



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 17, 2021 – 08:18 AM EDT

PDB ID : 1I8I  
Title : CRYSTAL STRUCTURE OF DSFV MR1 IN COMPLEX WITH THE PEPTIDE ANTIGEN OF THE MUTANT EPIDERMAL GROWTH FACTOR RECEPTOR, EGFRVIII, AT ROOM TEMPERATURE  
Authors : Landry, R.C.; Klimowicz, A.C.; Lavictoire, S.J.; Borisova, S.; Kottachchi, D.T.; Lorimer, I.A.; Evans, S.V.  
Deposited on : 2001-03-14  
Resolution : 2.40 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 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 : 2.23.2  
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) : 2.23.2

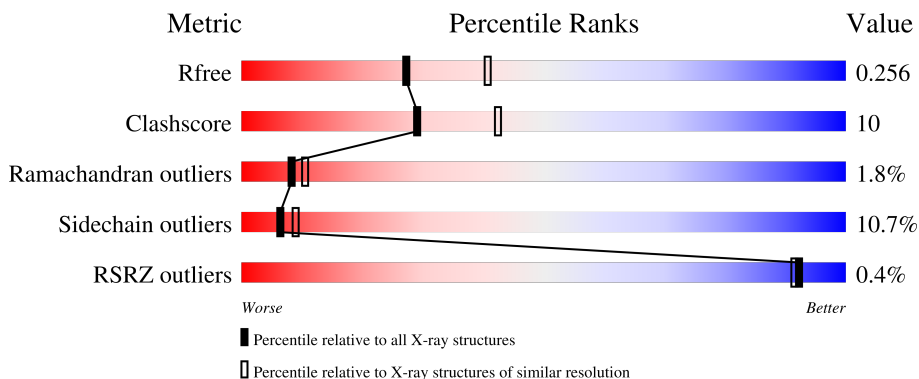
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	107	 67% 24% 7% ..
2	B	124	 % 71% 19% 6% ..
3	C	12	 25% 42% 8% 25%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1896 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 EPIDERMAL GROWTH FACTOR RECEPTOR ANTIBODY MR1SCFV LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	106	813	510	128	170	5	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	CYS	ASP	engineered mutation	UNP Q8R028

- Molecule 2 is a protein called EPIDERMAL GROWTH FACTOR RECEPTOR ANTIBODY MR1SCFV HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	119	925	582	155	182	6	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	344	CYS	ARG	engineered mutation	UNP P18529
B	420	GLY	SER	conflict	UNP P18529
B	421	ILE	SER	conflict	UNP P18529
B	422	GLU	GLY	conflict	UNP P18529
B	424	ARG	GLY	conflict	UNP P18529

- Molecule 3 is a protein called EPIDERMAL GROWTH FACTOR RECEPTOR, EGFRVIII PEPTIDE ANTIGEN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	9	73	45	13	15	0	0	0

- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	29	Total 29	O 29	0	0
4	B	53	Total 53	O 53	0	0
4	C	3	Total 3	O 3	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.

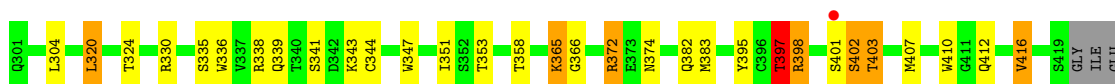
- Molecule 1: EPIDERMAL GROWTH FACTOR RECEPTOR ANTIBODY MR1SCFV LIGHT CHAIN

Chain A:  67% 24% 7% ..




- Molecule 2: EPIDERMAL GROWTH FACTOR RECEPTOR ANTIBODY MR1SCFV HEAVY CHAIN

Chain B:  % 71% 19% 6% ..



GLY  
ARG

- Molecule 3: EPIDERMAL GROWTH FACTOR RECEPTOR, EGFRVIII PEPTIDE ANTI-GEN

Chain C:  25% 42% 8% 25%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.60Å 45.30Å 110.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.40 6.00 – 2.43	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.00-2.40) 93.4 (6.00-2.43)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.78 (at 2.43Å)	Xtrriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.156 , 0.262 0.161 , 0.256	Depositor DCC
$R_{free}$ test set	997 reflections (10.30%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.2	Xtrriage
Anisotropy	0.323	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 72.8	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	0.96	EDS
Total number of atoms	1896	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.95% 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	0.81	0/830	1.57	12/1129 (1.1%)
2	B	0.89	1/944 (0.1%)	1.75	20/1274 (1.6%)
3	C	1.12	0/74	1.94	2/98 (2.0%)
All	All	0.87	1/1848 (0.1%)	1.68	34/2501 (1.4%)

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
3	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	SER	CA-CB	5.51	1.61	1.52

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	372	ARG	NE-CZ-NH2	-11.60	114.50	120.30
2	B	372	ARG	NE-CZ-NH1	11.13	125.86	120.30
2	B	398	ARG	NE-CZ-NH1	10.15	125.37	120.30
2	B	338	ARG	NE-CZ-NH2	-9.91	115.35	120.30
2	B	347	TRP	CD1-CG-CD2	8.59	113.17	106.30
1	A	35	TRP	CD1-CG-CD2	8.52	113.12	106.30
2	B	410	TRP	CD1-CG-CD2	8.01	112.70	106.30
1	A	35	TRP	CE2-CD2-CG	-7.46	101.33	107.30
2	B	410	TRP	CE2-CD2-CG	-6.96	101.73	107.30
2	B	402	SER	N-CA-C	6.93	129.71	111.00
1	A	36	TYR	CB-CG-CD1	-6.73	116.96	121.00
2	B	336	TRP	CE2-CD2-CG	-6.72	101.92	107.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	336	TRP	CD1-CG-CD2	6.66	111.63	106.30
1	A	86	TYR	CB-CG-CD1	-6.53	117.08	121.00
2	B	347	TRP	CG-CD1-NE1	-6.47	103.63	110.10
2	B	403	THR	N-CA-CB	-6.40	98.14	110.30
2	B	398	ARG	NE-CZ-NH2	-6.35	117.13	120.30
1	A	104	LEU	CA-CB-CG	6.33	129.87	115.30
1	A	103	LYS	CA-CB-CG	6.24	127.12	113.40
2	B	347	TRP	CE2-CD2-CG	-6.16	102.37	107.30
2	B	397	THR	N-CA-CB	-6.16	98.60	110.30
2	B	416	VAL	CG1-CB-CG2	6.12	120.69	110.90
1	A	1	ASP	N-CA-C	-6.05	94.66	111.00
1	A	35	TRP	CG-CD2-CE3	5.91	139.22	133.90
2	B	338	ARG	NE-CZ-NH1	5.71	123.15	120.30
1	A	53	THR	OG1-CB-CG2	-5.53	97.28	110.00
2	B	402	SER	CB-CA-C	-5.50	99.66	110.10
2	B	395	TYR	CB-CG-CD2	-5.37	117.78	121.00
1	A	35	TRP	CG-CD1-NE1	-5.24	104.86	110.10
3	C	507	VAL	CG1-CB-CG2	5.22	119.25	110.90
2	B	336	TRP	CG-CD2-CE3	5.20	138.58	133.90
1	A	29	ILE	CA-CB-CG1	-5.17	101.18	111.00
1	A	35	TRP	CB-CG-CD1	-5.07	120.41	127.00
3	C	505	TYR	CA-CB-CG	5.05	123.00	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	505	TYR	Sidechain

## 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	813	0	777	17	0
2	B	925	0	897	16	0
3	C	73	0	66	5	0
4	A	29	0	0	1	0

Continued on next page...



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	53	0	0	0	0
4	C	3	0	0	0	0
All	All	1896	0	1740	34	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 10.

All (34) 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:2:ILE:HD13	1:A:29:ILE:HD11	1.75	0.69
1:A:13:VAL:HG13	1:A:17:GLU:HB2	1.73	0.69
1:A:92:PHE:O	3:C:502:LYS:HB3	1.97	0.65
1:A:38:GLN:HE22	2:B:339:GLN:HE22	1.48	0.61
1:A:39:LYS:O	1:A:42:GLU:HG2	2.02	0.60
2:B:372:ARG:HH11	2:B:374:ASN:HD21	1.51	0.58
2:B:372:ARG:HD3	2:B:374:ASN:ND2	2.19	0.58
1:A:12:SER:HA	1:A:105:GLU:HA	1.86	0.55
2:B:304:LEU:HD22	2:B:324:THR:HG22	1.87	0.55
2:B:351:ILE:HD13	2:B:372:ARG:HG2	1.89	0.54
1:A:46:PHE:CD1	1:A:55:ARG:HD2	2.43	0.54
2:B:397:THR:CG2	2:B:407:MET:HG2	2.38	0.54
2:B:372:ARG:HH11	2:B:374:ASN:ND2	2.06	0.53
1:A:83:VAL:HA	1:A:104:LEU:HD12	1.91	0.52
2:B:351:ILE:HG13	2:B:358:THR:HG22	1.91	0.52
2:B:320:LEU:HD13	2:B:383:MET:HE1	1.93	0.51
2:B:365:LYS:HA	2:B:365:LYS:HZ2	1.77	0.50
1:A:2:ILE:HD13	1:A:29:ILE:CD1	2.41	0.48
1:A:13:VAL:H	1:A:105:GLU:HG3	1.81	0.46
2:B:397:THR:HG21	2:B:407:MET:HG2	1.97	0.46
2:B:365:LYS:HZ2	2:B:366:GLY:H	1.63	0.45
1:A:4:LEU:HG	1:A:25:THR:HG22	1.98	0.45
2:B:335:SER:HB2	2:B:397:THR:HG22	1.99	0.45
2:B:330:ARG:O	2:B:353:THR:HB	2.18	0.44
3:C:502:LYS:HB2	3:C:505:TYR:CZ	2.53	0.44
2:B:365:LYS:NZ	2:B:366:GLY:H	2.16	0.43
1:A:86:TYR:O	1:A:101:GLY:HA2	2.18	0.43
1:A:91:SER:O	3:C:502:LYS:HG2	2.20	0.42
1:A:89:LEU:HD12	1:A:97:THR:O	2.20	0.41
2:B:398:ARG:O	2:B:407:MET:HA	2.20	0.41
1:A:50:GLU:OE1	3:C:502:LYS:HE2	2.19	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:55:ARG:HD3	4:A:653:HOH:O	2.21	0.41
1:A:104:LEU:HD13	1:A:105:GLU:HB2	2.03	0.40
3:C:502:LYS:HB3	3:C:503:GLY:H	1.83	0.40

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	104/107 (97%)	98 (94%)	5 (5%)	1 (1%)	15 23
2	B	117/124 (94%)	110 (94%)	5 (4%)	2 (2%)	9 11
3	C	7/12 (58%)	5 (71%)	1 (14%)	1 (14%)	0 0
All	All	228/243 (94%)	213 (93%)	11 (5%)	4 (2%)	8 10

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	402	SER
1	A	68	GLY
2	B	343	LYS
3	C	504	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	95/96 (99%)	83 (87%)	12 (13%)	4	5
2	B	102/105 (97%)	93 (91%)	9 (9%)	10	15
3	C	8/11 (73%)	7 (88%)	1 (12%)	4	5
All	All	205/212 (97%)	183 (89%)	22 (11%)	6	9

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	LEU
1	A	13	VAL
1	A	27	THR
1	A	29	ILE
1	A	31	ASP
1	A	46	PHE
1	A	55	ARG
1	A	63	SER
1	A	69	THR
1	A	94	VAL
1	A	97	THR
1	A	105	GLU
2	B	320	LEU
2	B	341	SER
2	B	344	CYS
2	B	365	LYS
2	B	382	GLN
2	B	397	THR
2	B	403	THR
2	B	412	GLN
2	B	416	VAL
3	C	506	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	38	GLN
2	B	374	ASN
2	B	382	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	106/107 (99%)	-0.80	0 <a href="#">100</a>   <a href="#">100</a>	14, 31, 52, 65	0
2	B	119/124 (95%)	-1.09	1 (0%) <a href="#">86</a>   <a href="#">84</a>	9, 21, 38, 47	0
3	C	9/12 (75%)	-0.24	0 <a href="#">100</a>   <a href="#">100</a>	11, 20, 54, 55	0
All	All	234/243 (96%)	-0.93	1 (0%) <a href="#">92</a>   <a href="#">91</a>	9, 25, 50, 65	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	401	SER	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.