W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 19, 2023 – 03:01 AM EDT

PDB ID : 2O73
Title : Structure of OHCU decarboxylase in complex with allantoin
Authors : Cendron, L.; Berni, R.; Folli, C.; Ramazzina, I.; Percudani, R.; Zanotti, G.
Deposited on : 2006-12-10
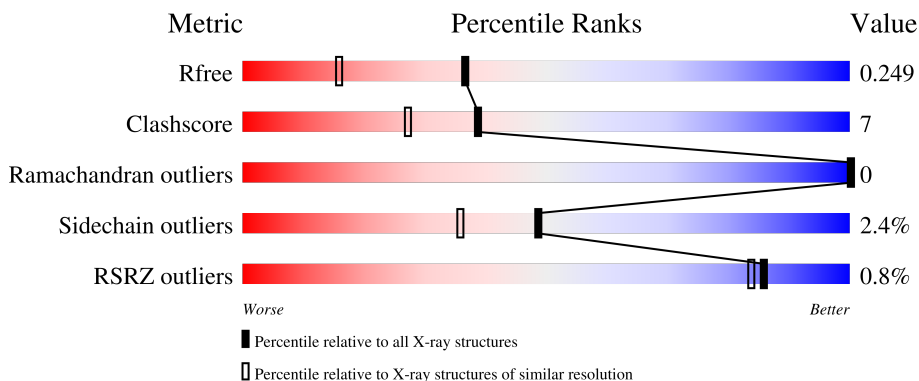Resolution : 1.80 Å(reported)

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 1.80 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	Percentile Ranks	Value
Rfree		0.249
Clashscore		7
Ramachandran outliers		0
Sidechain outliers		2.4%
RSRZ outliers		0.8%

Worse .... *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	174	2% 73% 21% • 5%
1	B	174	74% 20% • 6%
1	C	174	% 84% 11% • •
1	D	174	82% 11% • 6%
1	E	174	2% 81% 11% • 6%

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

| Mol	Chain	Length	Quality of chain
1	F	174	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8511 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 OHCU decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	165	Total	C	N	O	S	0	0	0
			1313	826	231	249	7			
1	B	164	Total	C	N	O	S	0	0	0
			1306	821	230	248	7			
1	C	167	Total	C	N	O	S	0	0	0
			1329	837	233	251	8			
1	D	164	Total	C	N	O	S	0	0	0
			1306	821	230	248	7			
1	E	164	Total	C	N	O	S	0	0	0
			1306	821	230	248	7			
1	F	168	Total	C	N	O	S	0	0	0
			1335	840	234	253	8			

- Molecule 2 is 1-(2,5-DIOXO-2,5-DIHYDRO-1H-IMIDAZOL-4-YL)UREA (three-letter code: 2AL) (formula: $C_4H_4N_4O_3$).
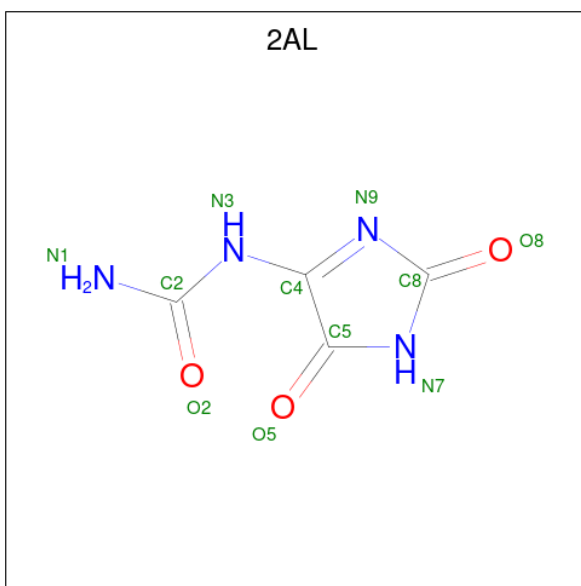
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 11	C 4	N 4	O 3	0	0
2	B	1	Total 11	C 4	N 4	O 3	0	0
2	C	1	Total 11	C 4	N 4	O 3	0	0
2	D	1	Total 11	C 4	N 4	O 3	0	0
2	E	1	Total 11	C 4	N 4	O 3	0	0
2	F	1	Total 11	C 4	N 4	O 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	69	Total 69	O 69	0	0
3	B	67	Total 67	O 67	0	0
3	C	133	Total 133	O 133	0	0
3	D	88	Total 88	O 88	0	0
3	E	63	Total 63	O 63	0	0
3	F	130	Total 130	O 130	0	0

3 Residue-property plots ⓘ

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 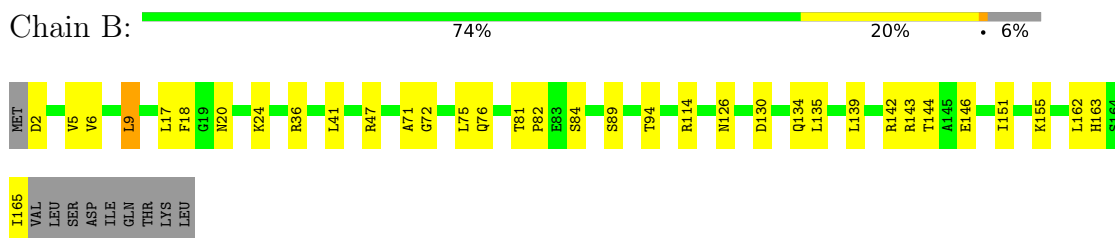outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: OHCU decarboxylase
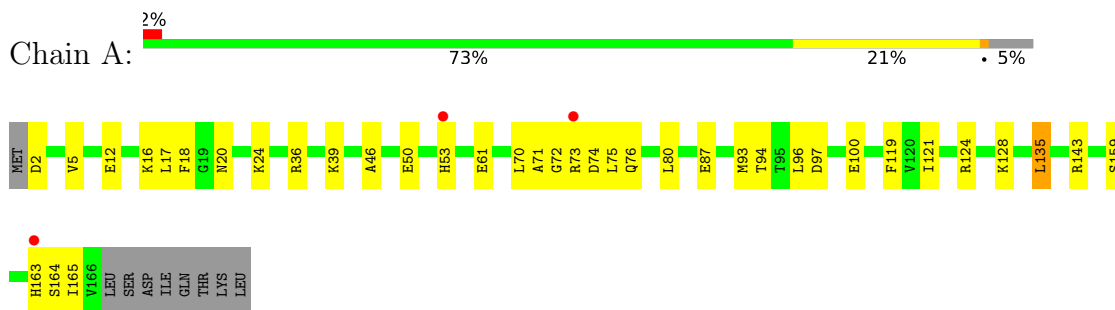
Chain A:

- Molecule 1: OHCU decarboxylase

Chain B:

- Molecule 1: OHCU decarboxylase

Chain C:

- Molecule 1: OHCU decarboxylase

Chain D:

- Molecule 1: OHCU decarboxylase

Chain E:

MET D2 V5 N20 K24 L27 I28 S29 L41 H53 K60 R73 D74 L75 Q76 S77 L80 T81 P82 Q85 E86 T94 A99 L107 R114 N126 D130 L135 R142 R143 H163 S164 I165 VAL LEU SER ASP ILE GLN THR LYS
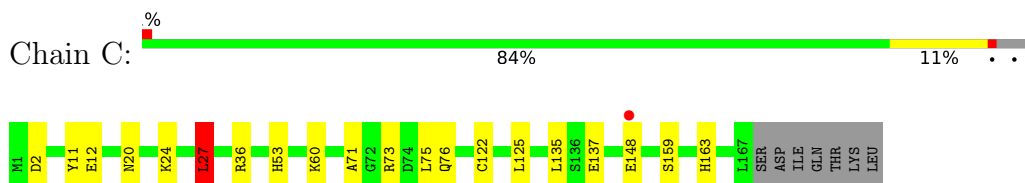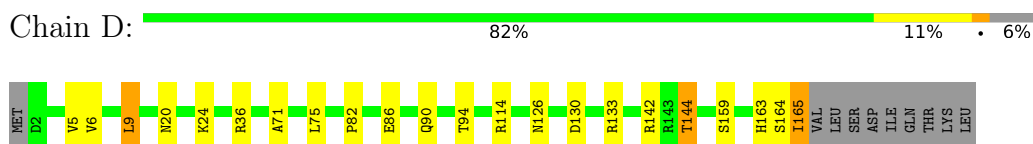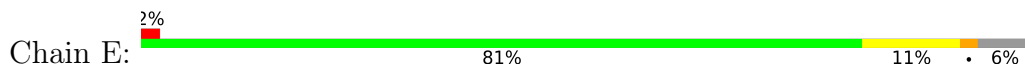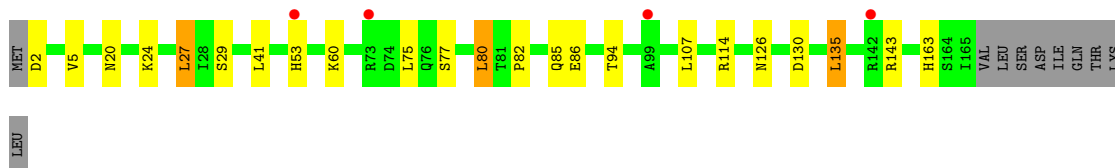
LEU

• Molecule 1: OHCU decarboxylase

Chain F:

86%   10%

M1 N20 K24 I32 I48 A71 L75 T94 T95 L107 F119 N126 D130 R133 Q134 L135 S136 R143 S168 ASP ILE GLN THR LYS LEU

4   Data and refinement statistics   (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	102.21Å 102.21Å 103.87Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	83.00 – 1.80 67.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	93.0 (83.00-1.80) 93.0 (67.37-1.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.70 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.217 , 0.252 0.216 , 0.249	Depositor DCC
R_{free} test set	5205 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	24.3	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 41.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -h,-k,l 0.024 for h,-h-k,-l 0.019 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8511	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.74% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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: 2AL

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	0.45	0/1335	0.54	0/1801
1	B	0.45	0/1328	0.58	0/1791
1	C	0.53	0/1351	0.61	1/1822 (0.1%)
1	D	0.50	0/1328	0.56	0/1791
1	E	0.45	0/1328	0.58	1/1791 (0.1%)
1	F	0.53	0/1357	0.58	1/1830 (0.1%)
All	All	0.49	0/8027	0.58	3/10826 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	27	LEU	CA-CB-CG	7.32	132.14	115.30
1	E	27	LEU	CA-CB-CG	7.06	131.54	115.30
1	F	135	LEU	CA-CB-CG	5.06	126.94	115.30

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	1313	0	1314	35	0
1	B	1306	0	1305	22	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1329	0	1337	18	0
1	D	1306	0	1305	18	0
1	E	1306	0	1305	12	0
1	F	1335	0	1342	14	0
2	A	11	0	4	0	0
2	B	11	0	4	0	0
2	C	11	0	4	0	0
2	D	11	0	4	0	0
2	E	11	0	4	0	0
2	F	11	0	4	0	0
3	A	69	0	0	5	0
3	B	67	0	0	4	0
3	C	133	0	0	8	0
3	D	88	0	0	4	0
3	E	63	0	0	0	0
3	F	130	0	0	4	0
All	All	8511	0	7932	119	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 7.

The worst 5 of 119 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:F:136:SER:HB3	3:F:3109:HOH:O	1.41	1.20
1:C:76:GLN:HG3	3:C:4005:HOH:O	1.65	0.96
1:A:96:LEU:HD21	3:A:6068:HOH:O	1.70	0.91
1:A:53:HIS:HD2	3:A:6055:HOH:O	1.55	0.89
1:F:95:THR:HG23	3:F:3016:HOH:O	1.73	0.89

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	163/174 (94%)	162 (99%)	1 (1%)	0	100	100
1	B	162/174 (93%)	162 (100%)	0	0	100	100
1	C	165/174 (95%)	163 (99%)	2 (1%)	0	100	100
1	D	162/174 (93%)	161 (99%)	1 (1%)	0	100	100
1	E	162/174 (93%)	161 (99%)	1 (1%)	0	100	100
1	F	166/174 (95%)	164 (99%)	2 (1%)	0	100	100
All	All	980/1044 (94%)	973 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	145/154 (94%)	141 (97%)	4 (3%)	43	30
1	B	144/154 (94%)	140 (97%)	4 (3%)	43	30
1	C	147/154 (96%)	144 (98%)	3 (2%)	55	44
1	D	144/154 (94%)	141 (98%)	3 (2%)	53	42
1	E	144/154 (94%)	139 (96%)	5 (4%)	36	21
1	F	148/154 (96%)	146 (99%)	2 (1%)	67	59
All	All	872/924 (94%)	851 (98%)	21 (2%)	49	36

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	27	LEU
1	E	80	LEU
1	F	135	LEU
1	E	135	LEU
1	E	77	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	126	ASN
1	E	20	ASN
1	E	4	ASN
1	E	103	HIS
1	B	103	HIS

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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

6 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	$\#|Z| > 2$
2	2AL	C	4001	-	9,11,11	3.66	3 (33%)	7,15,15	1.92	2 (28%)
2	2AL	F	3001	-	9,11,11	3.37	4 (44%)	7,15,15	1.97	2 (28%)
2	2AL	E	2001	-	9,11,11	3.59	3 (33%)	7,15,15	1.95	2 (28%)
2	2AL	D	5001	-	9,11,11	3.81	4 (44%)	7,15,15	1.88	2 (28%)
2	2AL	B	1001	-	9,11,11	3.80	5 (55%)	7,15,15	2.20	2 (28%)
2	2AL	A	6001	-	9,11,11	3.49	5 (55%)	7,15,15	2.01	2 (28%)

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	2AL	C	4001	-	-	3/4/16/16	0/1/1/1
2	2AL	F	3001	-	-	3/4/16/16	0/1/1/1
2	2AL	E	2001	-	-	3/4/16/16	0/1/1/1
2	2AL	D	5001	-	-	3/4/16/16	0/1/1/1
2	2AL	B	1001	-	-	3/4/16/16	0/1/1/1
2	2AL	A	6001	-	-	3/4/16/16	0/1/1/1

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	5001	2AL	C4-N9	8.65	1.48	1.33
2	B	1001	2AL	C4-N9	8.51	1.48	1.33
2	C	4001	2AL	C4-N9	8.46	1.48	1.33
2	E	2001	2AL	C4-N9	7.94	1.47	1.33
2	F	3001	2AL	C4-N9	7.49	1.46	1.33

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1001	2AL	C4-C5-N7	4.39	109.09	104.47
2	D	5001	2AL	C4-C5-N7	4.10	108.79	104.47
2	A	6001	2AL	C4-C5-N7	3.89	108.56	104.47
2	E	2001	2AL	C4-C5-N7	3.80	108.47	104.47
2	C	4001	2AL	C4-C5-N7	3.77	108.44	104.47

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	6001	2AL	N1-C2-N3-C4
2	A	6001	2AL	O2-C2-N3-C4
2	B	1001	2AL	N1-C2-N3-C4
2	B	1001	2AL	O2-C2-N3-C4
2	C	4001	2AL	N1-C2-N3-C4

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, 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	165/174 (94%)	0.20	3 (1%)	68	64	15, 25, 32, 38	0
1	B	164/174 (94%)	-0.01	0	100	100	15, 24, 32, 36	0
1	C	167/174 (95%)	0.08	1 (0%)	89	87	12, 19, 30, 40	0
1	D	164/174 (94%)	-0.02	0	100	100	12, 22, 31, 36	0
1	E	164/174 (94%)	0.09	4 (2%)	59	54	14, 24, 33, 35	0
1	F	168/174 (96%)	0.05	0	100	100	11, 20, 28, 31	0
All	All	992/1044 (95%)	0.07	8 (0%)	86	84	11, 22, 32, 40	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	73	ARG	2.7
1	A	53	HIS	2.2
1	E	53	HIS	2.2
1	E	142	ARG	2.2
1	A	163	HIS	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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, 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	B-factors(Å^2)	Q<0.9
2	2AL	A	6001	11/11	0.88	0.12	21,23,24,24	0
2	2AL	E	2001	11/11	0.91	0.10	20,22,23,24	0
2	2AL	B	1001	11/11	0.94	0.09	20,22,23,23	0
2	2AL	F	3001	11/11	0.94	0.11	17,19,20,20	0
2	2AL	C	4001	11/11	0.96	0.09	17,20,20,20	0
2	2AL	D	5001	11/11	0.96	0.08	18,20,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.