wwPDB EM Validation Summary Report

Nov 13, 2024 – 10:55 AM JST

PDB ID	:	8JHZ
EMDB ID	:	EMD-36301
Title	:	Cryo-EM structure of the TcsH-TMPRSS2 complex
Authors	:	Zhou, R.; Liang, T.; Zhan, X.
Deposited on	:	2023-05-25
Resolution	:	3.20 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

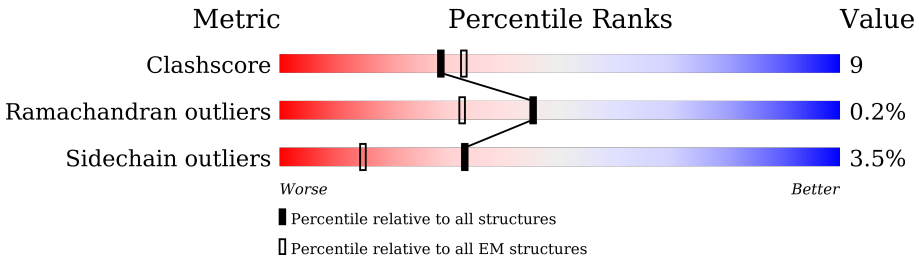
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	A	2626	81% 17% .
2	B	424	55% 22% .. 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 23814 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 Hemorrhagic toxin.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	2615	Total	C	N	O	P	S	0	0
			21144	13557	3425	4120	1	41		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1833	SER	LEU	conflict	UNP M9ZTT7
A	2619	HIS	-	expression tag	UNP M9ZTT7
A	2620	HIS	-	expression tag	UNP M9ZTT7
A	2621	HIS	-	expression tag	UNP M9ZTT7
A	2622	HIS	-	expression tag	UNP M9ZTT7
A	2623	HIS	-	expression tag	UNP M9ZTT7
A	2624	HIS	-	expression tag	UNP M9ZTT7
A	2625	HIS	-	expression tag	UNP M9ZTT7
A	2626	HIS	-	expression tag	UNP M9ZTT7

- Molecule 2 is a protein called Transmembrane protease serine 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	341	Total	C	N	O	S	0	0
			2669	1688	462	496	23		

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	69	MET	-	initiating methionine	UNP O15393
B	70	GLY	-	expression tag	UNP O15393
B	71	ILE	-	expression tag	UNP O15393
B	72	LEU	-	expression tag	UNP O15393
B	73	PRO	-	expression tag	UNP O15393
B	74	SER	-	expression tag	UNP O15393
B	75	PRO	-	expression tag	UNP O15393
B	76	GLY	-	expression tag	UNP O15393

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	77	MET	-	expression tag	UNP O15393
B	78	PRO	-	expression tag	UNP O15393
B	79	ALA	-	expression tag	UNP O15393
B	80	LEU	-	expression tag	UNP O15393
B	81	LEU	-	expression tag	UNP O15393
B	82	SER	-	expression tag	UNP O15393
B	83	LEU	-	expression tag	UNP O15393
B	84	VAL	-	expression tag	UNP O15393
B	85	SER	-	expression tag	UNP O15393
B	86	LEU	-	expression tag	UNP O15393
B	87	LEU	-	expression tag	UNP O15393
B	88	SER	-	expression tag	UNP O15393
B	89	VAL	-	expression tag	UNP O15393
B	90	LEU	-	expression tag	UNP O15393
B	91	LEU	-	expression tag	UNP O15393
B	92	MET	-	expression tag	UNP O15393
B	93	GLY	-	expression tag	UNP O15393
B	94	CYS	-	expression tag	UNP O15393
B	95	VAL	-	expression tag	UNP O15393
B	96	ALA	-	expression tag	UNP O15393
B	97	GLU	-	expression tag	UNP O15393
B	98	THR	-	expression tag	UNP O15393
B	99	GLY	-	expression tag	UNP O15393
B	100	HIS	-	expression tag	UNP O15393
B	101	HIS	-	expression tag	UNP O15393
B	102	HIS	-	expression tag	UNP O15393
B	103	HIS	-	expression tag	UNP O15393
B	104	HIS	-	expression tag	UNP O15393
B	105	HIS	-	expression tag	UNP O15393
B	255	GLN	ARG	conflict	UNP O15393

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

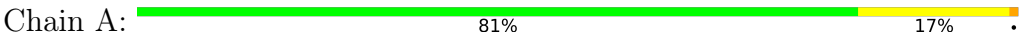
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
3	A	1	1	1	0

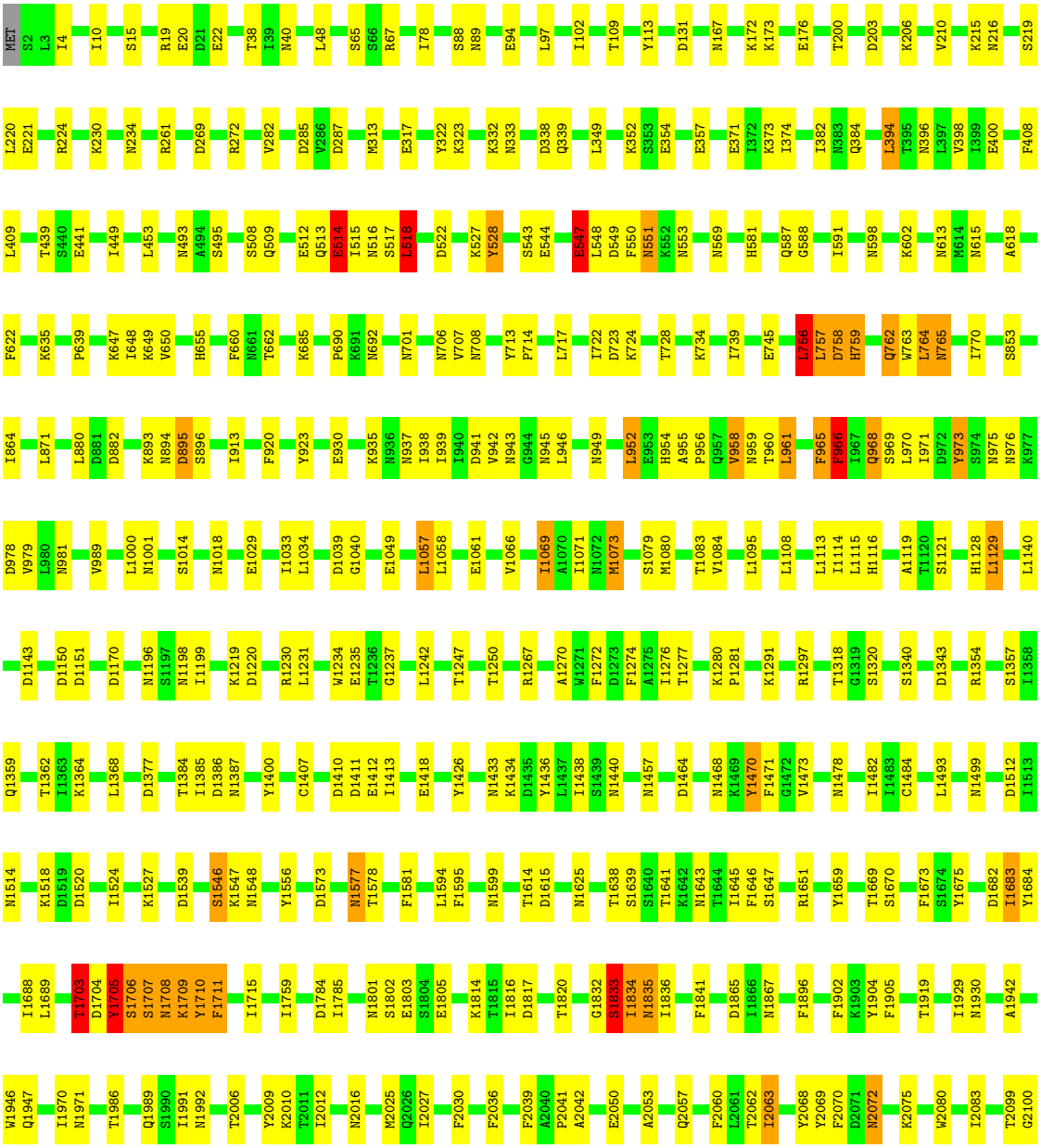
3 Residue-property plots (i)

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. 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.

- Molecule 1: Hemorrhagic toxin

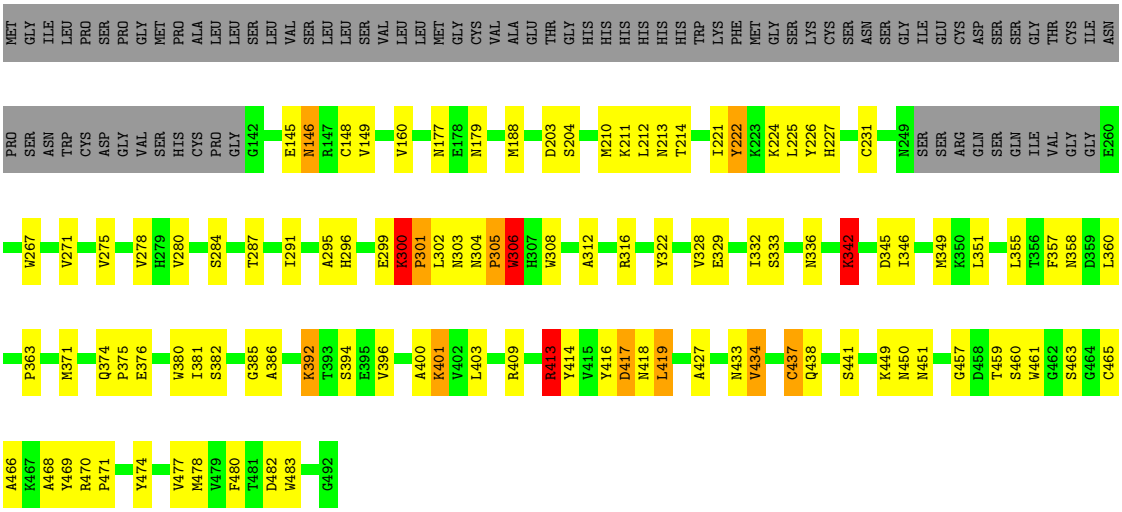
Chain A:





• Molecule 2: Transmembrane protease serine 2

Chain B:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	87378	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50.00	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SEP

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	0.50	22/21600 (0.1%)	0.57	22/29255 (0.1%)
2	B	0.55	8/2740 (0.3%)	0.74	13/3724 (0.3%)
All	All	0.51	30/24340 (0.1%)	0.59	35/32979 (0.1%)

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	3
2	B	0	2
All	All	0	5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	304	ASN	C-O	-9.72	1.04	1.23
1	A	1833	SER	C-O	-9.71	1.04	1.23
1	A	1832	GLY	C-O	-9.40	1.08	1.23
1	A	1471	PHE	C-O	-9.06	1.06	1.23
1	A	1710	TYR	C-O	-8.11	1.07	1.23

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	408	PHE	CB-CA-C	-14.12	82.17	110.40
1	A	959	ASN	CB-CA-C	-10.30	89.80	110.40
1	A	1471	PHE	CB-CA-C	-9.20	92.01	110.40
1	A	960	THR	CA-CB-OG1	-8.96	90.19	109.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1835	ASN	CB-CA-C	-8.86	92.69	110.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1703	THR	Mainchain
1	A	1833	SER	Mainchain
1	A	756	LEU	Mainchain
2	B	300	LYS	Mainchain
2	B	306	TRP	Mainchain

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	21144	0	20628	374	0
2	B	2669	0	2561	58	0
3	A	1	0	0	0	0
All	All	23814	0	23189	431	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 9.

The worst 5 of 431 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:2308:PHE:CZ	1:A:2331:ILE:HD13	1.37	1.56
1:A:2326:THR:CB	1:A:2340:PRO:HA	1.49	1.42
1:A:2308:PHE:HZ	1:A:2331:ILE:CD1	1.32	1.39
1:A:2326:THR:OG1	1:A:2340:PRO:CA	1.75	1.33
1:A:655:HIS:HE1	1:A:759:HIS:CE1	1.51	1.28

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	2612/2626 (100%)	2499 (96%)	109 (4%)	4 (0%)	44 75
2	B	337/424 (80%)	300 (89%)	34 (10%)	3 (1%)	14 49
All	All	2949/3050 (97%)	2799 (95%)	143 (5%)	7 (0%)	45 75

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1707	SER
2	B	300	LYS
1	A	1703	THR
2	B	145	GLU
1	A	1705	TYR

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	2352/2362 (100%)	2290 (97%)	62 (3%)	41 70
2	B	290/360 (81%)	259 (89%)	31 (11%)	5 24
All	All	2642/2722 (97%)	2549 (96%)	93 (4%)	33 63

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2460	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	306	TRP
1	A	2565	ASP
2	B	204	SER
2	B	342	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1909	ASN
1	A	2047	ASN
1	A	1959	ASN
1	A	1993	ASN
1	A	2058	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)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	SEP	A	517	1	8,9,10	1.54	2 (25%)	8,12,14	2.01	2 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SEP	A	517	1	-	4/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	517	SEP	P-OG	2.40	1.67	1.60
1	A	517	SEP	P-O3P	-2.22	1.46	1.54

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	517	SEP	OG-P-O1P	3.87	117.34	106.47
1	A	517	SEP	OG-CB-CA	2.95	111.02	108.14

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	517	SEP	N-CA-CB-OG
1	A	517	SEP	CB-OG-P-O2P
1	A	517	SEP	CB-OG-P-O3P
1	A	517	SEP	CB-OG-P-O1P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.