



# wwPDB X-ray Structure Validation Summary Report

Dec 5, 2023 – 02:51 am GMT

PDB ID : 2VJB  
Title : Torpedo Californica Acetylcholinesterase In Complex With A Non Hydrolysable Substrate Analogue, 4-Oxo-N,N,N- Trimethylpentanaminium - Orthorhombic space group - Dataset D at 100K  
Authors : Colletier, J.P.; Bourgeois, D.; Fournier, D.; Silman, I.; Sussman, J.L.; Weik, M.  
Deposited on : 2007-12-09  
Resolution : 2.39 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

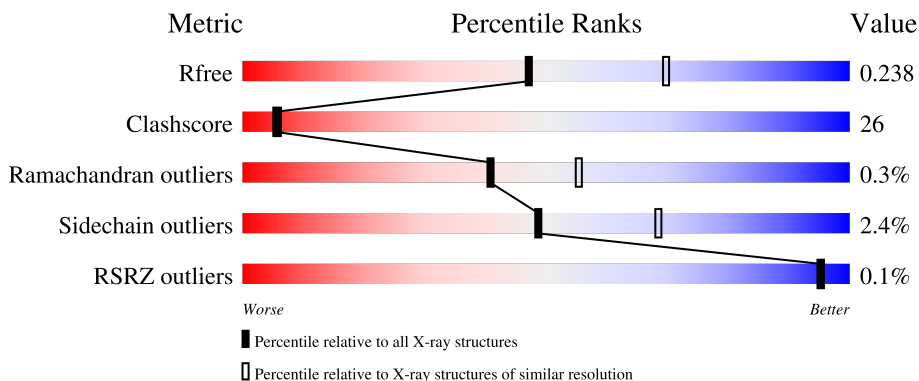
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.39 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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  $\geq 3$ , 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	537	
1	B	537	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	CCD	A	1540	X	-	-	-
4	CCD	B	1540	X	-	-	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 10185 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 ACETYLCHOLINESTERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	528	Total	C	N	O	S	0	0	0
			4210	2702	713	773	22			
1	B	532	Total	C	N	O	S	0	0	0
			4244	2721	720	781	22			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

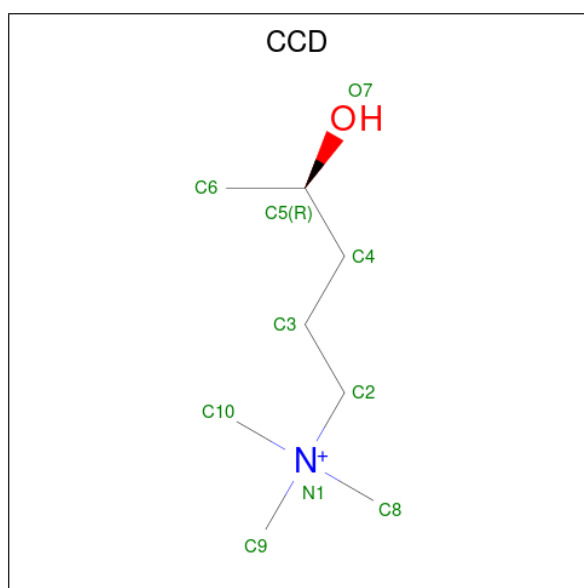


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	Cl	0	0
			3	3		
3	B	1	Total	Cl	0	0
			1	1		

- Molecule 4 is (4R)-4-HYDROXY-N,N,N-TRIMETHYLPENTAN-1-AMINIUM (three-letter code: CCD) (formula: C<sub>8</sub>H<sub>20</sub>NO).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			10	8	1	1		
4	A	1	Total	C	N	O	0	0
			10	8	1	1		
4	B	1	Total	C	N	O	0	0
			10	8	1	1		
4	B	1	Total	C	N	O	0	0
			10	8	1	1		

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 10 6 4	0	0
5	A	1	Total C O 10 6 4	0	0
5	B	1	Total C O 10 6 4	0	0

- Molecule 6 is water.

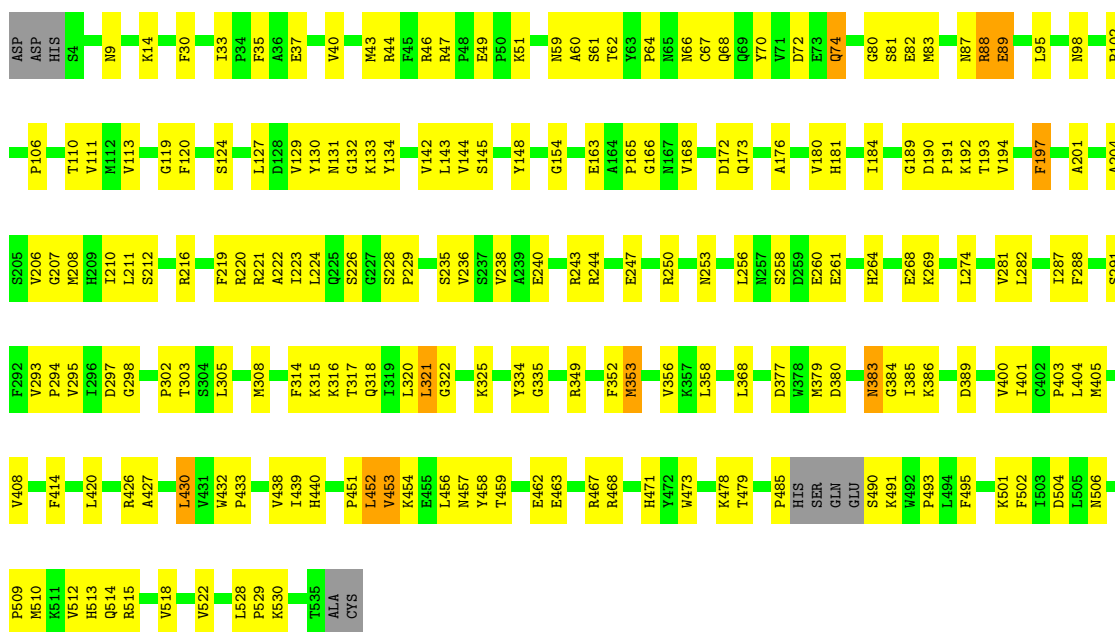
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	857	Total O 857 857	0	0
6	B	744	Total O 744 744	0	0

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

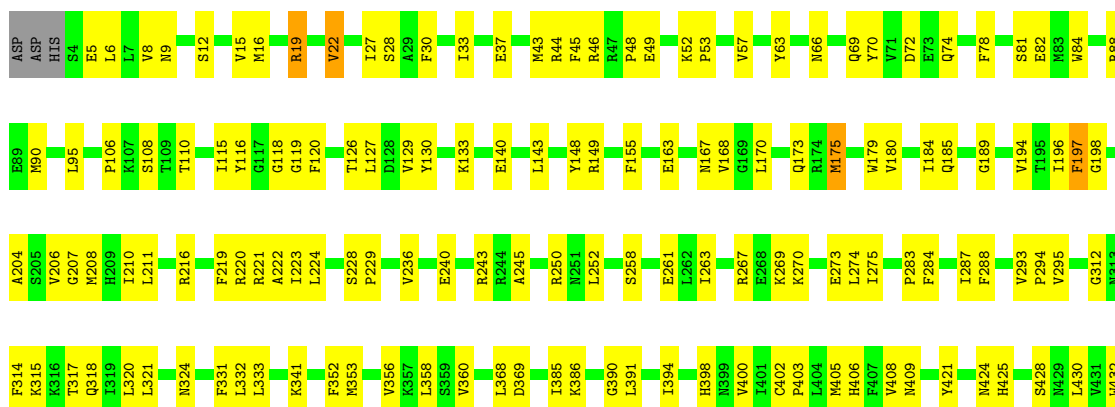
#### • Molecule 1: ACETYLCHOLINESTERASE

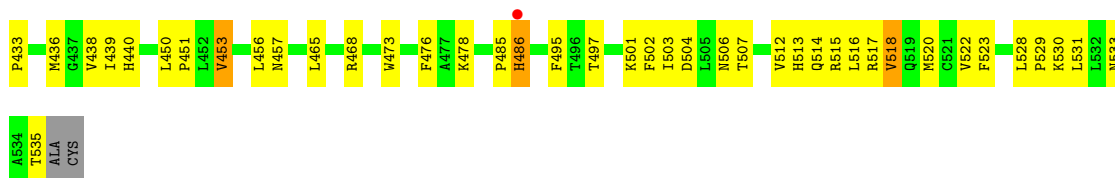
Chain A: 



#### • Molecule 1: ACETYLCHOLINESTERASE

Chain B: 







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.07Å 107.29Å 151.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.35 – 2.39 87.56 – 2.39	Depositor EDS
% Data completeness (in resolution range)	98.4 (46.35-2.39) 98.6 (87.56-2.39)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.20 (at 2.40Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.190 , 0.244 0.184 , 0.238	Depositor DCC
$R_{free}$ test set	2964 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtrriage
Anisotropy	0.778	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 55.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10185	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.50 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.7814e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: PGE, CL, CCD, NAG

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	0.34	0/4331	0.58	0/5879
1	B	0.35	0/4367	0.59	0/5929
All	All	0.34	0/8698	0.59	0/11808

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4210	0	4068	202	0
1	B	4244	0	4095	221	0
2	A	28	0	26	2	0
2	B	28	0	26	1	0
3	A	3	0	0	1	0
3	B	1	0	0	0	0
4	A	20	0	38	7	0
4	B	20	0	38	6	0
5	A	20	0	28	1	0
5	B	10	0	14	3	0
6	A	857	0	0	126	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	744	0	0	143	0
All	All	10185	0	8333	436	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 26.

The worst 5 of 436 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:493:PRO:HB2	6:A:2769:HOH:O	1.64	0.97
1:B:184:ILE:HG13	6:B:2301:HOH:O	1.72	0.89
1:A:168:VAL:HG13	6:A:2377:HOH:O	1.73	0.88
1:A:172:ASP:HB3	6:A:2356:HOH:O	1.73	0.88
1:A:243:ARG:HD3	6:A:2483:HOH:O	1.73	0.88

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	524/537 (98%)	493 (94%)	30 (6%)	1 (0%)	47 62
1	B	530/537 (99%)	505 (95%)	23 (4%)	2 (0%)	34 48
All	All	1054/1074 (98%)	998 (95%)	53 (5%)	3 (0%)	41 55

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	380	ASP
1	B	486	HIS
1	B	22	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	461/469 (98%)	448 (97%)	13 (3%)	43	63
1	B	465/469 (99%)	456 (98%)	9 (2%)	57	75
All	All	926/938 (99%)	904 (98%)	22 (2%)	49	68

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	74	GLN
1	B	284	PHE
1	B	197	PHE
1	B	288	PHE
1	A	353	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	74	GLN
1	B	457	ASN
1	B	514	GLN
1	A	383	ASN
1	A	514	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 4 are monoatomic - leaving 11 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
5	PGE	A	1542	-	9,9,9	1.03	0	8,8,8	1.00	0
4	CCD	B	1539	1	9,9,9	0.60	0	11,12,12	1.43	2 (18%)
2	NAG	B	1537	1	14,14,15	0.55	0	17,19,21	0.65	0
2	NAG	A	1537	1	14,14,15	0.63	0	17,19,21	0.66	1 (5%)
4	CCD	A	1539	1	9,9,9	0.58	0	11,12,12	1.79	2 (18%)
4	CCD	A	1540	-	9,9,9	0.78	0	11,12,12	1.28	2 (18%)
4	CCD	B	1540	-	9,9,9	0.71	0	11,12,12	1.30	2 (18%)
2	NAG	B	1536	1	14,14,15	0.50	0	17,19,21	0.75	1 (5%)
2	NAG	A	1536	1	14,14,15	0.48	0	17,19,21	0.68	1 (5%)
5	PGE	B	1541	-	9,9,9	1.42	1 (11%)	8,8,8	1.75	3 (37%)
5	PGE	A	1541	-	9,9,9	0.95	0	8,8,8	0.88	1 (12%)

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
5	PGE	A	1542	-	-	1/7/7/7	-
4	CCD	B	1539	1	-	1/7/7/7	-
4	CCD	A	1540	-	1/1/2/2	3/7/7/7	-
2	NAG	A	1537	1	-	4/6/23/26	0/1/1/1
4	CCD	B	1540	-	1/1/2/2	3/7/7/7	-
2	NAG	B	1537	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	CCD	A	1539	1	-	0/7/7/7	-
2	NAG	B	1536	1	-	2/6/23/26	0/1/1/1
2	NAG	A	1536	1	-	4/6/23/26	0/1/1/1
5	PGE	B	1541	-	-	5/7/7/7	-
5	PGE	A	1541	-	-	1/7/7/7	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1541	PGE	O3-C5	2.20	1.51	1.42

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1539	CCD	C6-C5-C4	-4.34	81.90	113.02
5	B	1541	PGE	C3-O2-C2	2.92	125.95	113.29
5	B	1541	PGE	O3-C5-C6	2.61	121.56	110.07
4	A	1539	CCD	O7-C5-C6	2.59	120.61	109.38
4	B	1539	CCD	C6-C5-C4	-2.53	94.90	113.02

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	1540	CCD	C5
4	B	1540	CCD	C5

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1536	NAG	C8-C7-N2-C2
2	A	1536	NAG	O7-C7-N2-C2
4	A	1540	CCD	C3-C4-C5-O7
4	B	1540	CCD	C2-C3-C4-C5
2	B	1536	NAG	C8-C7-N2-C2

There are no ring outliers.

8 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1539	CCD	5	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1537	NAG	1	0
4	A	1539	CCD	4	0
4	A	1540	CCD	3	0
4	B	1540	CCD	1	0
2	A	1536	NAG	2	0
5	B	1541	PGE	3	0
5	A	1541	PGE	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	528/537 (98%)	-0.18	0 <b>100</b>   <b>100</b>	22, 34, 51, 66	0
1	B	532/537 (99%)	-0.17	1 (0%) <b>95</b>   <b>94</b>	22, 34, 52, 77	0
All	All	1060/1074 (98%)	-0.17	1 (0%) <b>95</b>   <b>95</b>	22, 34, 51, 77	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	486	HIS	2.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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(Å <sup>2</sup> )	Q<0.9
5	PGE	B	1541	10/10	0.72	0.23	86,87,89,90	0
5	PGE	A	1541	10/10	0.77	0.23	83,85,85,85	0
5	PGE	A	1542	10/10	0.80	0.33	82,84,88,89	0
4	CCD	A	1540	10/10	0.83	0.27	89,89,90,90	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	1538	1/1	0.87	0.10	71,71,71,71	0
2	NAG	A	1536	14/15	0.88	0.10	59,63,65,66	0
2	NAG	A	1537	14/15	0.91	0.11	51,53,56,56	0
3	CL	A	1543	1/1	0.91	0.15	86,86,86,86	0
2	NAG	B	1537	14/15	0.91	0.12	43,46,50,53	0
4	CCD	B	1540	10/10	0.92	0.25	83,84,85,85	0
3	CL	A	1544	1/1	0.92	0.19	65,65,65,65	0
4	CCD	A	1539	10/10	0.93	0.20	37,43,48,49	0
4	CCD	B	1539	10/10	0.93	0.19	32,41,44,45	0
3	CL	B	1538	1/1	0.94	0.07	64,64,64,64	0
2	NAG	B	1536	14/15	0.94	0.10	57,59,62,62	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.