



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2023 – 01:16 PM JST

PDB ID : 7FI8  
Title : Crystal structure of human MICA mutants in complex with natural killer cell receptor NKG2D  
Authors : Cai, W.; Peng, S.; Xu, T.; Tian, Y.; Li, Y.; Liu, J.  
Deposited on : 2021-07-30  
Resolution : 2.80 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

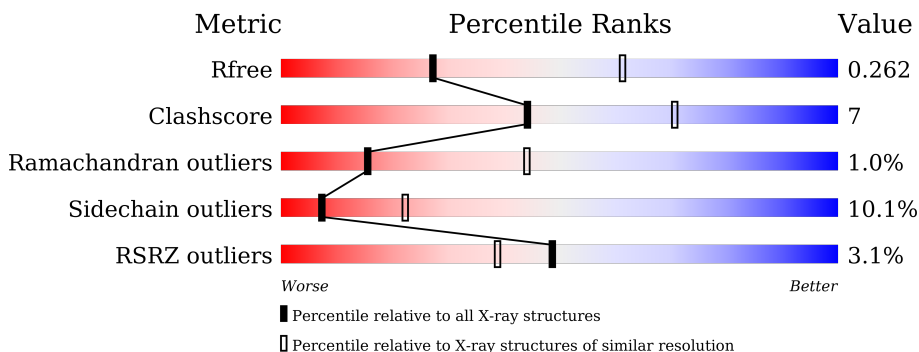
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	139	 5% 71% 17% • 12%
1	B	139	 % 73% 16% • 10%
2	C	275	 3% 68% 24% • 5%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 4150 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 NKG2-D type II integral membrane protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	123	996	633	160	191	12	0	0	0
1	B	125	1011	643	162	194	12	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	78	MET	-	initiating methionine	UNP P26718
A	79	GLU	-	expression tag	UNP P26718
B	78	MET	-	initiating methionine	UNP P26718
B	79	GLU	-	expression tag	UNP P26718

- Molecule 2 is a protein called MHC class I polypeptide-related sequence A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	262	2132	1321	393	403	15	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP Q29983
C	13	ILE	SER	engineered mutation	UNP Q29983
C	108	ILE	GLN	engineered mutation	UNP Q29983
C	120	ASN	GLN	engineered mutation	UNP Q29983
C	146	TRP	LEU	engineered mutation	UNP Q29983
C	157	TRP	TYR	engineered mutation	UNP Q29983

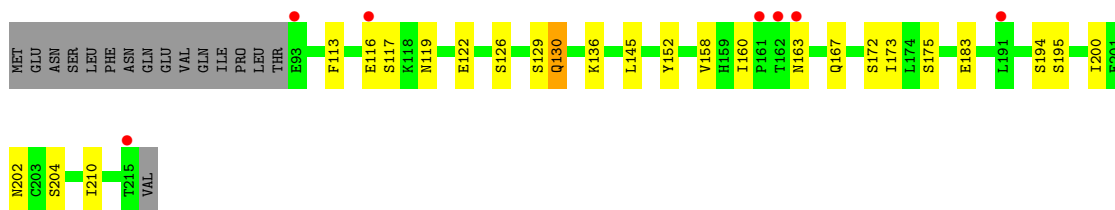
- Molecule 3 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
3	B	1	Total O 1 1	0	0
3	C	10	Total O 10 10	0	0

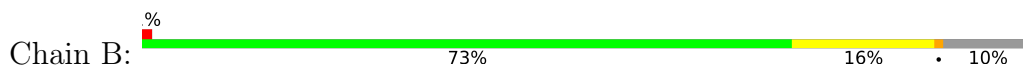
### 3 Residue-property plots [i](#)

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.

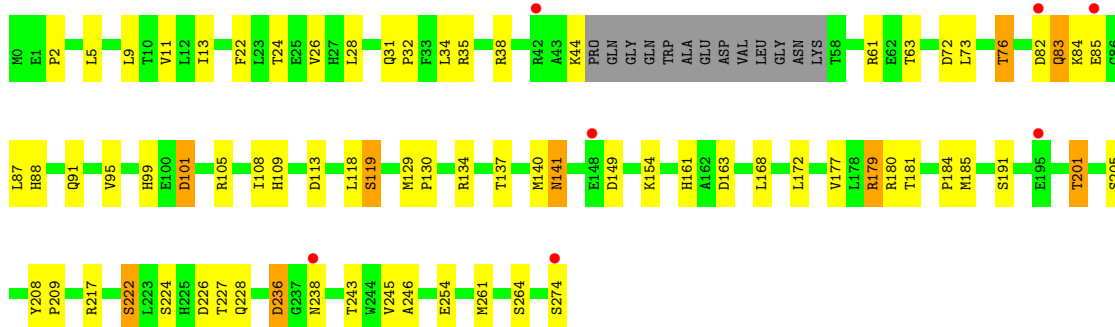
- Molecule 1: NKG2-D type II integral membrane protein



- Molecule 1: NKG2-D type II integral membrane protein



- Molecule 2: MHC class I polypeptide-related sequence A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.48Å 121.48Å 102.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.97 – 2.80 19.86 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.97-2.80) 99.9 (19.86-2.80)	Depositor EDS
$R_{merge}$	0.25	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.84 (at 2.79Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.204 , 0.267 0.207 , 0.262	Depositor DCC
$R_{free}$ test set	948 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.0	Xtrriage
Anisotropy	0.517	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 41.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4150	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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](#)

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	0.72	0/1024	0.87	0/1388
1	B	0.72	0/1039	0.85	0/1409
2	C	0.73	0/2184	0.97	0/2957
All	All	0.72	0/4247	0.92	0/5754

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	996	0	933	8	0
1	B	1011	0	951	9	1
2	C	2132	0	2032	36	0
3	B	1	0	0	0	0
3	C	10	0	0	2	0
All	All	4150	0	3916	53	1

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 7.

All (53) 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)
2:C:236:ASP:HB3	3:C:301:HOH:O	1.35	1.23
1:A:136:LYS:NZ	1:A:172:SER:OG	2.22	0.72
2:C:109:HIS:CD2	2:C:119:SER:HB3	2.32	0.64
1:B:100:PRO:HD3	1:B:213:GLN:OE1	1.99	0.63
1:A:113:PHE:CD1	1:A:210:ILE:HG12	2.35	0.62
2:C:154:LYS:HA	2:C:154:LYS:HE2	1.80	0.62
2:C:101:ASP:N	2:C:101:ASP:OD1	2.34	0.61
2:C:5:LEU:HD13	2:C:28:LEU:HD13	1.81	0.61
2:C:73:LEU:HA	2:C:76:THR:HG23	1.84	0.59
1:A:183:GLU:HA	1:A:200:ILE:HD12	1.84	0.59
2:C:154:LYS:HA	2:C:154:LYS:CE	2.32	0.59
2:C:201:THR:HB	2:C:246:ALA:HB2	1.85	0.58
2:C:181:THR:HG21	2:C:264:SER:CB	2.33	0.57
1:A:113:PHE:CE1	1:A:210:ILE:HG12	2.44	0.53
2:C:88:HIS:HA	2:C:113:ASP:OD1	2.10	0.52
2:C:228:GLN:NE2	2:C:245:VAL:HG21	2.25	0.52
2:C:2:PRO:HB3	2:C:99:HIS:CD2	2.46	0.50
1:A:126:SER:O	1:A:130:GLN:NE2	2.46	0.49
2:C:141:ASN:HD22	2:C:141:ASN:N	2.11	0.48
2:C:217:ARG:HG2	2:C:222:SER:HA	1.94	0.48
2:C:236:ASP:CB	3:C:301:HOH:O	2.19	0.48
2:C:28:LEU:HG	2:C:177:VAL:HG22	1.96	0.48
2:C:184:PRO:HG2	2:C:261:MET:CE	2.45	0.47
2:C:208:TYR:CG	2:C:209:PRO:HA	2.51	0.46
2:C:35:ARG:CD	2:C:44:LYS:HG3	2.46	0.46
2:C:72:ASP:O	2:C:76:THR:HG22	2.15	0.46
2:C:26:VAL:HB	2:C:34:LEU:HB2	1.97	0.46
2:C:9:LEU:O	2:C:91:GLN:HA	2.15	0.45
2:C:168:LEU:O	2:C:172:LEU:HG	2.17	0.45
2:C:31:GLN:HA	2:C:32:PRO:HD2	1.80	0.45
1:B:98:PRO:HD2	1:B:215:THR:HB	1.99	0.45
2:C:181:THR:HG21	2:C:264:SER:HB2	1.98	0.44
2:C:11:VAL:HG12	2:C:22:PHE:HB3	1.99	0.43
2:C:108:ILE:HD12	2:C:161:HIS:CE1	2.53	0.43
2:C:9:LEU:HD13	2:C:24:THR:OG1	2.19	0.43
1:B:155:MET:SD	1:B:192:TYR:HB2	2.59	0.42
2:C:82:ASP:O	2:C:84:LYS:N	2.53	0.42
2:C:83:GLN:NE2	2:C:83:GLN:HA	2.35	0.42
1:B:152:TYR:CD2	1:B:193:ALA:HB2	2.55	0.42
1:A:202:ASN:OD1	1:A:204:SER:OG	2.38	0.42
1:B:160:ILE:HD11	1:B:167:GLN:CD	2.39	0.41
1:A:194:SER:HA	1:A:195:SER:HA	1.86	0.41

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:119:ASN:HD21	1:A:122:GLU:HG3	1.84	0.41
2:C:118:LEU:HD12	2:C:118:LEU:HA	1.81	0.41
1:B:116:GLU:HA	1:B:116:GLU:OE1	2.20	0.41
2:C:228:GLN:HE21	2:C:245:VAL:CG2	2.34	0.41
2:C:13:ILE:O	2:C:87:LEU:HA	2.21	0.41
2:C:35:ARG:HD2	2:C:44:LYS:HG3	2.02	0.41
1:B:160:ILE:HD11	1:B:167:GLN:CG	2.51	0.41
1:B:170:ASP:OD1	1:B:171:GLY:N	2.54	0.41
1:B:178:LEU:HD23	1:B:178:LEU:HA	1.89	0.41
2:C:137:THR:O	2:C:140:MET:HB3	2.21	0.40
2:C:179:ARG:HD2	2:C:179:ARG:HA	1.84	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:93:GLU:O	1:B:93:GLU:O[8_554]	2.16	0.04

## 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	121/139 (87%)	109 (90%)	12 (10%)	0	100 100
1	B	123/139 (88%)	114 (93%)	8 (6%)	1 (1%)	19 49
2	C	258/275 (94%)	232 (90%)	22 (8%)	4 (2%)	9 31
All	All	502/553 (91%)	455 (91%)	42 (8%)	5 (1%)	15 44

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	83	GLN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	85	GLU
1	B	131	ASN
2	C	149	ASP
2	C	130	PRO

### 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	113/129 (88%)	101 (89%)	12 (11%)	6	20
1	B	115/129 (89%)	107 (93%)	8 (7%)	15	40
2	C	236/246 (96%)	209 (89%)	27 (11%)	5	18
All	All	464/504 (92%)	417 (90%)	47 (10%)	7	22

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	116	GLU
1	A	117	SER
1	A	129	SER
1	A	130	GLN
1	A	145	LEU
1	A	152	TYR
1	A	158	VAL
1	A	160	ILE
1	A	163	ASN
1	A	167	GLN
1	A	173	ILE
1	A	175	SER
1	B	115	ASP
1	B	117	SER
1	B	129	SER
1	B	130	GLN
1	B	131	ASN
1	B	145	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	175	SER
1	B	181	ILE
2	C	38	ARG
2	C	61	ARG
2	C	63	THR
2	C	76	THR
2	C	95	VAL
2	C	101	ASP
2	C	105	ARG
2	C	119	SER
2	C	129	MET
2	C	134	ARG
2	C	141	ASN
2	C	163	ASP
2	C	179	ARG
2	C	180	ARG
2	C	185	MET
2	C	191	SER
2	C	201	THR
2	C	205	SER
2	C	222	SER
2	C	224	SER
2	C	226	ASP
2	C	227	THR
2	C	236	ASP
2	C	238	ASN
2	C	243	THR
2	C	254	GLU
2	C	274	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	102	ASN
1	A	130	GLN
1	A	167	GLN
2	C	83	GLN
2	C	109	HIS
2	C	197	ASN
2	C	228	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 monosaccharides 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	123/139 (88%)	0.23	7 (5%) 23 15	61, 92, 133, 156	0
1	B	125/139 (89%)	-0.02	2 (1%) 72 66	45, 81, 118, 161	0
2	C	262/275 (95%)	-0.33	7 (2%) 54 44	36, 66, 111, 134	0
All	All	510/553 (92%)	-0.12	16 (3%) 49 39	36, 77, 120, 161	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	274	SER	4.3
1	A	215	THR	3.5
1	A	93	GLU	3.0
1	B	93	GLU	2.9
1	A	162	THR	2.8
1	B	162	THR	2.6
2	C	195	GLU	2.6
1	A	161	PRO	2.5
1	A	116	GLU	2.4
2	C	85	GLU	2.3
1	A	163	ASN	2.2
2	C	82	ASP	2.2
2	C	238	ASN	2.1
1	A	191	LEU	2.1
2	C	148	GLU	2.1
2	C	42	ARG	2.0

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.