wwPDB X-ray Structure Validation Summary Report (i)

Sep 12, 2023 – 06:06 AM EDT

PDB ID	:	4L5T
Title	:	Crystal structure of the tetrameric p202 HIN2
Authors	:	Yin, Q.; Tian, Y.; Wu, H.
Deposited on	:	2013-06-11
Resolution	:	3.40 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:

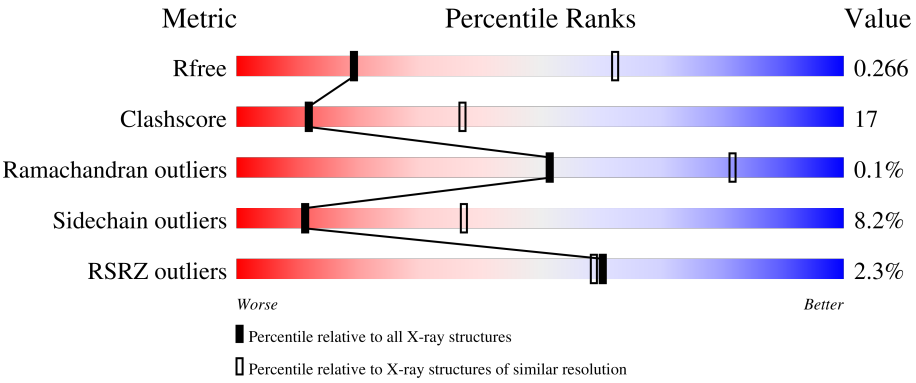
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	203	2% 62% 24% • 11%
1	B	203	% 65% 25% • 9%
1	C	203	56% 31% • 10%
1	D	203	5% 58% 28% 5% 10%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 5748 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 Interferon-activable protein 202.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1414	905	229	271	9			
1	B	185	Total	C	N	O	S	0	0	0
			1463	935	240	277	11			
1	C	182	Total	C	N	O	S	0	0	0
			1438	919	236	272	11			
1	D	183	Total	C	N	O	S	0	0	0
			1433	913	237	272	11			

There are 20 discrepancies between the modelled and reference sequences:

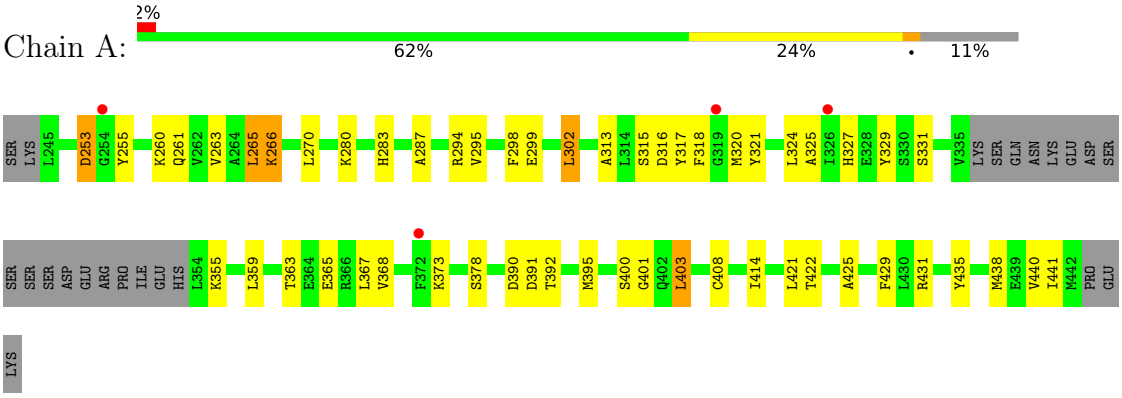
Chain	Residue	Modelled	Actual	Comment	Reference
A	243	SER	-	expression tag	UNP Q9R002
A	350	PRO	LEU	variant	UNP Q9R002
A	364	GLU	LYS	variant	UNP Q9R002
A	379	SER	THR	variant	UNP Q9R002
A	432	ALA	SER	variant	UNP Q9R002
B	243	SER	-	expression tag	UNP Q9R002
B	350	PRO	LEU	variant	UNP Q9R002
B	364	GLU	LYS	variant	UNP Q9R002
B	379	SER	THR	variant	UNP Q9R002
B	432	ALA	SER	variant	UNP Q9R002
C	243	SER	-	expression tag	UNP Q9R002
C	350	PRO	LEU	variant	UNP Q9R002
C	364	GLU	LYS	variant	UNP Q9R002
C	379	SER	THR	variant	UNP Q9R002
C	432	ALA	SER	variant	UNP Q9R002
D	243	SER	-	expression tag	UNP Q9R002
D	350	PRO	LEU	variant	UNP Q9R002
D	364	GLU	LYS	variant	UNP Q9R002
D	379	SER	THR	variant	UNP Q9R002
D	432	ALA	SER	variant	UNP Q9R002

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.

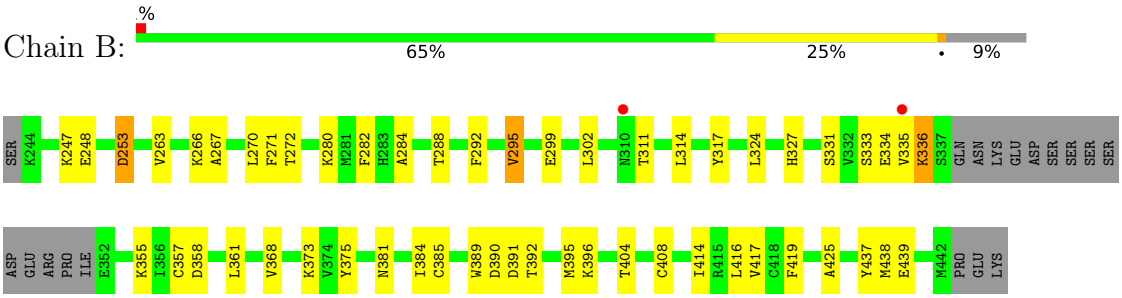
• Molecule 1: Interferon-activable protein 202

Chain A:



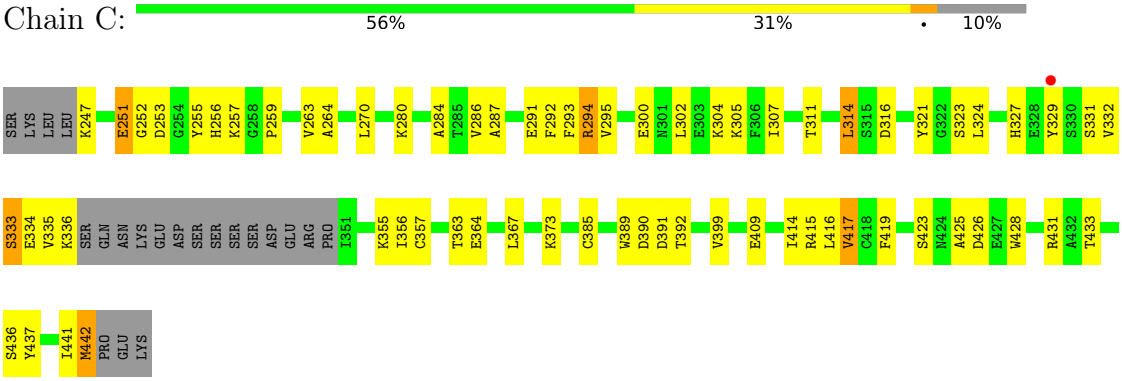
• Molecule 1: Interferon-activable protein 202

Chain B:



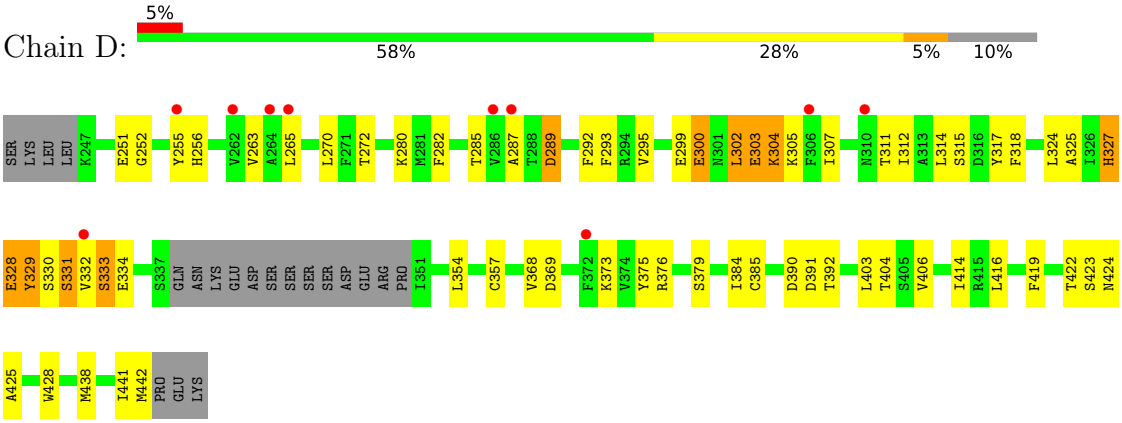
• Molecule 1: Interferon-activable protein 202

Chain C:



• Molecule 1: Interferon-activable protein 202

Chain D:


5%
58%
28%
5%
10%
SER LYS LEU LEU K247 E251 G252 Y255 H256 V262 V263 A264 L265 L270 F271 T272 K280 M281 F282 T285 V286 A287 T288 D289 F292 F293 R294 V295 E299 E300 N301 L302 E303 K304 K305 F306 I307 N310 T311 I312 A313 L314 S315 D316 Y317 F318 L324 A325 I326 H327
E328 Y329 S330 S331 V332 S333 E334 S337 GLN ASN LYS GLU ASP SER SER SER SER ASP GLU ARG PRO I351 L364 C357 V368 D369 F372 K373 V374 Y375 R376 S379 I384 C385 D390 D391 T392 L403 T404 S405 V406 I414 R415 L416 F419 T422 S423 N424
A425 W428 M438 I441 M442 PRO GLU LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	61.89Å 71.86Å 95.47Å 90.00° 102.38° 90.00°	Depositor
Resolution (Å)	46.39 – 3.40 46.39 – 3.41	Depositor EDS
% Data completeness (in resolution range)	97.9 (46.39-3.40) 94.0 (46.39-3.41)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.206 , 0.268 0.208 , 0.266	Depositor DCC
R_{free} test set	1114 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å^2)	80.4	Xtriage
Anisotropy	0.615	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 75.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5748	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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

5.1 Standard geometry

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	0.39	0/1441	0.58	0/1943
1	B	0.41	0/1491	0.60	0/2007
1	C	0.40	0/1466	0.60	0/1975
1	D	0.38	0/1460	0.58	1/1966 (0.1%)
All	All	0.40	0/5858	0.59	1/7891 (0.0%)

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.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	304	LYS	C-N-CA	6.26	137.34	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	329	TYR	Peptide

5.2 Too-close contacts

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	1414	0	1375	47	0
1	B	1463	0	1436	32	0
1	C	1438	0	1402	41	0
1	D	1433	0	1392	91	0
All	All	5748	0	5605	196	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 17.

The worst 5 of 196 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:315:SER:N	1:D:331:SER:OG	1.58	1.35
1:D:302:LEU:HD11	1:D:328:GLU:O	1.33	1.26
1:D:282:PHE:CE1	1:D:303:GLU:CB	2.23	1.21
1:D:302:LEU:CD1	1:D:328:GLU:O	1.91	1.18
1:A:255:TYR:HB2	1:D:327:HIS:ND1	1.68	1.08

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	176/203 (87%)	164 (93%)	12 (7%)	0	100 100
1	B	181/203 (89%)	170 (94%)	11 (6%)	0	100 100
1	C	178/203 (88%)	165 (93%)	12 (7%)	1 (1%)	25 57
1	D	179/203 (88%)	166 (93%)	13 (7%)	0	100 100
All	All	714/812 (88%)	665 (93%)	48 (7%)	1 (0%)	51 82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	321	TYR

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	153/182 (84%)	143 (94%)	10 (6%)	17 46
1	B	160/182 (88%)	148 (92%)	12 (8%)	13 41
1	C	156/182 (86%)	138 (88%)	18 (12%)	5 20
1	D	155/182 (85%)	144 (93%)	11 (7%)	14 44
All	All	624/728 (86%)	573 (92%)	51 (8%)	11 37

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	327	HIS
1	C	399	VAL
1	D	354	LEU
1	C	331	SER
1	C	363	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	327	HIS
1	D	360	HIS

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

6.1 Protein, DNA and RNA chains

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^{th} 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	180/203 (88%)	0.20	4 (2%)	62	60	73, 114, 145, 153	0
1	B	185/203 (91%)	-0.06	2 (1%)	80	79	60, 88, 116, 135	0
1	C	182/203 (89%)	-0.03	1 (0%)	91	90	66, 89, 119, 133	0
1	D	183/203 (90%)	0.32	10 (5%)	25	25	84, 116, 142, 158	0
All	All	730/812 (89%)	0.11	17 (2%)	60	59	60, 99, 141, 158	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	287	ALA	4.3
1	D	264	ALA	4.0
1	D	306	PHE	3.7
1	D	265	LEU	3.7
1	D	255	TYR	3.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.