Full wwPDB X-ray Structure Validation Report ⓘ

Jan 8, 2024 – 10:33 am GMT

PDB ID	:	5OJK
Title	:	Crystal structure of the human neuroligin 1 cholinesterase domain containing spliced sequence B (SSB) (NL1(-A+B))
Authors	:	Elegheert, J.; Aricescu, A.R.
Deposited on	:	2017-07-21
Resolution	:	2.55 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 ⓘ 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 ⓘ) were used in the production of this report:

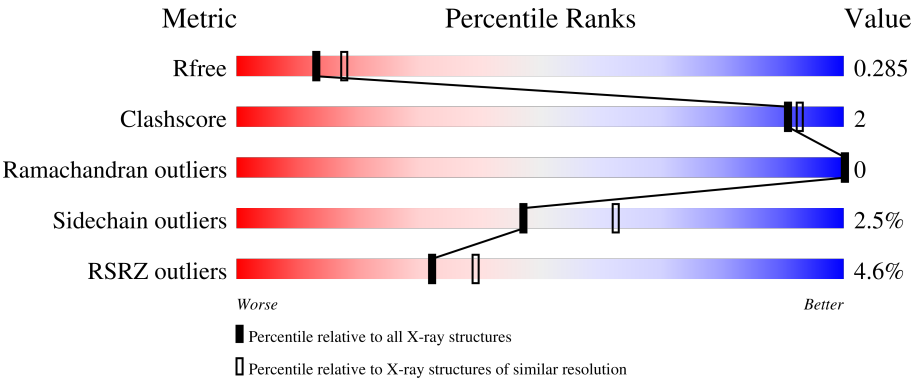
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	585	5% 86% 5% 8%
1	B	585	4% 85% 6% • 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8583 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.

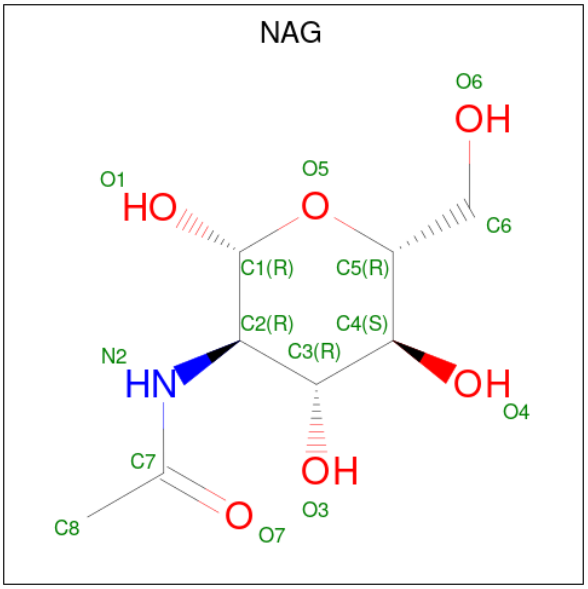
- Molecule 1 is a protein called Neuroligin-1,Neuroligin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	537	Total	C	N	O	S	0	1	0
			4199	2694	700	788	17			
1	B	537	Total	C	N	O	S	0	2	0
			4207	2704	700	786	17			

There are 24 discrepancies between the modelled and reference sequences:

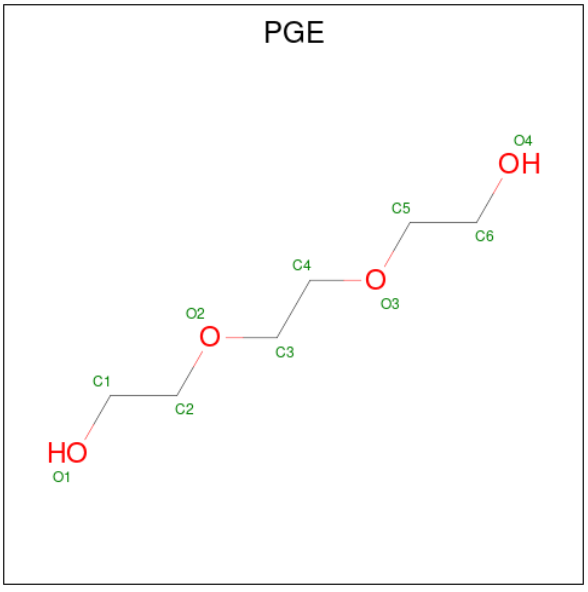
Chain	Residue	Modelled	Actual	Comment	Reference
A	43	GLU	-	expression tag	UNP Q8N2Q7
A	44	THR	-	expression tag	UNP Q8N2Q7
A	45	GLY	-	expression tag	UNP Q8N2Q7
A	636	ARG	-	expression tag	UNP Q8N2Q7
A	637	THR	-	expression tag	UNP Q8N2Q7
A	638	LYS	-	expression tag	UNP Q8N2Q7
A	639	HIS	-	expression tag	UNP Q8N2Q7
A	640	HIS	-	expression tag	UNP Q8N2Q7
A	641	HIS	-	expression tag	UNP Q8N2Q7
A	642	HIS	-	expression tag	UNP Q8N2Q7
A	643	HIS	-	expression tag	UNP Q8N2Q7
A	644	HIS	-	expression tag	UNP Q8N2Q7
B	43	GLU	-	expression tag	UNP Q8N2Q7
B	44	THR	-	expression tag	UNP Q8N2Q7
B	45	GLY	-	expression tag	UNP Q8N2Q7
B	636	ARG	-	expression tag	UNP Q8N2Q7
B	637	THR	-	expression tag	UNP Q8N2Q7
B	638	LYS	-	expression tag	UNP Q8N2Q7
B	639	HIS	-	expression tag	UNP Q8N2Q7
B	640	HIS	-	expression tag	UNP Q8N2Q7
B	641	HIS	-	expression tag	UNP Q8N2Q7
B	642	HIS	-	expression tag	UNP Q8N2Q7
B	643	HIS	-	expression tag	UNP Q8N2Q7
B	644	HIS	-	expression tag	UNP Q8N2Q7

- 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
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 6 4	0	0
3	A	1	Total C O 10 6 4	0	0
3	B	1	Total C O 10 6 4	0	0
3	B	1	Total C O 10 6 4	0	0

- Molecule 4 is water.

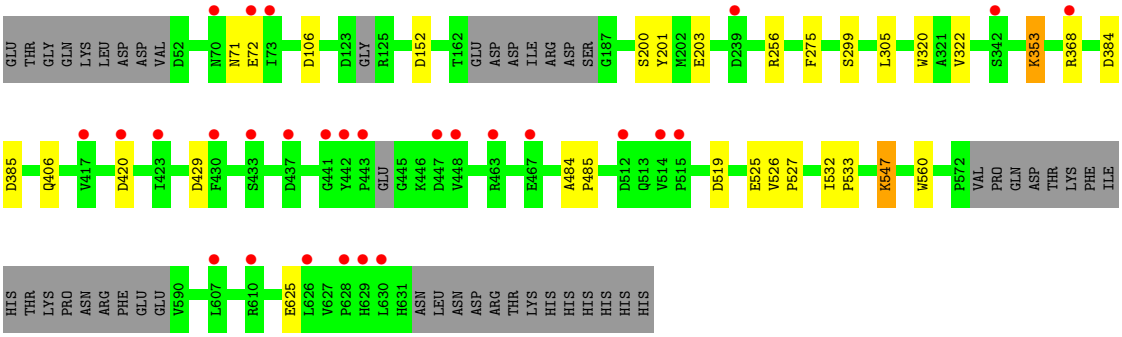
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	55	Total O 55 55	0	0
4	B	54	Total O 54 54	0	0

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.

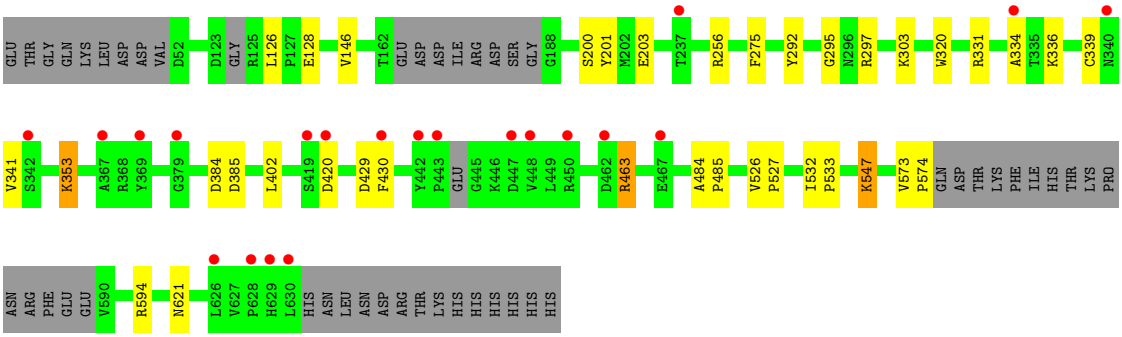
• Molecule 1: Neuroligin-1,Neuroligin-1

Chain A:



• Molecule 1: Neuroligin-1,Neuroligin-1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.57Å 118.05Å 214.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.70 – 2.55 51.70 – 2.55	Depositor EDS
% Data completeness (in resolution range)	96.7 (51.70-2.55) 88.0 (51.70-2.55)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.55Å)	Xtriage
Refinement program	PHENIX (dev_2044)	Depositor
R, R_{free}	0.236 , 0.278 0.244 , 0.285	Depositor DCC
R_{free} test set	2008 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	40.8	Xtriage
Anisotropy	0.444	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.4	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	8583	wwPDB-VP
Average B, all atoms (Å^2)	50.0	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 53.04 % 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 4.4623e-05. 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PGE, 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	0.23	0/4316	0.38	0/5893
1	B	0.23	0/4331	0.40	0/5913
All	All	0.23	0/8647	0.39	0/11806

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

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	4199	0	4017	14	1
1	B	4207	0	4053	16	1
2	A	14	0	13	0	0
2	B	14	0	13	0	0
3	A	20	0	28	1	0
3	B	20	0	28	1	0
4	A	55	0	0	2	0
4	B	54	0	0	1	0
All	All	8583	0	8152	29	1

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 2.

All (29) 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:331:ARG:NH1	1:B:341:VAL:O	2.19	0.74
1:A:625:GLU:OE2	1:B:463:ARG:NH2	2.29	0.65
1:A:71:ASN:O	4:A:801:HOH:O	2.15	0.64
1:A:353:LYS:NZ	3:A:702:PGE:O2	2.33	0.57
1:A:547:LYS:H	1:A:547:LYS:HD2	1.71	0.55
1:A:406:GLN:NE2	1:A:519:ASP:OD2	2.40	0.54
1:A:368:ARG:N	4:A:802:HOH:O	2.42	0.51
1:B:621:ASN:ND2	4:B:804:HOH:O	2.45	0.50
1:A:384:ASP:OD1	1:A:385:ASP:N	2.44	0.47
1:B:353:LYS:HE2	3:B:702:PGE:H2	1.97	0.46
1:B:334:ALA:O	1:B:339:CYS:N	2.38	0.46
1:A:200:SER:O	1:A:201:TYR:HB2	2.16	0.45
1:B:256:ARG:NH2	1:B:292:TYR:O	2.49	0.45
1:B:200:SER:O	1:B:201:TYR:HB2	2.17	0.45
1:B:384:ASP:OD1	1:B:385:ASP:N	2.48	0.44
1:A:525:GLU:OE1	1:A:525:GLU:N	2.46	0.44
1:A:532:ILE:N	1:A:533:PRO:CD	2.81	0.43
1:B:146:VAL:HG12	1:B:146:VAL:O	2.18	0.43
1:A:484:ALA:HB3	1:A:485:PRO:HD3	2.00	0.43
1:B:126:LEU:O	1:B:128:GLU:N	2.53	0.42
1:A:71:ASN:OD1	1:A:72:GLU:N	2.52	0.42
1:B:295:GLY:N	1:B:303:LYS:O	2.32	0.42
1:B:532:ILE:N	1:B:533:PRO:CD	2.82	0.42
1:B:484:ALA:HB3	1:B:485:PRO:HD3	2.02	0.42
1:B:526:VAL:N	1:B:527:PRO:HD2	2.35	0.42
1:A:256[B]:ARG:HG2	1:A:305:LEU:HD11	2.02	0.42
1:A:526:VAL:N	1:A:527:PRO:HD2	2.35	0.41
1:B:573:VAL:HG23	1:B:574:PRO:HA	2.03	0.41
1:B:547:LYS:H	1:B:547:LYS:HG2	1.72	0.40

All (1) 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:A:152:ASP:OD2	1:B:297:ARG:NH2[4_655]	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	528/585 (90%)	500 (95%)	28 (5%)	0	100	100
1	B	529/585 (90%)	502 (95%)	27 (5%)	0	100	100
All	All	1057/1170 (90%)	1002 (95%)	55 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	451/503 (90%)	440 (98%)	11 (2%)	49	64
1	B	455/503 (90%)	443 (97%)	12 (3%)	46	61
All	All	906/1006 (90%)	883 (98%)	23 (2%)	47	62

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	106	ASP
1	A	203	GLU
1	A	275	PHE
1	A	299	SER
1	A	320	TRP
1	A	322	VAL
1	A	353	LYS
1	A	420	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	429	ASP
1	A	547	LYS
1	A	560	TRP
1	B	203	GLU
1	B	275	PHE
1	B	320	TRP
1	B	336	LYS
1	B	353	LYS
1	B	402	LEU
1	B	420	ASP
1	B	429	ASP
1	B	430	PHE
1	B	463	ARG
1	B	547	LYS
1	B	594	ARG

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)

6 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	PGE	A	703	-	9,9,9	0.34	0	8,8,8	0.21	0
2	NAG	A	701	1	14,14,15	0.20	0	17,19,21	0.51	0
3	PGE	B	702	-	9,9,9	0.34	0	8,8,8	0.46	0
2	NAG	B	701	1	14,14,15	0.26	0	17,19,21	0.33	0
3	PGE	B	703	-	9,9,9	0.33	0	8,8,8	0.31	0
3	PGE	A	702	-	9,9,9	0.34	0	8,8,8	0.21	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	PGE	A	703	-	-	2/7/7/7	-
2	NAG	A	701	1	-	2/6/23/26	0/1/1/1
3	PGE	B	702	-	-	4/7/7/7	-
2	NAG	B	701	1	-	0/6/23/26	0/1/1/1
3	PGE	B	703	-	-	2/7/7/7	-
3	PGE	A	702	-	-	2/7/7/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	NAG	O5-C5-C6-O6
2	A	701	NAG	C4-C5-C6-O6
3	B	702	PGE	O2-C3-C4-O3
3	A	703	PGE	O1-C1-C2-O2
3	B	703	PGE	O3-C5-C6-O4
3	A	703	PGE	O2-C3-C4-O3
3	A	702	PGE	O2-C3-C4-O3
3	A	702	PGE	C6-C5-O3-C4
3	B	702	PGE	C1-C2-O2-C3
3	B	703	PGE	C6-C5-O3-C4
3	B	702	PGE	C4-C3-O2-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	B	702	PGE	C3-C4-O3-C5

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	702	PGE	1	0
3	A	702	PGE	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

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	537/585 (91%)	0.42	28 (5%)	27	32	30, 50, 75, 112	0
1	B	537/585 (91%)	0.36	21 (3%)	39	45	29, 47, 72, 107	0
All	All	1074/1170 (91%)	0.39	49 (4%)	32	39	29, 49, 74, 112	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	442	TYR	8.1
1	B	442	TYR	7.3
1	B	443	PRO	5.8
1	B	626	LEU	5.2
1	A	441	GLY	5.0
1	B	419	SER	4.8
1	A	630	LEU	4.4
1	A	443	PRO	4.3
1	B	447	ASP	4.1
1	A	626	LEU	4.1
1	A	70	ASN	4.1
1	A	629	HIS	3.9
1	B	342	SER	3.7
1	A	628	PRO	3.5
1	A	430	PHE	3.5
1	A	73	ILE	3.3
1	A	342	SER	3.2
1	B	450	ARG	3.2
1	A	512	ASP	3.2
1	A	515	PRO	3.1
1	A	420	ASP	3.1
1	A	447	ASP	3.0
1	A	607	LEU	2.9
1	B	628	PRO	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	423	ILE	2.8
1	B	367	ALA	2.8
1	B	430	PHE	2.7
1	A	368	ARG	2.7
1	B	237	THR	2.7
1	A	448	VAL	2.5
1	A	463	ARG	2.5
1	B	340	ASN	2.5
1	A	437	ASP	2.5
1	B	448	VAL	2.5
1	A	72	GLU	2.4
1	B	629	HIS	2.3
1	B	369	TYR	2.3
1	A	433	SER	2.3
1	B	630	LEU	2.3
1	A	467	GLU	2.2
1	B	462	ASP	2.2
1	A	610	ARG	2.2
1	A	514	VAL	2.2
1	B	379	GLY	2.1
1	B	420	ASP	2.1
1	A	417	VAL	2.1
1	B	334	ALA	2.0
1	A	239	ASP	2.0
1	B	467	GLU	2.0

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, 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
2	NAG	A	701	14/15	0.76	0.17	62,73,83,85	0
3	PGE	A	703	10/10	0.77	0.21	46,59,74,80	0
3	PGE	A	702	10/10	0.87	0.26	44,48,65,65	0
3	PGE	B	703	10/10	0.88	0.15	44,51,56,58	0
3	PGE	B	702	10/10	0.92	0.18	42,54,58,63	0
2	NAG	B	701	14/15	0.92	0.20	59,70,77,82	0

6.5 Other polymers (i)

There are no such residues in this entry.