



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:17 PM GMT

PDB ID : 4HSI
Title : Glycoprotein B from Herpes simplex virus type 1, A504P/R505G/Q507G/N511G mutant, low-pH
Authors : Heldwein, E.E.
Deposited on : 2012-10-30
Resolution : 3.10 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

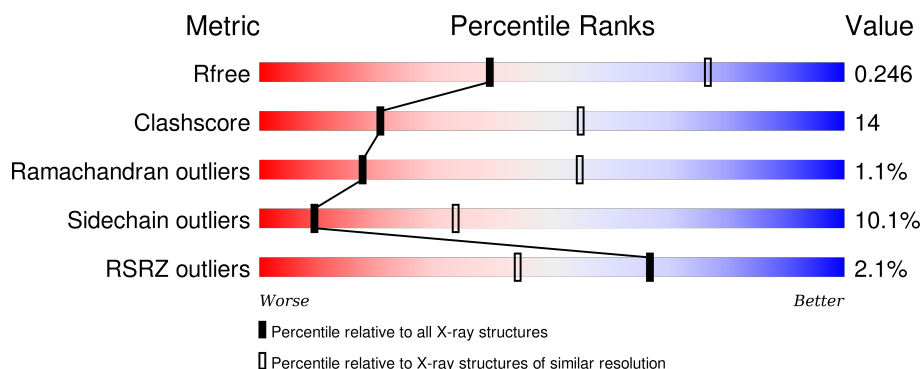
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	703	<div> <div>2%</div> <div>56%</div> <div>26%</div> <div>•</div> <div>15%</div> </div>
1	B	703	<div> <div>3%</div> <div>50%</div> <div>31%</div> <div>5%</div> <div>14%</div> </div>
1	C	703	<div> <div>3%</div> <div>52%</div> <div>30%</div> <div>•</div> <div>14%</div> </div>
1	D	703	<div> <div>•</div> <div>57%</div> <div>27%</div> <div>•</div> <div>14%</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	C	801	X	-	-	-
3	NAG	C	803	X	-	-	-
5	CL	B	803	-	-	-	X
5	CL	D	803	-	-	-	X

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 19653 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 Envelope glycoprotein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	601	Total	C	N	O	S	0	0	0
			4846	3057	851	916	22			
1	B	606	Total	C	N	O	S	0	0	0
			4891	3085	862	922	22			
1	C	602	Total	C	N	O	S	0	0	0
			4856	3063	854	917	22			
1	D	605	Total	C	N	O	S	0	0	0
			4882	3079	862	919	22			

There are 156 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	ASP	-	EXPRESSION TAG	UNP P06437
A	29	PRO	-	EXPRESSION TAG	UNP P06437
A	30	ALA	-	EXPRESSION TAG	UNP P06437
A	31	ALA	-	EXPRESSION TAG	UNP P06437
A	32	PRO	-	EXPRESSION TAG	UNP P06437
A	33	THR	-	EXPRESSION TAG	UNP P06437
A	34	SER	-	EXPRESSION TAG	UNP P06437
A	35	PRO	-	EXPRESSION TAG	UNP P06437
A	36	GLY	-	EXPRESSION TAG	UNP P06437
A	37	THR	-	EXPRESSION TAG	UNP P06437
A	38	PRO	-	EXPRESSION TAG	UNP P06437
A	39	GLY	-	EXPRESSION TAG	UNP P06437
A	40	VAL	-	EXPRESSION TAG	UNP P06437
A	41	ALA	-	EXPRESSION TAG	UNP P06437
A	42	ALA	-	EXPRESSION TAG	UNP P06437
A	43	ALA	-	EXPRESSION TAG	UNP P06437
A	44	THR	-	EXPRESSION TAG	UNP P06437
A	45	GLN	-	EXPRESSION TAG	UNP P06437
A	46	ALA	-	EXPRESSION TAG	UNP P06437
A	47	ALA	-	EXPRESSION TAG	UNP P06437
A	48	ASN	-	EXPRESSION TAG	UNP P06437

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	GLY	-	EXPRESSION TAG	UNP P06437
A	50	GLY	-	EXPRESSION TAG	UNP P06437
A	51	PRO	-	EXPRESSION TAG	UNP P06437
A	52	ALA	-	EXPRESSION TAG	UNP P06437
A	53	THR	-	EXPRESSION TAG	UNP P06437
A	54	PRO	-	EXPRESSION TAG	UNP P06437
A	55	ALA	-	EXPRESSION TAG	UNP P06437
A	56	PRO	-	EXPRESSION TAG	UNP P06437
A	57	PRO	-	EXPRESSION TAG	UNP P06437
A	58	PRO	-	EXPRESSION TAG	UNP P06437
A	59	LEU	-	EXPRESSION TAG	UNP P06437
A	60	GLY	-	EXPRESSION TAG	UNP P06437
A	313	SER	THR	CONFLICT	UNP P06437
A	443	LEU	GLN	CONFLICT	UNP P06437
A	504	PRO	ALA	ENGINEERED MUTATION	UNP P06437
A	505	GLY	ARG	ENGINEERED MUTATION	UNP P06437
A	507	GLY	GLN	ENGINEERED MUTATION	UNP P06437
A	511	GLY	ASN	ENGINEERED MUTATION	UNP P06437
B	28	ASP	-	EXPRESSION TAG	UNP P06437
B	29	PRO	-	EXPRESSION TAG	UNP P06437
B	30	ALA	-	EXPRESSION TAG	UNP P06437
B	31	ALA	-	EXPRESSION TAG	UNP P06437
B	32	PRO	-	EXPRESSION TAG	UNP P06437
B	33	THR	-	EXPRESSION TAG	UNP P06437
B	34	SER	-	EXPRESSION TAG	UNP P06437
B	35	PRO	-	EXPRESSION TAG	UNP P06437
B	36	GLY	-	EXPRESSION TAG	UNP P06437
B	37	THR	-	EXPRESSION TAG	UNP P06437
B	38	PRO	-	EXPRESSION TAG	UNP P06437
B	39	GLY	-	EXPRESSION TAG	UNP P06437
B	40	VAL	-	EXPRESSION TAG	UNP P06437
B	41	ALA	-	EXPRESSION TAG	UNP P06437
B	42	ALA	-	EXPRESSION TAG	UNP P06437
B	43	ALA	-	EXPRESSION TAG	UNP P06437
B	44	THR	-	EXPRESSION TAG	UNP P06437
B	45	GLN	-	EXPRESSION TAG	UNP P06437
B	46	ALA	-	EXPRESSION TAG	UNP P06437
B	47	ALA	-	EXPRESSION TAG	UNP P06437
B	48	ASN	-	EXPRESSION TAG	UNP P06437
B	49	GLY	-	EXPRESSION TAG	UNP P06437
B	50	GLY	-	EXPRESSION TAG	UNP P06437
B	51	PRO	-	EXPRESSION TAG	UNP P06437

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	52	ALA	-	EXPRESSION TAG	UNP P06437
B	53	THR	-	EXPRESSION TAG	UNP P06437
B	54	PRO	-	EXPRESSION TAG	UNP P06437
B	55	ALA	-	EXPRESSION TAG	UNP P06437
B	56	PRO	-	EXPRESSION TAG	UNP P06437
B	57	PRO	-	EXPRESSION TAG	UNP P06437
B	58	PRO	-	EXPRESSION TAG	UNP P06437
B	59	LEU	-	EXPRESSION TAG	UNP P06437
B	60	GLY	-	EXPRESSION TAG	UNP P06437
B	313	SER	THR	CONFLICT	UNP P06437
B	443	LEU	GLN	CONFLICT	UNP P06437
B	504	PRO	ALA	ENGINEERED MUTATION	UNP P06437
B	505	GLY	ARG	ENGINEERED MUTATION	UNP P06437
B	507	GLY	GLN	ENGINEERED MUTATION	UNP P06437
B	511	GLY	ASN	ENGINEERED MUTATION	UNP P06437
C	28	ASP	-	EXPRESSION TAG	UNP P06437
C	29	PRO	-	EXPRESSION TAG	UNP P06437
C	30	ALA	-	EXPRESSION TAG	UNP P06437
C	31	ALA	-	EXPRESSION TAG	UNP P06437
C	32	PRO	-	EXPRESSION TAG	UNP P06437
C	33	THR	-	EXPRESSION TAG	UNP P06437
C	34	SER	-	EXPRESSION TAG	UNP P06437
C	35	PRO	-	EXPRESSION TAG	UNP P06437
C	36	GLY	-	EXPRESSION TAG	UNP P06437
C	37	THR	-	EXPRESSION TAG	UNP P06437
C	38	PRO	-	EXPRESSION TAG	UNP P06437
C	39	GLY	-	EXPRESSION TAG	UNP P06437
C	40	VAL	-	EXPRESSION TAG	UNP P06437
C	41	ALA	-	EXPRESSION TAG	UNP P06437
C	42	ALA	-	EXPRESSION TAG	UNP P06437
C	43	ALA	-	EXPRESSION TAG	UNP P06437
C	44	THR	-	EXPRESSION TAG	UNP P06437
C	45	GLN	-	EXPRESSION TAG	UNP P06437
C	46	ALA	-	EXPRESSION TAG	UNP P06437
C	47	ALA	-	EXPRESSION TAG	UNP P06437
C	48	ASN	-	EXPRESSION TAG	UNP P06437
C	49	GLY	-	EXPRESSION TAG	UNP P06437
C	50	GLY	-	EXPRESSION TAG	UNP P06437
C	51	PRO	-	EXPRESSION TAG	UNP P06437
C	52	ALA	-	EXPRESSION TAG	UNP P06437
C	53	THR	-	EXPRESSION TAG	UNP P06437
C	54	PRO	-	EXPRESSION TAG	UNP P06437

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	55	ALA	-	EXPRESSION TAG	UNP P06437
C	56	PRO	-	EXPRESSION TAG	UNP P06437
C	57	PRO	-	EXPRESSION TAG	UNP P06437
C	58	PRO	-	EXPRESSION TAG	UNP P06437
C	59	LEU	-	EXPRESSION TAG	UNP P06437
C	60	GLY	-	EXPRESSION TAG	UNP P06437
C	313	SER	THR	CONFLICT	UNP P06437
C	443	LEU	GLN	CONFLICT	UNP P06437
C	504	PRO	ALA	ENGINEERED MUTATION	UNP P06437
C	505	GLY	ARG	ENGINEERED MUTATION	UNP P06437
C	507	GLY	GLN	ENGINEERED MUTATION	UNP P06437
C	511	GLY	ASN	ENGINEERED MUTATION	UNP P06437
D	28	ASP	-	EXPRESSION TAG	UNP P06437
D	29	PRO	-	EXPRESSION TAG	UNP P06437
D	30	ALA	-	EXPRESSION TAG	UNP P06437
D	31	ALA	-	EXPRESSION TAG	UNP P06437
D	32	PRO	-	EXPRESSION TAG	UNP P06437
D	33	THR	-	EXPRESSION TAG	UNP P06437
D	34	SER	-	EXPRESSION TAG	UNP P06437
D	35	PRO	-	EXPRESSION TAG	UNP P06437
D	36	GLY	-	EXPRESSION TAG	UNP P06437
D	37	THR	-	EXPRESSION TAG	UNP P06437
D	38	PRO	-	EXPRESSION TAG	UNP P06437
D	39	GLY	-	EXPRESSION TAG	UNP P06437
D	40	VAL	-	EXPRESSION TAG	UNP P06437
D	41	ALA	-	EXPRESSION TAG	UNP P06437
D	42	ALA	-	EXPRESSION TAG	UNP P06437
D	43	ALA	-	EXPRESSION TAG	UNP P06437
D	44	THR	-	EXPRESSION TAG	UNP P06437
D	45	GLN	-	EXPRESSION TAG	UNP P06437
D	46	ALA	-	EXPRESSION TAG	UNP P06437
D	47	ALA	-	EXPRESSION TAG	UNP P06437
D	48	ASN	-	EXPRESSION TAG	UNP P06437
D	49	GLY	-	EXPRESSION TAG	UNP P06437
D	50	GLY	-	EXPRESSION TAG	UNP P06437
D	51	PRO	-	EXPRESSION TAG	UNP P06437
D	52	ALA	-	EXPRESSION TAG	UNP P06437
D	53	THR	-	EXPRESSION TAG	UNP P06437
D	54	PRO	-	EXPRESSION TAG	UNP P06437
D	55	ALA	-	EXPRESSION TAG	UNP P06437
D	56	PRO	-	EXPRESSION TAG	UNP P06437
D	57	PRO	-	EXPRESSION TAG	UNP P06437

Continued on next page...

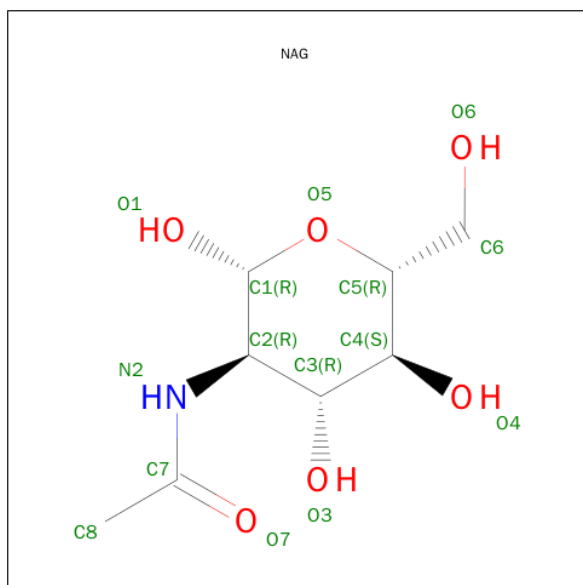
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	58	PRO	-	EXPRESSION TAG	UNP P06437
D	59	LEU	-	EXPRESSION TAG	UNP P06437
D	60	GLY	-	EXPRESSION TAG	UNP P06437
D	313	SER	THR	CONFLICT	UNP P06437
D	443	LEU	GLN	CONFLICT	UNP P06437
D	504	PRO	ALA	ENGINEERED MUTATION	UNP P06437
D	505	GLY	ARG	ENGINEERED MUTATION	UNP P06437
D	507	GLY	GLN	ENGINEERED MUTATION	UNP P06437
D	511	GLY	ASN	ENGINEERED MUTATION	UNP P06437

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	C	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



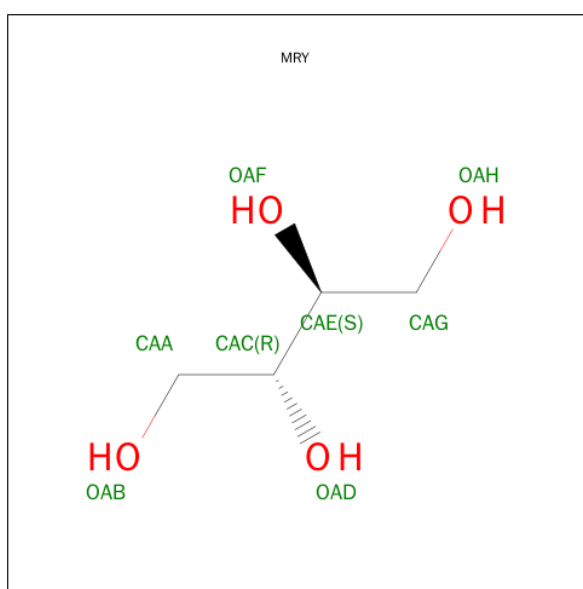
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

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

- Molecule 4 is MESO-ERYTHRITOL (three-letter code: MRY) (formula: $C_4H_{10}O_4$).



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

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Cl	0	0
			1	1		
5	A	1	Total	Cl	0	0
			1	1		
5	D	1	Total	Cl	0	0
			1	1		

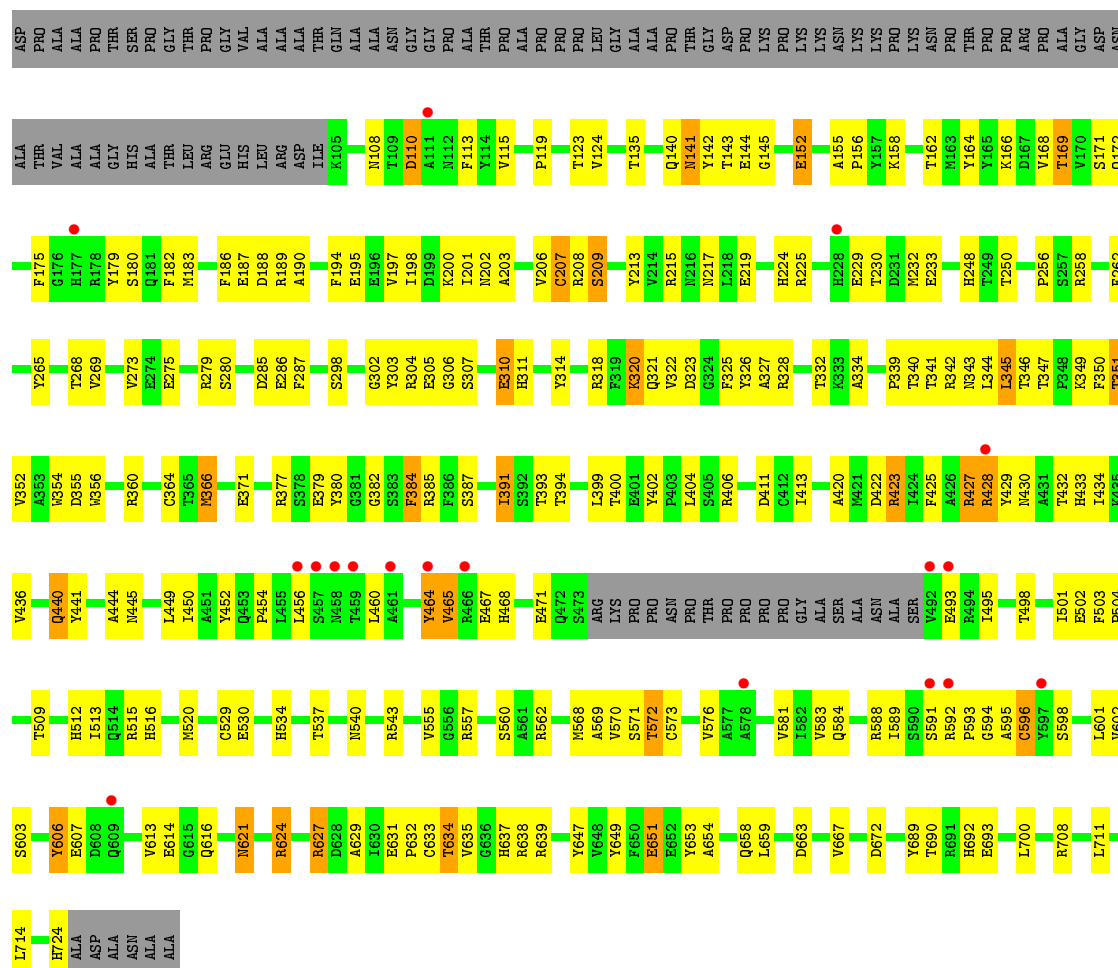
Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	1	Total	Cl	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	8	Total	O	0	0
			8	8		
6	B	1	Total	O	0	0
			1	1		
6	C	3	Total	O	0	0
			3	3		
6	D	6	Total	O	0	0
			6	6		



● Molecule 1: Envelope glycoprotein B



ASP	PRO	ALA	ALA	PRO	THR	SER	PRO	GLY	THR	PRO	GLY	VAL	ALA	ALA	THR	GLN	ALA	ALA	ASN	GLY	PRO	PRO	PRO	PRO	LEU	GLY	ALA	ALA	PRO	PRO	PRO	PRO	LYS	PRO	LYS	LYS	ASN	LYS	PRO	LYS	ASN	PRO	THR	PRO	PRO	ARG	PRO	ALA	GLY	ASP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
		ALA	THR	VAL	ALA	ALA	ALA	GLY	HIS	THR	LEU	ARG	GLU	HIS	LEU	ALA	ALA	ALA	THR	D103	T104	K105																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																

4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α , β , γ	117.30Å 117.30Å 321.87Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.91 – 3.10 45.91 – 3.10	Depositor EDS
% Data completeness (in resolution range)	82.8 (45.91-3.10) 75.7 (45.91-3.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.16 (at 3.12Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.205 , 0.255 0.200 , 0.246	Depositor DCC
R_{free} test set	3442 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	52.6	Xtriage
Anisotropy	0.378	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 50.9	EDS
Estimated twinning fraction	0.020 for -h,-k,l 0.459 for h,-h-k,-l 0.026 for -k,-h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 74374 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	19653	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.84 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.7087e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, MRY, 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	0.49	0/4966	0.64	2/6747 (0.0%)
1	B	0.45	0/5013	0.62	0/6809
1	C	0.44	0/4977	0.62	0/6762
1	D	0.46	0/5004	0.62	0/6797
All	All	0.46	0/19960	0.63	2/27115 (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
2	C	1	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	515	ARG	NE-CZ-NH1	6.76	123.68	120.30
1	A	515	ARG	NE-CZ-NH2	-5.60	117.50	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	801	NAG	C1

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	4846	0	4657	117	1
1	B	4891	0	4703	168	0
1	C	4856	0	4664	151	1
1	D	4882	0	4694	121	0
2	A	28	0	25	0	0
2	C	28	0	25	0	0
3	A	14	0	13	0	0
3	B	28	0	26	1	0
3	C	14	0	13	0	0
3	D	28	0	26	1	0
4	A	8	0	10	0	0
4	C	8	0	10	0	0
5	A	1	0	0	1	0
5	B	1	0	0	0	0
5	C	1	0	0	1	0
5	D	1	0	0	0	0
6	A	8	0	0	0	0
6	B	1	0	0	0	0
6	C	3	0	0	0	0
6	D	6	0	0	0	0
All	All	19653	0	18866	557	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 14.

The worst 5 of 557 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:515:ARG:NH2	5:A:805:CL:CL	2.36	0.95
1:C:515:ARG:NH2	5:C:805:CL:CL	2.40	0.91
1:C:614:GLU:HB3	1:C:627:ARG:HH21	1.37	0.87
1:D:116:CYS:HB3	1:D:560:SER:HB3	1.63	0.81
1:B:587:MET:HB3	1:B:653:TYR:HD2	1.48	0.79

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:152:GLU:O	1:A:689:TYR:OH[3_555]	2.13	0.07
1:C:152:GLU:O	1:C:689:TYR:OH[3_555]	2.18	0.02

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	597/703 (85%)	529 (89%)	66 (11%)	2 (0%)	46	80
1	B	600/703 (85%)	537 (90%)	54 (9%)	9 (2%)	13	46
1	C	598/703 (85%)	527 (88%)	64 (11%)	7 (1%)	16	52
1	D	599/703 (85%)	537 (90%)	54 (9%)	8 (1%)	15	50
All	All	2394/2812 (85%)	2130 (89%)	238 (10%)	26 (1%)	17	55

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	237	ALA
1	B	413	ILE
1	C	413	ILE
1	D	462	GLU
1	C	465	VAL

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	522/592 (88%)	463 (89%)	59 (11%)	7	28
1	B	528/592 (89%)	466 (88%)	62 (12%)	7	26
1	C	523/592 (88%)	470 (90%)	53 (10%)	9	33
1	D	526/592 (89%)	489 (93%)	37 (7%)	19	54
All	All	2099/2368 (89%)	1888 (90%)	211 (10%)	9	33

5 of 211 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	459	THR
1	B	663	ASP
1	D	432	THR
1	B	471	GLU
1	B	616	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	440	GLN
1	A	658	GLN
1	C	468	HIS
1	D	440	GLN

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](#)

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	$\# Z > 2$
2	NAG	A	801	1,2	14,14,15	0.60	0	15,19,21	1.12	2 (13%)
2	NAG	A	802	2	14,14,15	0.54	0	15,19,21	1.07	2 (13%)
2	NAG	C	801	1,2	14,14,15	0.60	0	15,19,21	1.04	1 (6%)
2	NAG	C	802	2	14,14,15	0.54	0	15,19,21	1.32	2 (13%)

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	801	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	802	2	-	0/6/23/26	0/1/1/1
2	NAG	C	801	1,2	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	C	802	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	802	NAG	C2-N2-C7	-2.32	120.06	123.04
2	A	802	NAG	C4-C3-C2	-2.09	107.98	111.23
2	A	802	NAG	C1-O5-C5	2.12	114.93	112.25
2	C	801	NAG	C4-C3-C2	2.32	114.84	111.23
2	A	801	NAG	C3-C4-C5	2.38	114.35	110.20

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	C	801	NAG	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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$
3	NAG	A	803	1	14,14,15	0.69	0	15,19,21	1.04	1 (6%)
4	MRY	A	804	-	7,7,7	0.37	0	6,8,8	0.75	0
3	NAG	B	801	1	14,14,15	0.54	0	15,19,21	1.87	3 (20%)
3	NAG	B	802	1	14,14,15	0.58	0	15,19,21	1.39	1 (6%)
3	NAG	C	803	1	14,14,15	0.45	0	15,19,21	1.11	1 (6%)
4	MRY	C	804	-	7,7,7	0.31	0	6,8,8	1.01	0
3	NAG	D	801	1	14,14,15	0.41	0	15,19,21	1.51	1 (6%)
3	NAG	D	802	1	14,14,15	0.57	0	15,19,21	1.37	2 (13%)

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	803	1	-	0/6/23/26	0/1/1/1
4	MRY	A	804	-	-	0/8/8/8	0/0/0/0
3	NAG	B	801	1	-	0/6/23/26	0/1/1/1
3	NAG	B	802	1	-	0/6/23/26	0/1/1/1
3	NAG	C	803	1	1/1/5/7	0/6/23/26	0/1/1/1
4	MRY	C	804	-	-	0/8/8/8	0/0/0/0
3	NAG	D	801	1	-	0/6/23/26	0/1/1/1
3	NAG	D	802	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	802	NAG	C2-N2-C7	-4.21	117.63	123.04

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	802	NAG	C2-N2-C7	-3.94	117.98	123.04
3	D	802	NAG	C4-C3-C2	-2.51	107.33	111.23
3	C	803	NAG	C2-N2-C7	-2.50	119.83	123.04
3	B	801	NAG	O7-C7-N2	2.17	126.29	121.86

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	803	NAG	C1

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	801	NAG	1	0
3	D	801	NAG	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	601/703 (85%)	-0.06	11 (1%) 71 50	1, 37, 84, 140	0
1	B	606/703 (86%)	0.09	18 (2%) 54 29	3, 44, 95, 185	0
1	C	602/703 (85%)	0.07	18 (2%) 54 29	7, 44, 98, 157	0
1	D	605/703 (86%)	-0.07	4 (0%) 89 78	2, 39, 78, 147	1 (0%)
All	All	2414/2812 (85%)	0.01	51 (2%) 67 44	1, 41, 89, 185	1 (0%)

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	108	ASN	5.1
1	B	493	GLU	4.6
1	B	110	ASP	4.5
1	A	264	ARG	4.1
1	D	109	THR	3.9

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [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	A	802	14/15	0.88	0.19	-	86,86,86,86	0
2	NAG	C	802	14/15	0.81	0.26	-	92,92,92,92	0
2	NAG	C	801	14/15	0.80	0.25	-	88,88,88,88	0
2	NAG	A	801	14/15	0.83	0.24	-	92,92,92,92	0

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
5	CL	B	803	1/1	0.88	0.44	7.55	67,67,67,67	0
5	CL	D	803	1/1	0.95	0.49	5.06	71,71,71,71	0
4	MRY	A	804	8/8	0.89	0.24	0.93	44,44,44,44	0
4	MRY	C	804	8/8	0.88	0.22	0.02	34,34,34,34	0
3	NAG	D	801	14/15	0.89	0.17	-0.55	71,71,71,71	0
5	CL	C	805	1/1	0.90	0.17	-0.93	32,32,32,32	0
3	NAG	B	801	14/15	0.90	0.15	-1.03	69,69,69,69	0
5	CL	A	805	1/1	0.92	0.15	-1.27	30,30,30,30	0
3	NAG	A	803	14/15	0.86	0.32	-	71,71,71,71	0
3	NAG	C	803	14/15	0.82	0.34	-	71,71,71,71	0
3	NAG	D	802	14/15	0.94	0.22	-	27,27,27,27	0
3	NAG	B	802	14/15	0.92	0.25	-	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.