



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2021 – 01:03 AM EDT

PDB ID : 2QFJ  
Title : Crystal Structure of First Two RRM Domains of FIR Bound to ssDNA from a Portion of FUSE  
Authors : Crichlow, G.V.; Yang, Y.; Fan, C.; Lolis, E.; Braddock, D.  
Deposited on : 2007-06-27  
Resolution : 2.10 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  
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

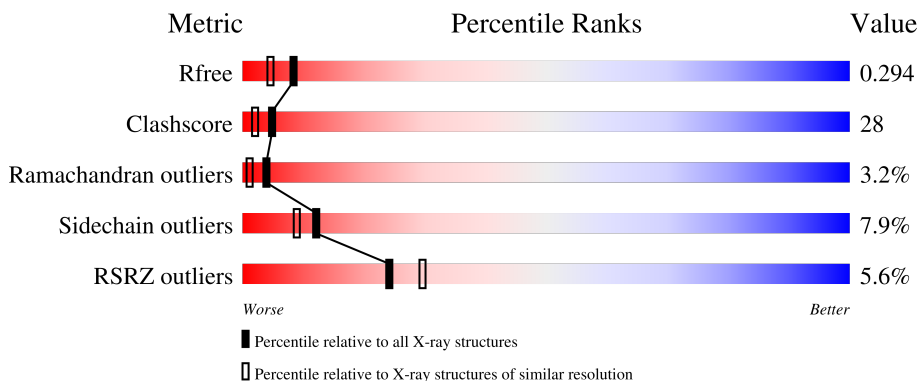
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	C	25	
2	A	216	
2	B	216	

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 3010 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 DNA chain called DNA (5'-D(\*DTP\*DCP\*DGP\*DGP\*DGP\*DAP\*DTP\*DTP\*DTP\*DTP\*DTP\*DAP\*DTP\*DTP\*DTP\*DTP\*DGP\*DTP\*DGP\*DTP\*DTP\*DAP\*DTP\*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	4	41	19	8	12	2	0	0	1

- Molecule 2 is a protein called FBP-interacting repressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	193	1472	932	251	283	6	0	0	0
2	B	194	1475	931	253	284	7	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	84	GLY	-	expression tag	UNP Q9NZA0
A	85	SER	-	expression tag	UNP Q9NZA0
A	86	HIS	-	expression tag	UNP Q9NZA0
A	87	MET	-	expression tag	UNP Q9NZA0
A	88	ALA	-	expression tag	UNP Q9NZA0
A	89	SER	-	expression tag	UNP Q9NZA0
A	90	MET	-	expression tag	UNP Q9NZA0
A	91	THR	-	expression tag	UNP Q9NZA0
A	92	GLY	-	expression tag	UNP Q9NZA0
A	93	GLY	-	expression tag	UNP Q9NZA0
A	94	GLN	-	expression tag	UNP Q9NZA0
A	95	GLN	-	expression tag	UNP Q9NZA0
A	96	MET	-	expression tag	UNP Q9NZA0
A	97	GLY	-	expression tag	UNP Q9NZA0
A	98	ARG	-	expression tag	UNP Q9NZA0
A	99	GLY	-	expression tag	UNP Q9NZA0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	SER	-	expression tag	UNP Q9NZA0
A	106	GLY	ARG	engineered mutation	UNP Q9NZA0
A	112	SER	CYS	engineered mutation	UNP Q9NZA0
A	238	ALA	CYS	engineered mutation	UNP Q9NZA0
B	84	GLY	-	expression tag	UNP Q9NZA0
B	85	SER	-	expression tag	UNP Q9NZA0
B	86	HIS	-	expression tag	UNP Q9NZA0
B	87	MET	-	expression tag	UNP Q9NZA0
B	88	ALA	-	expression tag	UNP Q9NZA0
B	89	SER	-	expression tag	UNP Q9NZA0
B	90	MET	-	expression tag	UNP Q9NZA0
B	91	THR	-	expression tag	UNP Q9NZA0
B	92	GLY	-	expression tag	UNP Q9NZA0
B	93	GLY	-	expression tag	UNP Q9NZA0
B	94	GLN	-	expression tag	UNP Q9NZA0
B	95	GLN	-	expression tag	UNP Q9NZA0
B	96	MET	-	expression tag	UNP Q9NZA0
B	97	GLY	-	expression tag	UNP Q9NZA0
B	98	ARG	-	expression tag	UNP Q9NZA0
B	99	GLY	-	expression tag	UNP Q9NZA0
B	100	SER	-	expression tag	UNP Q9NZA0
B	106	GLY	ARG	engineered mutation	UNP Q9NZA0
B	112	SER	CYS	engineered mutation	UNP Q9NZA0
B	238	ALA	CYS	engineered mutation	UNP Q9NZA0

- Molecule 3 is water.

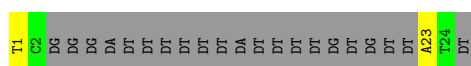
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	13	Total O 13 13	0	0
3	B	9	Total O 9 9	0	0

### 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.

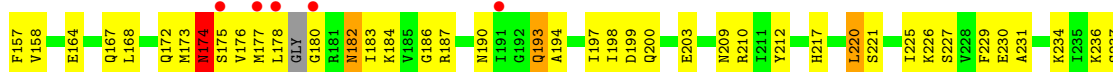
- Molecule 1: DNA (5'-D>(\*DTP\*DCP\*DGP\*DGP\*DGP\*DAP\*DTP\*DTP\*DTP\*DTP\*DTP\*DT\*TP\*DAP\*DTP\*DTP\*DTP\*DTP\*DTP\*DGP\*DTP\*DGP\*DTP\*DTP\*DAP\*DTP\*DT)-3')

Chain C: 



- Molecule 2: FBP-interacting repressor

Chain A: 



- Molecule 2: FBP-interacting repressor

Chain B: 



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.13Å 63.13Å 82.59Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.59 – 2.10 45.59 – 2.10	Depositor EDS
% Data completeness (in resolution range)	91.3 (45.59-2.10) 91.4 (45.59-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.56 (at 2.10Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.258 , 0.295 0.260 , 0.294	Depositor DCC
$R_{free}$ test set	953 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	47.2	Xtrriage
Anisotropy	0.126	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 42.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.038 for -h,-k,l 0.318 for h,-h-k,-l 0.046 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3010	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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	C	1.49	1/44 (2.3%)	1.68	1/66 (1.5%)
2	A	0.59	0/1500	0.86	1/2023 (0.0%)
2	B	0.61	0/1504	0.88	0/2029
All	All	0.62	1/3048 (0.0%)	0.89	2/4118 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	1	DT	O3'-P	-6.74	1.53	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	23	DA	C3'-C2'-C1'	-7.00	94.10	102.50
2	A	177	MET	C-N-CA	5.21	134.73	121.70

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	C	41	0	21	0	0
2	A	1472	0	1426	77	0
2	B	1475	0	1438	91	0
3	A	13	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	9	0	0	1	0
All	All	3010	0	2885	162	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 28.

The worst 5 of 162 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:B:175:SER:HA	2:B:181:ARG:NH1	1.59	1.15
2:B:175:SER:CA	2:B:181:ARG:HH12	1.70	1.04
2:A:294:THR:H	2:A:295:PRO:HD2	1.24	1.01
2:A:176:VAL:HB	2:A:180:GLY:O	1.58	1.01
2:A:176:VAL:HB	2:A:180:GLY:C	1.86	0.94

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
2	A	185/216 (86%)	164 (89%)	16 (9%)	5 (3%)	<b>5</b> <b>1</b>
2	B	192/216 (89%)	167 (87%)	18 (9%)	7 (4%)	<b>3</b> <b>1</b>
All	All	377/432 (87%)	331 (88%)	34 (9%)	12 (3%)	<b>4</b> <b>1</b>

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	292	LEU
2	A	293	LEU
2	B	147	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	B	150	MET
2	B	178	LEU

### 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	
2	A	151/171 (88%)	140 (93%)	11 (7%)	14	11
2	B	151/171 (88%)	138 (91%)	13 (9%)	10	7
All	All	302/342 (88%)	278 (92%)	24 (8%)	12	9

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	159	GLU
2	B	220	LEU
2	B	184	LYS
2	B	239	THR
2	A	184	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	278	GLN
2	B	264	GLN
2	B	278	GLN
2	A	152	HIS
2	A	130	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	C	4/25 (16%)	0.52	0 <b>100</b> <b>100</b>	91, 92, 97, 99	0
2	A	193/216 (89%)	0.15	11 (5%) 23 29	38, 52, 76, 87	4 (2%)
2	B	194/216 (89%)	0.07	11 (5%) 23 29	39, 51, 78, 85	0
All	All	391/457 (85%)	0.11	22 (5%) 24 29	38, 52, 79, 99	4 (1%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	149	THR	6.1
2	A	148	VAL	4.5
2	A	178	LEU	4.1
2	A	191	ILE	4.0
2	A	180	GLY	3.6

### 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.