



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 13, 2024 – 06:56 pm GMT

PDB ID : 6QFO  
Title : EngBF DARPin Fusion 9b 3G124  
Authors : Ernst, P.; Pluckthun, A.; Mittl, P.R.E.  
Deposited on : 2019-01-10  
Resolution : 2.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

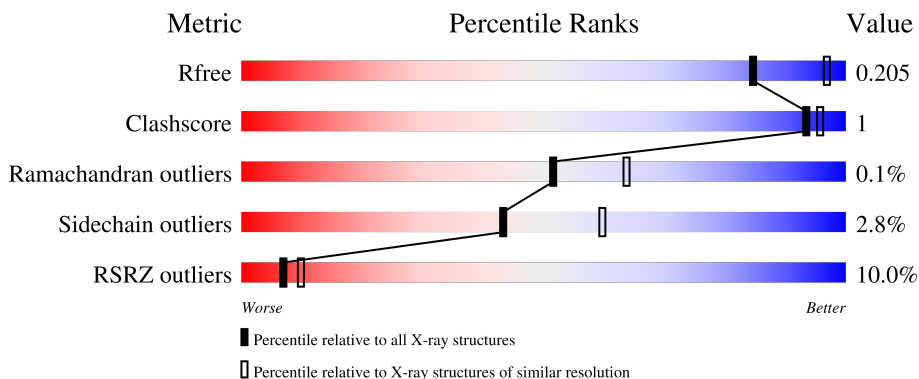
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (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 of 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	1362	

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 11376 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 PEGA domain-containing protein, PEGA domain-containing protein, EngBF DARPin fusion 9b 3G124.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1348	10339	6432	1782	2089	36	0	1	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	334	GLY	-	expression tag	UNP A0A413AG90
A	335	PRO	-	expression tag	UNP A0A413AG90
A	336	LEU	-	expression tag	UNP A0A413AG90
A	337	GLY	-	expression tag	UNP A0A413AG90
A	338	SER	-	expression tag	UNP A0A413AG90
A	339	MET	-	expression tag	UNP A0A413AG90
A	1064	CYS	GLU	conflict	UNP A0A413AG90
A	1090	CYS	GLU	conflict	UNP A0A413AG90
A	1118	CYS	ASP	conflict	UNP A0A413AG90
A	1309	ARG	GLN	conflict	UNP A0A413AG90
A	1406	CYS	VAL	conflict	UNP A0A413AG90
A	1476	CYS	GLU	conflict	UNP A0A413AG90
A	1488	CYS	THR	conflict	UNP A0A413AG90

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	Mn	0	0
			4	4		

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	12	6	1	4	1	0	0

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



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

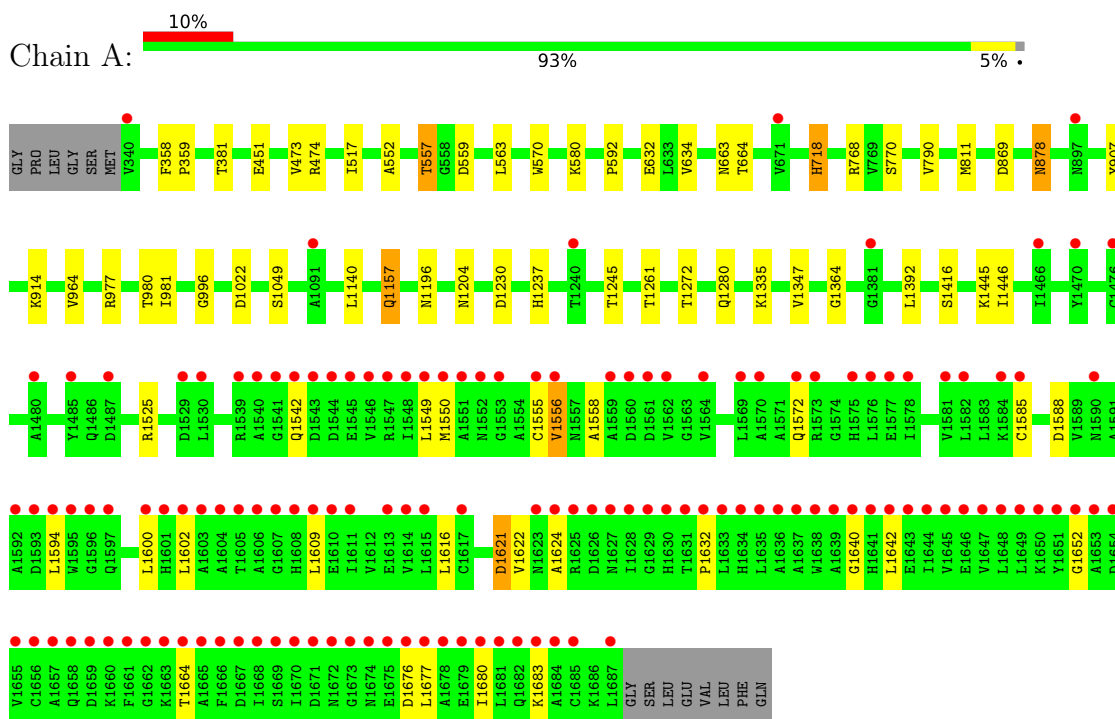
- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	1013	Total 1013	O 1013	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: PEGA domain-containing protein, PEGA domain-containing protein, EngBF DARPin fusion 9b 3G124



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	193.47Å 193.47Å 123.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.51 – 2.30 48.37 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.51-2.30) 100.0 (48.37-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.16 (at 2.29Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.171 , 0.204 0.176 , 0.205	Depositor DCC
$R_{free}$ test set	5833 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.0	Xtrriage
Anisotropy	0.253	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 56.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.021 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	11376	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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](#)

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

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	0.48	0/10536	0.67	0/14286

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 [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	10339	0	9926	28	0
2	A	4	0	0	0	0
3	A	12	0	13	0	0
4	A	8	0	14	0	0
5	A	1013	0	0	2	0
All	All	11376	0	9953	28	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 1.

All (28) 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:1621:ASP:HB3	1:A:1624:ALA:HB2	1.85	0.59
1:A:1364:GLY:HA2	1:A:1392:LEU:HD13	1.85	0.58
1:A:570:TRP:CD2	1:A:592:PRO:HB3	2.41	0.56
1:A:964:VAL:HG12	1:A:981:ILE:HG12	1.87	0.55
1:A:474:ARG:HG2	1:A:557:THR:HG21	1.91	0.53
1:A:1600:LEU:HD23	1:A:1632:PRO:HG2	1.91	0.53
1:A:1572:GLN:HB3	1:A:1602:LEU:HB3	1.92	0.51
1:A:473:VAL:HG13	1:A:517:ILE:HD12	1.93	0.49
1:A:1555:CYS:HB3	1:A:1558:ALA:HB2	1.95	0.48
1:A:1196:ASN:HB3	1:A:1204:ASN:HA	1.95	0.48
1:A:1140:LEU:HB2	1:A:1245:THR:HG23	1.97	0.47
1:A:1616:LEU:HD22	1:A:1622:VAL:HG12	1.96	0.47
1:A:1157:GLN:HG3	5:A:1938:HOH:O	2.15	0.46
1:A:1272:THR:HG22	1:A:1445:LYS:HG3	1.98	0.45
1:A:1556:VAL:HG11	1:A:1585:CYS:HB3	1.99	0.44
1:A:552:ALA:HB1	1:A:563:LEU:HD11	2.00	0.44
1:A:878:ASN:ND2	1:A:977:ARG:HH11	2.15	0.44
1:A:1335:LYS:HB2	1:A:1416:SER:HB3	2.00	0.44
1:A:634:VAL:HA	1:A:914:LYS:HG3	1.99	0.43
1:A:1347:VAL:HG21	1:A:1446:ILE:HD13	2.00	0.43
1:A:358:PHE:CG	1:A:359:PRO:HD2	2.55	0.42
1:A:557:THR:O	1:A:557:THR:HG23	2.19	0.42
1:A:718:HIS:CD2	1:A:718:HIS:C	2.93	0.42
1:A:664:THR:HA	1:A:907:TYR:OH	2.19	0.41
1:A:1640:GLY:HA2	1:A:1677:LEU:HD11	2.01	0.41
1:A:869:ASP:O	1:A:996:GLY:HA3	2.21	0.41
1:A:768:ARG:HD2	1:A:811:MET:HE1	2.02	0.40
1:A:980:THR:HG23	5:A:2396:HOH:O	2.22	0.40

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	1347/1362 (99%)	1295 (96%)	50 (4%)	2 (0%)	51 64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	790	VAL
1	A	1652	GLY

### 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	1098/1108 (99%)	1067 (97%)	31 (3%)	43 60

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	381	THR
1	A	451	GLU
1	A	557	THR
1	A	559	ASP
1	A	580	LYS
1	A	632	GLU
1	A	663	ASN
1	A	718	HIS
1	A	770	SER
1	A	878	ASN
1	A	1022	ASP
1	A	1049	SER
1	A	1157	GLN
1	A	1230	ASP
1	A	1237	HIS
1	A	1261	THR
1	A	1280	GLN
1	A	1525	ARG
1	A	1542	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1549	LEU
1	A	1550	MET
1	A	1556	VAL
1	A	1588	ASP
1	A	1594	LEU
1	A	1609	LEU
1	A	1621	ASP
1	A	1642	LEU
1	A	1664	THR
1	A	1676	ASP
1	A	1680	ILE
1	A	1683	LYS

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

Mol	Chain	Res	Type
1	A	492	GLN
1	A	878	ASN
1	A	1277	ASN
1	A	1597	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](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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
4	MPD	A	1706	-	7,7,7	0.65	0	9,10,10	0.49	0
3	MES	A	1705	-	12,12,12	0.74	0	14,16,16	0.45	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
4	MPD	A	1706	-	-	1/5/5/5	-
3	MES	A	1705	-	-	0/6/14/14	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1706	MPD	C1-C2-C3-C4

There are no ring outliers.

No monomer is involved in short contacts.

## 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	1348/1362 (98%)	0.25	135 (10%) <b>7</b> <b>10</b>	45, 62, 145, 246	0

All (135) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1677	LEU	15.9
1	A	1673	GLY	14.0
1	A	1670	ILE	13.8
1	A	1681	LEU	12.2
1	A	1680	ILE	11.3
1	A	1642	LEU	11.1
1	A	1684	ALA	10.1
1	A	1664	THR	9.5
1	A	1666	PHE	8.9
1	A	1674	ASN	8.7
1	A	1651	TYR	8.5
1	A	1548	ILE	8.4
1	A	1661	PHE	8.3
1	A	1683	LYS	7.8
1	A	1628	ILE	7.8
1	A	1647	VAL	7.5
1	A	1676	ASP	7.3
1	A	1632	PRO	7.2
1	A	1594	LEU	7.2
1	A	1665	ALA	7.1
1	A	1675	GLU	7.1
1	A	1668	ILE	7.1
1	A	1682	GLN	6.7
1	A	1595	TRP	6.7
1	A	1669	SER	6.6
1	A	1644	ILE	6.5
1	A	1624	ALA	6.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	1654	ASP	6.3
1	A	1685	CYS	6.2
1	A	1633	LEU	6.2
1	A	1672	ASN	6.2
1	A	1544	ASP	5.9
1	A	1678	ALA	5.9
1	A	1549	LEU	5.8
1	A	1655	VAL	5.8
1	A	1637	ALA	5.7
1	A	1541	GLY	5.7
1	A	1638	TRP	5.7
1	A	1552	ASN	5.7
1	A	1540	ALA	5.6
1	A	340	VAL	5.4
1	A	1657	ALA	5.4
1	A	1667	ASP	5.4
1	A	1671	ASP	5.4
1	A	1659	ASP	5.3
1	A	1578	ILE	5.2
1	A	1679	GLU	5.2
1	A	1658	GLN	5.2
1	A	1662	GLY	5.2
1	A	1600	LEU	5.2
1	A	1543	ASP	5.1
1	A	1547	ARG	5.0
1	A	1635	LEU	5.0
1	A	1603	ALA	4.9
1	A	1575	HIS	4.9
1	A	1663	LYS	4.9
1	A	1629	GLY	4.8
1	A	1625	ARG	4.8
1	A	1581	VAL	4.8
1	A	1608	HIS	4.7
1	A	1653	ALA	4.7
1	A	1582	LEU	4.7
1	A	1597	GLN	4.6
1	A	1636	ALA	4.6
1	A	1650	LYS	4.6
1	A	1555	CYS	4.5
1	A	1623	ASN	4.4
1	A	1611	ILE	4.4
1	A	1609	LEU	4.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	1556	VAL	4.3
1	A	1640	GLY	4.2
1	A	1596	GLY	4.2
1	A	1576	LEU	4.2
1	A	1631	THR	4.2
1	A	1639	ALA	4.2
1	A	1476	CYS	4.1
1	A	1656	CYS	4.1
1	A	1660	LYS	4.0
1	A	1604	ALA	4.0
1	A	1585	CYS	3.9
1	A	1634	HIS	3.9
1	A	1592	ALA	3.9
1	A	1553	GLY	3.8
1	A	1627	ASN	3.8
1	A	1648	LEU	3.8
1	A	1573	ARG	3.8
1	A	1643	GLU	3.7
1	A	1641	HIS	3.7
1	A	1687	LEU	3.7
1	A	1626	ASP	3.7
1	A	1530	LEU	3.6
1	A	1649	LEU	3.6
1	A	1570	ALA	3.6
1	A	1593	ASP	3.5
1	A	1559	ALA	3.5
1	A	1614	VAL	3.5
1	A	1645	VAL	3.4
1	A	1607	GLY	3.4
1	A	1646	GLU	3.4
1	A	1546	VAL	3.4
1	A	1652	GLY	3.3
1	A	1610	GLU	3.3
1	A	1590	ASN	3.2
1	A	1605	THR	3.2
1	A	1577	GLU	3.1
1	A	1584	LYS	3.1
1	A	1381	GLY	3.0
1	A	1466	ILE	2.9
1	A	1545	GLU	2.9
1	A	1617	CYS	2.8
1	A	1529	ASP	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1562	VAL	2.7
1	A	1550	MET	2.7
1	A	1542	GLN	2.6
1	A	1606	ALA	2.6
1	A	1569	LEU	2.6
1	A	1485	TYR	2.5
1	A	1560	ASP	2.5
1	A	1564	VAL	2.5
1	A	1480	ALA	2.4
1	A	1572	GLN	2.4
1	A	1561	ASP	2.4
1	A	1551	ALA	2.4
1	A	1470	TYR	2.3
1	A	1602	LEU	2.3
1	A	897[A]	ASN	2.3
1	A	1630	HIS	2.3
1	A	1601	HIS	2.3
1	A	1613	GLU	2.2
1	A	671	VAL	2.1
1	A	1615	LEU	2.1
1	A	1487	ASP	2.1
1	A	1091	ALA	2.1
1	A	1539	ARG	2.0
1	A	1240	THR	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
4	MPD	A	1706	8/8	0.89	0.24	103,106,107,107	0
3	MES	A	1705	12/12	0.96	0.14	67,89,101,103	0
2	MN	A	1704	1/1	0.98	0.04	74,74,74,74	0
2	MN	A	1702	1/1	0.99	0.06	66,66,66,66	0
2	MN	A	1701	1/1	0.99	0.07	67,67,67,67	0
2	MN	A	1703	1/1	1.00	0.11	68,68,68,68	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.