



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2024 – 10:18 AM EDT

PDB ID : 9DPF  
Title : Crystal structure of TMPRSS11D S368A complexed with its own zymogen activation motif  
Authors : Fraser, B.J.; Dong, A.; Ilyassov, O.; Seitova, A.; Li, Y.; Edwards, A.; Benard, F.; Arrowsmith, C.; Structural Genomics Consortium (SGC)  
Deposited on : 2024-09-21  
Resolution : 1.90 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

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:

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

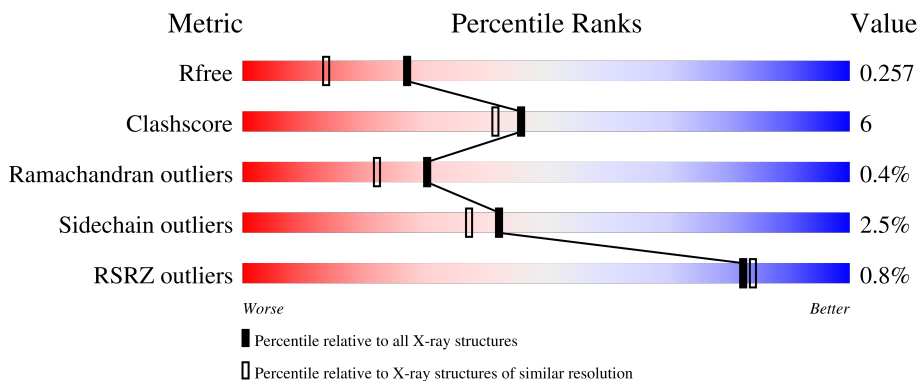
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	232	88% 11%
1	C	232	89% 11%
2	B	18	17% 61% 33% 6%
2	D	18	61% 22% 6% 11%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4128 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 Transmembrane protease serine 11D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	232	1802	1132	326	333	11	0	6	0
1	C	232	1812	1136	326	339	11	0	7	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	368	ALA	SER	engineered mutation	UNP O60235
C	368	ALA	SER	engineered mutation	UNP O60235

- Molecule 2 is a protein called Transmembrane protease serine 11D non-catalytic chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	18	127	76	23	27	1	0	0	0
2	D	16	114	67	21	25	1	0	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0


- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	122	Total O 122 122	0	0
4	B	5	Total O 5 5	0	0
4	C	129	Total O 129 129	0	0
4	D	9	Total O 9 9	0	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 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: Transmembrane protease serine 11D

Chain A:  88% 11%



- Molecule 1: Transmembrane protease serine 11D

Chain C:  89% 11%



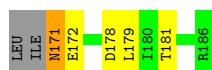
- Molecule 2: Transmembrane protease serine 11D non-catalytic chain

Chain B:  17% 61% 33% 6%



- Molecule 2: Transmembrane protease serine 11D non-catalytic chain

Chain D:  61% 22% 6% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.03Å 76.64Å 88.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.31 – 1.90 38.31 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.2 (38.31-1.90) 99.2 (38.31-1.90)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 1.91Å)	Xtrriage
Refinement program	REFMAC 5.8.0352	Depositor
R, $R_{free}$	0.199 , 0.254 0.208 , 0.257	Depositor DCC
$R_{free}$ test set	1999 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.8	Xtrriage
Anisotropy	0.835	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.009 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4128	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.76 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.6020e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

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

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	0.45	0/1865	0.76	0/2546
1	C	0.47	0/1874	0.74	0/2558
2	B	0.60	0/127	1.04	0/170
2	D	0.57	0/114	0.93	0/152
All	All	0.47	0/3980	0.77	0/5426

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	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	236	ARG	Sidechain
1	C	230	ARG	Sidechain

## 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	1802	0	1732	21	0
1	C	1812	0	1737	17	0
2	B	127	0	119	5	0
2	D	114	0	105	6	0
3	C	8	0	12	0	0
4	A	122	0	0	3	0
4	B	5	0	0	0	0
4	C	129	0	0	2	0
4	D	9	0	0	1	0
All	All	4128	0	3705	43	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 6.

All (43) 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:292:CYS:HG	2:B:173:CYS:HG	0.87	0.85
1:C:379:ARG:HH12	2:D:178:ASP:HB3	1.48	0.77
2:D:171:ASN:HD22	2:D:171:ASN:N	1.91	0.68
1:A:298:GLN:NE2	1:A:382[B]:PHE:CE2	2.67	0.61
2:D:181:THR:HG23	4:D:205:HOH:O	2.03	0.59
1:C:262[B]:ASN:ND2	1:C:263:ASN:OD1	2.37	0.57
1:C:201:GLN:NE2	1:C:309:THR:OG1	2.37	0.56
1:C:379:ARG:HH12	2:D:178:ASP:CB	2.16	0.56
1:A:383:ILE:HG23	1:A:383:ILE:O	2.06	0.54
1:A:211[B]:HIS:NE2	4:A:501:HOH:O	2.26	0.53
1:C:334:ASN:O	1:C:338:ASN:HB2	2.08	0.52
1:C:201:GLN:HE22	1:C:371:PRO:HD2	1.75	0.52
1:A:292:CYS:HG	2:B:173:CYS:CB	2.22	0.51
1:A:370:GLY:HA3	4:A:550:HOH:O	2.12	0.49
1:A:411:TRP:O	1:A:415:GLN:HG2	2.13	0.48
1:A:298:GLN:NE2	1:A:382[B]:PHE:CD2	2.82	0.48
1:A:337:CYS:SG	1:A:399:PRO:HG2	2.54	0.47
2:B:170:ILE:C	2:B:172:GLU:H	2.18	0.47

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:270:GLU:HG2	1:A:348:LEU:HD12	1.97	0.46
1:A:358:GLN:OE1	1:A:358:GLN:N	2.34	0.46
1:A:316:TYR:O	1:A:317:ALA:HB3	2.15	0.46
1:A:286:LYS:HA	2:B:180:ILE:HD13	1.97	0.45
1:C:330:ARG:HH11	1:C:357:PRO:HD3	1.82	0.45
1:C:370:GLY:HA3	4:C:682:HOH:O	2.17	0.45
1:C:379:ARG:NH1	2:D:178:ASP:HB3	2.25	0.45
1:C:352:LEU:C	1:C:352:LEU:HD23	2.37	0.45
1:A:206:LEU:HB3	1:A:211[B]:HIS:CD2	2.51	0.45
2:B:170:ILE:O	2:B:172:GLU:N	2.47	0.45
2:D:171:ASN:N	2:D:171:ASN:ND2	2.62	0.45
1:A:212[B]:CYS:HB2	1:A:368:ALA:O	2.17	0.44
1:C:294:PRO:HD3	1:C:383:ILE:O	2.18	0.44
1:C:413:ARG:HG2	1:C:413:ARG:HH11	1.83	0.43
1:C:188:LEU:O	1:C:361:VAL:HA	2.19	0.43
1:C:200:TRP:CG	1:C:291:VAL:HB	2.54	0.43
1:A:200:TRP:CG	1:A:291:VAL:HB	2.54	0.42
1:C:323:GLU:HG3	4:C:714:HOH:O	2.19	0.42
1:A:221:TRP:CE2	1:A:416:THR:HG22	2.55	0.42
1:C:378:ARG:HB2	1:C:380:LEU:HD22	2.02	0.42
1:A:236:ARG:CZ	1:A:236:ARG:CB	2.99	0.41
1:A:221:TRP:CD2	1:A:416:THR:HG22	2.56	0.41
1:A:256:ARG:HD2	4:A:579:HOH:O	2.21	0.41
1:A:352:LEU:C	1:A:352:LEU:HD23	2.41	0.41
1:C:212[B]:CYS:CB	1:C:228:CYS:SG	3.09	0.41

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 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	236/232 (102%)	232 (98%)	4 (2%)	0	<a href="#">100</a> <a href="#">100</a>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	237/232 (102%)	231 (98%)	6 (2%)	0	100	100
2	B	16/18 (89%)	13 (81%)	2 (12%)	1 (6%)	1	0
2	D	14/18 (78%)	13 (93%)	0	1 (7%)	1	0
All	All	503/500 (101%)	489 (97%)	12 (2%)	2 (0%)	30	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	172	GLU
2	B	171	ASN

### 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 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	193/190 (102%)	190 (98%)	3 (2%)	58	56
1	C	195/190 (103%)	193 (99%)	2 (1%)	73	74
2	B	13/15 (87%)	10 (77%)	3 (23%)	0	0
2	D	12/15 (80%)	10 (83%)	2 (17%)	2	0
All	All	413/410 (101%)	403 (98%)	10 (2%)	42	39

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	352	LEU
1	A	378	ARG
1	A	380	LEU
2	B	170	ILE
2	B	179	LEU
2	B	182	LEU
1	C	333	SER
1	C	380	LEU
2	D	171	ASN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	179	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	262	ASN
1	A	299	ASN
1	A	374	GLN
1	C	201	GLN
1	C	374	GLN
2	D	185	GLN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands are 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	$\# Z  > 2$
3	EDO	C	501	-	3,3,3	0.18	0	2,2,2	0.34	0
3	EDO	C	502	-	3,3,3	0.14	0	2,2,2	0.08	0

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
3	EDO	C	501	-	-	1/1/1/1	-
3	EDO	C	502	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	501	EDO	O1-C1-C2-O2
3	C	502	EDO	O1-C1-C2-O2

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.

## 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	232/232 (100%)	0.05	0 <a href="#">100</a>   <a href="#">100</a>	14, 32, 48, 62	6 (2%)
1	C	232/232 (100%)	0.01	1 (0%) <a href="#">89</a>   <a href="#">90</a>	16, 31, 49, 60	7 (3%)
2	B	18/18 (100%)	0.91	3 (16%) <a href="#">5</a>   <a href="#">5</a>	30, 37, 77, 77	0
2	D	16/18 (88%)	0.34	0 <a href="#">100</a>   <a href="#">100</a>	26, 35, 45, 61	0
All	All	498/500 (99%)	0.07	4 (0%) <a href="#">82</a>   <a href="#">84</a>	14, 32, 49, 77	13 (2%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	169	LEU	4.5
2	B	170	ILE	4.2
2	B	171	ASN	2.3
1	C	418	ILE	2.2

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	C	502	4/4	0.77	0.19	66,68,71,72	0
3	EDO	C	501	4/4	0.80	0.15	46,59,59,63	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.