



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 14, 2020 – 04:24 pm BST

PDB ID : 2PQF  
Title : Human Poly(ADP-Ribose) Polymerase 12, Catalytic fragment in complex with an inhibitor 3-Aminobenzoic acid  
Authors : Karlberg, T.; Lehtio, L.; Arrowsmith, C.H.; Berglund, H.; Busam, R.D.; Collins, R.; Dahlgren, L.G.; Edwards, A.; Flodin, S.; Flores, A.; Graslund, S.; Hammarstrom, M.; Hogbom, M.; Johansson, I.; Kallas, A.; Kotenyova, T.; Moche, M.; Nordlund, P.; Nyman, T.; Persson, C.; Sagemark, J.; Sundstrom, M.; Thorsell, A.G.; Van Den Berg, S.; Weigelt, J.; Holmberg-Schiavone, L.; Structural Genomics Consortium (SGC)  
Deposited on : 2007-05-02  
Resolution : 2.20 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 : 2.11  
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) : 2.11

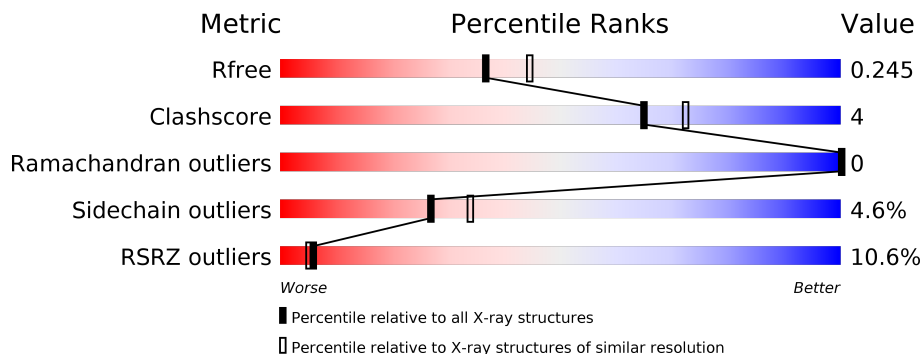
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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.



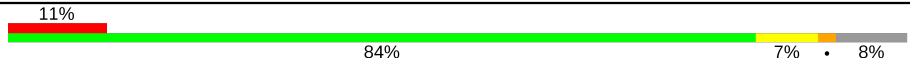
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	198	 11% 76% 16% • 7%
1	B	198	 8% 81% 10% • 9%
1	C	198	 10% 84% 9% • 7%
1	D	198	 8% 81% 10% • 8%
1	E	198	 11% 82% 10% • 7%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	198	 A horizontal bar chart showing the quality of chain. The bar is divided into four segments: a red segment on the left labeled '11%', a large green segment labeled '84%', a yellow segment labeled '7%', and a small grey segment on the right labeled '8%'. The segments are separated by thin white lines.

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9471 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 Poly [ADP-ribose] polymerase 12.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	185	1519	976	260	278	3	2	0	1	0
1	B	181	1479	950	252	272	3	2	0	0	0
1	C	185	1513	972	259	277	3	2	0	0	0
1	D	182	1498	967	254	272	3	2	0	1	0
1	E	185	1513	972	259	277	3	2	0	0	0
1	F	183	1493	961	254	273	3	2	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

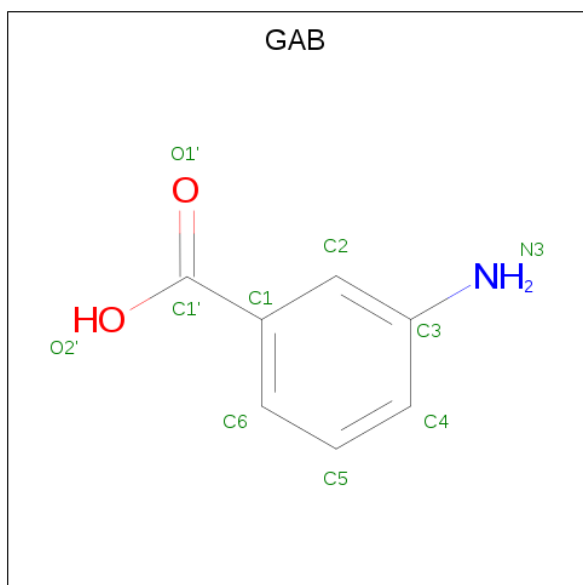
Chain	Residue	Modelled	Actual	Comment	Reference
A	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9
A	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
A	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
A	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
B	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9
B	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
B	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
B	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
C	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9
C	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
C	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
C	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
D	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9
D	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
D	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
D	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
E	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9

*Continued on next page...*

Continued from previous page...

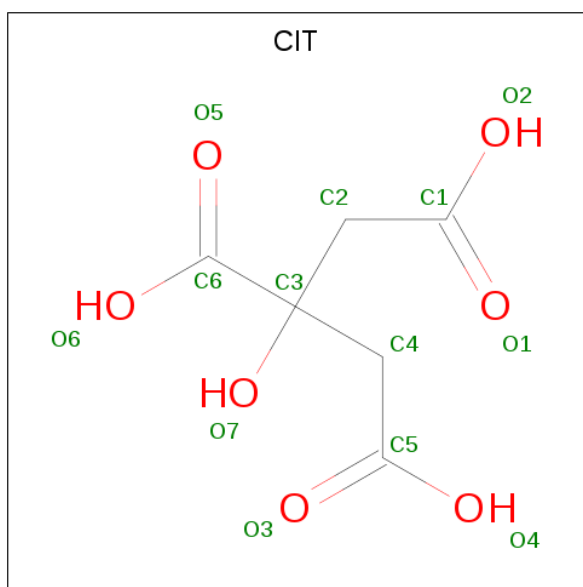
Chain	Residue	Modelled	Actual	Comment	Reference
E	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
E	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
E	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
F	487	SER	-	CLONING ARTIFACT	UNP Q9H0J9
F	488	MSE	-	CLONING ARTIFACT	UNP Q9H0J9
F	548	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9
F	617	MSE	MET	MODIFIED RESIDUE	UNP Q9H0J9

- Molecule 2 is 3-AMINOBENZOIC ACID (three-letter code: GAB) (formula: C<sub>7</sub>H<sub>7</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	10	7	1	2	0	0
2	B	1	10	7	1	2	0	0
2	C	1	10	7	1	2	0	0
2	D	1	10	7	1	2	0	0
2	E	1	10	7	1	2	0	0
2	F	1	10	7	1	2	0	0

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	C O	0	0
			13	6 7		
3	B	1	Total	C O	0	0
			13	6 7		
3	C	1	Total	C O	0	0
			13	6 7		
3	D	1	Total	C O	0	0
			13	6 7		
3	E	1	Total	C O	0	0
			13	6 7		
3	F	1	Total	C O	0	0
			13	6 7		

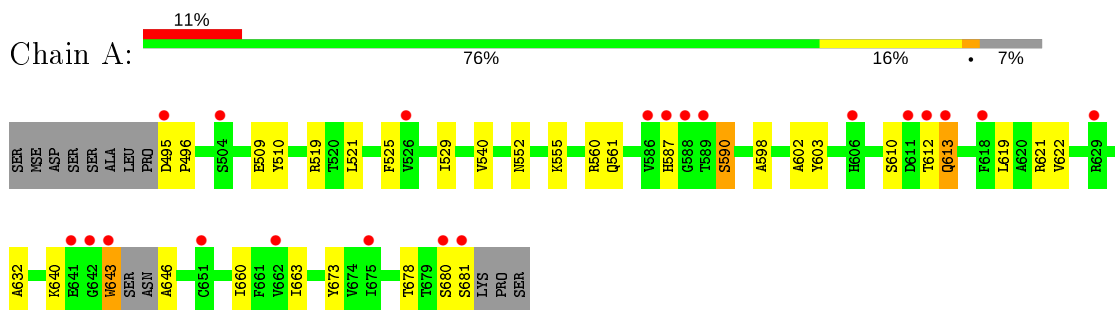
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	66	Total	O	0	0
			66	66		
4	B	51	Total	O	0	0
			51	51		
4	C	62	Total	O	0	0
			62	62		
4	D	42	Total	O	0	0
			42	42		
4	E	56	Total	O	0	0
			56	56		
4	F	41	Total	O	0	0
			41	41		

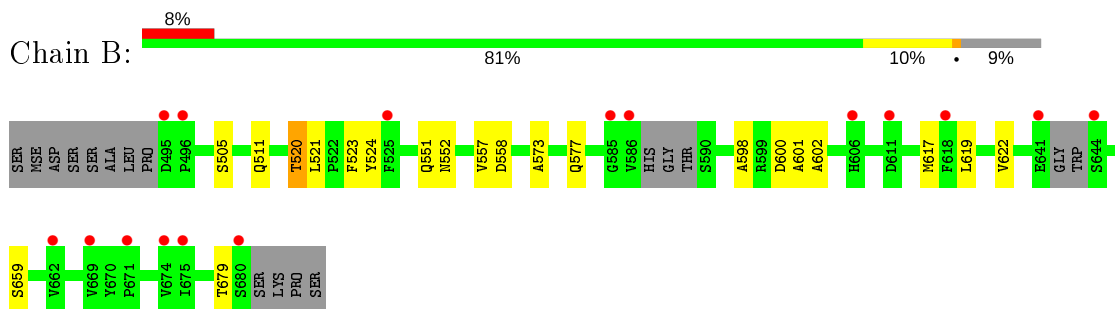
### 3 Residue-property plots [i](#)

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.

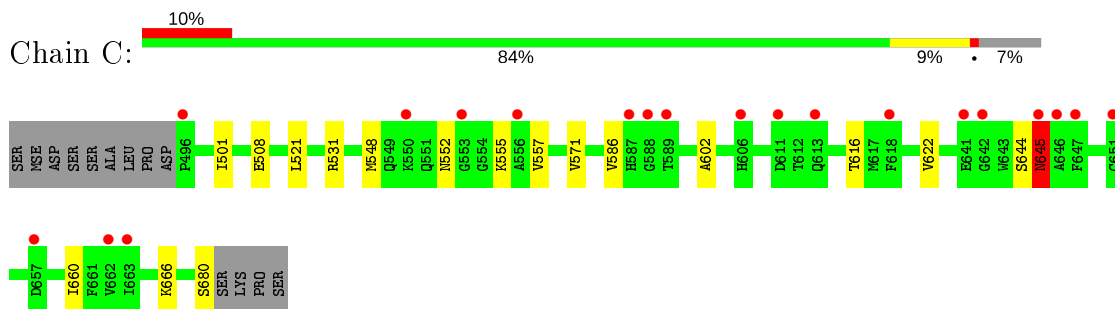
- Molecule 1: Poly [ADP-ribose] polymerase 12



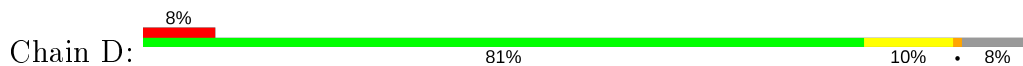
- Molecule 1: Poly [ADP-ribose] polymerase 12

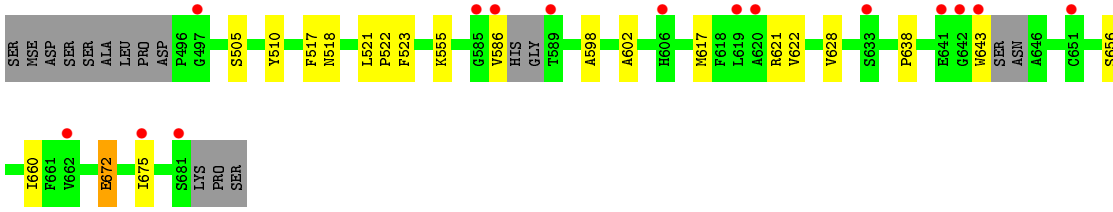


- Molecule 1: Poly [ADP-ribose] polymerase 12

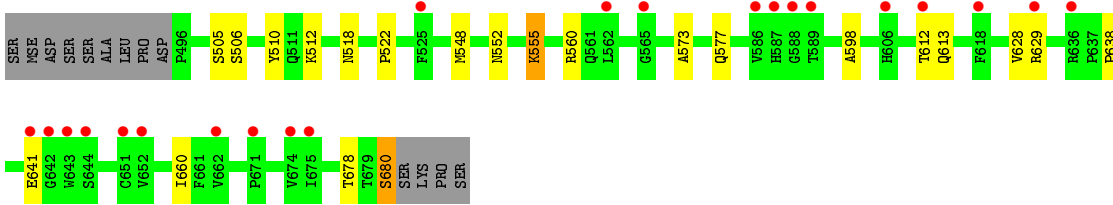
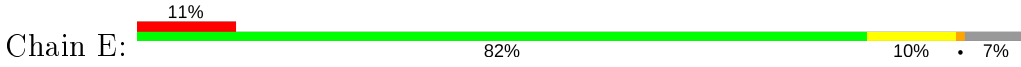


- Molecule 1: Poly [ADP-ribose] polymerase 12

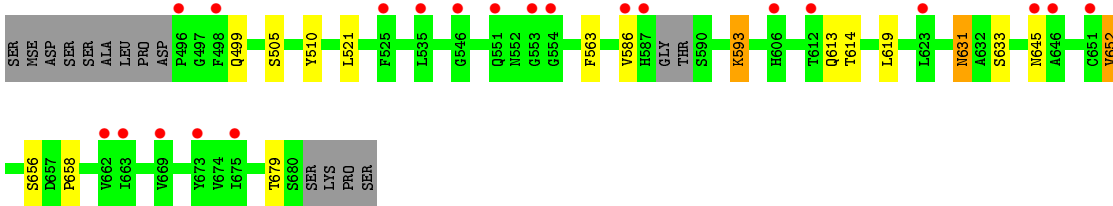
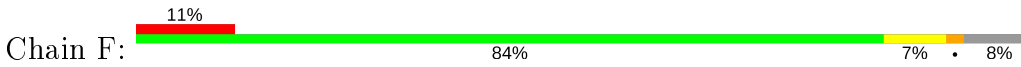




● Molecule 1: Poly [ADP-ribose] polymerase 12



● Molecule 1: Poly [ADP-ribose] polymerase 12





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	206.59Å 206.59Å 84.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.60 – 2.20 19.60 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.60-2.20) 100.0 (19.60-2.20)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 2.19Å)	Xtrriage
Refinement program	REFMAC 5.3.0032	Depositor
R, $R_{free}$	0.198 , 0.244 0.199 , 0.245	Depositor DCC
$R_{free}$ test set	1810 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.3	Xtrriage
Anisotropy	0.485	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 59.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtrriage
Estimated twinning fraction	0.044 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9471	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.51% 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: GAB, CIT

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	0.83	0/1562	0.76	1/2112 (0.0%)
1	B	0.80	1/1518 (0.1%)	0.79	1/2050 (0.0%)
1	C	0.77	0/1557	0.75	1/2106 (0.0%)
1	D	0.86	2/1540 (0.1%)	0.77	1/2080 (0.0%)
1	E	0.85	0/1557	0.77	1/2106 (0.0%)
1	F	0.67	0/1536	0.74	0/2078
All	All	0.80	3/9270 (0.0%)	0.76	5/12532 (0.0%)

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
1	C	1	1
All	All	1	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	622	VAL	CB-CG2	8.44	1.70	1.52
1	B	622	VAL	CB-CG2	6.68	1.66	1.52
1	D	672	GLU	CB-CG	-5.16	1.42	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	560	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	C	531	ARG	NE-CZ-NH2	5.17	122.88	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	558	ASP	CB-CG-OD1	5.12	122.91	118.30
1	D	621	ARG	NE-CZ-NH2	-5.12	117.74	120.30
1	E	560	ARG	NE-CZ-NH2	-5.03	117.78	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	645	ASN	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	680	SER	Peptide
1	C	645	ASN	Peptide

## 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	1519	0	1434	16	0
1	B	1479	0	1402	8	0
1	C	1513	0	1433	10	0
1	D	1498	0	1419	11	0
1	E	1513	0	1433	13	0
1	F	1493	0	1400	8	0
2	A	10	0	6	2	0
2	B	10	0	6	1	0
2	C	10	0	6	1	0
2	D	10	0	6	3	0
2	E	10	0	6	2	0
2	F	10	0	6	0	0
3	A	13	0	5	1	0
3	B	13	0	5	1	0
3	C	13	0	5	1	0
3	D	13	0	5	2	0
3	E	13	0	5	1	0
3	F	13	0	5	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	66	0	0	0	0
4	B	51	0	0	0	0
4	C	62	0	0	0	0
4	D	42	0	0	0	0
4	E	56	0	0	0	0
4	F	41	0	0	0	0
All	All	9471	0	8587	72	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 4.

All (72) 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:C:644:SER:O	1:C:645:ASN:HB3	1.47	1.13
1:B:520:THR:HG21	1:B:600:ASP:HB3	1.65	0.77
1:B:598:ALA:HB2	2:B:701:GAB:HN31	1.52	0.75
1:D:517:PHE:HD1	1:D:617:MSE:HE1	1.51	0.74
1:D:598:ALA:HB2	2:D:701:GAB:HN31	1.53	0.73
1:D:517:PHE:CD1	1:D:617:MSE:HE1	2.24	0.72
3:D:801:CIT:O1	3:D:801:CIT:C6	2.39	0.70
1:A:612:THR:HG22	1:A:613:GLN:H	1.59	0.67
1:A:612:THR:HG22	1:A:613:GLN:N	2.12	0.65
1:F:652:VAL:HG21	1:F:658:PRO:HB3	1.80	0.64
1:E:573:ALA:O	1:E:577:GLN:HG2	2.01	0.61
3:C:801:CIT:C6	3:C:801:CIT:O3	2.49	0.59
1:F:586:VAL:HG11	1:F:593:LYS:HG3	1.83	0.59
1:E:612:THR:O	1:E:613:GLN:HB2	2.02	0.59
1:C:571:VAL:CG2	1:C:616:THR:HG21	2.33	0.59
1:F:586:VAL:CG1	1:F:593:LYS:HG3	2.33	0.59
1:E:598:ALA:HB2	2:E:701:GAB:HN31	1.69	0.57
3:E:801:CIT:O5	3:E:801:CIT:C1	2.49	0.57
1:E:678:THR:HG22	1:E:680:SER:H	1.70	0.56
1:B:505:SER:O	1:B:511:GLN:NE2	2.38	0.55
1:A:552:ASN:ND2	1:A:555:LYS:O	2.39	0.55
1:A:525:PHE:HB2	1:A:678:THR:HG22	1.88	0.54
1:A:590:SER:HB3	1:A:632:ALA:HA	1.89	0.53
1:C:501:ILE:N	1:C:501:ILE:HD12	2.23	0.53
1:C:552:ASN:ND2	1:C:555:LYS:O	2.41	0.53
1:C:644:SER:O	1:C:645:ASN:CB	2.31	0.53
1:A:598:ALA:HB2	2:A:701:GAB:HN31	1.74	0.52

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:617:MSE:HE3	1:D:675:ILE:HD12	1.93	0.51
1:E:678:THR:HG22	1:E:680:SER:N	2.25	0.51
1:F:631:ASN:HD22	1:F:633:SER:H	1.58	0.51
1:A:660:ILE:HG12	2:A:701:GAB:HN32	1.76	0.50
3:B:801:CIT:O4	3:B:801:CIT:C6	2.61	0.49
1:C:571:VAL:HG21	1:C:616:THR:HG21	1.93	0.49
1:E:548:MSE:O	1:E:552:ASN:HB2	2.13	0.48
1:D:660:ILE:HG12	2:D:701:GAB:HN32	1.79	0.48
1:F:652:VAL:CG2	1:F:658:PRO:HB3	2.43	0.48
1:A:509:GLU:OE2	1:A:673:TYR:OH	2.26	0.47
1:F:613:GLN:HB3	1:F:679:THR:HG22	1.98	0.46
1:A:612:THR:CG2	1:A:613:GLN:N	2.79	0.46
1:E:628:VAL:HG11	1:E:638:PRO:HG2	1.97	0.46
1:E:518:ASN:ND2	1:E:522:PRO:HA	2.31	0.46
3:A:801:CIT:C6	3:A:801:CIT:O3	2.64	0.45
1:A:521:LEU:HD21	1:A:602:ALA:HA	1.98	0.45
1:A:561:GLN:HB3	1:A:619:LEU:HD11	1.98	0.45
1:B:573:ALA:O	1:B:577:GLN:HG2	2.17	0.45
1:D:617:MSE:HB3	1:D:617:MSE:HE2	1.83	0.45
1:C:521:LEU:HD21	1:C:602:ALA:HA	1.98	0.44
1:C:660:ILE:HG12	2:C:701:GAB:HN32	1.81	0.44
1:A:510:TYR:CD1	1:A:529:ILE:HD12	2.52	0.44
1:E:612:THR:O	1:E:613:GLN:CB	2.64	0.44
1:E:552:ASN:ND2	1:E:555:LYS:HB3	2.32	0.43
1:A:643:TRP:CE3	1:A:646:ALA:CB	3.01	0.43
1:A:495:ASP:CB	1:A:496:PRO:CD	2.97	0.42
1:D:521:LEU:HD21	1:D:602:ALA:HA	2.01	0.42
1:E:505:SER:HA	1:E:510:TYR:CG	2.54	0.42
3:D:801:CIT:O6	3:D:801:CIT:C1	2.64	0.42
1:C:552:ASN:HD21	1:C:555:LYS:HG2	1.85	0.42
1:A:622:VAL:HG11	1:A:663:ILE:HD11	2.02	0.42
1:A:603:TYR:CZ	1:A:660:ILE:HD11	2.55	0.41
1:B:520:THR:CG2	1:B:600:ASP:HB3	2.42	0.41
1:D:505:SER:HA	1:D:510:TYR:CG	2.56	0.41
1:D:518:ASN:ND2	1:D:522:PRO:HA	2.36	0.41
1:F:505:SER:HA	1:F:510:TYR:CG	2.55	0.41
1:B:521:LEU:HD21	1:B:602:ALA:HA	2.02	0.41
1:E:628:VAL:HG11	1:E:638:PRO:CG	2.51	0.41
1:B:601:ALA:O	1:B:617:MSE:HE1	2.21	0.41
1:C:548:MSE:O	1:C:552:ASN:HB3	2.21	0.41
1:D:628:VAL:HG11	1:D:638:PRO:HG2	2.03	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:660:ILE:HG12	2:E:701:GAB:HN32	1.86	0.40
1:B:524:TYR:CZ	1:B:679:THR:HG22	2.56	0.40
1:D:598:ALA:HB2	2:D:701:GAB:N3	2.30	0.40
1:F:563:PHE:CE1	1:F:619:LEU:HD13	2.57	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	182/198 (92%)	179 (98%)	3 (2%)	0	100	100
1	B	175/198 (88%)	173 (99%)	2 (1%)	0	100	100
1	C	183/198 (92%)	178 (97%)	5 (3%)	0	100	100
1	D	177/198 (89%)	171 (97%)	6 (3%)	0	100	100
1	E	183/198 (92%)	180 (98%)	3 (2%)	0	100	100
1	F	179/198 (90%)	177 (99%)	2 (1%)	0	100	100
All	All	1079/1188 (91%)	1058 (98%)	21 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	162/171 (95%)	152 (94%)	10 (6%)	18	21
1	B	159/171 (93%)	152 (96%)	7 (4%)	28	35
1	C	162/171 (95%)	155 (96%)	7 (4%)	29	36
1	D	160/171 (94%)	154 (96%)	6 (4%)	33	42
1	E	162/171 (95%)	156 (96%)	6 (4%)	34	43
1	F	158/171 (92%)	150 (95%)	8 (5%)	24	29
All	All	963/1026 (94%)	919 (95%)	44 (5%)	27	34

All (44) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	519	ARG
1	A	540	VAL
1	A	587	HIS
1	A	590	SER
1	A	610	SER
1	A	613	GLN
1	A	621	ARG
1	A	640	LYS
1	A	643	TRP
1	A	681	SER
1	B	520	THR
1	B	523	PHE
1	B	551	GLN
1	B	552	ASN
1	B	557	VAL
1	B	619	LEU
1	B	659	SER
1	C	508	GLU
1	C	557	VAL
1	C	586	VAL
1	C	622	VAL
1	C	645	ASN
1	C	666	LYS
1	C	680	SER
1	D	523	PHE
1	D	555	LYS
1	D	586	VAL
1	D	643	TRP
1	D	656	SER
1	D	672	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	506	SER
1	E	512	LYS
1	E	555	LYS
1	E	629	ARG
1	E	641	GLU
1	E	680	SER
1	F	499	GLN
1	F	521	LEU
1	F	593	LYS
1	F	614	THR
1	F	631	ASN
1	F	645	ASN
1	F	652	VAL
1	F	656	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	613	GLN
1	B	552	ASN
1	B	615	HIS
1	C	515	ASN
1	D	518	ASN
1	E	518	ASN
1	F	527	GLN
1	F	576	GLN
1	F	577	GLN
1	F	631	ASN
1	F	676	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

12 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	GAB	E	701	-	8,10,10	0.95	1 (12%)	10,13,13	1.88	3 (30%)
3	CIT	C	801	-	3,12,12	2.23	1 (33%)	3,17,17	1.93	1 (33%)
2	GAB	A	701	-	8,10,10	1.18	0	10,13,13	1.89	4 (40%)
3	CIT	F	801	-	3,12,12	1.50	0	3,17,17	3.21	2 (66%)
3	CIT	D	801	-	3,12,12	0.87	0	3,17,17	2.86	2 (66%)
2	GAB	C	701	-	8,10,10	1.32	1 (12%)	10,13,13	1.94	2 (20%)
3	CIT	A	801	-	3,12,12	2.56	1 (33%)	3,17,17	2.45	1 (33%)
3	CIT	B	801	-	3,12,12	2.35	1 (33%)	3,17,17	2.07	1 (33%)
2	GAB	F	701	-	8,10,10	1.06	0	10,13,13	1.86	2 (20%)
2	GAB	D	701	-	8,10,10	1.70	2 (25%)	10,13,13	1.65	3 (30%)
2	GAB	B	701	-	8,10,10	1.12	1 (12%)	10,13,13	2.07	5 (50%)
3	CIT	E	801	-	3,12,12	1.87	2 (66%)	3,17,17	8.96	3 (100%)

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	GAB	E	701	-	-	0/0/4/4	0/1/1/1
3	CIT	C	801	-	-	3/6/16/16	-
2	GAB	A	701	-	-	0/0/4/4	0/1/1/1
3	CIT	F	801	-	-	2/6/16/16	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	CIT	D	801	-	-	6/6/16/16	-
2	GAB	C	701	-	-	0/0/4/4	0/1/1/1
3	CIT	A	801	-	-	3/6/16/16	-
3	CIT	B	801	-	-	3/6/16/16	-
2	GAB	F	701	-	-	0/0/4/4	0/1/1/1
2	GAB	D	701	-	-	0/0/4/4	0/1/1/1
2	GAB	B	701	-	-	0/0/4/4	0/1/1/1
3	CIT	E	801	-	-	4/6/16/16	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	801	CIT	O7-C3	4.30	1.49	1.43
2	D	701	GAB	C1-C1'	4.15	1.51	1.47
3	B	801	CIT	O7-C3	3.95	1.49	1.43
3	C	801	CIT	O7-C3	3.79	1.49	1.43
3	E	801	CIT	C2-C3	2.34	1.58	1.54
2	C	701	GAB	C1-C1'	2.33	1.49	1.47
2	B	701	GAB	C3-N3	-2.16	1.30	1.38
2	D	701	GAB	C3-N3	-2.07	1.30	1.38
3	E	801	CIT	O7-C3	2.06	1.46	1.43
2	E	701	GAB	C3-N3	-2.04	1.30	1.38

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	801	CIT	C4-C3-C2	13.11	144.38	109.33
3	E	801	CIT	C3-C4-C5	-7.24	103.39	114.98
3	F	801	CIT	C3-C4-C5	-4.76	107.36	114.98
3	E	801	CIT	C3-C2-C1	-4.08	108.45	114.98
2	E	701	GAB	C1-C2-C3	-4.06	117.15	120.66
3	D	801	CIT	C4-C3-C2	4.03	120.11	109.33
3	A	801	CIT	C3-C4-C5	3.99	121.37	114.98
2	C	701	GAB	C1-C2-C3	-3.84	117.34	120.66
2	C	701	GAB	C4-C3-C2	3.80	123.56	118.62
2	F	701	GAB	C4-C3-C2	3.65	123.36	118.62
3	B	801	CIT	C3-C4-C5	3.54	120.65	114.98
2	F	701	GAB	C1-C2-C3	-3.47	117.66	120.66
2	B	701	GAB	C1-C2-C3	-3.36	117.76	120.66
3	C	801	CIT	C3-C4-C5	3.34	120.34	114.98
2	D	701	GAB	C4-C3-C2	3.32	122.94	118.62

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	GAB	C4-C3-C2	3.25	122.85	118.62
2	B	701	GAB	C4-C3-C2	3.16	122.72	118.62
2	E	701	GAB	C4-C3-C2	3.03	122.56	118.62
2	A	701	GAB	C1-C2-C3	-2.87	118.18	120.66
3	F	801	CIT	C4-C3-C2	2.86	116.98	109.33
2	B	701	GAB	C2-C1-C1'	-2.86	116.60	120.36
2	A	701	GAB	C2-C3-N3	-2.73	115.85	120.57
3	D	801	CIT	C3-C4-C5	-2.60	110.82	114.98
2	D	701	GAB	C1-C2-C3	-2.51	118.49	120.66
2	B	701	GAB	C2-C3-N3	-2.45	116.34	120.57
2	E	701	GAB	C6-C1-C2	2.25	121.34	118.16
2	B	701	GAB	C6-C1-C2	2.10	121.14	118.16
2	A	701	GAB	C6-C1-C2	2.07	121.09	118.16
2	D	701	GAB	C2-C3-N3	-2.02	117.08	120.57

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	801	CIT	C6-C3-C4-C5
3	D	801	CIT	C1-C2-C3-C6
3	D	801	CIT	C2-C3-C4-C5
3	D	801	CIT	O7-C3-C4-C5
3	D	801	CIT	C6-C3-C4-C5
3	A	801	CIT	C6-C3-C4-C5
3	B	801	CIT	C6-C3-C4-C5
3	E	801	CIT	C1-C2-C3-O7
3	E	801	CIT	C1-C2-C3-C6
3	E	801	CIT	C6-C3-C4-C5
3	C	801	CIT	O7-C3-C4-C5
3	A	801	CIT	O7-C3-C4-C5
3	C	801	CIT	C2-C3-C4-C5
3	A	801	CIT	C2-C3-C4-C5
3	D	801	CIT	C1-C2-C3-C4
3	B	801	CIT	C2-C3-C4-C5
3	B	801	CIT	O7-C3-C4-C5
3	E	801	CIT	O7-C3-C4-C5
3	F	801	CIT	C1-C2-C3-C6
3	F	801	CIT	C1-C2-C3-O7
3	D	801	CIT	C1-C2-C3-O7

There are no ring outliers.

10 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	701	GAB	2	0
3	C	801	CIT	1	0
2	A	701	GAB	2	0
3	D	801	CIT	2	0
2	C	701	GAB	1	0
3	A	801	CIT	1	0
3	B	801	CIT	1	0
2	D	701	GAB	3	0
2	B	701	GAB	1	0
3	E	801	CIT	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

### 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	183/198 (92%)	0.55	21 (11%) 4 4	46, 55, 68, 81	0
1	B	179/198 (90%)	0.48	16 (8%) 9 8	45, 55, 64, 67	0
1	C	183/198 (92%)	0.56	20 (10%) 5 5	47, 56, 65, 72	0
1	D	180/198 (90%)	0.49	15 (8%) 11 10	44, 55, 64, 69	0
1	E	183/198 (92%)	0.52	22 (12%) 4 3	44, 55, 65, 69	0
1	F	181/198 (91%)	0.54	21 (11%) 4 4	47, 56, 65, 71	0
All	All	1089/1188 (91%)	0.52	115 (10%) 6 5	44, 55, 65, 81	0

All (115) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	553	GLY	5.8
1	A	641	GLU	5.6
1	A	587	HIS	5.6
1	B	680	SER	5.6
1	C	588	GLY	5.2
1	A	589	THR	5.1
1	E	589	THR	5.1
1	A	642	GLY	5.1
1	D	681	SER	5.0
1	E	588	GLY	4.9
1	D	589	THR	4.8
1	A	643	TRP	4.8
1	D	643	TRP	4.7
1	E	587	HIS	4.5
1	E	644	SER	4.5
1	C	645	ASN	4.5
1	B	641	GLU	4.2
1	B	675	ILE	4.2
1	A	675	ILE	4.1

*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	588	GLY	4.1
1	E	643	TRP	4.0
1	F	675	ILE	4.0
1	F	535	LEU	4.0
1	C	496	PRO	4.0
1	F	645	ASN	3.9
1	A	680	SER	3.9
1	E	586	VAL	3.8
1	E	675	ILE	3.7
1	F	587	HIS	3.6
1	B	606	HIS	3.6
1	C	662	VAL	3.6
1	D	633	SER	3.6
1	A	606	HIS	3.5
1	F	496	PRO	3.4
1	E	641	GLU	3.4
1	A	613	GLN	3.3
1	A	662	VAL	3.3
1	A	681	SER	3.3
1	F	586	VAL	3.3
1	F	662	VAL	3.3
1	B	586	VAL	3.3
1	F	651	CYS	3.3
1	A	495	ASP	3.3
1	E	642	GLY	3.2
1	E	606	HIS	3.2
1	B	674	VAL	3.2
1	D	662	VAL	3.1
1	F	553	GLY	3.1
1	A	612	THR	3.1
1	B	585	GLY	3.1
1	F	646	ALA	3.0
1	B	618	PHE	3.0
1	D	606	HIS	3.0
1	C	646	ALA	3.0
1	C	647	PHE	3.0
1	B	495	ASP	3.0
1	B	662	VAL	3.0
1	D	675	ILE	3.0
1	A	618	PHE	2.9
1	D	642	GLY	2.9
1	E	652	VAL	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	E	674	VAL	2.8
1	F	663	ILE	2.8
1	F	551	GLN	2.7
1	C	589	THR	2.7
1	A	586	VAL	2.7
1	C	657	ASP	2.7
1	B	644	SER	2.6
1	D	585	GLY	2.6
1	F	606	HIS	2.6
1	C	550	LYS	2.6
1	D	651	CYS	2.6
1	E	651	CYS	2.6
1	C	618	PHE	2.6
1	C	613	GLN	2.6
1	B	611	ASP	2.5
1	D	620	ALA	2.5
1	C	641	GLU	2.5
1	F	554	GLY	2.5
1	C	587	HIS	2.4
1	D	497	GLY	2.4
1	E	618	PHE	2.4
1	C	611	ASP	2.4
1	E	662	VAL	2.4
1	C	606	HIS	2.4
1	F	612	THR	2.4
1	A	629	ARG	2.4
1	D	619	LEU	2.3
1	E	565	GLY	2.3
1	C	663	ILE	2.3
1	F	525	PHE	2.3
1	B	671	PRO	2.3
1	A	504	SER	2.3
1	C	556	ALA	2.3
1	C	651	CYS	2.2
1	E	525	PHE	2.2
1	F	669	VAL	2.2
1	B	669	VAL	2.2
1	E	629	ARG	2.2
1	A	651	CYS	2.2
1	D	641	GLU	2.2
1	D	586	VAL	2.1
1	E	612	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	623	LEU	2.1
1	F	673	TYR	2.1
1	E	562	LEU	2.1
1	C	642	GLY	2.1
1	F	498	PHE	2.1
1	E	671	PRO	2.1
1	B	496	PRO	2.1
1	A	611	ASP	2.0
1	E	636	ARG	2.0
1	F	546	GLY	2.0
1	A	526	VAL	2.0
1	B	525	PHE	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](#)

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(Å <sup>2</sup> )	Q<0.9
3	CIT	F	801	13/13	0.62	0.28	86,92,96,96	0
3	CIT	B	801	13/13	0.79	0.22	60,67,70,71	0
2	GAB	B	701	10/10	0.83	0.17	51,57,57,59	0
3	CIT	C	801	13/13	0.84	0.20	60,68,73,74	0
3	CIT	E	801	13/13	0.84	0.26	61,66,68,69	0
3	CIT	D	801	13/13	0.85	0.18	70,76,79,80	0
3	CIT	A	801	13/13	0.88	0.12	59,67,71,72	0
2	GAB	D	701	10/10	0.90	0.19	54,57,59,60	0
2	GAB	E	701	10/10	0.91	0.14	47,54,55,59	0
2	GAB	F	701	10/10	0.92	0.18	52,54,57,60	0
2	GAB	C	701	10/10	0.92	0.16	46,47,49,53	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GAB	A	701	10/10	0.94	0.16	48,54,56,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.