



Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 09:27 PM GMT

PDB ID : 3HS0
Title : Cobra Venom Factor (CVF) in complex with human factor B
Authors : Janssen, B.J.C.; Gomes, L.; Koning, R.I.; Svergun, D.I.; Koster, A.J.;
Fritzinger, D.C.; Vogel, C.-W.; Gros, P.
Deposited on : 2009-06-10
Resolution : 3.00 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

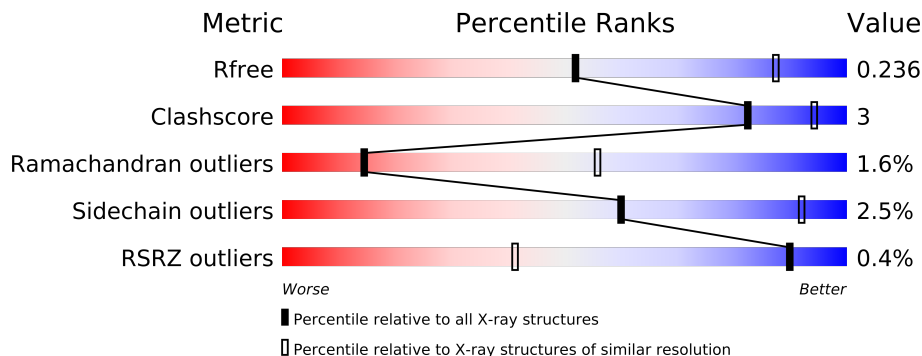
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.00 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}	66092	1216 (3.00-3.00)
Clashscore	79885	1594 (3.00-3.00)
Ramachandran outliers	78287	1537 (3.00-3.00)
Sidechain outliers	78261	1540 (3.00-3.00)
RSRZ outliers	66119	1217 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	627	
1	F	627	
2	B	252	
2	G	252	
3	C	379	
3	H	379	
4	D	741	
4	I	741	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	NAG	C	9324	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Geometry	Electron density
5	NAG	D	9117	-	X
5	NAG	D	9353	-	X
5	NAG	H	9324	-	X
5	NAG	I	9117	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 30435 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cobra venom factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	613	Total	C	N	O	S	0	0	0
			4794	3069	804	906	15			
1	F	617	Total	C	N	O	S	0	0	0
			4826	3085	811	915	15			

- Molecule 2 is a protein called Cobra venom factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	233	Total	C	N	O	S	0	0	0
			1856	1194	311	346	5			
2	G	233	Total	C	N	O	S	0	0	0
			1856	1194	311	346	5			

- Molecule 3 is a protein called Cobra venom factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	359	Total	C	N	O	S	0	0	0
			2900	1831	484	566	19			
3	H	366	Total	C	N	O	S	0	0	0
			2957	1864	496	578	19			

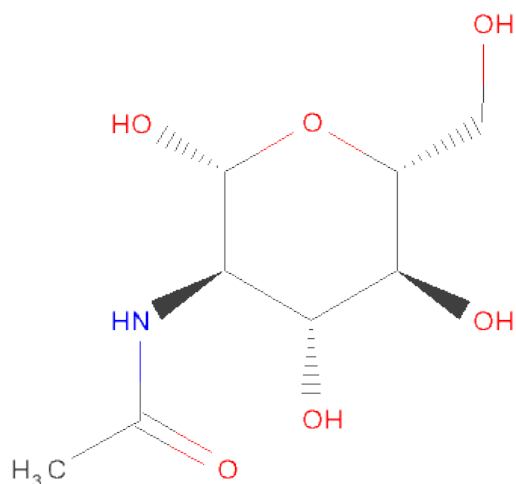
- Molecule 4 is a protein called Complement factor B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	699	Total	C	N	O	S	0	0	0
			5513	3474	954	1052	33			
4	I	704	Total	C	N	O	S	0	0	0
			5567	3506	972	1056	33			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	254	GLY	ASP	ENGINEERED	UNP P00751
D	260	ASP	ASN	ENGINEERED	UNP P00751
D	740	ALA	-	INSERTION	UNP P00751
D	741	ALA	-	INSERTION	UNP P00751
I	254	GLY	ASP	ENGINEERED	UNP P00751
I	260	ASP	ASN	ENGINEERED	UNP P00751
I	740	ALA	-	INSERTION	UNP P00751
I	741	ALA	-	INSERTION	UNP P00751

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	D	1	Total	C	N	O	0	0
			14	8	1	5		
5	D	1	Total	C	N	O	0	0
			14	8	1	5		
5	F	1	Total	C	N	O	0	0
			14	8	1	5		
5	H	1	Total	C	N	O	0	0
			14	8	1	5		
5	I	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	I	1	Total Mg 1 1	0	0
6	A	1	Total Mg 1 1	0	0
6	D	1	Total Mg 1 1	0	0
6	F	1	Total Mg 1 1	0	0

- Molecule 7 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	2	Total C N O 28 16 2 10	0	0
7	I	2	Total C N O 28 16 2 10	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	254	GLY	ASP	ENGINEERED	UNP P00751
D	260	ASP	ASN	ENGINEERED	UNP P00751
D	740	ALA	-	INSERTION	UNP P00751
D	741	ALA	-	INSERTION	UNP P00751
I	254	GLY	ASP	ENGINEERED	UNP P00751
I	260	ASP	ASN	ENGINEERED	UNP P00751
I	740	ALA	-	INSERTION	UNP P00751
I	741	ALA	-	INSERTION	UNP P00751

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	2	Total O 2 2	0	0
8	C	1	Total O 1 1	0	0
8	D	1	Total O 1 1	0	0
8	F	2	Total O 2 2	0	0
8	H	1	Total O 1 1	0	0

Continued on next page...

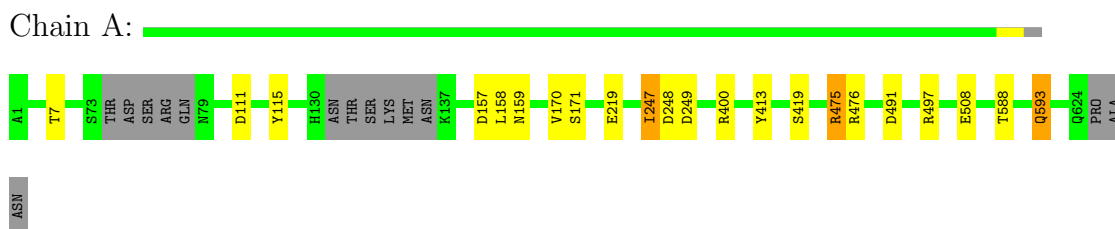
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	1	Total	O	0	0
			1	1		

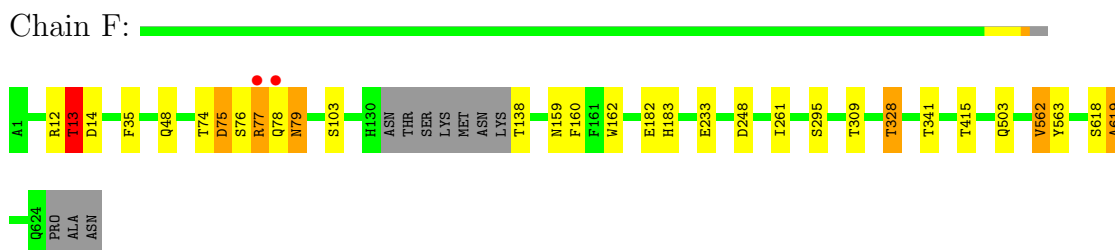
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

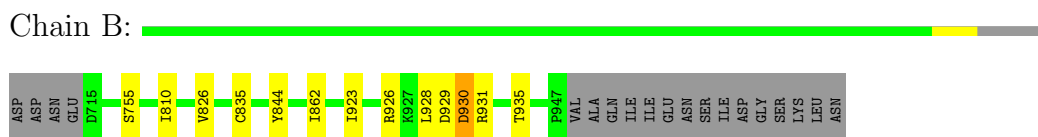
- Molecule 1: Cobra venom factor



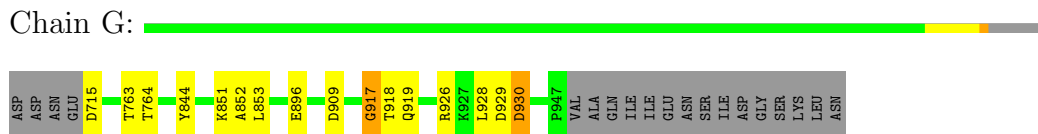
- Molecule 1: Cobra venom factor



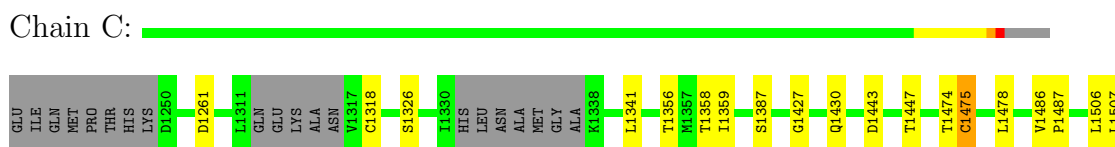
- Molecule 2: Cobra venom factor

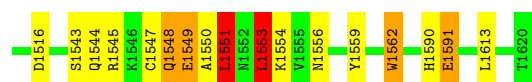


- Molecule 2: Cobra venom factor



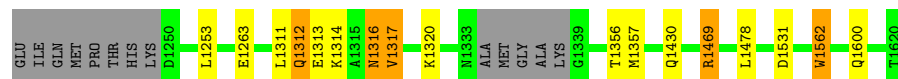
- Molecule 3: Cobra venom factor





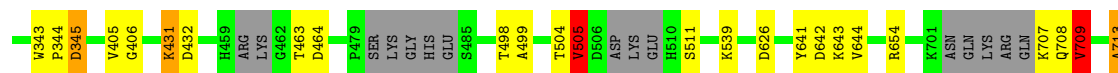
- Molecule 3: Cobra venom factor

Chain H:



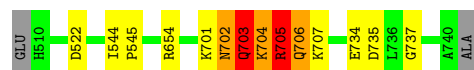
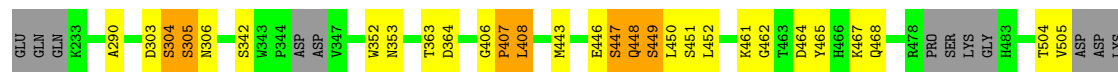
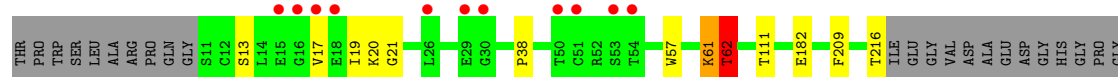
- Molecule 4: Complement factor B

Chain D:



- Molecule 4: Complement factor B

Chain I:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	134.03Å 136.97Å 283.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.88 – 3.00 34.88 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (34.88-3.00) 99.8 (34.88-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 3.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.189 , 0.243 0.179 , 0.236	Depositor DCC
R_{free} test set	2095 reflections (2.00%)	DCC
Wilson B-factor (Å ²)	72.0	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 26.3	EDS
Estimated twinning fraction	0.019 for k,h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 104864 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	30435	wwPDB-VP
Average B, all atoms (Å ²)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.03% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 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	0.27	0/4902	0.46	0/6668
1	F	0.27	0/4935	0.49	2/6715 (0.0%)
2	B	0.27	0/1894	0.46	0/2570
2	G	0.28	0/1894	0.51	1/2570 (0.0%)
3	C	0.26	0/2950	0.51	1/3989 (0.0%)
3	H	0.26	0/3009	0.47	0/4071
4	D	0.26	0/5636	0.48	2/7629 (0.0%)
4	I	0.25	0/5691	0.46	0/7699
All	All	0.26	0/30911	0.48	6/41911 (0.0%)

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	F	1	0
3	C	0	1
All	All	1	1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1553	LEU	N-CA-C	5.97	127.12	111.00
1	F	13	THR	C-N-CA	5.71	135.96	121.70
4	D	739	LEU	CA-CB-CG	5.49	127.93	115.30
2	G	917	GLY	N-CA-C	5.23	126.18	113.10
1	F	13	THR	CA-C-N	5.19	128.62	117.20
4	D	44	TYR	CA-CB-CG	-5.17	103.57	113.40

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	F	13	THR	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	1551	LEU	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4794	0	0	8	0
1	F	4826	0	0	14	0
2	B	1856	0	7	3	0
2	G	1856	0	7	5	0
3	C	2900	0	3	13	0
3	H	2957	0	3	8	0
4	D	5513	0	0	19	0
4	I	5567	0	0	27	0
5	A	14	0	13	0	0
5	C	14	0	13	0	0
5	D	28	0	0	0	0
5	F	14	0	13	0	0
5	H	14	0	13	0	0
5	I	14	0	0	0	0
6	A	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
6	I	1	0	0	0	0
7	D	28	0	0	0	0
7	I	28	0	0	0	0
8	A	2	0	0	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
8	F	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	1	0	0	0	0
8	I	1	0	0	0	0
All	All	30435	0	72	94	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (94) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:I:701:LYS:N	4:I:702:ASN:O	2.20	0.74
4:I:449:SER:O	4:I:451:SER:N	2.22	0.73
4:I:446:GLU:O	4:I:448:GLN:N	2.22	0.72
3:H:1312:GLN:O	3:H:1314:LYS:N	2.24	0.71
4:I:447:SER:O	4:I:448:GLN:CB	2.39	0.70
2:G:763:THR:CG2	2:G:764:THR:N	2.54	0.70
3:C:1549:GLU:C	3:C:1551:LEU:N	2.48	0.68
2:B:930:ASP:N	2:B:930:ASP:OD1	2.30	0.63
2:G:930:ASP:OD1	2:G:930:ASP:N	2.32	0.62
4:D:738:PHE:N	4:D:738:PHE:CD2	2.69	0.61
4:D:44:TYR:OH	4:D:74:ARG:NH1	2.33	0.61
4:I:448:GLN:CG	4:I:449:SER:N	2.65	0.59
3:H:1469:ARG:CG	3:H:1469:ARG:NH1	2.64	0.59
3:C:1562:TRP:CD1	3:C:1562:TRP:O	2.55	0.59
3:C:1553:LEU:CD1	3:C:1554:LYS:N	2.67	0.57
4:I:705:ARG:O	4:I:706:GLN:CB	2.52	0.57
1:F:74:THR:O	1:F:76:SER:N	2.37	0.57
4:D:714:ARG:CG	4:D:714:ARG:NH1	2.66	0.57
1:F:562:VAL:CG2	1:F:563:TYR:N	2.67	0.57
3:C:1590:HIS:O	3:C:1591:GLU:CB	2.53	0.56
1:F:78:GLN:O	1:F:79:ASN:CG	2.43	0.56
4:D:343:TRP:O	4:D:345:ASP:N	2.39	0.56
2:G:852:ALA:O	2:G:853:LEU:CB	2.53	0.56
1:A:247:ILE:CG2	1:A:248:ASP:N	2.67	0.55
3:C:1553:LEU:CG	3:C:1559:TYR:CE1	2.90	0.55
4:D:463:THR:CG2	4:D:464:ASP:N	2.69	0.55
1:F:13:THR:CG2	1:F:14:ASP:CB	2.85	0.55
1:A:475:ARG:CG	1:A:475:ARG:NH1	2.70	0.55
3:C:1545:ARG:O	3:C:1548:GLN:N	2.40	0.54
1:A:419:SER:OG	1:A:508:GLU:OE2	2.26	0.53
4:D:48:THR:O	4:D:48:THR:OG1	2.27	0.52
3:H:1263:GLU:N	4:I:182:GLU:OE2	2.45	0.50
4:I:303:ASP:O	4:I:304:SER:C	2.49	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:77:ARG:C	1:F:79:ASN:N	2.65	0.49
4:I:702:ASN:O	4:I:704:LYS:N	2.45	0.49
4:I:703:GLN:O	4:I:704:LYS:CB	2.60	0.49
4:D:713:ALA:O	4:D:714:ARG:CB	2.60	0.49
1:F:160:PHE:CE2	1:F:162:TRP:CE2	3.01	0.48
4:I:303:ASP:O	4:I:305:SER:N	2.47	0.47
4:I:407:PRO:O	4:I:408:LEU:CB	2.61	0.47
2:G:844:TYR:CE2	3:H:1430:GLN:NE2	2.82	0.47
4:I:734:GLU:O	4:I:735:ASP:CB	2.63	0.47
1:F:618:SER:O	1:F:619:ALA:CB	2.62	0.47
1:F:13:THR:CG2	1:F:103:SER:OG	2.63	0.47
4:I:304:SER:O	4:I:306:ASN:N	2.47	0.47
3:H:1316:ASN:O	3:H:1317:VAL:O	2.32	0.47
3:C:1516:ASP:N	3:C:1543:SER:O	2.48	0.46
4:D:47:GLN:NE2	4:D:47:GLN:N	2.62	0.46
4:I:352:TRP:CG	4:I:353:ASN:N	2.83	0.46
4:D:504:THR:O	4:D:505:VAL:CB	2.63	0.46
4:I:544:ILE:CG1	4:I:544:ILE:O	2.63	0.46
3:H:1562:TRP:O	3:H:1562:TRP:CD1	2.69	0.46
4:I:209:PHE:CD2	4:I:443:MET:CE	2.99	0.46
3:H:1316:ASN:O	3:H:1317:VAL:C	2.54	0.45
4:I:13:SER:O	4:I:57:TRP:NE1	2.49	0.45
3:H:1562:TRP:C	3:H:1562:TRP:CD1	2.90	0.45
1:F:77:ARG:O	1:F:78:GLN:CB	2.66	0.44
4:I:449:SER:O	4:I:452:LEU:N	2.50	0.44
1:F:35:PHE:CE1	1:F:48:GLN:NE2	2.86	0.44
1:F:328:THR:CG2	1:F:341:THR:O	2.66	0.44
1:A:111:ASP:OD1	1:A:115:TYR:OH	2.36	0.43
3:C:1506:LEU:O	3:C:1556:ASN:N	2.51	0.43
4:D:707:LYS:O	4:D:707:LYS:CD	2.66	0.43
1:A:497:ARG:NH2	1:A:588:THR:O	2.51	0.43
4:D:35:TYR:CD2	4:D:47:GLN:O	2.72	0.43
3:C:1474:THR:CG2	3:C:1475:CYS:N	2.81	0.43
4:D:642:ASP:O	4:D:644:VAL:N	2.51	0.43
1:A:157:ASP:O	1:A:159:ASN:N	2.51	0.43
4:I:522:ASP:OD1	4:I:522:ASP:N	2.52	0.43
2:B:844:TYR:CE2	3:C:1430:GLN:NE2	2.87	0.43
4:I:20:LYS:O	4:I:21:GLY:C	2.57	0.43
4:I:61:LYS:O	4:I:62:THR:OG1	2.37	0.42
4:I:464:ASP:O	4:I:467:LYS:N	2.52	0.42
1:F:12:ARG:O	1:F:13:THR:O	2.37	0.42
4:D:405:VAL:CG2	4:D:406:GLY:N	2.83	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:C:1358:THR:CG2	3:C:1359:ILE:N	2.82	0.42
4:D:708:GLN:O	4:D:709:VAL:CG1	2.67	0.42
4:D:626:ASP:OD1	4:D:626:ASP:N	2.52	0.42
4:D:724:LEU:N	4:D:725:PRO:CD	2.82	0.42
1:F:182:GLU:O	1:F:183:HIS:CB	2.67	0.42
2:B:931:ARG:NH1	2:B:935:THR:O	2.53	0.41
4:I:363:THR:OG1	4:I:364:ASP:N	2.53	0.41
4:I:706:GLN:N	4:I:707:LYS:CA	2.83	0.41
1:A:170:VAL:CG2	1:A:171:SER:N	2.83	0.41
4:D:498:THR:OG1	4:D:499:ALA:N	2.54	0.41
3:C:1486:VAL:N	3:C:1487:PRO:CD	2.83	0.41
4:I:449:SER:C	4:I:451:SER:N	2.74	0.41
1:A:413:TYR:CD1	1:A:593:GLN:O	2.73	0.41
2:G:909:ASP:N	2:G:919:GLN:NE2	2.69	0.41
1:F:562:VAL:O	1:F:563:TYR:CB	2.69	0.40
4:I:406:GLY:O	4:I:407:PRO:O	2.39	0.40
3:C:1545:ARG:C	3:C:1547:CYS:N	2.73	0.40
4:D:431:LYS:CG	4:D:432:ASP:N	2.84	0.40
4:D:641:TYR:CD2	4:D:641:TYR:N	2.90	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	607/627 (97%)	577 (95%)	26 (4%)	4 (1%)	30	78
1	F	613/627 (98%)	571 (93%)	36 (6%)	6 (1%)	22	70
2	B	231/252 (92%)	218 (94%)	13 (6%)	0	100	100
2	G	231/252 (92%)	218 (94%)	12 (5%)	1 (0%)	43	87
3	C	353/379 (93%)	318 (90%)	29 (8%)	6 (2%)	14	54
3	H	362/379 (96%)	333 (92%)	24 (7%)	5 (1%)	16	60
4	D	687/741 (93%)	608 (88%)	63 (9%)	16 (2%)	10	43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	I	694/741 (94%)	605 (87%)	65 (9%)	24 (4%)	6	30
All	All	3778/3998 (94%)	3448 (91%)	268 (7%)	62 (2%)	14	56

All (62) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	247	ILE
4	D	44	TYR
4	D	46	VAL
4	D	344	PRO
4	D	505	VAL
4	D	539	LYS
4	D	713	ALA
1	F	13	THR
1	F	75	ASP
1	F	79	ASN
3	H	1313	GLU
3	H	1317	VAL
4	I	19	ILE
4	I	38	PRO
4	I	61	LYS
4	I	305	SER
4	I	407	PRO
4	I	447	SER
4	I	448	GLN
4	I	449	SER
4	I	450	LEU
4	I	545	PRO
4	I	702	ASN
4	I	704	LYS
1	A	593	GLN
3	C	1549	GLU
3	C	1591	GLU
4	D	345	ASP
4	D	431	LYS
4	D	709	VAL
4	D	737	GLY
1	F	619	ALA
2	G	917	GLY
3	H	1357	MET
4	I	290	ALA
4	I	304	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	I	462	GLY
4	I	465	TYR
4	I	703	GLN
4	I	706	GLN
4	I	737	GLY
1	A	158	LEU
4	D	58	SER
4	D	290	ALA
4	D	643	LYS
1	F	248	ASP
3	C	1318	CYS
4	D	714	ARG
1	F	77	ARG
3	H	1600	GLN
4	I	62	THR
4	I	408	LEU
4	I	705	ARG
1	A	249	ASP
3	C	1443	ASP
3	C	1550	ALA
4	D	20	LYS
3	H	1531	ASP
4	I	504	THR
3	C	1427	GLY
4	D	511	SER
4	I	17	VAL

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	535/548 (98%)	529 (99%)	6 (1%)	84	97
1	F	539/548 (98%)	528 (98%)	11 (2%)	68	94
2	B	210/227 (92%)	200 (95%)	10 (5%)	35	79
2	G	210/227 (92%)	202 (96%)	8 (4%)	44	85
3	C	329/345 (95%)	314 (95%)	15 (5%)	37	80

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	H	335/345 (97%)	326 (97%)	9 (3%)	57	91
4	D	610/643 (95%)	596 (98%)	14 (2%)	63	93
4	I	615/643 (96%)	605 (98%)	10 (2%)	75	95
All	All	3383/3526 (96%)	3300 (98%)	83 (2%)	60	92

All (83) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	THR
1	A	219	GLU
1	A	400	ARG
1	A	475	ARG
1	A	476	ARG
1	A	491	ASP
2	B	755	SER
2	B	810	ILE
2	B	826	VAL
2	B	835	CYS
2	B	862	ILE
2	B	923	ILE
2	B	926	ARG
2	B	928	LEU
2	B	929	ASP
2	B	930	ASP
3	C	1261	ASP
3	C	1326	SER
3	C	1341	LEU
3	C	1356	THR
3	C	1387	SER
3	C	1447	THR
3	C	1475	CYS
3	C	1478	LEU
3	C	1507	LEU
3	C	1544	GLN
3	C	1548	GLN
3	C	1551	LEU
3	C	1553	LEU
3	C	1562	TRP
3	C	1613	LEU
4	D	44	TYR
4	D	47	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	48	THR
4	D	64	ASP
4	D	98	VAL
4	D	111	THR
4	D	175	SER
4	D	203	GLN
4	D	505	VAL
4	D	654	ARG
4	D	709	VAL
4	D	728	LYS
4	D	738	PHE
4	D	739	LEU
1	F	75	ASP
1	F	138	THR
1	F	159	ASN
1	F	233	GLU
1	F	261	ILE
1	F	295	SER
1	F	309	THR
1	F	328	THR
1	F	415	THR
1	F	503	GLN
1	F	562	VAL
2	G	715	ASP
2	G	851	LYS
2	G	896	GLU
2	G	918	THR
2	G	926	ARG
2	G	928	LEU
2	G	929	ASP
2	G	930	ASP
3	H	1253	LEU
3	H	1311	LEU
3	H	1312	GLN
3	H	1316	ASN
3	H	1320	LYS
3	H	1356	THR
3	H	1469	ARG
3	H	1478	LEU
3	H	1562	TRP
4	I	62	THR
4	I	111	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	I	216	THR
4	I	342	SER
4	I	461	LYS
4	I	468	GLN
4	I	505	VAL
4	I	654	ARG
4	I	703	GLN
4	I	705	ARG

Some 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 chains 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 ⓘ

4 carbohydrates 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	NAG	D	9097	4,7	12,14,15	0.60	0	15,19,21	1.89	4 (26%)
7	NAG	D	9098	7	12,14,15	0.60	0	15,19,21	1.07	1 (6%)
7	NAG	I	9097	4,7	12,14,15	0.60	0	15,19,21	1.06	1 (6%)
7	NAG	I	9098	7	12,14,15	0.56	0	15,19,21	1.78	1 (6%)

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
7	NAG	D	9097	4,7	-	0/6/23/26	0/1/1/1
7	NAG	D	9098	7	-	2/6/23/26	0/1/1/1
7	NAG	I	9097	4,7	-	0/6/23/26	0/1/1/1
7	NAG	I	9098	7	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	I	9098	NAG	O5-C5-C6	6.61	113.92	106.98
7	D	9097	NAG	O5-C5-C4	4.54	116.41	110.65
7	D	9098	NAG	O5-C5-C6	3.26	110.40	106.98
7	D	9097	NAG	C3-C4-C5	2.83	115.26	110.20
7	D	9097	NAG	C3-C2-N2	2.67	115.82	111.76
7	D	9097	NAG	C2-N2-C7	2.27	126.90	123.09
7	I	9097	NAG	O5-C5-C6	2.03	109.11	106.98

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	I	9098	NAG	C3-C2-N2-C7
7	D	9098	NAG	C8-C7-N2-C2
7	D	9098	NAG	O7-C7-N2-C2

There are no ring outliers.

5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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
5	NAG	A	9187	1	12,14,15	0.54	0	15,19,21	1.18	2 (13%)
5	NAG	C	9324	3	12,14,15	0.59	0	15,19,21	0.88	1 (6%)
5	NAG	D	9117	4	12,14,15	0.48	0	15,19,21	1.60	1 (6%)
5	NAG	D	9353	4	12,14,15	0.63	0	15,19,21	0.82	0
5	NAG	F	9187	1	12,14,15	0.51	0	15,19,21	1.51	3 (20%)
5	NAG	H	9324	3	12,14,15	0.62	0	15,19,21	0.89	1 (6%)
5	NAG	I	9117	4	12,14,15	0.58	0	15,19,21	0.89	1 (6%)

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	NAG	A	9187	1	-	0/6/23/26	0/1/1/1
5	NAG	C	9324	3	-	0/6/23/26	0/1/1/1
5	NAG	D	9117	4	-	0/6/23/26	0/1/1/1
5	NAG	D	9353	4	-	0/6/23/26	0/1/1/1
5	NAG	F	9187	1	-	0/6/23/26	0/1/1/1
5	NAG	H	9324	3	-	0/6/23/26	0/1/1/1
5	NAG	I	9117	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	9117	NAG	O5-C5-C4	5.36	117.46	110.65
5	F	9187	NAG	O5-C5-C6	3.29	110.43	106.98
5	F	9187	NAG	O5-C5-C4	3.26	114.79	110.65
5	A	9187	NAG	O5-C5-C6	2.87	109.99	106.98
5	H	9324	NAG	O5-C5-C6	2.41	109.51	106.98
5	F	9187	NAG	C3-C4-C5	2.36	114.42	110.20
5	I	9117	NAG	O5-C5-C4	2.22	113.47	110.65
5	A	9187	NAG	C2-N2-C7	-2.11	119.54	123.09
5	C	9324	NAG	O5-C5-C6	2.09	109.18	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å ²)	Q<0.9
1	A	613/627 (97%)	-0.44	0 100 100	31, 68, 125, 182	0
1	F	617/627 (98%)	-0.41	2 (0%) 91 48	37, 75, 132, 187	0
2	B	233/252 (92%)	-0.49	0 100 100	34, 63, 111, 123	0
2	G	233/252 (92%)	-0.47	0 100 100	33, 66, 115, 151	0
3	C	359/379 (94%)	-0.30	0 100 100	40, 94, 210, 291	0
3	H	366/379 (96%)	-0.42	0 100 100	33, 81, 135, 216	0
4	D	699/741 (94%)	-0.27	2 (0%) 91 48	35, 88, 189, 233	0
4	I	704/741 (95%)	-0.25	11 (1%) 68 16	37, 91, 191, 244	0
All	All	3824/3998 (95%)	-0.36	15 (0%) 90 41	31, 79, 168, 291	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	I	30	GLY	4.5
4	I	29	GLU	4.2
4	I	51	CYS	3.8
4	I	16	GLY	3.7
4	I	26	LEU	3.2
4	I	50	THR	3.1
4	I	18	GLU	2.7
4	D	51	CYS	2.5
4	I	15	GLU	2.5
4	I	17	VAL	2.3
4	I	53	SER	2.2
4	D	27	LEU	2.2
1	F	77	ARG	2.1
4	I	54	THR	2.0
1	F	78	GLN	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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
7	NAG	I	9097	14/15	0.27	3.78	126,139,157,173	0
7	NAG	D	9097	14/15	0.24	2.32	136,146,153,163	0
7	NAG	D	9098	14/15	0.50	-	161,168,182,183	0
7	NAG	I	9098	14/15	0.68	-	173,184,194,201	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	NAG	I	9117	14/15	0.35	4.02	116,129,142,143	0
5	NAG	D	9117	14/15	0.23	3.41	98,117,131,132	0
5	NAG	C	9324	14/15	0.28	3.35	111,123,130,134	0
5	NAG	H	9324	14/15	0.38	3.01	129,140,144,146	0
5	NAG	D	9353	14/15	0.29	2.04	122,136,142,146	0
5	NAG	A	9187	14/15	0.18	0.46	87,102,110,114	0
5	NAG	F	9187	14/15	0.30	0.46	106,116,125,127	0
6	MG	I	742	1/1	0.13	-0.72	97,97,97,97	0
6	MG	D	742	1/1	0.10	-1.85	107,107,107,107	0
6	MG	A	628	1/1	0.08	-2.22	105,105,105,105	0
6	MG	F	628	1/1	0.04	-3.08	73,73,73,73	0

6.5 Other polymers ⓘ

There are no such residues in this entry.