



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 15, 2021 – 12:03 pm GMT

PDB ID : 7BHE  
Title : DARPin\_D5/Her3 domain 4 complex, monoclinic crystals  
Authors : Mittl, P.R.E.; Radom, F.; Pluckthun, A.  
Deposited on : 2021-01-11  
Resolution : 2.30 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  
Mogul : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

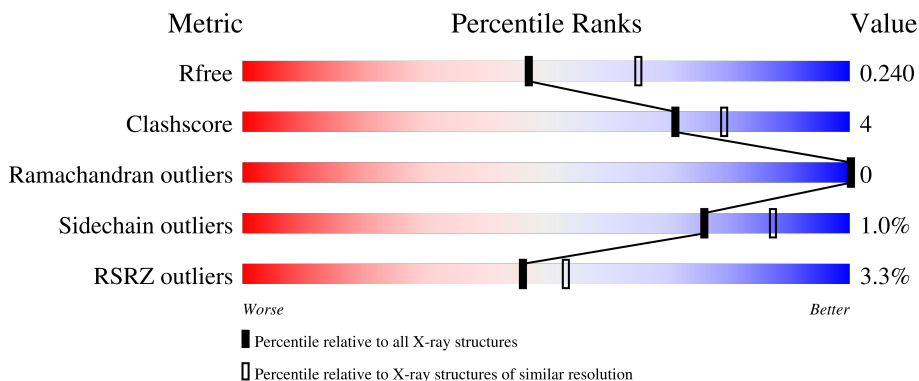
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	138	 5% 79% 12% 9%
1	C	138	 1% 84% 7% 9%
2	B	150	 5% 83% 5% 13%
2	D	150	 6% 81% 6% 13%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	ACT	A	202	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4319 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 DARPin\_D5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	126	950	593	166	190	1	0	0	0
1	C	126	950	593	166	190	1	0	0	0

- Molecule 2 is a protein called Receptor tyrosine-protein kinase erbB-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	131	959	574	176	188	21	0	0	0
2	D	131	967	578	178	190	21	0	1	0

There are 12 discrepancies between the modelled and reference sequences:

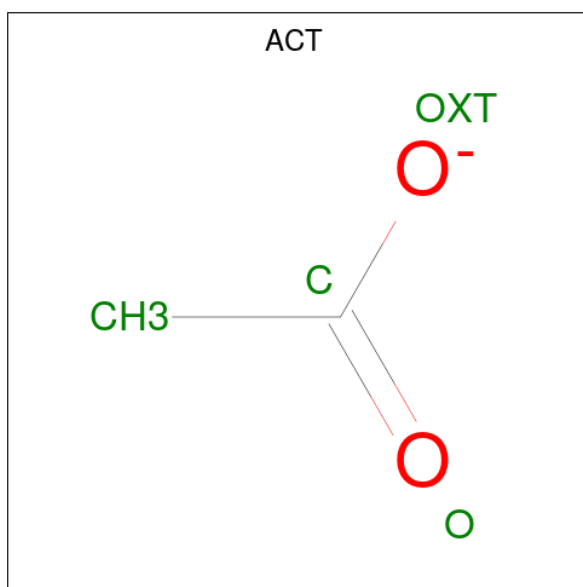
Chain	Residue	Modelled	Actual	Comment	Reference
B	475	HIS	-	expression tag	UNP P21860
B	476	HIS	-	expression tag	UNP P21860
B	477	HIS	-	expression tag	UNP P21860
B	478	HIS	-	expression tag	UNP P21860
B	479	HIS	-	expression tag	UNP P21860
B	480	HIS	-	expression tag	UNP P21860
D	475	HIS	-	expression tag	UNP P21860
D	476	HIS	-	expression tag	UNP P21860
D	477	HIS	-	expression tag	UNP P21860
D	478	HIS	-	expression tag	UNP P21860
D	479	HIS	-	expression tag	UNP P21860
D	480	HIS	-	expression tag	UNP P21860

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2^-$ ).



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

*Continued on next page...*

*Continued from previous page...*

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


- Molecule 5 is water.

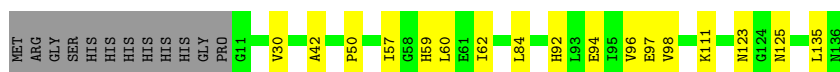
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	106	Total	O	0	0
			106	106		
5	B	130	Total	O	0	0
			130	130		
5	C	102	Total	O	0	0
			102	102		
5	D	125	Total	O	0	0
			125	125		

### 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: DARPin\_D5

Chain A: 




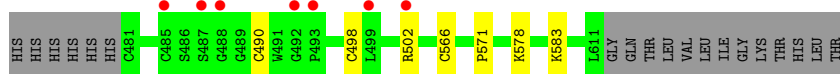
- Molecule 1: DARPin\_D5

Chain C: 




- Molecule 2: Receptor tyrosine-protein kinase erbB-3

Chain B: 



- Molecule 2: Receptor tyrosine-protein kinase erbB-3

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.87Å 62.25Å 74.55Å 90.00° 102.41° 90.00°	Depositor
Resolution (Å)	44.40 – 2.30 44.40 – 2.30	Depositor EDS
% Data completeness (in resolution range)	93.2 (44.40-2.30) 93.2 (44.40-2.30)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.26 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.10.4 (11-DEC-2020)	Depositor
R, $R_{free}$	0.184 , 0.254 0.178 , 0.240	Depositor DCC
$R_{free}$ test set	1185 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.9	Xtriage
Anisotropy	0.703	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4319	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.67% 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: ACT, GOL

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	0/964	0.58	0/1308
1	C	0.43	0/964	0.56	0/1308
2	B	0.39	0/986	0.57	0/1337
2	D	0.42	0/994	0.60	0/1348
All	All	0.42	0/3908	0.57	0/5301

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	950	0	921	13	0
1	C	950	0	921	6	0
2	B	959	0	839	3	0
2	D	967	0	844	4	0
3	A	12	0	16	4	0
3	D	6	0	8	0	0
4	A	4	0	3	3	0
4	B	8	0	6	1	0
5	A	106	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	130	0	0	0	1
5	C	102	0	0	0	1
5	D	125	0	0	0	0
All	All	4319	0	3558	26	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 4.

All (26) 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:111:LYS:NZ	4:A:202:ACT:H1	1.92	0.85
1:C:41:ASN:HD21	1:C:72:ASP:H	1.21	0.83
1:A:111:LYS:HZ3	4:A:202:ACT:H1	1.57	0.67
1:A:111:LYS:HZ1	4:A:202:ACT:H1	1.61	0.66
1:C:60:LEU:HD11	1:C:98:VAL:HG21	1.82	0.61
1:C:79:TRP:HB3	1:C:111:LYS:HE3	1.83	0.59
1:A:97:GLU:HG2	1:A:135:LEU:HD11	1.85	0.59
1:A:94:GLU:H	3:A:201:GOL:H2	1.68	0.58
2:B:571:PRO:HD2	2:B:583:LYS:HB2	1.87	0.57
1:C:41:ASN:ND2	1:C:72:ASP:H	1.99	0.56
2:D:597:ASN:HB2	2:D:607:LEU:HD13	1.89	0.55
1:A:125:ASN:HB2	3:A:203:GOL:H11	1.89	0.54
2:B:490:CYS:HB2	2:B:498:CYS:HA	1.89	0.53
1:A:92:HIS:HA	3:A:201:GOL:H31	1.93	0.51
1:A:42:ALA:O	1:A:50:PRO:HD3	2.10	0.50
2:D:571:PRO:HD2	2:D:583:LYS:HB2	1.93	0.50
2:D:522:ARG:NH1	2:D:540:MET:O	2.45	0.49
1:A:57:ILE:HG23	1:A:59:HIS:CD2	2.49	0.47
1:A:84:LEU:HD11	1:A:96:VAL:HG13	1.97	0.47
1:A:123:ASN:HB3	3:A:203:GOL:H12	1.98	0.45
1:A:60:LEU:HD11	1:A:98:VAL:HG21	1.98	0.45
2:B:566:CYS:SG	4:B:701:ACT:H3	2.56	0.45
2:D:557:CYS:SG	2:D:561:ARG:HG3	2.57	0.44
1:A:30:VAL:HG11	1:A:62:ILE:HG23	2.00	0.44
1:C:120:SER:HB3	1:C:128:LEU:HD23	2.02	0.41
1:C:42:ALA:O	1:C:50:PRO:HD3	2.21	0.41

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)
5:B:918:HOH:O	5:C:243:HOH:O[2_655]	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	124/138 (90%)	124 (100%)	0	0	100	100
1	C	124/138 (90%)	121 (98%)	3 (2%)	0	100	100
2	B	129/150 (86%)	125 (97%)	4 (3%)	0	100	100
2	D	130/150 (87%)	124 (95%)	6 (5%)	0	100	100
All	All	507/576 (88%)	494 (97%)	13 (3%)	0	100	100

There are no Ramachandran outliers to report.

#### 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	97/107 (91%)	97 (100%)	0	100	100
1	C	97/107 (91%)	96 (99%)	1 (1%)	76	87
2	B	108/125 (86%)	106 (98%)	2 (2%)	57	73
2	D	109/125 (87%)	108 (99%)	1 (1%)	78	89
All	All	411/464 (89%)	407 (99%)	4 (1%)	76	87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	502	ARG
2	B	578	LYS
1	C	111	LYS
2	D	600	GLN

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	36	ASN
1	C	36	ASN
1	C	41	ASN
1	C	109	GLN
2	D	556	GLN
2	D	589	ASN
2	D	600	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](#)

6 ligands are modelled in this entry.

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.

## 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	126/138 (91%)	-0.19	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	20, 28, 39, 47	0
1	C	126/138 (91%)	-0.11	1 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">89</span>	21, 29, 40, 53	0
2	B	131/150 (87%)	0.27	7 (5%) <span style="border: 1px solid red; padding: 2px;">26</span> <span style="border: 1px solid red; padding: 2px;">33</span>	20, 33, 71, 77	0
2	D	131/150 (87%)	0.12	9 (6%) <span style="border: 1px solid red; padding: 2px;">16</span> <span style="border: 1px solid red; padding: 2px;">22</span>	22, 34, 59, 65	0
All	All	514/576 (89%)	0.02	17 (3%) <span style="border: 1px solid gray; padding: 2px;">46</span> <span style="border: 1px solid gray; padding: 2px;">53</span>	20, 30, 60, 77	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	487	SER	6.6
2	B	488	GLY	5.2
2	B	499	LEU	3.8
2	B	492	GLY	3.4
1	C	45	HIS	3.2
2	B	502	ARG	3.1
2	B	493	PRO	2.9
2	D	493	PRO	2.8
2	D	499	LEU	2.7
2	D	541	GLU	2.5
2	D	587	VAL	2.5
2	D	488	GLY	2.5
2	B	485	CYS	2.4
2	D	487	SER	2.4
2	D	542	GLY	2.4
2	D	503[A]	ASN	2.2
2	D	543	THR	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(Å <sup>2</sup> )	Q<0.9
4	ACT	B	701	4/4	0.52	0.35	73,73,73,73	0
3	GOL	A	201	6/6	0.64	0.26	46,46,46,47	0
3	GOL	A	203	6/6	0.70	0.33	63,63,63,63	0
4	ACT	A	202	4/4	0.76	0.18	37,37,37,37	0
4	ACT	B	702	4/4	0.79	0.26	50,51,51,52	0
3	GOL	D	701	6/6	0.90	0.21	50,51,51,52	0

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

There are no such residues in this entry.