



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 9, 2017 – 01:10 AM EST

PDB ID : 4LRX
Title : Crystal Structure of the E.coli DhaR(N)-DhaK complex
Authors : Shi, R.; McDonald, L.; Cygler, M.; Ekiel, I.
Deposited on : unknown
Resolution : 3.25 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20030345

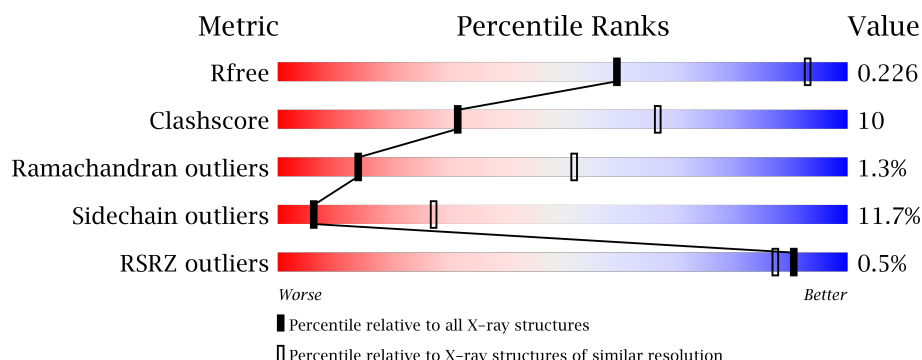
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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}	100719	1852 (3.32-3.20)
Clashscore	112137	2036 (3.32-3.20)
Ramachandran outliers	110173	2000 (3.32-3.20)
Sidechain outliers	110143	1998 (3.32-3.20)
RSRZ outliers	101464	1861 (3.32-3.20)

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 ≥ 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	356	<div> <div>76%</div> <div>22%</div> <div>.</div> </div>
1	B	356	<div> <div>77%</div> <div>22%</div> <div>.</div> </div>
2	C	318	<div> <div>%</div> <div>59%</div> <div>30%</div> <div>7%</div> <div>.</div> </div>
2	D	318	<div> <div>62%</div> <div>24%</div> <div>5%</div> <div>7%</div> <div>.</div> </div>

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
3	GOL	A	401	-	-	X	-
3	GOL	B	401	-	-	X	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9912 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.

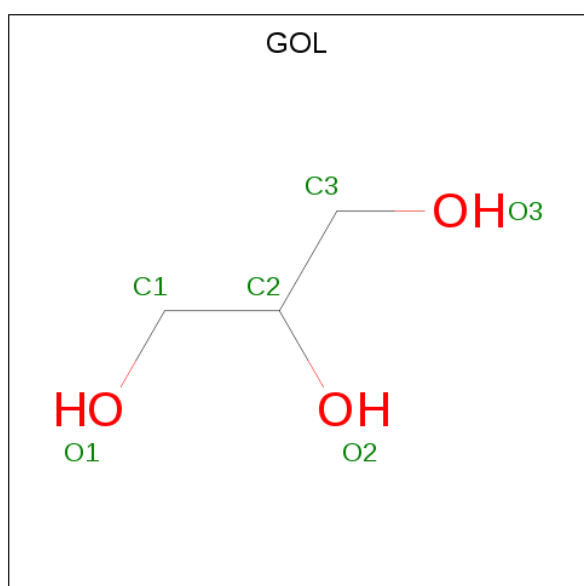
- Molecule 1 is a protein called PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	356	Total	C	N	O	S	0	0	0
			2681	1679	460	529	13			
1	B	356	Total	C	N	O	S	0	0	0
			2681	1679	460	529	13			

- Molecule 2 is a protein called PTS-dependent dihydroxyacetone kinase operon regulatory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	295	Total	C	N	O	S	0	0	0
			2269	1438	393	426	12			
2	D	295	Total	C	N	O	S	0	0	0
			2269	1438	393	426	12			

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).

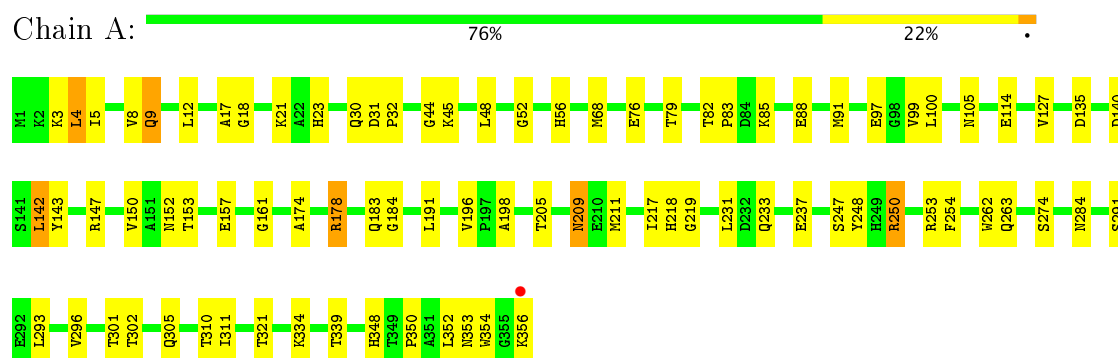


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

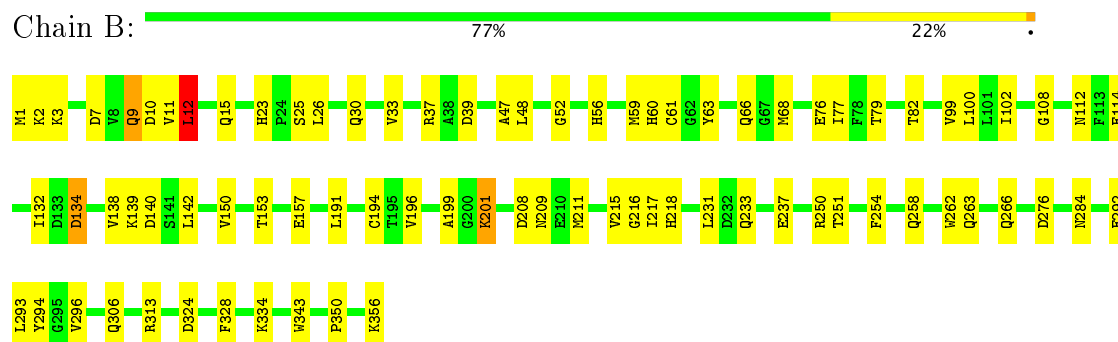
3 Residue-property plots [i](#)

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.

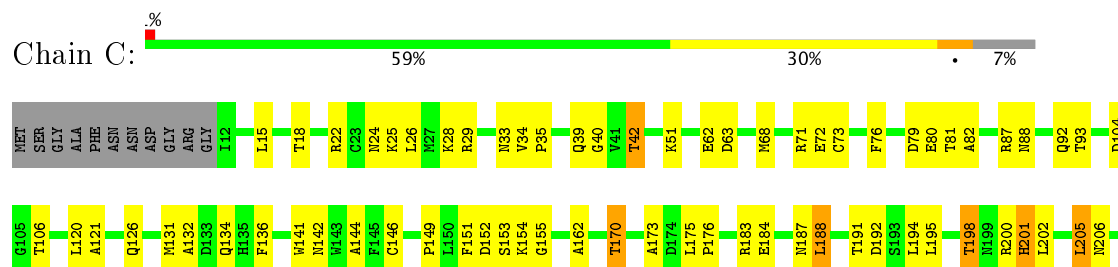
- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK

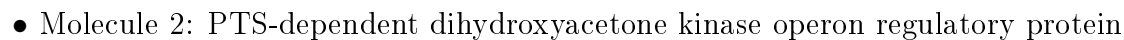


- Molecule 1: PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK



- Molecule 2: PTS-dependent dihydroxyacetone kinase operon regulatory protein





LEU	GLY	GLN	SER	THR	MET
V216	N229	F228	Q226	L225	W219
L236	R237	L238	D239	A240	T241
A242	R246	A247	I248	L252	T253
V257	L258	K262	L268	A273	S277
Q278	L282	V285	L286	T287	L288
K289	P290	T294	Q295	I300	H304
P305	GLU	GLN	MET	ARG	GLY
F136	V139	W140	F145	L150	F151
S153	K154	A155	A156	L157	T158
A164	E168	Q169	T170	A173	P176
E184	M187	L188	L189	T191	D192
S193	L194	L195	A196	T198	H199
R200	H201	L202	L205	L208	L209
E210	S211	M212	D213	D214	C215
N216	E217	L218	S219	L118	S119
L120	L123	V123	Q125	Q126	Q127
E128	E129	E130	E131	E132	E133
E134	E135	E136	E137	E138	E139
E140	E141	E142	E143	E144	E145
E146	E147	E148	E149	E150	E151
E152	E153	E154	E155	E156	E157
E158	E159	E160	E161	E162	E163
E164	E165	E166	E167	E168	E169
E170	E171	E172	E173	E174	E175
E176	E177	E178	E179	E180	E181
E182	E183	E184	E185	E186	E187
E188	E189	E190	E191	E192	E193
E194	E195	E196	E197	E198	E199
E200	E201	E202	E203	E204	E205
E206	E207	E208	E209	E210	E211
E212	E213	E214	E215	E216	E217
E218	E219	E220	E221	E222	E223
E224	E225	E226	E227	E228	E229
E230	E231	E232	E233	E234	E235
E236	E237	E238	E239	E240	E241
E242	E243	E244	E245	E246	E247
E248	E249	E250	E251	E252	E253
E254	E255	E256	E257	E258	E259
E260	E261	E262	E263	E264	E265
E266	E267	E268	E269	E270	E271
E272	E273	E274	E275	E276	E277
E278	E279	E280	E281	E282	E283
E284	E285	E286	E287	E288	E289
E290	E291	E292	E293	E294	E295
E296	E297	E298	E299	E300	E301
E302	E303	E304	E305	E306	E307
E308	E309	E310	E311	E312	E313
E314	E315	E316	E317	E318	E319
E320	E321	E322	E323	E324	E325
E326	E327	E328	E329	E330	E331
E332	E333	E334	E335	E336	E337
E338	E339	E340	E341	E342	E343
E344	E345	E346	E347	E348	E349
E350	E351	E352	E353	E354	E355
E356	E357	E358	E359	E360	E361
E362	E363	E364	E365	E366	E367
E368	E369	E370	E371	E372	E373
E374	E375	E376	E377	E378	E379
E380	E381	E382	E383	E384	E385
E386	E387	E388	E389	E390	E391
E392	E393	E394	E395	E396	E397
E398	E399	E400	E401	E402	E403
E404	E405	E406	E407	E408	E409
E410	E411	E412	E413	E414	E415
E416	E417	E418	E419	E420	E421
E422	E423	E424	E425	E426	E427
E428	E429	E430	E431	E432	E433
E434	E435	E436	E437	E438	E439
E440	E441	E442	E443	E444	E445
E446	E447	E448	E449	E450	E451
E452	E453	E454	E455	E456	E457
E458	E459	E460	E461	E462	E463
E464	E465	E466	E467	E468	E469
E470	E471	E472	E473	E474	E475
E476	E477	E478	E479	E480	E481
E482	E483	E484	E485	E486	E487
E488	E489	E490	E491	E492	E493
E494	E495	E496	E497	E498	E4

4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α , β , γ	232.13 Å 232.13 Å 79.89 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.25 46.95 – 3.25	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-3.25) 99.6 (46.95-3.25)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.66 (at 3.25 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.191 , 0.228 0.188 , 0.226	Depositor DCC
R_{free} test set	1956 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	58.0	Xtriage
Anisotropy	0.775	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 15.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.034 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9912	wwPDB-VP
Average B, all atoms (Å ²)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% 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, 0.333 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: GOL

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	0.65	0/2730	0.74	0/3710
1	B	0.63	0/2730	0.74	2/3710 (0.1%)
2	C	0.77	2/2311 (0.1%)	0.83	1/3147 (0.0%)
2	D	0.77	1/2311 (0.0%)	0.86	3/3147 (0.1%)
All	All	0.70	3/10082 (0.0%)	0.79	6/13714 (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
2	D	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	304	HIS	CB-CG	-6.30	1.38	1.50
2	C	62	GLU	CG-CD	5.76	1.60	1.51
2	D	273	ALA	C-O	-5.63	1.12	1.23

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	300	ILE	CB-CA-C	-5.98	99.64	111.60
1	B	12	LEU	CA-CB-CG	-5.82	101.92	115.30
2	D	152	ASP	N-CA-CB	-5.75	100.26	110.60
2	C	271	VAL	CA-CB-CG2	-5.50	102.66	110.90
1	B	61	CYS	CA-CB-SG	-5.49	104.12	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	151	PHE	Peptide

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	2681	0	2637	46	0
1	B	2681	0	2637	48	0
2	C	2269	0	2282	68	0
2	D	2269	0	2282	72	0
3	A	6	0	8	6	0
3	B	6	0	8	5	0
All	All	9912	0	9854	202	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 10.

The worst 5 of 202 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:B:218:HIS:NE2	3:B:401:GOL:H2	1.50	1.25
1:A:218:HIS:NE2	3:A:401:GOL:H2	1.66	1.10
2:D:68:MET:HE2	2:D:71:ARG:HH11	1.01	1.09
2:C:211:SER:HB3	2:D:300:ILE:HD11	1.08	1.07
1:B:218:HIS:NE2	3:B:401:GOL:C2	2.20	1.04

There are no symmetry-related clashes.

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	354/356 (99%)	327 (92%)	24 (7%)	3 (1%)	22	62
1	B	354/356 (99%)	322 (91%)	30 (8%)	2 (1%)	28	68
2	C	293/318 (92%)	267 (91%)	19 (6%)	7 (2%)	7	37
2	D	293/318 (92%)	274 (94%)	14 (5%)	5 (2%)	11	46
All	All	1294/1348 (96%)	1190 (92%)	87 (7%)	17 (1%)	14	52

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	220	ASP
1	A	140	ASP
2	C	40	GLY
2	C	230	ALA
2	C	265	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	287/288 (100%)	264 (92%)	23 (8%)	14	47
1	B	287/288 (100%)	268 (93%)	19 (7%)	19	56
2	C	241/260 (93%)	203 (84%)	38 (16%)	3	14
2	D	241/260 (93%)	197 (82%)	44 (18%)	2	9
All	All	1056/1096 (96%)	932 (88%)	124 (12%)	6	27

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	170	THR
2	C	246	ARG
2	D	239	ASP
2	C	188	LEU
2	C	205	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	183	GLN
2	C	88	ASN
2	D	229	ASN
1	B	185	HIS
1	B	305	GLN

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 ⓘ

2 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
3	GOL	A	401	-	5,5,5	0.40	0	5,5,5	0.54	0
3	GOL	B	401	-	5,5,5	0.25	0	5,5,5	0.50	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
3	GOL	A	401	-	-	0/4/4/4	0/0/0/0
3	GOL	B	401	-	-	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.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	GOL	6	0
3	B	401	GOL	5	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	356/356 (100%)	-0.33	1 (0%) 93 92	50, 66, 86, 96	0
1	B	356/356 (100%)	-0.31	0 100 100	47, 66, 86, 97	0
2	C	295/318 (92%)	-0.15	4 (1%) 75 67	41, 65, 107, 121	0
2	D	295/318 (92%)	-0.18	1 (0%) 93 92	44, 63, 100, 119	0
All	All	1302/1348 (96%)	-0.25	6 (0%) 90 87	41, 65, 94, 121	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	294	THR	3.0
2	C	292	ILE	3.0
2	C	296	GLY	2.7
2	D	294	THR	2.6
1	A	356	LYS	2.2

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(\AA^2)	Q<0.9
3	GOL	B	401	6/6	0.93	0.20	2.01	76,79,81,83	0
3	GOL	A	401	6/6	0.92	0.20	1.57	69,71,72,75	0

6.5 Other polymers [i](#)

There are no such residues in this entry.