



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2016 – 04:12 PM EDT

PDB ID : 5L5K
Title : Plexin A4 full extracellular region, domains 1 to 10, data to 7.5 angstrom, spacegroup P4(1)
Authors : Janssen, B.J.C.; Kong, Y.; Malinauskas, T.; Vangoor, V.R.; Coles, C.H.; Kaufmann, R.; Ni, T.; Gilbert, R.J.C.; Padilla-Parra, S.; Pasterkamp, R.J.; Jones, E.Y.
Deposited on : 2016-05-28
Resolution : 7.50 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

<http://wwpdb.org/validation/2016/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
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

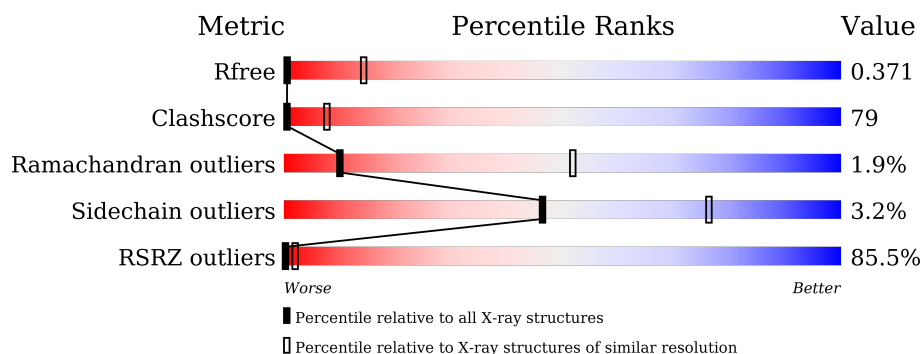
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.50 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}	91344	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
Sidechain outliers	100360	1005 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)

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. 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 $\leq 5\%$. 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	1207	<div> <div>83%</div> <div> <div>33%</div> <div>58%</div> <div>6%</div> </div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9134 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-A4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1168	Total	C	N	O	S	0	0	0
			9134	5761	1572	1729	72			

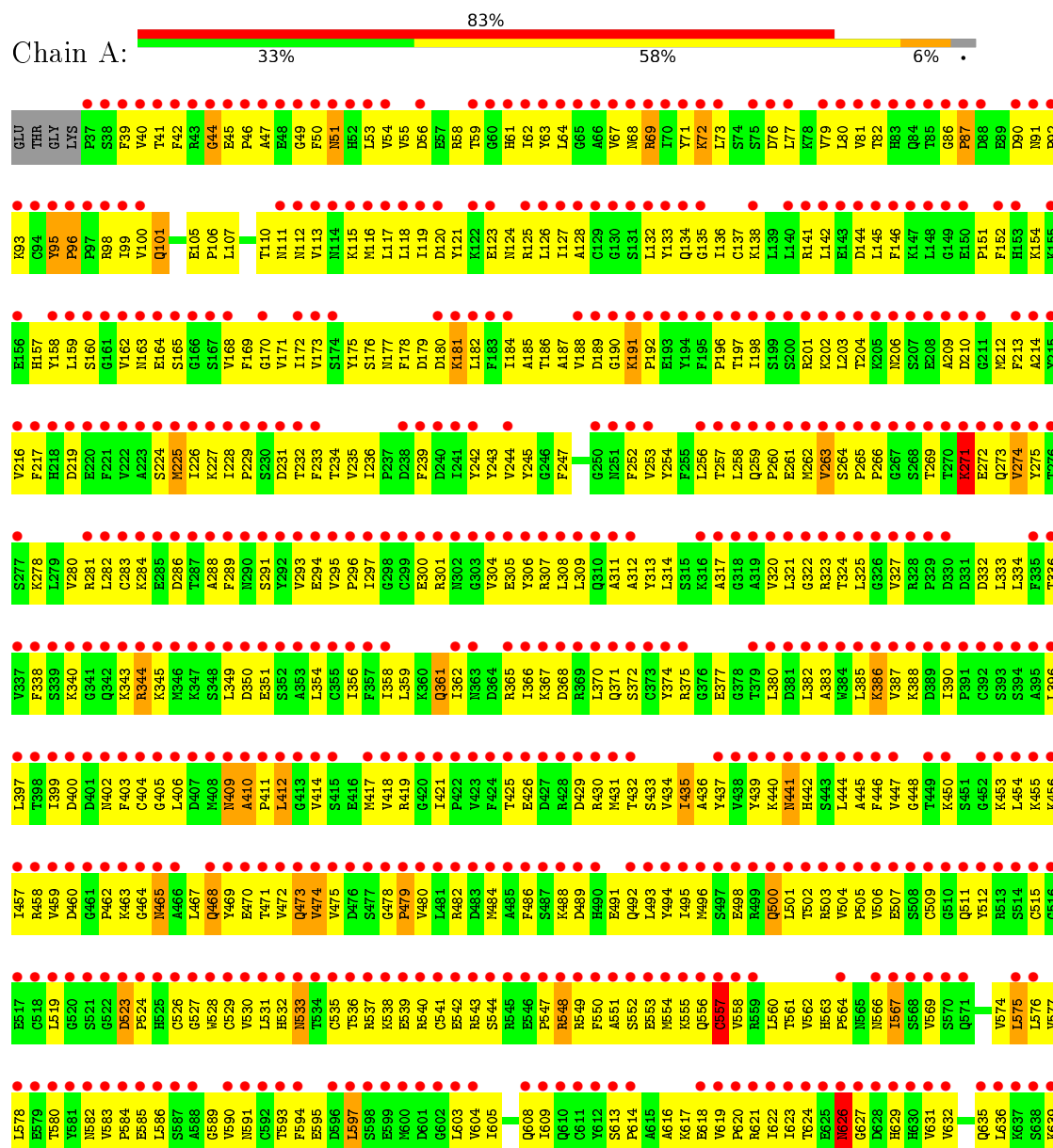
There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	GLU	-	expression tag	UNP Q80UG2
A	34	THR	-	expression tag	UNP Q80UG2
A	35	GLY	-	expression tag	UNP Q80UG2
A	1230	GLY	-	expression tag	UNP Q80UG2
A	1231	ARG	-	expression tag	UNP Q80UG2
A	1232	THR	-	expression tag	UNP Q80UG2
A	1233	LYS	-	expression tag	UNP Q80UG2
A	1234	HIS	-	expression tag	UNP Q80UG2
A	1235	HIS	-	expression tag	UNP Q80UG2
A	1236	HIS	-	expression tag	UNP Q80UG2
A	1237	HIS	-	expression tag	UNP Q80UG2
A	1238	HIS	-	expression tag	UNP Q80UG2
A	1239	HIS	-	expression tag	UNP Q80UG2

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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: Plexin-A4



V1191	L1130	P1070	S1010	Y949	D888	P828	N763	G700	E840
T1192	N1131	Q1071	E1011	P950	I889	G829	T764	F701	T641
V1193	K1132	I1072	E1012	N951	A890	Q830	S765	Q702	G642
S1194	T1133	R1073	V1013	T952	S891	C831	Y766	L703	M643
D1195	N1134	A1074	L1014	L953	H892	T832	L704	L704	L644
V1196	F1135	K1075	D1015	T954	V893	L833	Y767	R705	F645
T1136	T1136	H1076	K1016	L955	K894	H834	T768	V706	A646
Y1137	Y1137	G1077	K1017	A956	V895	Q835	S647	D707	S647
L1198	Y1138	G1078	V1018	D957	A896	H836	T769	K708	T648
L1199	P1139	K1079	T1019	D958	A897	C837	T770	L709	S649
C1200	N1140	E1080	K959	L958	V898	P838	L776	L710	F650
E1201	P1141	H1081	Q1020	K959	V898	P838	L776	L710	F650
S1202	P1142	I1082	Q1021	E999	E899	A839	P777	V711	V651
P1203	F1142	H1082	V1022	N961	C900	H840	V778	P712	F652
N1204	F1143	M1083	D1023	R862	S901	E841	E779	V713	Y653
LEU	A1144	I1084	A1024	G963	P902	S842	L780	E714	M654
ILE	E1145	C1085	A1025	P964	L903	R843	L781	V715	C655
GLY	PHE	E1086	I1026	N965	V904	W844	V782	L716	S656
ARG	SER	V1087	I1027	S966	D905	L845	V783	L716	V657
HIS	PRO	L1088	R1028	G967	G906	E846	W784	K717	H658
K1210	SER	H1089	Q1029	G968	V907	L847	N785		H659
V1211	GLY	A1090	D1030	T969	I908	S848	L721		S660
M1212	ILE	T1091	L1031	Q970	P909	G849	K722		C661
A1213	LEU	E1092	V1032	V971	A910	A850	A723		L662
R1214	GLU	M1093	F1033	T972	E911	N851	K724		S663
V1215	LEU	T1094	Q1034	T973	Q912	S852	N725		C664
G1216	LYS	C1095	Y1035	T974	T913	K853	L726		V665
G1217	PRO	Q1096	V1036	G975	T914	C854	P727		S666
M1218	GLY	A1097	V1037	T976	C915	T855	Q728		E667
E1219	T1158	P1098	D1038	N977	E916	N856	P729		P668
Y1220	P1159	A1099	P1039	L978	N917	P857	Q730		Y669
S1221	I1160	L1103	T1040	N979	G918	R858	S731		R670
PRO	I1161	A1101	L1041	A980	E919	I859	Q732		C671
GLY	L1162	L1102	V1042	G981	A920	T860	Q733		H672
MET	K1163	G1103	R1043	S982	R921	E861	R734		C673
VAL	G1164	P1104	I1044	N983	P922	I862	Q735		C674
TYR	K1165	D1105	E1045	V984	S923	I863	E737		K675
ILE	N1166	H1106	P1046	V985	Q924	P864	G738		Y676
ALA	L1167	Q1107	E1047	V986	H925	V865	G804		R677
PRO	I1168	S1108	W1048	V987	A926	T866	L739		H678
GLY	P1169	D1109	S1049	F988	G927	R867	L740		V679
ARG	P1170	L1110	I1050	G989	F928	P868	N741		C680
THR	V1171	T1111	V1051	S990	V929	R869	I742		T681
LYS	A1172	E1112	S1052	Q991	E930	E870	Q743		H682
HIS		R1113	G1053	P992	T931	G871	S809		D683
HIS	N1175	P1114	N1054	G993	C932	G872	G745		P684
HIS	V1176	E1115	T1055	L994	V933	T873	E746		M685
HIS	K1177	E1116	P1056	F995	A934	K874	Q747		T686
HIS	L1178	F1117	I1057	H996	V935	V875	R748		C687
HIS	N1179	G1118	A1058	R997	C936	T876	A751		S688
HIS	Y1180	F1119	V1059	R998	R937	I877	F689		Q690
	T1181	L1120	W1060	S999	P938	R878	L752		F689
	V1182	L1121	G1061	E939	P938	G879	R753		Q691
	L1183	D1122	T1062	F940	F940	E880	F754		G692
	V1184	N1123	H1063	I1003	N941	N881	N755		G693
	G1185	V1124	L1064	I1004	A942	L882	S756		V694
	E1186	Q1125	D1065	C1005	R943	G883	S757		K695
	K1187	S1126	L1066	M1006		L884	S758		L696
	P1188	L1127	L1067	Q946		E885	W759		P697
	C1189	L1128	Q1068	L947		F886	Q760		E698
	T1190	I1129	N1069	Y948		R887	Q762		D699

4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	220.59Å 220.59Å 65.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.15 – 7.50 61.18 – 7.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (55.15-7.50) 98.5 (61.18-7.50)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.89 (at 7.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.353 , 0.371 0.350 , 0.371	Depositor DCC
R_{free} test set	196 reflections (4.61%)	DCC
Wilson B-factor (Å ²)	344.7	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 399.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.067 for h,-k,-l	Xtriage
F_o, F_c correlation	0.71	EDS
Total number of atoms	9134	wwPDB-VP
Average B, all atoms (Å ²)	250.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.45% 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](#)

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	1.04	30/9324 (0.3%)	1.23	19/12638 (0.2%)

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	4

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	854	CYS	C-N	-21.05	0.85	1.34
1	A	700	CYS	C-N	20.20	1.72	1.34
1	A	557	CYS	C-N	-7.97	1.15	1.34
1	A	49	GLY	CA-C	6.34	1.62	1.51
1	A	626	ASN	CG-OD1	5.79	1.36	1.24

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	747	GLN	CG-CD-NE2	-30.31	43.95	116.70
1	A	557	CYS	CA-C-N	-25.68	60.70	117.20
1	A	557	CYS	C-N-CA	-18.83	74.63	121.70
1	A	747	GLN	CG-CD-OE1	-13.94	93.71	121.60
1	A	479	PRO	N-CA-C	8.16	133.31	112.10

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	557	CYS	Mainchain
1	A	854	CYS	Mainchain
1	A	863	ILE	Peptide
1	A	95	TYR	Peptide

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	9134	0	8998	1436	8
All	All	9134	0	8998	1436	8

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 79.

The worst 5 of 1436 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:440:LYS:CD	1:A:538:LYS:HD2	1.28	1.55
1:A:700:CYS:C	1:A:701:PRO:N	1.72	1.43
1:A:440:LYS:HD3	1:A:538:LYS:CD	1.51	1.41
1:A:530:VAL:HG11	1:A:584:PRO:CD	1.50	1.38
1:A:854:CYS:C	1:A:855:THR:CA	1.93	1.37

The worst 5 of 8 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:146:PHE:CE1	1:A:730:GLN:NE2[4_555]	0.96	1.24
1:A:146:PHE:CD1	1:A:730:GLN:OE1[4_555]	1.76	0.44
1:A:269:THR:CG2	1:A:377:GLU:CD[2_455]	1.90	0.30
1:A:146:PHE:CE1	1:A:730:GLN:CD[4_555]	1.91	0.29
1:A:146:PHE:CD1	1:A:730:GLN:NE2[4_555]	1.96	0.24

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	1150/1207 (95%)	1071 (93%)	57 (5%)	22 (2%)	10	52

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	PRO
1	A	181	LYS
1	A	191	LYS
1	A	410	ALA
1	A	465	ASN

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	1035/1067 (97%)	1002 (97%)	33 (3%)	46	76

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	621	ARG
1	A	773	ILE
1	A	1107	GLN
1	A	626	ASN
1	A	670	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no

such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	9

The worst 5 of 9 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1139:PRO	C	1140:ASN	N	4.11
1	A	506:VAL	C	507:GLU	N	2.75
1	A	951:MET	C	952:THR	N	2.26
1	A	802:LYS	C	803:CYS	N	2.24
1	A	1036:VAL	C	1037:GLU	N	2.02

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	1168/1207 (96%)	5.30	999 (85%) 0 2	185, 272, 310, 310	0

The worst 5 of 999 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	527	GLY	21.4
1	A	551	ALA	21.2
1	A	526	CYS	21.1
1	A	550	PHE	18.3
1	A	413	GLY	17.8

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.