wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2020 – 03:23 PM BST

PDB ID	:	3V64
Title	:	Crystal Structure of agrin and LRP4
Authors	:	Zong, Y.; Zhang, B.; Gu, S.; Lee, K.; Zhou, J.; Yao, G.; Figueiedo, D.; Perry, K.; Mei, L.; Jin, R.
Deposited on	:	2011-12-18
Resolution	:	2.85 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 following versions of software and data (see references (1)) were used in the production of this report:

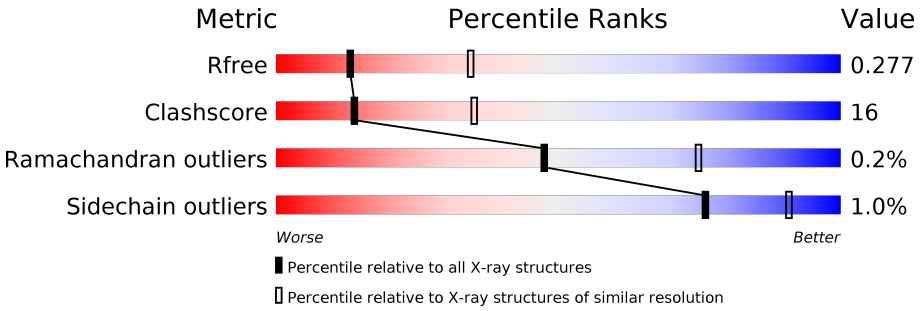
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)

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 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%

Mol	Chain	Length	Quality of chain
1	C	349	67% 30%
1	D	349	69% 26%
2	A	191	80% 19%
2	B	191	63% 36%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8471 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 Low-density lipoprotein receptor-related protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	338	Total	C	N	O	S	0	0	0
			2700	1697	498	490	15			
1	D	334	Total	C	N	O	S	0	0	0
			2673	1683	492	483	15			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	738	LEU	-	expression tag	UNP Q9QYP1
C	739	GLU	-	expression tag	UNP Q9QYP1
C	740	VAL	-	expression tag	UNP Q9QYP1
C	741	LEU	-	expression tag	UNP Q9QYP1
C	742	PHE	-	expression tag	UNP Q9QYP1
C	743	GLN	-	expression tag	UNP Q9QYP1
C	744	GLY	-	expression tag	UNP Q9QYP1
D	738	LEU	-	expression tag	UNP Q9QYP1
D	739	GLU	-	expression tag	UNP Q9QYP1
D	740	VAL	-	expression tag	UNP Q9QYP1
D	741	LEU	-	expression tag	UNP Q9QYP1
D	742	PHE	-	expression tag	UNP Q9QYP1
D	743	GLN	-	expression tag	UNP Q9QYP1
D	744	GLY	-	expression tag	UNP Q9QYP1

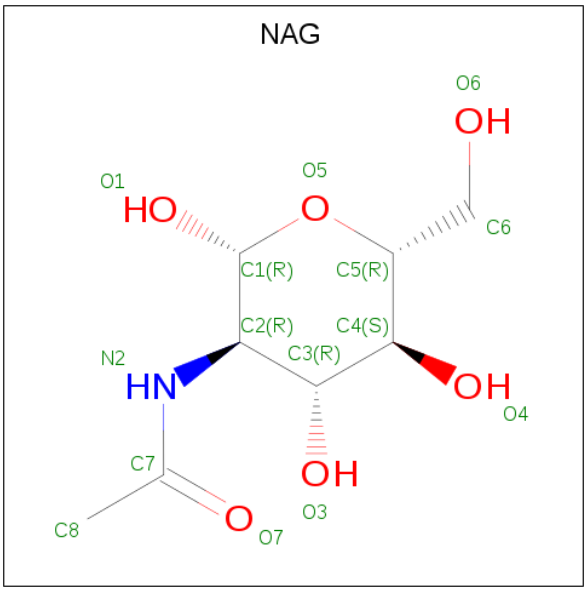
- Molecule 2 is a protein called agrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	191	Total	C	N	O	S	0	0	0
			1470	927	263	277	3			
2	B	190	Total	C	N	O	S	0	0	0
			1465	924	262	276	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1758	ALA	-	expression tag	UNP P25304
B	1758	ALA	-	expression tag	UNP P25304

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

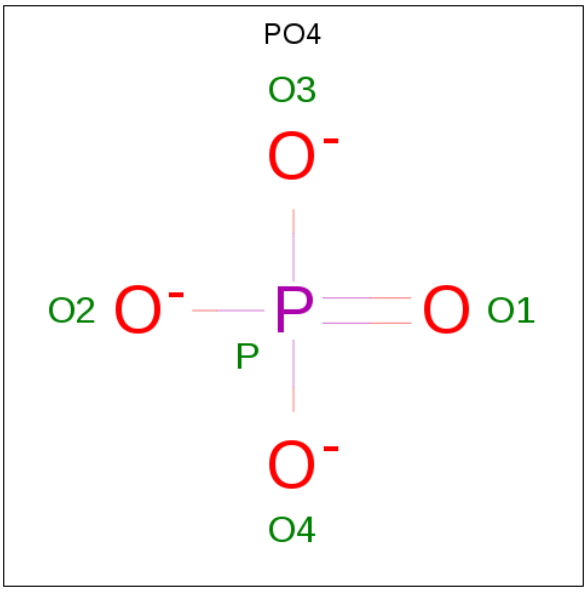


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	A	1	Total Ca 1 1	0	0
4	D	2	Total Ca 2 2	0	0
4	C	1	Total Ca 1 1	0	0

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total O P 5 4 1	0	0
5	C	1	Total O P 5 4 1	0	0
5	A	1	Total O P 5 4 1	0	0
5	B	1	Total O P 5 4 1	0	0

- Molecule 6 is water.

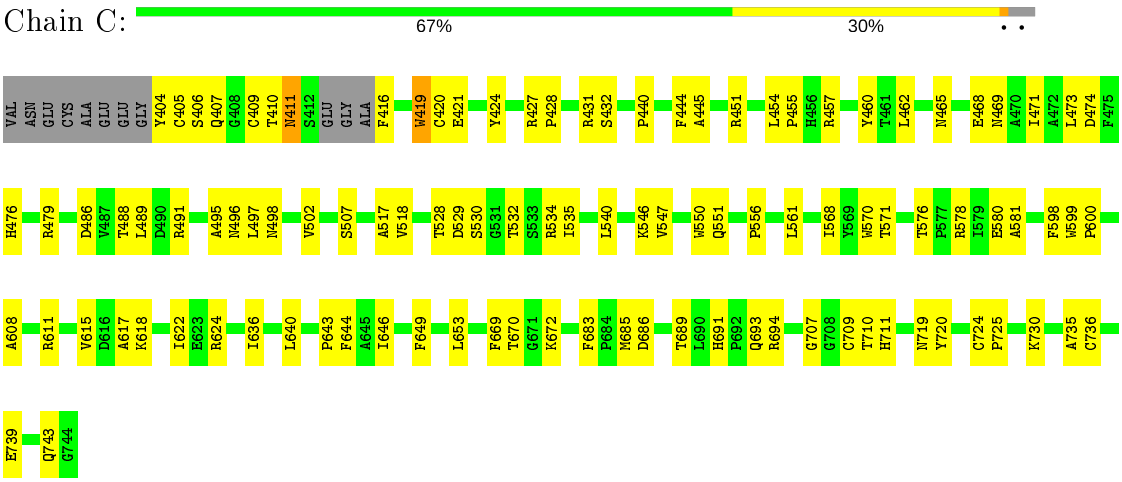
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	39	Total O 39 39	0	0
6	A	25	Total O 25 25	0	0
6	B	18	Total O 18 18	0	0
6	D	28	Total O 28 28	0	0

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. 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. 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.

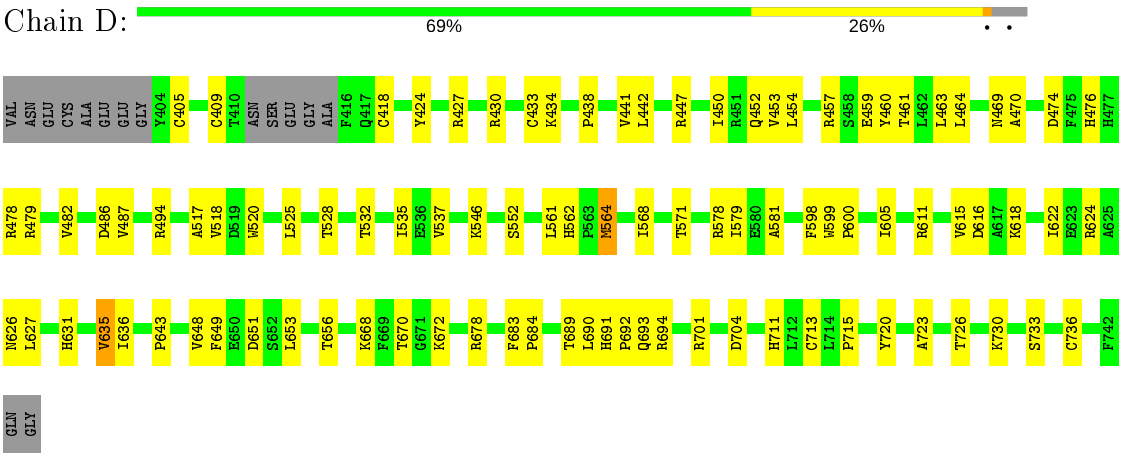
- Molecule 1: Low-density lipoprotein receptor-related protein 4

Chain C:



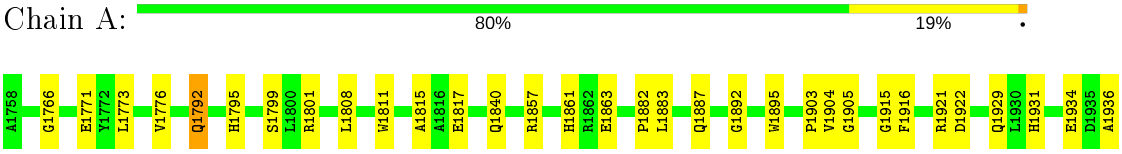
- Molecule 1: Low-density lipoprotein receptor-related protein 4

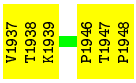
Chain D:



- Molecule 2: agrin

Chain A:





- Molecule 2: agrin

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.46Å 106.07Å 112.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.85 77.03 – 2.51	Depositor EDS
% Data completeness (in resolution range)	98.5 (50.00-2.85) 97.9 (77.03-2.51)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.203 , 0.273 0.235 , 0.277	Depositor DCC
R_{free} test set	2034 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	44.5	Xtriage
Anisotropy	0.382	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 61.0	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	0.91	EDS
Total number of atoms	8471	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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: PO4, CA, 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	C	0.50	0/2773	0.50	0/3768
1	D	0.49	0/2746	0.51	0/3732
2	A	0.40	0/1499	0.51	0/2041
2	B	0.41	0/1494	0.57	0/2034
All	All	0.46	0/8512	0.52	0/11575

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	C	2700	0	2595	87	0
1	D	2673	0	2574	70	0
2	A	1470	0	1474	34	0
2	B	1465	0	1469	73	0
3	C	28	0	26	8	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	5	0	0	1	0
5	B	5	0	0	0	0
5	C	10	0	0	0	0
6	A	25	0	0	0	0
6	B	18	0	0	2	0
6	C	39	0	0	0	0
6	D	28	0	0	0	0
All	All	8471	0	8138	260	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 16.

The worst 5 of 260 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)
2:A:1817:GLU:HG2	2:A:1904:VAL:HB	1.29	1.09
1:C:498:ASN:HB3	3:C:801:NAG:HN2	1.24	1.00
1:D:711:HIS:CD2	1:D:736:CYS:HB2	1.97	1.00
1:D:454:LEU:HD12	1:D:457:ARG:HH21	1.29	0.96
1:D:611:ARG:HD3	1:D:624:ARG:HD2	1.50	0.93

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	C	334/349 (96%)	308 (92%)	25 (8%)	1 (0%)	41 68
1	D	330/349 (95%)	300 (91%)	30 (9%)	0	100 100
2	A	189/191 (99%)	179 (95%)	10 (5%)	0	100 100
2	B	188/191 (98%)	179 (95%)	8 (4%)	1 (0%)	29 57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1041/1080 (96%)	966 (93%)	73 (7%)	2 (0%)	47 75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	411	ASN
2	B	1900	GLN

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	C	290/297 (98%)	285 (98%)	5 (2%)	60 83
1	D	287/297 (97%)	284 (99%)	3 (1%)	76 91
2	A	156/156 (100%)	155 (99%)	1 (1%)	86 95
2	B	156/156 (100%)	156 (100%)	0	100 100
All	All	889/906 (98%)	880 (99%)	9 (1%)	76 91

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	683	PHE
1	D	683	PHE
1	D	564	MET
1	C	540	LEU
2	A	1792	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1832	GLN
2	B	1861	HIS
1	D	688	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	1929	GLN
1	D	691	HIS

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)

Of 11 ligands modelled in this entry, 5 are monoatomic - leaving 6 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	$\#\|Z\| > 2$
5	PO4	A	2002	-	4,4,4	1.03	0	6,6,6	0.42	0
5	PO4	C	805	-	4,4,4	0.89	0	6,6,6	0.45	0
3	NAG	C	801	1	14,14,15	0.61	0	17,19,21	0.86	0
5	PO4	B	2002	-	4,4,4	0.92	0	6,6,6	0.41	0
3	NAG	C	802	-	14,14,15	0.56	0	17,19,21	0.85	0
5	PO4	C	804	-	4,4,4	0.90	0	6,6,6	0.39	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	802	-	-	2/6/23/26	0/1/1/1
3	NAG	C	801	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	802	NAG	O5-C5-C6-O6
3	C	802	NAG	C4-C5-C6-O6
3	C	801	NAG	C4-C5-C6-O6
3	C	801	NAG	C1-C2-N2-C7
3	C	801	NAG	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	2002	PO4	1	0
3	C	801	NAG	4	0
3	C	802	NAG	4	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.