



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 02:52 AM GMT

PDB ID : 4MVD  
Title : Crystal Structure of a Mammalian Cytidylyltransferase  
Authors : Lee, J.; Cornell, R.B.  
Deposited on : 2013-09-23  
Resolution : 8.00 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

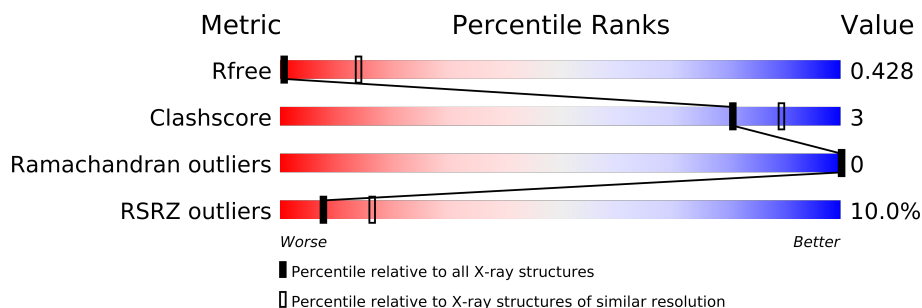
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance









The reported resolution of this entry is 8.00 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))
$R_{free}$	66092	1106 (11.50-3.50)
Clashscore	79885	1007 (12.20-3.54)
Ramachandran outliers	78287	1302 (12.20-3.50)
RSRZ outliers	66119	1105 (11.50-3.50)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	332	
1	B	332	
1	C	332	
1	D	332	
1	E	332	
1	F	332	
1	G	332	
1	H	332	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CDC	B	401	-	X
2	CDC	H	401	-	X

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10261 atoms, of which 0 are hydrogen and 0 are deuterium.

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-phosphate cytidylyltransferase A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	B	253	Total	C	N	O	0	0	0
			1253	747	253	253			
1	A	252	Total	C	N	O	0	0	0
			1249	745	252	252			
1	D	253	Total	C	N	O	0	0	0
			1253	747	253	253			
1	C	253	Total	C	N	O	0	0	0
			1254	748	253	253			
1	F	253	Total	C	N	O	0	0	0
			1253	747	253	253			
1	E	252	Total	C	N	O	0	0	0
			1249	745	252	252			
1	H	253	Total	C	N	O	0	0	0
			1253	747	253	253			
1	G	252	Total	C	N	O	0	0	0
			1249	745	252	252			

There are 160 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	EXPRESSION TAG	UNP P19836
B	-18	GLY	-	EXPRESSION TAG	UNP P19836
B	-17	SER	-	EXPRESSION TAG	UNP P19836
B	-16	SER	-	EXPRESSION TAG	UNP P19836
B	-15	HIS	-	EXPRESSION TAG	UNP P19836
B	-14	HIS	-	EXPRESSION TAG	UNP P19836
B	-13	HIS	-	EXPRESSION TAG	UNP P19836
B	-12	HIS	-	EXPRESSION TAG	UNP P19836
B	-11	HIS	-	EXPRESSION TAG	UNP P19836
B	-10	HIS	-	EXPRESSION TAG	UNP P19836
B	-9	SER	-	EXPRESSION TAG	UNP P19836
B	-8	SER	-	EXPRESSION TAG	UNP P19836
B	-7	GLY	-	EXPRESSION TAG	UNP P19836

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	LEU	-	EXPRESSION TAG	UNP P19836
B	-5	VAL	-	EXPRESSION TAG	UNP P19836
B	-4	PRO	-	EXPRESSION TAG	UNP P19836
B	-3	ARG	-	EXPRESSION TAG	UNP P19836
B	-2	GLY	-	EXPRESSION TAG	UNP P19836
B	-1	SER	-	EXPRESSION TAG	UNP P19836
B	0	HIS	-	EXPRESSION TAG	UNP P19836
A	-19	MET	-	EXPRESSION TAG	UNP P19836
A	-18	GLY	-	EXPRESSION TAG	UNP P19836
A	-17	SER	-	EXPRESSION TAG	UNP P19836
A	-16	SER	-	EXPRESSION TAG	UNP P19836
A	-15	HIS	-	EXPRESSION TAG	UNP P19836
A	-14	HIS	-	EXPRESSION TAG	UNP P19836
A	-13	HIS	-	EXPRESSION TAG	UNP P19836
A	-12	HIS	-	EXPRESSION TAG	UNP P19836
A	-11	HIS	-	EXPRESSION TAG	UNP P19836
A	-10	HIS	-	EXPRESSION TAG	UNP P19836
A	-9	SER	-	EXPRESSION TAG	UNP P19836
A	-8	SER	-	EXPRESSION TAG	UNP P19836
A	-7	GLY	-	EXPRESSION TAG	UNP P19836
A	-6	LEU	-	EXPRESSION TAG	UNP P19836
A	-5	VAL	-	EXPRESSION TAG	UNP P19836
A	-4	PRO	-	EXPRESSION TAG	UNP P19836
A	-3	ARG	-	EXPRESSION TAG	UNP P19836
A	-2	GLY	-	EXPRESSION TAG	UNP P19836
A	-1	SER	-	EXPRESSION TAG	UNP P19836
A	0	HIS	-	EXPRESSION TAG	UNP P19836
D	-19	MET	-	EXPRESSION TAG	UNP P19836
D	-18	GLY	-	EXPRESSION TAG	UNP P19836
D	-17	SER	-	EXPRESSION TAG	UNP P19836
D	-16	SER	-	EXPRESSION TAG	UNP P19836
D	-15	HIS	-	EXPRESSION TAG	UNP P19836
D	-14	HIS	-	EXPRESSION TAG	UNP P19836
D	-13	HIS	-	EXPRESSION TAG	UNP P19836
D	-12	HIS	-	EXPRESSION TAG	UNP P19836
D	-11	HIS	-	EXPRESSION TAG	UNP P19836
D	-10	HIS	-	EXPRESSION TAG	UNP P19836
D	-9	SER	-	EXPRESSION TAG	UNP P19836
D	-8	SER	-	EXPRESSION TAG	UNP P19836
D	-7	GLY	-	EXPRESSION TAG	UNP P19836
D	-6	LEU	-	EXPRESSION TAG	UNP P19836
D	-5	VAL	-	EXPRESSION TAG	UNP P19836

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-4	PRO	-	EXPRESSION TAG	UNP P19836
D	-3	ARG	-	EXPRESSION TAG	UNP P19836
D	-2	GLY	-	EXPRESSION TAG	UNP P19836
D	-1	SER	-	EXPRESSION TAG	UNP P19836
D	0	HIS	-	EXPRESSION TAG	UNP P19836
C	-19	MET	-	EXPRESSION TAG	UNP P19836
C	-18	GLY	-	EXPRESSION TAG	UNP P19836
C	-17	SER	-	EXPRESSION TAG	UNP P19836
C	-16	SER	-	EXPRESSION TAG	UNP P19836
C	-15	HIS	-	EXPRESSION TAG	UNP P19836
C	-14	HIS	-	EXPRESSION TAG	UNP P19836
C	-13	HIS	-	EXPRESSION TAG	UNP P19836
C	-12	HIS	-	EXPRESSION TAG	UNP P19836
C	-11	HIS	-	EXPRESSION TAG	UNP P19836
C	-10	HIS	-	EXPRESSION TAG	UNP P19836
C	-9	SER	-	EXPRESSION TAG	UNP P19836
C	-8	SER	-	EXPRESSION TAG	UNP P19836
C	-7	GLY	-	EXPRESSION TAG	UNP P19836
C	-6	LEU	-	EXPRESSION TAG	UNP P19836
C	-5	VAL	-	EXPRESSION TAG	UNP P19836
C	-4	PRO	-	EXPRESSION TAG	UNP P19836
C	-3	ARG	-	EXPRESSION TAG	UNP P19836
C	-2	GLY	-	EXPRESSION TAG	UNP P19836
C	-1	SER	-	EXPRESSION TAG	UNP P19836
C	0	HIS	-	EXPRESSION TAG	UNP P19836
F	-19	MET	-	EXPRESSION TAG	UNP P19836
F	-18	GLY	-	EXPRESSION TAG	UNP P19836
F	-17	SER	-	EXPRESSION TAG	UNP P19836
F	-16	SER	-	EXPRESSION TAG	UNP P19836
F	-15	HIS	-	EXPRESSION TAG	UNP P19836
F	-14	HIS	-	EXPRESSION TAG	UNP P19836
F	-13	HIS	-	EXPRESSION TAG	UNP P19836
F	-12	HIS	-	EXPRESSION TAG	UNP P19836
F	-11	HIS	-	EXPRESSION TAG	UNP P19836
F	-10	HIS	-	EXPRESSION TAG	UNP P19836
F	-9	SER	-	EXPRESSION TAG	UNP P19836
F	-8	SER	-	EXPRESSION TAG	UNP P19836
F	-7	GLY	-	EXPRESSION TAG	UNP P19836
F	-6	LEU	-	EXPRESSION TAG	UNP P19836
F	-5	VAL	-	EXPRESSION TAG	UNP P19836
F	-4	PRO	-	EXPRESSION TAG	UNP P19836
F	-3	ARG	-	EXPRESSION TAG	UNP P19836

*Continued on next page...*

*Continued from previous page...*

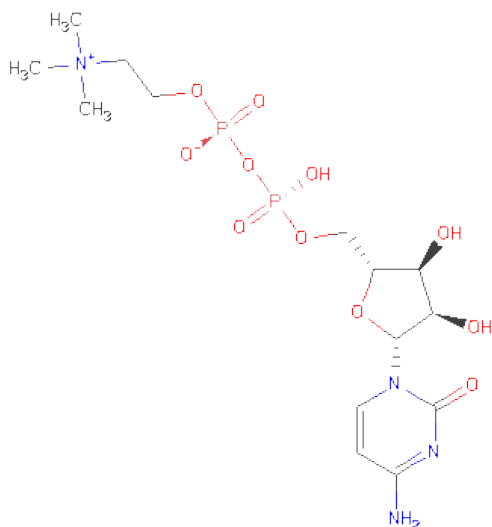
Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	EXPRESSION TAG	UNP P19836
F	-1	SER	-	EXPRESSION TAG	UNP P19836
F	0	HIS	-	EXPRESSION TAG	UNP P19836
E	-19	MET	-	EXPRESSION TAG	UNP P19836
E	-18	GLY	-	EXPRESSION TAG	UNP P19836
E	-17	SER	-	EXPRESSION TAG	UNP P19836
E	-16	SER	-	EXPRESSION TAG	UNP P19836
E	-15	HIS	-	EXPRESSION TAG	UNP P19836
E	-14	HIS	-	EXPRESSION TAG	UNP P19836
E	-13	HIS	-	EXPRESSION TAG	UNP P19836
E	-12	HIS	-	EXPRESSION TAG	UNP P19836
E	-11	HIS	-	EXPRESSION TAG	UNP P19836
E	-10	HIS	-	EXPRESSION TAG	UNP P19836
E	-9	SER	-	EXPRESSION TAG	UNP P19836
E	-8	SER	-	EXPRESSION TAG	UNP P19836
E	-7	GLY	-	EXPRESSION TAG	UNP P19836
E	-6	LEU	-	EXPRESSION TAG	UNP P19836
E	-5	VAL	-	EXPRESSION TAG	UNP P19836
E	-4	PRO	-	EXPRESSION TAG	UNP P19836
E	-3	ARG	-	EXPRESSION TAG	UNP P19836
E	-2	GLY	-	EXPRESSION TAG	UNP P19836
E	-1	SER	-	EXPRESSION TAG	UNP P19836
E	0	HIS	-	EXPRESSION TAG	UNP P19836
H	-19	MET	-	EXPRESSION TAG	UNP P19836
H	-18	GLY	-	EXPRESSION TAG	UNP P19836
H	-17	SER	-	EXPRESSION TAG	UNP P19836
H	-16	SER	-	EXPRESSION TAG	UNP P19836
H	-15	HIS	-	EXPRESSION TAG	UNP P19836
H	-14	HIS	-	EXPRESSION TAG	UNP P19836
H	-13	HIS	-	EXPRESSION TAG	UNP P19836
H	-12	HIS	-	EXPRESSION TAG	UNP P19836
H	-11	HIS	-	EXPRESSION TAG	UNP P19836
H	-10	HIS	-	EXPRESSION TAG	UNP P19836
H	-9	SER	-	EXPRESSION TAG	UNP P19836
H	-8	SER	-	EXPRESSION TAG	UNP P19836
H	-7	GLY	-	EXPRESSION TAG	UNP P19836
H	-6	LEU	-	EXPRESSION TAG	UNP P19836
H	-5	VAL	-	EXPRESSION TAG	UNP P19836
H	-4	PRO	-	EXPRESSION TAG	UNP P19836
H	-3	ARG	-	EXPRESSION TAG	UNP P19836
H	-2	GLY	-	EXPRESSION TAG	UNP P19836
H	-1	SER	-	EXPRESSION TAG	UNP P19836

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	0	HIS	-	EXPRESSION TAG	UNP P19836
G	-19	MET	-	EXPRESSION TAG	UNP P19836
G	-18	GLY	-	EXPRESSION TAG	UNP P19836
G	-17	SER	-	EXPRESSION TAG	UNP P19836
G	-16	SER	-	EXPRESSION TAG	UNP P19836
G	-15	HIS	-	EXPRESSION TAG	UNP P19836
G	-14	HIS	-	EXPRESSION TAG	UNP P19836
G	-13	HIS	-	EXPRESSION TAG	UNP P19836
G	-12	HIS	-	EXPRESSION TAG	UNP P19836
G	-11	HIS	-	EXPRESSION TAG	UNP P19836
G	-10	HIS	-	EXPRESSION TAG	UNP P19836
G	-9	SER	-	EXPRESSION TAG	UNP P19836
G	-8	SER	-	EXPRESSION TAG	UNP P19836
G	-7	GLY	-	EXPRESSION TAG	UNP P19836
G	-6	LEU	-	EXPRESSION TAG	UNP P19836
G	-5	VAL	-	EXPRESSION TAG	UNP P19836
G	-4	PRO	-	EXPRESSION TAG	UNP P19836
G	-3	ARG	-	EXPRESSION TAG	UNP P19836
G	-2	GLY	-	EXPRESSION TAG	UNP P19836
G	-1	SER	-	EXPRESSION TAG	UNP P19836
G	0	HIS	-	EXPRESSION TAG	UNP P19836

- Molecule 2 is [2-CYTIDYLATE-O'-PHOSPHONYLOXYL]-ETHYL-TRIMETHYL-AMMONIUM (three-letter code: CDC) (formula:  $C_{14}H_{26}N_4O_{11}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	A	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	D	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	C	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	F	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	E	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	H	1	Total	C	N	O	P	0	0
			31	14	4	11	2		
2	G	1	Total	C	N	O	P	0	0
			31	14	4	11	2		

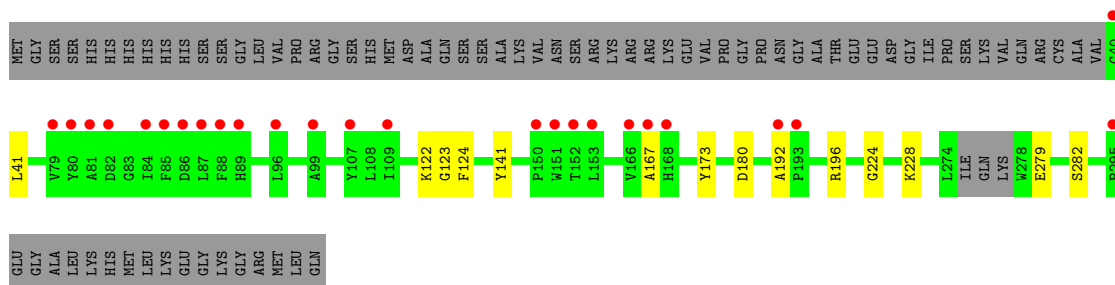


### 3 Residue-property plots

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.

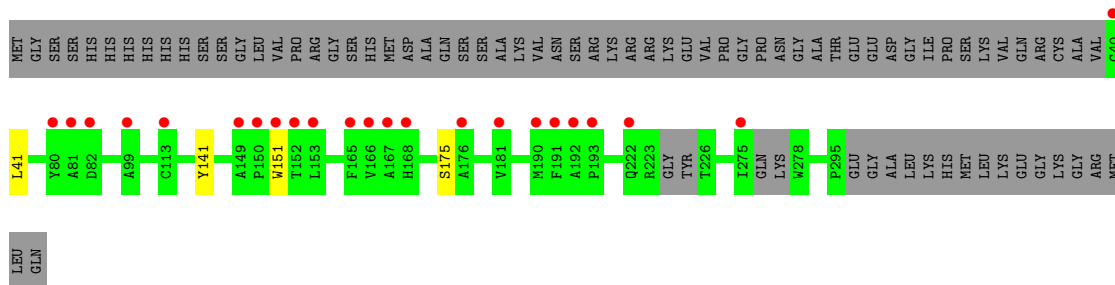
- Molecule 1: Choline-phosphate cytidyltransferase A

Chain B: 



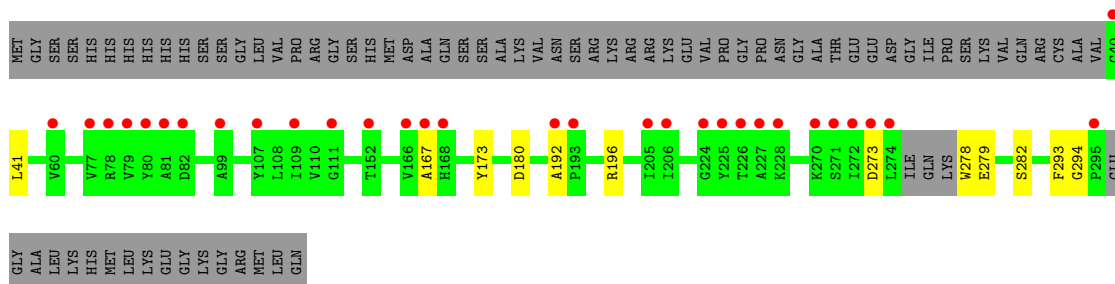
- Molecule 1: Choline-phosphate cytidyltransferase A

Chain A: 



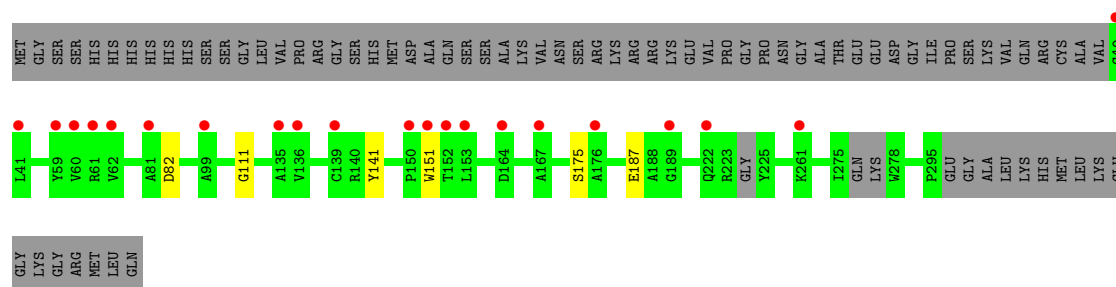
- Molecule 1: Choline-phosphate cytidyltransferase A

Chain D: 



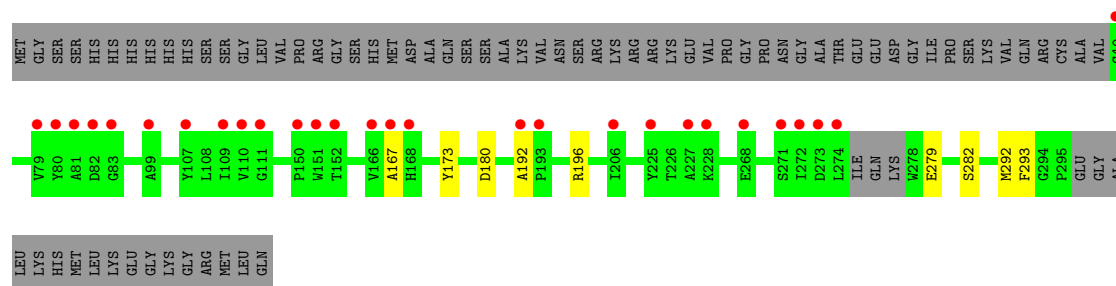
- Molecule 1: Choline-phosphate cytidyltransferase A

Chain C:



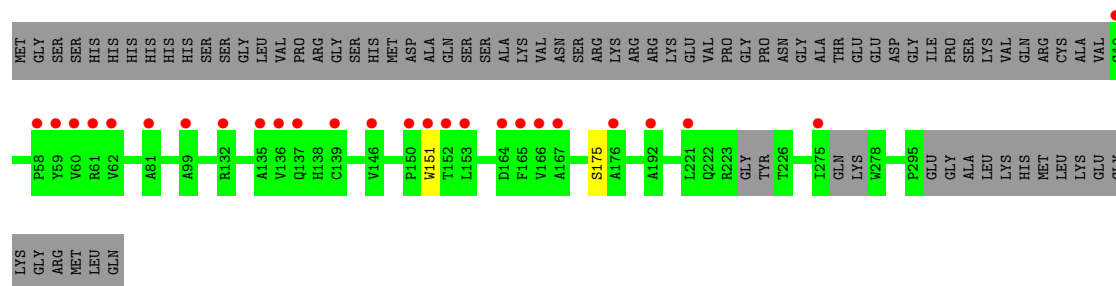
- Molecule 1: Choline-phosphate cytidylyltransferase A

Chain F:



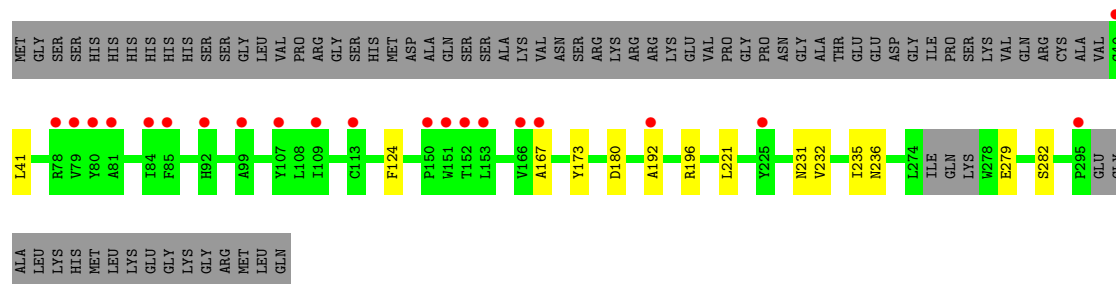
- Molecule 1: Choline-phosphate cytidylyltransferase A

Chain E:



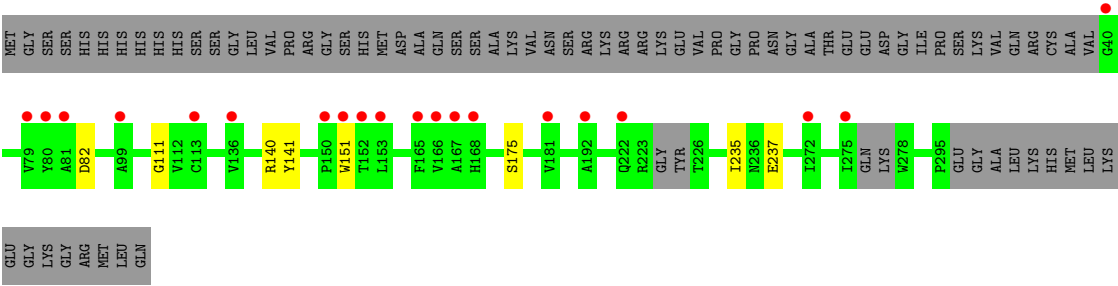
- Molecule 1: Choline-phosphate cytidylyltransferase A

Chain H:



- Molecule 1: Choline-phosphate cytidylyltransferase A

Chain G:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.78Å 107.78Å 575.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	107.80 – 8.00 107.78 – 8.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (107.80-8.00) 99.9 (107.78-8.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.91 (at 8.43Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.379 , 0.407 0.386 , 0.428	Depositor DCC
$R_{free}$ test set	329 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	388.8	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 454.4	EDS
Estimated twinning fraction	0.410 for h,-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 6916 reflections	Xtriage
$F_o, F_c$ correlation	0.58	EDS
Total number of atoms	10261	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.74% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CDC

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	0.37	0/1246	0.62	0/1733
1	B	0.31	0/1251	0.58	0/1741
1	C	0.34	0/1251	0.60	0/1740
1	D	0.32	0/1251	0.58	0/1741
1	E	0.33	0/1246	0.61	0/1733
1	F	0.31	0/1251	0.57	0/1741
1	G	0.41	1/1246 (0.1%)	0.64	0/1733
1	H	0.34	0/1251	0.59	1/1741 (0.1%)
All	All	0.34	1/9993 (0.0%)	0.60	1/13903 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
1	E	0	1
1	G	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	235	ILE	C-N	-6.49	1.19	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	221	LEU	CB-CA-C	-5.83	99.13	110.20

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	SER	Peptide
1	C	175	SER	Peptide
1	E	175	SER	Peptide
1	G	175	SER	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1249	0	546	3	0
1	B	1253	0	550	18	0
1	C	1254	0	548	3	1
1	D	1253	0	550	17	0
1	E	1249	0	546	1	0
1	F	1253	0	550	6	0
1	G	1249	0	545	4	1
1	H	1253	0	550	10	0
2	A	31	0	25	1	0
2	B	31	0	25	2	0
2	C	31	0	25	2	0
2	D	31	0	25	2	0
2	E	31	0	25	1	0
2	F	31	0	25	2	0
2	G	31	0	25	1	0
2	H	31	0	25	2	0
All	All	10261	0	4585	49	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (49) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:123:GLY:HA3	1:D:293:PHE:C	1.68	1.13
1:B:123:GLY:HA3	1:D:293:PHE:CA	1.92	0.98
1:B:123:GLY:HA3	1:D:293:PHE:HA	1.51	0.90
1:B:123:GLY:HA3	1:D:293:PHE:O	1.76	0.85
1:B:123:GLY:CA	1:D:293:PHE:HA	2.07	0.84
1:B:124:PHE:H	1:D:293:PHE:HA	1.42	0.83
1:B:124:PHE:N	1:D:293:PHE:HA	1.95	0.81
1:B:41:LEU:O	1:A:141:TYR:HA	1.92	0.68
1:H:232:VAL:O	1:H:236:ASN:N	2.31	0.60
1:B:224:GLY:O	1:B:228:LYS:N	2.34	0.59
1:B:123:GLY:CA	1:D:293:PHE:O	2.50	0.56
1:B:141:TYR:HA	1:A:41:LEU:O	2.07	0.54
1:B:122:LYS:O	1:D:294:GLY:HA2	2.08	0.52
1:B:123:GLY:C	1:D:293:PHE:HA	2.30	0.52
1:F:279:GLU:O	1:F:282:SER:CB	2.58	0.52
1:B:279:GLU:O	1:B:282:SER:CB	2.58	0.51
1:D:279:GLU:O	1:D:282:SER:CB	2.58	0.51
1:H:279:GLU:O	1:H:282:SER:CB	2.58	0.51
1:D:196:ARG:HA	2:D:401:CDC:O2	2.12	0.50
1:F:196:ARG:HA	2:F:401:CDC:O2	2.12	0.50
1:B:196:ARG:HA	2:B:401:CDC:O2	2.12	0.50
1:H:196:ARG:HA	2:H:401:CDC:O2	2.12	0.50
1:D:41:LEU:O	1:C:141:TYR:HA	2.12	0.49
1:G:151:TRP:O	2:G:401:CDC:H171	2.13	0.49
1:C:151:TRP:O	2:C:401:CDC:H171	2.13	0.49
1:E:151:TRP:O	2:E:401:CDC:H171	2.13	0.49
1:H:231:ASN:O	1:H:235:ILE:N	2.38	0.48
1:F:292:MET:O	1:H:124:PHE:N	2.40	0.48
1:A:151:TRP:O	2:A:401:CDC:H171	2.13	0.48
1:B:167:ALA:HA	1:B:192:ALA:O	2.14	0.48
1:D:167:ALA:HA	1:D:192:ALA:O	2.14	0.48
1:H:167:ALA:HA	1:H:192:ALA:O	2.14	0.47
1:F:167:ALA:HA	1:F:192:ALA:O	2.14	0.47
2:H:401:CDC:H162	2:H:401:CDC:H141	1.65	0.47
1:H:41:LEU:O	1:G:140:ARG:O	2.33	0.47
2:B:401:CDC:H162	2:B:401:CDC:H141	1.65	0.46
1:F:173:TYR:O	1:F:180:ASP:HA	2.20	0.42
1:F:293:PHE:HA	1:H:124:PHE:N	2.34	0.42
2:C:401:CDC:H162	2:C:401:CDC:H141	1.70	0.42
2:D:401:CDC:H141	2:D:401:CDC:H162	1.65	0.42
1:H:173:TYR:O	1:H:180:ASP:HA	2.20	0.42
1:H:41:LEU:O	1:G:141:TYR:HA	2.19	0.41
2:F:401:CDC:H162	2:F:401:CDC:H141	1.65	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:173:TYR:O	1:D:180:ASP:HA	2.20	0.41
1:B:173:TYR:O	1:B:180:ASP:HA	2.20	0.41
1:B:122:LYS:O	1:D:294:GLY:CA	2.68	0.41
1:D:273:ASP:O	1:D:278:TRP:N	2.54	0.41
1:G:82:ASP:HA	1:G:111:GLY:O	2.22	0.40
1:C:82:ASP:HA	1:C:111:GLY:O	2.22	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	Distance(Å)	Clash(Å)
1:C:187:GLU:O	1:G:237:GLU:CB[3_554]	1.68	0.52

## 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	246/332 (74%)	240 (98%)	6 (2%)	0	100	100
1	B	249/332 (75%)	236 (95%)	13 (5%)	0	100	100
1	C	247/332 (74%)	241 (98%)	6 (2%)	0	100	100
1	D	249/332 (75%)	235 (94%)	14 (6%)	0	100	100
1	E	246/332 (74%)	240 (98%)	6 (2%)	0	100	100
1	F	249/332 (75%)	235 (94%)	14 (6%)	0	100	100
1	G	246/332 (74%)	240 (98%)	6 (2%)	0	100	100
1	H	249/332 (75%)	235 (94%)	14 (6%)	0	100	100
All	All	1981/2656 (75%)	1902 (96%)	79 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

There are no protein residues with a non-rotameric sidechain to report in this entry.



### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

8 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	CDC	A	401	-	32,32,32	1.49	2 (6%)	46,49,49	1.11	3 (6%)
2	CDC	B	401	-	32,32,32	1.46	2 (6%)	46,49,49	1.18	4 (8%)
2	CDC	C	401	-	32,32,32	1.46	2 (6%)	46,49,49	1.11	3 (6%)
2	CDC	D	401	-	32,32,32	1.46	2 (6%)	46,49,49	1.20	4 (8%)
2	CDC	E	401	-	32,32,32	1.48	2 (6%)	46,49,49	1.12	3 (6%)
2	CDC	F	401	-	32,32,32	1.46	2 (6%)	46,49,49	1.19	4 (8%)
2	CDC	G	401	-	32,32,32	1.47	2 (6%)	46,49,49	1.11	3 (6%)
2	CDC	H	401	-	32,32,32	1.46	2 (6%)	46,49,49	1.17	4 (8%)

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	CDC	A	401	-	-	0/22/40/40	0/2/2/2
2	CDC	B	401	-	-	0/22/40/40	0/2/2/2
2	CDC	C	401	-	-	0/22/40/40	0/2/2/2
2	CDC	D	401	-	-	0/22/40/40	0/2/2/2
2	CDC	E	401	-	-	0/22/40/40	0/2/2/2
2	CDC	F	401	-	-	0/22/40/40	0/2/2/2
2	CDC	G	401	-	-	0/22/40/40	0/2/2/2
2	CDC	H	401	-	-	0/22/40/40	0/2/2/2

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	401	CDC	O2-C2	6.45	1.37	1.23
2	B	401	CDC	O2-C2	6.44	1.37	1.23
2	H	401	CDC	O2-C2	6.44	1.37	1.23
2	D	401	CDC	O2-C2	6.43	1.37	1.23
2	A	401	CDC	O2-C2	6.41	1.37	1.23
2	G	401	CDC	O2-C2	6.36	1.36	1.23
2	E	401	CDC	O2-C2	6.35	1.36	1.23
2	C	401	CDC	O2-C2	6.35	1.36	1.23
2	A	401	CDC	C2-N1	-3.67	1.34	1.38
2	E	401	CDC	C2-N1	-3.63	1.34	1.38
2	G	401	CDC	C2-N1	-3.59	1.34	1.38
2	C	401	CDC	C2-N1	-3.48	1.34	1.38
2	D	401	CDC	C2-N1	-3.29	1.34	1.38
2	B	401	CDC	C2-N1	-3.25	1.34	1.38
2	F	401	CDC	C2-N1	-3.22	1.34	1.38
2	H	401	CDC	C2-N1	-3.22	1.34	1.38

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	401	CDC	C6-C5-C4	3.60	118.97	117.47
2	F	401	CDC	C2-N3-C4	3.43	120.54	115.57
2	B	401	CDC	C2-N3-C4	3.41	120.50	115.57
2	D	401	CDC	C2-N3-C4	3.41	120.50	115.57
2	H	401	CDC	C2-N3-C4	3.41	120.50	115.57
2	B	401	CDC	C6-C5-C4	3.37	118.87	117.47
2	F	401	CDC	C6-C5-C4	3.36	118.86	117.47
2	H	401	CDC	C6-C5-C4	3.23	118.81	117.47
2	G	401	CDC	C2-N3-C4	3.15	120.14	115.57
2	C	401	CDC	C2-N3-C4	3.15	120.13	115.57
2	E	401	CDC	C2-N3-C4	3.13	120.11	115.57

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	401	CDC	C6-C5-C4	3.12	118.77	117.47
2	A	401	CDC	C2-N3-C4	3.12	120.08	115.57
2	G	401	CDC	C6-C5-C4	3.00	118.72	117.47
2	C	401	CDC	C6-C5-C4	2.98	118.71	117.47
2	A	401	CDC	C6-C5-C4	2.96	118.70	117.47
2	B	401	CDC	PB-O3A-PA	-2.88	120.59	132.95
2	F	401	CDC	PB-O3A-PA	-2.88	120.59	132.95
2	D	401	CDC	PB-O3A-PA	-2.88	120.59	132.95
2	H	401	CDC	PB-O3A-PA	-2.88	120.59	132.95
2	A	401	CDC	PB-O3A-PA	-2.82	120.84	132.95
2	C	401	CDC	PB-O3A-PA	-2.82	120.85	132.95
2	G	401	CDC	PB-O3A-PA	-2.82	120.85	132.95
2	E	401	CDC	PB-O3A-PA	-2.79	120.97	132.95
2	H	401	CDC	C3'-C2'-C1'	2.32	104.54	100.91
2	D	401	CDC	C3'-C2'-C1'	2.31	104.52	100.91
2	F	401	CDC	C3'-C2'-C1'	2.30	104.51	100.91
2	B	401	CDC	C3'-C2'-C1'	2.28	104.47	100.91

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	252/332 (75%)	0.51	23 (9%) 9 18	2, 30, 72, 94	0
1	B	253/332 (76%)	0.77	25 (9%) 8 16	2, 30, 75, 115	1 (0%)
1	C	253/332 (76%)	0.79	21 (8%) 11 20	2, 30, 72, 94	0
1	D	253/332 (76%)	0.77	31 (12%) 5 13	2, 30, 75, 115	1 (0%)
1	E	252/332 (75%)	0.81	26 (10%) 7 16	2, 30, 72, 94	0
1	F	253/332 (76%)	0.81	28 (11%) 6 14	2, 30, 75, 115	1 (0%)
1	G	252/332 (75%)	0.47	20 (7%) 13 21	2, 30, 72, 94	0
1	H	253/332 (76%)	0.72	21 (8%) 11 20	2, 30, 75, 115	1 (0%)
All	All	2021/2656 (76%)	0.71	195 (9%) 8 17	2, 30, 73, 115	4 (0%)

All (195) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	167	ALA	6.4
1	F	167	ALA	6.0
1	A	151	TRP	5.6
1	G	167	ALA	5.6
1	H	167	ALA	5.6
1	D	81	ALA	5.4
1	F	81	ALA	5.2
1	B	152	THR	5.2
1	G	152	THR	5.2
1	B	151	TRP	5.0
1	G	151	TRP	4.9
1	D	167	ALA	4.8
1	B	167	ALA	4.6
1	A	192	ALA	4.5
1	F	80	TYR	4.4
1	C	151	TRP	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	274	LEU	4.4
1	H	152	THR	4.4
1	A	166	VAL	4.3
1	E	59	TYR	4.3
1	E	151	TRP	4.3
1	E	61	ARG	4.2
1	D	274	LEU	4.1
1	D	80	TYR	4.1
1	E	167	ALA	4.1
1	C	167	ALA	4.0
1	E	150	PRO	3.9
1	B	192	ALA	3.9
1	E	152	THR	3.9
1	A	150	PRO	3.9
1	G	166	VAL	3.8
1	B	150	PRO	3.8
1	C	152	THR	3.8
1	D	192	ALA	3.8
1	C	59	TYR	3.7
1	G	153	LEU	3.7
1	A	152	THR	3.7
1	H	151	TRP	3.6
1	E	60	VAL	3.6
1	B	81	ALA	3.6
1	C	135	ALA	3.5
1	H	80	TYR	3.5
1	B	153	LEU	3.5
1	G	192	ALA	3.5
1	C	150	PRO	3.5
1	F	79	VAL	3.4
1	H	295	PRO	3.4
1	D	79	VAL	3.4
1	F	272	ILE	3.4
1	E	136	VAL	3.3
1	F	192	ALA	3.3
1	E	164	ASP	3.3
1	H	84	ILE	3.3
1	C	153	LEU	3.3
1	H	81	ALA	3.3
1	C	61	ARG	3.3
1	E	192	ALA	3.3
1	F	99	ALA	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	192	ALA	3.2
1	B	80	TYR	3.2
1	E	81	ALA	3.2
1	H	153	LEU	3.2
1	A	81	ALA	3.2
1	A	153	LEU	3.1
1	D	40	GLY	3.1
1	F	152	THR	3.1
1	B	99	ALA	3.0
1	F	166	VAL	3.0
1	C	81	ALA	3.0
1	A	168	HIS	3.0
1	A	176	ALA	2.9
1	F	271	SER	2.9
1	H	150	PRO	2.9
1	G	80	TYR	2.9
1	G	165	PHE	2.9
1	F	40	GLY	2.9
1	B	86	ASP	2.9
1	B	166	VAL	2.9
1	H	85	PHE	2.9
1	E	58	PRO	2.9
1	H	166	VAL	2.8
1	A	275	ILE	2.8
1	G	150	PRO	2.8
1	F	151	TRP	2.8
1	E	153	LEU	2.8
1	H	40	GLY	2.8
1	F	273	ASP	2.8
1	E	135	ALA	2.8
1	C	40	GLY	2.8
1	E	176	ALA	2.8
1	D	225	TYR	2.8
1	H	79	VAL	2.8
1	D	227	ALA	2.8
1	E	139	CYS	2.7
1	C	164	ASP	2.7
1	B	84	ILE	2.7
1	C	60	VAL	2.7
1	F	111	GLY	2.7
1	B	79	VAL	2.7
1	E	99	ALA	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	107	TYR	2.6
1	F	268	GLU	2.6
1	A	99	ALA	2.6
1	F	227	ALA	2.6
1	G	40	GLY	2.6
1	D	272	ILE	2.6
1	D	107	TYR	2.6
1	A	165	PHE	2.5
1	F	193	PRO	2.5
1	H	99	ALA	2.5
1	B	85	PHE	2.5
1	B	89	HIS	2.5
1	A	193	PRO	2.5
1	D	166	VAL	2.5
1	C	41	LEU	2.5
1	D	273	ASP	2.5
1	G	168	HIS	2.5
1	F	109	ILE	2.5
1	D	224	GLY	2.4
1	D	271	SER	2.4
1	H	225	TYR	2.4
1	H	113	CYS	2.4
1	B	82	ASP	2.4
1	G	99	ALA	2.4
1	C	99	ALA	2.4
1	E	221	LEU	2.4
1	C	222	GLN	2.4
1	B	87	LEU	2.4
1	D	206	ILE	2.4
1	D	82	ASP	2.4
1	F	83	GLY	2.3
1	B	88	PHE	2.3
1	H	109	ILE	2.3
1	F	110	VAL	2.3
1	F	228	LYS	2.3
1	D	168	HIS	2.3
1	C	136	VAL	2.3
1	D	111	GLY	2.3
1	G	181	VAL	2.3
1	B	295	PRO	2.3
1	A	80	TYR	2.3
1	A	222	GLN	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	62	VAL	2.3
1	E	132	ARG	2.3
1	C	62	VAL	2.3
1	C	139	CYS	2.3
1	B	168	HIS	2.2
1	G	113	CYS	2.2
1	D	295	PRO	2.2
1	G	272	ILE	2.2
1	G	81	ALA	2.2
1	D	99	ALA	2.2
1	F	82	ASP	2.2
1	B	109	ILE	2.2
1	C	176	ALA	2.2
1	B	96	LEU	2.2
1	E	166	VAL	2.2
1	D	109	ILE	2.2
1	F	150	PRO	2.2
1	E	137	GLN	2.2
1	G	275	ILE	2.2
1	D	228	LYS	2.1
1	E	165	PHE	2.1
1	H	107	TYR	2.1
1	G	136	VAL	2.1
1	A	181	VAL	2.1
1	D	270	LYS	2.1
1	A	113	CYS	2.1
1	D	78	ARG	2.1
1	E	40	GLY	2.1
1	G	222	GLN	2.1
1	F	225	TYR	2.1
1	H	92	HIS	2.1
1	D	226	THR	2.1
1	E	146	VAL	2.1
1	H	78	ARG	2.1
1	D	205	ILE	2.1
1	D	60	VAL	2.1
1	A	191	PHE	2.1
1	A	82	ASP	2.1
1	C	261	LYS	2.1
1	E	275	ILE	2.1
1	B	193	PRO	2.1
1	A	190	MET	2.1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	206	ILE	2.1
1	G	79	VAL	2.1
1	D	152	THR	2.0
1	B	40	GLY	2.0
1	C	189	GLY	2.0
1	D	77	VAL	2.0
1	D	193	PRO	2.0
1	F	168	HIS	2.0
1	A	40	GLY	2.0
1	B	107	TYR	2.0
1	A	149	ALA	2.0

## 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

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<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	CDC	H	401	31/31	0.64	0.39	12,32,49,64	0
2	CDC	D	401	31/31	0.45	0.18	12,32,49,64	0
2	CDC	B	401	31/31	0.56	0.01	12,32,49,64	0
2	CDC	F	401	31/31	0.46	-0.03	12,32,49,64	0
2	CDC	G	401	31/31	0.37	-0.12	9,28,48,66	0
2	CDC	C	401	31/31	0.44	-0.15	9,28,48,66	0
2	CDC	E	401	31/31	0.42	-0.27	9,28,48,66	0
2	CDC	A	401	31/31	0.35	-0.28	9,28,48,66	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.