wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 10:10 PM EDT

PDB ID	:	3WLW
Title	:	Molecular Architecture of the ErbB2 Extracellular Domain Homodimer
Authors	:	Hu, S.; Lou, Z.Y.; Guo, Y.J.
Deposited on	:	2013-11-15
Resolution	:	3.09 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 (1)) were used in the production of this report:

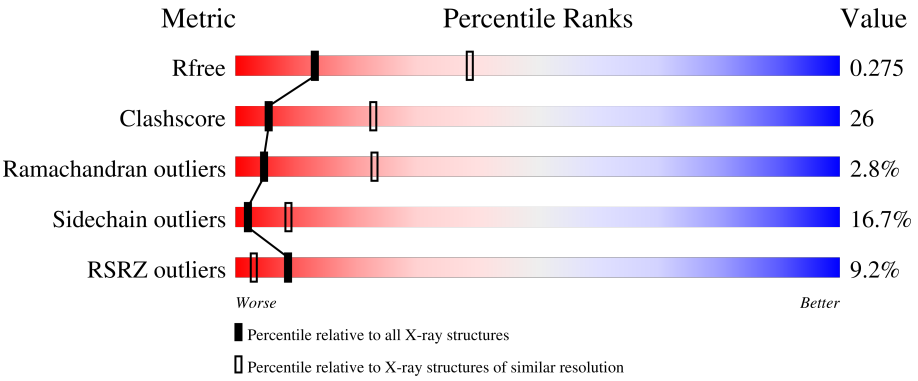
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 3.09 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.



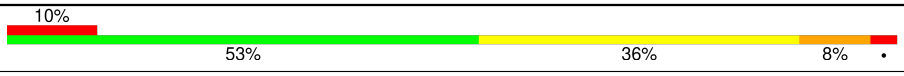
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1447 (3.10-3.06)
Clashscore	141614	1546 (3.10-3.06)
Ramachandran outliers	138981	1487 (3.10-3.06)
Sidechain outliers	138945	1486 (3.10-3.06)
RSRZ outliers	127900	1416 (3.10-3.06)

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 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	564	3% 60% 30% 7%
1	B	564	2% 55% 34% 9%
2	C	217	24% 43% 46% 10%
2	H	217	4% 54% 35% 9%
3	D	217	34% 43% 37% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	L	217	
4	E	2	

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
4	NAG	E	1	-	-	X	-
5	NAG	B	1001	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 15155 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 Receptor tyrosine-protein kinase erbB-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	555	Total	C	N	O	S	0	0	0
			4277	2656	769	805	47			
1	B	555	Total	C	N	O	S	0	0	0
			4277	2656	769	805	47			

- Molecule 2 is a protein called Antibody H Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	217	Total	C	N	O	S	0	0	0
			1624	1021	269	327	7			
2	C	217	Total	C	N	O	S	0	0	0
			1624	1021	269	327	7			

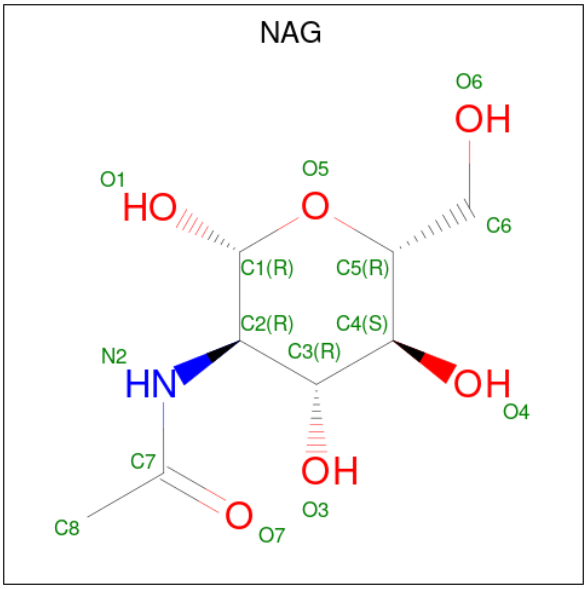
- Molecule 3 is a protein called Antibody L Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	217	Total	C	N	O	S	0	0	0
			1616	1001	269	340	6			
3	D	217	Total	C	N	O	S	0	0	0
			1616	1001	269	340	6			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 6 is water.

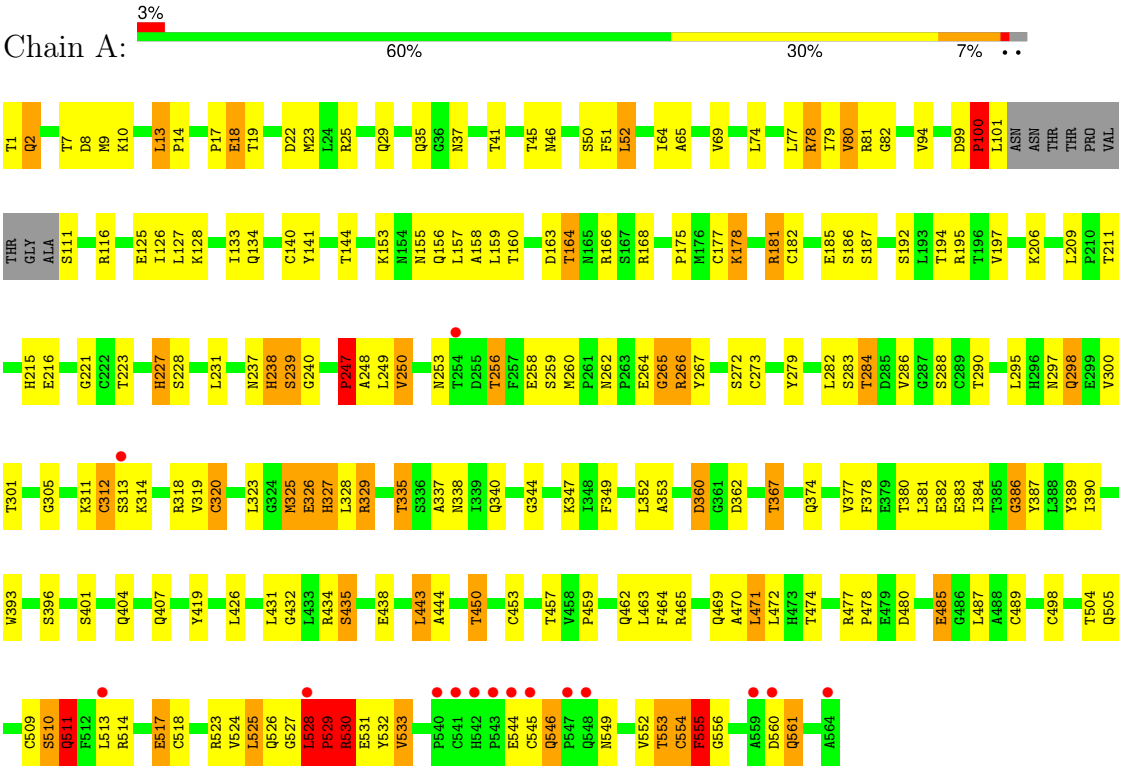
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	27	Total O 27 27	0	0
6	B	14	Total O 14 14	0	0
6	H	2	Total O 2 2	0	0
6	L	5	Total O 5 5	0	0
6	C	12	Total O 12 12	0	0
6	D	19	Total O 19 19	0	0

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.

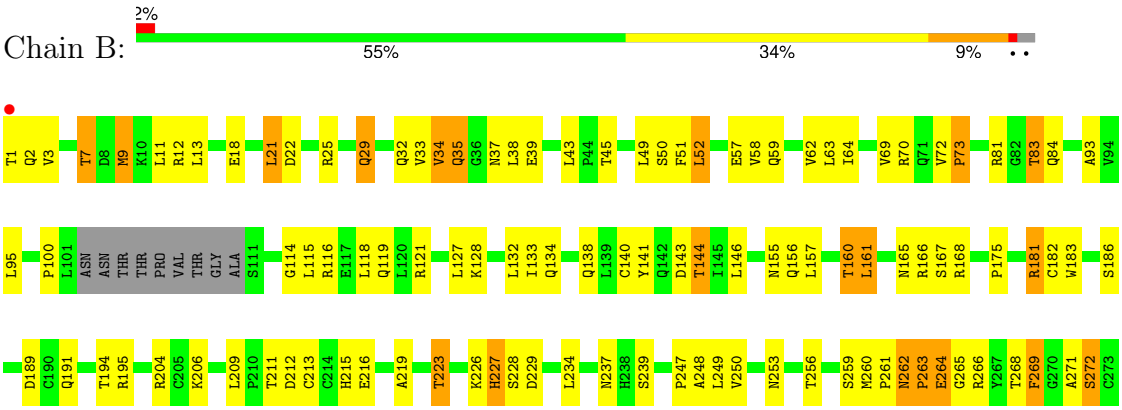
• Molecule 1: Receptor tyrosine-protein kinase erbB-2

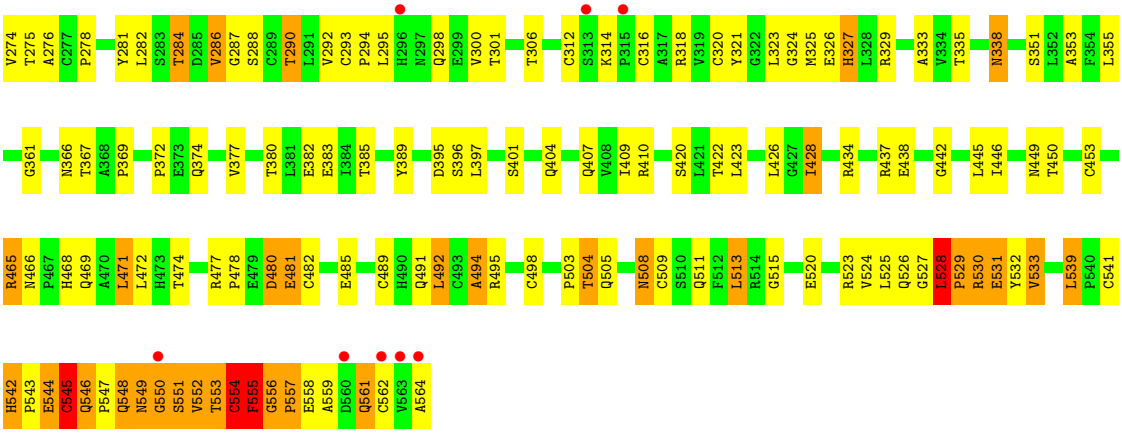
Chain A:



• Molecule 1: Receptor tyrosine-protein kinase erbB-2

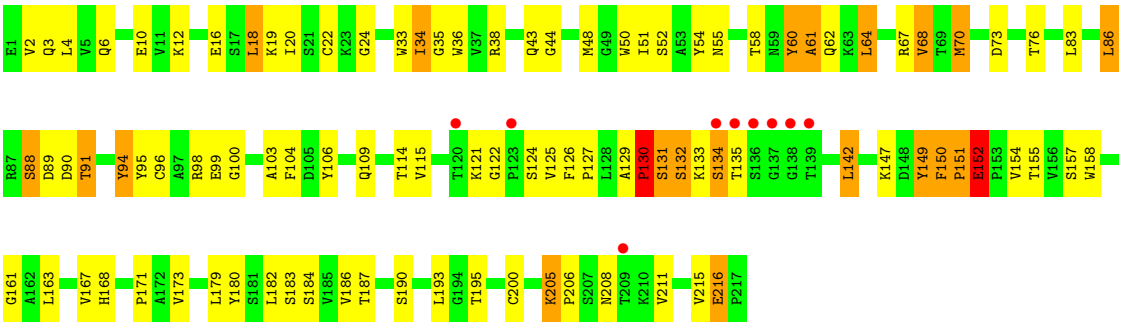
Chain B:





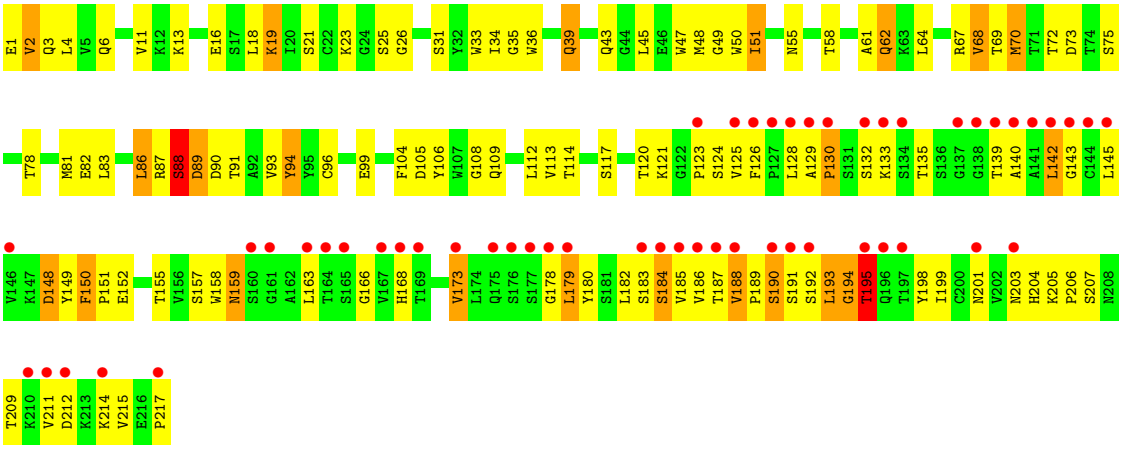
• Molecule 2: Antibody H Chain

Chain H:



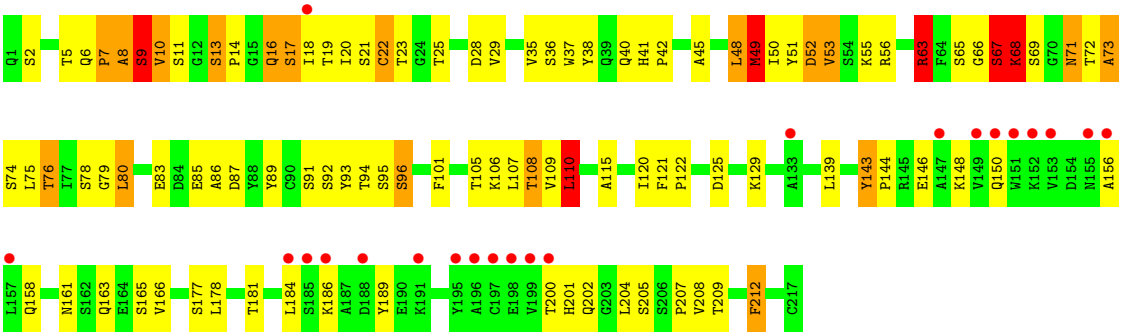
• Molecule 2: Antibody H Chain

Chain C:



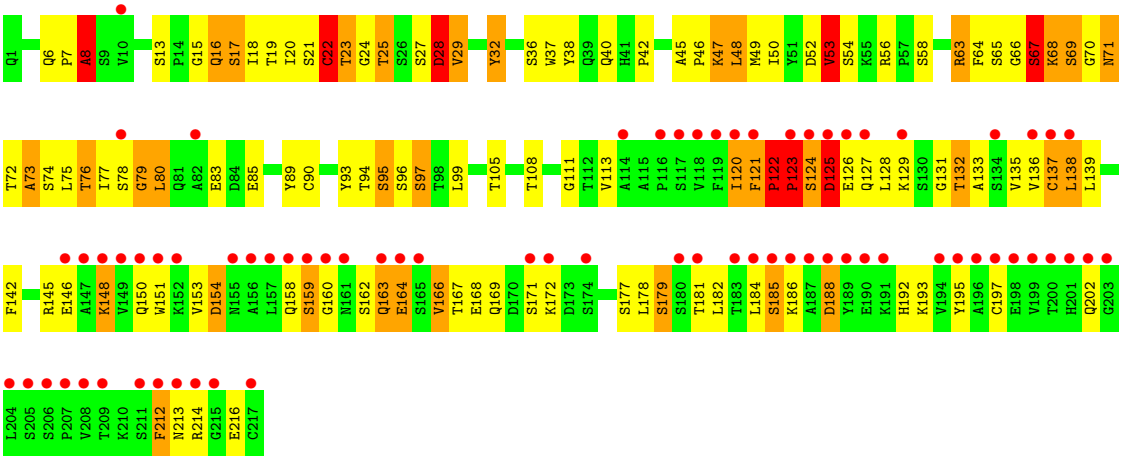
• Molecule 3: Antibody L Chain

Chain L:



• Molecule 3: Antibody L Chain

Chain D:



• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	84.72Å 104.19Å 116.71Å 106.89° 99.65° 111.12°	Depositor
Resolution (Å)	32.26 – 3.09 48.29 – 3.09	Depositor EDS
% Data completeness (in resolution range)	97.2 (32.26-3.09) 92.3 (48.29-3.09)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.14 (at 3.07Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.232 , 0.269 0.242 , 0.275	Depositor DCC
R_{free} test set	3059 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	62.5	Xtriage
Anisotropy	0.581	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 68.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	15155	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% 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: 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	0.58	3/4374 (0.1%)	0.91	24/5950 (0.4%)
1	B	0.59	2/4374 (0.0%)	0.88	13/5950 (0.2%)
2	C	0.46	0/1664	0.78	5/2268 (0.2%)
2	H	0.50	1/1664 (0.1%)	0.91	10/2268 (0.4%)
3	D	0.60	0/1649	1.08	18/2241 (0.8%)
3	L	0.55	0/1649	1.02	12/2241 (0.5%)
All	All	0.56	6/15374 (0.0%)	0.92	82/20918 (0.4%)

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	A	0	3
1	B	0	1
2	C	0	2
2	H	0	1
3	D	0	6
3	L	0	4
All	All	0	17

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	539	LEU	C-N	10.37	1.53	1.34
1	A	265	GLY	C-N	8.59	1.53	1.34
1	A	247	PRO	C-N	6.51	1.49	1.34
1	B	557	PRO	N-CD	5.63	1.55	1.47
1	A	529	PRO	N-CD	5.28	1.55	1.47

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	326	GLU	CB-CA-C	-11.98	86.43	110.40
1	B	269	PHE	CB-CA-C	11.84	134.09	110.40
2	H	62	GLN	N-CA-CB	-10.85	91.07	110.60
1	A	510	SER	CB-CA-C	-10.46	90.22	110.10
3	L	48	LEU	N-CA-CB	-10.30	89.81	110.40

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	111	SER	Peptide
1	A	247	PRO	Mainchain
1	A	77	LEU	Peptide
1	B	529	PRO	Peptide
2	H	152	GLU	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	4277	0	4106	142	0
1	B	4277	0	4108	216	3
2	C	1624	0	1567	110	2
2	H	1624	0	1567	81	1
3	D	1616	0	1548	148	0
3	L	1616	0	1548	89	0
4	E	28	0	25	8	0
5	B	14	0	13	10	0
6	A	27	0	0	4	0
6	B	14	0	0	1	0
6	C	12	0	0	7	0
6	D	19	0	0	4	0
6	H	2	0	0	0	0
6	L	5	0	0	1	0
All	All	15155	0	14482	766	3

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 26.

The worst 5 of 766 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:B:546:GLN:CG	1:B:548:GLN:HG3	1.37	1.53
3:D:123:PRO:HB2	3:D:127:GLN:CD	1.29	1.48
1:A:237:ASN:HD21	4:E:1:NAG:C1	1.21	1.47
3:D:123:PRO:CB	3:D:127:GLN:NE2	1.79	1.44
1:B:250:VAL:CB	1:B:260:MET:O	1.69	1.39

All (3) 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:B:160:THR:O	2:C:191:SER:O[1_654]	1.78	0.42
1:B:160:THR:C	2:C:191:SER:O[1_654]	2.02	0.18
1:B:532:TYR:OH	2:H:195:THR:O[1_566]	2.11	0.09

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	551/564 (98%)	453 (82%)	86 (16%)	12 (2%)	6	27
1	B	551/564 (98%)	445 (81%)	90 (16%)	16 (3%)	4	22
2	C	215/217 (99%)	176 (82%)	35 (16%)	4 (2%)	8	31
2	H	215/217 (99%)	192 (89%)	18 (8%)	5 (2%)	6	26
3	D	215/217 (99%)	164 (76%)	41 (19%)	10 (5%)	2	13
3	L	215/217 (99%)	165 (77%)	43 (20%)	7 (3%)	4	19
All	All	1962/1996 (98%)	1595 (81%)	313 (16%)	54 (3%)	5	23

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	528	LEU
1	A	530	ARG
1	A	555	PHE
1	B	262	ASN
1	B	528	LEU

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	478/485 (99%)	401 (84%)	77 (16%)	2 10
1	B	478/485 (99%)	404 (84%)	74 (16%)	2 11
2	C	180/181 (99%)	149 (83%)	31 (17%)	2 8
2	H	180/181 (99%)	155 (86%)	25 (14%)	3 14
3	D	185/186 (100%)	144 (78%)	41 (22%)	1 3
3	L	185/186 (100%)	152 (82%)	33 (18%)	2 7
All	All	1686/1704 (99%)	1405 (83%)	281 (17%)	2 9

5 of 281 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	159	ASN
2	C	212	ASP
3	D	123	PRO
1	B	213	CYS
1	B	181	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	298	GLN
3	D	155	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	546	GLN
3	D	169	GLN
2	C	203	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)

2 monosaccharides 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	$\#\|Z\| > 2$
4	NAG	E	1	4,1	14,14,15	0.38	0	17,19,21	0.72	1 (5%)
4	NAG	E	2	4	14,14,15	0.48	0	17,19,21	1.84	3 (17%)

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
4	NAG	E	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	E	2	4	-	5/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	2	NAG	C1-O5-C5	6.15	120.42	112.19
4	E	2	NAG	C4-C3-C2	3.06	115.50	111.02
4	E	2	NAG	O5-C1-C2	2.04	114.44	111.29
4	E	1	NAG	C3-C4-C5	2.02	113.89	110.23

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	2	NAG	C3-C2-N2-C7
4	E	2	NAG	C8-C7-N2-C2
4	E	2	NAG	O7-C7-N2-C2
4	E	2	NAG	O5-C5-C6-O6
4	E	1	NAG	O5-C5-C6-O6

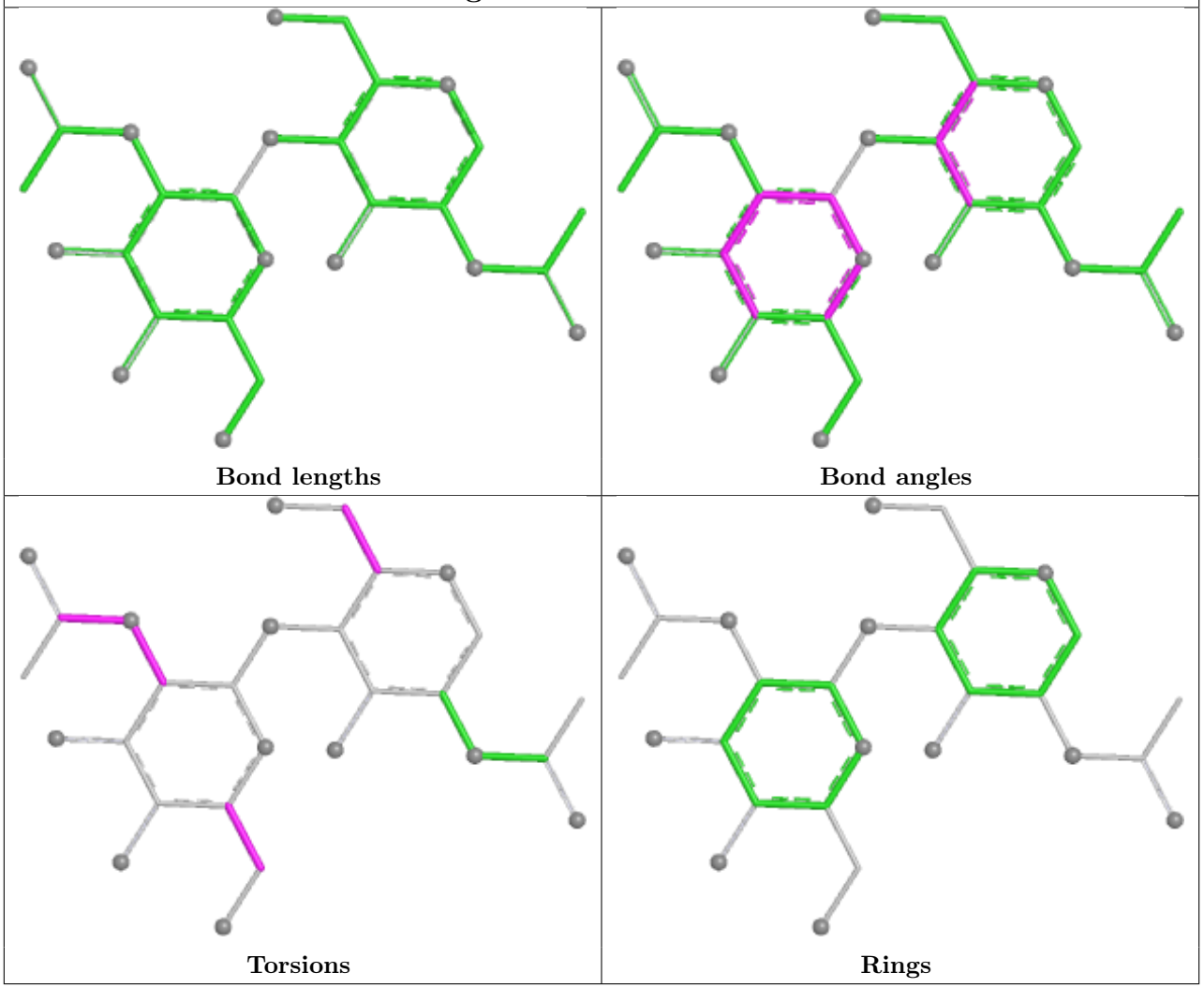
There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	2	NAG	2	0
4	E	1	NAG	7	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain E



5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
5	NAG	B	1001	1	14,14,15	0.37	0	17,19,21	0.72	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
5	NAG	B	1001	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	1001	NAG	C8-C7-N2-C2
5	B	1001	NAG	O7-C7-N2-C2
5	B	1001	NAG	C4-C5-C6-O6
5	B	1001	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1001	NAG	10	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

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	555/564 (98%)	-0.23	15 (2%) 54 29	28, 52, 125, 178	0
1	B	555/564 (98%)	-0.26	9 (1%) 72 51	31, 54, 122, 152	0
2	C	217/217 (100%)	1.15	53 (24%) 0 0	39, 86, 170, 180	0
2	H	217/217 (100%)	0.05	9 (4%) 37 18	38, 70, 129, 178	0
3	D	217/217 (100%)	1.60	73 (33%) 0 0	45, 114, 187, 203	0
3	L	217/217 (100%)	0.34	22 (10%) 7 2	36, 102, 163, 185	0
All	All	1978/1996 (99%)	0.21	181 (9%) 9 3	28, 63, 166, 203	0

The worst 5 of 181 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	147	ALA	16.8
2	C	138	GLY	16.4
3	D	120	ILE	13.3
3	D	189	TYR	12.9
2	C	142	LEU	12.2

6.2 Non-standard residues in protein, DNA, RNA chains

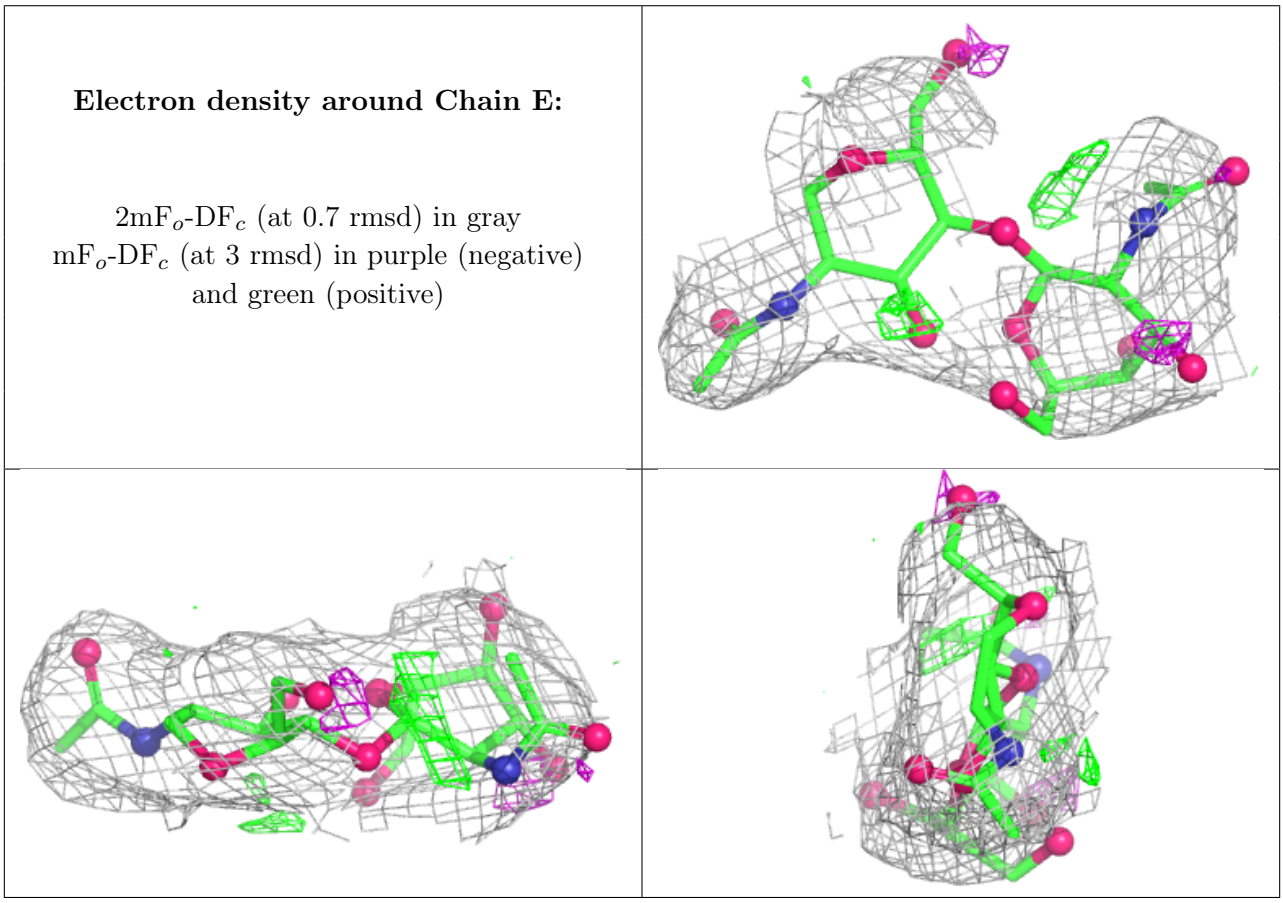
There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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
4	NAG	E	2	14/15	0.81	0.25	80,88,108,108	0
4	NAG	E	1	14/15	0.91	0.18	47,59,69,73	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



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
5	NAG	B	1001	14/15	0.56	0.27	57,60,76,82	0

6.5 Other polymers (i)

There are no such residues in this entry.