Full wwPDB X-ray Structure Validation Report (i)

Jan 14, 2024 – 01:04 pm GMT

PDB ID	:	6Z4V
Title	:	Crystal structure of the neurotensin receptor 1 (NTSR1-H4bmx) in complex with NTS8-13
Authors	:	Deluigi, M.; Klipp, A.; Hilge, M.; Merklinger, L.; Klenk, C.; Plueckthun, A.
Deposited on	:	2020-05-25
Resolution	:	2.60 Å(reported)

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

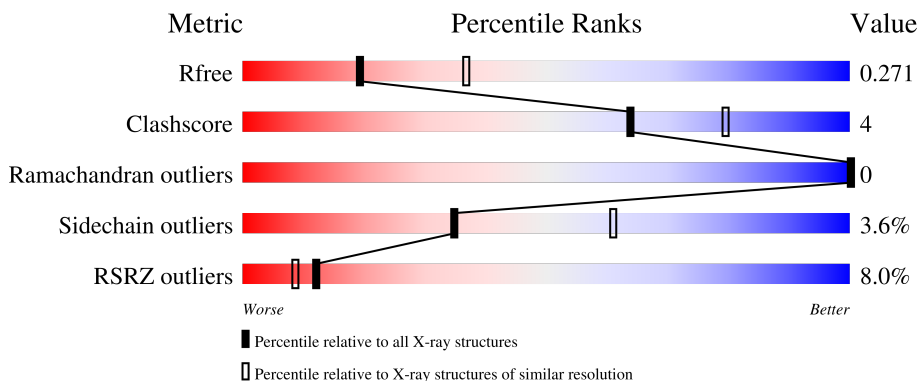|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	Percentile Ranks	Value
Rfree		0.271
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		3.6%
RSRZ outliers		8.0%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	AAA	482	8% — 83% / 11% / 5%
2	BBB	6	83% / 17%

2 Entity composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6855 atoms, of which 3422 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 Neurotensin receptor type 1,Neurotensin receptor type 1,DARPin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	456	6757	2214	3371	552	608	12	182	0	0

There are 53 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	46	GLY	-	expression tag	UNP P20789
AAA	47	PRO	-	expression tag	UNP P20789
AAA	48	GLY	-	expression tag	UNP P20789
AAA	49	SER	-	expression tag	UNP P20789
AAA	83	GLY	SER	engineered mutation	UNP P20789
AAA	86	LEU	ALA	engineered mutation	UNP P20789
AAA	101	ARG	THR	engineered mutation	UNP P20789
AAA	103	ASP	HIS	engineered mutation	UNP P20789
AAA	105	TYR	HIS	engineered mutation	UNP P20789
AAA	119	PHE	LEU	engineered mutation	UNP P20789
AAA	121	LEU	MET	engineered mutation	UNP P20789
AAA	143	LYS	ARG	engineered mutation	UNP P20789
AAA	161	VAL	ALA	engineered mutation	UNP P20789
AAA	167	LEU	ARG	engineered mutation	UNP P20789
AAA	213	LEU	ARG	engineered mutation	UNP P20789
AAA	234	LEU	VAL	engineered mutation	UNP P20789
AAA	235	ARG	LYS	engineered mutation	UNP P20789
AAA	240	LEU	VAL	engineered mutation	UNP P20789
AAA	253	ALA	ILE	engineered mutation	UNP P20789
AAA	260	ALA	ILE	engineered mutation	UNP P20789
AAA	262	ARG	ASN	engineered mutation	UNP P20789
AAA	263	ARG	LYS	engineered mutation	UNP P20789
AAA	?	-	GLU	deletion	UNP P20789
AAA	?	-	GLN	deletion	UNP P20789
AAA	?	-	GLY	deletion	UNP P20789
AAA	?	-	ARG	deletion	UNP P20789
AAA	?	-	VAL	deletion	UNP P20789

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	?	-	CYS	deletion	UNP P20789
AAA	?	-	THR	deletion	UNP P20789
AAA	?	-	VAL	deletion	UNP P20789
AAA	?	-	GLY	deletion	UNP P20789
AAA	?	-	THR	deletion	UNP P20789
AAA	?	-	HIS	deletion	UNP P20789
AAA	?	-	ASN	deletion	UNP P20789
AAA	?	-	GLY	deletion	UNP P20789
AAA	?	-	LEU	deletion	UNP P20789
AAA	?	-	GLU	deletion	UNP P20789
AAA	?	-	HIS	deletion	UNP P20789
AAA	?	-	SER	deletion	UNP P20789
AAA	?	-	THR	deletion	UNP P20789
AAA	305	ARG	HIS	engineered mutation	UNP P20789
AAA	332	VAL	CYS	engineered mutation	UNP P20789
AAA	342	ALA	PHE	engineered mutation	UNP P20789
AAA	354	SER	THR	engineered mutation	UNP P20789
AAA	372	ALA	-	linker	UNP P20789
AAA	373	GLU	-	linker	UNP P20789
AAA	374	ASP	-	linker	UNP P20789
AAA	375	LEU	-	linker	UNP P20789
AAA	376	VAL	-	linker	UNP P20789
AAA	377	GLU	-	linker	UNP P20789
AAA	378	ASP	-	linker	UNP P20789
AAA	379	TRP	-	linker	UNP P20789
AAA	380	GLU	-	linker	UNP P20789

- Molecule 2 is a protein called ARG-PRO-TYR-ILE-LEU.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	BBB	5	98	32	51	8	7	1	0	0

3 Residue-property plots ⓘ

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: Neurotensin receptor type 1,Neurotensin receptor type 1,DARPin

Chain AAA:
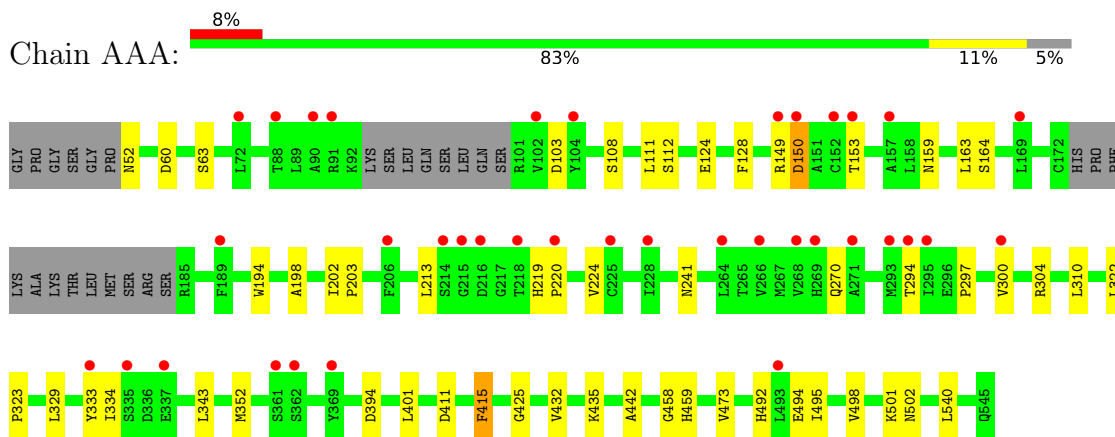
• Molecule 2: ARG-PRO-TYR-ILE-LEU

Chain BBB:
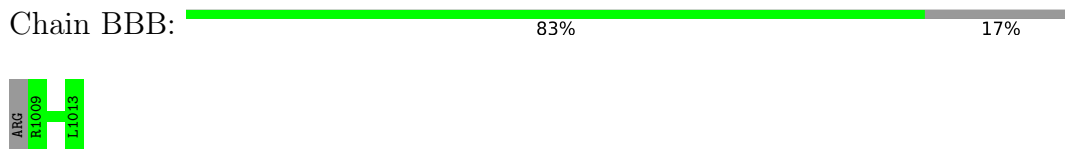
4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	69.22Å 77.53Å 111.86Å 90.00° 107.22° 90.00°	Depositor
Resolution (Å)	29.25 – 2.60 29.23 – 2.60	Depositor EDS
% Data completeness (in resolution range)	80.5 (29.25-2.60) 80.6 (29.23-2.60)	Depositor EDS
R_{merge}	0.45	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.64 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.256 , 0.272 0.260 , 0.271	Depositor DCC
R_{free} test set	723 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	35.2	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 34.4	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6855	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.94% 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 (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	AAA	0.64	0/3463	0.71	0/4755
2	BBB	0.59	0/48	0.68	0/63
All	All	0.64	0/3511	0.71	0/4818

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 (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	AAA	3386	3371	3272	28	0
2	BBB	47	51	50	0	0
All	All	3433	3422	3322	28	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 4.

All (28) 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:AAA:492:HIS:HB3	1:AAA:495:ILE:HD12	1.87	0.57
1:AAA:213:LEU:HD12	1:AAA:224:VAL:HG21	1.88	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:432:VAL:HA	1:AAA:435:LYS:HD2	1.89	0.54
1:AAA:108:SER:HB3	1:AAA:159:ASN:HD21	1.75	0.52
1:AAA:149:ARG:O	1:AAA:153:THR:HG23	2.10	0.52
1:AAA:498:VAL:O	1:AAA:502:ASN:ND2	2.37	0.51
1:AAA:150:ASP:OD2	1:AAA:241:ASN:ND2	2.44	0.51
1:AAA:198:ALA:O	1:AAA:202:ILE:HG13	2.14	0.48
1:AAA:124:GLU:HA	1:AAA:128:PHE:HD2	1.78	0.48
1:AAA:63:SER:OG	1:AAA:352:MET:SD	2.73	0.47
1:AAA:164:SER:HB3	1:AAA:310:LEU:HD21	1.97	0.47
1:AAA:219:HIS:CG	1:AAA:220:PRO:HD2	2.49	0.47
1:AAA:334:ILE:HD13	1:AAA:343:LEU:HD21	1.97	0.47
1:AAA:425:GLY:HA2	1:AAA:459:HIS:CD2	2.50	0.46
1:AAA:111:LEU:HD23	1:AAA:194:TRP:CZ2	2.52	0.45
1:AAA:294:THR:HG21	1:AAA:442:ALA:HA	1.99	0.44
1:AAA:60:ASP:O	1:AAA:63:SER:HB3	2.18	0.44
1:AAA:213:LEU:HD12	1:AAA:224:VAL:CG2	2.48	0.43
1:AAA:322:LEU:HB3	1:AAA:323:PRO:HD3	1.98	0.43
1:AAA:334:ILE:HG22	1:AAA:334:ILE:O	2.19	0.43
1:AAA:458:GLY:HA2	1:AAA:495:ILE:CD1	2.49	0.43
1:AAA:297:PRO:HD2	1:AAA:300:VAL:HG21	2.01	0.42
1:AAA:494:GLU:O	1:AAA:498:VAL:HG13	2.19	0.42
1:AAA:498:VAL:HA	1:AAA:501:LYS:HG2	2.03	0.41
1:AAA:329:LEU:O	1:AAA:333:TYR:HB2	2.20	0.41
1:AAA:202:ILE:N	1:AAA:203:PRO:CD	2.84	0.41
1:AAA:164:SER:HA	1:AAA:310:LEU:HD21	2.02	0.41
1:AAA:411:ASP:OD1	1:AAA:415:PHE:HB2	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	AAA	450/482 (93%)	434 (96%)	16 (4%)	0	100	100
2	BBB	3/6 (50%)	3 (100%)	0	0	100	100
All	All	453/488 (93%)	437 (96%)	16 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

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	AAA	333/398 (84%)	321 (96%)	12 (4%)	35	61
2	BBB	5/6 (83%)	5 (100%)	0	100	100
All	All	338/404 (84%)	326 (96%)	12 (4%)	35	61

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	52	ASN
1	AAA	103	ASP
1	AAA	112	SER
1	AAA	150	ASP
1	AAA	163	LEU
1	AAA	270	GLN
1	AAA	304	ARG
1	AAA	394	ASP
1	AAA	401	LEU
1	AAA	415	PHE
1	AAA	473	VAL
1	AAA	540	LEU

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)

There are no ligands in this entry.

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, 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	AAA	456/482 (94%)	0.55	37 (8%)	12	8	11, 37, 60, 75	0
2	BBB	5/6 (83%)	0.11	0	100	100	34, 37, 44, 56	0
All	All	461/488 (94%)	0.55	37 (8%)	12	9	11, 37, 60, 75	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	294	THR	6.3
1	AAA	335	SER	5.6
1	AAA	266	VAL	4.1
1	AAA	90	ALA	4.1
1	AAA	104	TYR	3.6
1	AAA	206	PHE	3.5
1	AAA	493	LEU	3.4
1	AAA	153	THR	3.3
1	AAA	102	VAL	3.3
1	AAA	333	TYR	3.2
1	AAA	150	ASP	3.2
1	AAA	268	VAL	3.2
1	AAA	215	GLY	3.2
1	AAA	269	HIS	3.1
1	AAA	295	ILE	3.1
1	AAA	152	CYS	2.8
1	AAA	216	ASP	2.8
1	AAA	361	SER	2.8
1	AAA	369	TYR	2.7
1	AAA	293	MET	2.6
1	AAA	91	ARG	2.6
1	AAA	220	PRO	2.5
1	AAA	88	THR	2.5
1	AAA	72	LEU	2.5

Continued on next page...

Mol	Chain	Res	Type	RSRZ
1	AAA	218	THR	2.4
1	AAA	225	CYS	2.4
1	AAA	157	ALA	2.3
1	AAA	337	GLU	2.3
1	AAA	362	SER	2.3
1	AAA	264	LEU	2.3
1	AAA	169	LEU	2.2
1	AAA	271	ALA	2.1
1	AAA	214	SER	2.1
1	AAA	300	VAL	2.1
1	AAA	189	PHE	2.1
1	AAA	149	ARG	2.1
1	AAA	228	ILE	2.0

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.