wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 03:35 AM EDT

PDB ID	:	1G88
Title	:	S4AFL3ARG515 MUTANT
Authors	:	Chako, B.M.; Qin, B.; Lam, S.S.; Correia, J.J.; Lin, K.
Deposited on	:	2000-11-16
Resolution	:	3.00 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 following versions of software and data (see references (1)) were used in the production of this report:

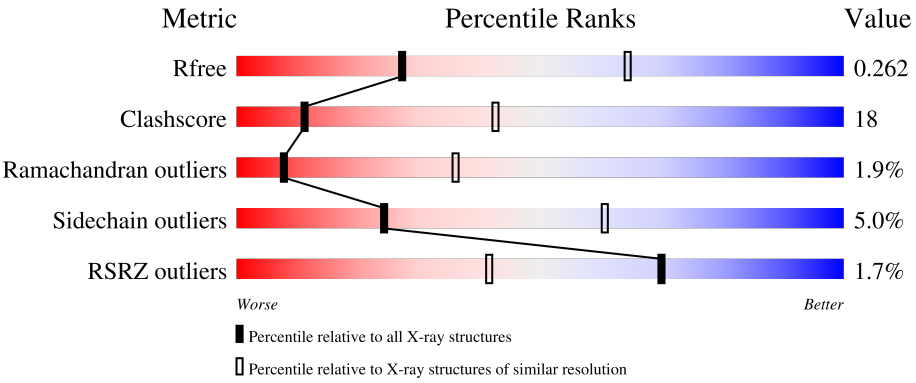
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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% 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	268	2% 50% 30% • 18%
1	B	268	% 54% 33% • 10%
1	C	268	% 59% 26% • 12%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 5431 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 SMAD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	0	0
			1733	1096	312	312	13			
1	B	242	Total	C	N	O	S	0	0	0
			1871	1184	335	339	13			
1	C	236	Total	C	N	O	S	0	0	0
			1827	1154	328	332	13			

There are 3 discrepancies between the modelled and reference sequences:

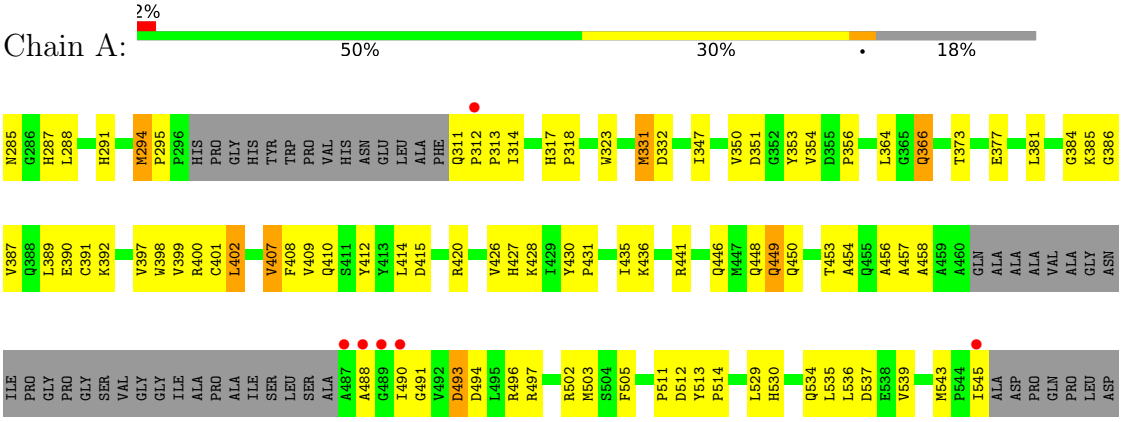
Chain	Residue	Modelled	Actual	Comment	Reference
A	515	SER	ARG	engineered mutation	UNP Q13485
B	515	SER	ARG	engineered mutation	UNP Q13485
C	515	SER	ARG	engineered mutation	UNP Q13485

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.

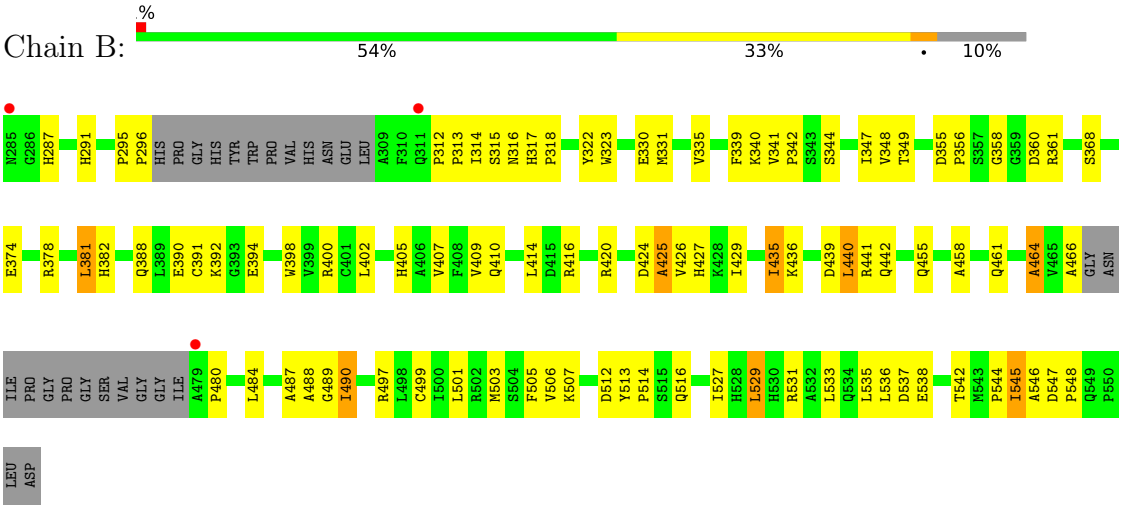
• Molecule 1: SMAD4

Chain A:



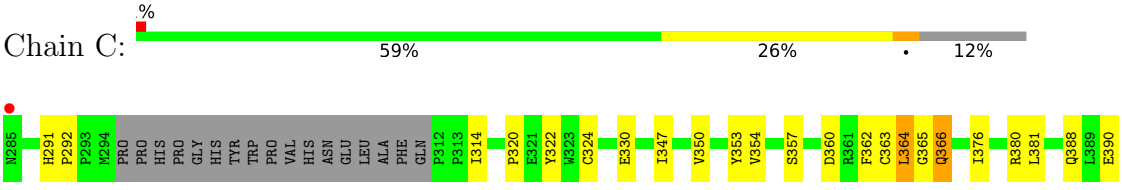
• Molecule 1: SMAD4

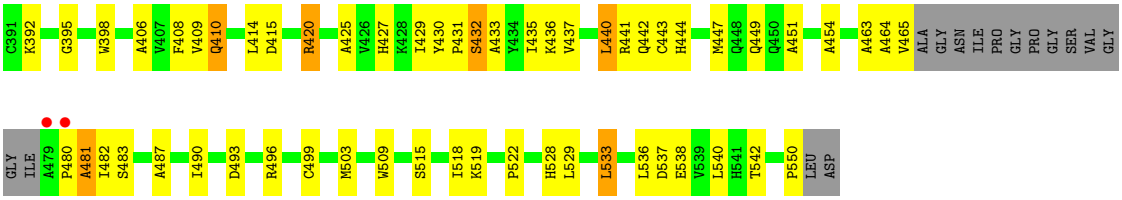
Chain B:



• Molecule 1: SMAD4

Chain C:




C391 K392 G395 W398 A406 V407 F408 V409 Q410 L414 D415 R420 A425 V426 H427 K428 I429 Y430 P431 S432 A433 Y434 I435 K436 V437 L440 R441 Q442 C443 H444 M447 Q448 Q449 Q450 A451 A454 A463 A464 V465 ALA GLY ASN ILE PRO GLY PRO GLY SER VAL GLY
GLY ILE A479 P480 A481 I482 S483 A487 I490 D493 R496 C499 M503 W509 S515 I518 K519 P522 H528 L529 L533 L536 D537 E538 V539 L540 H541 T542 P550 LEU ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α, β, γ	140.85Å 140.85Å 194.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.50 – 3.00 40.50 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.6 (40.50-3.00) 92.7 (40.50-3.00)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 3.01Å)	Xtriage
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.203 , 0.263 0.198 , 0.262	Depositor DCC
R_{free} test set	947 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	29.0	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 42.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.025 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k 0.026 for -1/2*h-1/2*k+1/2*l,-1/2*h-1/2*k-1/2*l,h-k	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5431	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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	0.35	0/1781	0.64	0/2421
1	B	0.37	0/1923	0.64	1/2619 (0.0%)
1	C	0.36	0/1876	0.62	0/2552
All	All	0.36	0/5580	0.63	1/7592 (0.0%)

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	B	416	ARG	NE-CZ-NH2	-5.10	117.75	120.30

There are no chirality outliers.

There are no planarity outliers.

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	1733	0	1681	70	0
1	B	1871	0	1817	72	0
1	C	1827	0	1777	59	0
All	All	5431	0	5275	192	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 18.

The worst 5 of 192 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:C:420:ARG:HH11	1:C:420:ARG:HB3	1.14	1.12
1:B:455:GLN:HE21	1:B:544:PRO:HD2	1.28	0.98
1:A:366:GLN:NE2	1:A:366:GLN:H	1.63	0.95
1:C:314:ILE:HD12	1:C:440:LEU:HD13	1.49	0.94
1:A:366:GLN:H	1:A:366:GLN:HE21	0.99	0.91

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	215/268 (80%)	195 (91%)	19 (9%)	1 (0%)	29 68
1	B	236/268 (88%)	206 (87%)	24 (10%)	6 (2%)	5 28
1	C	230/268 (86%)	212 (92%)	12 (5%)	6 (3%)	5 27
All	All	681/804 (85%)	613 (90%)	55 (8%)	13 (2%)	8 36

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	464	ALA
1	B	488	ALA
1	C	357	SER
1	A	493	ASP
1	B	425	ALA

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	185/217 (85%)	176 (95%)	9 (5%)	25	61
1	B	197/217 (91%)	188 (95%)	9 (5%)	27	64
1	C	193/217 (89%)	182 (94%)	11 (6%)	20	56
All	All	575/651 (88%)	546 (95%)	29 (5%)	24	60

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	490	ILE
1	C	529	LEU
1	B	545	ILE
1	C	441	ARG
1	B	529	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	334	GLN
1	C	427	HIS
1	C	449	GLN
1	B	287	HIS
1	A	455	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

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	221/268 (82%)	-0.52	6 (2%)	54	26	4, 19, 75, 98	1 (0%)
1	B	242/268 (90%)	-0.50	3 (1%)	79	54	4, 20, 80, 100	2 (0%)
1	C	236/268 (88%)	-0.59	3 (1%)	77	51	3, 21, 70, 93	1 (0%)
All	All	699/804 (86%)	-0.54	12 (1%)	70	41	3, 20, 77, 100	4 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	285	ASN	9.0
1	A	487	ALA	7.1
1	A	489	GLY	6.5
1	A	488	ALA	5.5
1	C	479	ALA	5.3

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.