



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2022 – 01:23 pm BST

PDB ID : 7OI5  
Title : Crystal structure of AP2 Mu2 - FCHO2 chimera (GST cleaved)  
Authors : Zaccai, N.R.; Kelly, B.T.; Evans, P.R.; Owen, D.J.  
Deposited on : 2021-05-11  
Resolution : 2.61 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  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.28.1  
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) : 2.28.1

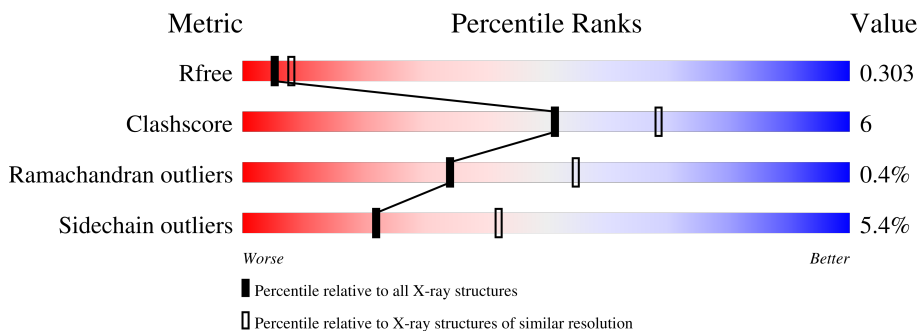
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.61 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	B	359	 64% 16% • 19%
1	D	359	 66% 12% • 22%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 4669 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 AP-2 complex subunit mu,F-BAR domain only protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	292	2362	1512	411	424	15	0	0	0
1	D	281	2273	1459	396	403	15	0	0	0

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	153	PRO	-	expression tag	UNP P84092
B	154	GLY	-	expression tag	UNP P84092
B	155	ILE	-	expression tag	UNP P84092
B	156	HIS	-	expression tag	UNP P84092
B	157	MET	-	expression tag	UNP P84092
B	436	GLY	-	linker	UNP P84092
B	437	ALA	-	linker	UNP P84092
B	438	SER	-	linker	UNP P84092
B	439	GLY	-	linker	UNP P84092
B	440	SER	-	linker	UNP P84092
B	441	ALA	-	linker	UNP P84092
B	442	GLY	-	linker	UNP P84092
B	443	SER	-	linker	UNP P84092
B	444	ALA	-	linker	UNP P84092
B	445	GLY	-	linker	UNP P84092
B	446	PRO	-	linker	UNP P84092
B	447	SER	-	linker	UNP P84092
B	448	GLY	-	linker	UNP P84092
B	449	ALA	-	linker	UNP P84092
B	450	GLY	-	linker	UNP P84092
B	451	SER	-	linker	UNP P84092
B	452	ALA	-	linker	UNP P84092
B	453	GLY	-	linker	UNP P84092
B	454	SER	-	linker	UNP P84092
B	455	ALA	-	linker	UNP P84092

*Continued on next page...*

*Continued from previous page...*

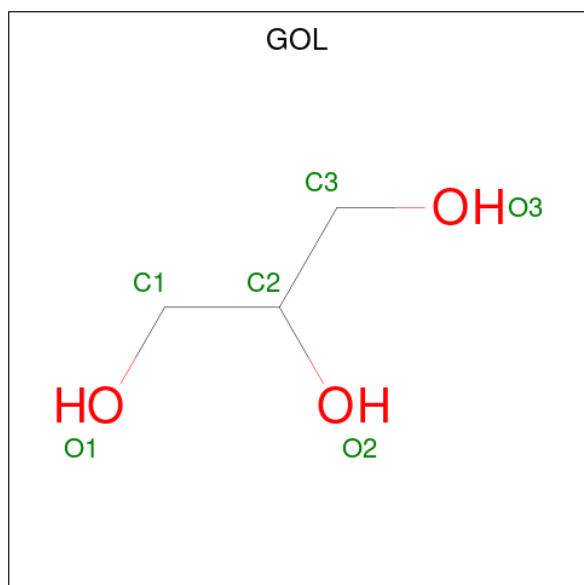
Chain	Residue	Modelled	Actual	Comment	Reference
B	456	GLY	-	linker	UNP P84092
B	457	PRO	-	linker	UNP P84092
B	458	SER	-	linker	UNP P84092
B	459	ALA	-	linker	UNP P84092
B	460	GLY	-	linker	UNP P84092
B	461	SER	-	linker	UNP P84092
B	462	ALA	-	linker	UNP P84092
B	463	GLY	-	linker	UNP P84092
B	464	SER	-	linker	UNP P84092
B	465	ALA	-	linker	UNP P84092
B	466	GLY	-	linker	UNP P84092
B	467	SER	-	linker	UNP P84092
B	468	GLY	-	linker	UNP P84092
B	469	SER	-	linker	UNP P84092
B	470	ALA	-	linker	UNP P84092
B	471	GLY	-	linker	UNP P84092
B	472	SER	-	linker	UNP P84092
B	473	ALA	-	linker	UNP P84092
B	474	PRO	-	linker	UNP P84092
B	475	GLY	-	linker	UNP P84092
D	153	PRO	-	expression tag	UNP P84092
D	154	GLY	-	expression tag	UNP P84092
D	155	ILE	-	expression tag	UNP P84092
D	156	HIS	-	expression tag	UNP P84092
D	157	MET	-	expression tag	UNP P84092
D	436	GLY	-	linker	UNP P84092
D	437	ALA	-	linker	UNP P84092
D	438	SER	-	linker	UNP P84092
D	439	GLY	-	linker	UNP P84092
D	440	SER	-	linker	UNP P84092
D	441	ALA	-	linker	UNP P84092
D	442	GLY	-	linker	UNP P84092
D	443	SER	-	linker	UNP P84092
D	444	ALA	-	linker	UNP P84092
D	445	GLY	-	linker	UNP P84092
D	446	PRO	-	linker	UNP P84092
D	447	SER	-	linker	UNP P84092
D	448	GLY	-	linker	UNP P84092
D	449	ALA	-	linker	UNP P84092
D	450	GLY	-	linker	UNP P84092
D	451	SER	-	linker	UNP P84092
D	452	ALA	-	linker	UNP P84092

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	453	GLY	-	linker	UNP P84092
D	454	SER	-	linker	UNP P84092
D	455	ALA	-	linker	UNP P84092
D	456	GLY	-	linker	UNP P84092
D	457	PRO	-	linker	UNP P84092
D	458	SER	-	linker	UNP P84092
D	459	ALA	-	linker	UNP P84092
D	460	GLY	-	linker	UNP P84092
D	461	SER	-	linker	UNP P84092
D	462	ALA	-	linker	UNP P84092
D	463	GLY	-	linker	UNP P84092
D	464	SER	-	linker	UNP P84092
D	465	ALA	-	linker	UNP P84092
D	466	GLY	-	linker	UNP P84092
D	467	SER	-	linker	UNP P84092
D	468	GLY	-	linker	UNP P84092
D	469	SER	-	linker	UNP P84092
D	470	ALA	-	linker	UNP P84092
D	471	GLY	-	linker	UNP P84092
D	472	SER	-	linker	UNP P84092
D	473	ALA	-	linker	UNP P84092
D	474	PRO	-	linker	UNP P84092
D	475	GLY	-	linker	UNP P84092

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	12	Total	O	0	0
			12	12		
3	D	4	Total	O	0	0
			4	4		



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.05Å 55.38Å 167.74Å 90.00° 109.05° 90.00°	Depositor
Resolution (Å)	79.27 – 2.61 79.27 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.5 (79.27-2.61) 99.7 (79.27-2.61)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 2.62Å)	Xtrriage
Refinement program	REFMAC 5.8.0257, PHENIX 1.19rc1_4016	Depositor
R, $R_{free}$	0.238 , 0.297 0.240 , 0.303	Depositor DCC
$R_{free}$ test set	1560 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.5	Xtrriage
Anisotropy	0.914	Xtrriage
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>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	4669	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.48% 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: 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	B	0.26	0/2410	0.54	0/3239
1	D	0.32	0/2321	0.58	0/3119
All	All	0.29	0/4731	0.56	0/6358

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
1	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	253	ARG	Sidechain

### 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	B	2362	0	2424	34	0
1	D	2273	0	2339	23	0
2	B	12	0	16	0	0
2	D	6	0	8	1	0
3	B	12	0	0	1	0
3	D	4	0	0	0	0
All	All	4669	0	4787	53	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 6.

The worst 5 of 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)
1:D:377:ASP:O	1:D:379:LYS:NZ	2.25	0.70
1:B:432:GLU:OE1	1:D:434:ARG:NH2	2.27	0.68
1:B:156:HIS:CD2	1:B:158:GLN:HE21	2.15	0.64
1:B:480:GLU:HG3	1:B:481:GLU:H	1.63	0.63
1:D:376:ASN:O	1:D:378:LYS:NZ	2.27	0.60

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	B	284/359 (79%)	263 (93%)	20 (7%)	1 (0%)	34	55
1	D	273/359 (76%)	257 (94%)	15 (6%)	1 (0%)	34	55
All	All	557/718 (78%)	520 (93%)	35 (6%)	2 (0%)	34	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	480	GLU
1	D	291	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	B	266/304 (88%)	251 (94%)	15 (6%)	21	40
1	D	255/304 (84%)	242 (95%)	13 (5%)	24	45
All	All	521/608 (86%)	493 (95%)	28 (5%)	22	42

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	495	LYS
1	D	480	GLU
1	D	193	SER
1	D	380	LYS
1	D	190	GLN

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

Mol	Chain	Res	Type
1	B	156	HIS
1	B	239	GLN
1	D	389	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](#)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GOL	B	602	-	5,5,5	0.94	0	5,5,5	0.92	0
2	GOL	D	601	-	5,5,5	1.02	0	5,5,5	0.97	0
2	GOL	B	601	-	5,5,5	0.93	0	5,5,5	0.96	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	602	-	-	2/4/4/4	-
2	GOL	D	601	-	-	4/4/4/4	-
2	GOL	B	601	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	602	GOL	C1-C2-C3-O3
2	B	602	GOL	O2-C2-C3-O3
2	D	601	GOL	O1-C1-C2-C3
2	D	601	GOL	C1-C2-C3-O3
2	D	601	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	601	GOL	1	0

## 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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.