wwPDB X-ray Structure Validation Summary Report (i)

Aug 30, 2023 – 04:10 AM EDT

PDB ID	:	3NWA
Title	:	Glycoprotein B from Herpes simplex virus type 1, W174R mutant, low-pH
Authors	:	Stampfer, S.D.; Lou, H.; Cohen, G.H.; Eisenberg, R.J.; Heldwein, E.E.
Deposited on	:	2010-07-09
Resolution	:	2.26 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
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 (i)) were used in the production of this report:

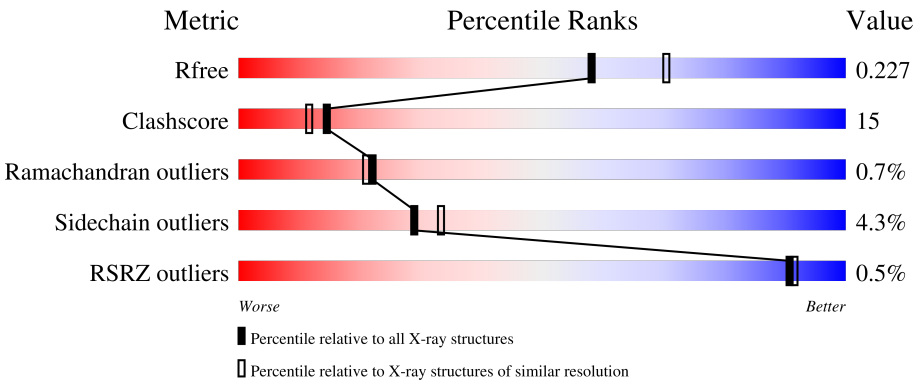
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.26 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	703	62% 23% • 14%
1	B	703	% 56% 29% • 14%
1	C	703	% 59% 26% • 14%
1	D	703	59% 26% • 14%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 21541 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
			Total	C	N	O	S			
1	B	607	4896	3086	867	920	23	0	1	0
1	A	608	4899	3088	867	922	22	0	1	0
1	C	608	4883	3080	864	917	22	0	0	0
1	D	607	4900	3087	867	923	23	0	1	0

There are 24 discrepancies between the modelled and reference sequences:

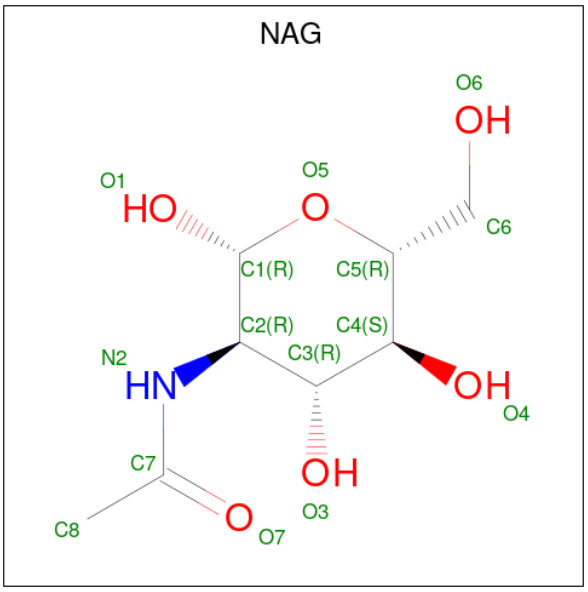
Chain	Residue	Modelled	Actual	Comment	Reference
B	28	ASP	-	expression tag	UNP P06437
B	29	PRO	-	expression tag	UNP P06437
B	58	ALA	PRO	SEE REMARK 999	UNP P06437
B	174	ARG	TRP	engineered mutation	UNP P06437
B	313	SER	THR	SEE REMARK 999	UNP P06437
B	443	LEU	GLN	SEE REMARK 999	UNP P06437
A	28	ASP	-	expression tag	UNP P06437
A	29	PRO	-	expression tag	UNP P06437
A	58	ALA	PRO	SEE REMARK 999	UNP P06437
A	174	ARG	TRP	engineered mutation	UNP P06437
A	313	SER	THR	SEE REMARK 999	UNP P06437
A	443	LEU	GLN	SEE REMARK 999	UNP P06437
C	28	ASP	-	expression tag	UNP P06437
C	29	PRO	-	expression tag	UNP P06437
C	58	ALA	PRO	SEE REMARK 999	UNP P06437
C	174	ARG	TRP	engineered mutation	UNP P06437
C	313	SER	THR	SEE REMARK 999	UNP P06437
C	443	LEU	GLN	SEE REMARK 999	UNP P06437
D	28	ASP	-	expression tag	UNP P06437
D	29	PRO	-	expression tag	UNP P06437
D	58	ALA	PRO	SEE REMARK 999	UNP P06437

Continued on next page...

Continued from previous page...

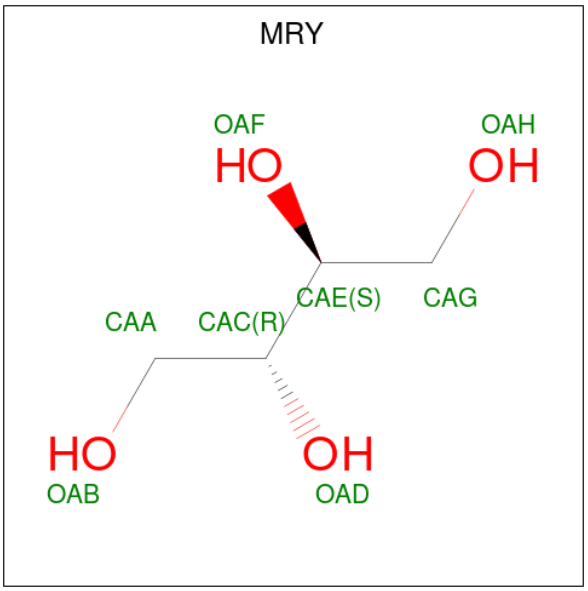
Chain	Residue	Modelled	Actual	Comment	Reference
D	174	ARG	TRP	engineered mutation	UNP P06437
D	313	SER	THR	SEE REMARK 999	UNP P06437
D	443	LEU	GLN	SEE REMARK 999	UNP P06437

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



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

- Molecule 3 is MESO-ERYTHRITOL (three-letter code: MRY) (formula: $C_4H_{10}O_4$).



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

- Molecule 4 is water.

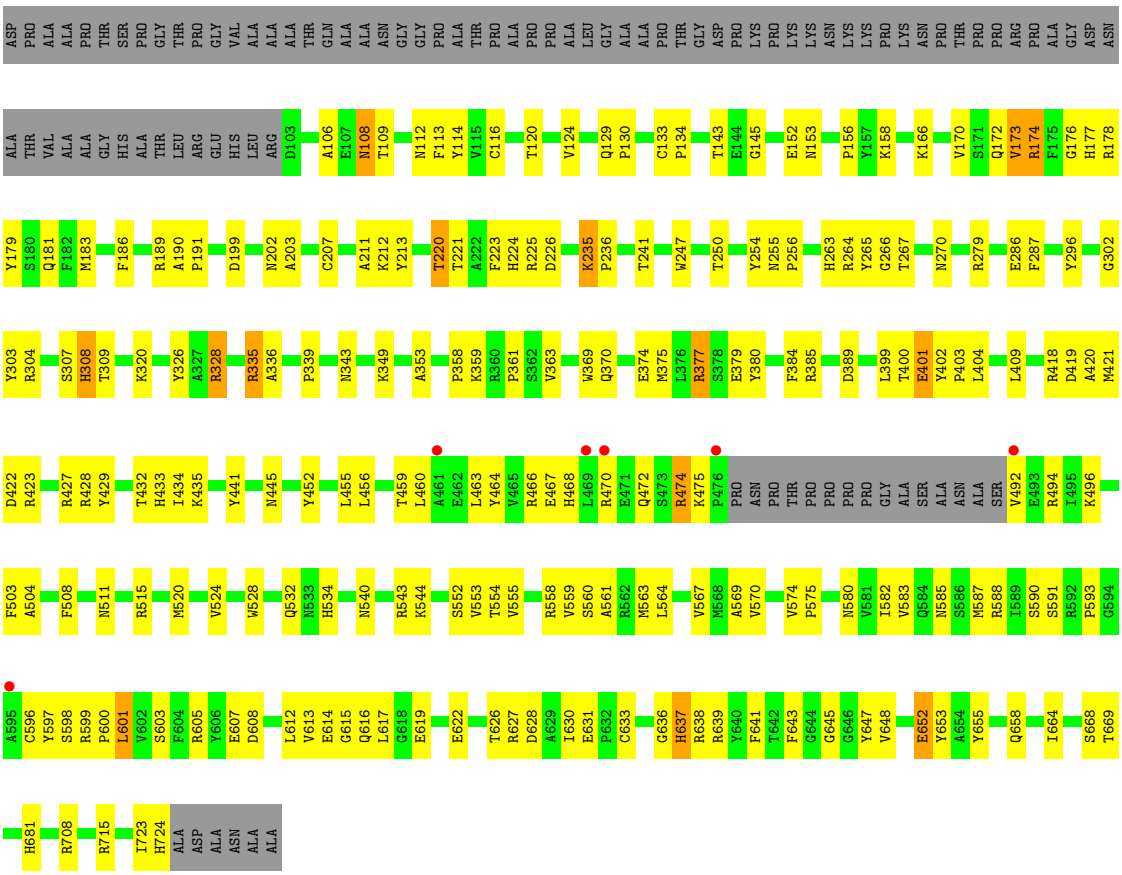
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	444	Total O 444 444	0	0
4	A	478	Total O 478 478	0	0
4	C	462	Total O 462 462	0	0
4	D	483	Total O 483 483	0	0

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.

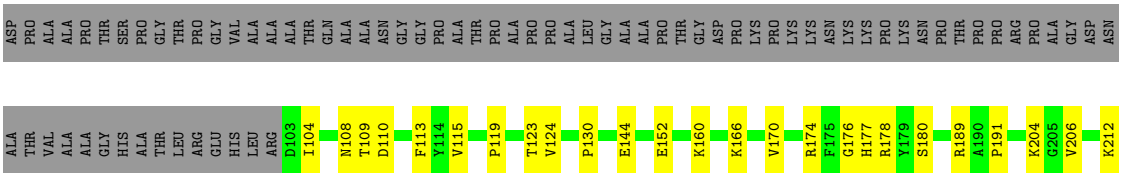
• Molecule 1: Envelope glycoprotein B

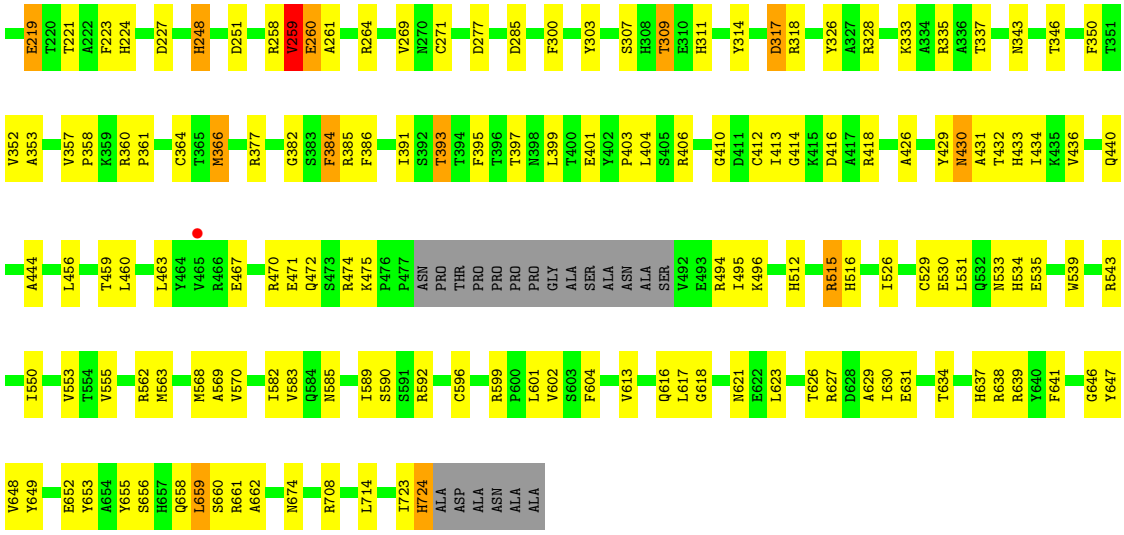
Chain B:



• Molecule 1: Envelope glycoprotein B

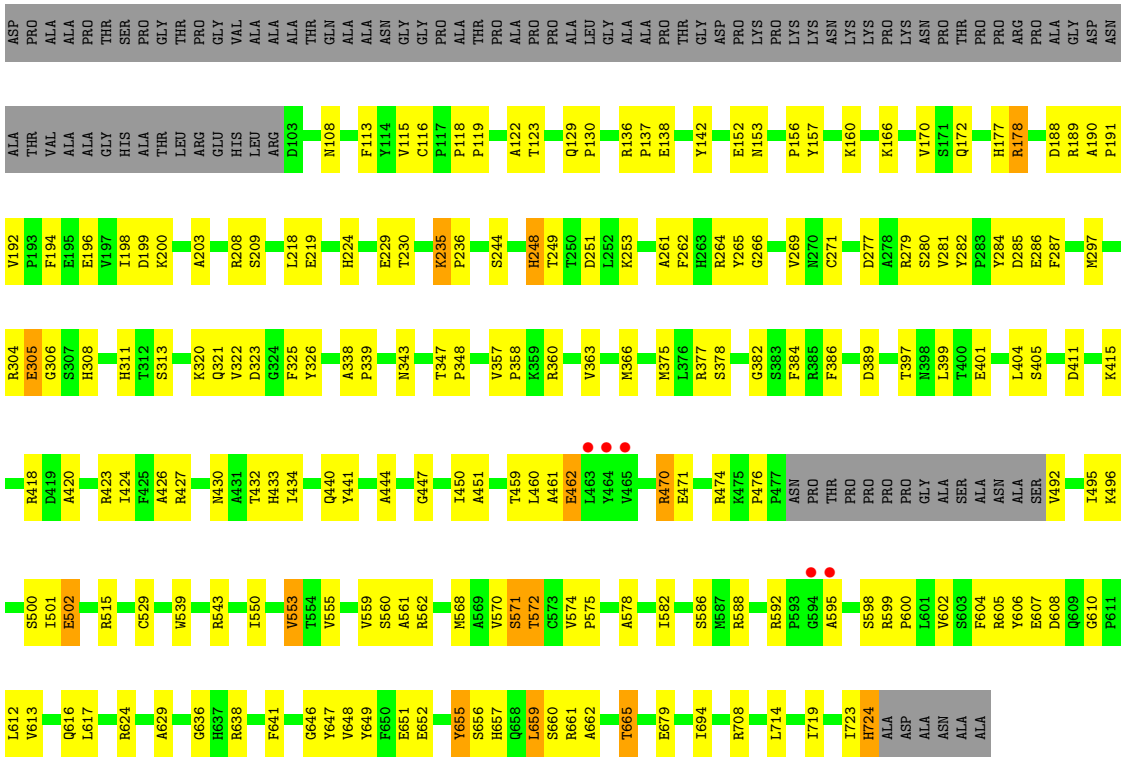
Chain A:





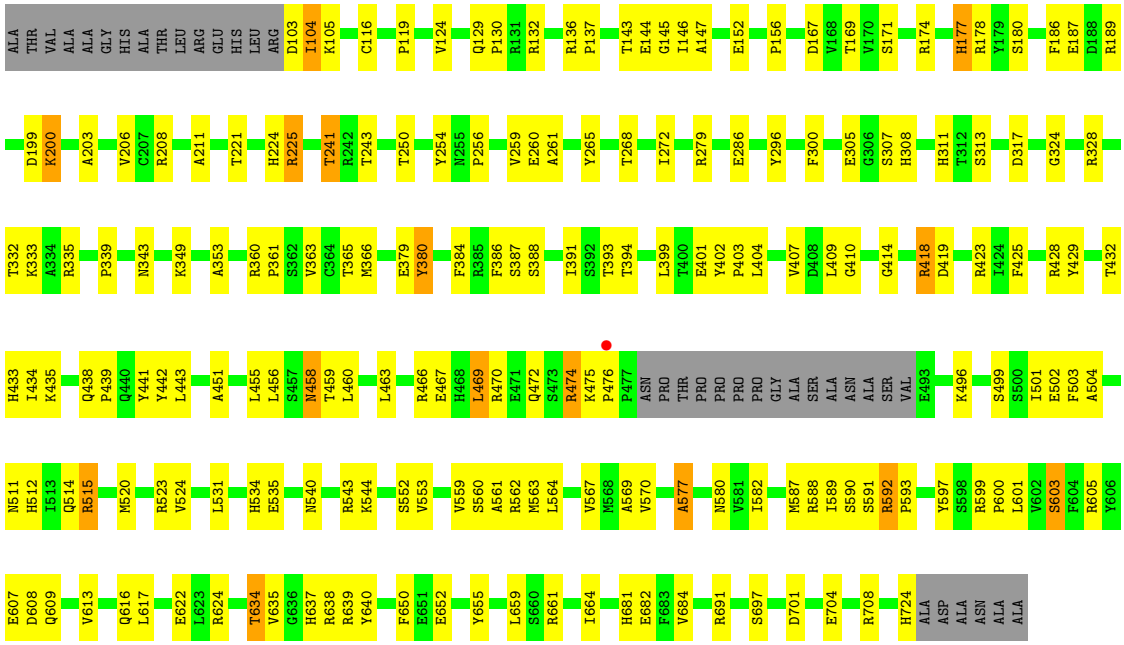
• Molecule 1: Envelope glycoprotein B

Chain C:



• Molecule 1: Envelope glycoprotein B

Chain D:


ALA THR VAL ALA ALA GLY HIS ALA THR LEU ARG GLU HIS LEU ARG D103 I104 K105 C116 P119 V124 Q129 P130 R131 R132 R136 P137 T143 E144 G145 I146 A147 E152 P156 D167 V168 T169 V170 S171 R174 H177 R178 Y179 S180 F186 E187 D188 R189
D199 K200 A203 V206 C207 R208 A211 T221 H224 R225 T241 R242 T243 T250 Y254 N255 P256 V259 E260 A261 Y265 T268 I272 R279 E286 Y296 F300 E305 G306 S307 H308 H311 T312 S313 D317 G324 R328
T332 K333 A334 R335 P339 N343 K349 A353 R360 P361 S362 V363 C364 T365 M366 E379 Y380 F384 R385 F386 S387 S388 I391 S392 T393 T394 L399 T400 E401 Y402 P403 L404 V407 D408 L409 G410 G414 R418 D419 R423 I424 F425 R428 Y429 T432
H433 I434 K435 Q438 P439 Q440 Y441 Y442 L443 A451 L455 L456 S457 N458 T459 L460 L463 R466 E467 H468 L469 R470 E471 Q472 S473 R474 K475 P476 P477 ASN PRO THR PRO PRO PRO PRO GLY ALA SER ALA ASN ALA SER VAL E493 K496 S499 S500 I501 E502 F503 A504
N511 H512 I513 Q514 R515 M520 R523 V524 L531 H534 E535 N540 R543 K544 S552 V553 V559 S560 A561 R562 M563 L564 V567 M568 A569 V570 A577 N580 V581 I582 M587 R588 I589 S590 S591 R592 P593 Y597 S598 R599 P600 L601 V602 S603 F604 R605 Y606
E607 D608 Q609 V613 Q616 L617 E622 L623 R624 T634 V635 G636 H637 R638 R639 Y640 F650 E651 E652 Y655 L659 S660 R661 I664 H681 E682 F683 V684 R691 S697 D701 E704 R708 H724 ALA ASP ALA ASN ALA ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	117.09Å 117.09Å 321.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.32 – 2.26 48.35 – 2.26	Depositor EDS
% Data completeness (in resolution range)	87.1 (47.32-2.26) 87.1 (48.35-2.26)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 2.27Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, R_{free}	0.172 , 0.227 0.171 , 0.227	Depositor DCC
R_{free} test set	11385 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	28.3	Xtriage
Anisotropy	0.493	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 44.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.020 for -h,-k,l 0.467 for h,-h-k,-l 0.021 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	21541	wwPDB-VP
Average B, all atoms (Å^2)	37.0	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 74.87 % 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 1.4096e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.38	0/5021	0.55	0/6822
1	B	0.37	0/5017	0.54	0/6814
1	C	0.37	0/5002	0.53	0/6798
1	D	0.38	0/5020	0.55	0/6820
All	All	0.37	0/20060	0.54	0/27254

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (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	4899	0	4713	133	0
1	B	4896	0	4717	183	0
1	C	4883	0	4691	134	0
1	D	4900	0	4708	149	0
2	A	14	0	13	0	0
2	B	14	0	13	1	0
2	C	14	0	13	1	0
2	D	14	0	13	1	0
3	A	8	0	10	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	16	0	20	0	0
3	C	8	0	10	0	0
3	D	8	0	10	0	0
4	A	478	0	0	13	0
4	B	444	0	0	12	0
4	C	462	0	0	19	0
4	D	483	0	0	14	0
All	All	21541	0	18931	595	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 15.

The worst 5 of 595 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:515:ARG:HG3	1:D:515:ARG:HH11	1.27	0.98
1:C:543:ARG:HB3	1:C:550:ILE:HG21	1.55	0.88
1:D:515:ARG:HG3	1:D:515:ARG:NH1	1.82	0.88
1:B:189:ARG:HB2	1:B:349:LYS:HE2	1.59	0.85
1:D:403:PRO:HG3	1:D:476:PRO:HB3	1.60	0.82

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	605/703 (86%)	558 (92%)	43 (7%)	4 (1%)	22 21
1	B	604/703 (86%)	555 (92%)	46 (8%)	3 (0%)	29 29
1	C	604/703 (86%)	543 (90%)	55 (9%)	6 (1%)	15 13
1	D	604/703 (86%)	571 (94%)	29 (5%)	4 (1%)	22 21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2417/2812 (86%)	2227 (92%)	173 (7%)	17 (1%)	22	21

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	259	VAL
1	A	260	GLU
1	D	177	HIS
1	C	118	PRO
1	C	655	TYR

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	526/593 (89%)	504 (96%)	22 (4%)	30	34
1	B	526/593 (89%)	501 (95%)	25 (5%)	25	28
1	C	522/593 (88%)	503 (96%)	19 (4%)	35	42
1	D	527/593 (89%)	502 (95%)	25 (5%)	26	29
All	All	2101/2372 (89%)	2010 (96%)	91 (4%)	29	33

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	502	GLU
1	D	305	GLU
1	C	553	VAL
1	C	724	HIS
1	D	418	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	308	HIS
1	D	511	ASN
1	D	706	GLN
1	D	681	HIS
1	A	585	ASN

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)

9 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	C	1141	1	14,14,15	0.57	0	17,19,21	1.12	1 (5%)
3	MRY	C	2000	-	7,7,7	0.20	0	8,8,8	0.95	0
3	MRY	B	2000	-	7,7,7	0.35	0	8,8,8	0.78	0
2	NAG	A	1398	1	14,14,15	0.52	0	17,19,21	1.41	1 (5%)
2	NAG	D	1398	1	14,14,15	0.55	0	17,19,21	0.96	1 (5%)
2	NAG	B	1398	1	14,14,15	0.46	0	17,19,21	1.00	1 (5%)
3	MRY	D	2000	-	7,7,7	0.48	0	8,8,8	0.81	0
3	MRY	A	2000	-	7,7,7	0.49	0	8,8,8	0.65	0
3	MRY	B	3000	-	7,7,7	0.91	0	8,8,8	1.38	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
2	NAG	C	1141	1	-	2/6/23/26	0/1/1/1
3	MRY	C	2000	-	-	0/8/8/8	-
3	MRY	B	2000	-	-	0/8/8/8	-
2	NAG	A	1398	1	-	0/6/23/26	0/1/1/1
2	NAG	D	1398	1	-	2/6/23/26	0/1/1/1
2	NAG	B	1398	1	-	2/6/23/26	0/1/1/1
3	MRY	D	2000	-	-	0/8/8/8	-
3	MRY	A	2000	-	-	0/8/8/8	-
3	MRY	B	3000	-	-	0/8/8/8	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1398	NAG	C1-O5-C5	4.71	118.58	112.19
2	B	1398	NAG	C1-O5-C5	3.13	116.44	112.19
2	C	1141	NAG	O5-C1-C2	2.56	115.34	111.29
2	D	1398	NAG	C1-O5-C5	2.05	114.97	112.19

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1141	NAG	O5-C5-C6-O6
2	B	1398	NAG	O5-C5-C6-O6
2	D	1398	NAG	C8-C7-N2-C2
2	B	1398	NAG	C4-C5-C6-O6
2	D	1398	NAG	O7-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1141	NAG	1	0
2	D	1398	NAG	1	0
2	B	1398	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	608/703 (86%)	-0.28	1 (0%)	95	96	21, 33, 55, 75	0
1	B	607/703 (86%)	-0.22	6 (0%)	82	84	21, 36, 63, 90	0
1	C	608/703 (86%)	-0.22	5 (0%)	86	87	22, 35, 64, 88	0
1	D	607/703 (86%)	-0.28	1 (0%)	95	96	23, 34, 55, 88	0
All	All	2430/2812 (86%)	-0.25	13 (0%)	91	91	21, 34, 60, 90	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	463	LEU	3.1
1	D	476	PRO	2.9
1	B	492	VAL	2.9
1	C	464	TYR	2.9
1	A	465	VAL	2.8

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)

There are no monosaccharides 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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	D	1398	14/15	0.85	0.13	40,58,64,64	0
2	NAG	A	1398	14/15	0.86	0.25	49,65,76,81	0
2	NAG	C	1141	14/15	0.93	0.23	84,89,92,95	0
2	NAG	B	1398	14/15	0.93	0.09	46,59,65,66	0
3	MRY	B	2000	8/8	0.95	0.13	29,35,39,44	0
3	MRY	B	3000	8/8	0.95	0.10	39,41,43,47	0
3	MRY	A	2000	8/8	0.97	0.14	27,33,36,37	0
3	MRY	C	2000	8/8	0.98	0.12	27,29,33,38	0
3	MRY	D	2000	8/8	0.98	0.12	27,27,32,32	0

6.5 Other polymers (i)

There are no such residues in this entry.