



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 04:14 am BST

PDB ID : 4IC0  
Title : Crystal Structure of PAI-1 in Complex with Gallate  
Authors : Hong, Z.B.; Lin, Z.H.; Gong, L.H.; Huang, M.D.  
Deposited on : 2012-12-09  
Resolution : 2.32 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

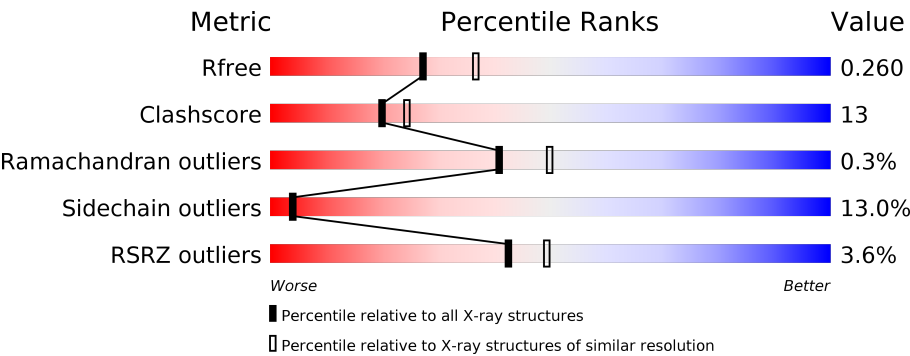
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.32 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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	379	
1	B	379	
1	C	379	
1	D	379	

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
2	GDE	A	401	-	-	X	-

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11613 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.

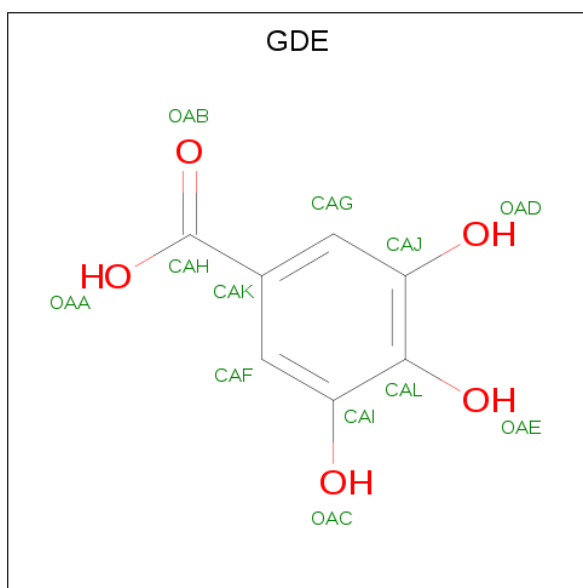
- Molecule 1 is a protein called Plasminogen activator inhibitor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	366	Total	C	N	O	S	0	0	0
			2924	1881	499	529	15			
1	B	358	Total	C	N	O	S	0	0	0
			2867	1845	488	520	14			
1	C	366	Total	C	N	O	S	0	0	0
			2924	1881	499	529	15			
1	D	359	Total	C	N	O	S	0	0	0
			2874	1850	489	521	14			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	150	HIS	ASN	ENGINEERED MUTATION	UNP P05121
A	154	THR	LYS	ENGINEERED MUTATION	UNP P05121
A	158	GLN	ASP	ENGINEERED MUTATION	UNP P05121
A	319	LEU	GLN	ENGINEERED MUTATION	UNP P05121
A	354	ILE	MET	ENGINEERED MUTATION	UNP P05121
B	150	HIS	ASN	ENGINEERED MUTATION	UNP P05121
B	154	THR	LYS	ENGINEERED MUTATION	UNP P05121
B	158	GLN	ASP	ENGINEERED MUTATION	UNP P05121
B	319	LEU	GLN	ENGINEERED MUTATION	UNP P05121
B	354	ILE	MET	ENGINEERED MUTATION	UNP P05121
C	150	HIS	ASN	ENGINEERED MUTATION	UNP P05121
C	154	THR	LYS	ENGINEERED MUTATION	UNP P05121
C	158	GLN	ASP	ENGINEERED MUTATION	UNP P05121
C	319	LEU	GLN	ENGINEERED MUTATION	UNP P05121
C	354	ILE	MET	ENGINEERED MUTATION	UNP P05121
D	150	HIS	ASN	ENGINEERED MUTATION	UNP P05121
D	154	THR	LYS	ENGINEERED MUTATION	UNP P05121
D	158	GLN	ASP	ENGINEERED MUTATION	UNP P05121
D	319	LEU	GLN	ENGINEERED MUTATION	UNP P05121
D	354	ILE	MET	ENGINEERED MUTATION	UNP P05121

- Molecule 2 is 3,4,5-trihydroxybenzoic acid (three-letter code: GDE) (formula:  $C_7H_6O_5$ ).

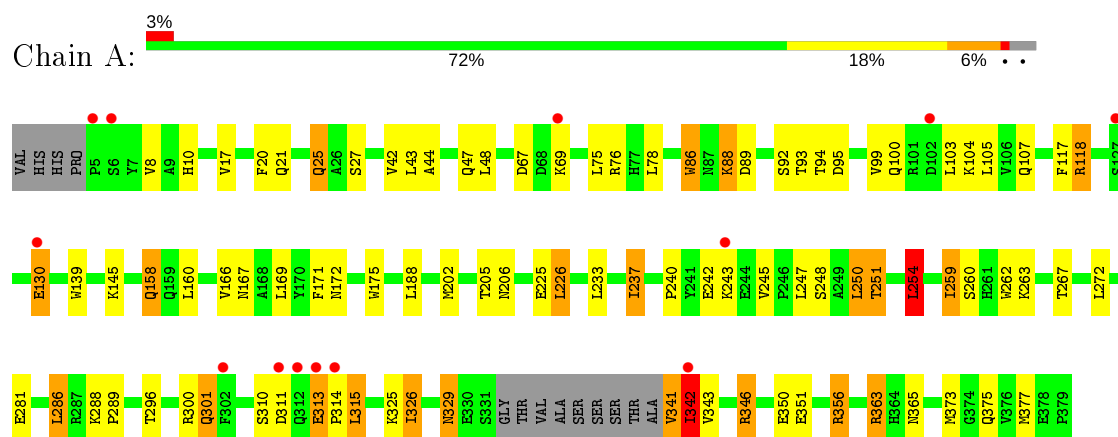


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	7	5		
2	C	1	Total	C	O	0	0
			12	7	5		

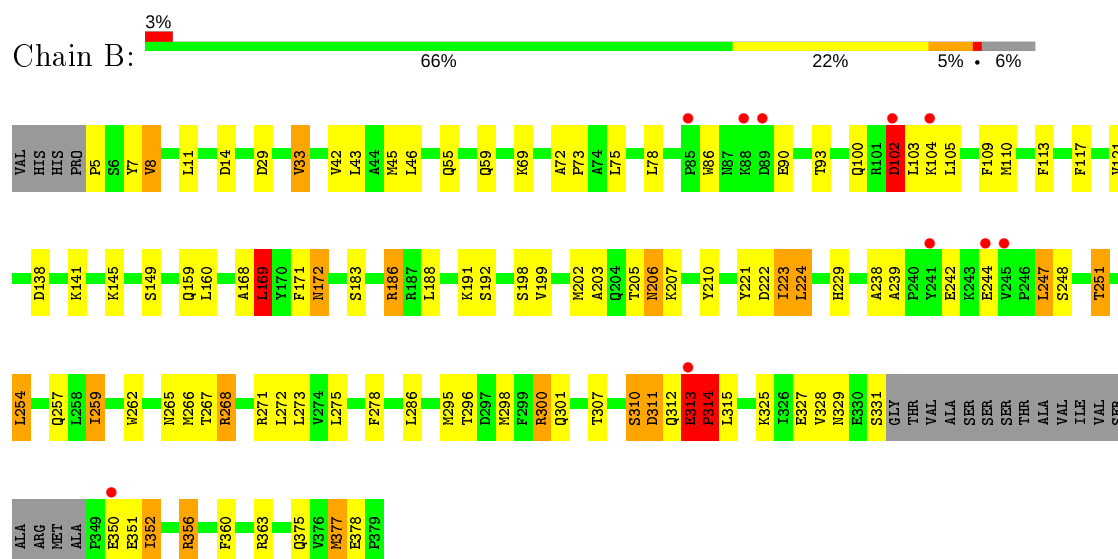
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Plasminogen activator inhibitor 1

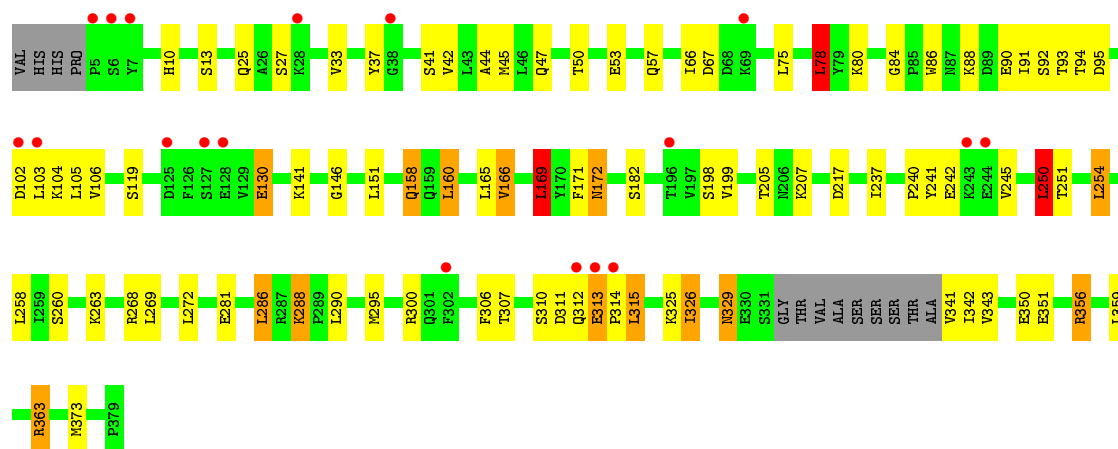


#### • Molecule 1: Plasminogen activator inhibitor 1



#### • Molecule 1: Plasminogen activator inhibitor 1





• Molecule 1: Plasminogen activator inhibitor 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.28Å 74.99Å 103.87Å 90.91° 93.29° 115.82°	Depositor
Resolution (Å)	38.25 – 2.32 38.25 – 2.32	Depositor EDS
% Data completeness (in resolution range)	94.2 (38.25-2.32) 94.3 (38.25-2.32)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.66 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.202 , 0.258 0.210 , 0.260	Depositor DCC
$R_{free}$ test set	3628 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.5	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 38.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.007 for -h,h+k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11613	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% 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: GDE

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	1.04	4/2995 (0.1%)	1.01	10/4060 (0.2%)
1	B	0.99	0/2938	0.95	4/3982 (0.1%)
1	C	1.06	1/2995 (0.0%)	0.98	8/4060 (0.2%)
1	D	0.95	0/2946	0.97	4/3994 (0.1%)
All	All	1.01	5/11874 (0.0%)	0.98	26/16096 (0.2%)

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	2
1	B	0	1
1	C	0	2
1	D	0	1
All	All	0	6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	86	TRP	NE1-CE2	-6.16	1.29	1.37
1	A	139	TRP	CD2-CE2	5.85	1.48	1.41
1	A	175	TRP	CD2-CE2	5.42	1.47	1.41
1	A	262	TRP	CD2-CE2	5.18	1.47	1.41
1	C	86	TRP	CD2-CE2	5.10	1.47	1.41

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	169	LEU	CA-CB-CG	8.42	134.67	115.30
1	C	250	LEU	CA-CB-CG	7.71	133.03	115.30
1	C	217	ASP	CB-CG-OD2	7.62	125.16	118.30
1	A	343	VAL	CB-CA-C	-7.45	97.24	111.40
1	A	314	PRO	N-CA-C	-6.85	94.29	112.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	313	GLU	Peptide
1	A	341	VAL	Peptide
1	B	102	ASP	Peptide
1	C	102	ASP	Peptide
1	C	313	GLU	Peptide

## 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	2924	0	2926	83	0
1	B	2867	0	2862	79	0
1	C	2924	0	2926	62	0
1	D	2874	0	2869	69	0
2	A	12	0	2	5	0
2	C	12	0	2	3	0
All	All	11613	0	11587	292	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 13.

The worst 5 of 292 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:346:ARG:CG	1:A:346:ARG:HH11	1.60	1.14
1:C:158:GLN:HE22	1:C:160:LEU:HB2	1.01	1.07

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:205:THR:HG22	1:A:272:LEU:HA	1.33	1.07
1:C:93:THR:HG22	1:C:169:LEU:HD13	1.26	1.06
1:A:93:THR:HG22	1:A:169:LEU:HD23	1.38	1.03

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	362/379 (96%)	355 (98%)	7 (2%)	0	100	100
1	B	354/379 (93%)	347 (98%)	5 (1%)	2 (1%)	25	30
1	C	362/379 (96%)	350 (97%)	11 (3%)	1 (0%)	41	50
1	D	355/379 (94%)	348 (98%)	5 (1%)	2 (1%)	25	30
All	All	1433/1516 (94%)	1400 (98%)	28 (2%)	5 (0%)	41	50

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	313	GLU
1	B	314	PRO
1	D	314	PRO
1	D	5	PRO
1	C	84	GLY

### 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	323/333 (97%)	286 (88%)	37 (12%)	5	6
1	B	317/333 (95%)	270 (85%)	47 (15%)	3	2
1	C	323/333 (97%)	282 (87%)	41 (13%)	4	4
1	D	318/333 (96%)	277 (87%)	41 (13%)	4	4
All	All	1281/1332 (96%)	1115 (87%)	166 (13%)	4	4

5 of 166 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	314	PRO
1	C	78	LEU
1	D	257	GLN
1	B	325	LYS
1	C	10	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	375	GLN
1	C	265	ASN
1	D	316	HIS
1	C	158	GLN
1	C	292	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

2 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	GDE	C	401	-	10,12,12	1.98	3 (30%)	14,17,17	3.34	9 (64%)
2	GDE	A	401	-	10,12,12	3.32	7 (70%)	14,17,17	3.64	5 (35%)

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	GDE	C	401	-	-	0/0/4/4	0/1/1/1
2	GDE	A	401	-	-	0/0/4/4	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	GDE	CAK-CAH	5.95	1.53	1.47
2	A	401	GDE	CAI-CAL	-5.07	1.33	1.40
2	A	401	GDE	CAJ-CAL	-5.02	1.33	1.40
2	C	401	GDE	CAJ-CAL	-4.21	1.34	1.40
2	A	401	GDE	OAE-CAL	-2.59	1.30	1.37

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	GDE	CAF-CAK-CAH	8.29	131.26	120.36
2	C	401	GDE	CAF-CAI-CAL	-7.02	115.64	120.47
2	A	401	GDE	CAG-CAK-CAH	-6.39	111.95	120.36
2	C	401	GDE	CAG-CAJ-CAL	6.31	124.81	120.47
2	A	401	GDE	CAF-CAI-CAL	-5.77	116.50	120.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	401	GDE	3	0
2	A	401	GDE	5	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	366/379 (96%)	0.08	13 (3%) 42 49	17, 30, 54, 85	0
1	B	358/379 (94%)	0.14	10 (2%) 53 60	17, 33, 57, 88	0
1	C	366/379 (96%)	0.17	18 (4%) 29 37	16, 29, 54, 91	0
1	D	359/379 (94%)	0.13	11 (3%) 49 56	16, 33, 55, 94	0
All	All	1449/1516 (95%)	0.13	52 (3%) 42 49	16, 31, 55, 94	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	5	PRO	5.6
1	A	312	GLN	5.5
1	B	313	GLU	4.7
1	C	244	GLU	4.1
1	D	313	GLU	4.0

### 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. 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
2	GDE	A	401	12/12	0.68	0.29	20,20,20,20	0
2	GDE	C	401	12/12	0.75	0.22	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.