



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:48 AM GMT

PDB ID : 2IS1
Title : Crystal structure of UvrD-DNA-SO4 complex
Authors : Yang, W.; Lee, J.Y.
Deposited on : 2006-10-16
Resolution : 2.90 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
<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

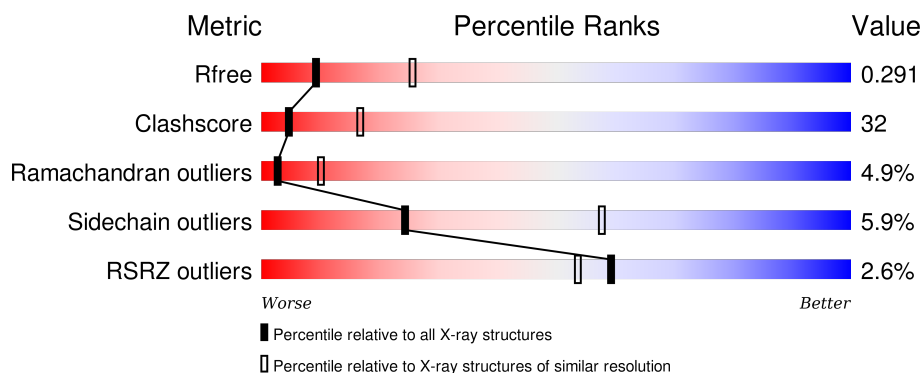
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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. 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	C	17	<div> <div>6%</div> <div>18% 65% 12% 6%</div> </div>
1	E	17	<div> <div>24% 65% 12%</div> </div>
2	D	10	<div> <div>10% 90%</div> </div>
2	F	10	<div> <div>20% 80%</div> </div>
3	A	680	<div> <div>56% 37% 5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	680	

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	GOL	A	686	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10896 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 DNA chain called 5'-D(*GP*CP*AP*GP*TP*GP*CP*TP*CP*GP*TP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	16	Total	C	N	O	P	0	0	0
			323	157	50	101	15			
1	E	15	Total	C	N	O	P	0	0	0
			300	147	48	92	13			

- Molecule 2 is a DNA chain called 5'-D(*CP*GP*AP*GP*CP*AP*CP*TP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	10	Total	C	N	O	P	0	0	0
			201	96	39	57	9			
2	F	10	Total	C	N	O	P	0	0	0
			200	96	39	56	9			

- Molecule 3 is a protein called DNA helicase II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	649	Total	C	N	O	S	0	0	0
			5058	3175	920	940	23			
3	B	624	Total	C	N	O	S	0	0	0
			4727	2967	848	889	23			

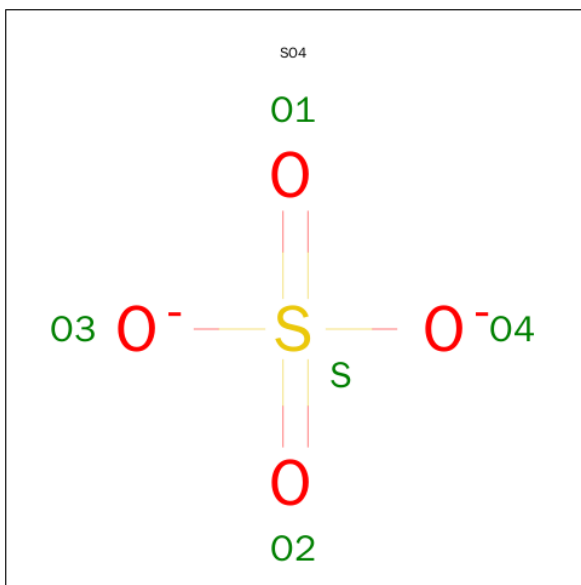
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	399	VAL	ALA	ENGINEERED	UNP P03018
B	399	VAL	ALA	ENGINEERED	UNP P03018

- Molecule 4 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	2	Total Hg 2 2	0	0
4	A	3	Total Hg 3 3	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	40	Total	O	0	0
			40	40		
7	B	15	Total	O	0	0
			15	15		

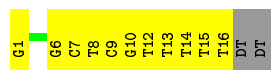
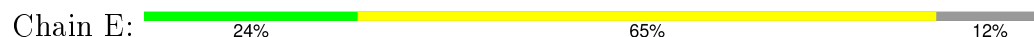
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

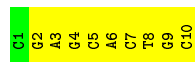
- Molecule 1: 5'-D(*GP*CP*AP*GP*TP*GP*CP*TP*CP*GP*TP*TP*TP*TP*TP*TP*T)-3',



- Molecule 1: 5'-D(*GP*CP*AP*GP*TP*GP*CP*TP*CP*GP*TP*TP*TP*TP*TP*TP*TP*T)-3',



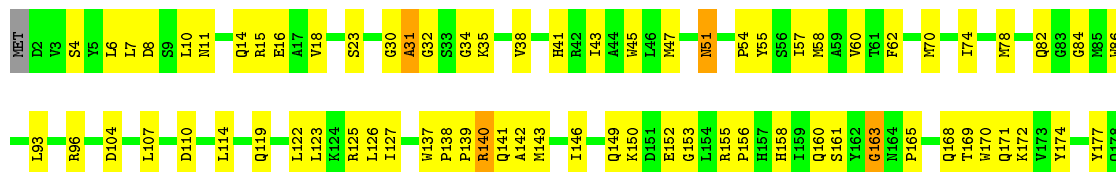
- Molecule 2: 5'-D(*CP*GP*AP*GP*CP*AP*CP*TP*GP*C)-3'



- Molecule 2: 5'-D(*CP*GP*AP*GP*CP*AP*CP*TP*GP*C)-3'



- Molecule 3: DNA helicase II





L592	D525	L435	E365	N294	D220	Y145	M64	MET
E536	T436	T436	A366	I297	E221	I146	R71	ASP
E594	D527	L437	L367	I297	F222		H72	V3
E595	L528	W438	E298	N299	D223	G153	H73	S4
E596	M529	Q439	L368	N299	D224	L154		Y5
E597	P530		Q370		T225	R155	L77	L6
L598	L531	L444	S371	R303	Q229	P156	M78	L7
E599	Q532	L450	M372	G305	L304	H157	T80	D8
Y600	H537	L451	P373		P373	H158	T90	S9
G601		A451	Y374		L233	I159	T90	L10
G602		A452	R375		R234	Q160	S81	M11
G603	E541	R453	I376	L308	L235	S161	Q82	D12
T604	A542		Y377	R309		TTR	G83	K13
R605	G545	L458	G378	T310	T240	GLY	V87	Q14
A606	G546	R460	G379	G312	W243	ASN		R15
M607	ALA	F461	M380	A313	W244	PRO		E16
	ASP	M462	R381	D314	I245	V166	F90	
L610	THR	E463	F382	I318	T246	E167	L93	A20
T611	TRP	L464	E384	S319	G247	Q168	A94	P21
T613	GLN	R465	R385	S320	D248	T169	H95	R22
Y614	D552	L466	Q386	Y321	D249	W170	S23	N24
A615	A553	A467	E387	C322	D250	Q171	R96	M24
E616	W554	L468	I388	A323	Q251	K372	L97	L25
Q555	Q555	A469		F324	S252	Y173	L26	L26
R618	L556	Q470	S393	R324	T253	Y174	R99	V27
R619	M557	E471	N325	N325	F254	Q175	A100	L28
T620	T558	T472	R396	R327	G255	Y177	H101	A29
Y621	L559		L397	D328	W256	Q178	H102	G30
G622	H560	P476		E329	R257		M103	A31
K623	A562	L477	M400	F332	G258	V187	A105	
E624		H478	R401	V333	A259	D188	M106	K35
		V479	N402	V334	Q260	L107	T36	R37
		Q480	D403		V261	E191	P108	R37
H627	L565				E262	Q109		V38
R628	E566	V484	F407	I337	I263	L194	D110	L39
P629	F567	L485	E408		L264	R195	V40	
S630	P688	R486	R409	W340	Q265	A196	Q112	H41
F631	Q569	D487	V410	Q341	R266	H197	I113	R42
F632	W570	S488	V411		F267	E198	L114	I43
I633	F571	Q489	N412	G344	L268	L199		
G634	I572		T413	G345		W200	D118	W45
E635	W573	R490	P414	A346	F271	L201	Q119	L46
L636	G574	R491	T415	L347		M202	L120	M47
P637	M575		T415	A348	A274	K203	R121	S48
E638	E576	V494	R416	E349		P204	L122	V49
E639	E577	E495	G417			W205	L123	E50
G640	G578		I418		L279	T206	R124	M51
V641	M579	R505		I352		L207	R125	C52
E642	F580	L506	R421	L353	N282	Q208	L126	M52
E643	P581	E507	T422	Y354	R283	Q208	S53	P54
V644	S582			R355	H209	Y210	A129	
ARG	Q583	E510	V425	S356	T285	Y210	Y55	S56
LEU	M584		V426	N357	T286	R211	W137	I57
ARG	SER	T516		A358	S287	E212	P138	M58
ALA	LEU	R517	T429	Q359	T288	R213	F139	A59
THR	ASP		S430	S360	I289			
VAL	GLY	Y521	R431	L290	L290	W216	A142	T61
SER	GLU	ASN	D432	V362	S29			

VAL
SER
HIS
GLN
ARG
MET
GLY
THR
PRO
MET
VAL
GLU
ASN
ASP
SER
GLY
TYR
LYS
LEU
GLY
GLN
ARG
VAL
ARG
HIS
ALA
LYS

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	152.45Å 94.49Å 136.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.76 – 2.90 47.62 – 2.79	Depositor EDS
% Data completeness (in resolution range)	96.5 (29.76-2.90) 92.3 (47.62-2.79)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.52 (at 2.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.236 , 0.295 0.233 , 0.291	Depositor DCC
R_{free} test set	4334 reflections (11.26%)	DCC
Wilson B-factor (Å ²)	54.0	Xtriage
Anisotropy	0.712	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 59.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 47996 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	10896	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.48% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, SO4, HG

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	C	0.63	0/359	0.99	3/553 (0.5%)
1	E	0.47	0/333	0.76	0/511
2	D	0.42	0/225	0.68	0/345
2	F	0.40	0/224	0.68	0/343
3	A	0.44	0/5153	0.68	1/6979 (0.0%)
3	B	0.43	0/4811	0.66	1/6529 (0.0%)
All	All	0.45	0/11105	0.68	5/15260 (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	C	1	0

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	13	DT	O5'-P-OP1	-6.12	100.19	105.70
3	B	105	ALA	N-CA-C	-5.92	95.01	111.00
1	C	11	DT	O4'-C1'-N1	5.55	111.88	108.00
3	A	653	PRO	N-CA-CB	5.40	109.78	103.30
1	C	11	DT	N1-C1'-C2'	5.19	122.47	112.60

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	11	DT	C1'

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	C	323	0	186	21	0
1	E	300	0	176	20	0
2	D	201	0	113	15	0
2	F	200	0	111	11	0
3	A	5058	0	4857	220	0
3	B	4727	0	4410	384	0
4	A	3	0	0	0	0
4	B	2	0	0	0	0
5	A	10	0	0	2	0
5	B	5	0	0	0	0
6	A	6	0	8	0	0
6	B	6	0	8	2	0
7	A	40	0	0	0	0
7	B	15	0	0	0	0
All	All	10896	0	9869	662	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 32.

The worst 5 of 662 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:E:12:DT:H3'	1:E:13:DT:C5'	1.70	1.21
1:E:12:DT:C3'	1:E:13:DT:H5''	1.75	1.17
1:C:11:DT:H3'	1:C:12:DT:C5'	1.72	1.16
1:C:11:DT:C3'	1:C:12:DT:H5''	1.75	1.15
3:A:23:SER:HA	3:A:242:LYS:HD2	1.15	1.10

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	
3	A	645/680 (95%)	575 (89%)	52 (8%)	18 (3%)	6	24
3	B	614/680 (90%)	471 (77%)	99 (16%)	44 (7%)	1	3
All	All	1259/1360 (93%)	1046 (83%)	151 (12%)	62 (5%)	3	10

5 of 62 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	263	ASN
3	A	303	ARG
3	A	521	TYR
3	A	547	ALA
3	A	583	GLN

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	
3	A	502/574 (88%)	481 (96%)	21 (4%)	36	73
3	B	450/574 (78%)	415 (92%)	35 (8%)	16	41
All	All	952/1148 (83%)	896 (94%)	56 (6%)	24	58

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	96	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	240	THR
3	B	597	ARG
3	B	106	ASN
3	B	121	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	95	HIS
3	B	147	ASN
3	B	412	ASN
3	B	106	ASN
3	B	158	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 10 ligands modelled in this entry, 5 are monoatomic - leaving 5 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	SO4	A	684	-	4,4,4	0.19	0	6,6,6	0.09	0
5	SO4	A	685	-	4,4,4	0.44	0	6,6,6	0.16	0
6	GOL	A	686	-	5,5,5	0.50	0	5,5,5	0.28	0
5	SO4	B	683	-	4,4,4	0.24	0	6,6,6	0.28	0
6	GOL	B	684	-	5,5,5	0.39	0	5,5,5	0.18	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
5	SO4	A	684	-	-	0/0/0/0	0/0/0/0
5	SO4	A	685	-	-	0/0/0/0	0/0/0/0
6	GOL	A	686	-	-	0/4/4/4	0/0/0/0
5	SO4	B	683	-	-	0/0/0/0	0/0/0/0
6	GOL	B	684	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	684	SO4	1	0
5	A	685	SO4	1	0
6	B	684	GOL	2	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	C	16/17 (94%)	0.40	1 (6%) 23 17	35, 68, 116, 117	0
1	E	15/17 (88%)	0.09	0 100 100	39, 78, 115, 124	0
2	D	10/10 (100%)	0.34	0 100 100	32, 85, 101, 103	0
2	F	10/10 (100%)	0.32	0 100 100	34, 64, 107, 108	0
3	A	649/680 (95%)	-0.16	3 (0%) 91 90	18, 43, 75, 95	0
3	B	624/680 (91%)	0.20	30 (4%) 34 28	21, 68, 96, 109	0
All	All	1324/1414 (93%)	0.03	34 (2%) 59 54	18, 53, 93, 124	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	587	ASP	4.8
3	B	633	ILE	4.8
3	B	630	SER	4.7
3	B	613	THR	4.3
3	B	579	MET	4.1

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, 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	LLDF	B-factors(Å ²)	Q<0.9
6	GOL	A	686	6/6	0.68	0.33	7.62	71,73,73,73	0
5	SO4	A	685	5/5	0.93	0.20	1.24	56,56,59,60	0
5	SO4	B	683	5/5	0.94	0.22	0.41	56,57,58,58	0
5	SO4	A	684	5/5	0.99	0.16	-0.67	44,45,46,47	0
4	HG	A	683	1/1	0.96	0.11	-2.52	107,107,107,107	1
4	HG	B	682	1/1	0.95	0.11	-	113,113,113,113	1
6	GOL	B	684	6/6	0.95	0.21	-	46,46,48,48	0
4	HG	A	681	1/1	0.94	0.11	-	84,84,84,84	1
4	HG	B	681	1/1	0.85	0.08	-	127,127,127,127	1
4	HG	A	682	1/1	0.99	0.09	-	75,75,75,75	1

6.5 Other polymers [i](#)

There are no such residues in this entry.