Full wwPDB X-ray Structure Validation Report (i)

Nov 15, 2021 – 12:02 pm GMT

PDB ID	:	7BHF
Title	:	DARPin_D5/Her3 domain 4 complex, orthorhombic crystals
Authors	:	Mittl, P.R.E.; Radom, F.; Pluckthun, A.
Deposited on	:	2021-01-11
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

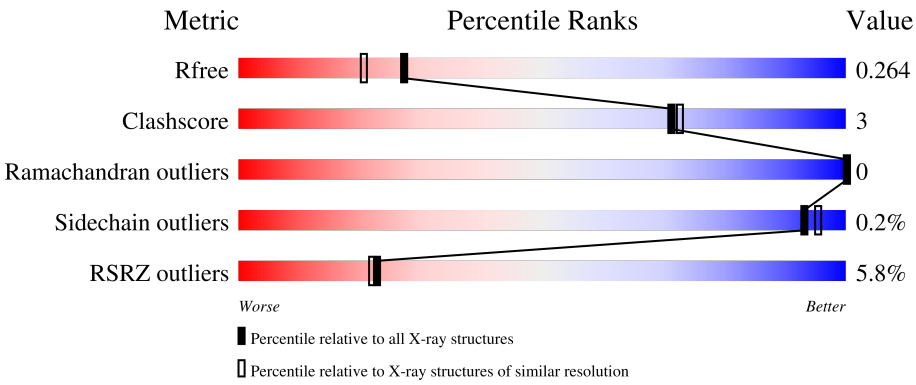
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	% 85% 7% 9%
1	C	138	84% 7% 9%
2	B	150	7% 80% 7% 13%
2	D	150	11% 79% 8% 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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACT	C	201	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4590 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	977	609	171	196	1	0	3	0
1	C	126	967	603	168	195	1	0	2	0

- Molecule 2 is a protein called Isoform 4 of Receptor tyrosine-protein kinase erbB-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	131	981	585	181	194	21	0	3	0
2	D	131	967	578	178	190	21	0	1	0

There are 38 discrepancies between the modelled and reference sequences:

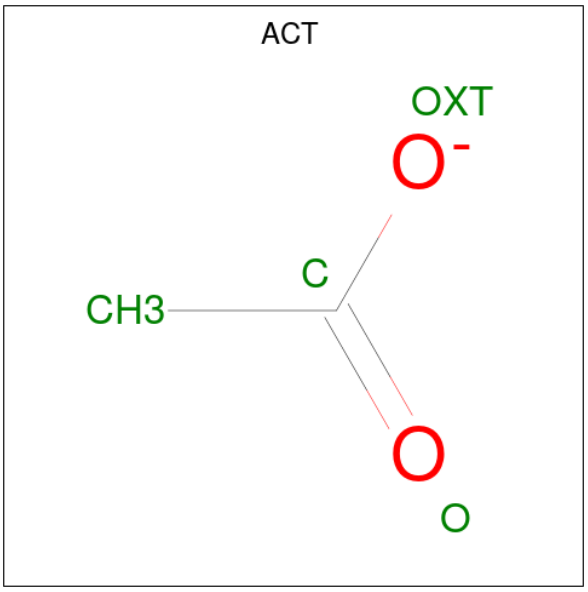
Chain	Residue	Modelled	Actual	Comment	Reference
B	475	HIS	-	expression tag	UNP P21860
B	476	HIS	-	expression tag	UNP P21860
B	477	HIS	-	expression tag	UNP P21860
B	478	HIS	-	expression tag	UNP P21860
B	479	HIS	-	expression tag	UNP P21860
B	480	HIS	-	expression tag	UNP P21860
B	612	GLY	-	expression tag	UNP P21860
B	613	GLN	-	expression tag	UNP P21860
B	614	THR	-	expression tag	UNP P21860
B	615	LEU	-	expression tag	UNP P21860
B	616	VAL	-	expression tag	UNP P21860
B	617	LEU	-	expression tag	UNP P21860
B	618	ILE	-	expression tag	UNP P21860
B	619	GLY	-	expression tag	UNP P21860
B	620	LYS	-	expression tag	UNP P21860
B	621	THR	-	expression tag	UNP P21860

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	622	HIS	-	expression tag	UNP P21860
B	623	LEU	-	expression tag	UNP P21860
B	624	THR	-	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
D	612	GLY	-	expression tag	UNP P21860
D	613	GLN	-	expression tag	UNP P21860
D	614	THR	-	expression tag	UNP P21860
D	615	LEU	-	expression tag	UNP P21860
D	616	VAL	-	expression tag	UNP P21860
D	617	LEU	-	expression tag	UNP P21860
D	618	ILE	-	expression tag	UNP P21860
D	619	GLY	-	expression tag	UNP P21860
D	620	LYS	-	expression tag	UNP P21860
D	621	THR	-	expression tag	UNP P21860
D	622	HIS	-	expression tag	UNP P21860
D	623	LEU	-	expression tag	UNP P21860
D	624	THR	-	expression tag	UNP P21860

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	154	154	154	0	0
			Total	O		
4	B	190	190	190	0	0
			Total	O		
4	C	168	168	168	0	0
			Total	O		
4	D	178	178	178	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: DARPin_D5

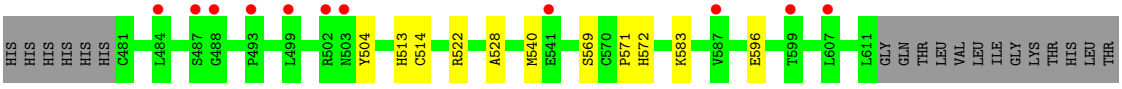
Chain A:

• Molecule 1: DARPin_D5

Chain C:

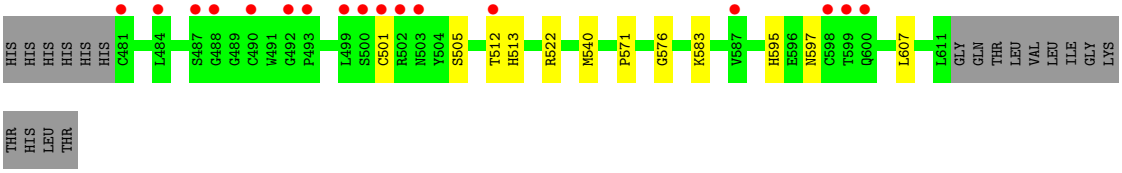
• Molecule 2: Isoform 4 of Receptor tyrosine-protein kinase erbB-3

Chain B:



• Molecule 2: Isoform 4 of Receptor tyrosine-protein kinase erbB-3

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	62.26Å 64.95Å 144.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.19 – 2.00 38.19 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.7 (38.19-2.00) 94.7 (38.19-2.00)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 2.00Å)	Xtriage
Refinement program	BUSTER 2.10.3 (18-SEP-2020)	Depositor
R, R_{free}	0.198 , 0.243 0.211 , 0.264	Depositor DCC
R_{free} test set	1842 reflections (4.77%)	wwPDB-VP
Wilson B-factor (Å^2)	21.4	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4590	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.45 % 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 6.4619e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.47	0/992	0.55	0/1345
1	C	0.46	0/981	0.56	0/1330
2	B	0.39	0/1008	0.58	0/1367
2	D	0.36	0/994	0.55	0/1348
All	All	0.42	0/3975	0.56	0/5390

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	977	0	940	5	0
1	C	967	0	934	8	0
2	B	981	0	853	6	0
2	D	967	0	844	7	0
3	A	4	0	3	0	0
3	C	4	0	3	5	0
4	A	154	0	0	0	1
4	B	190	0	0	1	1
4	C	168	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	178	0	0	0	0
All	All	4590	0	3577	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 3.

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:C:111:LYS:NZ	3:C:201:ACT:H1	1.81	0.95
2:D:512:THR:HG22	2:D:513:HIS:ND1	1.93	0.83
1:C:81:ASN:HD21	3:C:201:ACT:H3	1.58	0.68
1:C:111:LYS:HZ3	3:C:201:ACT:H1	1.59	0.65
1:C:57[B]:ILE:HD12	1:C:59:HIS:CD2	2.33	0.63
2:B:572:HIS:HD2	4:B:769:HOH:O	1.86	0.58
1:C:111:LYS:HZ2	3:C:201:ACT:H1	1.66	0.56
2:D:571:PRO:HD2	2:D:583:LYS:HB2	1.88	0.55
3:C:201:ACT:H2	2:D:576:GLY:O	2.08	0.54
1:A:93:LEU:HG	1:A:131:ILE:HD12	1.91	0.53
2:D:597:ASN:HB2	2:D:607:LEU:HD13	1.92	0.51
2:D:522:ARG:NH1	2:D:540:MET:O	2.45	0.49
1:A:90:VAL:HG23	1:A:92:HIS:CD2	2.47	0.49
1:A:30:VAL:HG11	1:A:62:ILE:HG23	1.94	0.49
2:B:504:TYR:HA	2:B:528:ALA:O	2.15	0.47
2:B:571:PRO:HD2	2:B:583:LYS:HB2	1.98	0.46
2:D:595:HIS:HE1	2:D:597:ASN:HD22	1.64	0.46
2:B:569:SER:HG	2:B:572:HIS:CE1	2.35	0.45
1:C:60:LEU:HD11	1:C:98:VAL:HG21	1.99	0.45
1:A:60:LEU:HD11	1:A:98:VAL:HG21	1.97	0.44
2:B:513:HIS:HD2	2:B:514:CYS:O	2.00	0.44
2:D:501:CYS:SG	2:D:505:SER:HB3	2.60	0.42
2:B:522:ARG:NH1	2:B:540:MET:O	2.54	0.41
1:C:42:ALA:O	1:C:50:PRO:HD3	2.21	0.41
1:A:55:ALA:O	1:A:92:HIS:HD2	2.03	0.41
1:C:106:VAL:HG11	1:C:136:ASN:HB3	2.03	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)
4:A:330:HOH:O	4:B:758:HOH:O[3_555]	2.18	0.02

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	127/138 (92%)	127 (100%)	0	0	100	100
1	C	126/138 (91%)	123 (98%)	3 (2%)	0	100	100
2	B	132/150 (88%)	126 (96%)	6 (4%)	0	100	100
2	D	130/150 (87%)	124 (95%)	6 (5%)	0	100	100
All	All	515/576 (89%)	500 (97%)	15 (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	100/107 (94%)	100 (100%)	0	100	100
1	C	99/107 (92%)	99 (100%)	0	100	100
2	B	111/125 (89%)	110 (99%)	1 (1%)	78	83
2	D	109/125 (87%)	109 (100%)	0	100	100
All	All	419/464 (90%)	418 (100%)	1 (0%)	93	95

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	596	GLU

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	36	ASN
1	A	92	HIS
2	B	597	ASN
1	C	46	ASN
1	C	133	GLN
2	D	597	ASN

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)

2 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, 95th 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^2)	Q<0.9
1	A	126/138 (91%)	-0.08	2 (1%) 72 70	15, 25, 35, 52	0
1	C	126/138 (91%)	-0.12	0 100 100	16, 25, 35, 51	0
2	B	131/150 (87%)	0.31	11 (8%) 11 10	18, 28, 57, 63	0
2	D	131/150 (87%)	0.41	17 (12%) 3 3	17, 29, 68, 72	0
All	All	514/576 (89%)	0.13	30 (5%) 23 22	15, 26, 57, 72	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	488	GLY	5.9
2	B	487	SER	5.5
2	D	502	ARG	4.9
2	B	499	LEU	4.8
2	D	487	SER	4.5
2	D	499	LEU	3.9
2	D	493	PRO	3.7
2	B	599	THR	3.5
2	D	599	THR	3.5
2	D	503	ASN	3.4
2	D	512	THR	3.0
2	B	503[A]	ASN	2.9
2	D	501	CYS	2.8
2	D	500	SER	2.6
2	D	484	LEU	2.6
2	B	541	GLU	2.6
2	B	484	LEU	2.5
2	D	481	CYS	2.5
2	B	502	ARG	2.5
2	B	493	PRO	2.4
2	D	492	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	490	CYS	2.4
2	B	607	LEU	2.3
2	D	488	GLY	2.3
2	D	587	VAL	2.3
1	A	136	ASN	2.2
1	A	45[A]	HIS	2.1
2	D	598	CYS	2.1
2	B	587	VAL	2.0
2	D	600	GLN	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)

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, 95th 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^2)	Q<0.9
3	ACT	A	201	4/4	0.55	0.26	38,39,39,40	0
3	ACT	C	201	4/4	0.58	0.28	39,39,40,40	0

6.5 Other polymers (i)

There are no such residues in this entry.