



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 08:46 pm BST

PDB ID : 3SQJ
Title : Recombinant human serum albumin from transgenic plant
Authors : He, Y.; Yang, D.
Deposited on : 2011-07-05
Resolution : 2.05 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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

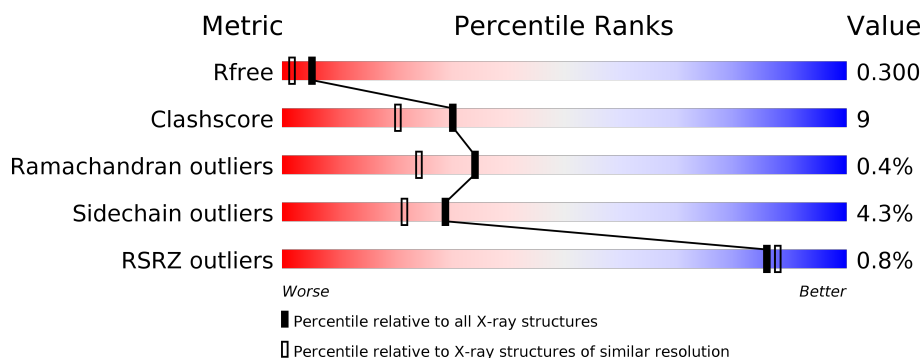
1 Overall quality at a glance ⓘ

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}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (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 ≥ 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	582	<div> <div>%</div> <div> <div></div> <div>79%</div> <div>19%</div> <div>•</div> </div> </div>
1	B	582	<div> <div>%</div> <div> <div></div> <div>81%</div> <div>17%</div> <div>•</div> </div> </div>

2 Entry composition [i](#)

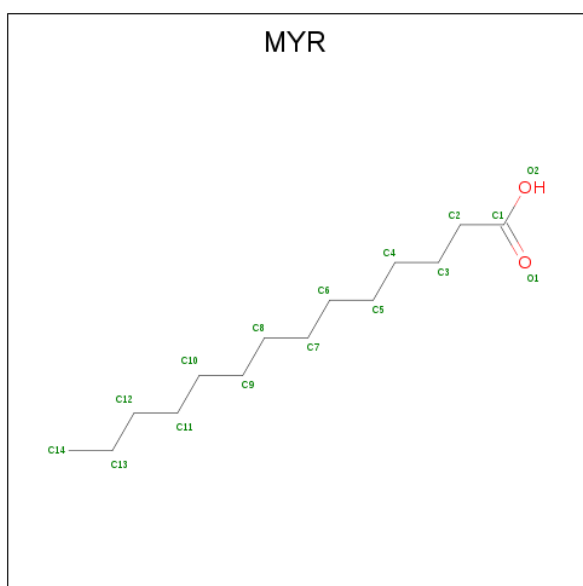
There are 3 unique types of molecules in this entry. The entry contains 9869 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Serum albumin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	582	Total	C	N	O	S	0	0	0
			4557	2880	759	877	41			
1	B	582	Total	C	N	O	S	0	0	0
			4538	2873	759	865	41			

- Molecule 2 is MYRISTIC ACID (three-letter code: MYR) (formula: $C_{14}H_{28}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			14	12	2		
2	A	1	Total	C	O	0	0
			14	12	2		
2	A	1	Total	C	O	0	0
			14	12	2		
2	A	1	Total	C	O	0	0
			13	11	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			16	14	2		
2	A	1	Total	C	O	0	0
			16	14	2		
2	A	1	Total	C	O	0	0
			16	14	2		
2	A	1	Total	C	O	0	0
			16	14	2		
2	B	1	Total	C	O	0	0
			14	12	2		
2	B	1	Total	C	O	0	0
			13	11	2		
2	B	1	Total	C	O	0	0
			14	12	2		
2	B	1	Total	C	O	0	0
			13	11	2		
2	B	1	Total	C	O	0	0
			16	14	2		
2	B	1	Total	C	O	0	0
			15	13	2		
2	B	1	Total	C	O	0	0
			14	12	2		

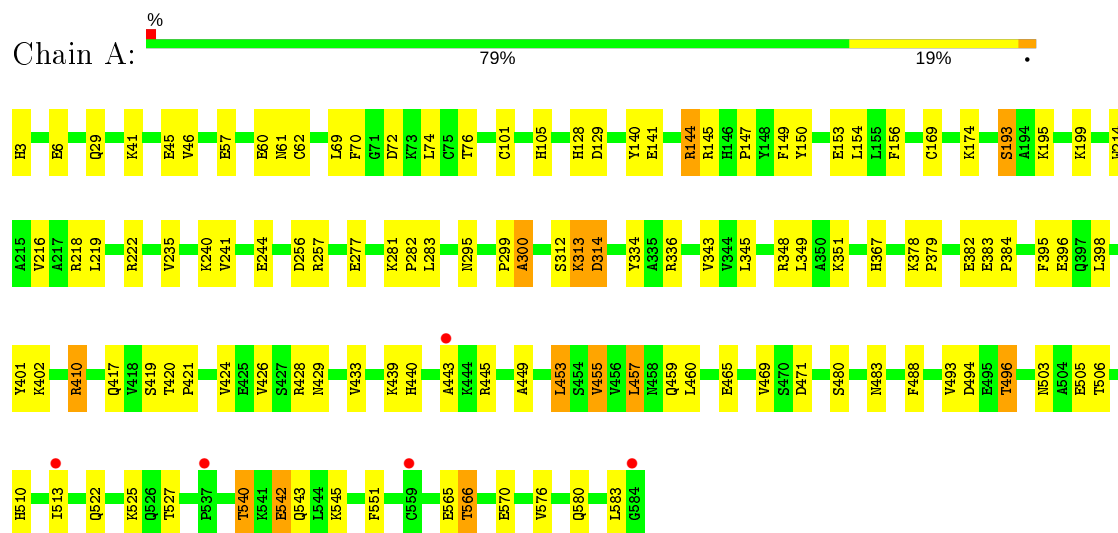
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	270	Total	O	0	0
			270	270		
3	B	286	Total	O	0	0
			286	286		

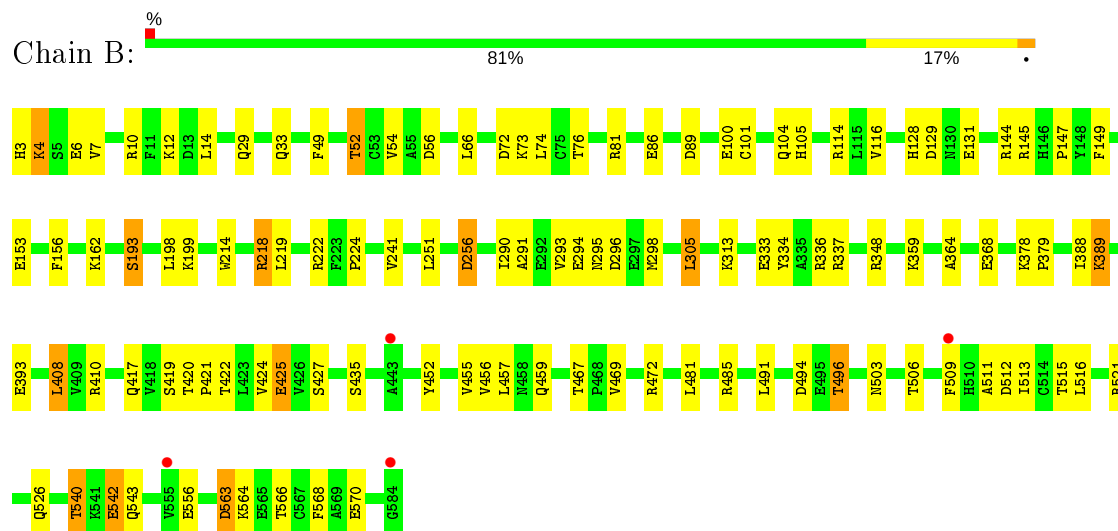
3 Residue-property plots

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: Serum albumin



• Molecule 1: Serum albumin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	95.61Å 38.37Å 184.04Å 90.00° 104.93° 90.00°	Depositor
Resolution (Å)	47.80 – 2.05 47.80 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.5 (47.80-2.05) 99.5 (47.80-2.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.44 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.235 , 0.303 0.234 , 0.300	Depositor DCC
R_{free} test set	4112 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 41.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9869	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.35 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.5503e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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: MYR

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	0.99	3/4647 (0.1%)	0.90	8/6288 (0.1%)
1	B	1.00	1/4628 (0.0%)	0.93	12/6260 (0.2%)
All	All	0.99	4/9275 (0.0%)	0.92	20/12548 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	6	GLU	CG-CD	7.41	1.63	1.51
1	A	62	CYS	CB-SG	-7.26	1.70	1.82
1	A	6	GLU	CG-CD	5.70	1.60	1.51
1	A	150	TYR	CD1-CE1	5.09	1.47	1.39

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	144	ARG	NE-CZ-NH2	-11.61	114.50	120.30
1	A	144	ARG	NE-CZ-NH2	-10.13	115.24	120.30
1	A	144	ARG	NE-CZ-NH1	9.15	124.88	120.30
1	B	348	ARG	NE-CZ-NH2	-8.78	115.91	120.30
1	B	472	ARG	NE-CZ-NH2	-8.02	116.29	120.30

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	4557	0	4354	89	0
1	B	4538	0	4345	83	0
2	A	119	0	186	14	0
2	B	99	0	145	3	0
3	A	270	0	0	9	0
3	B	286	0	0	14	0
All	All	9869	0	9030	170	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 9.

The worst 5 of 170 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:257:ARG:HH11	2:A:1007:MYR:H32	0.99	1.10
1:A:257:ARG:NH1	2:A:1007:MYR:H32	1.77	0.99
1:B:511:ALA:HB2	3:B:743:HOH:O	1.63	0.97
1:A:46:VAL:HG22	2:A:1008:MYR:H102	1.49	0.91
1:B:116:VAL:CG1	3:B:691:HOH:O	2.18	0.90

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	580/582 (100%)	558 (96%)	19 (3%)	3 (0%)	29	18
1	B	580/582 (100%)	555 (96%)	23 (4%)	2 (0%)	41	31
All	All	1160/1164 (100%)	1113 (96%)	42 (4%)	5 (0%)	34	24

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	300	ALA
1	B	563	ASP
1	B	564	LYS
1	A	314	ASP
1	A	443	ALA

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	483/509 (95%)	465 (96%)	18 (4%)	34	27
1	B	478/509 (94%)	455 (95%)	23 (5%)	25	18
All	All	961/1018 (94%)	920 (96%)	41 (4%)	29	22

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	4	LYS
1	B	74	LEU
1	B	540	THR
1	B	14	LEU
1	B	52	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	503	ASN
1	B	3	HIS
1	B	429	ASN
1	A	459	GLN
1	B	128	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

15 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	MYR	B	1005	-	12,15,15	0.37	0	11,15,15	0.88	0
2	MYR	B	1004	-	9,12,15	0.32	0	8,12,15	0.70	0
2	MYR	A	1002	-	10,13,15	0.40	0	9,13,15	0.56	0
2	MYR	A	1008	-	12,15,15	0.62	0	11,15,15	0.65	0
2	MYR	A	1001	-	10,13,15	0.53	0	9,13,15	0.82	0
2	MYR	A	1005	-	12,15,15	0.30	0	11,15,15	0.78	0
2	MYR	B	1001	-	10,13,15	0.43	0	9,13,15	1.00	1 (11%)
2	MYR	B	1002	-	9,12,15	0.24	0	8,12,15	0.98	0
2	MYR	A	1004	-	9,12,15	0.51	0	8,12,15	0.84	0
2	MYR	B	1006	-	11,14,15	0.39	0	10,14,15	0.78	0
2	MYR	A	1003	-	10,13,15	0.46	0	9,13,15	0.46	0
2	MYR	B	1003	-	10,13,15	0.62	0	9,13,15	0.42	0
2	MYR	A	1006	-	12,15,15	0.30	0	11,15,15	0.89	0
2	MYR	A	1007	-	12,15,15	0.53	0	11,15,15	0.73	0
2	MYR	B	1007	-	10,13,15	0.60	0	9,13,15	0.68	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	MYR	B	1005	-	-	4/11/13/13	-
2	MYR	B	1004	-	-	5/8/10/13	-
2	MYR	A	1002	-	-	5/9/11/13	-
2	MYR	A	1008	-	-	6/11/13/13	-
2	MYR	A	1001	-	-	6/9/11/13	-
2	MYR	A	1005	-	-	6/11/13/13	-
2	MYR	B	1001	-	-	6/9/11/13	-
2	MYR	B	1002	-	-	3/8/10/13	-
2	MYR	A	1004	-	-	5/8/10/13	-
2	MYR	B	1006	-	-	1/10/12/13	-
2	MYR	A	1003	-	-	4/9/11/13	-
2	MYR	B	1003	-	-	3/9/11/13	-
2	MYR	A	1006	-	-	5/11/13/13	-
2	MYR	A	1007	-	-	8/11/13/13	-
2	MYR	B	1007	-	-	6/9/11/13	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1001	MYR	C9-C8-C7	-2.32	102.66	114.42

There are no chirality outliers.

5 of 73 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1005	MYR	C1-C2-C3-C4
2	B	1004	MYR	C1-C2-C3-C4
2	A	1005	MYR	C1-C2-C3-C4
2	B	1001	MYR	C1-C2-C3-C4
2	B	1007	MYR	C1-C2-C3-C4

There are no ring outliers.

8 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1008	MYR	4	0
2	A	1005	MYR	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1002	MYR	1	0
2	A	1004	MYR	2	0
2	B	1003	MYR	1	0
2	A	1006	MYR	1	0
2	A	1007	MYR	4	0
2	B	1007	MYR	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, 95th 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 ²)	Q<0.9
1	A	582/582 (100%)	-0.06	5 (0%) 84 86	20, 38, 59, 75	0
1	B	582/582 (100%)	-0.16	4 (0%) 87 89	21, 35, 56, 80	0
All	All	1164/1164 (100%)	-0.11	9 (0%) 86 88	20, 36, 58, 80	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	584	GLY	4.8
1	B	509	PHE	4.1
1	B	443	ALA	3.1
1	A	443	ALA	2.8
1	B	584	GLY	2.7

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, 95th 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(\AA^2)	Q<0.9
2	MYR	A	1007	16/16	0.69	0.25	42,49,66,67	0
2	MYR	B	1007	14/16	0.69	0.25	37,50,59,62	0
2	MYR	B	1004	13/16	0.73	0.19	38,40,49,50	0
2	MYR	A	1004	13/16	0.77	0.18	41,46,50,51	0
2	MYR	A	1008	16/16	0.79	0.23	50,53,69,72	0
2	MYR	B	1005	16/16	0.82	0.16	35,39,56,58	0
2	MYR	A	1001	14/16	0.83	0.25	37,52,55,56	0
2	MYR	A	1005	16/16	0.87	0.20	37,46,61,64	0
2	MYR	B	1001	14/16	0.88	0.16	39,45,51,51	0
2	MYR	B	1002	13/16	0.89	0.14	27,39,48,51	0
2	MYR	A	1002	14/16	0.91	0.14	32,38,51,51	0
2	MYR	B	1006	15/16	0.92	0.12	28,33,44,46	0
2	MYR	A	1006	16/16	0.93	0.13	27,32,43,46	0
2	MYR	B	1003	14/16	0.94	0.15	22,35,43,44	0
2	MYR	A	1003	14/16	0.95	0.14	24,37,42,44	0

6.5 Other polymers [i](#)

There are no such residues in this entry.