W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 01:00 am BST

PDB ID	:	2QFI
Title	:	Structure of the zinc transporter YiiP
Authors	:	Lu, M.
Deposited on	:	2007-06-27
Resolution	:	3.80 Å(reported)

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

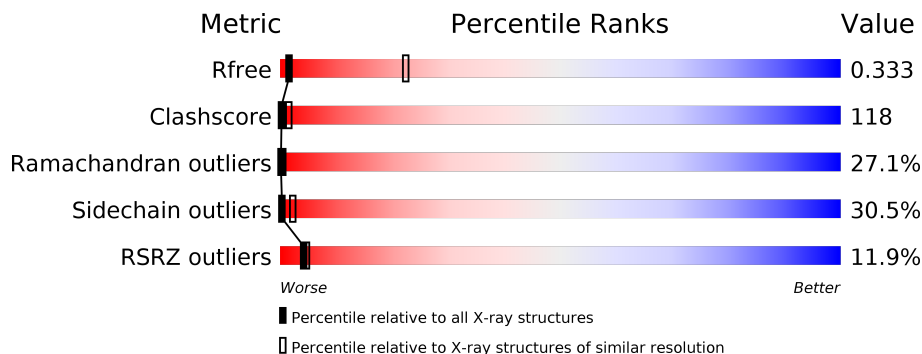|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.80 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	Percentile Ranks	Value
Rfree		0.333
Clashscore		118
Ramachandran outliers		27.1%
Sidechain outliers		30.5%
RSRZ outliers		11.9%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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 on 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	300	5% 12% 38% 33% 12% 5%
1	B	300	17% 15% 39% 29% 12% 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4417 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 Ferrous-iron efflux pump fieF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total 2205	C 1419	N 378	O 397	S 11	0	0	0
1	B	286	Total 2205	C 1419	N 378	O 397	S 11	0	0	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

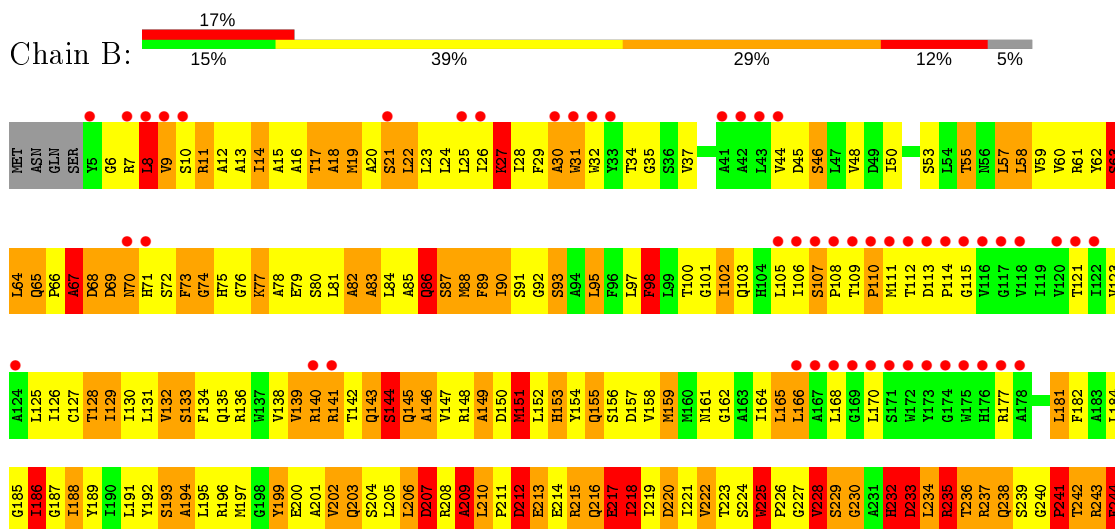| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total 3	Zn 3	0	0
2	A	4	Total 4	Zn 4	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: Ferrous-iron efflux pump fieF

- Molecule 1: Ferrous-iron efflux pump fieF

I245 Q246 I247 H248 L249 E250 M251 E252 D253 S254 L255 P256 L257 V258 Q259 A260 H261 M262 V263 A264 D265 Q266 V267 E268 Q269 A270 I271 L272 R273 R274 F275 P276 G277 S278 D279 V280 I281 I282 H283 Q284 D285 P286 C287 S288 V289 V290 PRO ARG GLU GLY LYS ARG SER MET LEU SER

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	106.66Å 110.84Å 130.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 3.80 39.29 – 3.79	Depositor EDS
% Data completeness (in resolution range)	100.0 (12.00-3.80) 97.4 (39.29-3.79)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 3.76Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.322 , 0.329 0.348 , 0.333	Depositor DCC
R_{free} test set	772 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	118.0	Xtriage
Anisotropy	0.553	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.16 , 155.3	EDS
L-test for twinning[2]	$< |L| > = 0.40, < L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.086 for k,h,-l	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	4417	wwPDB-VP
Average B, all atoms (Å^2)	200.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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: ZN

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	1.41	25/2252 (1.1%)	1.39	37/3069 (1.2%)
1	B	1.07	9/2252 (0.4%)	1.22	18/3069 (0.6%)
All	All	1.25	34/4504 (0.8%)	1.31	55/6138 (0.9%)

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	A	0	24
1	B	0	23
All	All	0	47

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	245	ILE	CA-CB	8.52	1.74	1.54
1	A	279	ASP	CB-CG	7.99	1.68	1.51
1	A	247	ILE	CA-CB	7.64	1.72	1.54
1	B	247	ILE	CA-CB	7.59	1.72	1.54
1	A	268	GLU	CG-CD	7.35	1.62	1.51

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	233	ASP	N-CA-C	8.50	133.95	111.00
1	B	279	ASP	CB-CG-OD1	7.66	125.19	118.30
1	A	247	ILE	N-CA-CB	7.59	128.26	110.80
1	A	249	LEU	CA-CB-CG	7.50	132.56	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	233	ASP	N-CA-C	7.44	131.09	111.00

There are no chirality outliers.

5 of 47 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	27	LYS	Peptide
1	A	6	GLY	Peptide
1	A	63	SER	Peptide
1	A	76	GLY	Peptide
1	A	97	LEU	Peptide

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	2205	0	2239	542	4
1	B	2205	0	2239	514	0
2	A	4	0	0	0	2
2	B	3	0	0	0	0
All	All	4417	0	4478	1052	4

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 118.

The worst 5 of 1052 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:B:232:HIS:CE1	1:B:245:ILE:HG21	1.36	1.61
1:B:245:ILE:CB	1:B:245:ILE:CA	1.74	1.56
1:B:232:HIS:CE1	1:B:245:ILE:CG2	1.96	1.49
1:B:232:HIS:NE2	1:B:245:ILE:CG2	1.79	1.41
1:B:232:HIS:NE2	1:B:245:ILE:HD12	1.47	1.30

All (4) 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)
1:A:232:HIS:NE2	2:A:306:ZN:ZN[4_556]	1.49	0.71
1:A:68:ASP:N	2:A:303:ZN:ZN[4_556]	1.54	0.66
1:A:261:HIS:CE1	1:A:285:ASP:OD1[4_556]	1.56	0.64
1:A:261:HIS:ND1	1:A:285:ASP:OD1[4_556]	2.09	0.11

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	284/300 (95%)	133 (47%)	76 (27%)	75 (26%)	0	0
1	B	284/300 (95%)	129 (45%)	76 (27%)	79 (28%)	0	0
All	All	568/600 (95%)	262 (46%)	152 (27%)	154 (27%)	0	0

5 of 154 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	ALA
1	A	21	SER
1	A	55	THR
1	A	67	ALA
1	A	69	ASP

5.3.2 Protein sidechains ⓘ

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	236/249 (95%)	163 (69%)	73 (31%)	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	236/249 (95%)	165 (70%)	71 (30%)	0	2
All	All	472/498 (95%)	328 (70%)	144 (30%)	0	2

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	278	SER
1	B	27	LYS
1	B	263	VAL
1	A	281	ILE
1	B	7	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	266	GLN
1	B	65	GLN
1	B	266	GLN
1	A	284	GLN
1	B	70	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 carbohydrates in this entry.

5.6   Ligand geometry (i)

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

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	286/300 (95%)	0.08	16 (5%)	24	20	34, 112, 307, 364	0
1	B	286/300 (95%)	0.77	52 (18%)	1	1	48, 212, 575, 683	0
All	All	572/600 (95%)	0.42	68 (11%)	4	5	34, 151, 522, 683	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	42	ALA	15.7
1	B	43	LEU	12.8
1	B	169	GLY	10.3
1	B	168	LEU	9.8
1	B	167	ALA	9.7

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 carbohydrates 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
2	ZN	A	302	1/1	0.95	0.14	28,28,28,28	0
2	ZN	A	307	1/1	0.96	0.07	109,109,109,109	0
2	ZN	B	301	1/1	0.96	0.23	25,25,25,25	0
2	ZN	B	305	1/1	0.96	0.09	77,77,77,77	0
2	ZN	A	306	1/1	0.97	0.20	28,28,28,28	0
2	ZN	B	304	1/1	0.97	0.14	8,8,8,8	0
2	ZN	A	303	1/1	0.98	0.02	26,26,26,26	0

6.5 Other polymers (i)

There are no such residues in this entry.