0	13	0	0
4	G	14	0	13	0	0
4	H	14	0	13	0	0
4	J	14	0	13	0	0
4	M	14	0	13	0	0
5	C	1	0	0	0	0
5	E	1	0	0	0	0
5	G	1	0	0	0	0
5	H	1	0	0	0	0
5	J	1	0	0	0	0
5	M	1	0	0	0	0
6	A	26	0	0	0	0
6	B	13	0	0	0	0
6	C	193	0	0	0	0
6	D	9	0	0	0	0
6	E	151	0	0	0	0
6	F	11	0	0	0	0
6	G	144	0	0	1	0
6	H	166	0	0	0	0
6	I	7	0	0	0	0
6	J	154	0	0	0	0
6	K	11	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	L	21	0	0	0	0
6	M	149	0	0	1	1
6	N	21	0	0	0	1
6	O	21	0	0	0	0
6	P	25	0	0	0	0
6	Q	20	0	0	0	0
6	R	21	0	0	0	0
All	All	16237	0	14647	46	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 2.

All (46) 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:H:139:THR:HG22	3:H:157:VAL:HG22	1.70	0.73
3:J:38:GLN:HE22	1:Q:25:ILE:H	1.37	0.73
2:L:22(A):LYS:HG3	2:L:22(B):THR:HG23	1.76	0.66
3:C:139:THR:HG22	3:C:157:VAL:HG22	1.78	0.64
3:E:105:LEU:HD11	3:E:239:ILE:HG23	1.84	0.59
3:E:61:GLU:HG2	3:E:87:LYS:HA	1.85	0.58
3:M:31:VAL:HG13	3:M:66:VAL:HG13	1.84	0.58
3:J:93:ARG:HH11	3:J:93:ARG:HG3	1.71	0.54
3:J:105:LEU:HD11	3:J:239:ILE:HG23	1.91	0.53
3:J:93:ARG:HB2	3:J:101:ARG:HD3	1.93	0.51
3:M:87:LYS:HG2	3:M:89:TYR:CZ	2.45	0.51
3:C:60:LEU:HD11	4:C:301:NAG:H82	1.93	0.49
3:C:185:LYS:HE2	3:C:186(A):ASP:OD1	2.13	0.49
3:G:36:LYS:HG2	3:G:65:LEU:HD23	1.95	0.48
3:G:53:LEU:HD11	3:G:103:ILE:HD11	1.96	0.48
3:J:145:LYS:HE3	3:J:149(B):ASN:HB3	1.96	0.48
3:M:53:LEU:HD11	3:M:103:ILE:HD11	1.95	0.47
3:M:31:VAL:CG1	3:M:66:VAL:HG13	2.43	0.47
3:J:38:GLN:NE2	1:Q:24:PRO:HA	2.29	0.46
3:J:53:LEU:HD11	3:J:103:ILE:HD11	1.96	0.46
3:H:61:GLU:HG2	3:H:87:LYS:HA	1.97	0.46
3:M:124:PRO:HG2	3:M:233:PHE:HB2	1.98	0.45
3:G:130:LEU:HD12	3:G:162:ILE:HD13	1.98	0.45
3:G:119:HIS:HD2	6:G:507:HOH:O	1.99	0.45
3:M:211:GLY:HA2	3:M:230:THR:O	2.16	0.44
3:J:145:LYS:HE2	3:J:149(B):ASN:HD22	1.83	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:H:211:GLY:HA2	3:H:229:THR:O	2.18	0.44
3:C:61:GLU:HG2	3:C:87:LYS:HA	2.01	0.43
3:J:38:GLN:HE21	3:J:38:GLN:HB3	1.64	0.43
3:M:36:LYS:HE2	3:M:84:MET:SD	2.58	0.43
3:J:93:ARG:HG3	3:J:93:ARG:NH1	2.33	0.43
3:M:225:LYS:NZ	6:M:407:HOH:O	2.51	0.43
3:G:211:GLY:HA2	3:G:230:THR:O	2.18	0.43
3:H:145:LYS:HE3	3:H:149(B):ASN:CB	2.49	0.42
3:C:211:GLY:HA2	3:C:230:THR:O	2.20	0.42
3:G:29:TRP:CG	3:G:121:VAL:HB	2.55	0.42
3:C:201:MET:SD	3:C:210:MET:HG3	2.60	0.41
3:G:145:LYS:HD3	3:G:145:LYS:HA	1.93	0.41
2:D:14:LEU:HA	2:D:18:LYS:HD2	2.02	0.41
3:C:29:TRP:CG	3:C:121:VAL:HB	2.56	0.41
3:J:109:LYS:HE3	3:J:109:LYS:HB2	1.84	0.41
3:C:145:LYS:HD3	3:C:145:LYS:HA	1.94	0.41
3:E:81:LYS:HE3	3:E:118:ILE:HD12	2.03	0.41
3:E:16:ILE:N	3:E:194:ASP:OD2	2.54	0.40
2:F:8:ASP:O	3:G:119:HIS:HE1	2.04	0.40
1:O:34:LEU:HD23	1:O:34:LEU:HA	1.98	0.40

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)
6:M:517:HOH:O	6:N:121:HOH:O[3_455]	2.18	0.02

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	29/36 (81%)	29 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	28/36 (78%)	28 (100%)	0	0	100	100
1	O	28/36 (78%)	27 (96%)	1 (4%)	0	100	100
1	P	27/36 (75%)	26 (96%)	1 (4%)	0	100	100
1	Q	28/36 (78%)	28 (100%)	0	0	100	100
1	R	28/36 (78%)	28 (100%)	0	0	100	100
2	B	29/36 (81%)	28 (97%)	1 (3%)	0	100	100
2	D	25/36 (69%)	24 (96%)	1 (4%)	0	100	100
2	F	25/36 (69%)	24 (96%)	1 (4%)	0	100	100
2	I	27/36 (75%)	24 (89%)	3 (11%)	0	100	100
2	K	25/36 (69%)	24 (96%)	1 (4%)	0	100	100
2	L	28/36 (78%)	27 (96%)	1 (4%)	0	100	100
3	C	255/259 (98%)	249 (98%)	6 (2%)	0	100	100
3	E	254/259 (98%)	249 (98%)	5 (2%)	0	100	100
3	G	255/259 (98%)	249 (98%)	6 (2%)	0	100	100
3	H	255/259 (98%)	248 (97%)	7 (3%)	0	100	100
3	J	255/259 (98%)	248 (97%)	7 (3%)	0	100	100
3	M	256/259 (99%)	249 (97%)	7 (3%)	0	100	100
All	All	1857/1986 (94%)	1809 (97%)	48 (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	24/28 (86%)	24 (100%)	0	100	100
1	N	24/28 (86%)	24 (100%)	0	100	100
1	O	24/28 (86%)	24 (100%)	0	100	100
1	P	22/28 (79%)	22 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	Q	23/28 (82%)	23 (100%)	0	100	100
1	R	24/28 (86%)	23 (96%)	1 (4%)	30	28
2	B	26/31 (84%)	26 (100%)	0	100	100
2	D	21/31 (68%)	21 (100%)	0	100	100
2	F	24/31 (77%)	24 (100%)	0	100	100
2	I	25/31 (81%)	25 (100%)	0	100	100
2	K	23/31 (74%)	23 (100%)	0	100	100
2	L	25/31 (81%)	25 (100%)	0	100	100
3	C	219/225 (97%)	218 (100%)	1 (0%)	88	92
3	E	214/225 (95%)	212 (99%)	2 (1%)	78	83
3	G	213/225 (95%)	211 (99%)	2 (1%)	78	83
3	H	216/225 (96%)	215 (100%)	1 (0%)	88	92
3	J	218/225 (97%)	217 (100%)	1 (0%)	88	92
3	M	220/225 (98%)	217 (99%)	3 (1%)	67	72
All	All	1585/1704 (93%)	1574 (99%)	11 (1%)	84	89

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	182	CYS
3	C	182	CYS
3	E	154	VAL
3	E	182	CYS
3	G	182	CYS
3	G	234	ARG
3	J	182	CYS
3	M	157	VAL
3	M	182	CYS
3	M	234	ARG
1	R	35	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	131	GLN
3	G	119	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	151	GLN
3	J	38	GLN
1	P	31	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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	$\#|Z| > 2$
4	NAG	C	301	3	14,14,15	0.54	0	17,19,21	0.68	1 (5%)
4	NAG	E	301	3	14,14,15	0.39	0	17,19,21	0.64	1 (5%)
4	NAG	J	301	3	14,14,15	0.42	0	17,19,21	0.55	0
4	NAG	M	301	3	14,14,15	0.37	0	17,19,21	0.44	0
4	NAG	G	301	3	14,14,15	0.45	0	17,19,21	0.42	0
4	NAG	H	301	3	14,14,15	0.31	0	17,19,21	0.54	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	C	301	3	-	<mark>2/6/23/26</mark>	0/1/1/1
4	NAG	E	301	3	-	0/6/23/26	0/1/1/1
4	NAG	J	301	3	-	<mark>2/6/23/26</mark>	0/1/1/1
4	NAG	M	301	3	-	<mark>2/6/23/26</mark>	0/1/1/1
4	NAG	G	301	3	-	<mark>4/6/23/26</mark>	0/1/1/1
4	NAG	H	301	3	-	<mark>2/6/23/26</mark>	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	301	NAG	C1-O5-C5	2.39	115.43	112.19
4	E	301	NAG	C1-O5-C5	2.25	115.24	112.19

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	M	301	NAG	O5-C5-C6-O6
4	H	301	NAG	O5-C5-C6-O6
4	G	301	NAG	O5-C5-C6-O6
4	H	301	NAG	C4-C5-C6-O6
4	M	301	NAG	C4-C5-C6-O6
4	G	301	NAG	C4-C5-C6-O6
4	J	301	NAG	O5-C5-C6-O6
4	C	301	NAG	O5-C5-C6-O6
4	G	301	NAG	C1-C2-N2-C7
4	C	301	NAG	C4-C5-C6-O6
4	J	301	NAG	C4-C5-C6-O6
4	G	301	NAG	C3-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	301	NAG	1	0

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, 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(Å²)	Q<0.9
1	A	31/36 (86%)	0.26	0	100	100	30, 41, 68, 76	0
1	N	30/36 (83%)	0.14	0	100	100	29, 43, 70, 79	0
1	O	30/36 (83%)	-0.20	0	100	100	29, 44, 72, 81	0
1	P	29/36 (80%)	-0.08	0	100	100	34, 42, 67, 71	0
1	Q	30/36 (83%)	-0.04	0	100	100	31, 40, 60, 70	0
1	R	30/36 (83%)	-0.10	0	100	100	33, 42, 66, 73	0
2	B	31/36 (86%)	0.11	3 (9%)	7	11	33, 42, 77, 83	0
2	D	27/36 (75%)	-0.29	1 (3%)	41	49	40, 46, 57, 73	0
2	F	27/36 (75%)	-0.13	0	100	100	43, 50, 63, 66	0
2	I	29/36 (80%)	0.08	1 (3%)	45	53	41, 51, 74, 83	0
2	K	27/36 (75%)	0.14	2 (7%)	14	20	46, 52, 74, 79	0
2	L	30/36 (83%)	0.03	1 (3%)	46	55	34, 41, 69, 73	0
3	C	257/259 (99%)	-0.17	1 (0%)	92	94	26, 36, 56, 79	0
3	E	256/259 (98%)	-0.25	3 (1%)	79	83	28, 38, 55, 83	0
3	G	257/259 (99%)	-0.19	2 (0%)	86	89	33, 41, 60, 82	0
3	H	257/259 (99%)	-0.13	2 (0%)	86	89	26, 37, 59, 80	0
3	J	257/259 (99%)	-0.10	1 (0%)	92	94	29, 40, 59, 81	0
3	M	258/259 (99%)	-0.16	1 (0%)	92	94	29, 43, 71, 83	0
All	All	1893/1986 (95%)	-0.13	18 (0%)	82	86	26, 40, 66, 83	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	23	ILE	4.5
3	E	149(A)	ALA	4.5
3	G	149(A)	ALA	3.8

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	E	149	THR	3.7
2	K	22(G)	LEU	2.9
3	J	204(A)	PHE	2.8
2	B	6	GLU	2.8
3	E	149(B)	ASN	2.5
3	G	242	VAL	2.5
2	B	22(M)	GLY	2.5
2	I	7	ALA	2.5
3	H	149(A)	ALA	2.4
2	B	5	GLY	2.4
3	M	204(A)	PHE	2.3
2	L	6	GLU	2.3
3	H	149(B)	ASN	2.2
3	C	149(A)	ALA	2.1
2	D	22(K)	ILE	2.1

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 ⓘ

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(Å²)	Q<0.9
4	NAG	G	301	14/15	0.73	0.23	84,94,98,100	0
4	NAG	M	301	14/15	0.78	0.15	63,71,75,79	0
4	NAG	C	301	14/15	0.83	0.20	61,75,85,89	0
4	NAG	J	301	14/15	0.85	0.21	67,78,95,97	0
5	NA	M	302	1/1	0.85	0.12	34,34,34,34	0
4	NAG	H	301	14/15	0.88	0.26	55,66,75,76	0
4	NAG	E	301	14/15	0.92	0.20	52,64,66,76	0
5	NA	E	302	1/1	0.96	0.07	31,31,31,31	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å²)	Q<0.9
5	NA	G	302	1/1	0.96	0.05	38,38,38,38	0
5	NA	C	302	1/1	0.96	0.08	33,33,33,33	0
5	NA	J	302	1/1	0.98	0.07	35,35,35,35	0
5	NA	H	302	1/1	0.98	0.07	31,31,31,31	0

6.5　Other polymers ⓘ

There are no such residues in this entry.