



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 04:41 PM GMT

PDB ID : 3C9G  
Title : Crystal structure of uncharacterized UPF0201 protein AF\_135  
Authors : Sugadev, R.; Ozyurt, S.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2008-02-15  
Resolution : 2.30 Å(reported)

This is a full wwPDB validation 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

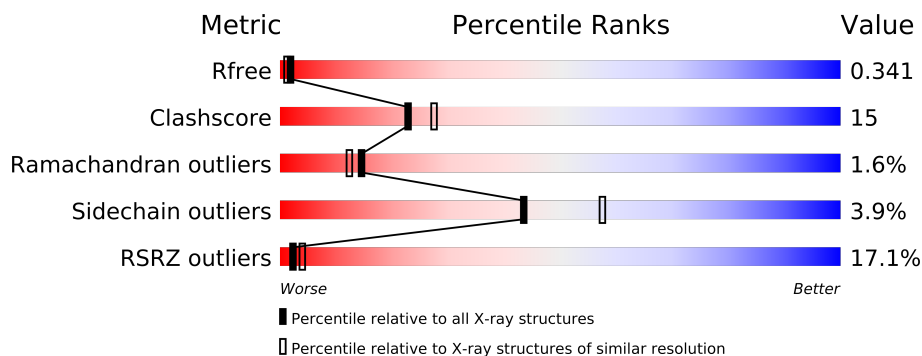
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance



The reported resolution of this entry is 2.30 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))
$R_{free}$	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	142	
1	B	142	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2109 atoms, of which 0 are hydrogen and 0 are deuterium.

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 UPF0200/UPF0201 protein AF\_1395.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	130	Total	C	N	O	S	0	0	0
			1050	660	188	201	1			
1	B	128	Total	C	N	O	S	0	0	0
			1039	656	184	198	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	SER	-	EXPRESSION TAG	UNP O28876
B	182	SER	-	EXPRESSION TAG	UNP O28876

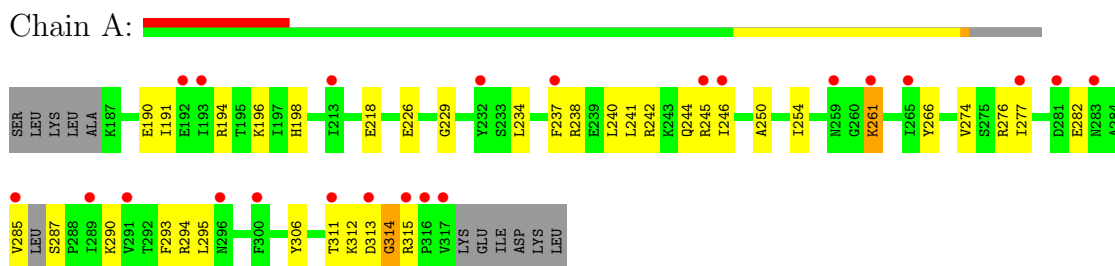
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	13	Total	O	0	0
			13	13		
2	B	7	Total	O	0	0
			7	7		

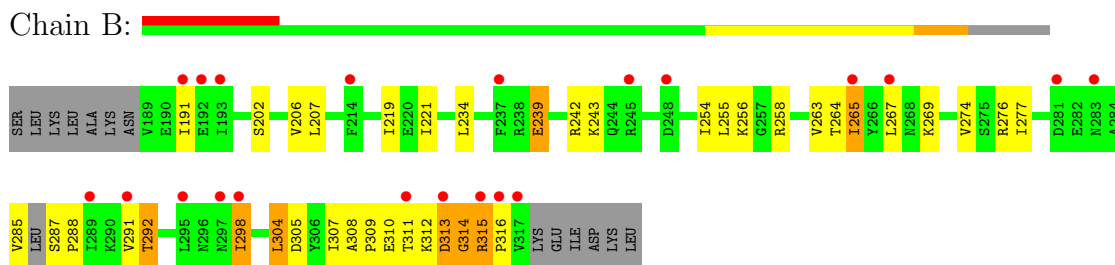
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: UPF0200/UPF0201 protein AF\_1395



- Molecule 1: UPF0200/UPF0201 protein AF\_1395



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.99Å 79.99Å 95.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.45 – 2.30 27.11 – 2.29	Depositor EDS
% Data completeness (in resolution range)	85.3 (24.45-2.30) 96.8 (27.11-2.29)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.68 (at 2.29Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.269 , 0.324 0.284 , 0.341	Depositor DCC
$R_{free}$ test set	617 reflections (4.84%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.8	Xtriage
Anisotropy	0.417	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 34.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 26567 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	2109	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	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 27.33 % 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 2.2436e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.40	0/1064	0.65	0/1433
1	B	0.41	0/1054	0.70	0/1420
All	All	0.40	0/2118	0.68	0/2853

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1050	0	1065	29	0
1	B	1039	0	1053	36	0
2	A	13	0	0	4	0
2	B	7	0	0	0	0
All	All	2109	0	2118	62	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

All (62) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:313:ASP:OD1	1:B:314:GLY:N	2.13	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:242:ARG:HE	1:A:315:ARG:HE	1.29	0.80
1:A:285:VAL:HG13	1:B:285:VAL:HA	1.66	0.77
1:B:242:ARG:HE	1:B:315:ARG:HE	1.33	0.76
1:A:190:GLU:HB3	1:A:294:ARG:CB	2.18	0.74
1:A:250:ALA:O	1:A:254:ILE:HG12	1.89	0.73
1:B:267:LEU:HD22	1:B:277:ILE:HG23	1.74	0.68
1:B:242:ARG:HE	1:B:315:ARG:NE	1.92	0.66
1:A:242:ARG:NE	1:A:315:ARG:HE	1.93	0.65
1:B:287:SER:HB2	1:B:288:PRO:HD2	1.81	0.61
1:A:312:LYS:HE2	2:A:9:HOH:O	2.02	0.60
1:A:285:VAL:HA	1:B:285:VAL:HG22	1.84	0.59
1:A:237:PHE:O	1:A:241:LEU:HG	2.02	0.59
1:A:194:ARG:HB2	1:A:290:LYS:HE3	1.85	0.59
1:A:312:LYS:CE	2:A:9:HOH:O	2.50	0.57
1:B:239:GLU:O	1:B:243:LYS:HG3	2.06	0.56
1:B:191:ILE:HD11	1:B:234:LEU:HD13	1.87	0.56
1:A:266:TYR:CZ	1:A:282:GLU:HG2	2.44	0.52
1:B:255:LEU:HD21	1:B:304:LEU:HD11	1.91	0.52
1:B:256:LYS:HB3	1:B:256:LYS:NZ	2.25	0.51
1:A:242:ARG:HE	1:A:315:ARG:NE	2.03	0.51
1:B:274:VAL:O	1:B:276:ARG:HG2	2.11	0.51
1:B:274:VAL:HG23	1:B:276:ARG:HG2	1.94	0.50
1:A:261:LYS:HE3	1:A:295:LEU:O	2.12	0.50
1:A:312:LYS:CG	2:A:9:HOH:O	2.60	0.50
1:A:245:ARG:NH1	2:A:9:HOH:O	2.45	0.49
1:B:206:VAL:HG13	1:B:269:LYS:HD2	1.93	0.49
1:A:196:LYS:HG3	1:A:198:HIS:CE1	2.48	0.49
1:B:313:ASP:OD1	1:B:313:ASP:C	2.50	0.49
1:B:265:ILE:CD1	1:B:291:VAL:HB	2.43	0.49
1:B:311:THR:C	1:B:313:ASP:H	2.16	0.49
1:A:244:GLN:O	1:A:246:ILE:HG13	2.13	0.49
1:B:219:ILE:N	1:B:219:ILE:HD12	2.28	0.48
1:B:265:ILE:HD11	1:B:291:VAL:HB	1.95	0.48
1:A:240:LEU:HB3	1:A:277:ILE:HD13	1.96	0.47
1:B:315:ARG:HA	1:B:316:PRO:HD3	1.76	0.47
1:A:242:ARG:HH11	1:A:242:ARG:HG3	1.80	0.46
1:B:265:ILE:HD13	1:B:265:ILE:H	1.81	0.45
1:B:254:ILE:HG23	1:B:263:VAL:HG11	1.98	0.45
1:A:274:VAL:HG23	1:A:276:ARG:HG3	1.99	0.45
1:A:218:GLU:O	1:A:229:GLY:HA2	2.18	0.44
1:A:238:ARG:HD2	1:A:306:TYR:CE1	2.52	0.44
1:B:313:ASP:OD1	1:B:314:GLY:CA	2.66	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:255:LEU:CD2	1:B:304:LEU:HD11	2.49	0.43
1:A:196:LYS:HE3	1:A:287:SER:OG	2.18	0.42
1:B:307:ILE:HG22	1:B:308:ALA:N	2.34	0.42
1:B:207:LEU:HD11	1:B:221:ILE:HD11	2.00	0.42
1:B:258:ARG:HG2	1:B:258:ARG:HH11	1.85	0.42
1:A:285:VAL:CG1	1:B:285:VAL:HA	2.43	0.42
1:B:264:THR:OG1	1:B:292:THR:HG22	2.19	0.42
1:B:298:ILE:N	1:B:298:ILE:HD13	2.34	0.42
1:B:256:LYS:NZ	1:B:256:LYS:CB	2.82	0.42
1:A:313:ASP:O	1:A:314:GLY:C	2.58	0.41
1:B:276:ARG:HA	1:B:276:ARG:HD2	1.90	0.41
1:B:305:ASP:O	1:B:309:PRO:HB3	2.21	0.41
1:B:311:THR:HG22	1:B:312:LYS:H	1.86	0.41
1:A:194:ARG:NH2	1:A:226:GLU:HB3	2.35	0.41
1:B:202:SER:O	1:B:206:VAL:HG23	2.21	0.41
1:A:191:ILE:HG12	1:A:293:PHE:CD2	2.56	0.41
1:A:266:TYR:CE2	1:A:282:GLU:HG2	2.56	0.40
1:B:256:LYS:HB3	1:B:256:LYS:HZ3	1.86	0.40
1:A:246:ILE:HG22	1:A:246:ILE:O	2.20	0.40

There are no symmetry-related clashes.

## 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	126/142 (89%)	117 (93%)	7 (6%)	2 (2%)	14	12
1	B	124/142 (87%)	112 (90%)	10 (8%)	2 (2%)	14	12
All	All	250/284 (88%)	229 (92%)	17 (7%)	4 (2%)	14	12

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	311	THR

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	314	GLY
1	B	314	GLY
1	B	310	GLU

### 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	116/129 (90%)	114 (98%)	2 (2%)	73	87
1	B	115/129 (89%)	108 (94%)	7 (6%)	26	34
All	All	231/258 (90%)	222 (96%)	9 (4%)	43	57

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	234	LEU
1	A	261	LYS
1	B	239	GLU
1	B	265	ILE
1	B	292	THR
1	B	298	ILE
1	B	304	LEU
1	B	313	ASP
1	B	315	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	198	HIS
1	A	259	ASN
1	A	297	ASN
1	B	278	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

### 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	130/142 (91%)	0.85	23 (17%) 2 4	24, 43, 73, 83	0
1	B	128/142 (90%)	0.76	21 (16%) 2 4	23, 43, 72, 85	0
All	All	258/284 (90%)	0.80	44 (17%) 2 4	23, 43, 74, 85	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	311	THR	5.7
1	A	281	ASP	4.6
1	A	291	VAL	4.5
1	A	232	TYR	4.4
1	B	291	VAL	4.3
1	B	311	THR	3.8
1	A	313	ASP	3.7
1	A	300	PHE	3.5
1	A	237	PHE	3.4
1	A	265	ILE	3.4
1	B	315	ARG	3.4
1	A	315	ARG	3.3
1	A	285	VAL	3.1
1	B	316	PRO	3.0
1	A	277	ILE	2.8
1	B	245	ARG	2.8
1	B	298	ILE	2.8
1	A	296	ASN	2.8
1	A	261	LYS	2.8
1	B	265	ILE	2.7
1	B	295	LEU	2.7
1	A	289	ILE	2.7
1	B	297	ASN	2.6
1	B	289	ILE	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	237	PHE	2.5
1	A	245	ARG	2.5
1	A	317	VAL	2.4
1	B	313	ASP	2.3
1	A	316	PRO	2.3
1	B	214	PHE	2.2
1	B	283	ASN	2.2
1	A	193	ILE	2.2
1	A	246	ILE	2.2
1	B	248	ASP	2.2
1	B	281	ASP	2.2
1	A	283	ASN	2.2
1	A	213	ILE	2.2
1	B	267	LEU	2.2
1	A	259	ASN	2.1
1	B	191	ILE	2.1
1	B	193	ILE	2.1
1	B	192	GLU	2.1
1	B	317	VAL	2.0
1	A	192	GLU	2.0

## 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.