



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 08:32 pm GMT

PDB ID : 4CQ4  
Title : C-terminal fragment of Af1503-sol: transmembrane receptor Af1503 from Archaeoglobus fulgidus engineered for solubility  
Authors : Hartmann, M.D.; Dunin-Horkawicz, S.; Hulko, M.; Martin, J.; Coles, M.; Lupas, A.N.  
Deposited on : 2014-02-11  
Resolution : 1.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

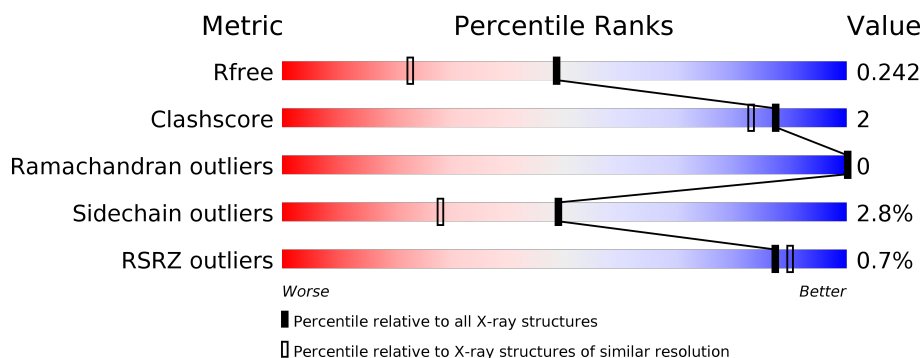
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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))
$R_{free}$	100719	3453 (1.70-1.70)
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)
RSRZ outliers	101464	3491 (1.70-1.70)

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  $\geq 3$ , 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	309	
1	B	309	
1	C	309	
1	D	309	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3465 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 ENGINEERED VERSION OF TRANSMEMBRANE RECEPTOR AF1503.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	0	0	0
			782	480	138	163	1			
1	B	102	Total	C	N	O	S	0	0	0
			747	462	129	155	1			
1	C	102	Total	C	N	O	S	0	2	0
			768	471	136	160	1			
1	D	102	Total	C	N	O	S	0	4	0
			786	483	138	164	1			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	255	LYS	TYR	ENGINEERED MUTATION	UNP O28769
A	256	ASN	TYR	ENGINEERED MUTATION	UNP O28769
A	257	LEU	ALA	ENGINEERED MUTATION	UNP O28769
A	259	THR	GLY	ENGINEERED MUTATION	UNP O28769
A	260	LEU	ILE	ENGINEERED MUTATION	UNP O28769
A	263	ASP	ALA	ENGINEERED MUTATION	UNP O28769
A	264	ARG	ILE	ENGINEERED MUTATION	UNP O28769
A	266	GLU	ILE	ENGINEERED MUTATION	UNP O28769
A	267	GLN	VAL	ENGINEERED MUTATION	UNP O28769
A	268	ILE	PHE	ENGINEERED MUTATION	UNP O28769
A	270	ASN	ILE	ENGINEERED MUTATION	UNP O28769
A	271	ASP	VAL	ENGINEERED MUTATION	UNP O28769
A	274	SER	VAL	ENGINEERED MUTATION	UNP O28769
A	275	THR	PHE	ENGINEERED MUTATION	UNP O28769
B	255	LYS	TYR	ENGINEERED MUTATION	UNP O28769
B	256	ASN	TYR	ENGINEERED MUTATION	UNP O28769
B	257	LEU	ALA	ENGINEERED MUTATION	UNP O28769
B	259	THR	GLY	ENGINEERED MUTATION	UNP O28769
B	260	LEU	ILE	ENGINEERED MUTATION	UNP O28769
B	263	ASP	ALA	ENGINEERED MUTATION	UNP O28769

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	264	ARG	ILE	ENGINEERED MUTATION	UNP O28769
B	266	GLU	ILE	ENGINEERED MUTATION	UNP O28769
B	267	GLN	VAL	ENGINEERED MUTATION	UNP O28769
B	268	ILE	PHE	ENGINEERED MUTATION	UNP O28769
B	270	ASN	ILE	ENGINEERED MUTATION	UNP O28769
B	271	ASP	VAL	ENGINEERED MUTATION	UNP O28769
B	274	SER	VAL	ENGINEERED MUTATION	UNP O28769
B	275	THR	PHE	ENGINEERED MUTATION	UNP O28769
C	255	LYS	TYR	ENGINEERED MUTATION	UNP O28769
C	256	ASN	TYR	ENGINEERED MUTATION	UNP O28769
C	257	LEU	ALA	ENGINEERED MUTATION	UNP O28769
C	259	THR	GLY	ENGINEERED MUTATION	UNP O28769
C	260	LEU	ILE	ENGINEERED MUTATION	UNP O28769
C	263	ASP	ALA	ENGINEERED MUTATION	UNP O28769
C	264	ARG	ILE	ENGINEERED MUTATION	UNP O28769
C	266	GLU	ILE	ENGINEERED MUTATION	UNP O28769
C	267	GLN	VAL	ENGINEERED MUTATION	UNP O28769
C	268	ILE	PHE	ENGINEERED MUTATION	UNP O28769
C	270	ASN	ILE	ENGINEERED MUTATION	UNP O28769
C	271	ASP	VAL	ENGINEERED MUTATION	UNP O28769
C	274	SER	VAL	ENGINEERED MUTATION	UNP O28769
C	275	THR	PHE	ENGINEERED MUTATION	UNP O28769
D	255	LYS	TYR	ENGINEERED MUTATION	UNP O28769
D	256	ASN	TYR	ENGINEERED MUTATION	UNP O28769
D	257	LEU	ALA	ENGINEERED MUTATION	UNP O28769
D	259	THR	GLY	ENGINEERED MUTATION	UNP O28769
D	260	LEU	ILE	ENGINEERED MUTATION	UNP O28769
D	263	ASP	ALA	ENGINEERED MUTATION	UNP O28769
D	264	ARG	ILE	ENGINEERED MUTATION	UNP O28769
D	266	GLU	ILE	ENGINEERED MUTATION	UNP O28769
D	267	GLN	VAL	ENGINEERED MUTATION	UNP O28769
D	268	ILE	PHE	ENGINEERED MUTATION	UNP O28769
D	270	ASN	ILE	ENGINEERED MUTATION	UNP O28769
D	271	ASP	VAL	ENGINEERED MUTATION	UNP O28769
D	274	SER	VAL	ENGINEERED MUTATION	UNP O28769
D	275	THR	PHE	ENGINEERED MUTATION	UNP O28769

- Molecule 2 is water.

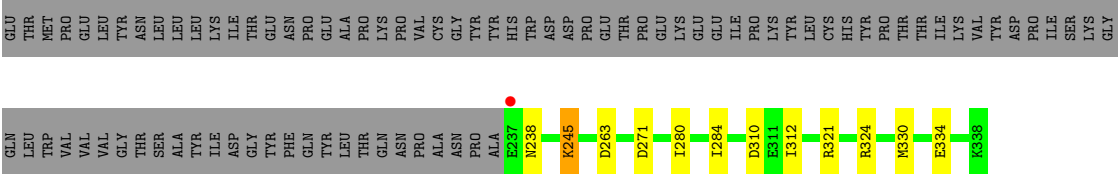
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	87	Total O 87 87	0	0

*Continued on next page...*

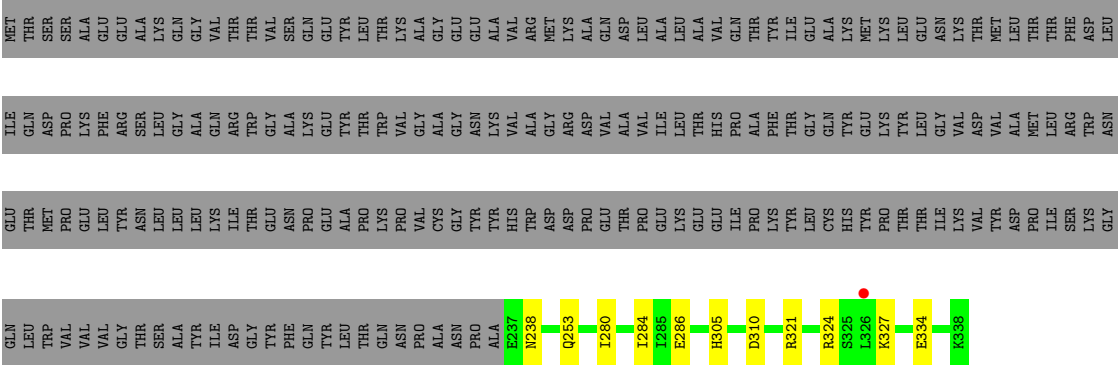
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	96	Total 96	O 96	0	0
2	C	103	Total 103	O 103	0	0
2	D	96	Total 96	O 96	0	0





● Molecule 1: ENGINEERED VERSION OF TRANSMEMBRANE RECEPTOR AF1503



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	44.07Å 48.04Å 95.19Å 90.00° 98.04° 90.00°	Depositor
Resolution (Å)	37.67 – 1.70 37.64 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.2 (37.67-1.70) 99.2 (37.64-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.193 , 0.243 0.198 , 0.242	Depositor DCC
$R_{free}$ test set	2185 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	19.4	Xtriage
Anisotropy	0.122	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3465	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.71% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 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	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.80	0/785	0.88	3/1063 (0.3%)
1	B	0.83	0/749	0.92	3/1013 (0.3%)
1	C	0.87	0/776	1.02	5/1048 (0.5%)
1	D	0.93	0/800	1.01	3/1080 (0.3%)
All	All	0.86	0/3110	0.96	14/4204 (0.3%)

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	321	ARG	NE-CZ-NH2	-8.51	116.05	120.30
1	C	321	ARG	NE-CZ-NH1	7.74	124.17	120.30
1	B	321	ARG	NE-CZ-NH2	-6.79	116.91	120.30
1	B	321	ARG	NE-CZ-NH1	6.79	123.69	120.30
1	D	321	ARG	NE-CZ-NH2	-6.60	117.00	120.30
1	D	321	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	C	263	ASP	CB-CG-OD2	5.41	123.17	118.30
1	A	271	ASP	CB-CG-OD2	-5.39	113.45	118.30
1	A	308	ARG	NE-CZ-NH2	-5.36	117.62	120.30
1	B	292	ASP	CB-CG-OD1	5.31	123.08	118.30
1	D	310	ASP	CB-CG-OD2	-5.25	113.57	118.30
1	A	321	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	C	324	ARG	NE-CZ-NH1	5.14	122.87	120.30
1	C	324	ARG	NE-CZ-NH2	-5.09	117.76	120.30

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	782	0	793	4	0
1	B	747	0	757	5	0
1	C	768	0	789	7	0
1	D	786	0	816	3	0
2	A	87	0	0	1	0
2	B	96	0	0	0	0
2	C	103	0	0	1	0
2	D	96	0	0	1	0
All	All	3465	0	3155	15	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 2.

All (15) 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:335:GLU:OE2	2:A:2085:HOH:O	2.14	0.65
1:B:239:ILE:HD11	1:C:312:ILE:HG12	1.87	0.56
1:C:280:ILE:HG22	1:C:284:ILE:HD12	1.89	0.54
1:A:280:ILE:HD13	1:B:280:ILE:HG21	1.90	0.54
1:D:286:GLU:OE1	1:D:305:HIS:NE2	2.36	0.48
1:D:280:ILE:HG22	1:D:284[A]:ILE:HD12	1.96	0.48
1:A:241:SER:O	1:A:245:LYS:HD3	2.15	0.47
1:C:330:MET:O	1:C:334:GLU:HG3	2.16	0.46
1:B:282:ARG:CB	1:B:283:PRO:HD3	2.46	0.46
1:D:324:ARG:NH2	2:D:2063:HOH:O	2.50	0.45
1:B:237:GLU:O	1:C:310:ASP:HB3	2.18	0.43
1:C:271:ASP:OD2	2:C:2032:HOH:O	2.22	0.42
1:B:311:GLU:OE2	1:C:238:ASN:HB2	2.20	0.41
1:A:284:ILE:HD11	1:A:312:ILE:HD12	2.03	0.41

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	104/309 (34%)	103 (99%)	1 (1%)	0	100	100
1	B	100/309 (32%)	99 (99%)	1 (1%)	0	100	100
1	C	102/309 (33%)	101 (99%)	1 (1%)	0	100	100
1	D	104/309 (34%)	103 (99%)	1 (1%)	0	100	100
All	All	410/1236 (33%)	406 (99%)	4 (1%)	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	A	83/261 (32%)	82 (99%)	1 (1%)	75	64
1	B	78/261 (30%)	75 (96%)	3 (4%)	38	16
1	C	84/261 (32%)	83 (99%)	1 (1%)	75	64
1	D	88/261 (34%)	84 (96%)	4 (4%)	32	12
All	All	333/1044 (32%)	324 (97%)	9 (3%)	49	30

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	322	LEU
1	B	274	SER
1	B	326	LEU
1	B	338	LYS
1	C	245	LYS
1	D	238	ASN
1	D	253	GLN
1	D	327	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	334	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	270	ASN

### 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](#)

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	106/309 (34%)	0.03	1 (0%) 84 87	16, 28, 42, 56	0
1	B	102/309 (33%)	-0.13	0 100 100	16, 26, 39, 46	0
1	C	102/309 (33%)	-0.24	1 (0%) 82 86	15, 23, 37, 59	0
1	D	102/309 (33%)	-0.12	1 (0%) 82 86	13, 22, 34, 46	0
All	All	412/1236 (33%)	-0.12	3 (0%) 87 90	13, 25, 40, 59	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	233	ALA	4.9
1	C	237	GLU	2.5
1	D	326	LEU	2.1

### 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.