wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 02:24 AM EDT

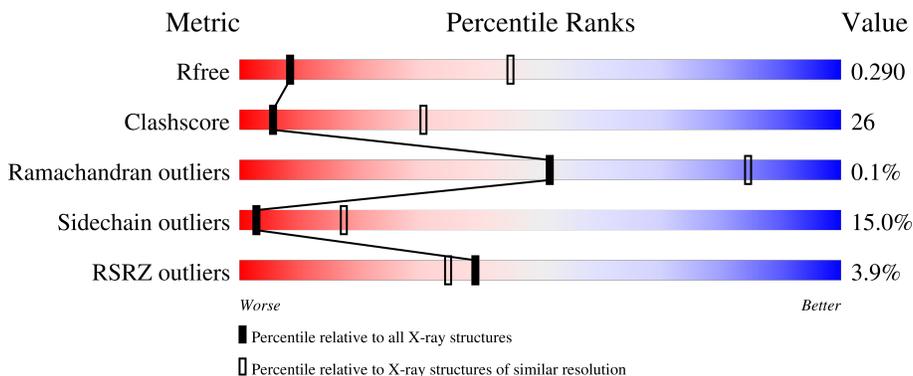|		
PDB ID	:	4FRW
Title	:	Crystal structure of human nectin-4 extracellular fragment D1-D2
Authors	:	Harrison, O.J.; Jin, X.; Brasch, J.; Shapiro, L.
Deposited on	:	2012-06-26
Resolution	:	3.50 Å(reported)

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 3.50 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	Percentile Ranks	Value
Rfree		0.290
Clashscore		26
Ramachandran outliers		0.1%
Sidechain outliers		15.0%
RSRZ outliers		3.9%

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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	218	60% 33% 5% •
1	B	218	55% 34% 9% •
1	C	218	61% 29% 6% •
1	D	218	4% 60% 32% 6% •
1	E	218	9% 56% 33% 6% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	218	

9%

42%   34%   10%   13%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9257 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 Poliovirus receptor-related protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	0	0
			1611	1002	291	313	5			
1	B	213	Total	C	N	O	S	0	0	0
			1601	996	289	311	5			
1	C	211	Total	C	N	O	S	0	0	0
			1571	978	280	308	5			
1	D	212	Total	C	N	O	S	0	0	0
			1561	973	278	305	5			
1	E	208	Total	C	N	O	S	0	0	0
			1533	954	276	298	5			
1	F	189	Total	C	N	O	S	0	0	0
			1380	860	246	269	5			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	244	HIS	-	expression tag	UNP Q96NY8
A	245	HIS	-	expression tag	UNP Q96NY8
A	246	HIS	-	expression tag	UNP Q96NY8
A	247	HIS	-	expression tag	UNP Q96NY8
A	248	HIS	-	expression tag	UNP Q96NY8
A	249	HIS	-	expression tag	UNP Q96NY8
B	244	HIS	-	expression tag	UNP Q96NY8
B	245	HIS	-	expression tag	UNP Q96NY8
B	246	HIS	-	expression tag	UNP Q96NY8
B	247	HIS	-	expression tag	UNP Q96NY8
B	248	HIS	-	expression tag	UNP Q96NY8
B	249	HIS	-	expression tag	UNP Q96NY8
C	244	HIS	-	expression tag	UNP Q96NY8
C	245	HIS	-	expression tag	UNP Q96NY8
C	246	HIS	-	expression tag	UNP Q96NY8
C	247	HIS	-	expression tag	UNP Q96NY8
C	248	HIS	-	expression tag	UNP Q96NY8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	249	HIS	-	expression tag	UNP Q96NY8
D	244	HIS	-	expression tag	UNP Q96NY8
D	245	HIS	-	expression tag	UNP Q96NY8
D	246	HIS	-	expression tag	UNP Q96NY8
D	247	HIS	-	expression tag	UNP Q96NY8
D	248	HIS	-	expression tag	UNP Q96NY8
D	249	HIS	-	expression tag	UNP Q96NY8
E	244	HIS	-	expression tag	UNP Q96NY8
E	245	HIS	-	expression tag	UNP Q96NY8
E	246	HIS	-	expression tag	UNP Q96NY8
E	247	HIS	-	expression tag	UNP Q96NY8
E	248	HIS	-	expression tag	UNP Q96NY8
E	249	HIS	-	expression tag	UNP Q96NY8
F	244	HIS	-	expression tag	UNP Q96NY8
F	245	HIS	-	expression tag	UNP Q96NY8
F	246	HIS	-	expression tag	UNP Q96NY8
F	247	HIS	-	expression tag	UNP Q96NY8
F	248	HIS	-	expression tag	UNP Q96NY8
F	249	HIS	-	expression tag	UNP Q96NY8

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.
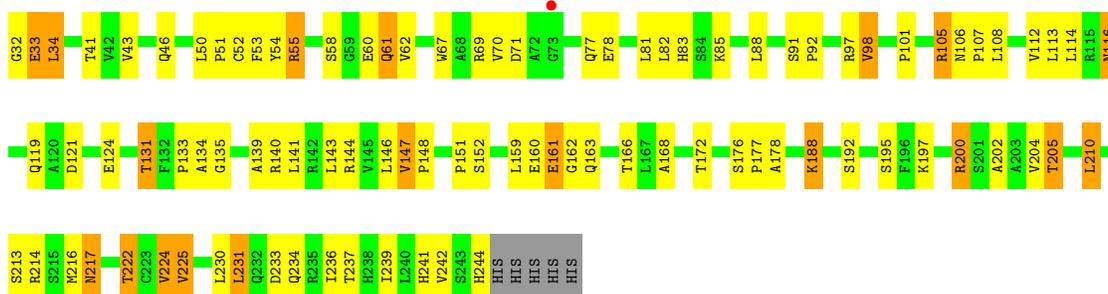
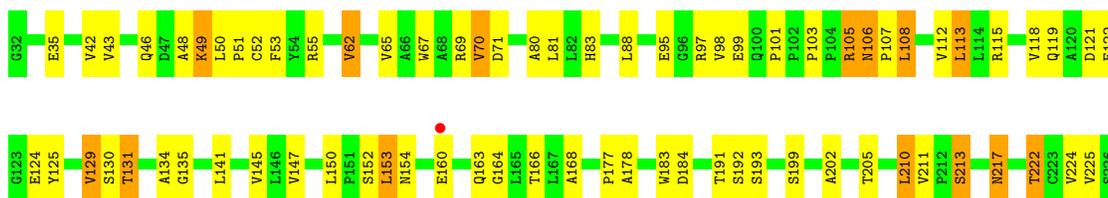- Molecule 1: Poliovirus receptor-related protein 4

Chain A:

- Molecule 1: Poliovirus receptor-related protein 4
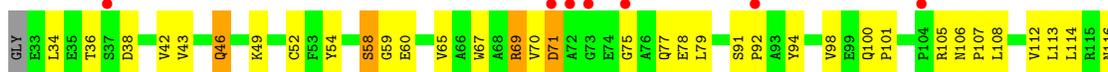
Chain B:

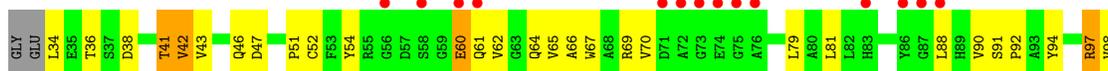- Molecule 1: Poliovirus receptor-related protein 4

Chain C:

H227 P228 T237 V242 SER HIS HIS HIS HIS HIS HIS

• Molecule 1: Poliovirus receptor-related protein 4

Chain D:

4%

60% 32% 6% •

GLY E33 L34 E35 T36 S37 D38 V42 V43 Q46 K49 C52 F53 Y54 S58 G59 E60 V65 A66 W67 A68 R69 V70 D71 A72 G73 E74 G75 A76 Q77 E78 L79 S91 P92 A93 Y94 V98 E99 Q100 P101 P104 R105 N106 P107 L108 V112 L113 L114 R115 N116

A117 V118 Q119 A120 D121 E124 R128 V129 S130 T131 F132 G135 Q138 A139 R140 L141 R142 V145 L146 V147 P151 N154 L159 E160 E161 T172 A173 E174 G175 S176 D184 T185 E186 T190 R194 K197 R200 S201 A202 A203 L210 S213 R214 S215

M216 Q219 T222 C223 V224 V225 S226 H227 L230 Q234 T237 V242 S243 H244 HIS HIS HIS HIS

• Molecule 1: Poliovirus receptor-related protein 4

Chain E:

9%

56% 33% 6% 5%

GLY GLU L34 E35 T36 S37 D38 T41 V42 V43 Q46 D47 P51 C52 F53 Y54 R55 G56 D57 S58 G59 E60 Q61 V62 G63 Q64 V65 A66 W67 A68 R69 V70 D71 A72 G73 E74 G75 A76 L79 A80 L81 L82 H83 Y86 G87 L88 H89 V90 S91 P92 A93 Y94 R97 V98

E99 Q100 P101 P102 P103 P104 R105 N106 P107 V112 L113 L114 R115 V118 Q119 A120 D121 E124 Y125 R128 V129 S130 T131 F132 P133 Q138 A139 R140 L141 R142 L143 R144 V147 P148 P149 L150 P151 S152 L153 L159 Q163 G164 L165 T166 L167 S170 A178 T182 W183

D184 S195 F196 K197 R200 S201 A202 A203 V204 T205 L210 T222 C223 V224 V225 S226 H227 P228 G229 L230 L231 Q232 D233 Q234 R235 I236 T237 H238 I239 L240 I241 VAL SER HIS HIS HIS HIS HIS HIS

• Molecule 1: Poliovirus receptor-related protein 4

Chain F:

9%

42% 34% 10% 13%

GLY GLU LEU GLU T36 S37 D38 V39 V40 T41 V42 V43 L44 A48 K49 L50 P51 C52 F53 TYR ARG GLY ASP SER GLY GLU GLN VAL GLY Q64 V65 A66 W67 A68 R69 V70 D71 A72 G73 E74 G75 A76 Q77 E78 L79 A80 L81 L82 HIS SER LYS TYR GLY LEU H89 V90 S91 P92 A93

Y94 E95 G96 R97 V98 P101 P102 P103 P104 R105 N106 P107 L108 D109 G110 S111 V112 L113 L114 R115 V118 Y125 V129 S130 T131 PHE PRO ALA GLY SER PHE Q138 A139 R140 L141 R142 V147 P148 P149 L150 P151 S152 L153 L159 E160 E161 G162 Q163 G164 L165 T166 S170 C171

T172 A173 S176 P177 A178 T182 W183 D184 T185 E186 V187 T190 T191 S192 S193 R194 S195 S201 A202 A203 V204 T205 S206 E207 F208 H209 L210 V211 P212 S213 R214 T222 C223 V224 V225 S226 H227 P228 L231 Q232 D233 Q234 R235 I236 T237 L240 H244 H245 H246 HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	85.97Å 142.81Å 341.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.50 20.01 – 3.50	Depositor EDS
% Data completeness (in resolution range)	98.7 (20.00-3.50) 99.4 (20.01-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 3.52Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.257 , 0.290 0.260 , 0.290	Depositor DCC
R_{free} test set	1334 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	96.9	Xtriage
Anisotropy	0.312	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 38.1	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.054 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.055 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	9257	wwPDB-VP
Average B, all atoms (Å^2)	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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	A	0.71	0/1649	0.91	4/2251 (0.2%)
1	B	0.71	0/1639	0.86	1/2237 (0.0%)
1	C	0.70	0/1608	0.84	1/2199 (0.0%)
1	D	0.65	0/1598	0.84	1/2187 (0.0%)
1	E	0.68	1/1569 (0.1%)	0.89	4/2148 (0.2%)
1	F	0.68	1/1409 (0.1%)	0.86	3/1930 (0.2%)
All	All	0.69	2/9472 (0.0%)	0.87	14/12952 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	183	TRP	CD2-CE2	5.49	1.48	1.41
1	F	183	TRP	CD2-CE2	5.25	1.47	1.41

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	200	ARG	NE-CZ-NH1	8.38	124.49	120.30
1	C	105	ARG	NE-CZ-NH1	6.79	123.70	120.30
1	A	71	ASP	CB-CG-OD1	6.00	123.70	118.30
1	F	240	LEU	CA-CB-CG	5.74	128.49	115.30
1	B	200	ARG	NE-CZ-NH1	5.67	123.14	120.30

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	A	1611	0	1566	77	0
1	B	1601	0	1562	63	0
1	C	1571	0	1517	64	0
1	D	1561	0	1505	64	0
1	E	1533	0	1478	87	0
1	F	1380	0	1328	118	0
All	All	9257	0	8956	469	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 26.

The worst 5 of 469 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:E:66:ALA:HB1	1:E:128:ARG:CZ	1.22	1.62
1:E:66:ALA:CB	1:E:128:ARG:CZ	2.00	1.40
1:E:66:ALA:HB1	1:E:128:ARG:NH2	1.43	1.33
1:E:101:PRO:O	1:E:105:ARG:HD3	1.38	1.19
1:E:66:ALA:CB	1:E:128:ARG:NE	2.06	1.18

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	A	213/218 (98%)	206 (97%)	7 (3%)	0	100	100
1	B	211/218 (97%)	204 (97%)	7 (3%)	0	100	100
1	C	209/218 (96%)	199 (95%)	10 (5%)	0	100	100
1	D	210/218 (96%)	197 (94%)	12 (6%)	1 (0%)	29	68
1	E	206/218 (94%)	190 (92%)	16 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	181/218 (83%)	167 (92%)	14 (8%)	0	100	100
All	All	1230/1308 (94%)	1163 (95%)	66 (5%)	1 (0%)	51	84

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	59	GLY

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	A	176/181 (97%)	160 (91%)	16 (9%)	9	36
1	B	176/181 (97%)	145 (82%)	31 (18%)	2	10
1	C	171/181 (94%)	146 (85%)	25 (15%)	3	18
1	D	169/181 (93%)	142 (84%)	27 (16%)	2	14
1	E	165/181 (91%)	143 (87%)	22 (13%)	4	21
1	F	150/181 (83%)	120 (80%)	30 (20%)	1	7
All	All	1007/1086 (93%)	856 (85%)	151 (15%)	3	17

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	205	THR
1	F	204	VAL
1	E	230	LEU
1	F	131	THR
1	F	234	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	232	GLN
1	F	198	HIS
1	F	234	GLN
1	F	77	GLN
1	D	219	GLN

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 monosaccharides 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 ⓘ

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	A	215/218 (98%)	-0.43	0	100	100	35, 59, 99, 131	0
1	B	213/218 (97%)	-0.42	1 (0%)	91	88	40, 58, 89, 133	0
1	C	211/218 (96%)	-0.34	1 (0%)	91	88	39, 58, 121, 148	0
1	D	212/218 (97%)	-0.08	8 (3%)	40	36	68, 107, 148, 189	0
1	E	208/218 (95%)	0.24	19 (9%)	9	9	55, 118, 206, 260	0
1	F	189/218 (86%)	0.55	20 (10%)	6	7	114, 145, 192, 263	0
All	All	1248/1308 (95%)	-0.09	49 (3%)	39	35	35, 84, 168, 263	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	107	PRO	5.3
1	E	87	GLY	4.8
1	E	86	TYR	4.7
1	E	56	GLY	4.4
1	E	102	PRO	4.0

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.