Full wwPDB X-ray Structure Validation Report (i)

Oct 23, 2016 – 03:50 PM EDT

PDB ID	:	5FAU
Title	:	wild-type choline TMA lyase in complex with choline
Authors	:	Funk, M.A.; Drennan, C.L.
Deposited on	:	2015-12-12
Resolution	:	1.90 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:

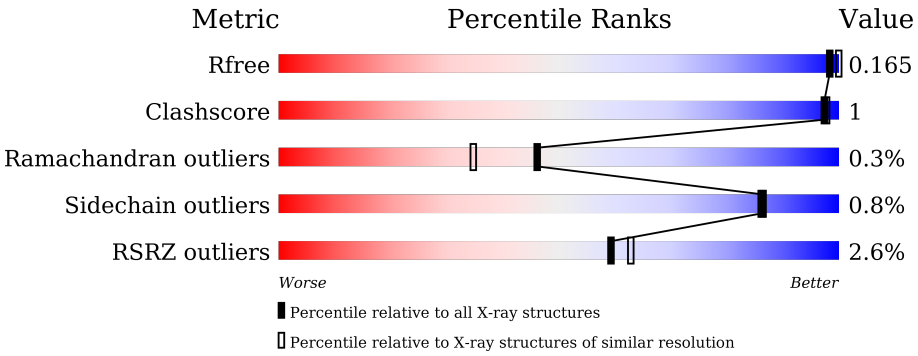
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

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 1.90 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	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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 <=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	849	7% 89% · 6%
1	B	849	% 90% · 6%
1	C	849	91% · 6%
1	D	849	% 90% · 6%

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	B	902	-	-	-	X
3	GOL	C	904	-	-	-	X
3	GOL	D	903	-	-	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 52391 atoms, of which 24722 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 Choline trimethylamine-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	794	Total	C	H	N	O	S	0	0	0
			12401	3971	6138	1062	1185	45			
1	B	795	Total	C	H	N	O	S	0	2	0
			12428	3979	6150	1064	1190	45			
1	C	795	Total	C	H	N	O	S	0	2	0
			12426	3978	6149	1064	1190	45			
1	D	795	Total	C	H	N	O	S	0	3	0
			12454	3985	6165	1068	1191	45			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	initiating methionine	UNP Q30W70
A	-1	GLY	-	expression tag	UNP Q30W70
A	0	SER	-	expression tag	UNP Q30W70
A	1	SER	-	expression tag	UNP Q30W70
A	2	HIS	-	expression tag	UNP Q30W70
A	3	HIS	-	expression tag	UNP Q30W70
A	4	HIS	-	expression tag	UNP Q30W70
A	5	HIS	-	expression tag	UNP Q30W70
A	6	HIS	-	expression tag	UNP Q30W70
A	7	HIS	-	expression tag	UNP Q30W70
A	8	SER	-	expression tag	UNP Q30W70
A	9	SER	-	expression tag	UNP Q30W70
A	10	GLY	-	expression tag	UNP Q30W70
A	11	LEU	-	expression tag	UNP Q30W70
A	12	VAL	-	expression tag	UNP Q30W70
A	13	PRO	-	expression tag	UNP Q30W70
A	14	ARG	-	expression tag	UNP Q30W70
A	15	GLY	-	expression tag	UNP Q30W70
A	16	SER	-	expression tag	UNP Q30W70
A	17	HIS	-	expression tag	UNP Q30W70
A	18	MET	-	expression tag	UNP Q30W70

Continued on next page...

Continued from previous page...

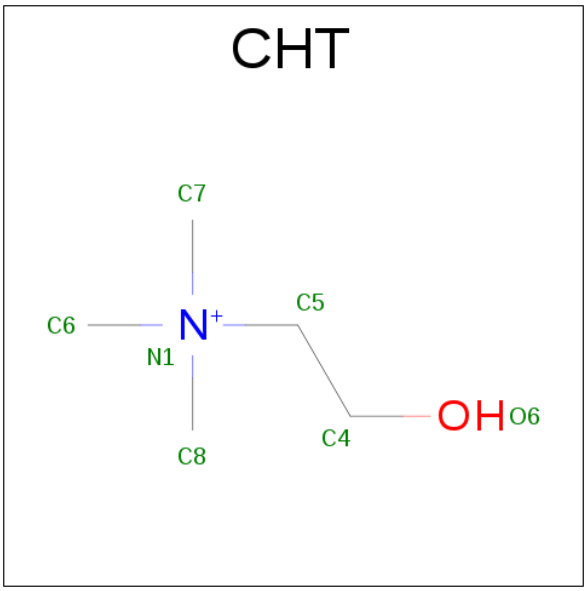
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	initiating methionine	UNP Q30W70
B	-1	GLY	-	expression tag	UNP Q30W70
B	0	SER	-	expression tag	UNP Q30W70
B	1	SER	-	expression tag	UNP Q30W70
B	2	HIS	-	expression tag	UNP Q30W70
B	3	HIS	-	expression tag	UNP Q30W70
B	4	HIS	-	expression tag	UNP Q30W70
B	5	HIS	-	expression tag	UNP Q30W70
B	6	HIS	-	expression tag	UNP Q30W70
B	7	HIS	-	expression tag	UNP Q30W70
B	8	SER	-	expression tag	UNP Q30W70
B	9	SER	-	expression tag	UNP Q30W70
B	10	GLY	-	expression tag	UNP Q30W70
B	11	LEU	-	expression tag	UNP Q30W70
B	12	VAL	-	expression tag	UNP Q30W70
B	13	PRO	-	expression tag	UNP Q30W70
B	14	ARG	-	expression tag	UNP Q30W70
B	15	GLY	-	expression tag	UNP Q30W70
B	16	SER	-	expression tag	UNP Q30W70
B	17	HIS	-	expression tag	UNP Q30W70
B	18	MET	-	expression tag	UNP Q30W70
C	-2	MET	-	initiating methionine	UNP Q30W70
C	-1	GLY	-	expression tag	UNP Q30W70
C	0	SER	-	expression tag	UNP Q30W70
C	1	SER	-	expression tag	UNP Q30W70
C	2	HIS	-	expression tag	UNP Q30W70
C	3	HIS	-	expression tag	UNP Q30W70
C	4	HIS	-	expression tag	UNP Q30W70
C	5	HIS	-	expression tag	UNP Q30W70
C	6	HIS	-	expression tag	UNP Q30W70
C	7	HIS	-	expression tag	UNP Q30W70
C	8	SER	-	expression tag	UNP Q30W70
C	9	SER	-	expression tag	UNP Q30W70
C	10	GLY	-	expression tag	UNP Q30W70
C	11	LEU	-	expression tag	UNP Q30W70
C	12	VAL	-	expression tag	UNP Q30W70
C	13	PRO	-	expression tag	UNP Q30W70
C	14	ARG	-	expression tag	UNP Q30W70
C	15	GLY	-	expression tag	UNP Q30W70
C	16	SER	-	expression tag	UNP Q30W70
C	17	HIS	-	expression tag	UNP Q30W70
C	18	MET	-	expression tag	UNP Q30W70

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	MET	-	initiating methionine	UNP Q30W70
D	-1	GLY	-	expression tag	UNP Q30W70
D	0	SER	-	expression tag	UNP Q30W70
D	1	SER	-	expression tag	UNP Q30W70
D	2	HIS	-	expression tag	UNP Q30W70
D	3	HIS	-	expression tag	UNP Q30W70
D	4	HIS	-	expression tag	UNP Q30W70
D	5	HIS	-	expression tag	UNP Q30W70
D	6	HIS	-	expression tag	UNP Q30W70
D	7	HIS	-	expression tag	UNP Q30W70
D	8	SER	-	expression tag	UNP Q30W70
D	9	SER	-	expression tag	UNP Q30W70
D	10	GLY	-	expression tag	UNP Q30W70
D	11	LEU	-	expression tag	UNP Q30W70
D	12	VAL	-	expression tag	UNP Q30W70
D	13	PRO	-	expression tag	UNP Q30W70
D	14	ARG	-	expression tag	UNP Q30W70
D	15	GLY	-	expression tag	UNP Q30W70
D	16	SER	-	expression tag	UNP Q30W70
D	17	HIS	-	expression tag	UNP Q30W70
D	18	MET	-	expression tag	UNP Q30W70

- Molecule 2 is CHOLINE ION (three-letter code: CHT) (formula: $C_5H_{14}NO$).



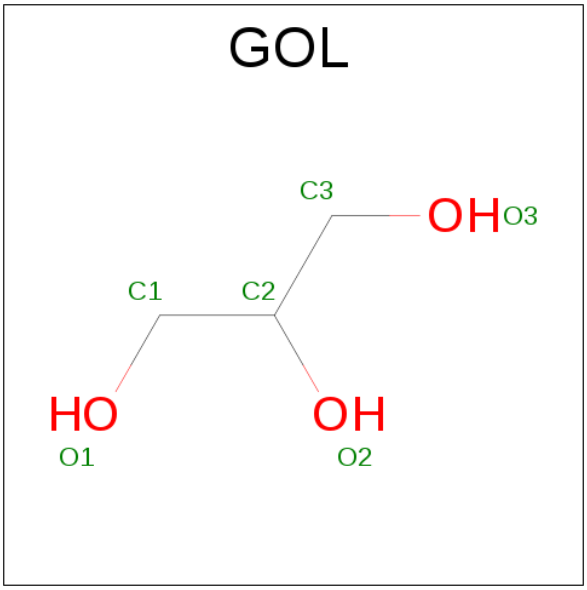
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	A	1	21	5	14	1	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	H	N	O	0	0
			21	5	14	1	1		
2	C	1	Total	C	H	N	O	0	0
			21	5	14	1	1		
2	D	1	Total	C	H	N	O	0	0
			21	5	14	1	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is water.

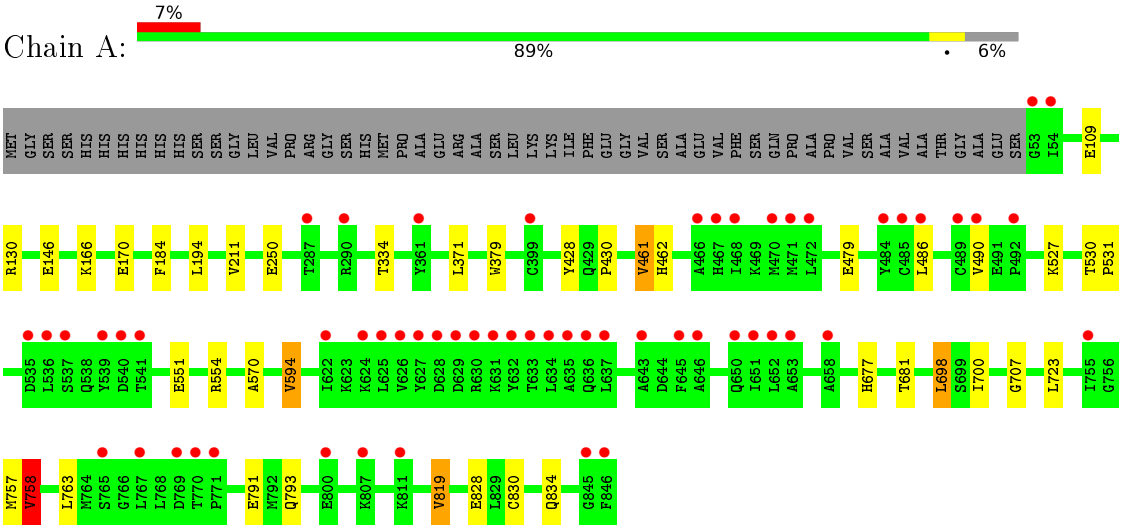
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	477	Total O 477 477	0	0
4	B	634	Total O 634 634	0	0
4	C	666	Total O 666 666	0	0
4	D	709	Total O 709 709	0	0

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.

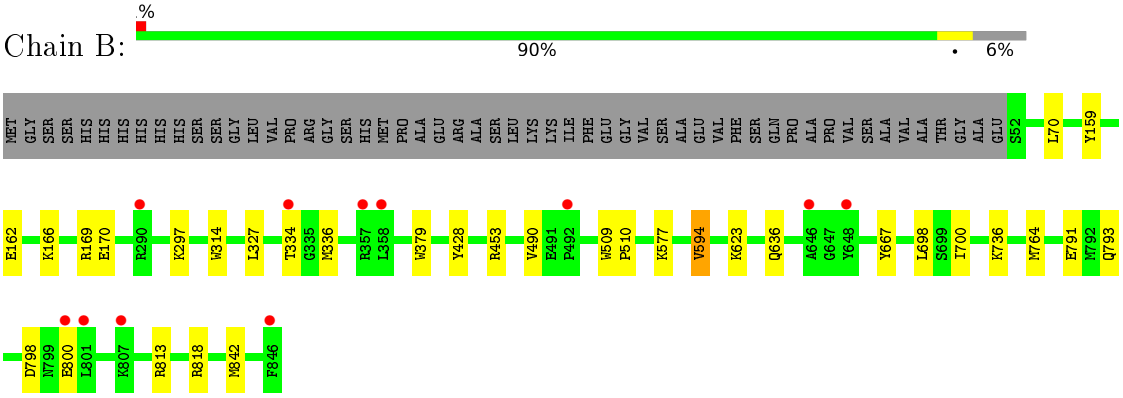
- Molecule 1: Choline trimethylamine-lyase

Chain A:



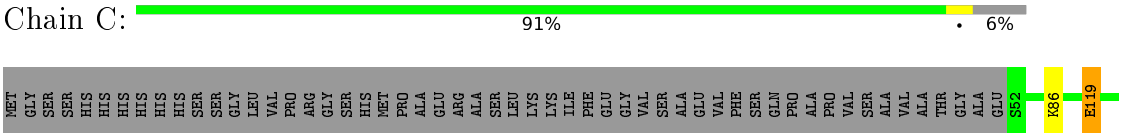
- Molecule 1: Choline trimethylamine-lyase

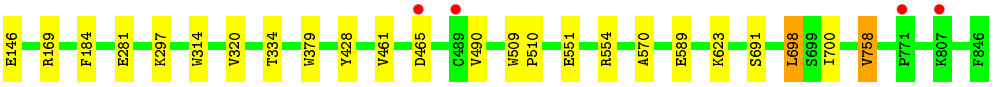
Chain B:



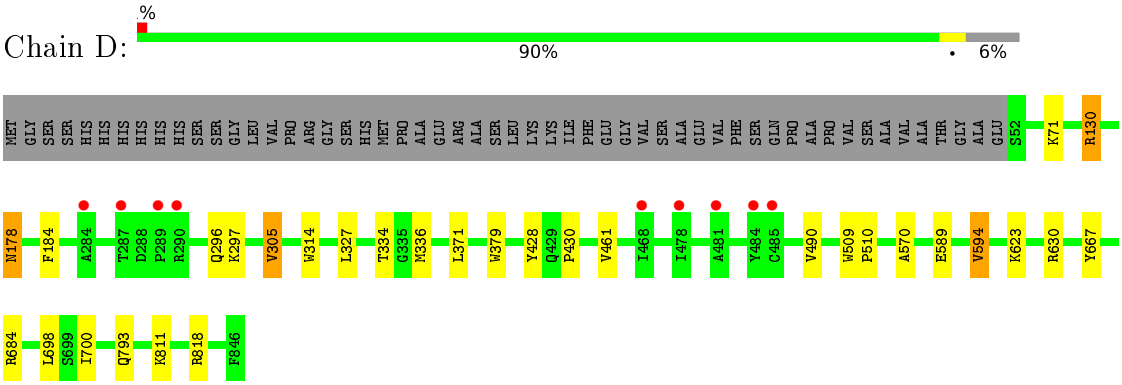
- Molecule 1: Choline trimethylamine-lyase

Chain C:





- Molecule 1: Choline trimethylamine-lyase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	79.61Å 234.87Å 104.99Å 90.00° 109.62° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 49.65 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.00-1.90) 93.4 (49.65-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 1.90Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.145 , 0.169 0.141 , 0.165	Depositor DCC
R_{free} test set	7936 reflections (2.99%)	DCC
Wilson B-factor (Å^2)	26.7	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 40.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.022 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	52391	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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: GOL, CHT

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	0.40	0/6408	0.58	5/8678 (0.1%)
1	B	0.45	0/6431	0.57	2/8709 (0.0%)
1	C	0.48	2/6427 (0.0%)	0.59	3/8704 (0.0%)
1	D	0.49	1/6443 (0.0%)	0.62	9/8724 (0.1%)
All	All	0.46	3/25709 (0.0%)	0.59	19/34815 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	119	GLU	CB-CG	-7.03	1.38	1.52
1	C	589	GLU	CG-CD	6.64	1.61	1.51
1	D	589	GLU	CG-CD	5.10	1.59	1.51

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	461	VAL	CG1-CB-CG2	9.57	126.21	110.90
1	A	758	VAL	CG1-CB-CG2	8.36	124.27	110.90
1	D	305	VAL	CG1-CB-CG2	8.17	123.97	110.90
1	C	461	VAL	CG1-CB-CG2	8.16	123.95	110.90
1	A	594	VAL	CG1-CB-CG2	8.12	123.90	110.90
1	A	819	VAL	CG1-CB-CG2	8.01	123.71	110.90
1	D	594	VAL	CG1-CB-CG2	7.63	123.11	110.90
1	C	758	VAL	CG1-CB-CG2	7.60	123.06	110.90
1	D	461	VAL	CG1-CB-CG2	7.54	122.97	110.90
1	B	594	VAL	CG1-CB-CG2	7.16	122.35	110.90
1	D	130	ARG	NE-CZ-NH2	-6.53	117.04	120.30
1	B	594	VAL	CA-CB-CG2	5.92	119.78	110.90
1	D	130	ARG	NE-CZ-NH1	5.43	123.02	120.30
1	D	305	VAL	CA-CB-CG2	5.37	118.96	110.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	684	ARG	NE-CZ-NH2	-5.26	117.67	120.30
1	D	594	VAL	CA-CB-CG2	5.20	118.70	110.90
1	C	758	VAL	CA-CB-CG1	5.18	118.68	110.90
1	D	461	VAL	CA-CB-CG1	5.18	118.67	110.90
1	A	594	VAL	CA-CB-CG2	5.13	118.60	110.90

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	6263	6138	6137	19	0
1	B	6278	6150	6137	16	1
1	C	6277	6149	6141	12	1
1	D	6289	6165	6152	15	0
2	A	7	14	14	0	0
2	B	7	14	14	0	0
2	C	7	14	14	0	0
2	D	7	14	14	0	0
3	A	6	8	8	0	0
3	B	12	16	16	0	0
3	C	18	24	24	0	0
3	D	12	16	16	0	0
4	A	477	0	0	5	0
4	B	634	0	0	5	0
4	C	666	0	0	6	0
4	D	709	0	0	7	1
All	All	27669	24722	24687	62	2

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 1.

All (62) 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:169:ARG:NH1	4:B:1002:HOH:O	2.15	0.78
1:D:296:GLN:OE1	4:D:1001:HOH:O	2.09	0.71
1:C:465[A]:ASP:OD1	4:C:1001:HOH:O	2.08	0.70
1:C:86:LYS:NZ	4:C:1003:HOH:O	2.24	0.70
1:C:623:LYS:NZ	4:C:1004:HOH:O	2.25	0.69
1:B:736:LYS:NZ	4:B:1003:HOH:O	2.25	0.69
1:A:146:GLU:OE1	4:A:1001:HOH:O	2.11	0.68
1:C:551:GLU:OE2	1:C:554:ARG:NH1	2.26	0.68
1:B:636:GLN:NE2	4:B:1007:HOH:O	2.28	0.65
1:B:162:GLU:OE1	4:B:1001:HOH:O	2.14	0.64
1:C:169:ARG:NH2	4:C:1007:HOH:O	2.33	0.61
1:D:178:ASN:ND2	4:D:1005:HOH:O	2.30	0.60
1:D:811:LYS:NZ	4:D:1012:HOH:O	2.41	0.53
1:A:479:GLU:OE1	4:A:1002:HOH:O	2.18	0.53
1:A:250:GLU:O	4:A:1003:HOH:O	2.18	0.50
1:B:798:ASP:OD2	1:B:800:GLU:HB2	2.11	0.50
1:A:551:GLU:OE2	1:A:554:ARG:NH1	2.45	0.49
1:B:166:LYS:O	1:B:170:GLU:HG2	2.12	0.49
1:D:130:ARG:CD	4:D:1194:HOH:O	2.60	0.49
1:C:146:GLU:OE1	4:C:1002:HOH:O	2.19	0.49
1:D:623:LYS:HE3	1:D:667:TYR:CZ	2.48	0.49
1:C:698:LEU:HA	1:C:758:VAL:HG12	1.94	0.49
1:A:698:LEU:HA	1:A:758:VAL:HG12	1.94	0.49
1:A:211:VAL:O	1:A:211:VAL:HG12	2.13	0.48
1:B:623:LYS:HE3	1:B:667:TYR:CZ	2.47	0.48
1:B:791:GLU:OE2	1:B:793:GLN:NE2	2.45	0.48
1:C:119:GLU:HB2	4:C:1406:HOH:O	2.15	0.46
1:C:184:PHE:CD1	1:C:570:ALA:HB2	2.51	0.46
1:A:707:GLY:HA2	1:A:723:LEU:HD13	1.96	0.46
1:A:334:THR:HB	1:A:379:TRP:CG	2.50	0.46
1:D:371:LEU:HD22	1:D:430:PRO:HD2	1.98	0.45
1:A:530:THR:HB	1:A:531:PRO:CD	2.47	0.45
1:A:130:ARG:NH2	1:A:828:GLU:OE2	2.45	0.45
1:B:297:LYS:HG2	1:B:314:TRP:CZ3	2.52	0.45
1:D:130:ARG:HD2	4:D:1194:HOH:O	2.17	0.45
1:A:677:HIS:O	1:A:681:THR:HG23	2.18	0.44
1:B:453:ARG:NH2	4:B:1028:HOH:O	2.47	0.44
1:A:109:GLU:OE1	4:A:1004:HOH:O	2.21	0.44
1:A:830:CYS:O	1:A:834:GLN:HG2	2.16	0.44
1:D:327:LEU:HB3	1:D:336:MET:SD	2.57	0.44
1:A:166:LYS:O	1:A:170:GLU:HG2	2.18	0.43
1:D:297:LYS:HG2	1:D:314:TRP:CZ3	2.52	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:71:LYS:NZ	4:D:1031:HOH:O	2.50	0.43
1:A:371:LEU:HD22	1:A:430:PRO:HD2	2.00	0.43
1:A:184:PHE:CD1	1:A:570:ALA:HB2	2.54	0.43
1:B:509:TRP:N	1:B:510:PRO:CD	2.82	0.42
1:B:327:LEU:HB3	1:B:336:MET:SD	2.60	0.42
1:D:630:ARG:NH2	4:D:1033:HOH:O	2.52	0.42
1:A:462:HIS:CE1	1:A:486:LEU:HD22	2.54	0.42
1:C:334:THR:HB	1:C:379:TRP:CG	2.54	0.42
1:D:793:GLN:NE2	1:D:818:ARG:HH21	2.18	0.41
1:A:130:ARG:CD	4:A:1032:HOH:O	2.68	0.41
1:B:793:GLN:NE2	1:B:818:ARG:HH21	2.19	0.41
1:A:791:GLU:OE2	1:A:793:GLN:NE2	2.53	0.41
1:C:509:TRP:N	1:C:510:PRO:HD2	2.35	0.41
1:B:334:THR:HB	1:B:379:TRP:CG	2.55	0.41
1:D:334:THR:HB	1:D:379:TRP:CG	2.56	0.41
1:D:509:TRP:N	1:D:510:PRO:CD	2.84	0.41
1:D:184:PHE:CD1	1:D:570:ALA:HB2	2.56	0.41
1:B:70:LEU:O	1:B:813:ARG:NH1	2.54	0.40
1:C:297:LYS:HG2	1:C:314:TRP:CZ3	2.57	0.40
1:B:764:MET:HE2	1:B:842:MET:HG2	2.03	0.40

All (2) 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)
4:D:1407:HOH:O	4:D:1447:HOH:O[1_655]	2.13	0.07
1:B:159:TYR:HH	1:C:281:GLU:OE1[1_556]	1.56	0.04

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	792/849 (93%)	768 (97%)	22 (3%)	2 (0%)	46	35
1	B	795/849 (94%)	774 (97%)	19 (2%)	2 (0%)	46	35
1	C	795/849 (94%)	771 (97%)	21 (3%)	3 (0%)	39	27
1	D	796/849 (94%)	775 (97%)	19 (2%)	2 (0%)	46	35
All	All	3178/3396 (94%)	3088 (97%)	81 (2%)	9 (0%)	46	35

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	700	ILE
1	B	700	ILE
1	C	700	ILE
1	D	700	ILE
1	A	490	VAL
1	B	490	VAL
1	C	691	SER
1	C	490	VAL
1	D	490	VAL

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	669/712 (94%)	659 (98%)	10 (2%)	72	69
1	B	672/712 (94%)	668 (99%)	4 (1%)	90	90
1	C	671/712 (94%)	668 (100%)	3 (0%)	93	94
1	D	673/712 (94%)	668 (99%)	5 (1%)	88	88
All	All	2685/2848 (94%)	2663 (99%)	22 (1%)	86	86

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	194	LEU
1	A	428	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	461	VAL
1	A	527	LYS
1	A	594	VAL
1	A	698	LEU
1	A	757	MET
1	A	758	VAL
1	A	763	LEU
1	A	819	VAL
1	B	428	TYR
1	B	577	LYS
1	B	594	VAL
1	B	698	LEU
1	C	320	VAL
1	C	428	TYR
1	C	698	LEU
1	D	178	ASN
1	D	305	VAL
1	D	428	TYR
1	D	594	VAL
1	D	698	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

12 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	CHT	A	901	-	6,6,6	1.34	1 (16%)	8,8,8	0.62	0
3	GOL	A	902	-	5,5,5	0.32	0	5,5,5	0.24	0
2	CHT	B	901	-	6,6,6	1.16	0	8,8,8	0.62	0
3	GOL	B	902	-	5,5,5	0.40	0	5,5,5	0.14	0
3	GOL	B	903	-	5,5,5	0.35	0	5,5,5	0.31	0
2	CHT	C	901	-	6,6,6	1.25	0	8,8,8	0.30	0
3	GOL	C	902	-	5,5,5	0.38	0	5,5,5	0.32	0
3	GOL	C	903	-	5,5,5	0.35	0	5,5,5	0.47	0
3	GOL	C	904	-	5,5,5	0.57	0	5,5,5	0.50	0
2	CHT	D	901	-	6,6,6	1.33	1 (16%)	8,8,8	0.52	0
3	GOL	D	902	-	5,5,5	0.36	0	5,5,5	0.24	0
3	GOL	D	903	-	5,5,5	0.36	0	5,5,5	0.25	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	CHT	A	901	-	-	0/4/4/4	0/0/0/0
3	GOL	A	902	-	-	0/4/4/4	0/0/0/0
2	CHT	B	901	-	-	0/4/4/4	0/0/0/0
3	GOL	B	902	-	-	0/4/4/4	0/0/0/0
3	GOL	B	903	-	-	0/4/4/4	0/0/0/0
2	CHT	C	901	-	-	0/4/4/4	0/0/0/0
3	GOL	C	902	-	-	0/4/4/4	0/0/0/0
3	GOL	C	903	-	-	0/4/4/4	0/0/0/0
3	GOL	C	904	-	-	0/4/4/4	0/0/0/0
2	CHT	D	901	-	-	0/4/4/4	0/0/0/0
3	GOL	D	902	-	-	0/4/4/4	0/0/0/0
3	GOL	D	903	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	901	CHT	C8-N1	-2.27	1.43	1.50
2	A	901	CHT	C8-N1	-2.11	1.43	1.50

There are no bond angle outliers.

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	794/849 (93%)	0.16	58 (7%)	18	20	25, 40, 62, 85	0
1	B	795/849 (93%)	-0.10	11 (1%)	78	80	22, 33, 50, 62	0
1	C	795/849 (93%)	-0.25	4 (0%)	91	92	20, 30, 46, 59	0
1	D	795/849 (93%)	-0.30	9 (1%)	82	84	20, 28, 44, 70	0
All	All	3179/3396 (93%)	-0.12	82 (2%)	59	63	20, 33, 54, 85	0

All (82) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	630	ARG	4.9
1	A	625	LEU	4.7
1	A	807	LYS	4.6
1	A	290	ARG	3.9
1	A	631	LYS	3.8
1	A	637	LEU	3.8
1	A	540	ASP	3.7
1	A	651	ILE	3.7
1	A	536	LEU	3.6
1	A	468	ILE	3.5
1	D	287	THR	3.4
1	A	53	GLY	3.4
1	A	632	TYR	3.4
1	A	629	ASP	3.4
1	A	486	LEU	3.3
1	A	627	TYR	3.3
1	A	645	PHE	3.2
1	A	535	ASP	3.0
1	A	634	LEU	3.0
1	A	539	TYR	3.0
1	B	846	PHE	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	767	LEU	2.9
1	A	492	PRO	2.9
1	D	468	ILE	2.9
1	A	770	THR	2.8
1	A	771	PRO	2.7
1	A	54	ILE	2.7
1	A	490	VAL	2.7
1	D	289	PRO	2.7
1	A	811	LYS	2.6
1	C	771	PRO	2.6
1	A	652	LEU	2.6
1	A	646	ALA	2.6
1	B	801	LEU	2.6
1	A	633	THR	2.6
1	A	636	GLN	2.6
1	A	846	PHE	2.5
1	D	284	ALA	2.5
1	A	472	LEU	2.5
1	A	845	GLY	2.5
1	A	765	SER	2.5
1	A	484	TYR	2.5
1	B	807	LYS	2.5
1	A	489	CYS	2.5
1	B	357	ARG	2.4
1	D	485	CYS	2.4
1	D	290	ARG	2.3
1	A	635	ALA	2.3
1	A	287	THR	2.3
1	A	485	CYS	2.3
1	A	470	MET	2.3
1	A	626	VAL	2.3
1	A	755	ILE	2.3
1	B	646	ALA	2.3
1	B	800	GLU	2.3
1	A	653	ALA	2.2
1	D	478	ILE	2.2
1	B	648	TYR	2.2
1	A	650	GLN	2.2
1	A	399	CYS	2.2
1	A	466	ALA	2.2
1	A	658	ALA	2.2
1	A	643	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	467	HIS	2.1
1	D	481	ALA	2.1
1	D	484	TYR	2.1
1	C	489	CYS	2.1
1	A	628	ASP	2.1
1	B	334	THR	2.1
1	A	800	GLU	2.1
1	B	492	PRO	2.1
1	A	471	MET	2.1
1	A	622	ILE	2.1
1	A	541	THR	2.1
1	A	537	SER	2.1
1	B	358	LEU	2.1
1	A	624	LYS	2.1
1	A	769	ASP	2.0
1	C	807	LYS	2.0
1	A	361	TYR	2.0
1	B	290	ARG	2.0
1	C	465[A]	ASP	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 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
3	GOL	B	902	6/6	0.82	0.25	5.95	58,69,73,74	0
3	GOL	D	903	6/6	0.63	0.22	4.97	72,89,104,107	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	GOL	C	904	6/6	0.75	0.27	4.24	62,74,77,78	0
3	GOL	B	903	6/6	0.85	0.20	1.63	41,50,50,51	0
3	GOL	A	902	6/6	0.89	0.28	1.51	58,70,72,72	0
3	GOL	D	902	6/6	0.95	0.15	1.51	35,42,44,44	0
3	GOL	C	903	6/6	0.92	0.17	1.01	52,62,63,63	0
2	CHT	B	901	7/7	0.96	0.17	0.48	27,33,34,34	0
3	GOL	C	902	6/6	0.95	0.13	0.33	40,48,52,52	0
2	CHT	A	901	7/7	0.98	0.16	0.22	34,41,43,43	0
2	CHT	D	901	7/7	0.98	0.13	0.11	27,33,35,35	0
2	CHT	C	901	7/7	0.99	0.16	-0.09	24,29,31,31	0

6.5 Other polymers (i)

There are no such residues in this entry.