



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:43 pm BST

PDB ID : 1ECQ  
Title : E. COLI GLUCARATE DEHYDRATASE BOUND TO 4-DEOXYGLUCARATE  
Authors : Gulick, A.M.; Hubbard, B.K.; Gerlt, J.A.; Rayment, I.  
Deposited on : 2000-01-25  
Resolution : 2.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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 : 2.11  
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) : 2.11



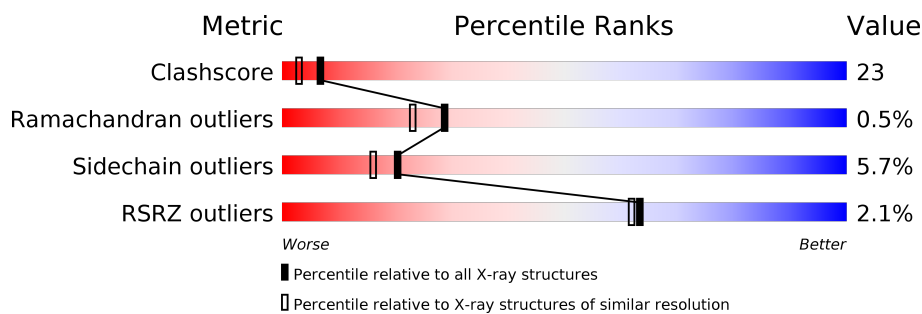
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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  $\leq 5\%$ . 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	446	
1	B	446	
1	C	446	
1	D	446	

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	DXG	D	502	-	-	X	-
4	IPA	C	604	-	-	X	-



## 2 Entry composition [i](#)

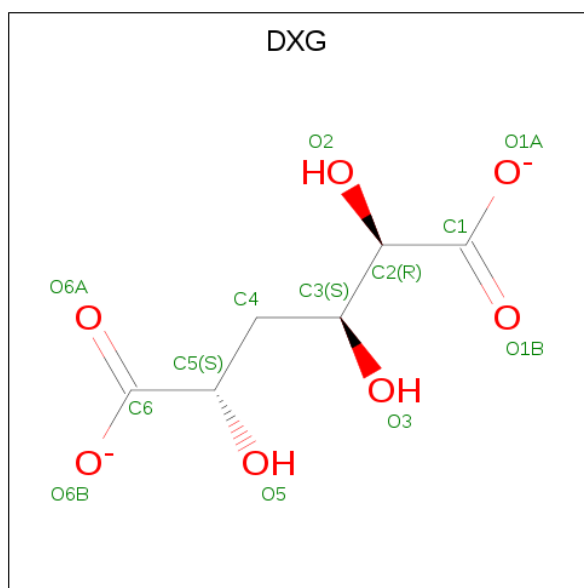
There are 5 unique types of molecules in this entry. The entry contains 14643 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 GLUCARATE DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	444	Total	C	N	O	S	0	0	0
			3436	2171	602	642	21			
1	B	442	Total	C	N	O	S	0	0	0
			3408	2153	598	636	21			
1	C	443	Total	C	N	O	S	0	1	0
			3421	2161	599	640	21			
1	D	442	Total	C	N	O	S	0	0	0
			3405	2153	595	636	21			

- Molecule 2 is 4-DEOXYGLUCARATE (three-letter code: DXG) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		

*Continued on next page...*



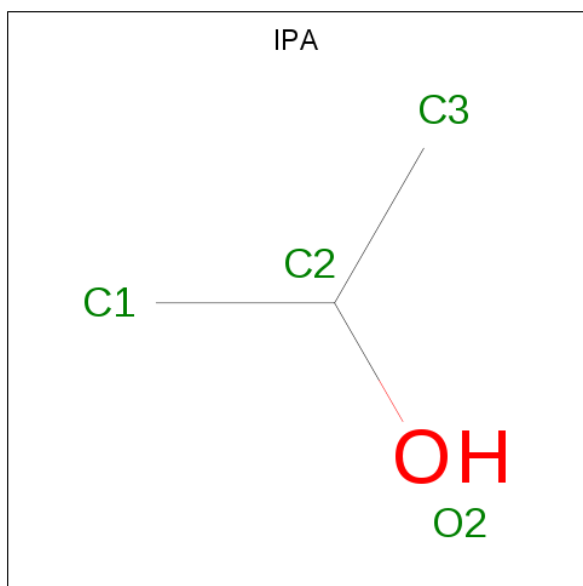
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			13	6	7		
2	D	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	D	1	Total	Mg	0	0
			1	1		
3	C	1	Total	Mg	0	0
			1	1		

- Molecule 4 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C<sub>3</sub>H<sub>8</sub>O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	3	1		
4	B	1	Total	C	O	0	0
			4	3	1		
4	C	1	Total	C	O	0	0
			4	3	1		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			4	3	1		

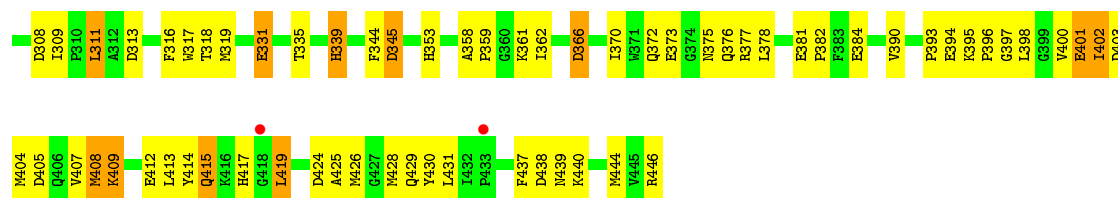
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	229	Total	O	0	0
			229	229		
5	B	168	Total	O	0	0
			168	168		
5	C	276	Total	O	0	0
			276	276		
5	D	228	Total	O	0	0
			228	228		

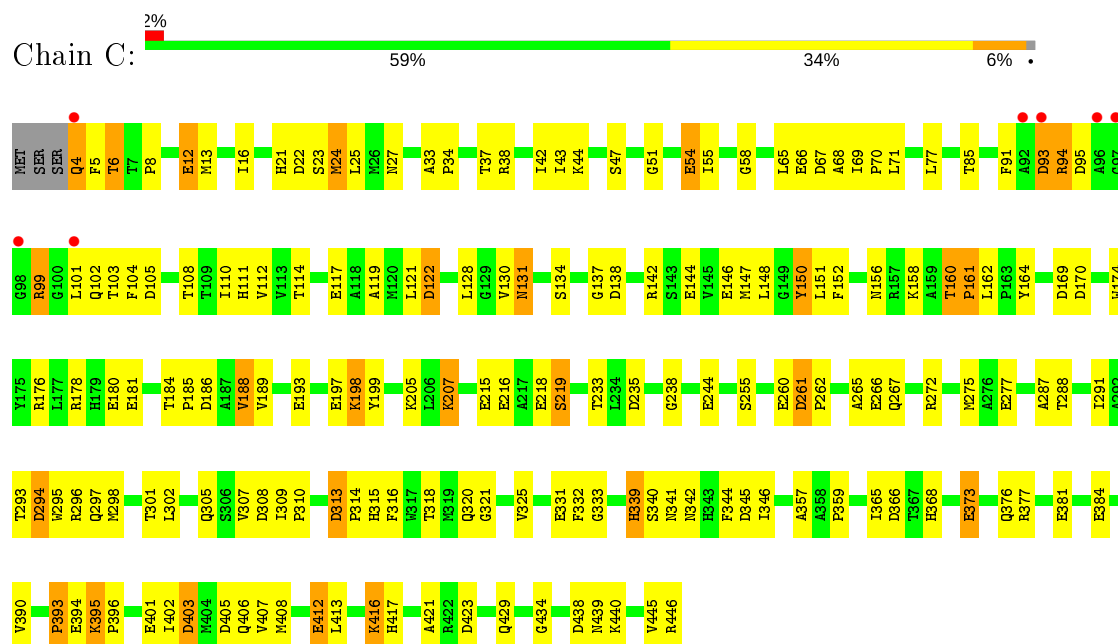




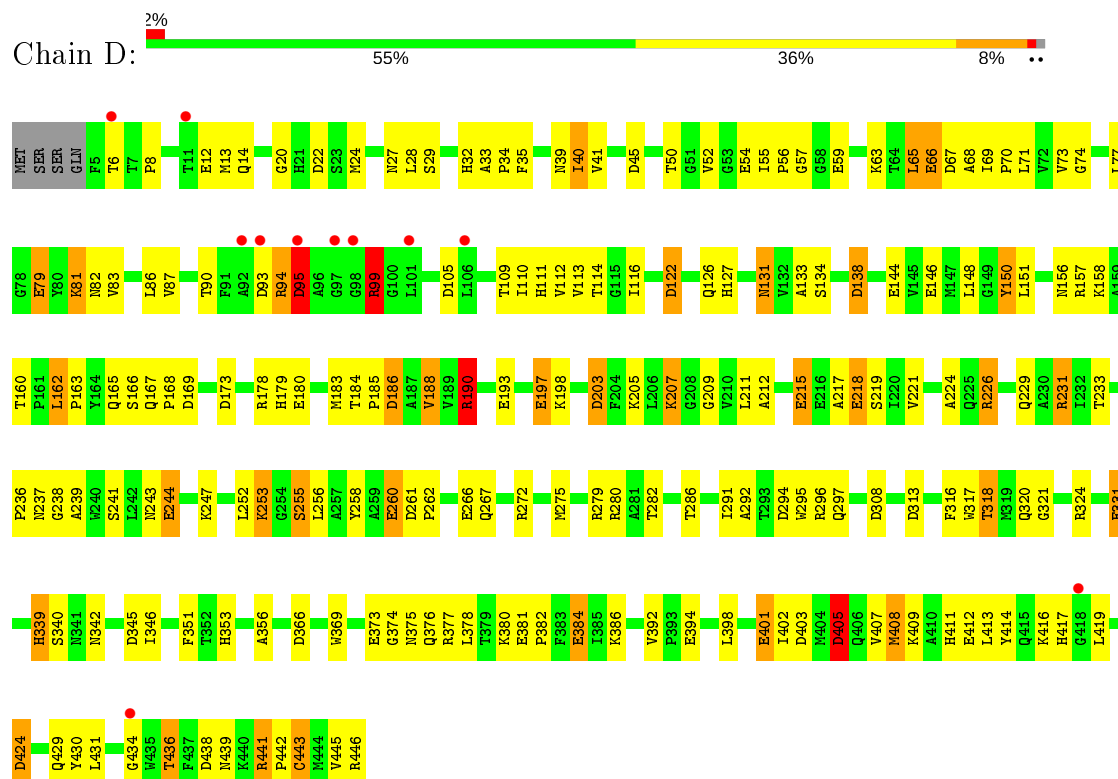




● Molecule 1: GLUCARATE DEHYDRATASE



● Molecule 1: GLUCARATE DEHYDRATASE





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.11Å 84.60Å 98.78Å 103.46° 93.97° 113.03°	Depositor
Resolution (Å)	30.00 – 2.00 20.03 – 2.01	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.00) 93.9 (20.03-2.01)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.90 (at 2.01Å)	Xtriage
Refinement program	TNT 5E	Depositor
R, $R_{free}$	0.209 , 0.305 0.197 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 89.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14643	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.70% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: DXG, MG, IPA

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	1.03	21/3515 (0.6%)	1.28	39/4763 (0.8%)
1	B	1.00	24/3486 (0.7%)	1.26	39/4725 (0.8%)
1	C	1.04	19/3500 (0.5%)	1.25	34/4746 (0.7%)
1	D	1.04	21/3484 (0.6%)	1.30	45/4724 (1.0%)
All	All	1.03	85/13985 (0.6%)	1.27	157/18958 (0.8%)

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	D	1	0

The worst 5 of 85 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	190	ARG	NE-CZ	12.02	1.48	1.33
1	C	412	GLU	CD-OE2	7.37	1.33	1.25
1	C	394	GLU	CD-OE2	7.32	1.33	1.25
1	B	331	GLU	CD-OE2	7.18	1.33	1.25
1	B	193	GLU	CD-OE2	7.15	1.33	1.25

The worst 5 of 157 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	294	ASP	CB-CG-OD2	-9.85	109.44	118.30
1	B	67	ASP	CB-CG-OD2	-9.58	109.68	118.30
1	B	366	ASP	CB-CG-OD2	-9.20	110.02	118.30
1	B	366	ASP	CB-CG-OD1	9.07	126.46	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	203	ASP	CB-CG-OD2	-8.85	110.33	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	351	PHE	CA

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

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	3436	0	3368	162	0
1	B	3408	0	3332	191	0
1	C	3421	0	3333	138	0
1	D	3405	0	3318	165	0
2	A	13	0	8	5	0
2	B	13	0	8	3	0
2	C	13	0	8	4	0
2	D	13	0	8	6	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	4	0	8	0	0
4	B	4	0	8	0	0
4	C	4	0	8	5	0
4	D	4	0	8	1	0
5	A	229	0	0	10	0
5	B	168	0	0	11	0
5	C	276	0	0	12	0
5	D	228	0	0	11	0
All	All	14643	0	13415	632	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 23.



The worst 5 of 632 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:C:339:HIS:NE2	2:C:501:DXG:H51	1.66	1.11
1:B:287:ALA:HB1	1:B:311:LEU:HD21	1.29	1.10
1:A:339:HIS:NE2	2:A:499:DXG:H51	1.70	1.07
1:B:24:MET:HG2	1:B:26:MET:HE3	1.40	1.03
1:B:339:HIS:NE2	2:B:500:DXG:H51	1.76	0.99

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	442/446 (99%)	414 (94%)	28 (6%)	0	100	100
1	B	440/446 (99%)	405 (92%)	30 (7%)	5 (1%)	14	8
1	C	442/446 (99%)	412 (93%)	28 (6%)	2 (0%)	29	23
1	D	440/446 (99%)	409 (93%)	30 (7%)	1 (0%)	47	44
All	All	1764/1784 (99%)	1640 (93%)	116 (7%)	8 (0%)	29	23

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	107	ARG
1	B	102	GLN
1	B	80	TYR
1	C	24	MET
1	D	443	CYS



### 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	359/362 (99%)	341 (95%)	18 (5%)	24	20
1	B	354/362 (98%)	331 (94%)	23 (6%)	17	12
1	C	355/362 (98%)	338 (95%)	17 (5%)	25	22
1	D	353/362 (98%)	330 (94%)	23 (6%)	17	12
All	All	1421/1448 (98%)	1340 (94%)	81 (6%)	20	16

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	409	LYS
1	C	131	ASN
1	D	255	SER
1	B	415	GLN
1	C	4	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	111	HIS
1	C	225	GLN
1	D	339	HIS
1	C	140	GLN
1	C	243	ASN

### 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](#)

Of 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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	$\# Z  > 2$
4	IPA	A	602	-	3,3,3	0.68	0	3,3,3	0.24	0
2	DXG	B	500	3	6,12,12	0.41	0	7,16,16	1.80	3 (42%)
4	IPA	B	603	-	3,3,3	0.22	0	3,3,3	0.57	0
2	DXG	A	499	3	6,12,12	1.40	1 (16%)	7,16,16	1.95	1 (14%)
4	IPA	C	604	-	3,3,3	0.55	0	3,3,3	0.25	0
2	DXG	D	502	3	6,12,12	1.22	1 (16%)	7,16,16	1.40	2 (28%)
2	DXG	C	501	3	6,12,12	0.72	0	7,16,16	1.98	2 (28%)
4	IPA	D	601	-	3,3,3	0.70	0	3,3,3	0.30	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	DXG	B	500	3	-	3/8/16/16	-
2	DXG	D	502	3	-	4/8/16/16	-
2	DXG	A	499	3	-	3/8/16/16	-
2	DXG	C	501	3	-	4/8/16/16	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	499	DXG	C3-C2	-3.16	1.50	1.53
2	D	502	DXG	C3-C2	-2.49	1.51	1.53



The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	499	DXG	O5-C5-C4	4.55	118.20	108.78
2	C	501	DXG	O5-C5-C4	3.55	116.14	108.78
2	C	501	DXG	O3-C3-C2	2.93	113.19	109.26
2	B	500	DXG	O5-C5-C4	2.81	114.60	108.78
2	B	500	DXG	O3-C3-C4	2.46	114.05	109.11

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	500	DXG	C3-C4-C5-O5
2	B	500	DXG	C3-C4-C5-C6
2	A	499	DXG	C3-C4-C5-O5
2	D	502	DXG	C3-C4-C5-O5
2	D	502	DXG	C3-C4-C5-C6

There are no ring outliers.

6 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	500	DXG	3	0
2	A	499	DXG	5	0
4	C	604	IPA	5	0
2	D	502	DXG	6	0
2	C	501	DXG	4	0
4	D	601	IPA	1	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	444/446 (99%)	-0.16	7 (1%) 72 70	14, 33, 66, 100	1 (0%)
1	B	442/446 (99%)	0.03	12 (2%) 54 53	17, 43, 72, 100	0
1	C	443/446 (99%)	-0.26	7 (1%) 72 70	13, 31, 59, 100	0
1	D	442/446 (99%)	-0.09	11 (2%) 57 56	18, 35, 70, 99	0
All	All	1771/1784 (99%)	-0.12	37 (2%) 63 62	13, 35, 68, 100	1 (0%)

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	97	GLY	4.9
1	A	101	LEU	4.4
1	C	93	ASP	4.1
1	C	101	LEU	3.9
1	C	97	GLY	3.9

### 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. 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	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	IPA	B	603	4/4	0.90	0.18	22,26,26,96	0
2	DXG	D	502	13/13	0.93	0.11	30,42,97,100	0
2	DXG	B	500	13/13	0.94	0.10	30,46,67,68	0
4	IPA	D	601	4/4	0.94	0.16	23,37,44,56	0
2	DXG	C	501	13/13	0.95	0.10	25,35,70,100	0
4	IPA	A	602	4/4	0.95	0.16	17,33,41,54	0
2	DXG	A	499	13/13	0.96	0.09	19,38,62,81	0
4	IPA	C	604	4/4	0.97	0.19	19,46,67,100	0
3	MG	A	498	1/1	0.97	0.05	28,28,28,28	0
3	MG	B	498	1/1	0.99	0.05	38,38,38,38	0
3	MG	D	498	1/1	0.99	0.04	34,34,34,34	0
3	MG	C	498	1/1	0.99	0.04	29,29,29,29	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.