



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 10, 2020 – 11:07 AM BST

PDB ID : 5Z4W  
Title : Crystal structure of signalling protein from buffalo (SPB-40) with an altered conformation of Trp78 at 1.79 Å resolution  
Authors : Singh, P.K.; Chaudhary, A.; Tyagi, T.K.; Kaur, P.; Sharma, S.; Singh, T.P.  
Deposited on : 2018-01-15  
Resolution : 1.79 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

<https://www.wwpdb.org/validation/2017/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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

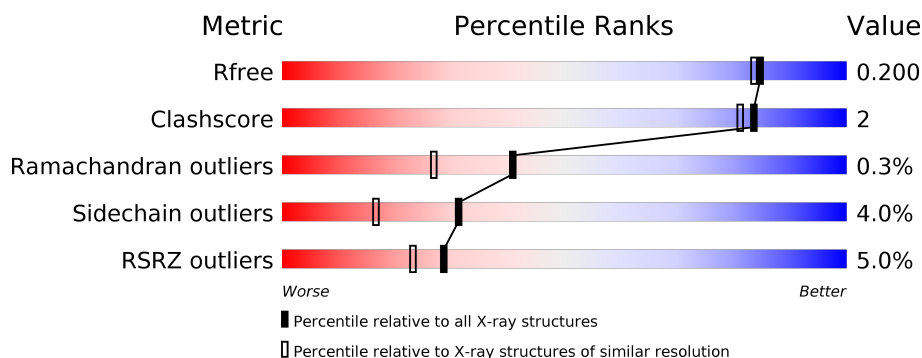
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.79 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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  $\leq 5\%$ . 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	361	<div> <div>5%</div> <div>86%</div> <div>12%</div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3284 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 Chitinase-3-like protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	361	Total	C	N	O	S	0	0	0
			2874	1841	499	525	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ASP	deletion	UNP Q7YS85

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



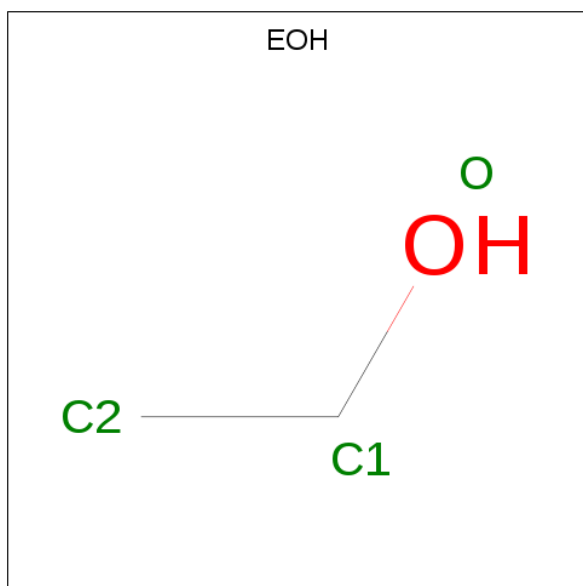
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:  $C_6H_{14}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	6	2		

- Molecule 4 is ETHANOL (three-letter code: EOH) (formula:  $C_2H_6O$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			3	2	1		

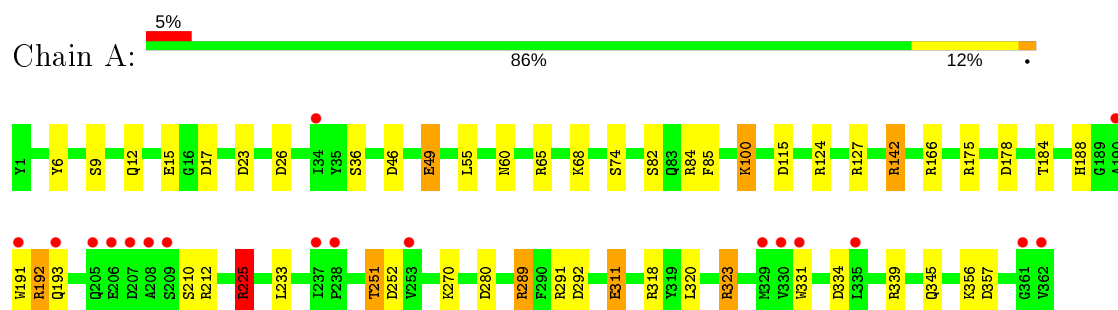
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	385	Total 385	O 385	0	0

### 3 Residue-property plots

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: Chitinase-3-like protein 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.07Å 66.73Å 106.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.49 – 1.79 31.24 – 1.79	Depositor EDS
% Data completeness (in resolution range)	98.6 (56.49-1.79) 98.6 (31.24-1.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.39 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.150 , 0.186 0.163 , 0.200	Depositor DCC
$R_{free}$ test set	2058 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.6	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3284	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 5.57% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MPD, EOH, NAG

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	1.50	22/2952 (0.7%)	1.54	39/4005 (1.0%)

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	A	0	1

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	311	GLU	CD-OE1	11.55	1.38	1.25
1	A	311	GLU	CG-CD	8.95	1.65	1.51
1	A	166	ARG	CZ-NH1	8.03	1.43	1.33
1	A	289	ARG	CD-NE	-7.57	1.33	1.46
1	A	82	SER	CB-OG	7.33	1.51	1.42
1	A	323	ARG	CD-NE	-7.26	1.34	1.46
1	A	60	ASN	CB-CG	7.23	1.67	1.51
1	A	191	TRP	CE3-CZ3	7.20	1.50	1.38
1	A	289	ARG	NE-CZ	-7.01	1.24	1.33
1	A	49	GLU	CD-OE1	-6.77	1.18	1.25
1	A	127	ARG	NE-CZ	6.52	1.41	1.33
1	A	49	GLU	CD-OE2	-6.18	1.18	1.25
1	A	311	GLU	CD-OE2	6.15	1.32	1.25
1	A	191	TRP	CE2-CZ2	5.95	1.49	1.39
1	A	127	ARG	CZ-NH1	5.82	1.40	1.33
1	A	210	SER	C-N	5.77	1.47	1.34
1	A	331	TRP	CE3-CZ3	5.34	1.47	1.38
1	A	100	LYS	CE-NZ	5.31	1.62	1.49

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	6	TYR	CZ-OH	5.21	1.46	1.37
1	A	60	ASN	CG-ND2	5.13	1.45	1.32
1	A	65	ARG	CZ-NH2	5.08	1.39	1.33
1	A	233	LEU	N-CA	-5.07	1.36	1.46

All (39) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	289	ARG	NE-CZ-NH2	-21.97	109.32	120.30
1	A	323	ARG	NE-CZ-NH2	-19.90	110.35	120.30
1	A	318	ARG	NE-CZ-NH2	-17.56	111.52	120.30
1	A	166	ARG	NE-CZ-NH1	16.62	128.61	120.30
1	A	318	ARG	NE-CZ-NH1	15.07	127.83	120.30
1	A	127	ARG	NE-CZ-NH1	14.01	127.31	120.30
1	A	166	ARG	NE-CZ-NH2	-13.50	113.55	120.30
1	A	289	ARG	CD-NE-CZ	13.09	141.93	123.60
1	A	323	ARG	NE-CZ-NH1	11.77	126.19	120.30
1	A	339	ARG	NE-CZ-NH1	11.55	126.07	120.30
1	A	178	ASP	CB-CG-OD2	10.60	127.84	118.30
1	A	289	ARG	NE-CZ-NH1	-10.21	115.19	120.30
1	A	84	ARG	NE-CZ-NH2	-8.61	115.99	120.30
1	A	339	ARG	NE-CZ-NH2	-8.28	116.16	120.30
1	A	334	ASP	CB-CG-OD2	-7.41	111.63	118.30
1	A	84	ARG	NE-CZ-NH1	7.20	123.90	120.30
1	A	292	ASP	CB-CG-OD1	7.18	124.76	118.30
1	A	55	LEU	CB-CG-CD2	7.17	123.19	111.00
1	A	26	ASP	CB-CG-OD2	-6.93	112.06	118.30
1	A	127	ARG	NE-CZ-NH2	-6.89	116.86	120.30
1	A	175	ARG	NE-CZ-NH1	6.86	123.73	120.30
1	A	289	ARG	CA-CB-CG	6.36	127.39	113.40
1	A	142	ARG	NE-CZ-NH1	6.21	123.41	120.30
1	A	23	ASP	CB-CG-OD1	6.17	123.85	118.30
1	A	280	ASP	CB-CG-OD2	6.04	123.74	118.30
1	A	292	ASP	CB-CG-OD2	-5.98	112.91	118.30
1	A	252	ASP	N-CA-CB	-5.86	100.06	110.60
1	A	318	ARG	CG-CD-NE	-5.74	99.74	111.80
1	A	142	ARG	NE-CZ-NH2	-5.61	117.50	120.30
1	A	17	ASP	CB-CG-OD2	-5.59	113.27	118.30
1	A	36	SER	CA-CB-OG	-5.49	96.38	111.20
1	A	318	ARG	CD-NE-CZ	5.48	131.27	123.60
1	A	60	ASN	CB-CA-C	5.46	121.32	110.40
1	A	225	ARG	NE-CZ-NH1	5.44	123.02	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	124	ARG	NE-CZ-NH1	-5.39	117.61	120.30
1	A	225	ARG	CD-NE-CZ	5.34	131.08	123.60
1	A	356	LYS	CD-CE-NZ	-5.25	99.63	111.70
1	A	46	ASP	CB-CG-OD1	5.24	123.01	118.30
1	A	289	ARG	CB-CG-CD	-5.20	98.08	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	289	ARG	Sidechain

## 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	2874	0	2805	13	0
2	A	14	0	13	1	0
3	A	8	0	14	0	0
4	A	3	0	6	0	0
5	A	385	0	0	7	0
All	All	3284	0	2838	14	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 2.

All (14) 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:192:ARG:O	1:A:193:GLN:HB2	1.68	0.91
2:A:401:NAG:O3	5:A:501:HOH:O	2.00	0.80
1:A:311:GLU:HG2	5:A:787:HOH:O	1.92	0.68
1:A:251:THR:HG23	5:A:703:HOH:O	2.09	0.52
1:A:49:GLU:OE2	5:A:502:HOH:O	2.20	0.48
1:A:225:ARG:O	1:A:225:ARG:NH1	2.49	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:188:HIS:CE1	1:A:192:ARG:HD3	2.53	0.43
1:A:357:ASP:OD2	5:A:503:HOH:O	2.21	0.43
1:A:323:ARG:HD2	5:A:712:HOH:O	2.18	0.43
1:A:311:GLU:CG	5:A:787:HOH:O	2.60	0.42
1:A:9:SER:O	1:A:12:GLN:HG2	2.19	0.42
1:A:74:SER:HA	1:A:115:ASP:HB3	2.00	0.42
1:A:320:LEU:C	1:A:320:LEU:HD12	2.41	0.41
1:A:192:ARG:HH11	1:A:192:ARG:HD2	1.79	0.41

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	359/361 (99%)	350 (98%)	8 (2%)	1 (0%)	41	27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	212	ARG

### 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	302/302 (100%)	290 (96%)	12 (4%)	31	16

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	GLU
1	A	68	LYS
1	A	85	PHE
1	A	100	LYS
1	A	142	ARG
1	A	184	THR
1	A	192	ARG
1	A	225	ARG
1	A	251	THR
1	A	270	LYS
1	A	291	ARG
1	A	345	GLN

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

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

3 ligands 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	$\# Z  > 2$
2	NAG	A	401	1	14,14,15	1.19	2 (14%)	17,19,21	1.09	2 (11%)
3	MPD	A	402	-	7,7,7	0.59	0	9,10,10	0.87	0
4	EOH	A	403	-	2,2,2	0.54	0	1,1,1	0.22	0

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
2	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	MPD	A	402	-	-	2/5/5/5	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	NAG	O5-C1	-2.42	1.39	1.43
2	A	401	NAG	C2-N2	-2.09	1.42	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	NAG	O3-C3-C2	-2.14	105.04	109.47
2	A	401	NAG	C3-C4-C5	-2.07	106.54	110.24

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	402	MPD	C2-C3-C4-O4
3	A	402	MPD	C2-C3-C4-C5

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	NAG	1	0

## 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](#)

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	361/361 (100%)	0.09	18 (4%)	28 23	16, 26, 47, 126	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	209	SER	8.2
1	A	206	GLU	8.1
1	A	208	ALA	7.9
1	A	362	VAL	7.2
1	A	207	ASP	6.9
1	A	193	GLN	3.6
1	A	330	VAL	3.4
1	A	237	ILE	3.1
1	A	253	VAL	3.1
1	A	205	GLN	3.1
1	A	331	TRP	2.9
1	A	191	TRP	2.9
1	A	34	ILE	2.6
1	A	190	ALA	2.3
1	A	335	LEU	2.2
1	A	238	PRO	2.2
1	A	329	MET	2.0
1	A	361	GLY	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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MPD	A	402	8/8	0.90	0.13	20,28,29,37	0
2	NAG	A	401	14/15	0.92	0.13	32,36,43,44	0
4	EOH	A	403	3/3	0.95	0.25	38,38,40,40	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.