



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 10:54 pm GMT

PDB ID : 4S1V  
Title : Crystal structure of phosphoglycerate oxidoreductase from *Vibrio Cholerae* o395  
Authors : Tarique, K.F.; Rehman, S.A.A.; Devi, S.; Gourinath, S.  
Deposited on : 2015-01-15  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure 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/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

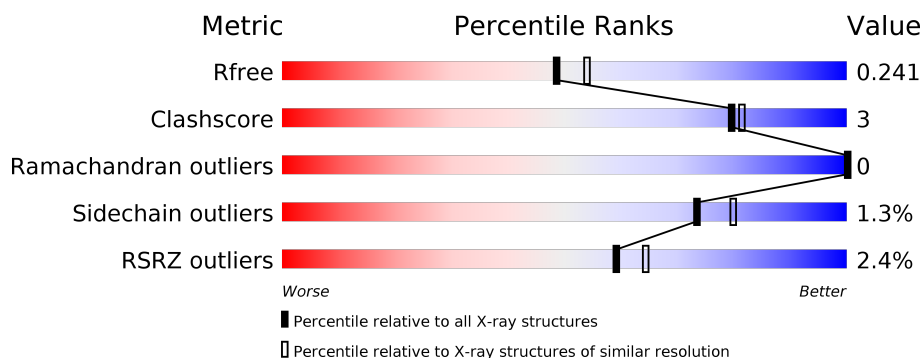
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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. 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  $\leq 5\%$ . 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	332	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>5% • 6%</div> </div> </div>
1	B	332	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>6% 5%</div> </div> </div>
1	C	332	<div> <div>%</div> <div> <div></div> <div>85%</div> <div>9% • 5%</div> </div> </div>
1	D	332	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>9% 5%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10222 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 D-3-phosphoglycerate dehydrogenase-related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	314	Total	C	N	O	S	0	0	0
			2435	1554	423	448	10			
1	A	313	Total	C	N	O	S	0	1	0
			2437	1556	425	446	10			
1	B	315	Total	C	N	O	S	0	1	0
			2448	1562	427	449	10			
1	C	315	Total	C	N	O	S	0	1	0
			2448	1562	427	449	10			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	EXPRESSION TAG	UNP A5F016
D	324	LEU	-	EXPRESSION TAG	UNP A5F016
D	325	GLU	-	EXPRESSION TAG	UNP A5F016
D	326	HIS	-	EXPRESSION TAG	UNP A5F016
D	327	HIS	-	EXPRESSION TAG	UNP A5F016
D	328	HIS	-	EXPRESSION TAG	UNP A5F016
D	329	HIS	-	EXPRESSION TAG	UNP A5F016
D	330	HIS	-	EXPRESSION TAG	UNP A5F016
D	331	HIS	-	EXPRESSION TAG	UNP A5F016
A	0	MET	-	EXPRESSION TAG	UNP A5F016
A	324	LEU	-	EXPRESSION TAG	UNP A5F016
A	325	GLU	-	EXPRESSION TAG	UNP A5F016
A	326	HIS	-	EXPRESSION TAG	UNP A5F016
A	327	HIS	-	EXPRESSION TAG	UNP A5F016
A	328	HIS	-	EXPRESSION TAG	UNP A5F016
A	329	HIS	-	EXPRESSION TAG	UNP A5F016
A	330	HIS	-	EXPRESSION TAG	UNP A5F016
A	331	HIS	-	EXPRESSION TAG	UNP A5F016
B	0	MET	-	EXPRESSION TAG	UNP A5F016
B	324	LEU	-	EXPRESSION TAG	UNP A5F016
B	325	GLU	-	EXPRESSION TAG	UNP A5F016

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	326	HIS	-	EXPRESSION TAG	UNP A5F016
B	327	HIS	-	EXPRESSION TAG	UNP A5F016
B	328	HIS	-	EXPRESSION TAG	UNP A5F016
B	329	HIS	-	EXPRESSION TAG	UNP A5F016
B	330	HIS	-	EXPRESSION TAG	UNP A5F016
B	331	HIS	-	EXPRESSION TAG	UNP A5F016
C	0	MET	-	EXPRESSION TAG	UNP A5F016
C	324	LEU	-	EXPRESSION TAG	UNP A5F016
C	325	GLU	-	EXPRESSION TAG	UNP A5F016
C	326	HIS	-	EXPRESSION TAG	UNP A5F016
C	327	HIS	-	EXPRESSION TAG	UNP A5F016
C	328	HIS	-	EXPRESSION TAG	UNP A5F016
C	329	HIS	-	EXPRESSION TAG	UNP A5F016
C	330	HIS	-	EXPRESSION TAG	UNP A5F016
C	331	HIS	-	EXPRESSION TAG	UNP A5F016

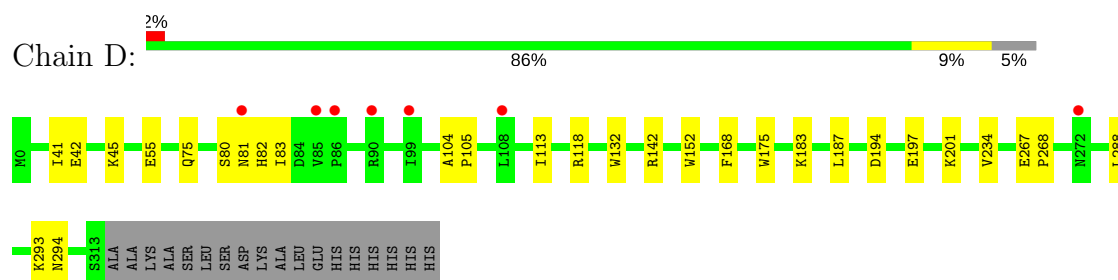
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	137	Total O 137 137	0	0
2	A	86	Total O 86 86	0	0
2	B	105	Total O 105 105	0	0
2	C	126	Total O 126 126	0	0

### 3 Residue-property plots [i](#)

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: D-3-phosphoglycerate dehydrogenase-related protein



A314

ALA
LYS
ALA
SER
LEU
SER
ASP
LYS
ALA
LEU
GLU
HIS
HIS
HIS
HIS
HIS
HIS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.92Å 83.26Å 85.99Å 64.39° 81.94° 76.03°	Depositor
Resolution (Å)	50.00 – 2.10 36.03 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.5 (50.00-2.10) 83.4 (36.03-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.202 , 0.243 0.203 , 0.241	Depositor DCC
$R_{free}$ test set	3543 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.8	Xtriage
Anisotropy	0.144	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 37.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10222	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.57% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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](#)

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	0.46	0/2492	0.53	1/3386 (0.0%)
1	B	0.48	0/2503	0.56	0/3401
1	C	0.48	2/2503 (0.1%)	0.57	1/3401 (0.0%)
1	D	0.51	0/2487	0.56	1/3380 (0.0%)
All	All	0.48	2/9985 (0.0%)	0.55	3/13568 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	175	TRP	CD2-CE2	5.10	1.47	1.41
1	C	152	TRP	CD2-CE2	5.08	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	288	LEU	CB-CA-C	-5.56	99.63	110.20
1	C	273	ASN	N-CA-CB	-5.31	101.05	110.60
1	A	288	LEU	CB-CA-C	-5.16	100.40	110.20

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	2437	0	2455	11	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2448	0	2465	13	0
1	C	2448	0	2465	29	0
1	D	2435	0	2447	16	0
2	A	86	0	0	1	0
2	B	105	0	0	2	0
2	C	126	0	0	3	0
2	D	137	0	0	4	0
All	All	10222	0	9832	63	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 3.

All (63) 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:34:VAL:HG22	1:C:272:ASN:O	1.51	1.08
1:B:313:SER:H	1:B:314:ALA:HA	1.30	0.97
1:C:142:ARG:HD2	1:C:147:ARG:HE	1.35	0.90
1:C:147:ARG:CG	1:C:147:ARG:HH11	1.98	0.77
2:D:408:HOH:O	1:C:128:HIS:HE1	1.71	0.73
1:D:293:LYS:HE2	2:D:424:HOH:O	1.88	0.72
1:C:142:ARG:NH1	1:C:147:ARG:HD2	2.06	0.70
1:C:147:ARG:HH11	1:C:147:ARG:HG2	1.59	0.67
1:D:294:ASN:HD21	1:C:147:ARG:HH12	1.46	0.63
1:A:112:LEU:HD22	1:A:283:LEU:HD21	1.86	0.58
1:D:55:GLU:HG2	1:D:82:HIS:CD2	2.39	0.58
1:A:11:ASN:ND2	1:A:31:ASN:OD1	2.38	0.57
1:C:147:ARG:CG	1:C:147:ARG:NH1	2.65	0.56
1:A:11:ASN:OD1	1:C:252:GLN:NE2	2.39	0.56
1:C:80:SER:OG	1:C:81:ASN:N	2.39	0.54
1:B:313:SER:N	1:B:314:ALA:HA	2.02	0.54
1:C:270:LEU:O	1:C:273:ASN:HB3	2.09	0.53
1:C:5:ILE:HG12	1:C:51:VAL:HB	1.90	0.53
1:D:118:ARG:HD3	2:C:405:HOH:O	2.08	0.52
1:A:87:LEU:HA	1:A:90:ARG:NH1	2.26	0.51
1:D:183:LYS:HE3	1:D:187:LEU:HD11	1.92	0.51
1:C:119:HIS:HD2	2:C:471:HOH:O	1.94	0.51
1:A:178:GLU:O	1:A:182:GLN:HG2	2.12	0.50
1:C:5:ILE:HD12	1:C:29:VAL:HG22	1.94	0.50
1:D:267:GLU:HA	1:D:268:PRO:C	2.34	0.48
1:D:152:TRP:O	2:D:455:HOH:O	2.20	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:197:GLU:HG3	1:D:201:LYS:HE2	1.96	0.46
2:D:408:HOH:O	1:C:128:HIS:CE1	2.57	0.46
1:C:175:TRP:CG	1:C:176:GLY:N	2.84	0.46
1:C:283:LEU:C	1:C:283:LEU:HD23	2.36	0.46
1:C:104:ALA:HB3	1:C:105:PRO:HD3	1.97	0.46
1:B:0:MET:HB3	1:B:25:HIS:CE1	2.51	0.45
1:A:274:GLU:HA	1:A:275:PRO:HD2	1.88	0.45
1:A:292:GLU:CD	1:B:143:THR:HG23	2.36	0.45
1:D:113:ILE:HG12	1:D:234:VAL:HG21	1.99	0.45
1:B:147:ARG:HE	1:B:203:ASP:HB3	1.81	0.45
1:C:6:LEU:HD21	1:C:44:LEU:HD21	1.99	0.45
1:C:147:ARG:HH11	1:C:147:ARG:HG3	1.79	0.45
1:A:252:GLN:HG3	1:A:253:ALA:N	2.32	0.45
1:B:6:LEU:HD21	1:B:44:LEU:HD21	1.98	0.44
1:B:267:GLU:HA	1:B:268:PRO:C	2.38	0.44
1:C:224:LEU:O	1:C:257:ARG:NH1	2.47	0.44
2:A:471:HOH:O	1:B:118:ARG:HD3	2.17	0.44
1:D:75:GLN:OE1	1:D:82:HIS:CE1	2.71	0.44
1:D:175:TRP:CZ3	1:D:194:ASP:HA	2.53	0.43
1:A:290:TYR:HB2	1:A:299:TYR:HE2	1.84	0.43
1:C:313:SER:HA	1:C:314:ALA:HA	1.68	0.42
1:D:168:PHE:CE2	1:C:107:GLU:HG2	2.54	0.42
1:D:41:ILE:O	1:D:45:LYS:HG2	2.20	0.42
1:B:14:ARG:HG3	2:B:406:HOH:O	2.18	0.42
1:C:119:HIS:HE1	2:C:470:HOH:O	2.02	0.42
1:B:288:LEU:O	1:B:289:GLY:C	2.58	0.41
1:B:80:SER:HB2	1:B:81:ASN:H	1.73	0.41
1:D:80:SER:OG	1:D:81:ASN:N	2.52	0.41
1:D:132:TRP:CG	1:C:268:PRO:HB3	2.56	0.41
1:C:21:CYS:SG	1:C:301:GLN:NE2	2.93	0.41
1:C:147:ARG:NH1	1:C:147:ARG:HG3	2.36	0.41
1:B:313:SER:N	1:B:314:ALA:CA	2.81	0.41
1:A:118:ARG:HD3	2:B:488:HOH:O	2.21	0.41
1:C:297:GLU:O	1:C:301:GLN:HB2	2.20	0.41
1:B:259:ALA:HB3	1:B:282:VAL:HG22	2.03	0.40
1:C:142:ARG:HB2	1:C:147:ARG:HH21	1.85	0.40
1:D:104:ALA:HB3	1:D:105:PRO:HD3	2.04	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	312/332 (94%)	300 (96%)	12 (4%)	0	100	100
1	B	314/332 (95%)	300 (96%)	14 (4%)	0	100	100
1	C	314/332 (95%)	304 (97%)	10 (3%)	0	100	100
1	D	312/332 (94%)	304 (97%)	8 (3%)	0	100	100
All	All	1252/1328 (94%)	1208 (96%)	44 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	261/275 (95%)	257 (98%)	4 (2%)	70	76
1	B	262/275 (95%)	258 (98%)	4 (2%)	70	76
1	C	262/275 (95%)	259 (99%)	3 (1%)	78	83
1	D	261/275 (95%)	258 (99%)	3 (1%)	78	83
All	All	1046/1100 (95%)	1032 (99%)	14 (1%)	73	80

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	42	GLU
1	D	83	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	142	ARG
1	A	11	ASN
1	A	31	ASN
1	A	33	SER
1	A	252	GLN
1	B	0	MET
1	B	90	ARG
1	B	252	GLN
1	B	256	MET
1	C	90	ARG
1	C	147	ARG
1	C	210	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	20	GLN
1	D	82	HIS
1	D	128	HIS
1	D	258	GLN
1	D	294	ASN
1	A	10	GLN
1	A	11	ASN
1	A	31	ASN
1	A	131	HIS
1	A	222	GLN
1	A	252	GLN
1	A	258	GLN
1	B	66	HIS
1	B	128	HIS
1	B	131	HIS
1	C	23	GLN
1	C	119	HIS
1	C	128	HIS
1	C	252	GLN

### 5.3.3 RNA ⓘ

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](#)

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<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	313/332 (94%)	0.19	13 (4%) 37 44	20, 33, 48, 66	0
1	B	315/332 (94%)	-0.09	6 (1%) 67 71	19, 26, 43, 70	0
1	C	315/332 (94%)	-0.01	4 (1%) 77 81	19, 29, 42, 64	0
1	D	314/332 (94%)	0.13	7 (2%) 62 67	20, 27, 41, 54	0
All	All	1257/1328 (94%)	0.05	30 (2%) 59 64	19, 28, 45, 70	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	214	ALA	4.7
1	B	314	ALA	4.5
1	A	23	GLN	4.4
1	C	0	MET	4.3
1	A	24	GLY	4.0
1	D	90	ARG	3.9
1	B	313	SER	3.4
1	D	81	ASN	3.4
1	A	90	ARG	3.4
1	B	312	HIS	3.3
1	D	86	PRO	3.0
1	A	312	HIS	2.9
1	C	99	ILE	2.9
1	C	314	ALA	2.9
1	D	272	ASN	2.9
1	A	179	ALA	2.8
1	D	99	ILE	2.8
1	A	85	VAL	2.7
1	A	0	MET	2.5
1	A	272	ASN	2.5
1	B	23	GLN	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	24	GLY	2.4
1	A	86	PRO	2.3
1	D	85	VAL	2.3
1	B	20	GLN	2.2
1	A	89	GLU	2.2
1	A	213	ASP	2.2
1	A	275	PRO	2.1
1	C	272	ASN	2.1
1	D	108	LEU	2.0

## 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.