



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:36 PM BST

PDB ID : 2NSU
EMDB ID: : EMD-1288
Title : Crystal structure of the ectodomain of human transferrin receptor fitted into a cryo-EM reconstruction of canine parvovirus and feline transferrin receptor complex
Authors : Hafenstein, S.; Kostyuchenko, V.A.; Rossmann, M.G.
Deposited on : 2006-11-06
Resolution : 27.00 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

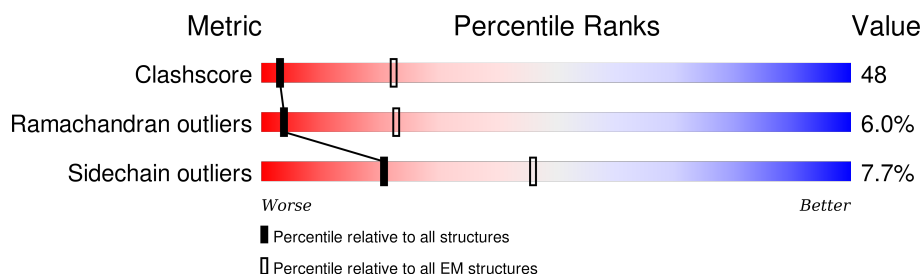
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 27.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	639	
1	B	639	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 10112 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Transferrin receptor protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	639	Total	C	N	O	S	0	0
			5056	3244	846	952	14		
1	B	639	Total	C	N	O	S	0	0
			5056	3244	846	952	14		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	142	SER	GLY	VARIANT	UNP P02786
A	172	GLU	GLN	SEE REMARK 999	UNP P02786
A	613	GLU	ARG	SEE REMARK 999	UNP P02786
B	142	SER	GLY	VARIANT	UNP P02786
B	172	GLU	GLN	SEE REMARK 999	UNP P02786
B	613	GLU	ARG	SEE REMARK 999	UNP P02786

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. 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.

Chain A: 40% 49% 9%

The visualization displays a hierarchical tree structure of 1000 nodes, color-coded by a 'Chain A' metric. The nodes are arranged in a grid-like pattern, with colors ranging from green (40%) to red (9%). The visualization includes a legend at the top and a detailed view of the tree structure below.

The nodes are organized into a grid of 10 columns and 100 rows. The columns are labeled with letters A through J, and the rows are labeled with numbers 1 through 100. The color of each node represents its 'Chain A' value, with a gradient from green (40%) to red (9%).

The visualization shows a complex pattern of colors, with green nodes (40%) concentrated in the leftmost columns (A through D) and red nodes (9%) concentrated in the rightmost columns (I through J). The middle columns (E through H) show a mix of green and red nodes, with a general trend of increasing redness from left to right.

The visualization also includes a legend at the top, which shows the color gradient for 'Chain A' and the corresponding percentages: 40% (green), 49% (yellow), and 9% (red).

Chain B:

State	Category
S199	Green
S200	Green
S201	Green
S202	Green
S203	Green
S204	Green
S205	Green
S206	Green
S207	Yellow
S208	Yellow
S209	Yellow
S210	Yellow
S211	Yellow
S212	Yellow
S213	Yellow
S214	Yellow
S215	Yellow
S216	Yellow
S217	Yellow
S218	Yellow
S219	Yellow
S220	Yellow
S221	Yellow
S222	Yellow
S223	Yellow
S224	Yellow
S225	Orange
S226	Orange
S227	Orange
S228	Orange
S229	Orange
S230	Orange
S231	Orange
S232	Orange
S233	Orange
S234	Orange
S235	Orange
S236	Orange
S237	Orange
S238	Orange
S239	Orange
S240	Orange
S241	Orange
S242	Orange
S243	Orange
S244	Orange
S245	Orange
S246	Orange
S247	Orange
S248	Orange
S249	Orange
S250	Orange
S251	Orange
S252	Orange
S253	Orange
S254	Orange
S255	Orange
S256	Orange
S257	Orange
S258	Orange
S259	Orange
S260	Orange
S261	Orange
S262	Orange
S263	Orange
S264	Orange
S265	Orange
S266	Orange
S267	Orange

P710	K633	C588	N483	P410	R339	E272
A711	L637	E559	L494	A411	A340	A342
L712	S638	D562	D495	W412	A341	1277
L716	L639	Y563	V488	G415	E278	G279
R719	Y643	P564	L489	A416	K344	V279
K720	R646	Y565	G490	W417	L345	L280
K721	R646	L566	T491	K418	F346	1281
N722	G647	G567	V496	L425	N348	Y282
N723	D648	T568	S497	L426	N349	D284
N726	F649	Y570	A498	L427	E350	D285
F726	F650	D571	S499	K428	G351	T286
L730	R651	E575	F500	L429	D352	K287
F731	A652	1580	L501	A430	C353	F288
R732	T653	P581	L502	Q431	P354	
N733	R655	E582	I506	W432	W357	V291
Q734	L656	L583	P507	F433	K358	V292
L735	T657	1584	K508	V437	T359	L295
W740	F660	K585	T509	D440	S360	L296
T741	E664	Y586	K508	G441	S361	F297
I742	K665	V587	M510	F442	T362	F298
Q743	T666	E588	N512	Q443	C363	G299
A746	T667	A589	W513	P444	R364	H300
L749	R668	A590	K514	R446	N365	A301
S750	F669	A591	H515		V366	H302
G751	V670	E592	P517	I449	T367	T305
D752	W671	A594	T518	F450	S368	G306
V753	K672	G595	G519	F450	S370	D307
W754	R673	Q596	Q520	A451	K371	P308
D755	L674	F597	F521	S452	N372	T309
I756	N675	Y598	L522	W453	V373	T310
D757	D676	1599	Y523	S454	K374	P311
N758	R677	T602	Q524	D457	L375	G312
F760	R680	H603	D525	F458	T376	F313
	V681	D604	S530	G459	V377	P314
	E682	V605	E533	S460	S378	S315
	Y683	E606	Y537	V461	N379	F316
	H684	L607	D538	Q462	V380	N317
	F685	W608	N539	A463	I386	H318
	L686	L609	A540	T464	L387	P322
	V690	D610	A541	E465	N388	F323
	S691	E613	L467	W466	V392	S324
	P692	Y614	F542	E468	V397	S326
	K693	Q617	P543	G469	S327	S327
	R698	L618	F544	Y470	E398	G328
	H699	L619	Y547	L471	P399	L329
	V700	S620	I550	S472		P330
	F701	1625	V553	S473	Y402	N331
	W702	N626	S554	L476	V403	I332
	H707	N629	F556	W477	V404	P333
	T708		C556	A478	W405	V334
	L709		C556	F479	G406	Q335
			F557	I482	A407	T336
					Q408	I337
					R409	S338

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	CTF correction of each particle	Depositor
Microscope	FEI/PHILIPS CM200 FEG	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	25.96	Depositor
Minimum defocus (nm)	1700	Depositor
Maximum defocus (nm)	3900	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO163 FILM	Depositor

5 Model quality

5.1 Standard geometry

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 >2	RMSZ	# Z >2
1	A	0.38	0/5177	0.61	1/7021 (0.0%)
1	B	0.38	0/5177	0.61	1/7021 (0.0%)
All	All	0.38	0/10354	0.61	2/14042 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	751	GLY	N-CA-C	-5.19	100.12	113.10
1	A	751	GLY	N-CA-C	-5.14	100.25	113.10

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	A	5056	0	4982	506	0
1	B	5056	0	4982	481	0
All	All	10112	0	9964	970	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 48.

The worst 5 of 970 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:153:VAL:HG13	1:A:154:PRO:HD3	1.20	1.17
1:B:153:VAL:HG13	1:B:154:PRO:HD3	1.21	1.09
1:B:210:VAL:HG13	1:B:211:TYR:H	1.22	1.05
1:B:359:THR:HG22	1:B:360:ASP:H	1.22	1.04
1:A:359:THR:HG22	1:A:360:ASP:H	1.25	0.99

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 PDB entries followed by that with respect to all EM entries.

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	637/639 (100%)	497 (78%)	102 (16%)	38 (6%)	2	26
1	B	637/639 (100%)	498 (78%)	100 (16%)	39 (6%)	2	26
All	All	1274/1278 (100%)	995 (78%)	202 (16%)	77 (6%)	4	26

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	ASN
1	A	210	VAL
1	A	251	ASN
1	A	330	PRO
1	A	361	SER

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 PDB entries followed by that with respect to all EM entries.

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	548/548 (100%)	505 (92%)	43 (8%)	16	51
1	B	548/548 (100%)	507 (92%)	41 (8%)	17	53
All	All	1096/1096 (100%)	1012 (92%)	84 (8%)	21	52

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	648	ASP
1	B	181	VAL
1	B	660	PHE
1	A	723	ASN
1	A	758	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	699	HIS
1	B	160	GLN
1	B	699	HIS
1	A	723	ASN
1	A	758	ASN

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

There are no chain breaks in this entry.