Full wwPDB X-ray Structure Validation Report (i)

Jun 23, 2024 – 09:47 AM EDT

PDB ID	:	5LJY
Title	:	Structure of hantavirus envelope glycoprotein Gc in complex with scFv A5
Authors	:	Guardado-Calvo, P.; Stettner, E.; Jeffers, S.A.; Rey, F.A.
Deposited on	:	2016-07-20
Resolution	:	3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

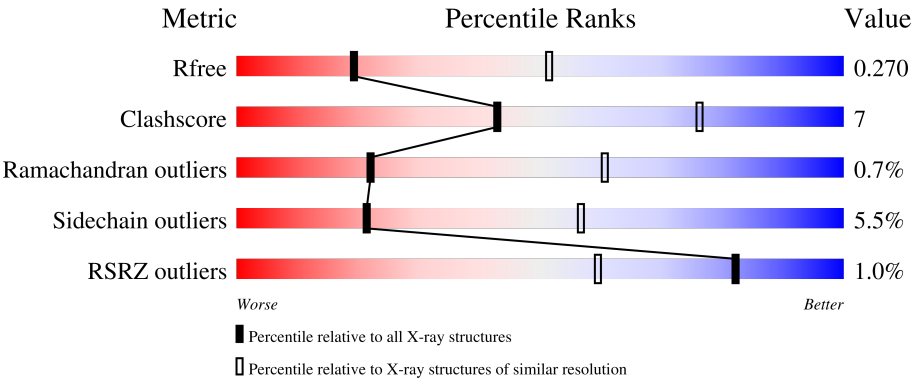
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 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	495	% 62% 12% · 24%
2	H	281	32% 7% · 60%
2	L	281	33% · · 62%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4594 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 Envelopment polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	374	Total	C	N	O	S	0	0	0
			2908	1835	495	550	28			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	458	GLY	-	expression tag	UNP P08668
A	459	GLY	-	expression tag	UNP P08668
A	460	SER	-	expression tag	UNP P08668
A	461	ASP	-	expression tag	UNP P08668
A	462	ASP	-	expression tag	UNP P08668
A	463	ASP	-	expression tag	UNP P08668
A	464	ASP	-	expression tag	UNP P08668
A	465	LYS	-	expression tag	UNP P08668
A	466	ALA	-	expression tag	UNP P08668
A	467	GLY	-	expression tag	UNP P08668
A	468	TRP	-	expression tag	UNP P08668
A	469	SER	-	expression tag	UNP P08668
A	470	HIS	-	expression tag	UNP P08668
A	471	PRO	-	expression tag	UNP P08668
A	472	GLN	-	expression tag	UNP P08668
A	473	PHE	-	expression tag	UNP P08668
A	474	GLU	-	expression tag	UNP P08668
A	475	LYS	-	expression tag	UNP P08668
A	476	GLY	-	expression tag	UNP P08668
A	477	GLY	-	expression tag	UNP P08668
A	478	GLY	-	expression tag	UNP P08668
A	479	SER	-	expression tag	UNP P08668
A	480	GLY	-	expression tag	UNP P08668
A	481	GLY	-	expression tag	UNP P08668
A	482	GLY	-	expression tag	UNP P08668
A	483	SER	-	expression tag	UNP P08668
A	484	GLY	-	expression tag	UNP P08668

Continued on next page...

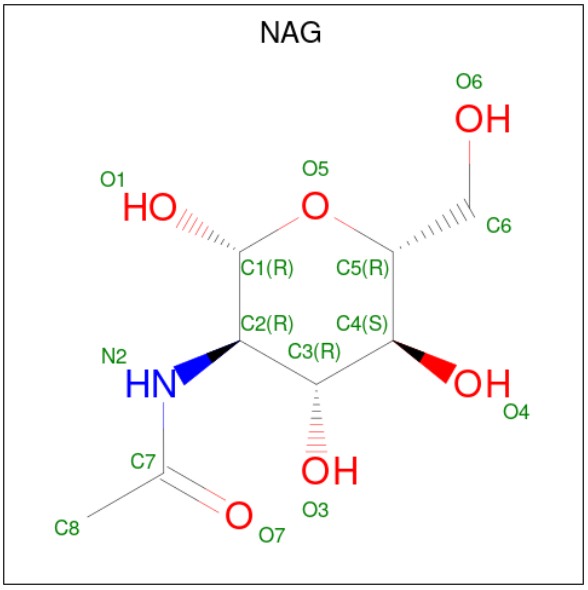
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	485	GLY	-	expression tag	UNP P08668
A	486	GLY	-	expression tag	UNP P08668
A	487	SER	-	expression tag	UNP P08668
A	488	TRP	-	expression tag	UNP P08668
A	489	SER	-	expression tag	UNP P08668
A	490	HIS	-	expression tag	UNP P08668
A	491	PRO	-	expression tag	UNP P08668
A	492	GLN	-	expression tag	UNP P08668
A	493	PHE	-	expression tag	UNP P08668
A	494	GLU	-	expression tag	UNP P08668
A	495	LYS	-	expression tag	UNP P08668

- Molecule 2 is a protein called scFvA5.

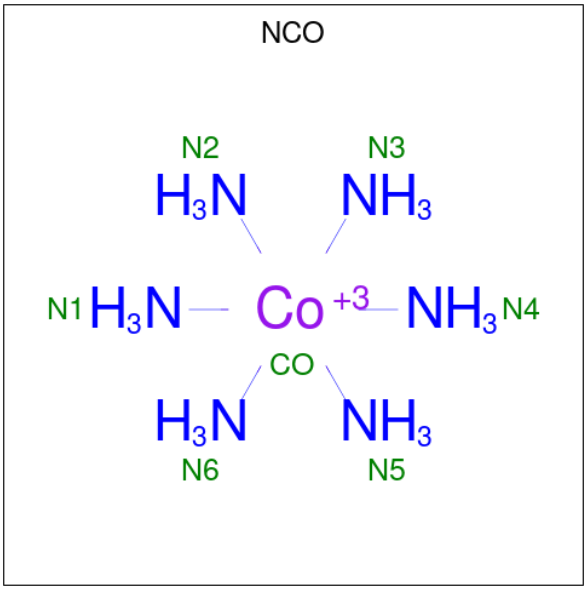
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	113	Total	C	N	O	S	0	0	0
			865	544	150	168	3			
2	L	107	Total	C	N	O	S	0	0	0
			798	489	141	166	2			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 4 is COBALT HEXAMMINE(III) (three-letter code: NCO) (formula: $CoH_{18}N_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Co N 7 1 6	0	0

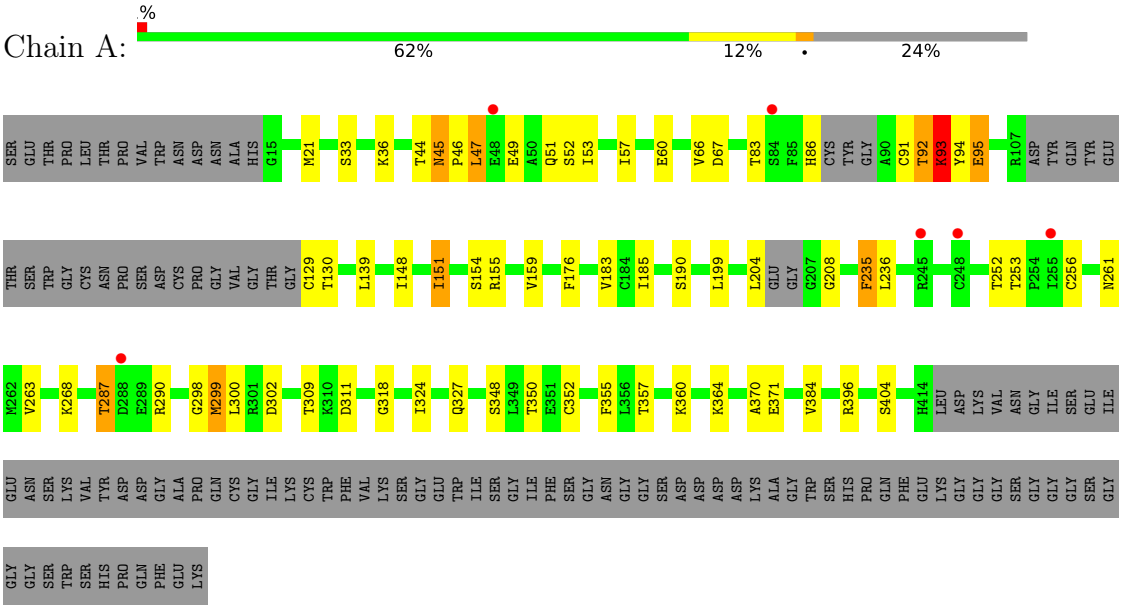
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total O 2 2	0	0

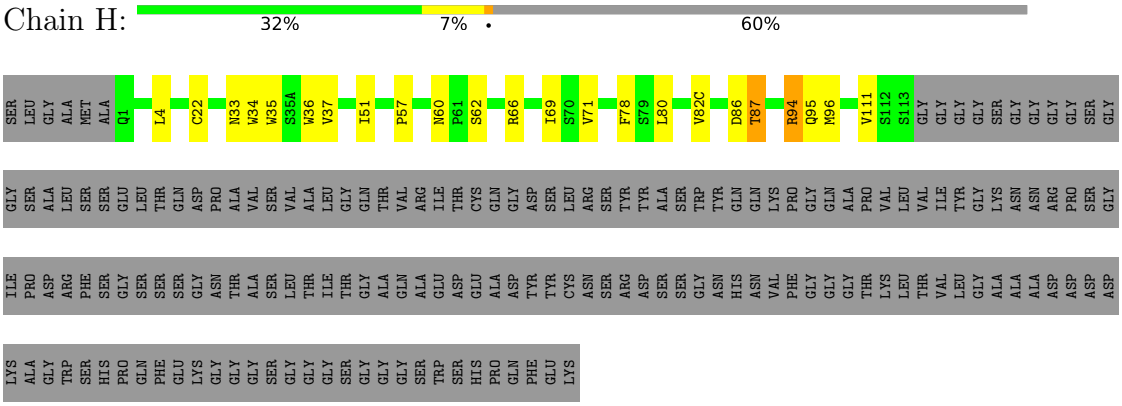
3 Residue-property plots (i)

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.

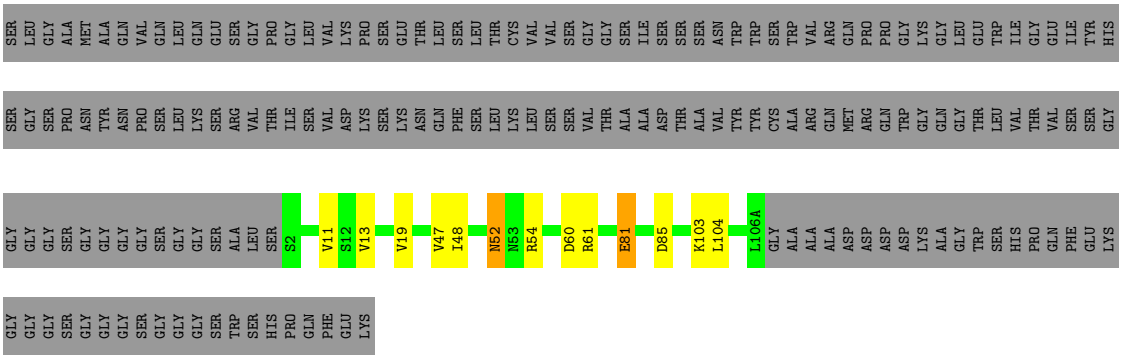
• Molecule 1: Envelopment polyprotein



• Molecule 2: scFvA5



• Molecule 2: scFvA5


SER LEU GLY ALA MET ALA GLN VAL GLN LEU GLN GLU SER GLY PRO GLY LEU VAL LYS PRO SER GLU THR LEU SER LEU THR CYS VAL VAL SER GLY GLY SER ILE SER SER SER ASN TRP TRP SER TRP VAL ARG GLN PRO PRO GLY LYS GLY LEU GLU TRP ILE GLY GLU ILE TYR HIS
SER GLY SER PRO ASN TYR ASN PRO SER LEU LYS SER ARG VAL THR ILE SER VAL ASP LYS SER LYS ASN GLN PHE SER LEU LYS LEU SER SER VAL THR ALA ALA ASP THR ALA VAL TYR TYR CYS ALA ARG GLN MET ARG GLN TRP GLY GLN GLY THR LEU VAL THR VAL SER SER GLY
GLY GLY GLY SER GLY GLY GLY SER GLY GLY SER ALA LEU SER S2 V11 S12 V13 V19 V47 I48 N52 N53 R54 D60 R61 E81 D85 K103 L104 L106A GLY ALA ALA ALA ASP ASP ASP ASP LYS ALA GLY TRP SER HIS PRO GLN PHE GLU LYS
GLY GLY GLY SER GLY GLY SER GLY GLY GLY SER TRP SER HIS PRO GLN PHE GLU LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	116.62Å 148.79Å 38.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.27 – 3.00 37.78 – 2.99	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.27-3.00) 99.9 (37.78-2.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.8.3_1471	Depositor
R, R_{free}	0.212 , 0.270 0.216 , 0.270	Depositor DCC
R_{free} test set	996 reflections (7.11%)	wwPDB-VP
Wilson B-factor (Å^2)	73.1	Xtriage
Anisotropy	0.655	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 48.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4594	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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: NCO, 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	0.31	0/2975	0.63	1/4018 (0.0%)
2	H	0.24	0/886	0.44	0/1206
2	L	0.26	0/813	0.44	0/1105
All	All	0.29	0/4674	0.57	1/6329 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	185	ILE	N-CA-C	-5.37	96.51	111.00

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	2908	0	2798	43	0
2	H	865	0	854	13	0
2	L	798	0	755	7	0
3	A	14	0	13	0	0
4	A	7	0	0	0	0
5	A	2	0	0	0	0
All	All	4594	0	4420	61	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 7.

All (61) 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:A:93:LYS:HA	1:A:93:LYS:HE2	1.54	0.87
1:A:91:CYS:O	1:A:92:THR:O	1.94	0.85
1:A:299:MET:SD	1:A:299:MET:N	2.51	0.82
1:A:45:ASN:ND2	1:A:46:PRO:O	2.16	0.79
2:H:22:CYS:HB3	2:H:78:PHE:HB3	1.70	0.72
2:H:51:ILE:HG12	2:H:57:PRO:HB3	1.73	0.69
1:A:93:LYS:O	1:A:93:LYS:NZ	2.20	0.69
2:L:47:VAL:HG12	2:L:48:ILE:HG12	1.81	0.63
1:A:287:THR:O	1:A:287:THR:OG1	2.14	0.63
1:A:91:CYS:O	1:A:92:THR:C	2.39	0.60
1:A:93:LYS:HA	1:A:93:LYS:CE	2.32	0.59
2:H:35:TRP:CZ3	2:H:94:ARG:HB3	2.41	0.56
2:H:35:TRP:CE3	2:H:94:ARG:HB3	2.42	0.55
1:A:49:GLU:HG3	1:A:51:GLN:HG2	1.90	0.54
2:L:13:VAL:HG21	2:L:19:VAL:HG22	1.90	0.54
1:A:355:PHE:CE2	1:A:357:THR:HB	2.43	0.54
1:A:46:PRO:O	1:A:47:LEU:HB3	2.08	0.53
1:A:190:SER:HA	1:A:300:LEU:HD22	1.90	0.53
2:H:82(C):VAL:HG13	2:H:86:ASP:HB2	1.91	0.53
1:A:309:THR:HG22	1:A:311:ASP:H	1.74	0.52
1:A:66:VAL:HG22	1:A:151:ILE:HG23	1.91	0.52
1:A:36:LYS:HG2	1:A:60:GLU:HA	1.90	0.52
1:A:44:THR:HG22	1:A:52:SER:HB3	1.91	0.52
2:H:87:THR:HB	2:H:111:VAL:H	1.75	0.51
1:A:91:CYS:C	1:A:92:THR:O	2.49	0.51
1:A:263:VAL:HB	1:A:268:LYS:HZ2	1.75	0.51
1:A:364:LYS:HZ1	1:A:396:ARG:HH12	1.58	0.51
1:A:66:VAL:HG11	1:A:148:ILE:HG12	1.93	0.50
1:A:93:LYS:O	1:A:93:LYS:HD3	2.12	0.49
2:H:66:ARG:NH1	2:H:86:ASP:OD2	2.37	0.49
1:A:364:LYS:NZ	1:A:396:ARG:HH12	2.11	0.49
1:A:199:LEU:O	1:A:208:GLY:HA2	2.12	0.48
1:A:21:MET:HB2	1:A:46:PRO:HD3	1.96	0.48
2:H:51:ILE:HD12	2:H:71:VAL:HG23	1.96	0.48
2:H:69:ILE:HG12	2:H:80:LEU:HD12	1.96	0.48
2:L:52:ASN:O	2:L:52:ASN:ND2	2.42	0.48
1:A:67:ASP:OD1	1:A:290:ARG:NH1	2.48	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:L:11:VAL:O	2:L:104:LEU:HD12	2.15	0.47
1:A:53:ILE:HG12	1:A:318:GLY:HA2	1.97	0.47
2:H:36:TRP:CE2	2:H:80:LEU:HB2	2.51	0.46
1:A:93:LYS:HE2	1:A:93:LYS:CA	2.36	0.46
1:A:33:SER:HA	1:A:155:ARG:NH2	2.30	0.46
1:A:57:ILE:HG12	1:A:159:VAL:HG22	1.98	0.45
2:L:85:ASP:OD1	2:L:103:LYS:HG3	2.16	0.45
1:A:396:ARG:HD3	1:A:404:SER:O	2.16	0.45
1:A:324:ILE:HA	1:A:348:SER:O	2.17	0.44
2:L:54:ARG:NH2	2:L:60:ASP:HA	2.32	0.44
1:A:176:PHE:HB3	1:A:183:VAL:HB	1.99	0.44
2:L:61:ARG:NH2	2:L:81:GLU:OE2	2.50	0.43
1:A:93:LYS:HZ2	1:A:95:GLU:HG3	1.84	0.43
1:A:261:ASN:OD1	1:A:268:LYS:NZ	2.34	0.42
1:A:83:THR:HG22	1:A:256:CYS:HB2	2.01	0.42
1:A:360:LYS:HG2	1:A:370:ALA:HB2	2.01	0.42
2:H:60:ASN:OD1	2:H:62:SER:OG	2.37	0.42
1:A:371:GLU:HB3	1:A:384:VAL:HG22	2.01	0.42
1:A:94:TYR:CD1	2:H:33:ASN:HB2	2.55	0.42
1:A:92:THR:O	1:A:94:TYR:N	2.53	0.42
1:A:93:LYS:CE	1:A:93:LYS:CA	2.95	0.42
1:A:93:LYS:O	1:A:94:TYR:C	2.55	0.41
1:A:95:GLU:O	2:H:33:ASN:ND2	2.52	0.41
1:A:139:LEU:HD13	1:A:235:PHE:CZ	2.57	0.40

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	366/495 (74%)	337 (92%)	25 (7%)	4 (1%)	14 50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	H	111/281 (40%)	106 (96%)	5 (4%)	0	100 100
2	L	105/281 (37%)	102 (97%)	3 (3%)	0	100 100
All	All	582/1057 (55%)	545 (94%)	33 (6%)	4 (1%)	22 60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	92	THR
1	A	93	LYS
1	A	298	GLY
1	A	253	THR

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	325/419 (78%)	306 (94%)	19 (6%)	20 55
2	H	99/218 (45%)	92 (93%)	7 (7%)	14 46
2	L	87/218 (40%)	85 (98%)	2 (2%)	50 80
All	All	511/855 (60%)	483 (94%)	28 (6%)	21 57

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	A	47	LEU
1	A	86	HIS
1	A	93	LYS
1	A	95	GLU
1	A	129	CYS
1	A	130	THR
1	A	151	ILE
1	A	154	SER
1	A	204	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	235	PHE
1	A	236	LEU
1	A	252	THR
1	A	287	THR
1	A	299	MET
1	A	302	ASP
1	A	327	GLN
1	A	350	THR
1	A	352	CYS
2	H	4	LEU
2	H	34	TRP
2	H	37	VAL
2	H	87	THR
2	H	94	ARG
2	H	95	GLN
2	H	96	MET
2	L	52	ASN
2	L	81	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

2 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	$\#\|Z\| > 2$
3	NAG	A	501	1	14,14,15	0.40	0	17,19,21	0.40	0
4	NCO	A	502	-	6,6,6	0.45	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
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	374/495 (75%)	-0.19	6 (1%) 72 44	40, 76, 146, 219	0
2	H	113/281 (40%)	-0.15	0 100 100	61, 89, 121, 139	0
2	L	107/281 (38%)	-0.21	0 100 100	47, 69, 104, 140	0
All	All	594/1057 (56%)	-0.19	6 (1%) 82 59	40, 78, 138, 219	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	SER	3.4
1	A	245	ARG	2.4
1	A	248	CYS	2.3
1	A	255	ILE	2.2
1	A	288	ASP	2.2
1	A	48	GLU	2.2

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, 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
3	NAG	A	501	14/15	0.86	0.19	77,88,92,96	0
4	NCO	A	502	7/7	0.92	0.18	60,101,105,161	0

6.5 Other polymers (i)

There are no such residues in this entry.