Full wwPDB X-ray Structure Validation Report ⓘ

Nov 7, 2023 – 12:34 PM JST

PDB ID	:	5XR2
Title	:	SAV0551
Authors	:	Kim, H.J.; Kwon, A.R.; Lee, B.J.
Deposited on	:	2017-06-07
Resolution	:	2.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

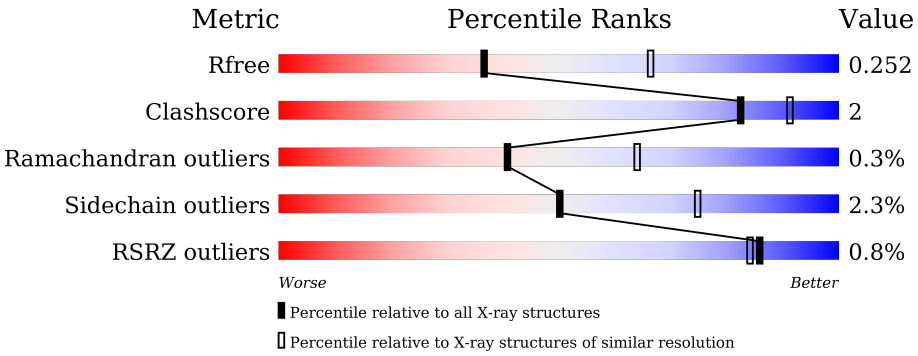
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 of 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 <=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	300	90% 6% ·
1	B	300	85% 9% · 5%
1	C	300	89% 9% ·
1	D	300	84% 9% · 5%
1	E	300	89% 6% ·
1	F	300	85% 9% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	300	94%
1	H	300	2% 93%

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
2	LAC	A	401	-	X	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 18184 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 Protein/nucleic acid deglycase HchA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	0	0
			2224	1418	365	432	9			
1	B	286	Total	C	N	O	S	0	0	0
			2214	1411	363	431	9			
1	C	295	Total	C	N	O	S	0	0	0
			2301	1463	386	443	9			
1	D	284	Total	C	N	O	S	0	0	0
			2198	1401	360	428	9			
1	E	287	Total	C	N	O	S	0	0	0
			2224	1416	366	433	9			
1	F	283	Total	C	N	O	S	0	0	0
			2190	1397	359	425	9			
1	G	295	Total	C	N	O	S	0	0	0
			2301	1463	386	443	9			
1	H	289	Total	C	N	O	S	0	0	0
			2239	1424	368	438	9			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	293	LEU	-	expression tag	UNP P64312
A	294	GLU	-	expression tag	UNP P64312
A	295	HIS	-	expression tag	UNP P64312
A	296	HIS	-	expression tag	UNP P64312
A	297	HIS	-	expression tag	UNP P64312
A	298	HIS	-	expression tag	UNP P64312
A	299	HIS	-	expression tag	UNP P64312
A	300	HIS	-	expression tag	UNP P64312
B	293	LEU	-	expression tag	UNP P64312
B	294	GLU	-	expression tag	UNP P64312
B	295	HIS	-	expression tag	UNP P64312
B	296	HIS	-	expression tag	UNP P64312
B	297	HIS	-	expression tag	UNP P64312

Continued on next page...

Continued from previous page...

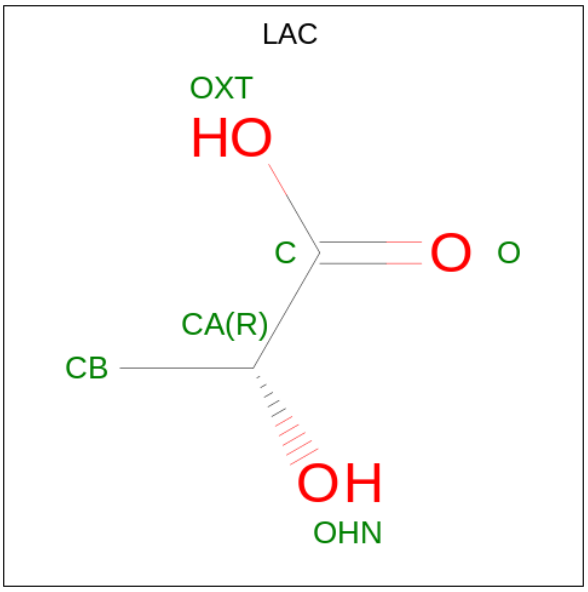
Chain	Residue	Modelled	Actual	Comment	Reference
B	298	HIS	-	expression tag	UNP P64312
B	299	HIS	-	expression tag	UNP P64312
B	300	HIS	-	expression tag	UNP P64312
C	293	LEU	-	expression tag	UNP P64312
C	294	GLU	-	expression tag	UNP P64312
C	295	HIS	-	expression tag	UNP P64312
C	296	HIS	-	expression tag	UNP P64312
C	297	HIS	-	expression tag	UNP P64312
C	298	HIS	-	expression tag	UNP P64312
C	299	HIS	-	expression tag	UNP P64312
C	300	HIS	-	expression tag	UNP P64312
D	293	LEU	-	expression tag	UNP P64312
D	294	GLU	-	expression tag	UNP P64312
D	295	HIS	-	expression tag	UNP P64312
D	296	HIS	-	expression tag	UNP P64312
D	297	HIS	-	expression tag	UNP P64312
D	298	HIS	-	expression tag	UNP P64312
D	299	HIS	-	expression tag	UNP P64312
D	300	HIS	-	expression tag	UNP P64312
E	293	LEU	-	expression tag	UNP P64312
E	294	GLU	-	expression tag	UNP P64312
E	295	HIS	-	expression tag	UNP P64312
E	296	HIS	-	expression tag	UNP P64312
E	297	HIS	-	expression tag	UNP P64312
E	298	HIS	-	expression tag	UNP P64312
E	299	HIS	-	expression tag	UNP P64312
E	300	HIS	-	expression tag	UNP P64312
F	293	LEU	-	expression tag	UNP P64312
F	294	GLU	-	expression tag	UNP P64312
F	295	HIS	-	expression tag	UNP P64312
F	296	HIS	-	expression tag	UNP P64312
F	297	HIS	-	expression tag	UNP P64312
F	298	HIS	-	expression tag	UNP P64312
F	299	HIS	-	expression tag	UNP P64312
F	300	HIS	-	expression tag	UNP P64312
G	293	LEU	-	expression tag	UNP P64312
G	294	GLU	-	expression tag	UNP P64312
G	295	HIS	-	expression tag	UNP P64312
G	296	HIS	-	expression tag	UNP P64312
G	297	HIS	-	expression tag	UNP P64312
G	298	HIS	-	expression tag	UNP P64312
G	299	HIS	-	expression tag	UNP P64312

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	300	HIS	-	expression tag	UNP P64312
H	293	LEU	-	expression tag	UNP P64312
H	294	GLU	-	expression tag	UNP P64312
H	295	HIS	-	expression tag	UNP P64312
H	296	HIS	-	expression tag	UNP P64312
H	297	HIS	-	expression tag	UNP P64312
H	298	HIS	-	expression tag	UNP P64312
H	299	HIS	-	expression tag	UNP P64312
H	300	HIS	-	expression tag	UNP P64312

- Molecule 2 is LACTIC ACID (three-letter code: LAC) (formula: $C_3H_6O_3$).



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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

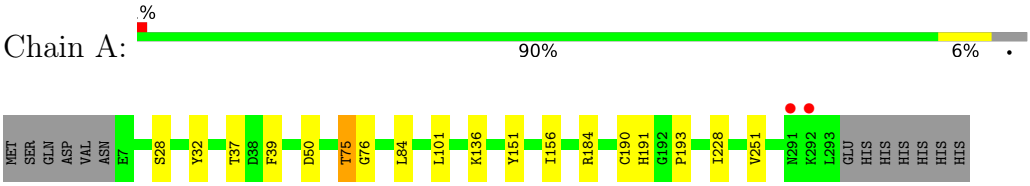
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	64	Total O 64 64	0	0
4	B	21	Total O 21 21	0	0
4	C	55	Total O 55 55	0	0
4	D	21	Total O 21 21	0	0
4	E	30	Total O 30 30	0	0
4	F	17	Total O 17 17	0	0
4	G	50	Total O 50 50	0	0
4	H	22	Total O 22 22	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

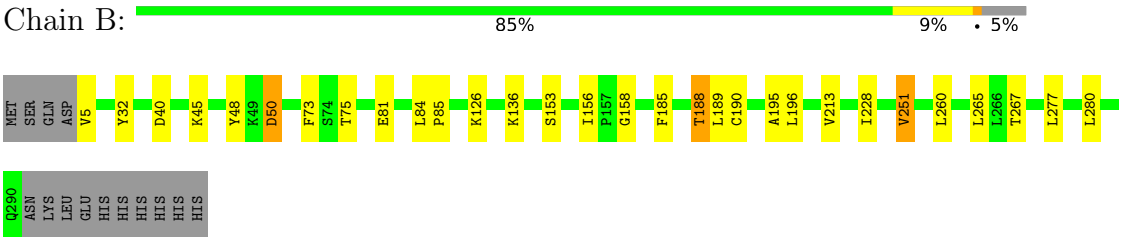
• Molecule 1: Protein/nucleic acid deglycase HchA

Chain A:



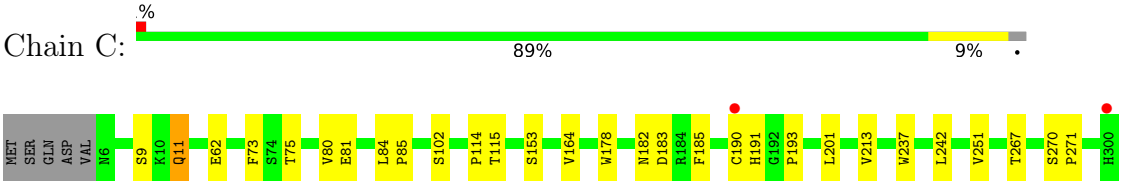
• Molecule 1: Protein/nucleic acid deglycase HchA

Chain B:



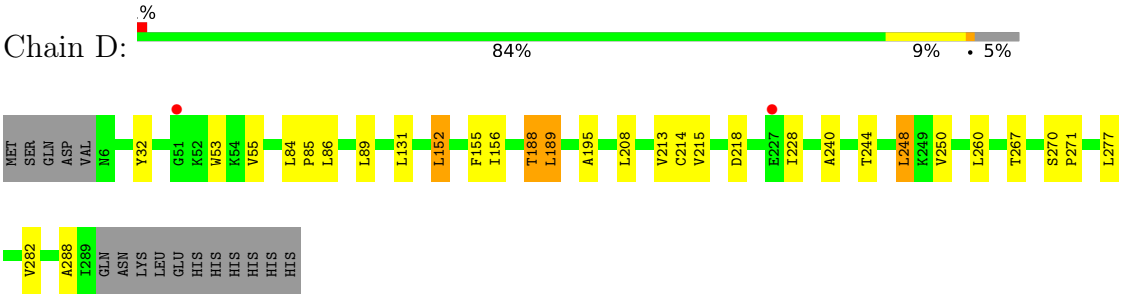
• Molecule 1: Protein/nucleic acid deglycase HchA

Chain C:



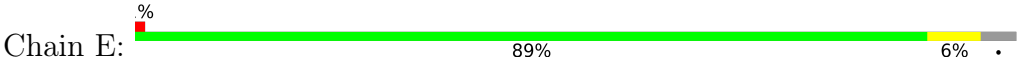
• Molecule 1: Protein/nucleic acid deglycase HchA

Chain D:

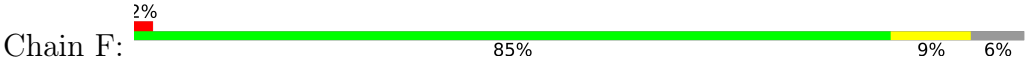


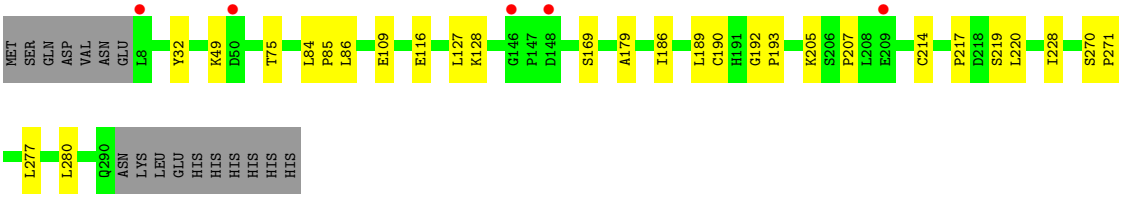
• Molecule 1: Protein/nucleic acid deglycase HchA

Chain E:



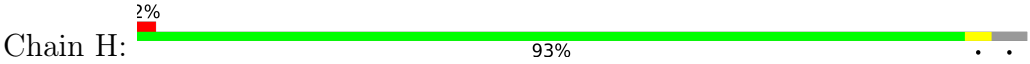
• Molecule 1: Protein/nucleic acid deglycase HchA

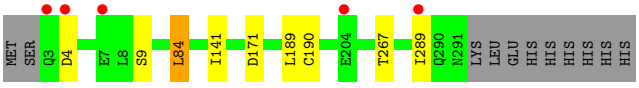




• Molecule 1: Protein/nucleic acid deglycase HchA

• Molecule 1: Protein/nucleic acid deglycase HchA





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	96.28Å 129.45Å 187.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.65 – 2.60 39.38 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.9 (48.65-2.60) 98.9 (39.38-2.60)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	5.67 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.181 , 0.254 0.187 , 0.252	Depositor DCC
R_{free} test set	3580 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	39.4	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	18184	wwPDB-VP
Average B, all atoms (Å^2)	41.0	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 29.55 % 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 1.5202e-03. 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, 0.333 respectively for untwinned datasets, and 0.375, 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: LAC, ZN

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	0.70	0/2271	0.80	0/3081
1	B	0.61	0/2261	0.77	0/3069
1	C	0.70	0/2354	0.80	0/3194
1	D	0.61	0/2245	0.77	0/3047
1	E	0.66	0/2271	0.79	0/3081
1	F	0.61	0/2237	0.74	0/3036
1	G	0.69	0/2354	0.80	0/3194
1	H	0.61	0/2286	0.75	0/3103
All	All	0.65	0/18279	0.78	0/24805

There are no bond length outliers.

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	2224	0	2221	11	0
1	B	2214	0	2206	15	0
1	C	2301	0	2275	15	0
1	D	2198	0	2189	19	0
1	E	2224	0	2216	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2190	0	2185	10	0
1	G	2301	0	2275	6	0
1	H	2239	0	2224	3	0
2	A	6	0	0	0	0
2	E	6	0	0	0	0
3	C	1	0	0	0	0
4	A	64	0	0	0	0
4	B	21	0	0	0	0
4	C	55	0	0	0	0
4	D	21	0	0	0	0
4	E	30	0	0	0	0
4	F	17	0	0	0	0
4	G	50	0	0	1	0
4	H	22	0	0	0	0
All	All	18184	0	17791	88	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 2.

All (88) 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:190:CYS:O	1:A:193:PRO:HD2	1.80	0.81
1:C:270:SER:HB2	1:C:271:PRO:HD2	1.69	0.72
1:G:75:THR:HG21	1:G:160:HIS:HB2	1.70	0.72
1:C:270:SER:HB2	1:C:271:PRO:CD	2.20	0.71
1:C:190:CYS:SG	1:C:191:HIS:N	2.64	0.71
1:B:158:GLY:HA3	1:B:190:CYS:HB3	1.74	0.68
1:F:189:LEU:HD22	1:F:277:LEU:HD23	1.86	0.57
1:D:218:ASP:OD1	1:D:240:ALA:HB3	2.05	0.56
1:D:53:TRP:CE3	1:D:152:LEU:HD23	2.41	0.56
1:F:179:ALA:HB1	1:F:186:ILE:HD11	1.88	0.55
1:A:32:TYR:CZ	1:A:228:ILE:HD11	2.41	0.55
1:D:213:VAL:HG12	1:D:267:THR:HG21	1.89	0.54
1:A:190:CYS:SG	1:A:191:HIS:N	2.81	0.54
1:A:75:THR:OG1	1:A:76:GLY:N	2.41	0.53
1:A:32:TYR:CE2	1:A:228:ILE:HD11	2.43	0.53
1:A:37:THR:HG21	1:A:84:LEU:HD22	1.90	0.53
1:A:39:PHE:HB2	1:A:84:LEU:HD11	1.90	0.53
1:F:179:ALA:CB	1:F:186:ILE:HD11	2.39	0.52
1:E:160:HIS:O	1:E:163:VAL:HG22	2.10	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:244:THR:HG23	1:D:250:VAL:CG2	2.41	0.50
1:D:189:LEU:HD22	1:D:277:LEU:CD2	2.41	0.50
1:A:28:SER:HA	1:A:228:ILE:HD12	1.93	0.50
1:E:188:THR:O	1:E:267:THR:HA	2.12	0.50
1:A:156:ILE:HD12	1:A:156:ILE:N	2.26	0.49
1:B:84:LEU:HB3	1:B:85:PRO:HD3	1.95	0.49
1:B:153:SER:HB3	1:B:185:PHE:HB2	1.95	0.49
1:C:62:GLU:HG2	1:C:164:VAL:HG11	1.95	0.49
1:F:75:THR:OG1	1:F:109:GLU:OE2	2.31	0.48
1:D:244:THR:HG23	1:D:250:VAL:HG23	1.95	0.48
1:E:58:ILE:HD11	1:E:175:THR:HG21	1.94	0.48
1:G:180:LEU:O	4:G:401:HOH:O	2.20	0.47
1:E:162:ALA:HB1	1:E:166:ILE:HD11	1.96	0.47
1:D:270:SER:HB2	1:D:271:PRO:HD2	1.98	0.46
1:C:190:CYS:HG	1:C:191:HIS:N	2.12	0.46
1:D:156:ILE:HB	1:D:188:THR:HG23	1.97	0.46
1:C:73:PHE:CE2	1:C:75:THR:HB	2.50	0.46
1:G:189:LEU:N	1:G:189:LEU:HD23	2.31	0.46
1:C:190:CYS:O	1:C:193:PRO:HD2	2.16	0.45
1:D:89:LEU:HD23	1:D:282:VAL:HG22	1.97	0.45
1:F:86:LEU:HD13	1:F:127:LEU:HD11	1.97	0.45
1:E:75:THR:OG1	1:E:76:GLY:N	2.48	0.45
1:F:217:PRO:HG2	1:F:220:LEU:HD12	1.97	0.45
1:B:153:SER:CB	1:B:185:PHE:HB2	2.47	0.45
1:D:32:TYR:CE2	1:D:228:ILE:HD11	2.52	0.45
1:D:215:VAL:CG2	1:D:240:ALA:HB2	2.47	0.44
1:B:73:PHE:CE2	1:B:75:THR:HB	2.52	0.44
1:D:189:LEU:HD23	1:D:189:LEU:N	2.33	0.44
1:D:156:ILE:HD13	1:D:195:ALA:CB	2.47	0.44
1:C:81:GLU:OE2	1:C:190:CYS:HB2	2.18	0.44
1:C:9:SER:OG	1:C:11:GLN:HB2	2.18	0.43
1:D:53:TRP:CD2	1:D:152:LEU:HD23	2.53	0.43
1:D:84:LEU:HB3	1:D:85:PRO:HD3	2.00	0.43
1:D:32:TYR:HE2	1:D:228:ILE:HD11	1.82	0.43
1:A:101:LEU:O	1:A:136:LYS:NZ	2.41	0.43
1:F:270:SER:HB2	1:F:271:PRO:HD2	2.01	0.43
1:D:86:LEU:HD21	1:D:155:PHE:CD2	2.54	0.43
1:B:32:TYR:CZ	1:B:228:ILE:HD11	2.54	0.42
1:H:141:ILE:HD11	1:H:171:ASP:O	2.19	0.42
1:B:189:LEU:HD22	1:B:277:LEU:CD2	2.49	0.42
1:A:151:TYR:O	1:A:184:ARG:HD2	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:237:TRP:CE3	1:C:242:LEU:HD22	2.54	0.42
1:G:176:LEU:HD22	1:G:186:ILE:HD13	2.02	0.42
1:B:156:ILE:HD13	1:B:195:ALA:CB	2.49	0.42
1:C:213:VAL:HG12	1:C:267:THR:HG21	2.01	0.42
1:B:213:VAL:CG1	1:B:265:LEU:HD21	2.50	0.42
1:B:32:TYR:CE2	1:B:228:ILE:HD11	2.54	0.42
1:B:188:THR:HG23	1:B:267:THR:HG22	2.02	0.42
1:E:190:CYS:SG	1:E:191:HIS:N	2.92	0.42
1:B:81:GLU:OE1	1:B:190:CYS:HB2	2.20	0.42
1:H:84:LEU:HD12	1:H:84:LEU:HA	1.81	0.42
1:H:189:LEU:N	1:H:189:LEU:HD23	2.35	0.42
1:F:192:GLY:N	1:F:193:PRO:CD	2.83	0.41
1:G:190:CYS:SG	1:G:191:HIS:N	2.93	0.41
1:C:80:VAL:HG22	1:C:114:PRO:HG3	2.02	0.41
1:B:251:VAL:HG11	1:B:260:LEU:HD22	2.01	0.41
1:C:153:SER:OG	1:C:185:PHE:HB2	2.21	0.41
1:D:208:LEU:O	1:D:248:LEU:HD22	2.20	0.41
1:B:81:GLU:CD	1:B:190:CYS:HB2	2.41	0.41
1:E:36:LYS:HA	1:E:117:ASP:HB2	2.02	0.41
1:G:138:ALA:HA	1:G:141:ILE:HD12	2.03	0.41
1:E:86:LEU:HB3	1:E:90:MET:HE2	2.03	0.41
1:F:32:TYR:CE2	1:F:228:ILE:HD11	2.56	0.41
1:C:84:LEU:HB3	1:C:85:PRO:HD3	2.03	0.41
1:F:84:LEU:HB3	1:F:85:PRO:HD3	2.03	0.41
1:C:178:TRP:CD1	1:C:182:ASN:ND2	2.90	0.40
1:E:213:VAL:HG12	1:E:267:THR:HG21	2.03	0.40
1:D:215:VAL:HG23	1:D:240:ALA:HB2	2.03	0.40
1:B:48:TYR:OH	1:B:50:ASP:HB2	2.21	0.40

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	285/300 (95%)	265 (93%)	20 (7%)	0	100	100
1	B	284/300 (95%)	266 (94%)	17 (6%)	1 (0%)	34	57
1	C	293/300 (98%)	278 (95%)	15 (5%)	0	100	100
1	D	282/300 (94%)	260 (92%)	20 (7%)	2 (1%)	22	43
1	E	285/300 (95%)	272 (95%)	12 (4%)	1 (0%)	34	57
1	F	281/300 (94%)	265 (94%)	14 (5%)	2 (1%)	22	43
1	G	293/300 (98%)	279 (95%)	14 (5%)	0	100	100
1	H	287/300 (96%)	274 (96%)	12 (4%)	1 (0%)	41	64
All	All	2290/2400 (95%)	2159 (94%)	124 (5%)	7 (0%)	41	64

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	190	CYS
1	B	45	LYS
1	D	288	ALA
1	H	190	CYS
1	D	260	LEU
1	E	190	CYS
1	F	207	PRO

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	244/257 (95%)	241 (99%)	3 (1%)	71	87
1	B	243/257 (95%)	234 (96%)	9 (4%)	34	60
1	C	252/257 (98%)	246 (98%)	6 (2%)	49	74
1	D	241/257 (94%)	234 (97%)	7 (3%)	42	68
1	E	244/257 (95%)	241 (99%)	3 (1%)	71	87
1	F	240/257 (93%)	232 (97%)	8 (3%)	38	64
1	G	252/257 (98%)	247 (98%)	5 (2%)	55	78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	H	246/257 (96%)	241 (98%)	5 (2%)	55 78
All	All	1962/2056 (95%)	1916 (98%)	46 (2%)	50 75

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	ASP
1	A	75	THR
1	A	251	VAL
1	B	5	VAL
1	B	40	ASP
1	B	50	ASP
1	B	126	LYS
1	B	136	LYS
1	B	188	THR
1	B	196	LEU
1	B	251	VAL
1	B	280	LEU
1	C	11	GLN
1	C	102	SER
1	C	115	THR
1	C	183	ASP
1	C	201	LEU
1	C	251	VAL
1	D	55	VAL
1	D	131	LEU
1	D	152	LEU
1	D	188	THR
1	D	189	LEU
1	D	214	CYS
1	D	248	LEU
1	E	6	ASN
1	E	49	LYS
1	E	65	VAL
1	F	49	LYS
1	F	116	GLU
1	F	128	LYS
1	F	169	SER
1	F	205	LYS
1	F	214	CYS
1	F	219	SER
1	F	280	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	75	THR
1	G	122	SER
1	G	128	LYS
1	G	154	VAL
1	G	277	LEU
1	H	4	ASP
1	H	9	SER
1	H	84	LEU
1	H	267	THR
1	H	289	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	182	ASN
1	D	78	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	LAC	A	401	-	5,5,5	1.33	1 (20%)	4,6,6	1.51	1 (25%)
2	LAC	E	401	-	5,5,5	0.81	0	4,6,6	0.76	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	LAC	A	401	-	-	4/4/4/4	-
2	LAC	E	401	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	LAC	O-C	2.23	1.28	1.22

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	LAC	OXT-C-O	2.46	129.67	124.09

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	LAC	O-C-CA-CB
2	A	401	LAC	O-C-CA-OHN
2	A	401	LAC	OXT-C-CA-CB
2	A	401	LAC	OXT-C-CA-OHN
2	E	401	LAC	O-C-CA-CB
2	E	401	LAC	OXT-C-CA-CB
2	E	401	LAC	OXT-C-CA-OHN
2	E	401	LAC	O-C-CA-OHN

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

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	287/300 (95%)	-0.53	2 (0%) 87 86	16, 30, 46, 81	0
1	B	286/300 (95%)	-0.28	0 100 100	22, 46, 64, 81	0
1	C	295/300 (98%)	-0.56	2 (0%) 87 86	18, 32, 52, 90	0
1	D	284/300 (94%)	-0.23	2 (0%) 87 86	28, 48, 70, 84	0
1	E	287/300 (95%)	-0.42	2 (0%) 87 86	21, 38, 62, 83	0
1	F	283/300 (94%)	-0.26	5 (1%) 68 64	23, 48, 70, 87	0
1	G	295/300 (98%)	-0.59	0 100 100	19, 32, 49, 86	0
1	H	289/300 (96%)	-0.18	5 (1%) 70 66	23, 47, 72, 96	0
All	All	2306/2400 (96%)	-0.38	18 (0%) 86 84	16, 39, 65, 96	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	3	GLN	4.3
1	F	146	GLY	3.7
1	F	8	LEU	3.1
1	H	4	ASP	2.7
1	F	148	ASP	2.5
1	H	204	GLU	2.4
1	D	51	GLY	2.4
1	E	6	ASN	2.4
1	D	227	GLU	2.3
1	C	300	HIS	2.3
1	H	289	ILE	2.2
1	E	292	LYS	2.1
1	H	7	GLU	2.1
1	A	292	LYS	2.1
1	F	50	ASP	2.0
1	A	291	ASN	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	190	CYS	2.0
1	F	209	GLU	2.0

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 monosaccharides 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, 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	B-factors(Å^2)	Q<0.9
2	LAC	E	401	6/6	0.89	0.17	34,36,37,38	0
2	LAC	A	401	6/6	0.91	0.16	30,35,37,37	0
3	ZN	C	401	1/1	0.97	0.06	52,52,52,52	0

6.5 Other polymers (i)

There are no such residues in this entry.