



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 10:32 pm GMT

PDB ID : 2Y1D  
Title : X-ray structure of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, DXR, Rv2870c, from Mycobacterium tuberculosis, in complex with a 3,4-dichlorophenyl-substituted fosmidomycin analogue and manganese.  
Authors : Henriksson, L.M.; Larsson, A.M.S.; Bergfors, T.; Bjorkelid, C.; Unge, T.; Mowbray, S.L.; Jones, T.A.  
Deposited on : 2010-12-08  
Resolution : 2.05 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
Xtriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

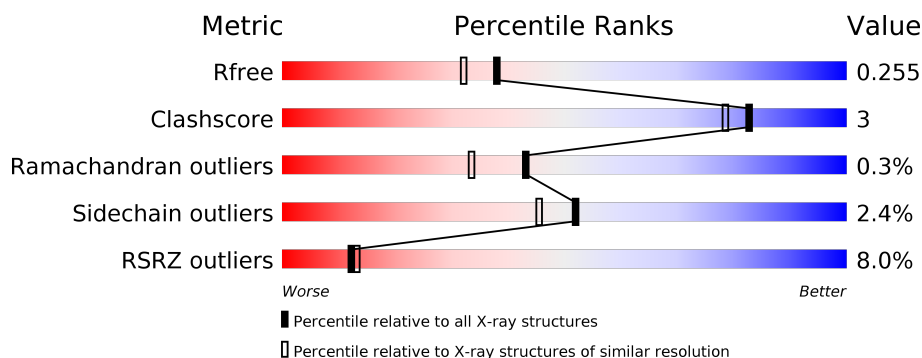
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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))
$R_{free}$	111664	1449 (2.04-2.04)
Clashscore	122126	1524 (2.04-2.04)
Ramachandran outliers	120053	1512 (2.04-2.04)
Sidechain outliers	120020	1512 (2.04-2.04)
RSRZ outliers	108989	1429 (2.04-2.04)

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	398	<div> <div>13%</div> <div> <div></div> <div>84%</div> <div>7%</div> <div>10%</div> </div> </div>
1	B	398	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>9%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5806 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.

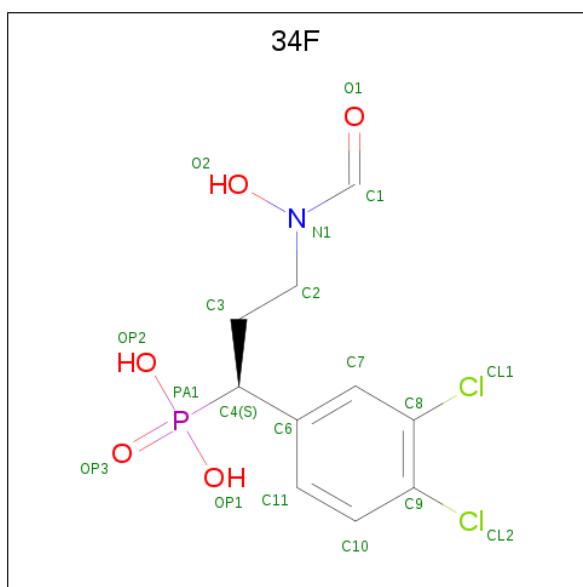
- Molecule 1 is a protein called 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	359	Total	C	N	O	S	0	0	0
			2632	1651	473	501	7			
1	B	379	Total	C	N	O	S	0	3	0
			2806	1757	508	532	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	THR	-	expression tag	UNP A2VLK3
A	-7	MET	-	expression tag	UNP A2VLK3
A	-6	ALA	-	expression tag	UNP A2VLK3
A	-5	HIS	-	expression tag	UNP A2VLK3
A	-4	HIS	-	expression tag	UNP A2VLK3
A	-3	HIS	-	expression tag	UNP A2VLK3
A	-2	HIS	-	expression tag	UNP A2VLK3
A	-1	HIS	-	expression tag	UNP A2VLK3
A	0	HIS	-	expression tag	UNP A2VLK3
B	-8	THR	-	expression tag	UNP A2VLK3
B	-7	MET	-	expression tag	UNP A2VLK3
B	-6	ALA	-	expression tag	UNP A2VLK3
B	-5	HIS	-	expression tag	UNP A2VLK3
B	-4	HIS	-	expression tag	UNP A2VLK3
B	-3	HIS	-	expression tag	UNP A2VLK3
B	-2	HIS	-	expression tag	UNP A2VLK3
B	-1	HIS	-	expression tag	UNP A2VLK3
B	0	HIS	-	expression tag	UNP A2VLK3

- Molecule 2 is (1S)-1-(3,4-DICHLOROPHENYL)-3-[FORMYL(HYDROXY)AMINO]PROPYL}PHOSPHONIC ACID (three-letter code: 34F) (formula: C<sub>10</sub>H<sub>12</sub>Cl<sub>2</sub>NO<sub>5</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	P	
			19	10	2	1	5	1	

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mn		
			1	1	0	0
3	A	1	Total	Mn		
			1	1	0	0

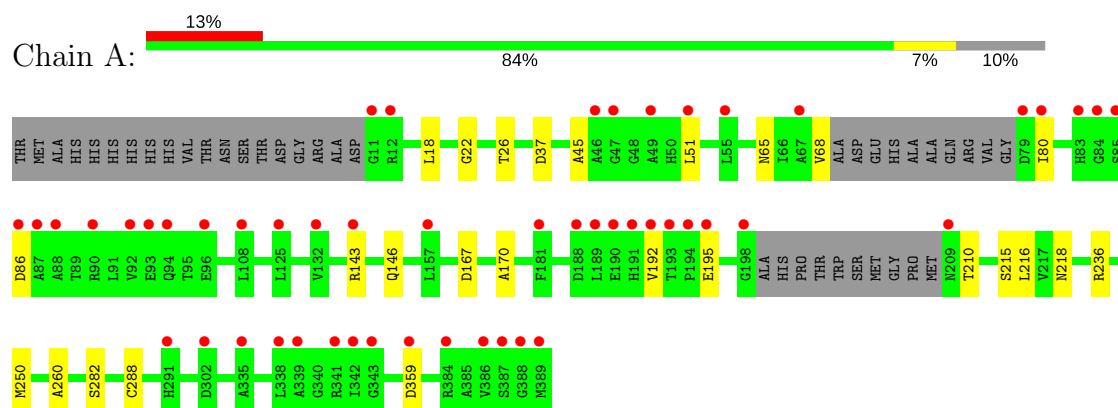
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	118	Total	O		
			118	118	0	0
4	B	229	Total	O		
			229	229	0	0

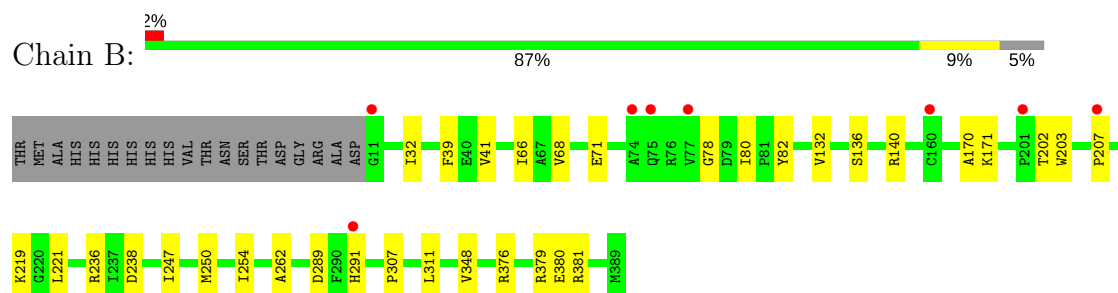
### 3 Residue-property plots [i](#)

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: 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE



#### • Molecule 1: 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.62Å 67.32Å 84.30Å 90.00° 107.38° 90.00°	Depositor
Resolution (Å)	40.00 – 2.05 31.05 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.00-2.05) 99.5 (31.05-2.05)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.209 , 0.261 0.209 , 0.255	Depositor DCC
$R_{free}$ test set	2352 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.0	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 47.1	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	5806	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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: 34F, MN

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	0/2678	0.50	0/3655
1	B	0.39	0/2860	0.53	0/3907
All	All	0.37	0/5538	0.51	0/7562

There are no bond length outliers.

There are no bond angle outliers.

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	2632	0	2632	11	0
1	B	2806	0	2794	19	0
2	A	19	0	9	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	118	0	0	1	0
4	B	229	0	0	1	0
All	All	5806	0	5435	29	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 (29) 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:B:376:ARG:O	1:B:380:GLU:HG2	1.90	0.72
1:B:170:ALA:O	1:B:236:ARG:HD2	1.97	0.65
1:A:22:GLY:O	1:A:26:THR:HG23	2.01	0.61
1:A:143:ARG:HG3	1:A:146:GLN:HB2	1.84	0.59
1:B:247:ILE:HD13	1:B:262:ALA:HB2	1.84	0.58
1:A:288:CYS:SG	4:A:2080:HOH:O	2.58	0.56
1:B:202[B]:THR:H	1:B:207:PRO:HG3	1.72	0.55
1:A:218:ASN:ND2	2:A:1390:34F:H3	2.23	0.54
1:A:250:MET:HG2	1:A:260:ALA:HB2	1.91	0.52
1:B:171:LYS:HD3	1:B:238:ASP:OD2	2.11	0.50
1:A:68:VAL:HG23	1:A:80:ILE:HG21	1.94	0.49
1:B:132:VAL:HG11	1:B:221:LEU:HD22	1.94	0.49
1:B:202[A]:THR:H	1:B:207:PRO:HG3	1.77	0.48
1:B:32:ILE:HG21	1:B:41:VAL:HG23	1.96	0.48
1:B:202[A]:THR:HG22	1:B:203:TRP:N	2.31	0.44
1:A:282:SER:HB2	1:B:254:ILE:HD12	1.99	0.43
1:B:379:ARG:NH2	4:B:2221:HOH:O	2.50	0.43
1:B:219:LYS:HD3	1:B:219:LYS:HA	1.86	0.43
1:B:68:VAL:HG23	1:B:80:ILE:HG21	1.99	0.43
1:A:250:MET:HG2	1:A:260:ALA:CB	2.49	0.42
1:A:18:LEU:HA	1:A:45:ALA:HB3	2.01	0.42
1:B:66:ILE:HD11	1:B:80:ILE:HD12	2.01	0.42
1:B:136:SER:O	1:B:140:ARG:HB2	2.19	0.42
1:B:289:ASP:OD2	1:B:291:HIS:HB2	2.19	0.42
1:A:170:ALA:O	1:A:236:ARG:HD2	2.20	0.41
1:A:215:SER:O	1:A:216:LEU:HB2	2.20	0.41
1:B:32:ILE:HG23	1:B:39:PHE:HB2	2.03	0.41
1:B:307:PRO:HG2	1:B:348:VAL:HG11	2.03	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	353/398 (89%)	339 (96%)	13 (4%)	1 (0%)	43	33
1	B	380/398 (96%)	374 (98%)	5 (1%)	1 (0%)	43	33
All	All	733/796 (92%)	713 (97%)	18 (2%)	2 (0%)	43	33

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	65	ASN
1	B	78	GLY

### 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	268/298 (90%)	260 (97%)	8 (3%)	44	37
1	B	285/298 (96%)	280 (98%)	5 (2%)	62	57
All	All	553/596 (93%)	540 (98%)	13 (2%)	52	45

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	ASP
1	A	51	LEU
1	A	86	ASP
1	A	167	ASP
1	A	192	VAL
1	A	195	GLU
1	A	210	THR
1	A	359	ASP
1	B	71	GLU
1	B	82	TYR
1	B	250	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	311	LEU
1	B	381	ARG

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

Mol	Chain	Res	Type
1	A	60	GLN
1	A	291	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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	34F	A	1390	3	16,19,19	4.12	5 (31%)	18,27,27	0.90	1 (5%)

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	34F	A	1390	3	-	0/15/17/17	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1390	34F	PA1-C4	-15.49	1.62	1.82
2	A	1390	34F	PA1-OP2	-2.10	1.51	1.54
2	A	1390	34F	PA1-OP1	-2.01	1.51	1.54
2	A	1390	34F	C8-CL1	2.65	1.80	1.73
2	A	1390	34F	C9-CL2	2.66	1.80	1.73

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1390	34F	OP2-PA1-OP1	2.07	113.29	107.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1390	34F	1	0

## 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	359/398 (90%)	0.82	51 (14%) <b>2</b> <b>2</b>	22, 41, 67, 75	0
1	B	379/398 (95%)	0.09	8 (2%) 63 68	16, 26, 44, 52	0
All	All	738/796 (92%)	0.45	59 (7%) <b>12</b> <b>13</b>	16, 32, 56, 75	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	GLY	8.8
1	A	389	MET	7.7
1	A	191	HIS	6.0
1	A	67	ALA	5.3
1	A	108	LEU	5.1
1	A	341	ARG	4.7
1	A	86	ASP	4.6
1	A	84	GLY	4.6
1	A	338	LEU	4.5
1	A	85	SER	3.9
1	A	343	GLY	3.8
1	A	83	HIS	3.8
1	A	80	ILE	3.7
1	A	388	GLY	3.7
1	B	201	PRO	3.6
1	A	291	HIS	3.6
1	A	342	ILE	3.4
1	A	198	GLY	3.4
1	A	386	VAL	3.4
1	A	209	ASN	3.3
1	A	51	LEU	3.3
1	B	11	GLY	3.3
1	A	188	ASP	3.3
1	A	96	GLU	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	195	GLU	3.2
1	A	90	ARG	3.1
1	A	49	ALA	3.0
1	A	384	ARG	3.0
1	A	88	ALA	3.0
1	A	335	ALA	2.9
1	A	125	LEU	2.9
1	A	47	GLY	2.8
1	A	79	ASP	2.7
1	A	387	SER	2.7
1	A	339	ALA	2.7
1	A	189	LEU	2.7
1	A	55	LEU	2.7
1	A	157	LEU	2.5
1	A	190	GLU	2.5
1	B	207	PRO	2.5
1	A	87	ALA	2.5
1	A	46	ALA	2.5
1	B	77	VAL	2.4
1	B	75	GLN	2.4
1	B	291	HIS	2.4
1	A	193	THR	2.4
1	A	12	ARG	2.4
1	A	359	ASP	2.4
1	B	74	ALA	2.3
1	A	302	ASP	2.2
1	A	132	VAL	2.2
1	A	194	PRO	2.1
1	A	181	PHE	2.1
1	A	93	GLU	2.1
1	A	143	ARG	2.1
1	A	192	VAL	2.1
1	B	160	CYS	2.1
1	A	94	GLN	2.0
1	A	92	VAL	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 [i](#)

There are no carbohydrates 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
2	34F	A	1390	19/19	0.86	0.19	35,41,46,48	0
3	MN	A	1391	1/1	0.99	0.14	30,30,30,30	0
3	MN	B	1390	1/1	0.99	0.07	30,30,30,30	0

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

There are no such residues in this entry.