wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 06:47 AM EST

PDB ID	:	6XNF
Title	:	GCN4-p1 Peptide Trimer with Tetrafluoroiodophenylalanine residue at position 16 (TFI-F16)
Authors	:	Rowe Hartje, R.K.; Czarny, R.S.; Ho, A.
Deposited on	:	2020-07-02
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

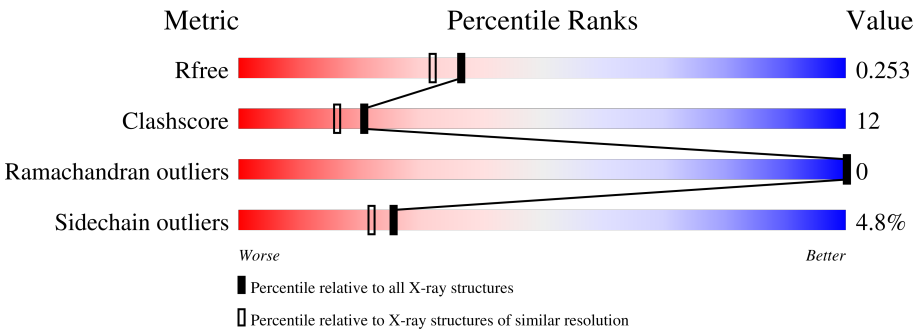
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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)

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%

Mol	Chain	Length	Quality of chain
1	A	30	70% 27% .
2	B	30	87% 13%
2	C	30	73% 27%

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
1	V61	A	16[A]	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 894 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 GCN4-p1 peptide with TFI-F16.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	F	I	N	O	S			
1	A	30	288	184	8	2	46	47	1	0	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	V61	ASN	engineered mutation	UNP P03069

- Molecule 2 is a protein called GCN4-p1 peptide with A16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	30	266	170	47	48	1	0	1	0
			Total	C	N	O	S			
2	C	30	266	170	49	46	1	0	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	16	ALA	ASN	engineered mutation	UNP P03069
C	16	ALA	ASN	engineered mutation	UNP P03069

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
3	A	1	1	1	0	0
			Total	Na		
3	B	2	2	2	0	0
			Total	Na		
3	C	1	1	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	26	26	26	0	0
			Total	O		
4	B	25	25	25	0	0
			Total	O		
4	C	19	19	19	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. 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. 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: GCN4-p1 peptide with TFI-F16

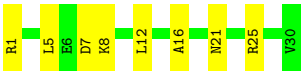
Chain A:

- Molecule 2: GCN4-p1 peptide with A16

Chain B:

- Molecule 2: GCN4-p1 peptide with A16

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	59.65Å 35.19Å 47.02Å 90.00° 100.65° 90.00°	Depositor
Resolution (Å)	46.21 – 2.00 46.21 – 1.38	Depositor EDS
% Data completeness (in resolution range)	95.2 (46.21-2.00) 45.1 (46.21-1.38)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	-0.33 (at 1.37Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.197 , 0.254 0.197 , 0.253	Depositor DCC
R_{free} test set	610 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	15.6	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	894	wwPDB-VP
Average B, all atoms (Å^2)	39.0	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 21.86 % 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 6.4150e-03. 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: NA, V61

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	0.38	0/261	0.49	0/343
2	B	0.41	0/265	0.62	0/349
2	C	0.44	0/273	0.53	0/360
All	All	0.41	0/799	0.55	0/1052

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	288	0	289	15	0
2	B	266	0	292	5	0
2	C	266	0	302	9	0
3	A	1	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
4	A	26	0	0	1	0
4	B	25	0	0	0	0
4	C	19	0	0	2	0
All	All	894	0	883	21	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 12.

The worst 5 of 21 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:5:LEU:HD23	2:C:5:LEU:HD13	1.67	0.75
1:A:12:LEU:O	1:A:16[A]:V61:F14	1.97	0.72
2:C:21:ASN:HB3	2:C:25:ARG:CZ	2.24	0.66
1:A:22:GLU:OE1	1:A:25:ARG:NH2	2.28	0.60
2:C:7:ASP:OD1	4:C:201:HOH:O	2.17	0.58

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	A	28/30 (93%)	28 (100%)	0	0	100	100
2	B	28/30 (93%)	26 (93%)	2 (7%)	0	100	100
2	C	29/30 (97%)	29 (100%)	0	0	100	100
All	All	85/90 (94%)	83 (98%)	2 (2%)	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	30/28 (107%)	25 (83%)	5 (17%)	2 1
2	B	30/28 (107%)	29 (97%)	1 (3%)	38 37
2	C	30/28 (107%)	30 (100%)	0	100 100
All	All	90/84 (107%)	84 (93%)	6 (7%)	25 11

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30[A]	VAL
1	A	30[B]	VAL
2	B	3	LYS
1	A	19[A]	LEU
1	A	3	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	V61	A	16[C]	-	15,16,17	0.96	1 (6%)	19,23,25	2.09	8 (42%)
1	V61	A	16[A]	-	15,16,17	0.82	0	19,23,25	1.33	2 (10%)

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
1	V61	A	16[C]	-	-	3/5/6/8	0/1/1/1
1	V61	A	16[A]	-	-	2/5/6/8	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	16[C]	V61	C03-C04	2.34	1.54	1.51

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	16[C]	V61	C03-C04-C05	-4.66	115.65	122.03
1	A	16[C]	V61	F14-C13-C04	3.90	124.05	119.32
1	A	16[C]	V61	C07-C09-C11	2.98	120.07	117.00
1	A	16[C]	V61	C03-C04-C13	2.69	125.72	122.03
1	A	16[C]	V61	C05-C04-C13	2.31	118.64	115.90

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	16[C]	V61	CA-C03-C04-C13
1	A	16[C]	V61	CA-C03-C04-C05
1	A	16[A]	V61	CA-C03-C04-C13
1	A	16[C]	V61	C04-C03-CA-C
1	A	16[A]	V61	CA-C03-C04-C05

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	16[C]	V61	3	0
1	A	16[A]	V61	7	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 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)

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

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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