



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:58 am BST

PDB ID : 4GZA
Title : Complex of mouse Plexin A2 - Semaphorin 3A - Neuropilin-1
Authors : Janssen, B.J.C.; Malinauskas, T.; Siebold, C.; Jones, E.Y.
Deposited on : 2012-09-06
Resolution : 7.00 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

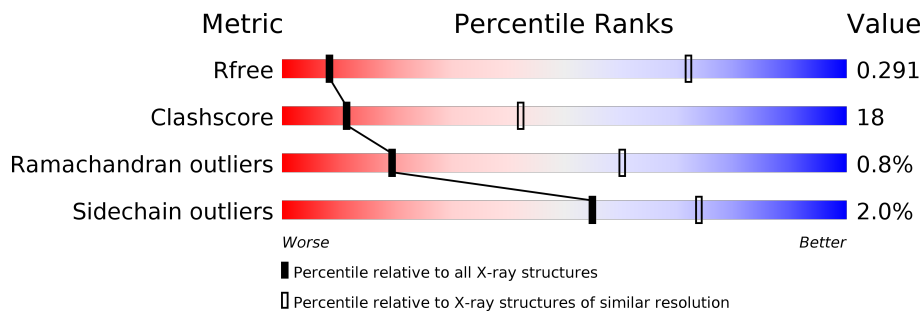
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.00 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	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	681	72% (green), 23% (yellow), 5% (orange), 0% (red), 0% (grey)
1	B	681	74% (green), 22% (yellow), 4% (orange), 0% (red), 0% (grey)
1	C	681	73% (green), 22% (yellow), 4% (orange), 0% (red), 0% (grey)
1	D	681	73% (green), 21% (yellow), 5% (orange), 0% (red), 0% (grey)
1	E	681	74% (green), 21% (yellow), 5% (orange), 0% (red), 0% (grey)
1	F	681	73% (green), 22% (yellow), 4% (orange), 0% (red), 0% (grey)
2	G	538	44% (green), 41% (yellow), 5% (orange), 10% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	577	 17% . 80%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 35655 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 Plexin-A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	656	5138	3255	876	972	35	0	0	0
1	B	656	5138	3255	876	972	35	0	0	0
1	C	656	5138	3255	876	972	35	0	0	0
1	D	656	5138	3255	876	972	35	0	0	0
1	E	656	5138	3255	876	972	35	0	0	0
1	F	656	5138	3255	876	972	35	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	GLU	-	EXPRESSION TAG	UNP P70207
A	704	GLY	-	EXPRESSION TAG	UNP P70207
A	705	THR	-	EXPRESSION TAG	UNP P70207
A	706	LYS	-	EXPRESSION TAG	UNP P70207
A	707	HIS	-	EXPRESSION TAG	UNP P70207
A	708	HIS	-	EXPRESSION TAG	UNP P70207
A	709	HIS	-	EXPRESSION TAG	UNP P70207
A	710	HIS	-	EXPRESSION TAG	UNP P70207
A	711	HIS	-	EXPRESSION TAG	UNP P70207
A	712	HIS	-	EXPRESSION TAG	UNP P70207
B	32	GLU	-	EXPRESSION TAG	UNP P70207
B	704	GLY	-	EXPRESSION TAG	UNP P70207
B	705	THR	-	EXPRESSION TAG	UNP P70207
B	706	LYS	-	EXPRESSION TAG	UNP P70207
B	707	HIS	-	EXPRESSION TAG	UNP P70207
B	708	HIS	-	EXPRESSION TAG	UNP P70207
B	709	HIS	-	EXPRESSION TAG	UNP P70207

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	710	HIS	-	EXPRESSION TAG	UNP P70207
B	711	HIS	-	EXPRESSION TAG	UNP P70207
B	712	HIS	-	EXPRESSION TAG	UNP P70207
C	32	GLU	-	EXPRESSION TAG	UNP P70207
C	704	GLY	-	EXPRESSION TAG	UNP P70207
C	705	THR	-	EXPRESSION TAG	UNP P70207
C	706	LYS	-	EXPRESSION TAG	UNP P70207
C	707	HIS	-	EXPRESSION TAG	UNP P70207
C	708	HIS	-	EXPRESSION TAG	UNP P70207
C	709	HIS	-	EXPRESSION TAG	UNP P70207
C	710	HIS	-	EXPRESSION TAG	UNP P70207
C	711	HIS	-	EXPRESSION TAG	UNP P70207
C	712	HIS	-	EXPRESSION TAG	UNP P70207
D	32	GLU	-	EXPRESSION TAG	UNP P70207
D	704	GLY	-	EXPRESSION TAG	UNP P70207
D	705	THR	-	EXPRESSION TAG	UNP P70207
D	706	LYS	-	EXPRESSION TAG	UNP P70207
D	707	HIS	-	EXPRESSION TAG	UNP P70207
D	708	HIS	-	EXPRESSION TAG	UNP P70207
D	709	HIS	-	EXPRESSION TAG	UNP P70207
D	710	HIS	-	EXPRESSION TAG	UNP P70207
D	711	HIS	-	EXPRESSION TAG	UNP P70207
D	712	HIS	-	EXPRESSION TAG	UNP P70207
E	32	GLU	-	EXPRESSION TAG	UNP P70207
E	704	GLY	-	EXPRESSION TAG	UNP P70207
E	705	THR	-	EXPRESSION TAG	UNP P70207
E	706	LYS	-	EXPRESSION TAG	UNP P70207
E	707	HIS	-	EXPRESSION TAG	UNP P70207
E	708	HIS	-	EXPRESSION TAG	UNP P70207
E	709	HIS	-	EXPRESSION TAG	UNP P70207
E	710	HIS	-	EXPRESSION TAG	UNP P70207
E	711	HIS	-	EXPRESSION TAG	UNP P70207
E	712	HIS	-	EXPRESSION TAG	UNP P70207
F	32	GLU	-	EXPRESSION TAG	UNP P70207
F	704	GLY	-	EXPRESSION TAG	UNP P70207
F	705	THR	-	EXPRESSION TAG	UNP P70207
F	706	LYS	-	EXPRESSION TAG	UNP P70207
F	707	HIS	-	EXPRESSION TAG	UNP P70207
F	708	HIS	-	EXPRESSION TAG	UNP P70207
F	709	HIS	-	EXPRESSION TAG	UNP P70207
F	710	HIS	-	EXPRESSION TAG	UNP P70207
F	711	HIS	-	EXPRESSION TAG	UNP P70207

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	712	HIS	-	EXPRESSION TAG	UNP P70207

- Molecule 2 is a protein called Semaphorin-3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	G	484	3882	2474	669	720	19	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	18	GLU	-	EXPRESSION TAG	UNP O08665
G	19	THR	-	EXPRESSION TAG	UNP O08665
G	20	GLY	-	EXPRESSION TAG	UNP O08665
G	475	VAL	ILE	SEE REMARK 999	UNP O08665

- Molecule 3 is a protein called Neuropilin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	117	944	611	152	176	5	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	19	GLU	-	EXPRESSION TAG	UNP P97333
H	20	THR	-	EXPRESSION TAG	UNP P97333
H	21	GLY	-	EXPRESSION TAG	UNP P97333
H	587	ARG	-	EXPRESSION TAG	UNP P97333
H	588	THR	-	EXPRESSION TAG	UNP P97333
H	589	LYS	-	EXPRESSION TAG	UNP P97333
H	590	HIS	-	EXPRESSION TAG	UNP P97333
H	591	HIS	-	EXPRESSION TAG	UNP P97333
H	592	HIS	-	EXPRESSION TAG	UNP P97333
H	593	HIS	-	EXPRESSION TAG	UNP P97333
H	594	HIS	-	EXPRESSION TAG	UNP P97333
H	595	HIS	-	EXPRESSION TAG	UNP P97333

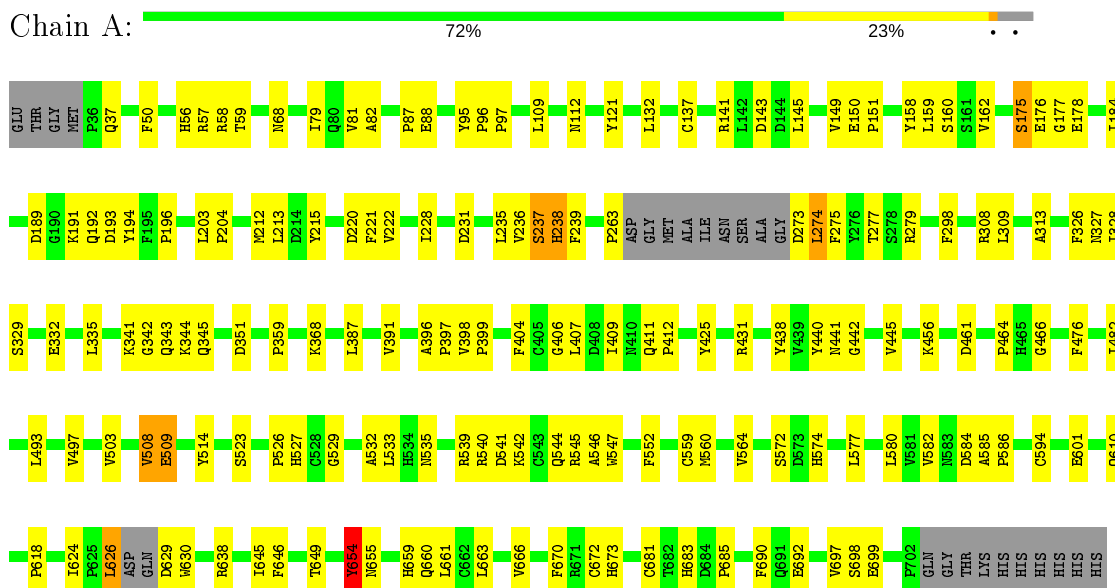
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	1	Total 1	Ca 1	0	0

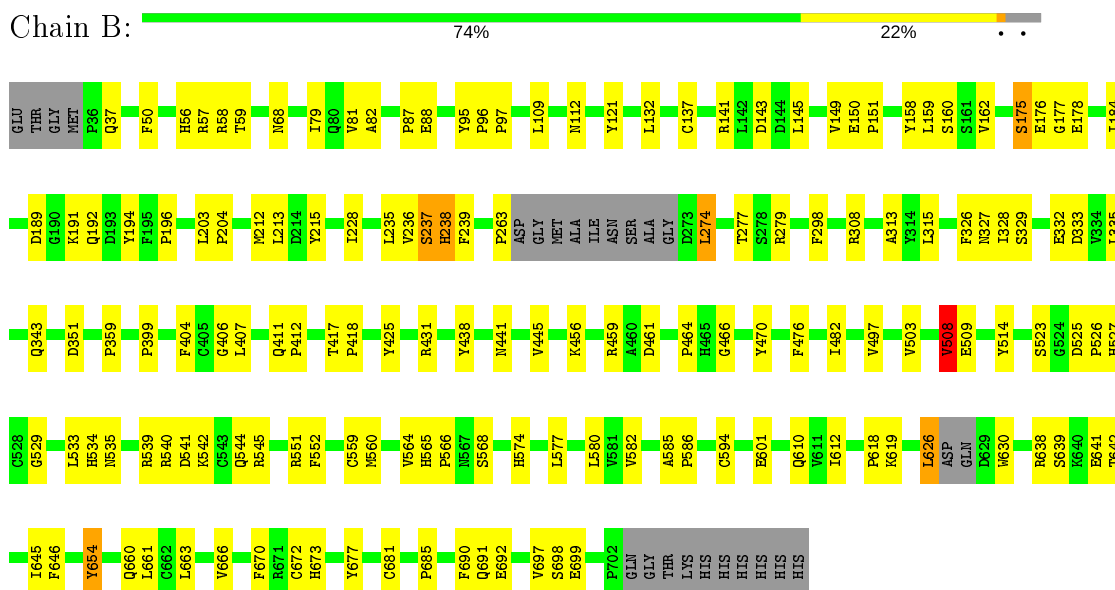
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

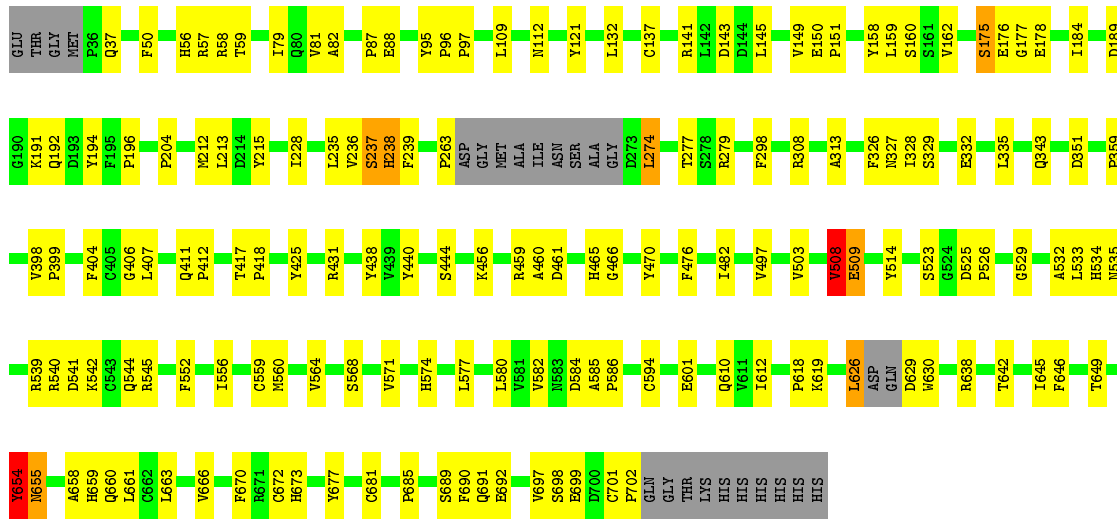
- Molecule 1: Plexin-A2



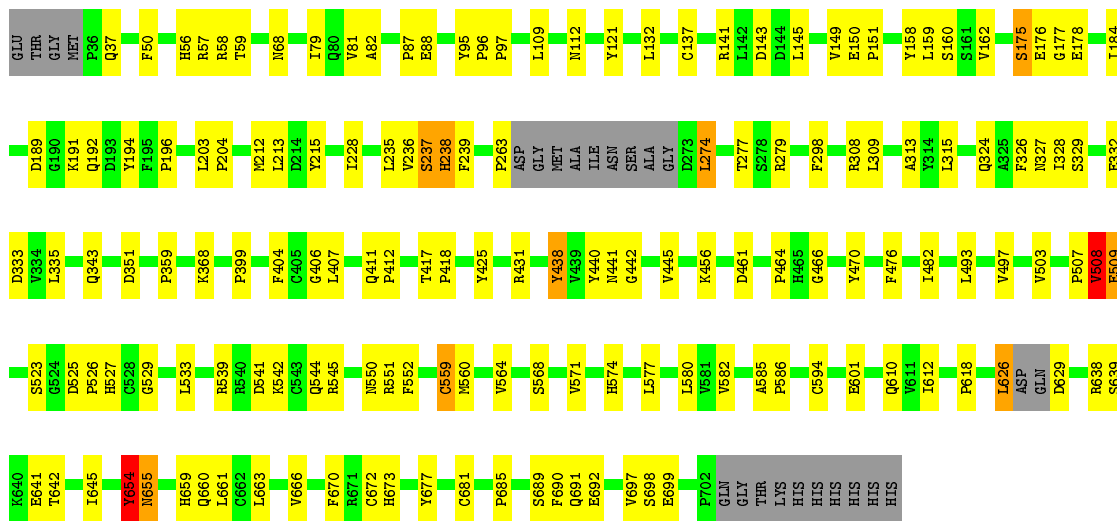
- Molecule 1: Plexin-A2



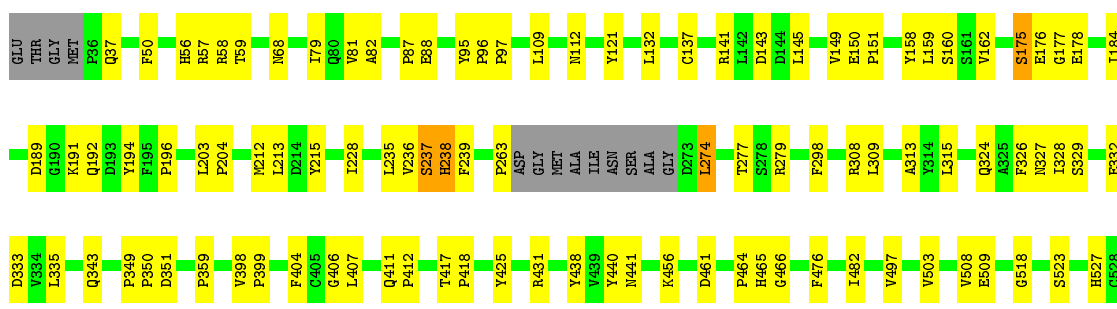
- Molecule 1: Plexin-A2

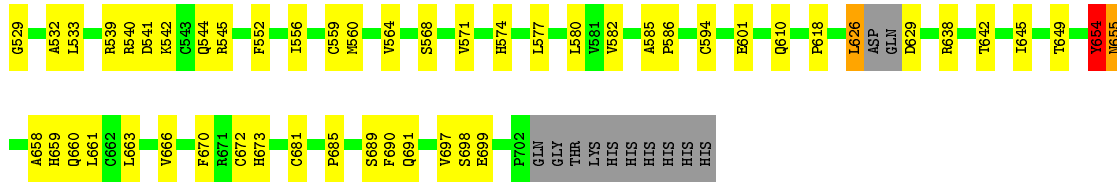


• Molecule 1: Plexin-A2

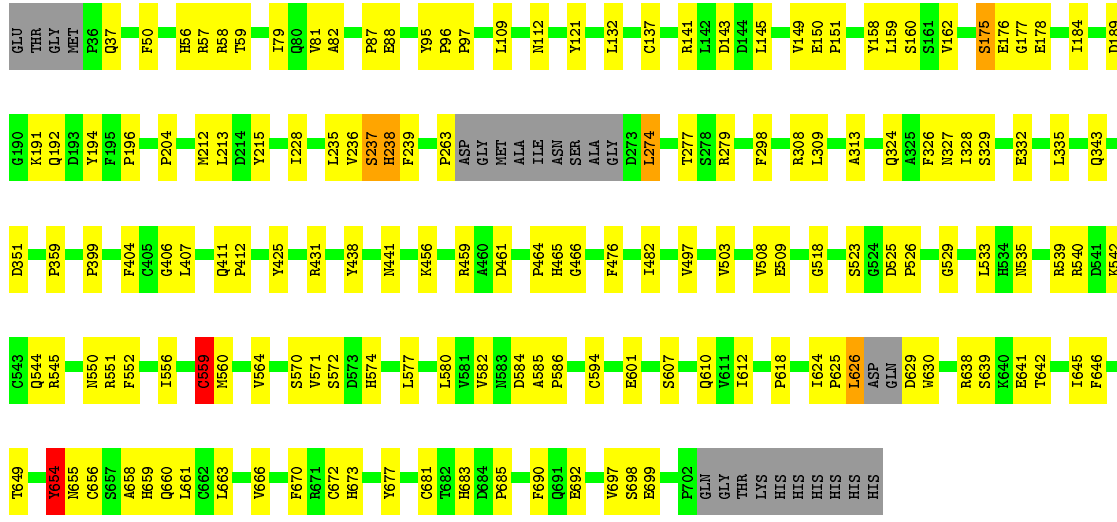


• Molecule 1: Plexin-A2

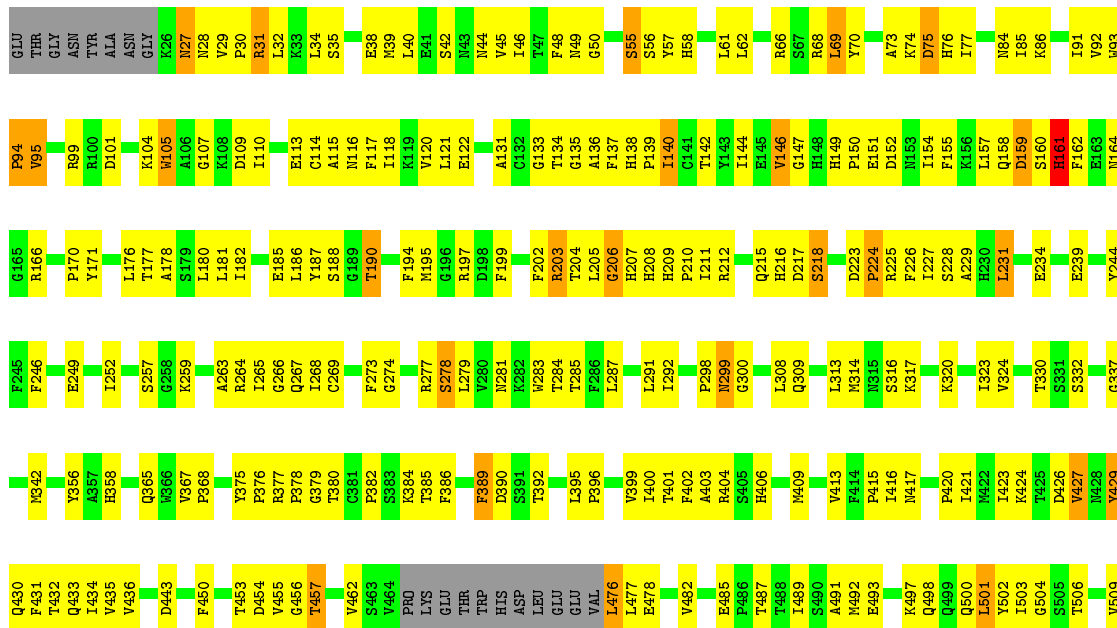




• Molecule 1: Plexin-A2



• Molecule 2: Semaphorin-3A



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	191.63Å 293.60Å 252.18Å 90.00° 106.38° 90.00°	Depositor
Resolution (Å)	125.51 – 7.00 125.51 – 7.00	Depositor EDS
% Data completeness (in resolution range)	95.6 (125.51-7.00) 95.5 (125.51-7.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 6.73Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.285 , 0.285 0.285 , 0.291	Depositor DCC
R_{free} test set	1022 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	293.7	Xtrriage
Anisotropy	0.196	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 423.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.76	EDS
Total number of atoms	35655	wwPDB-VP
Average B, all atoms (Å ²)	323.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.92% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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:
CA

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	0.46	2/5260 (0.0%)	0.48	2/7138 (0.0%)
1	B	0.30	1/5260 (0.0%)	0.53	4/7138 (0.1%)
1	C	0.42	2/5260 (0.0%)	0.72	3/7138 (0.0%)
1	D	0.46	3/5261 (0.1%)	0.69	5/7141 (0.1%)
1	E	0.30	1/5261 (0.0%)	0.72	7/7141 (0.1%)
1	F	0.35	1/5261 (0.0%)	0.54	8/7141 (0.1%)
2	G	0.53	0/3988	0.72	0/5411
3	H	0.52	0/977	0.64	0/1325
All	All	0.41	10/36528 (0.0%)	0.63	29/49573 (0.1%)

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	1
1	B	0	1
1	C	0	2
1	D	0	2
1	E	0	1
1	F	0	2
2	G	0	2
3	H	0	1
All	All	0	12

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	508	VAL	C-N	24.79	1.91	1.34
1	C	508	VAL	C-N	-21.01	0.85	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	559	CYS	C-N	19.66	1.79	1.34
1	F	508	VAL	C-N	12.94	1.63	1.34
1	D	654	TYR	C-N	-12.42	1.05	1.34

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	654	TYR	O-C-N	-44.56	51.41	122.70
1	D	654	TYR	O-C-N	-40.97	57.14	122.70
1	E	654	TYR	O-C-N	-31.49	72.32	122.70
1	E	654	TYR	CA-C-N	21.86	165.30	117.20
1	E	654	TYR	C-N-CA	20.96	174.09	121.70

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	654	TYR	Mainchain
1	B	508	VAL	Mainchain
1	C	508	VAL	Mainchain
1	C	654	TYR	Mainchain
1	D	508	VAL	Mainchain

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	5138	0	4989	288	34
1	B	5138	0	4991	179	18
1	C	5138	0	4990	229	0
1	D	5138	0	4989	188	0
1	E	5138	0	4993	158	29
1	F	5138	0	4995	174	23
2	G	3882	0	3747	291	11
3	H	944	0	875	44	5
4	H	1	0	0	0	0
All	All	35655	0	34569	1242	63

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 18.

The worst 5 of 1242 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:344:LYS:CD	3:H:80:ARG:HH11	1.00	1.58
1:A:344:LYS:HD3	3:H:80:ARG:CD	1.23	1.57
1:A:222:VAL:CG2	2:G:166:ARG:HH12	1.17	1.54
1:D:577:LEU:HD11	1:F:607:SER:CB	1.44	1.47
1:E:324:GLN:NE2	1:F:518:GLY:HA3	1.23	1.47

The worst 5 of 63 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:690:PHE:CE1	1:F:540:ARG:NE[3_455]	0.68	1.52
1:A:690:PHE:CE2	1:F:540:ARG:CG[3_455]	0.78	1.42
1:B:541:ASP:N	1:E:690:PHE:CE2[3_455]	0.79	1.41
1:A:577:LEU:CD1	1:E:610:GLN:OE1[3_455]	0.87	1.33
1:B:540:ARG:C	1:E:690:PHE:CZ[3_455]	0.87	1.33

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	648/681 (95%)	628 (97%)	18 (3%)	2 (0%)	41 77
1	B	648/681 (95%)	629 (97%)	16 (2%)	3 (0%)	29 69
1	C	648/681 (95%)	627 (97%)	17 (3%)	4 (1%)	25 66
1	D	650/681 (95%)	630 (97%)	16 (2%)	4 (1%)	25 66
1	E	650/681 (95%)	632 (97%)	15 (2%)	3 (0%)	29 69
1	F	650/681 (95%)	631 (97%)	17 (3%)	2 (0%)	41 77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	G	480/538 (89%)	408 (85%)	55 (12%)	17 (4%)	3	25
3	H	115/577 (20%)	106 (92%)	7 (6%)	2 (2%)	9	42
All	All	4489/5201 (86%)	4291 (96%)	161 (4%)	37 (1%)	19	60

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	655	ASN
2	G	161	HIS
2	G	207	HIS
2	G	257	SER
3	H	81	ASP

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	578/597 (97%)	571 (99%)	7 (1%)	71	83
1	B	578/597 (97%)	571 (99%)	7 (1%)	71	83
1	C	578/597 (97%)	571 (99%)	7 (1%)	71	83
1	D	578/597 (97%)	571 (99%)	7 (1%)	71	83
1	E	578/597 (97%)	571 (99%)	7 (1%)	71	83
1	F	578/597 (97%)	571 (99%)	7 (1%)	71	83
2	G	428/472 (91%)	397 (93%)	31 (7%)	14	39
3	H	102/506 (20%)	97 (95%)	5 (5%)	25	50
All	All	3998/4560 (88%)	3920 (98%)	78 (2%)	55	74

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	237	SER
2	G	31	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	G	501	LEU
1	F	238	HIS
1	F	438	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	659	HIS
2	G	27	ASN
2	G	417	ASN
1	F	324	GLN
1	B	534	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	3
1	C	2
1	A	2
1	E	1
1	B	1
1	F	1

The worst 5 of 10 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	559:CYS	C	560:MET	N	2.86
1	B	559:CYS	C	560:MET	N	2.56
1	C	559:CYS	C	560:MET	N	2.24
1	A	508:VAL	C	509:GLU	N	1.91
1	D	559:CYS	C	560:MET	N	1.79

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.