



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 07:01 pm BST

PDB ID : 3ML4  
Title : Crystal structure of a complex between Dok7 PH-PTB and the MuSK jux-  
tamembrane region  
Authors : Bergamin, E.; Hubbard, S.R.  
Deposited on : 2010-04-16  
Resolution : 2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

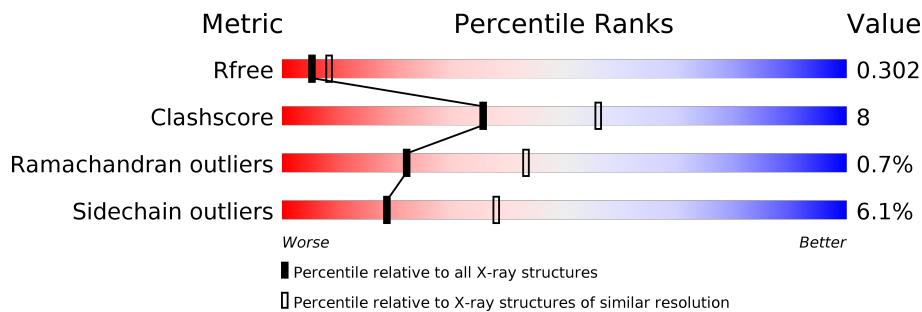
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.



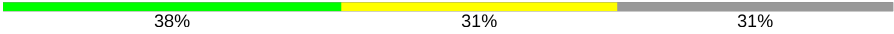
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	224	
1	B	224	
1	C	224	
1	D	224	
2	E	13	
2	F	13	
2	G	13	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	13	 38% 31% 31%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6712 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 Protein Dok-7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	205	Total 1562	C 1003	N 268	O 278	S 8	Se 5	0	0	0
1	B	207	Total 1587	C 1017	N 273	O 283	S 8	Se 6	0	0	0
1	C	208	Total 1588	C 1016	N 274	O 285	S 8	Se 5	0	0	0
1	D	206	Total 1585	C 1014	N 275	O 283	S 8	Se 5	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP Q18PE0
A	-2	SER	-	EXPRESSION TAG	UNP Q18PE0
A	-1	GLU	-	EXPRESSION TAG	UNP Q18PE0
A	0	PHE	-	EXPRESSION TAG	UNP Q18PE0
B	-3	GLY	-	EXPRESSION TAG	UNP Q18PE0
B	-2	SER	-	EXPRESSION TAG	UNP Q18PE0
B	-1	GLU	-	EXPRESSION TAG	UNP Q18PE0
B	0	PHE	-	EXPRESSION TAG	UNP Q18PE0
C	-3	GLY	-	EXPRESSION TAG	UNP Q18PE0
C	-2	SER	-	EXPRESSION TAG	UNP Q18PE0
C	-1	GLU	-	EXPRESSION TAG	UNP Q18PE0
C	0	PHE	-	EXPRESSION TAG	UNP Q18PE0
D	-3	GLY	-	EXPRESSION TAG	UNP Q18PE0
D	-2	SER	-	EXPRESSION TAG	UNP Q18PE0
D	-1	GLU	-	EXPRESSION TAG	UNP Q18PE0
D	0	PHE	-	EXPRESSION TAG	UNP Q18PE0

- Molecule 2 is a protein called Muscle, skeletal receptor tyrosine-protein kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
2	E	9	Total	C	N	O	P	S	0	0	0
			80	49	15	14	1	1			
2	F	9	Total	C	N	O	P	S	0	0	0
			80	49	15	14	1	1			
2	G	9	Total	C	N	O	P	S	0	0	0
			80	49	15	14	1	1			
2	H	9	Total	C	N	O	P	S	0	0	0
			80	49	15	14	1	1			

- Molecule 3 is water.

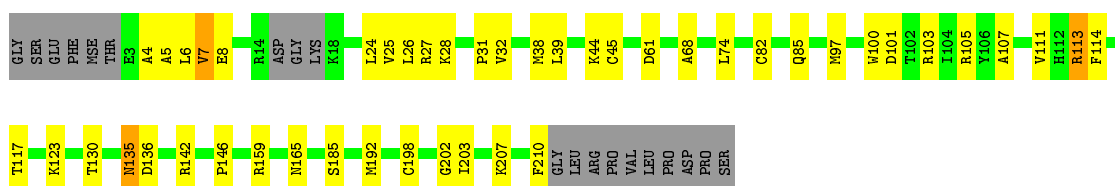
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	13	Total	O	0	0
			13	13		
3	B	22	Total	O	0	0
			22	22		
3	C	16	Total	O	0	0
			16	16		
3	D	16	Total	O	0	0
			16	16		
3	F	3	Total	O	0	0
			3	3		

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

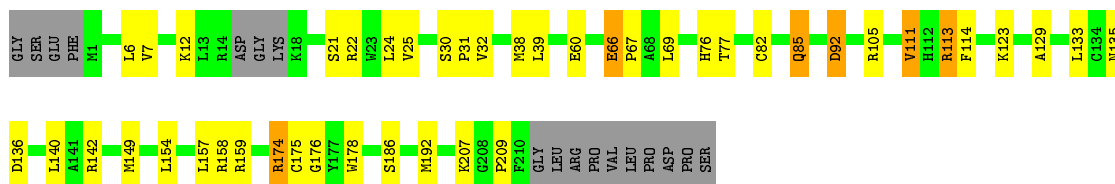
- Molecule 1: Protein Dok-7

Chain A: 



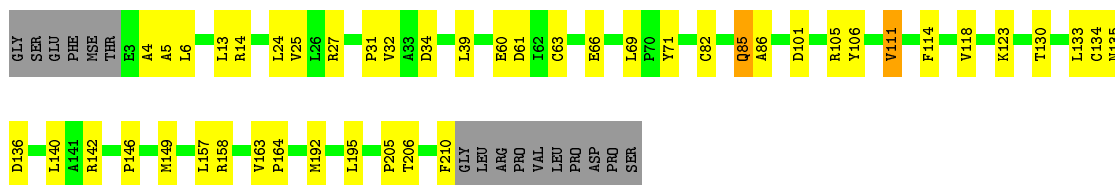
- Molecule 1: Protein Dok-7

Chain B: 



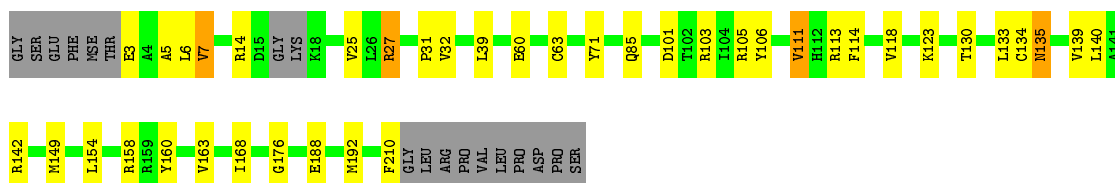
- Molecule 1: Protein Dok-7

Chain C: 



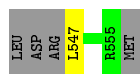
- Molecule 1: Protein Dok-7

Chain D: 



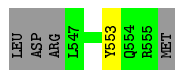
- Molecule 2: Muscle, skeletal receptor tyrosine-protein kinase

Chain E:  62% 8% 31%



- Molecule 2: Muscle, skeletal receptor tyrosine-protein kinase

Chain F:  62% 8% 31%




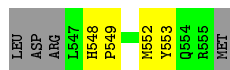
- Molecule 2: Muscle, skeletal receptor tyrosine-protein kinase

Chain G:  62% 8% 31%



- Molecule 2: Muscle, skeletal receptor tyrosine-protein kinase

Chain H:  38% 31% 31%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.60Å 134.60Å 121.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.60 45.11 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-2.60) 100.0 (45.11-2.60)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.29 (at 2.61Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.253 , 0.302 0.252 , 0.302	Depositor DCC
$R_{free}$ test set	1982 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtrriage
Anisotropy	0.186	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 12.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.269 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6712	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.36% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.36	0/1595	0.54	0/2158
1	B	0.38	0/1620	0.56	0/2190
1	C	0.38	0/1622	0.55	0/2195
1	D	0.38	0/1618	0.58	0/2187
2	E	0.38	0/65	0.48	0/86
2	F	0.35	0/65	0.53	0/86
2	G	0.37	0/65	0.53	0/86
2	H	0.34	0/65	0.58	0/86
All	All	0.37	0/6715	0.55	0/9074

There are no bond length outliers.

There are no bond angle outliers.

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	1562	0	1533	26	0
1	B	1587	0	1567	32	0
1	C	1588	0	1556	29	0
1	D	1585	0	1561	26	0
2	E	80	0	68	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	80	0	68	2	0
2	G	80	0	68	1	0
2	H	80	0	68	3	0
3	A	13	0	0	0	0
3	B	22	0	0	4	0
3	C	16	0	0	2	0
3	D	16	0	0	0	0
3	F	3	0	0	0	0
All	All	6712	0	6489	106	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 8.

The worst 5 of 106 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:B:69:LEU:HG	3:B:232:HOH:O	1.57	1.03
1:A:101:ASP:O	1:A:105:ARG:HG3	1.66	0.95
1:D:101:ASP:O	1:D:105:ARG:HG3	1.69	0.91
1:C:101:ASP:O	1:C:105:ARG:HG3	1.76	0.85
1:D:114:PHE:HB3	1:D:192:MSE:HE2	1.65	0.77

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	201/224 (90%)	190 (94%)	10 (5%)	1 (0%)	29 52
1	B	203/224 (91%)	195 (96%)	7 (3%)	1 (0%)	29 52
1	C	206/224 (92%)	195 (95%)	9 (4%)	2 (1%)	15 32

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	202/224 (90%)	195 (96%)	5 (2%)	2 (1%)	15	32
2	E	6/13 (46%)	5 (83%)	1 (17%)	0	100	100
2	F	6/13 (46%)	6 (100%)	0	0	100	100
2	G	6/13 (46%)	6 (100%)	0	0	100	100
2	H	6/13 (46%)	5 (83%)	1 (17%)	0	100	100
All	All	836/948 (88%)	797 (95%)	33 (4%)	6 (1%)	22	43

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	176	GLY
1	C	135	ASN
1	D	135	ASN
1	A	135	ASN
1	C	34	ASP

### 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	162/181 (90%)	152 (94%)	10 (6%)	18	37
1	B	166/181 (92%)	152 (92%)	14 (8%)	11	21
1	C	165/181 (91%)	157 (95%)	8 (5%)	25	49
1	D	166/181 (92%)	156 (94%)	10 (6%)	19	39
2	E	7/12 (58%)	7 (100%)	0	100	100
2	F	7/12 (58%)	7 (100%)	0	100	100
2	G	7/12 (58%)	7 (100%)	0	100	100
2	H	7/12 (58%)	7 (100%)	0	100	100
All	All	687/772 (89%)	645 (94%)	42 (6%)	18	38

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	136	ASP
1	B	175	CYS
1	D	123	LYS
1	B	154	LEU
1	B	159	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	85	GLN
2	G	548	HIS
1	C	165	ASN
1	B	94	HIS
1	D	94	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 non-standard protein/DNA/RNA residues 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
2	PTR	H	553	2	15,16,17	1.95	1 (6%)	19,22,24	0.54	0
2	PTR	F	553	2	15,16,17	1.98	1 (6%)	19,22,24	0.58	0
2	PTR	G	553	2	15,16,17	1.97	1 (6%)	19,22,24	0.62	0
2	PTR	E	553	2	15,16,17	1.94	1 (6%)	19,22,24	0.57	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
2	PTR	H	553	2	-	1/10/11/13	0/1/1/1
2	PTR	F	553	2	-	0/10/11/13	0/1/1/1
2	PTR	G	553	2	-	0/10/11/13	0/1/1/1
2	PTR	E	553	2	-	0/10/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	553	PTR	OH-CZ	-7.48	1.23	1.40
2	H	553	PTR	OH-CZ	-7.25	1.24	1.40
2	G	553	PTR	OH-CZ	-7.23	1.24	1.40
2	E	553	PTR	OH-CZ	-7.16	1.24	1.40

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	H	553	PTR	CZ-OH-P-O3P

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	553	PTR	1	0
2	F	553	PTR	2	0
2	G	553	PTR	1	0

## 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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.