Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2016 – 10:15 PM GMT

PDB ID	:	4Z9N
Title	:	ABC transporter / periplasmic binding protein from Brucella ovis with glutathione bound
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2015-04-10
Resolution	:	1.75 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

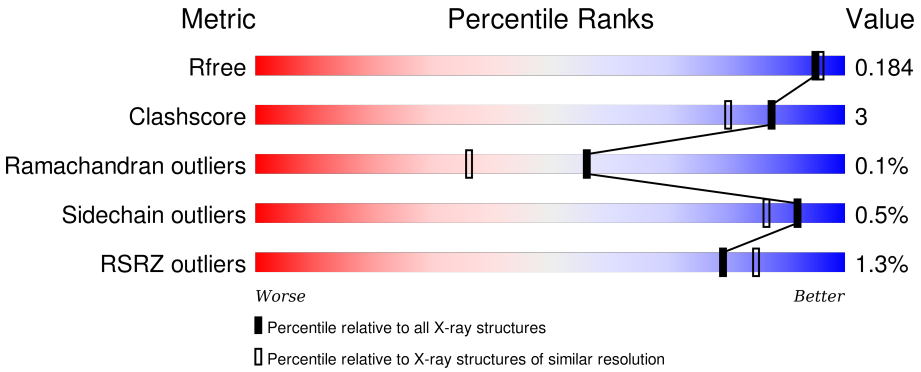
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 1.75 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.



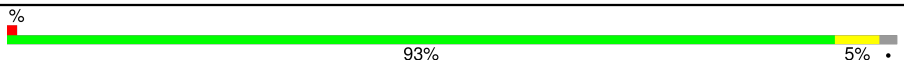
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2417 (1.76-1.72)
Clashscore	102246	2570 (1.76-1.72)
Ramachandran outliers	100387	2544 (1.76-1.72)
Sidechain outliers	100360	2544 (1.76-1.72)
RSRZ outliers	91569	2420 (1.76-1.72)

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 <=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	328	94%
1	B	328	92% 6%
1	C	328	94%
1	D	328	95%
1	E	328	92% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	328	

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
3	NA	A	402	-	-	-	X
3	NA	A	405	-	-	-	X
4	EDO	B	403	-	-	-	X
4	EDO	B	404	-	-	-	X
4	EDO	C	403	-	-	-	X
4	EDO	F	402	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16848 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 Amino acid ABC transporter, periplasmic amino acid-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	320	Total	C	N	O	S	0	5	0
			2447	1541	418	480	8			
1	B	324	Total	C	N	O	S	0	5	0
			2500	1570	431	490	9			
1	C	320	Total	C	N	O	S	0	5	0
			2471	1552	420	489	10			
1	D	321	Total	C	N	O	S	0	5	0
			2445	1537	416	483	9			
1	E	320	Total	C	N	O	S	0	3	0
			2439	1533	417	481	8			
1	F	321	Total	C	N	O	S	0	7	0
			2474	1553	421	491	9			

There are 48 discrepancies between the modelled and reference sequences:

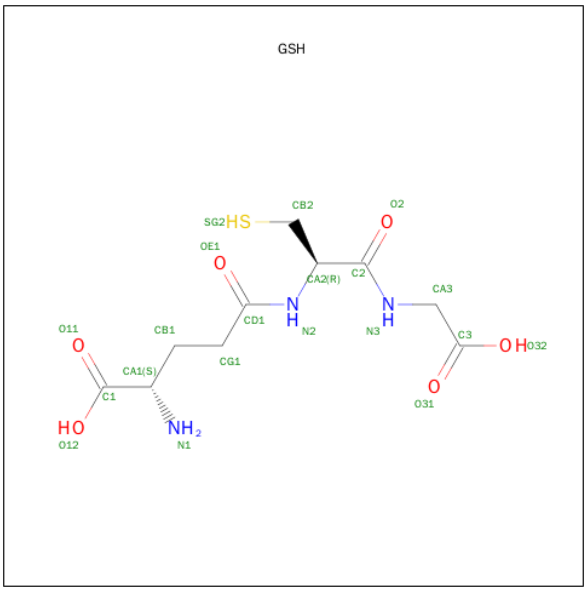
Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP A5VPS6
A	-5	ALA	-	expression tag	UNP A5VPS6
A	-4	HIS	-	expression tag	UNP A5VPS6
A	-3	HIS	-	expression tag	UNP A5VPS6
A	-2	HIS	-	expression tag	UNP A5VPS6
A	-1	HIS	-	expression tag	UNP A5VPS6
A	0	HIS	-	expression tag	UNP A5VPS6
A	1	HIS	-	expression tag	UNP A5VPS6
B	-6	MET	-	initiating methionine	UNP A5VPS6
B	-5	ALA	-	expression tag	UNP A5VPS6
B	-4	HIS	-	expression tag	UNP A5VPS6
B	-3	HIS	-	expression tag	UNP A5VPS6
B	-2	HIS	-	expression tag	UNP A5VPS6
B	-1	HIS	-	expression tag	UNP A5VPS6
B	0	HIS	-	expression tag	UNP A5VPS6
B	1	HIS	-	expression tag	UNP A5VPS6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	MET	-	initiating methionine	UNP A5VPS6
C	-5	ALA	-	expression tag	UNP A5VPS6
C	-4	HIS	-	expression tag	UNP A5VPS6
C	-3	HIS	-	expression tag	UNP A5VPS6
C	-2	HIS	-	expression tag	UNP A5VPS6
C	-1	HIS	-	expression tag	UNP A5VPS6
C	0	HIS	-	expression tag	UNP A5VPS6
C	1	HIS	-	expression tag	UNP A5VPS6
D	-6	MET	-	initiating methionine	UNP A5VPS6
D	-5	ALA	-	expression tag	UNP A5VPS6
D	-4	HIS	-	expression tag	UNP A5VPS6
D	-3	HIS	-	expression tag	UNP A5VPS6
D	-2	HIS	-	expression tag	UNP A5VPS6
D	-1	HIS	-	expression tag	UNP A5VPS6
D	0	HIS	-	expression tag	UNP A5VPS6
D	1	HIS	-	expression tag	UNP A5VPS6
E	-6	MET	-	initiating methionine	UNP A5VPS6
E	-5	ALA	-	expression tag	UNP A5VPS6
E	-4	HIS	-	expression tag	UNP A5VPS6
E	-3	HIS	-	expression tag	UNP A5VPS6
E	-2	HIS	-	expression tag	UNP A5VPS6
E	-1	HIS	-	expression tag	UNP A5VPS6
E	0	HIS	-	expression tag	UNP A5VPS6
E	1	HIS	-	expression tag	UNP A5VPS6
F	-6	MET	-	initiating methionine	UNP A5VPS6
F	-5	ALA	-	expression tag	UNP A5VPS6
F	-4	HIS	-	expression tag	UNP A5VPS6
F	-3	HIS	-	expression tag	UNP A5VPS6
F	-2	HIS	-	expression tag	UNP A5VPS6
F	-1	HIS	-	expression tag	UNP A5VPS6
F	0	HIS	-	expression tag	UNP A5VPS6
F	1	HIS	-	expression tag	UNP A5VPS6

- Molecule 2 is GLUTATHIONE (three-letter code: GSH) (formula: $C_{10}H_{17}N_3O_6S$).

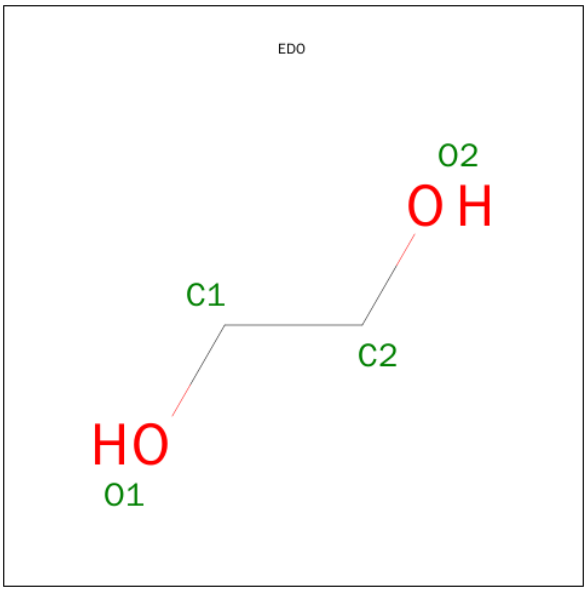


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			20	10	3	6	1		
2	B	1	Total	C	N	O	S	0	0
			20	10	3	6	1		
2	C	1	Total	C	N	O	S	0	0
			20	10	3	6	1		
2	D	1	Total	C	N	O	S	0	0
			20	10	3	6	1		
2	E	1	Total	C	N	O	S	0	0
			20	10	3	6	1		
2	F	1	Total	C	N	O	S	0	0
			20	10	3	6	1		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Na	0	0
			2	2		
3	A	2	Total	Na	0	0
			2	2		
3	C	2	Total	Na	0	0
			2	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	E	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	329	Total O 332 332	0	3
5	B	349	Total O 353 353	0	4
5	C	292	Total O 294 294	0	2
5	D	299	Total O 299 299	0	0

Continued on next page...

Continued from previous page...

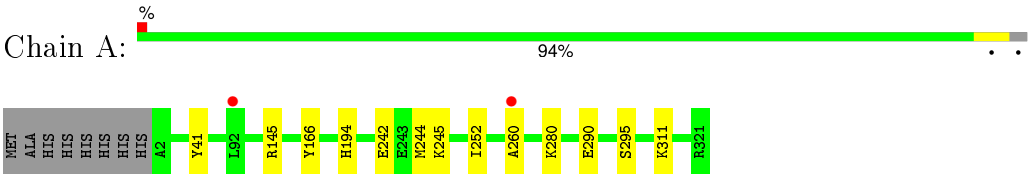
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	298	Total O 301 301	0	3
5	F	332	Total O 335 335	0	3

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 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.

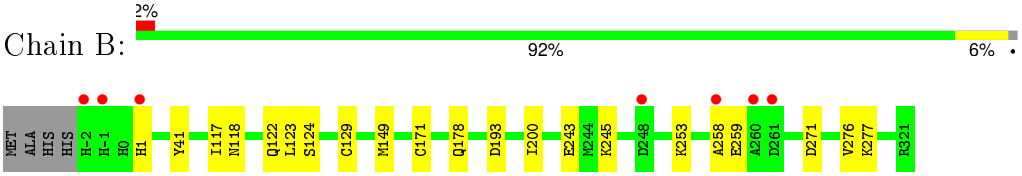
- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain A:



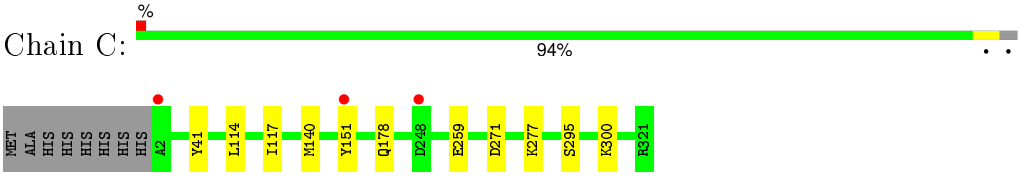
- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain B:



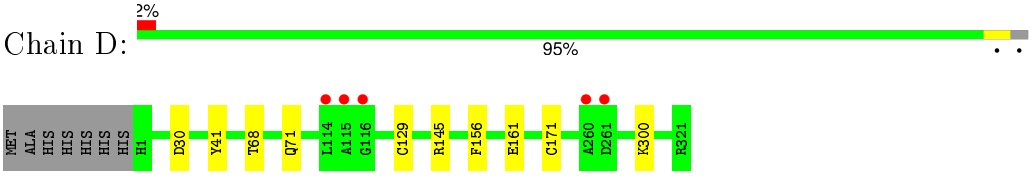
- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain C:



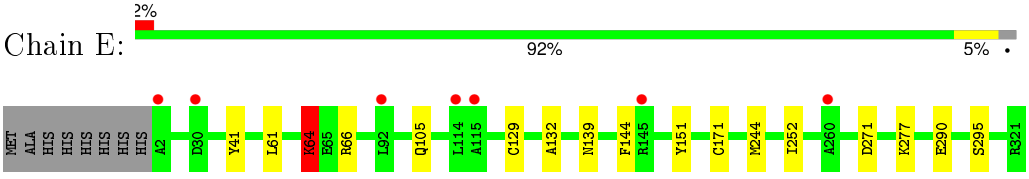
- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain D:



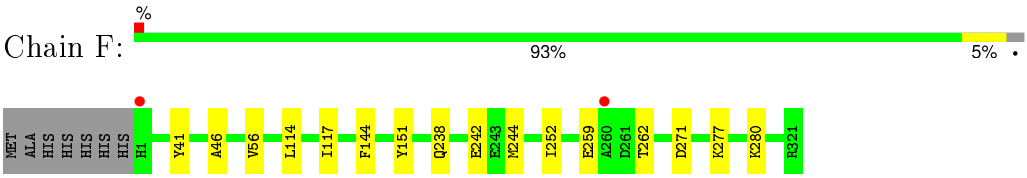
- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain E:



- Molecule 1: Amino acid ABC transporter, periplasmic amino acid-binding protein

Chain F:


%
93%
5%
MET
ALA
HIS
HIS
HIS
HIS
HIS
H1
Y41
A46
V56
L114
I117
F144
Y151
Q238
E242
E243
M244
I252
E259
A260
D261
T262
D271
K277
K280
R321

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.99Å 76.93Å 131.47Å 90.00° 94.07° 90.00°	Depositor
Resolution (Å)	31.66 – 1.75 46.88 – 1.74	Depositor EDS
% Data completeness (in resolution range)	98.7 (31.66-1.75) 95.2 (46.88-1.74)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.12 (at 1.75Å)	Xtriage
Refinement program	PHENIX (dev_2006: ???)	Depositor
R, R_{free}	0.152 , 0.183 0.157 , 0.184	Depositor DCC
R_{free} test set	3932 reflections (2.22%)	DCC
Wilson B-factor (Å^2)	15.6	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 54.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	1 of 187564 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	16848	wwPDB-VP
Average B, all atoms (Å^2)	22.0	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 56.39 % 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 2.7572e-05.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GSH, EDO, NA

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	0.42	0/2493	0.56	0/3388
1	B	0.44	0/2548	0.58	0/3458
1	C	0.39	0/2517	0.54	0/3415
1	D	0.39	0/2491	0.52	0/3385
1	E	0.43	1/2485 (0.0%)	0.56	0/3376
1	F	0.45	0/2520	0.58	0/3424
All	All	0.42	1/15054 (0.0%)	0.56	0/20446

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	64	LYS	CD-CE	-6.42	1.35	1.51

There are no bond angle outliers.

There are no chirality outliers.

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	A	2447	0	2353	9	0
1	B	2500	0	2408	21	0
1	C	2471	0	2382	8	1
1	D	2445	0	2337	6	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2439	0	2340	16	0
1	F	2474	0	2373	15	0
2	A	20	0	15	0	0
2	B	20	0	15	0	0
2	C	20	0	15	0	0
2	D	20	0	15	0	0
2	E	20	0	15	0	0
2	F	20	0	15	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
4	A	8	0	12	0	0
4	B	8	0	12	0	0
4	C	4	0	6	0	0
4	D	4	0	6	0	0
4	E	4	0	6	0	0
4	F	4	0	6	0	0
5	A	332	0	0	4	0
5	B	353	0	0	5	0
5	C	294	0	0	4	0
5	D	299	0	0	3	0
5	E	301	0	0	2	0
5	F	335	0	0	3	0
All	All	16848	0	14331	75	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 3.

All (75) 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:243:GLU:OE2	5:B:501:HOH:O	1.81	0.99
1:B:243:GLU:OE1	5:B:502:HOH:O	1.91	0.89
1:E:64:LYS:NZ	5:E:501:HOH:O	1.92	0.88
1:F:242:GLU:OE2	5:F:501:HOH:O	2.00	0.79
1:A:245:LYS:NZ	5:A:502:HOH:O	2.15	0.78
1:E:105:GLN:HE22	1:E:139:ASN:HB2	1.53	0.73
1:B:193:ASP:OD2	5:B:503:HOH:O	2.09	0.70
1:E:244:MET:HE1	1:E:252:ILE:HD13	1.74	0.69
1:F:238:GLN:O	5:F:502:HOH:O	2.11	0.69
1:E:244:MET:CE	1:E:252:ILE:HD13	2.22	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:271:ASP:OD2	1:E:277:LYS:NZ	2.27	0.67
1:B:118:ASN:H	1:B:122:GLN:HE22	1.40	0.67
1:A:311:LYS:O	5:A:501:HOH:O	2.13	0.66
1:F:259:GLU:OE2	5:F:503:HOH:O	2.14	0.66
1:A:244:MET:CE	1:A:252:ILE:HD13	2.27	0.65
1:C:259:GLU:OE2	5:C:501:HOH:O	2.13	0.65
1:F:242:GLU:H	1:F:242:GLU:CD	2.00	0.65
1:F:244:MET:HE1	1:F:252:ILE:HD13	1.80	0.63
1:A:242:GLU:OE2	1:A:280:LYS:NZ	2.24	0.62
1:F:244:MET:CE	1:F:252:ILE:HD13	2.30	0.61
1:F:259:GLU:HG3	1:F:262:THR:HB	1.82	0.60
1:C:151:TYR:HA	5:C:509:HOH:O	2.02	0.58
1:F:242:GLU:HG3	1:F:280:LYS:HE2	1.84	0.58
1:B:129[A]:CYS:HB3	1:B:171:CYS:SG	2.43	0.58
1:D:300:LYS:NZ	5:D:506:HOH:O	2.37	0.57
1:B:245:LYS:HE2	1:B:276:VAL:HG21	1.88	0.56
1:A:166:TYR:OH	1:A:194:HIS:HD2	1.88	0.54
1:B:124:SER:N	1:B:149:MET:HE2	2.23	0.54
1:B:178[A]:GLN:OE1	1:B:200:ILE:HD11	2.09	0.53
1:E:64:LYS:NZ	1:E:132:ALA:O	2.39	0.53
1:C:271:ASP:OD2	1:C:277:LYS:NZ	2.40	0.53
1:E:64:LYS:H	1:E:64:LYS:HD2	1.74	0.53
1:B:117:ILE:HA	1:B:122:GLN:NE2	2.24	0.52
1:E:105:GLN:NE2	1:E:139:ASN:HB2	2.24	0.52
1:B:243:GLU:OE2	1:B:243:GLU:N	2.39	0.52
1:F:271:ASP:OD2	1:F:277:LYS:NZ	2.41	0.51
1:E:64:LYS:HD2	1:E:64:LYS:N	2.26	0.51
1:A:194:HIS:HE1	5:A:573:HOH:O	1.95	0.50
1:B:271:ASP:OD2	1:B:277:LYS:NZ	2.45	0.49
1:D:161:GLU:OE1	5:D:501:HOH:O	2.19	0.49
1:C:259:GLU:OE2	5:C:502:HOH:O	2.20	0.49
1:A:145:ARG:NH1	5:A:507:HOH:O	2.46	0.47
1:F:144:PHE:CD2	1:F:151:TYR:HB3	2.50	0.47
1:D:129[A]:CYS:HB3	1:D:171:CYS:SG	2.55	0.46
1:B:117:ILE:HA	1:B:122:GLN:HE22	1.78	0.46
1:B:117:ILE:HG23	1:B:122:GLN:HE21	1.79	0.46
1:E:290:GLU:OE1	1:E:295:SER:OG	2.31	0.46
1:D:68[B]:THR:HA	1:D:71:GLN:HE21	1.81	0.46
1:D:30:ASP:N	5:D:512:HOH:O	2.42	0.46
1:B:123:LEU:H	1:B:149:MET:HE3	1.80	0.45
1:E:144:PHE:CD2	1:E:151:TYR:HB3	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:114:LEU:HD23	1:C:117:ILE:HD12	2.00	0.44
1:E:129:CYS:HB3	1:E:171:CYS:SG	2.57	0.44
1:F:46:ALA:HB2	1:F:56[B]:VAL:HG22	2.00	0.44
1:F:244:MET:CE	1:F:252:ILE:HG21	2.48	0.43
1:E:290:GLU:OE1	1:E:295:SER:N	2.50	0.43
1:E:64:LYS:HD2	5:E:507:HOH:O	2.19	0.43
1:B:259:GLU:OE1	5:B:504:HOH:O	2.21	0.43
1:F:114:LEU:HG	1:F:117:ILE:HD12	2.00	0.43
1:E:64:LYS:CE	1:E:132:ALA:O	2.67	0.42
1:A:245:LYS:HB3	1:A:245:LYS:HE3	1.73	0.42
1:B:124:SER:N	1:B:149:MET:CE	2.82	0.42
1:A:290:GLU:OE1	1:A:295:SER:N	2.43	0.42
1:B:118:ASN:N	1:B:122:GLN:HE22	2.13	0.42
1:F:244:MET:HE3	1:F:252:ILE:HG21	2.01	0.41
1:D:129[B]:CYS:SG	1:D:156:PHE:CE1	3.13	0.41
1:C:300:LYS:HB2	1:C:300:LYS:HE2	1.75	0.41
1:B:245:LYS:HD3	5:B:543:HOH:O	2.21	0.41
1:B:123:LEU:N	1:B:149:MET:HE3	2.35	0.41
1:B:253:LYS:HG3	1:B:258:ALA:HB3	2.01	0.41
1:C:178:GLN:NE2	5:C:516:HOH:O	2.53	0.41
1:E:61:LEU:HD13	1:E:66:ARG:HA	2.02	0.41
1:F:242:GLU:HG3	1:F:280:LYS:CE	2.51	0.41
1:C:140:MET:HG2	1:C:151:TYR:CE1	2.56	0.41
1:B:200:ILE:HA	1:B:200:ILE:HD13	1.91	0.40

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:295:SER:O	1:D:145:ARG:NH1[1_565]	2.06	0.14

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	323/328 (98%)	308 (95%)	13 (4%)	2 (1%)	30 11
1	B	327/328 (100%)	315 (96%)	12 (4%)	0	100 100
1	C	323/328 (98%)	311 (96%)	12 (4%)	0	100 100
1	D	324/328 (99%)	313 (97%)	11 (3%)	0	100 100
1	E	321/328 (98%)	308 (96%)	13 (4%)	0	100 100
1	F	326/328 (99%)	313 (96%)	13 (4%)	0	100 100
All	All	1944/1968 (99%)	1868 (96%)	74 (4%)	2 (0%)	56 36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	260[A]	ALA
1	A	260[B]	ALA

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	251/265 (95%)	250 (100%)	1 (0%)	93 90
1	B	259/265 (98%)	257 (99%)	2 (1%)	86 77
1	C	258/265 (97%)	257 (100%)	1 (0%)	93 90
1	D	251/265 (95%)	250 (100%)	1 (0%)	93 90
1	E	251/265 (95%)	249 (99%)	2 (1%)	86 77
1	F	256/265 (97%)	255 (100%)	1 (0%)	93 90
All	All	1526/1590 (96%)	1518 (100%)	8 (0%)	92 87

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1	HIS
1	B	41	TYR
1	C	41	TYR
1	D	41	TYR
1	E	41	TYR
1	E	64	LYS
1	F	41	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	178	GLN
1	A	194	HIS
1	B	122	GLN
1	B	238	GLN
1	C	152	ASN
1	C	178	GLN
1	D	71	GLN
1	D	178	GLN
1	E	105	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

Of 20 ligands modelled in this entry, 6 are monoatomic - leaving 14 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	$\#\|Z\| > 2$
2	GSH	A	401	-	13,19,19	2.37	3 (23%)	15,24,24	1.57	4 (26%)
4	EDO	A	403	-	3,3,3	0.53	0	2,2,2	0.21	0
4	EDO	A	404	-	3,3,3	0.45	0	2,2,2	0.31	0
2	GSH	B	401	-	13,19,19	2.16	2 (15%)	15,24,24	1.82	4 (26%)
4	EDO	B	403	-	3,3,3	0.51	0	2,2,2	0.36	0
4	EDO	B	404	-	3,3,3	0.50	0	2,2,2	0.36	0
2	GSH	C	401	-	13,19,19	2.19	2 (15%)	15,24,24	1.50	3 (20%)
4	EDO	C	403	-	3,3,3	0.49	0	2,2,2	0.46	0
2	GSH	D	401	-	13,19,19	2.31	3 (23%)	15,24,24	1.43	1 (6%)
4	EDO	D	402	-	3,3,3	0.46	0	2,2,2	0.15	0
2	GSH	E	401	-	13,19,19	2.17	2 (15%)	15,24,24	1.32	0
4	EDO	E	402	-	3,3,3	0.45	0	2,2,2	0.32	0
2	GSH	F	401	-	13,19,19	2.32	3 (23%)	15,24,24	1.68	3 (20%)
4	EDO	F	402	-	3,3,3	0.60	0	2,2,2	0.32	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	GSH	A	401	-	-	0/18/24/24	0/0/0/0
4	EDO	A	403	-	-	0/1/1/1	0/0/0/0
4	EDO	A	404	-	-	0/1/1/1	0/0/0/0
2	GSH	B	401	-	-	0/18/24/24	0/0/0/0
4	EDO	B	403	-	-	0/1/1/1	0/0/0/0
4	EDO	B	404	-	-	0/1/1/1	0/0/0/0
2	GSH	C	401	-	-	0/18/24/24	0/0/0/0
4	EDO	C	403	-	-	0/1/1/1	0/0/0/0
2	GSH	D	401	-	-	0/18/24/24	0/0/0/0
4	EDO	D	402	-	-	0/1/1/1	0/0/0/0
2	GSH	E	401	-	-	0/18/24/24	0/0/0/0
4	EDO	E	402	-	-	0/1/1/1	0/0/0/0
2	GSH	F	401	-	-	0/18/24/24	0/0/0/0
4	EDO	F	402	-	-	0/1/1/1	0/0/0/0

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	GSH	O2-C2	-2.91	1.17	1.23
2	F	401	GSH	O2-C2	-2.11	1.19	1.23
2	D	401	GSH	O2-C2	-2.09	1.19	1.23
2	B	401	GSH	CD1-N2	4.95	1.44	1.34
2	E	401	GSH	CD1-N2	5.12	1.44	1.34
2	C	401	GSH	CD1-N2	5.14	1.44	1.34
2	F	401	GSH	CD1-N2	5.16	1.44	1.34
2	E	401	GSH	C2-N3	5.21	1.44	1.33
2	C	401	GSH	C2-N3	5.25	1.44	1.33
2	D	401	GSH	CD1-N2	5.29	1.44	1.34
2	A	401	GSH	CD1-N2	5.37	1.44	1.34
2	B	401	GSH	C2-N3	5.38	1.44	1.33
2	A	401	GSH	C2-N3	5.51	1.45	1.33
2	D	401	GSH	C2-N3	5.59	1.45	1.33
2	F	401	GSH	C2-N3	5.65	1.45	1.33

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	GSH	CA2-CB2-SG2	-4.21	109.36	113.99
2	F	401	GSH	OE1-CD1-N2	-3.12	117.71	122.96
2	F	401	GSH	CA2-CB2-SG2	-3.11	110.57	113.99
2	A	401	GSH	OE1-CD1-N2	-2.89	118.10	122.96
2	B	401	GSH	CA2-N2-CD1	-2.76	115.72	121.72
2	C	401	GSH	CA2-CB2-SG2	-2.70	111.03	113.99
2	A	401	GSH	CB2-CA2-N2	-2.39	108.21	111.43
2	D	401	GSH	CG1-CB1-CA1	-2.24	108.89	114.18
2	B	401	GSH	CG1-CB1-CA1	-2.21	108.96	114.18
2	A	401	GSH	CA2-N2-CD1	-2.17	117.01	121.72
2	C	401	GSH	OE1-CD1-N2	-2.13	119.38	122.96
2	B	401	GSH	CB1-CG1-CD1	-2.04	108.60	113.26
2	F	401	GSH	CG1-CD1-N2	2.23	119.59	115.85
2	C	401	GSH	O2-C2-CA2	2.45	125.60	120.39
2	A	401	GSH	CG1-CD1-N2	2.59	120.20	115.85

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

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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	320/328 (97%)	-0.36	2 (0%)	90 93	10, 19, 36, 50	0
1	B	324/328 (98%)	-0.32	7 (2%)	65 72	8, 17, 37, 69	0
1	C	320/328 (97%)	-0.42	3 (0%)	85 90	11, 19, 38, 54	0
1	D	321/328 (97%)	-0.28	5 (1%)	74 81	11, 23, 40, 56	0
1	E	320/328 (97%)	-0.19	7 (2%)	65 72	8, 20, 42, 52	0
1	F	321/328 (97%)	-0.45	2 (0%)	90 93	9, 19, 37, 54	0
All	All	1926/1968 (97%)	-0.34	26 (1%)	79 85	8, 19, 39, 69	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	260	ALA	7.4
1	B	260	ALA	4.9
1	C	2	ALA	4.9
1	A	260[A]	ALA	3.8
1	E	2	ALA	3.7
1	B	-1	HIS	3.2
1	E	30	ASP	3.1
1	B	1	HIS	3.0
1	F	1	HIS	3.0
1	F	260[A]	ALA	2.9
1	E	114	LEU	2.8
1	E	115	ALA	2.6
1	E	260	ALA	2.6
1	B	248	ASP	2.6
1	D	114	LEU	2.4
1	E	92	LEU	2.4
1	B	261	ASP	2.4
1	E	145	ARG	2.4
1	B	-2	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	151	TYR	2.3
1	D	261	ASP	2.2
1	D	116	GLY	2.2
1	C	248	ASP	2.2
1	D	115	ALA	2.1
1	B	258	ALA	2.1
1	A	92	LEU	2.1

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, 95^{th} 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
4	EDO	C	403	4/4	0.87	0.16	8.26	20,22,25,26	0
4	EDO	B	403	4/4	0.93	0.09	4.00	22,23,23,24	0
4	EDO	F	402	4/4	0.90	0.12	3.47	26,28,29,30	0
3	NA	A	405	1/1	0.99	0.12	2.63	13,13,13,13	0
4	EDO	B	404	4/4	0.92	0.13	2.31	26,30,35,37	0
3	NA	A	402	1/1	0.99	0.14	2.27	12,12,12,12	0
4	EDO	A	404	4/4	0.89	0.12	1.50	31,36,37,38	0
4	EDO	A	403	4/4	0.92	0.10	0.75	23,24,30,32	0
4	EDO	D	402	4/4	0.97	0.08	0.66	21,22,22,23	0
4	EDO	E	402	4/4	0.97	0.10	0.41	28,29,30,30	0
2	GSH	C	401	20/20	0.97	0.07	0.24	10,14,20,21	0
2	GSH	A	401	20/20	0.96	0.07	-0.20	11,15,20,22	0
2	GSH	B	401	20/20	0.98	0.06	-0.23	6,11,17,18	0
3	NA	C	402	1/1	0.99	0.07	-0.24	15,15,15,15	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	GSH	D	401	20/20	0.97	0.06	-0.45	9,16,23,23	0
2	GSH	E	401	20/20	0.97	0.06	-0.62	10,15,21,28	0
2	GSH	F	401	20/20	0.98	0.06	-0.64	10,15,19,20	0
3	NA	C	404	1/1	0.99	0.07	-0.64	16,16,16,16	0
3	NA	B	402	1/1	0.99	0.07	-0.84	13,13,13,13	0
3	NA	B	405	1/1	1.00	0.10	-0.90	13,13,13,13	0

6.5 Other polymers

There are no such residues in this entry.