



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:08 PM GMT

PDB ID : 4WSS
Title : The crystal structure of hemagglutinin form A/chicken/New York/14677-13/1998 in complex with LSTa
Authors : Yang, H.; Carney, P.J.; Chang, J.C.; Villanueva, J.M.; Stevens, J.
Deposited on : 2014-10-28
Resolution : 2.80 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

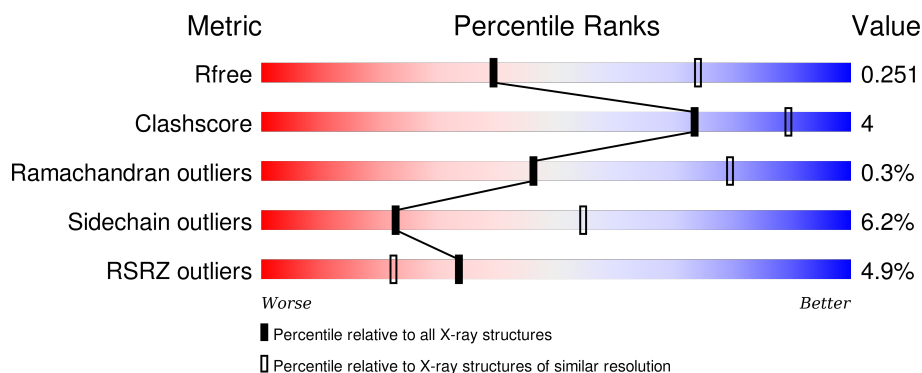
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	514	<div> <div>2%</div> <div>83% 10% • 5%</div> </div>
1	B	514	<div> <div>3%</div> <div>80% 12% • 6%</div> </div>
1	C	514	<div> <div>2%</div> <div>82% 9% • 6%</div> </div>
1	D	514	<div> <div>9%</div> <div>82% 12% • •</div> </div>
1	E	514	<div> <div>81% 12% • 5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	514	

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
2	NAG	B	601	X	-	-	-
2	NAG	B	602	-	-	-	X
2	NAG	C	602	X	-	-	-
2	NAG	D	900	X	-	-	-
2	NAG	E	601	X	-	-	-
2	NAG	F	601	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 23442 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	490	Total	C	N	O	S	0	0	0
			3895	2446	676	754	19			
1	B	481	Total	C	N	O	S	0	0	0
			3821	2399	663	740	19			
1	C	482	Total	C	N	O	S	0	0	0
			3832	2405	667	741	19			
1	E	489	Total	C	N	O	S	0	0	0
			3886	2441	675	751	19			
1	D	493	Total	C	N	O	S	0	0	0
			3909	2458	678	754	19			
1	F	479	Total	C	N	O	S	0	0	0
			3810	2391	663	737	19			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	ALA	-	expression tag	UNP Q0A3A5
A	-2	ASP	-	expression tag	UNP Q0A3A5
A	-1	PRO	-	expression tag	UNP Q0A3A5
A	0	GLY	-	expression tag	UNP Q0A3A5
A	317	THR	UNK	conflict	UNP Q0A3A5
A	499	GLU	GLN	conflict	UNP Q0A3A5
A	504	GLY	VAL	conflict	UNP Q0A3A5
A	505	ARG	MET	conflict	UNP Q0A3A5
A	507	VAL	-	expression tag	UNP Q0A3A5
A	508	PRO	-	expression tag	UNP Q0A3A5
A	509	ARG	-	expression tag	UNP Q0A3A5
A	510	GLY	-	expression tag	UNP Q0A3A5
B	-3	ALA	-	expression tag	UNP Q0A3A5
B	-2	ASP	-	expression tag	UNP Q0A3A5
B	-1	PRO	-	expression tag	UNP Q0A3A5
B	0	GLY	-	expression tag	UNP Q0A3A5
B	317	THR	UNK	conflict	UNP Q0A3A5

Continued on next page...

Continued from previous page...

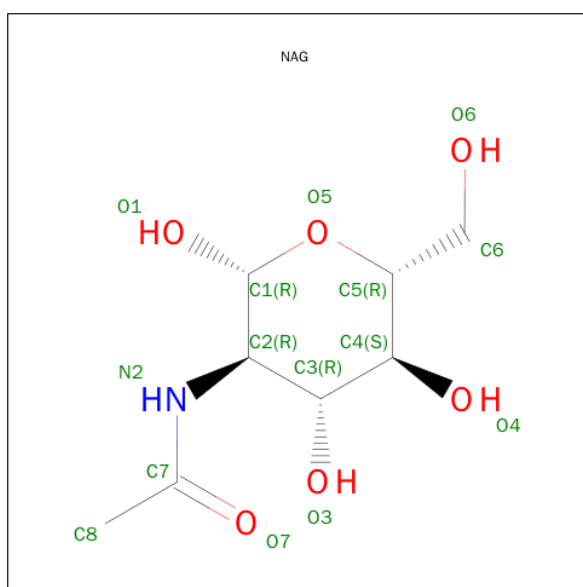
Chain	Residue	Modelled	Actual	Comment	Reference
B	499	GLU	GLN	conflict	UNP Q0A3A5
B	504	GLY	VAL	conflict	UNP Q0A3A5
B	505	ARG	MET	conflict	UNP Q0A3A5
B	507	VAL	-	expression tag	UNP Q0A3A5
B	508	PRO	-	expression tag	UNP Q0A3A5
B	509	ARG	-	expression tag	UNP Q0A3A5
B	510	GLY	-	expression tag	UNP Q0A3A5
C	-3	ALA	-	expression tag	UNP Q0A3A5
C	-2	ASP	-	expression tag	UNP Q0A3A5
C	-1	PRO	-	expression tag	UNP Q0A3A5
C	0	GLY	-	expression tag	UNP Q0A3A5
C	317	THR	UNK	conflict	UNP Q0A3A5
C	499	GLU	GLN	conflict	UNP Q0A3A5
C	504	GLY	VAL	conflict	UNP Q0A3A5
C	505	ARG	MET	conflict	UNP Q0A3A5
C	507	VAL	-	expression tag	UNP Q0A3A5
C	508	PRO	-	expression tag	UNP Q0A3A5
C	509	ARG	-	expression tag	UNP Q0A3A5
C	510	GLY	-	expression tag	UNP Q0A3A5
E	-3	ALA	-	expression tag	UNP Q0A3A5
E	-2	ASP	-	expression tag	UNP Q0A3A5
E	-1	PRO	-	expression tag	UNP Q0A3A5
E	0	GLY	-	expression tag	UNP Q0A3A5
E	317	THR	UNK	conflict	UNP Q0A3A5
E	499	GLU	GLN	conflict	UNP Q0A3A5
E	504	GLY	VAL	conflict	UNP Q0A3A5
E	505	ARG	MET	conflict	UNP Q0A3A5
E	507	VAL	-	expression tag	UNP Q0A3A5
E	508	PRO	-	expression tag	UNP Q0A3A5
E	509	ARG	-	expression tag	UNP Q0A3A5
E	510	GLY	-	expression tag	UNP Q0A3A5
D	-3	ALA	-	expression tag	UNP Q0A3A5
D	-2	ASP	-	expression tag	UNP Q0A3A5
D	-1	PRO	-	expression tag	UNP Q0A3A5
D	0	GLY	-	expression tag	UNP Q0A3A5
D	317	THR	UNK	conflict	UNP Q0A3A5
D	499	GLU	GLN	conflict	UNP Q0A3A5
D	504	GLY	VAL	conflict	UNP Q0A3A5
D	505	ARG	MET	conflict	UNP Q0A3A5
D	507	VAL	-	expression tag	UNP Q0A3A5
D	508	PRO	-	expression tag	UNP Q0A3A5
D	509	ARG	-	expression tag	UNP Q0A3A5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	510	GLY	-	expression tag	UNP Q0A3A5
F	-3	ALA	-	expression tag	UNP Q0A3A5
F	-2	ASP	-	expression tag	UNP Q0A3A5
F	-1	PRO	-	expression tag	UNP Q0A3A5
F	0	GLY	-	expression tag	UNP Q0A3A5
F	317	THR	UNK	conflict	UNP Q0A3A5
F	499	GLU	GLN	conflict	UNP Q0A3A5
F	504	GLY	VAL	conflict	UNP Q0A3A5
F	505	ARG	MET	conflict	UNP Q0A3A5
F	507	VAL	-	expression tag	UNP Q0A3A5
F	508	PRO	-	expression tag	UNP Q0A3A5
F	509	ARG	-	expression tag	UNP Q0A3A5
F	510	GLY	-	expression tag	UNP Q0A3A5

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



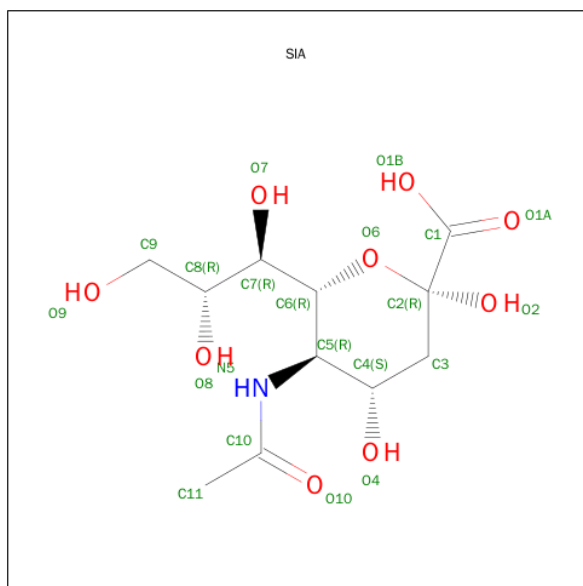
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is O-SIALIC ACID (three-letter code: SIA) (formula: C₁₁H₁₉NO₉).



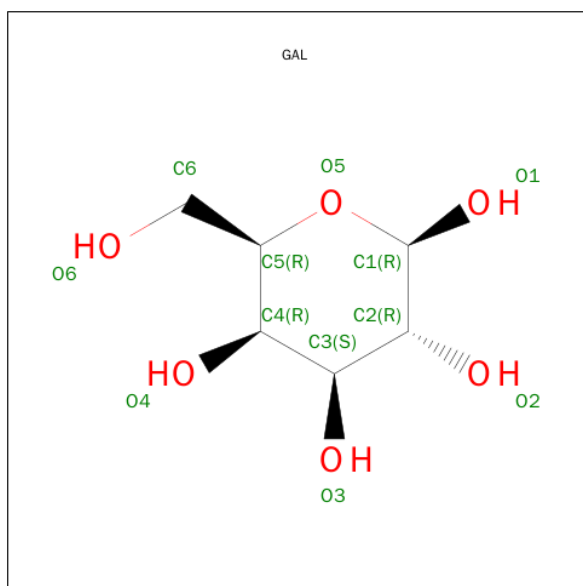
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	N	O	0	0
			20	11	1	8		
3	E	1	Total	C	N	O	0	0
			20	11	1	8		

Continued on next page...

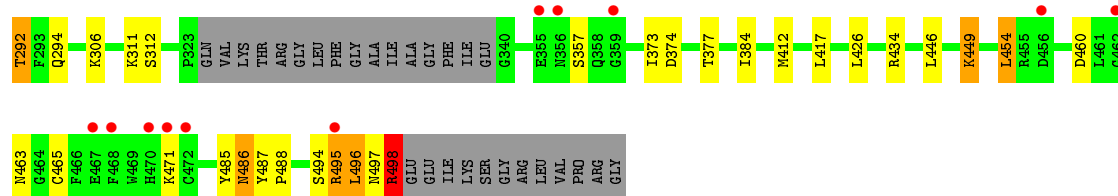
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	F	1	Total	C	N	O	0	0
			20	11	1	8		

- Molecule 4 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: C₆H₁₂O₆).

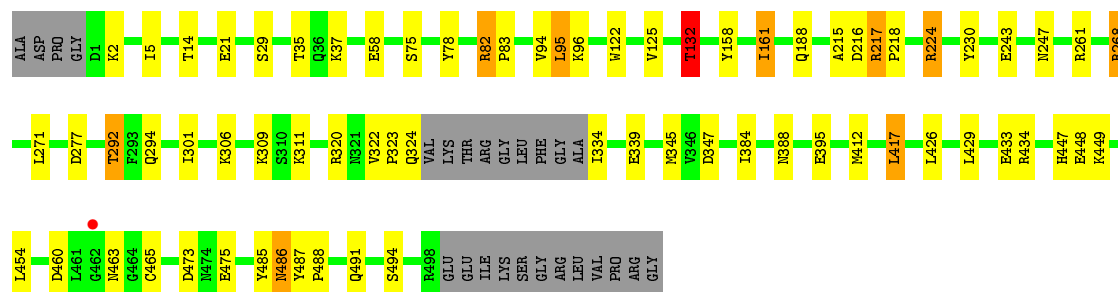


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		
4	F	1	Total	C	O	0	0
			11	6	5		



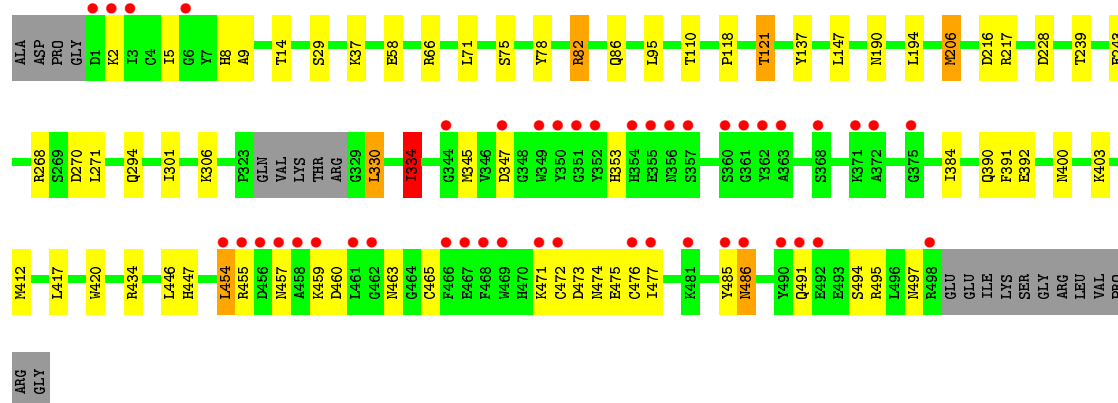
• Molecule 1: Hemagglutinin

Chain E: 81% 12% 5%



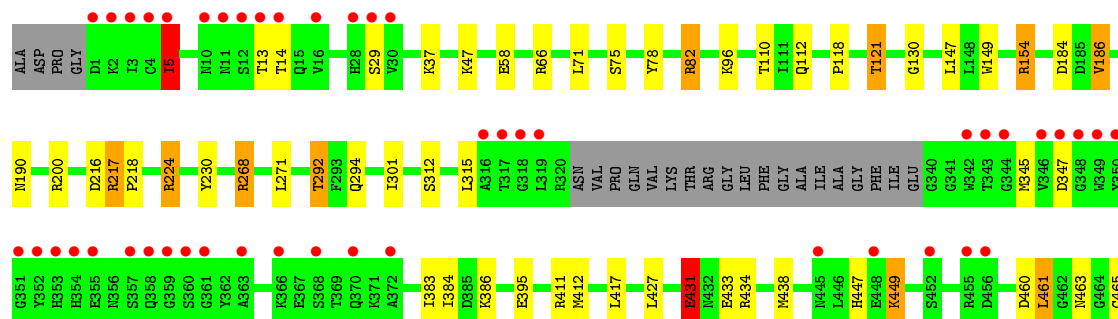
• Molecule 1: Hemagglutinin

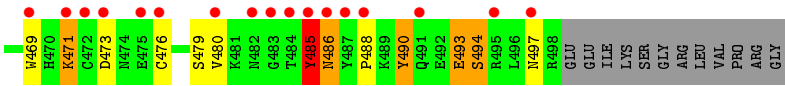
Chain D: 9% 82% 12%



• Molecule 1: Hemagglutinin

Chain F: 12% 80% 10% 7%





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	254.97Å 134.78Å 122.62Å 90.00° 112.20° 90.00°	Depositor
Resolution (Å)	45.20 – 2.80 45.20 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.5 (45.20-2.80) 98.5 (45.20-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.63 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.220 , 0.254 0.221 , 0.251	Depositor DCC
R_{free} test set	4650 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	52.3	Xtriage
Anisotropy	0.359	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 18.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 92840 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	23442	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: SIA, GAL, 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	0.64	0/3983	0.77	5/5397 (0.1%)
1	B	0.61	0/3908	0.76	5/5297 (0.1%)
1	C	0.62	0/3919	0.77	4/5311 (0.1%)
1	D	0.64	0/3998	0.80	7/5417 (0.1%)
1	E	0.64	0/3974	0.79	6/5385 (0.1%)
1	F	0.64	0/3896	0.81	7/5278 (0.1%)
All	All	0.63	0/23678	0.79	34/32085 (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	D	0	1
1	E	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	461	LEU	CA-CB-CG	9.70	137.62	115.30
1	E	277	ASP	CB-CG-OD2	8.01	125.51	118.30
1	F	485	TYR	N-CA-C	7.71	131.82	111.00
1	F	431	GLU	CB-CA-C	7.60	125.60	110.40
1	C	496	LEU	CA-CB-CG	-6.86	99.53	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	334	ILE	Peptide
1	E	323	PRO	Peptide

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	3895	0	3753	33	0
1	B	3821	0	3680	29	0
1	C	3832	0	3693	34	0
1	D	3909	0	3770	32	2
1	E	3886	0	3746	37	0
1	F	3810	0	3672	40	0
2	A	28	0	25	0	0
2	B	28	0	26	1	0
2	C	42	0	38	1	0
2	D	14	0	13	0	0
2	E	42	0	38	2	0
2	F	42	0	37	0	0
3	C	20	0	17	4	0
3	E	20	0	17	1	0
3	F	20	0	17	0	0
4	C	11	0	9	0	0
4	E	11	0	9	0	0
4	F	11	0	9	0	0
All	All	23442	0	22569	187	2

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 4.

The worst 5 of 187 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:457:ASN:ND2	1:D:485:TYR:OH	1.98	0.96
1:B:457:ASN:ND2	1:B:485:TYR:OH	1.98	0.96
1:D:473:ASP:O	1:D:476:CYS:SG	2.24	0.96
1:E:2:LYS:NZ	1:E:334:ILE:O	1.98	0.95

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:469:TRP:CD1	1:F:494:SER:OG	2.20	0.95

All (2) 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)
1:D:137:TYR:OH	1:D:137:TYR:OH[2_555]	1.91	0.29
1:D:137:TYR:CZ	1:D:137:TYR:OH[2_555]	2.19	0.01

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	486/514 (95%)	465 (96%)	19 (4%)	2 (0%)	39	74
1	B	477/514 (93%)	457 (96%)	19 (4%)	1 (0%)	52	84
1	C	478/514 (93%)	457 (96%)	19 (4%)	2 (0%)	39	74
1	D	489/514 (95%)	468 (96%)	19 (4%)	2 (0%)	39	74
1	E	485/514 (94%)	464 (96%)	20 (4%)	1 (0%)	52	84
1	F	475/514 (92%)	451 (95%)	22 (5%)	2 (0%)	39	74
All	All	2890/3084 (94%)	2762 (96%)	118 (4%)	10 (0%)	46	79

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	486	ASN
1	A	486	ASN
1	B	486	ASN
1	C	486	ASN
1	D	486	ASN

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	426/443 (96%)	402 (94%)	24 (6%)	26	59
1	B	419/443 (95%)	390 (93%)	29 (7%)	19	48
1	C	420/443 (95%)	394 (94%)	26 (6%)	23	54
1	D	426/443 (96%)	399 (94%)	27 (6%)	22	53
1	E	425/443 (96%)	403 (95%)	22 (5%)	29	62
1	F	417/443 (94%)	387 (93%)	30 (7%)	18	45
All	All	2533/2658 (95%)	2375 (94%)	158 (6%)	23	54

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	426	LEU
1	E	292	THR
1	F	417	LEU
1	C	454	LEU
1	E	75	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	86	GLN
1	E	112	GLN
1	D	457	ASN
1	B	457	ASN
1	F	8	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

17 non-standard protein/DNA/RNA residues 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$
2	NAG	A	2000	1,2	14,14,15	0.87	1 (7%)	15,19,21	1.24	1 (6%)
2	NAG	A	2001	2	14,14,15	0.52	0	15,19,21	1.19	2 (13%)
2	NAG	B	601	1	14,14,15	0.40	0	15,19,21	1.59	4 (26%)
2	NAG	B	602	1	14,14,15	1.04	1 (7%)	15,19,21	1.86	4 (26%)
2	NAG	C	601	1	14,14,15	0.68	0	15,19,21	1.49	2 (13%)
2	NAG	C	602	1	14,14,15	1.03	1 (7%)	15,19,21	1.54	4 (26%)
4	GAL	C	604	3,2	11,11,12	0.85	0	14,15,17	2.02	3 (21%)
2	NAG	C	605	4	14,14,15	0.71	0	15,19,21	1.43	2 (13%)
2	NAG	D	900	1	14,14,15	0.80	0	15,19,21	1.27	1 (6%)
2	NAG	E	601	1	14,14,15	1.55	2 (14%)	15,19,21	2.97	10 (66%)
2	NAG	E	602	1	14,14,15	0.55	0	15,19,21	1.91	1 (6%)
4	GAL	E	604	3,2	11,11,12	0.87	0	14,15,17	2.10	5 (35%)
2	NAG	E	605	4	14,14,15	0.87	1 (7%)	15,19,21	1.73	4 (26%)
2	NAG	F	601	1,2	14,14,15	0.57	0	15,19,21	1.08	1 (6%)
2	NAG	F	602	2	14,14,15	0.65	0	15,19,21	1.13	1 (6%)
4	GAL	F	604	3,2	11,11,12	0.79	0	14,15,17	2.49	5 (35%)
2	NAG	F	605	4	14,14,15	0.87	1 (7%)	15,19,21	2.03	6 (40%)

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
2	NAG	A	2000	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	2001	2	-	0/6/23/26	0/1/1/1
2	NAG	B	601	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	602	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	601	1	-	0/6/23/26	0/1/1/1
2	NAG	C	602	1	1/1/5/7	0/6/23/26	0/1/1/1
4	GAL	C	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	C	605	4	-	0/6/23/26	0/1/1/1
2	NAG	D	900	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	E	601	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	E	602	1	-	0/6/23/26	0/1/1/1
4	GAL	E	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	E	605	4	-	0/6/23/26	0/1/1/1
2	NAG	F	601	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	602	2	-	0/6/23/26	0/1/1/1
4	GAL	F	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	F	605	4	-	0/6/23/26	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	605	NAG	C3-C2	2.03	1.57	1.52
2	E	601	NAG	C3-C2	2.13	1.57	1.52
2	C	602	NAG	C1-C2	2.28	1.55	1.52
2	A	2000	NAG	C1-C2	2.48	1.55	1.52
2	E	605	NAG	C1-C2	2.60	1.56	1.52

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	601	NAG	C3-C4-C5	-5.23	101.08	110.20
4	F	604	GAL	O3-C3-C2	-5.00	100.97	110.00
4	C	604	GAL	O5-C1-C2	-4.35	103.79	110.86
4	F	604	GAL	O2-C2-C3	-4.26	101.55	110.12
4	F	604	GAL	O5-C1-C2	-4.23	104.00	110.86

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	602	NAG	C1
2	E	601	NAG	C1
2	B	601	NAG	C1
2	D	900	NAG	C1

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	NAG	1	0
2	C	602	NAG	1	0
2	E	601	NAG	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

20 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	$\# Z > 2$
2	NAG	A	2000	1,2	14,14,15	0.87	1 (7%)	15,19,21	1.24	1 (6%)
2	NAG	A	2001	2	14,14,15	0.52	0	15,19,21	1.19	2 (13%)
2	NAG	B	601	1	14,14,15	0.40	0	15,19,21	1.59	4 (26%)
2	NAG	B	602	1	14,14,15	1.04	1 (7%)	15,19,21	1.86	4 (26%)
2	NAG	C	601	1	14,14,15	0.68	0	15,19,21	1.49	2 (13%)
2	NAG	C	602	1	14,14,15	1.03	1 (7%)	15,19,21	1.54	4 (26%)
3	SIA	C	603	4	16,20,21	0.60	0	18,28,31	1.31	2 (11%)
4	GAL	C	604	3,2	11,11,12	0.85	0	14,15,17	2.02	3 (21%)
2	NAG	C	605	4	14,14,15	0.71	0	15,19,21	1.43	2 (13%)
2	NAG	D	900	1	14,14,15	0.80	0	15,19,21	1.27	1 (6%)
2	NAG	E	601	1	14,14,15	1.55	2 (14%)	15,19,21	2.97	10 (66%)
2	NAG	E	602	1	14,14,15	0.55	0	15,19,21	1.91	1 (6%)
3	SIA	E	603	4	16,20,21	0.86	1 (6%)	18,28,31	1.91	4 (22%)
4	GAL	E	604	3,2	11,11,12	0.87	0	14,15,17	2.10	5 (35%)
2	NAG	E	605	4	14,14,15	0.87	1 (7%)	15,19,21	1.73	4 (26%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	NAG	F	601	1,2	14,14,15	0.57	0	15,19,21	1.08	1 (6%)
2	NAG	F	602	2	14,14,15	0.65	0	15,19,21	1.13	1 (6%)
3	SIA	F	603	4	16,20,21	0.75	0	18,28,31	1.16	2 (11%)
4	GAL	F	604	3,2	11,11,12	0.79	0	14,15,17	2.49	5 (35%)
2	NAG	F	605	4	14,14,15	0.87	1 (7%)	15,19,21	2.03	6 (40%)

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
2	NAG	A	2000	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	2001	2	-	0/6/23/26	0/1/1/1
2	NAG	B	601	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	B	602	1	-	0/6/23/26	0/1/1/1
2	NAG	C	601	1	-	0/6/23/26	0/1/1/1
2	NAG	C	602	1	1/1/5/7	0/6/23/26	0/1/1/1
3	SIA	C	603	4	-	0/14/34/38	0/1/1/1
4	GAL	C	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	C	605	4	-	0/6/23/26	0/1/1/1
2	NAG	D	900	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	E	601	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	E	602	1	-	0/6/23/26	0/1/1/1
3	SIA	E	603	4	-	0/14/34/38	0/1/1/1
4	GAL	E	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	E	605	4	-	0/6/23/26	0/1/1/1
2	NAG	F	601	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	602	2	-	0/6/23/26	0/1/1/1
3	SIA	F	603	4	-	0/14/34/38	0/1/1/1
4	GAL	F	604	3,2	-	0/2/19/22	0/1/1/1
2	NAG	F	605	4	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	603	SIA	C5-N5	-2.37	1.41	1.45
2	F	605	NAG	C3-C2	2.03	1.57	1.52
2	E	601	NAG	C3-C2	2.13	1.57	1.52
2	C	602	NAG	C1-C2	2.28	1.55	1.52
2	A	2000	NAG	C1-C2	2.48	1.55	1.52

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	601	NAG	C3-C4-C5	-5.23	101.08	110.20
4	F	604	GAL	O3-C3-C2	-5.00	100.97	110.00
3	E	603	SIA	C3-C4-C5	-4.43	106.53	111.47
4	C	604	GAL	O5-C1-C2	-4.35	103.79	110.86
4	F	604	GAL	O2-C2-C3	-4.26	101.55	110.12

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	602	NAG	C1
2	E	601	NAG	C1
2	B	601	NAG	C1
2	D	900	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	NAG	1	0
2	C	602	NAG	1	0
3	C	603	SIA	4	0
2	E	601	NAG	2	0
3	E	603	SIA	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	490/514 (95%)	-0.16	11 (2%) 65 54	33, 52, 101, 128	0
1	B	481/514 (93%)	-0.11	13 (2%) 58 45	34, 56, 115, 140	0
1	C	482/514 (93%)	-0.19	11 (2%) 64 52	35, 52, 98, 133	0
1	D	493/514 (95%)	0.14	45 (9%) 11 6	32, 55, 160, 224	0
1	E	489/514 (95%)	-0.29	1 (0%) 95 94	36, 54, 76, 99	0
1	F	479/514 (93%)	0.36	63 (13%) 4 2	36, 58, 164, 200	0
All	All	2914/3084 (94%)	-0.04	144 (4%) 33 22	32, 54, 130, 224	0

The worst 5 of 144 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	357	SER	10.6
1	F	359	GLY	9.4
1	F	354	HIS	7.7
1	F	472	CYS	7.3
1	F	452	SER	7.0

6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	NAG	B	602	14/15	0.83	0.25	4.55	68,75,78,81	0
2	NAG	F	601	14/15	0.90	0.33	2.64	82,87,99,100	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	NAG	A	2000	14/15	0.91	0.20	1.40	68,84,92,94	0
4	GAL	F	604	11/12	0.93	0.09	-	57,65,74,77	0
2	NAG	C	602	14/15	0.85	0.32	-	57,80,94,100	0
2	NAG	E	601	14/15	0.78	0.25	-	55,85,93,97	0
2	NAG	B	601	14/15	0.73	0.26	-	106,117,129,130	0
2	NAG	C	605	14/15	0.87	0.16	-	73,94,107,113	0
2	NAG	E	605	14/15	0.84	0.17	-	85,102,118,127	0
2	NAG	E	602	14/15	0.79	0.21	-	97,114,122,122	0
2	NAG	C	601	14/15	0.77	0.25	-	104,117,132,133	0
2	NAG	F	602	14/15	0.83	0.38	-	88,109,115,115	0
4	GAL	E	604	11/12	0.94	0.09	-	59,65,74,76	0
2	NAG	A	2001	14/15	0.85	0.29	-	86,101,109,110	0
2	NAG	F	605	14/15	0.78	0.20	-	87,98,110,110	0
4	GAL	C	604	11/12	0.96	0.08	-	53,57,65,66	0
2	NAG	D	900	14/15	0.77	0.19	-	101,115,122,126	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	NAG	B	602	14/15	0.83	0.25	4.55	68,75,78,81	0
2	NAG	F	601	14/15	0.90	0.33	2.64	82,87,99,100	0
2	NAG	A	2000	14/15	0.91	0.20	1.40	68,84,92,94	0
3	SIA	F	603	20/21	0.96	0.18	0.22	49,52,64,67	0
3	SIA	C	603	20/21	0.96	0.15	-0.24	38,45,57,65	0
3	SIA	E	603	20/21	0.97	0.12	-0.65	42,46,53,55	0
2	NAG	C	602	14/15	0.85	0.32	-	57,80,94,100	0
2	NAG	E	601	14/15	0.78	0.25	-	55,85,93,97	0
2	NAG	B	601	14/15	0.73	0.26	-	106,117,129,130	0
4	GAL	F	604	11/12	0.93	0.09	-	57,65,74,77	0
2	NAG	F	605	14/15	0.78	0.20	-	87,98,110,110	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	NAG	E	605	14/15	0.84	0.17	-	85,102,118,127	0
2	NAG	E	602	14/15	0.79	0.21	-	97,114,122,122	0
2	NAG	C	601	14/15	0.77	0.25	-	104,117,132,133	0
2	NAG	F	602	14/15	0.83	0.38	-	88,109,115,115	0
4	GAL	E	604	11/12	0.94	0.09	-	59,65,74,76	0
4	GAL	C	604	11/12	0.96	0.08	-	53,57,65,66	0
2	NAG	C	605	14/15	0.87	0.16	-	73,94,107,113	0
2	NAG	A	2001	14/15	0.85	0.29	-	86,101,109,110	0
2	NAG	D	900	14/15	0.77	0.19	-	101,115,122,126	0

6.5 Other polymers [i](#)

There are no such residues in this entry.