



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:02 AM GMT

PDB ID : 3D2F  
Title : Crystal structure of a complex of Sse1p and Hsp70  
Authors : Polier, S.; Bracher, A.  
Deposited on : 2008-05-08  
Resolution : 2.30 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

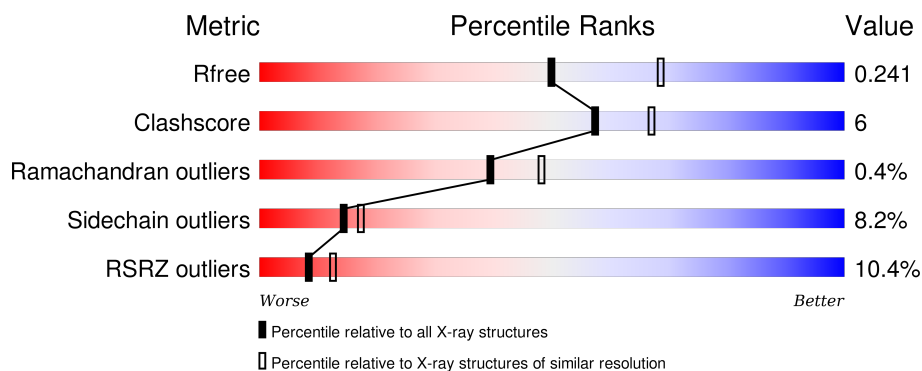
# 1 Overall quality at a glance

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))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (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 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	675	
1	C	675	
2	B	382	
2	D	382	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	GOL	A	3002	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 16199 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 Heat shock protein homolog SSE1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	629	Total	C	N	O	S	0	2	0
			4874	3085	817	960	12			
1	C	627	Total	C	N	O	S	0	1	0
			4828	3050	810	956	12			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	503	ALA	-	LINKER	UNP P32589
A	504	GLY	-	LINKER	UNP P32589
A	505	SER	-	LINKER	UNP P32589
A	506	ASP	-	LINKER	UNP P32589
C	503	ALA	-	LINKER	UNP P32589
C	504	GLY	-	LINKER	UNP P32589
C	505	SER	-	LINKER	UNP P32589
C	506	ASP	-	LINKER	UNP P32589

- Molecule 2 is a protein called Heat shock 70 kDa protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	379	Total	C	N	O	S	0	1	0
			2895	1824	505	559	7			
2	D	379	Total	C	N	O	S	0	3	0
			2865	1806	490	562	7			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	2	Total	Mg	0	0
			2	2		

*Continued on next page...*

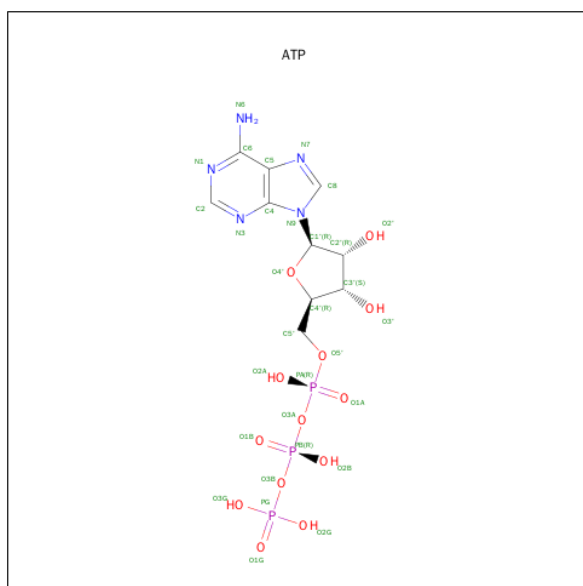
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Mg	0	0
			1	1		
3	C	2	Total	Mg	0	0
			2	2		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	K	0	0
			1	1		
4	C	1	Total	K	0	0
			1	1		

- Molecule 5 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
5	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).

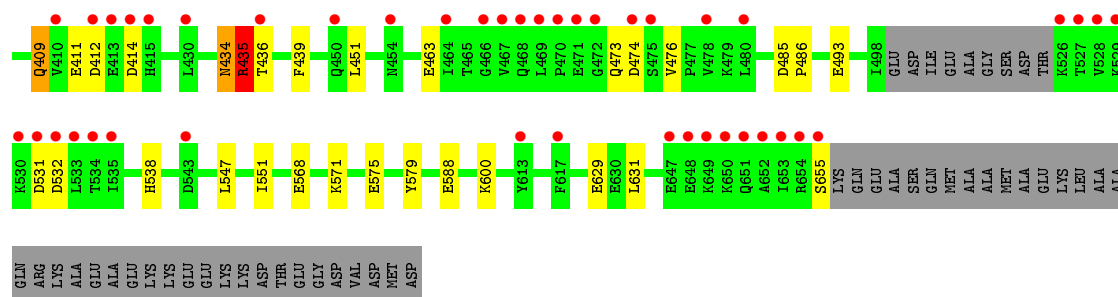


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		

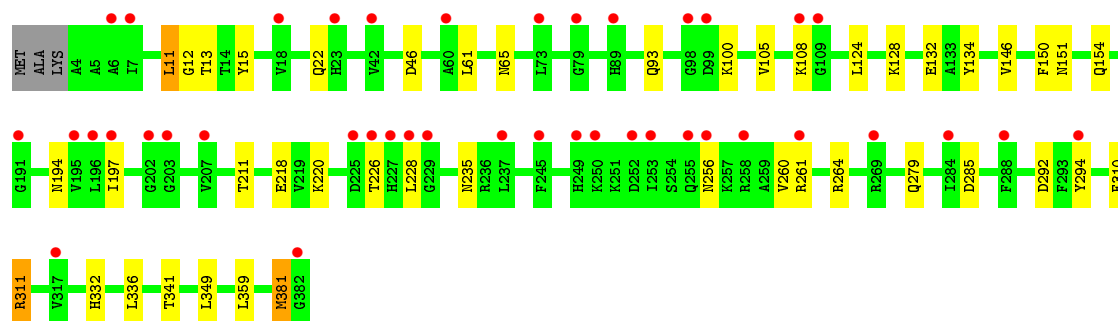
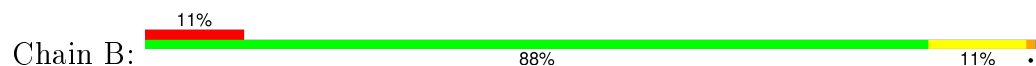
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	272	Total	O	0	0
			272	272		
7	B	102	Total	O	0	0
			102	102		
7	C	207	Total	O	0	0
			207	207		
7	D	80	Total	O	0	0
			80	80		

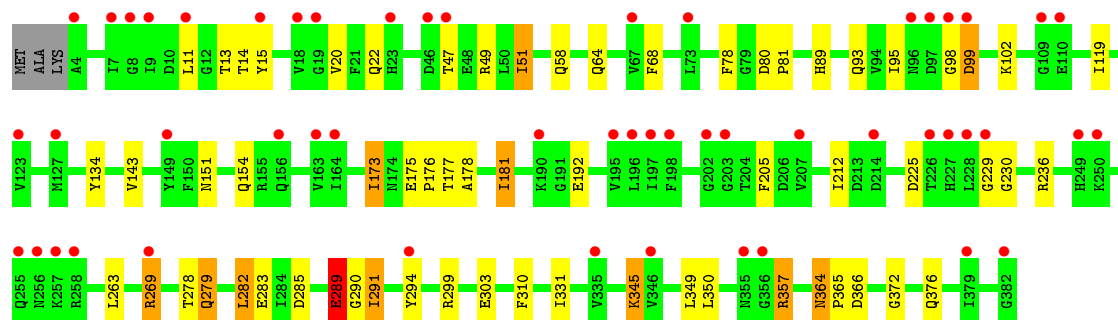
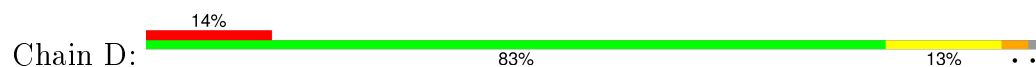




• Molecule 2: Heat shock 70 kDa protein 1



• Molecule 2: Heat shock 70 kDa protein 1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	129.50Å 141.90Å 151.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 19.93 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.30) 100.0 (19.93-2.30)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.197 , 0.244 0.196 , 0.241	Depositor DCC
$R_{free}$ test set	6248 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.7	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 53.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 124106 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16199	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.84% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, MG, K, ATP

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	0.59	1/4965 (0.0%)	0.73	5/6731 (0.1%)
1	C	0.79	13/4917 (0.3%)	0.75	4/6673 (0.1%)
2	B	0.55	0/2943	0.63	0/3988
2	D	1.21	7/2913 (0.2%)	0.67	5/3958 (0.1%)
All	All	0.80	21/15738 (0.1%)	0.71	14/21350 (0.1%)

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	C	0	1
2	B	0	1
All	All	0	2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	289	GLU	CD-OE1	55.35	1.86	1.25
1	C	435	ARG	CZ-NH1	19.98	1.59	1.33
1	C	411	GLU	CD-OE1	14.87	1.42	1.25
1	C	434	ASN	CG-OD1	14.74	1.56	1.24
2	D	289	GLU	CD-OE2	-13.50	1.10	1.25

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	435	ARG	NE-CZ-NH2	-18.75	110.92	120.30
1	C	435	ARG	NE-CZ-NH1	11.00	125.80	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	289	GLU	CG-CD-OE2	8.11	134.53	118.30
1	A	141	VAL	CB-CA-C	-7.16	97.80	111.40
1	A	34	ARG	NE-CZ-NH2	-6.34	117.13	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	226	THR	Peptide
1	C	434	ASN	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	4874	0	4770	70	0
1	C	4828	0	4693	55	1
2	B	2895	0	2846	27	0
2	D	2865	0	2748	36	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	A	31	0	12	0	0
5	C	31	0	12	0	0
6	A	6	0	8	0	0
7	A	272	0	0	9	1
7	B	102	0	0	0	0
7	C	207	0	0	7	0
7	D	80	0	0	1	0
All	All	16199	0	15089	185	1

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 6.

The worst 5 of 185 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)
2:D:289:GLU:CD	2:D:289:GLU:OE1	1.86	1.13
2:B:311:ARG:HH21	2:B:311:ARG:HG2	1.13	1.09
1:A:17:ALA:HB1	1:A:24[A]:ILE:HD11	1.40	1.02
1:C:61:ILE:O	1:C:89:THR:HG23	1.65	0.95
2:B:194:ASN:H	2:B:332:HIS:HD2	1.15	0.89

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:293:ASN:OD1	7:A:3056:HOH:O[4_467]	2.13	0.07

## 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	627/675 (93%)	600 (96%)	24 (4%)	3 (0%)	34	41
1	C	624/675 (92%)	594 (95%)	29 (5%)	1 (0%)	52	64
2	B	378/382 (99%)	365 (97%)	13 (3%)	0	100	100
2	D	380/382 (100%)	365 (96%)	11 (3%)	4 (1%)	17	18
All	All	2009/2114 (95%)	1924 (96%)	77 (4%)	8 (0%)	39	48

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	499	GLU
1	A	472	GLY
1	C	473	GLN
2	D	230	GLY
2	D	98	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	520/576 (90%)	469 (90%)	51 (10%)	10	11
1	C	512/576 (89%)	462 (90%)	50 (10%)	10	11
2	B	301/315 (96%)	292 (97%)	9 (3%)	48	65
2	D	291/315 (92%)	267 (92%)	24 (8%)	14	17
All	All	1624/1782 (91%)	1490 (92%)	134 (8%)	14	17

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	3	THR
1	C	100	THR
2	D	269	ARG
1	C	9	LEU
1	C	31	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	332	HIS
1	C	60	ASN
2	D	194	ASN
2	B	376	GLN
1	C	13	ASN

### 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 11 ligands modelled in this entry, 8 are monoatomic - leaving 3 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	$\# Z  > 2$
5	ATP	A	1001	3,4	24,33,33	1.00	2 (8%)	31,52,52	1.97	6 (19%)
6	GOL	A	3002	-	5,5,5	0.60	0	5,5,5	0.79	0
5	ATP	C	1001	3,4	24,33,33	0.98	1 (4%)	31,52,52	1.95	7 (22%)

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
5	ATP	A	1001	3,4	-	0/18/38/38	0/3/3/3
6	GOL	A	3002	-	-	0/4/4/4	0/0/0/0
5	ATP	C	1001	3,4	-	0/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1001	ATP	O4'-C1'	2.12	1.43	1.41
5	A	1001	ATP	C5-C4	3.06	1.47	1.40
5	C	1001	ATP	C5-C4	3.31	1.48	1.40

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1001	ATP	N3-C2-N1	-7.89	122.85	128.89
5	C	1001	ATP	N3-C2-N1	-7.40	123.23	128.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1001	ATP	C2'-C1'-N9	-3.54	108.88	114.29
5	C	1001	ATP	C1'-N9-C4	-3.21	122.09	126.94
5	C	1001	ATP	C4-C5-N7	-2.98	106.74	109.48

There are no chirality outliers.

There are no torsion outliers.

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 [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	629/675 (93%)	0.31	45 (7%)	18 26	34, 43, 54, 68	0
1	C	627/675 (92%)	0.53	71 (11%)	7 10	35, 44, 55, 76	0
2	B	379/382 (99%)	0.54	41 (10%)	8 11	35, 44, 54, 79	0
2	D	379/382 (99%)	0.69	52 (13%)	4 6	38, 45, 52, 72	0
All	All	2014/2114 (95%)	0.49	209 (10%)	8 12	34, 44, 54, 79	0

The worst 5 of 209 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	226	THR	8.6
1	C	654	ARG	7.9
1	A	525	THR	7.8
2	B	226	THR	7.1
2	D	229	GLY	6.9

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	GOL	A	3002	6/6	0.80	0.20	3.17	57,60,61,62	0
3	MG	A	2002	1/1	0.96	0.15	-0.57	42,42,42,42	0
3	MG	D	2001	1/1	0.91	0.14	-0.83	33,33,33,33	0
5	ATP	A	1001	31/31	0.99	0.10	-1.07	40,42,45,47	0
3	MG	B	2001	1/1	0.97	0.14	-1.11	37,37,37,37	0
5	ATP	C	1001	31/31	0.99	0.09	-1.21	36,41,43,44	0
3	MG	C	2002	1/1	0.89	0.11	-1.51	29,29,29,29	0
3	MG	A	2001	1/1	0.98	0.10	-1.74	40,40,40,40	0
3	MG	C	2001	1/1	0.94	0.06	-2.09	40,40,40,40	0
4	K	C	3001	1/1	0.99	0.14	-	45,45,45,45	0
4	K	A	3001	1/1	1.00	0.12	-	44,44,44,44	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.