



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 25, 2017 – 02:28 PM EDT

PDB ID : 3KRO  
Title : Mint heterotetrameric geranyl pyrophosphate synthase in complex with magnesium, IPP, and DMASPP (II)  
Authors : Chang, T.-H.; Hsieh, F.-L.; Ko, T.-P.; Wang, A.H.-J.  
Deposited on : unknown  
Resolution : 1.95 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20030345

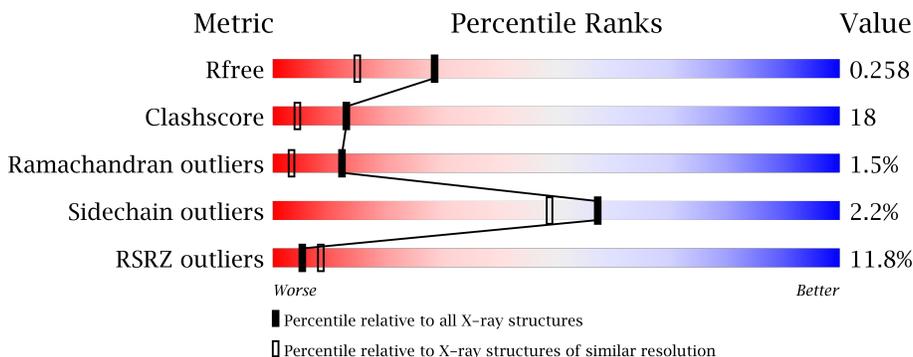
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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))
$R_{free}$	100719	2004 (1.96-1.96)
Clashscore	112137	2136 (1.96-1.96)
Ramachandran outliers	110173	2117 (1.96-1.96)
Sidechain outliers	110143	2117 (1.96-1.96)
RSRZ outliers	101464	2018 (1.96-1.96)

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. 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	295	 7% 74% 20% • 5%
1	D	295	 8% 73% 26% •
2	B	274	 18% 58% 33% • • 5%
2	C	274	 14% 65% 26% • 5%

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
6	EDO	C	2005	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 9441 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 Geranyl diphosphate synthase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	279	2120	1338	372	392	18	0	0	0
1	D	295	2232	1406	391	417	18	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	EXPRESSION TAG	UNP Q9SBR3
D	1	MET	-	EXPRESSION TAG	UNP Q9SBR3

- Molecule 2 is a protein called Geranyl diphosphate synthase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	259	1954	1233	339	368	14	0	0	0
2	C	259	1954	1233	339	368	14	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

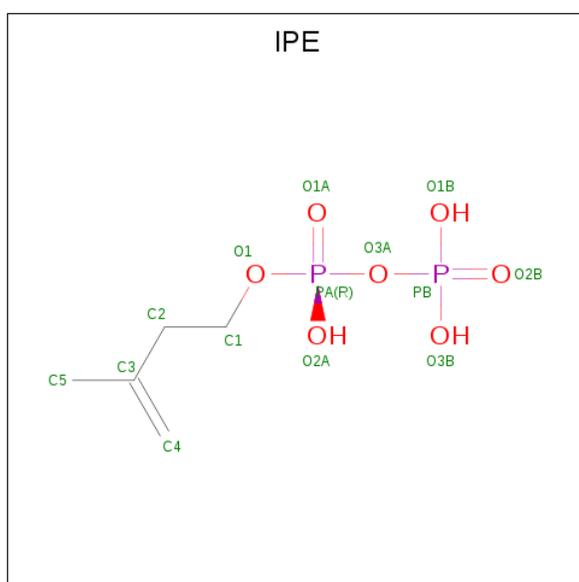
Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	EXPRESSION TAG	UNP Q9SBR4
B	267	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	268	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	269	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	270	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	271	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	272	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	273	HIS	-	EXPRESSION TAG	UNP Q9SBR4
B	274	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	1	MET	-	EXPRESSION TAG	UNP Q9SBR4

*Continued on next page...*

Continued from previous page...

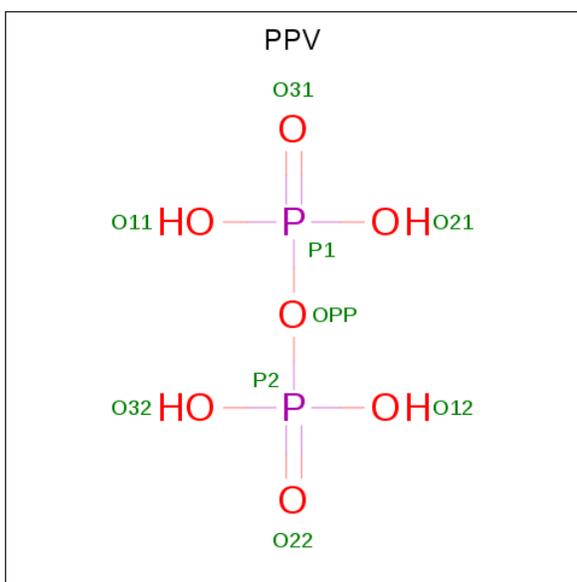
Chain	Residue	Modelled	Actual	Comment	Reference
C	267	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	268	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	269	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	270	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	271	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	272	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	273	HIS	-	EXPRESSION TAG	UNP Q9SBR4
C	274	HIS	-	EXPRESSION TAG	UNP Q9SBR4

- Molecule 3 is 3-METHYLBUT-3-ENYL TRIHYDROGEN DIPHOSPHATE (three-letter code: IPE) (formula:  $C_5H_{12}O_7P_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	P	0	0
			14	5	7	2		
3	D	1	Total	C	O	P	0	0
			14	5	7	2		

- Molecule 4 is PYROPHOSPHATE (three-letter code: PPV) (formula:  $H_4O_7P_2$ ).

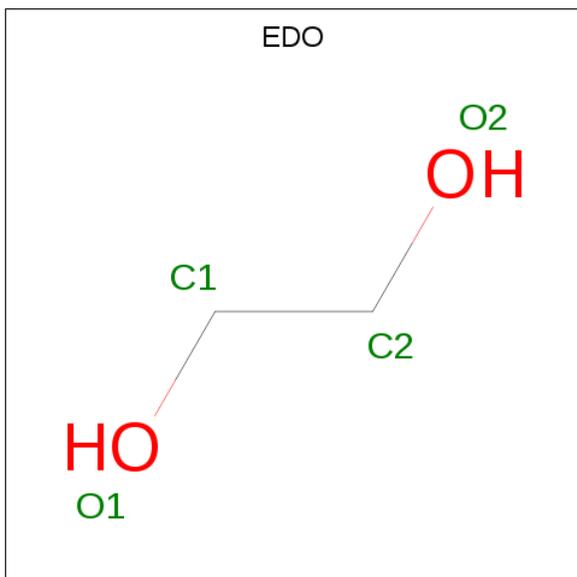


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 9 7 2	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

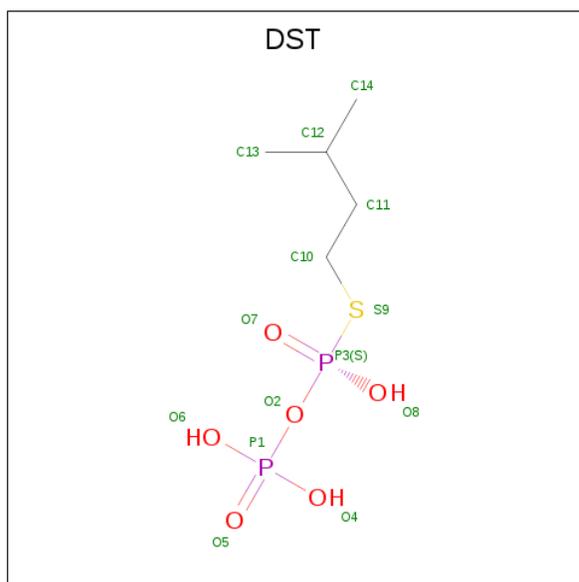
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Mg 2 2	0	0
5	D	2	Total Mg 2 2	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C O 4 2 2	0	0

- Molecule 7 is DIMETHYLALLYL S-THIOLODIPHOSPHATE (three-letter code: DST) (formula:  $C_5H_{14}O_6P_2S$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total C O P S 14 5 6 2 1	0	0

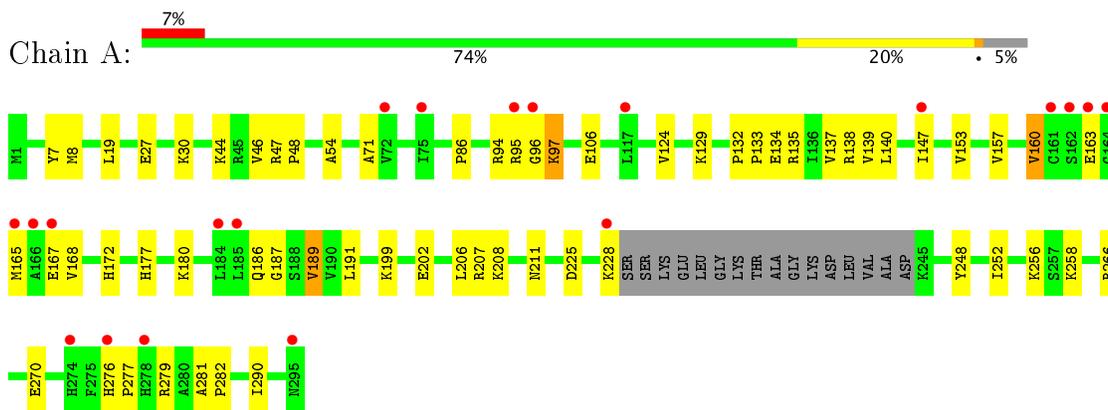
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	268	Total O 268 268	0	0
8	B	243	Total O 243 243	0	0
8	C	294	Total O 294 294	0	0
8	D	317	Total O 317 317	0	0

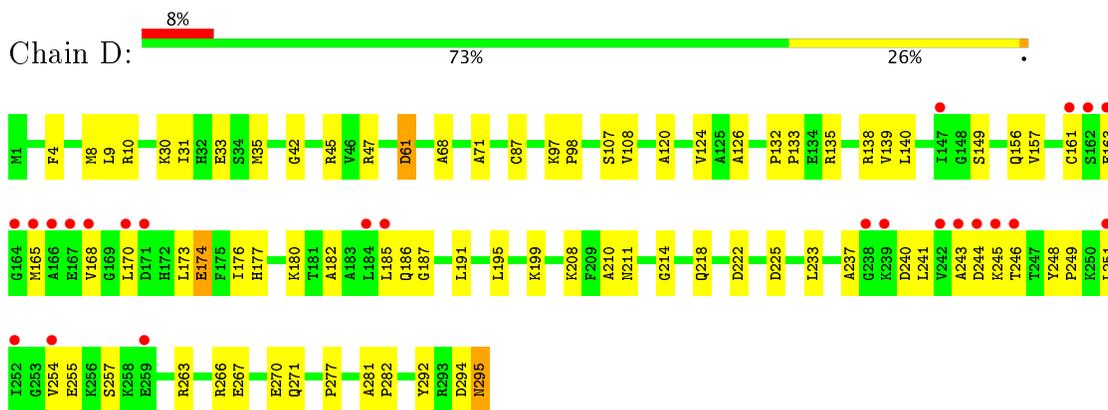
### 3 Residue-property plots [i](#)

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

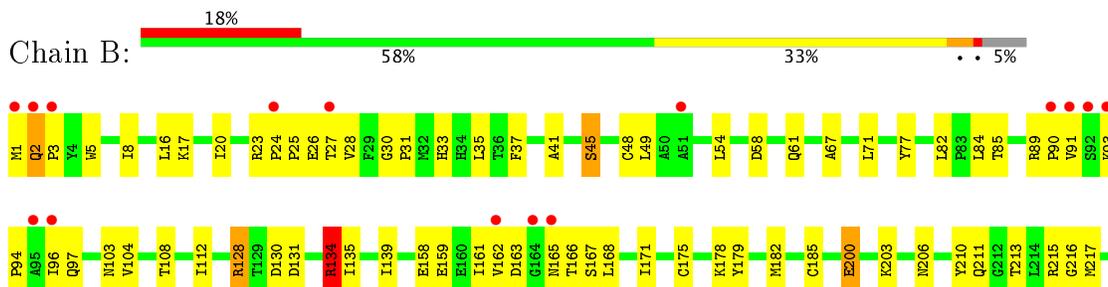
- Molecule 1: Geranyl diphosphate synthase large subunit



- Molecule 1: Geranyl diphosphate synthase large subunit

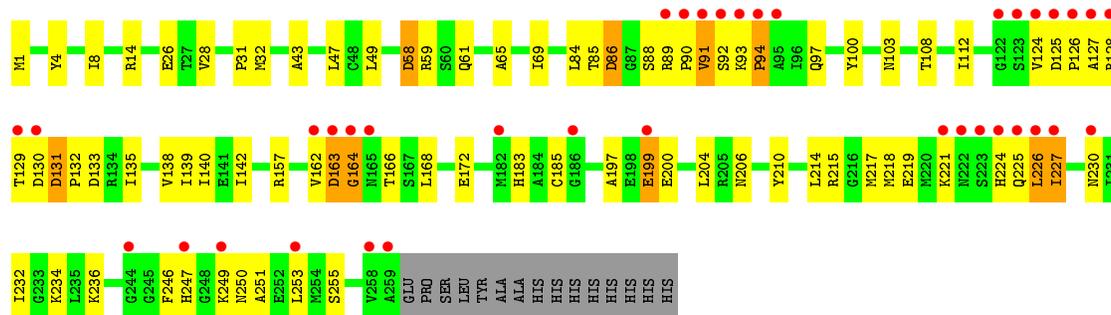


- Molecule 2: Geranyl diphosphate synthase small subunit





● Molecule 2: Geranyl diphosphate synthase small subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.44Å 108.70Å 182.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.95 29.02 – 1.94	Depositor EDS
% Data completeness (in resolution range)	90.1 (30.00-1.95) 89.6 (29.02-1.94)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.34 (at 1.95Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.205 , 0.256 0.205 , 0.258	Depositor DCC
$R_{free}$ test set	3588 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtriage
Anisotropy	0.820	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 57.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9441	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.96% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PPV, MG, EDO, DST, IPE

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	0.96	2/2155 (0.1%)	0.85	0/2903
1	D	0.93	0/2268	0.87	3/3055 (0.1%)
2	B	0.83	2/1993 (0.1%)	0.82	1/2695 (0.0%)
2	C	0.84	1/1993 (0.1%)	0.83	1/2695 (0.0%)
All	All	0.89	5/8409 (0.1%)	0.84	5/11348 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	175	CYS	CB-SG	5.74	1.92	1.82
1	A	27	GLU	CG-CD	5.23	1.59	1.51
1	A	46	VAL	CA-CB	5.20	1.65	1.54
2	B	200	GLU	CG-CD	5.15	1.59	1.51
2	C	185	CYS	CB-SG	-5.05	1.73	1.81

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	47	ARG	NE-CZ-NH2	-6.67	116.96	120.30
2	C	157	ARG	NE-CZ-NH1	-6.08	117.26	120.30
1	D	45	ARG	NE-CZ-NH2	-5.75	117.43	120.30
1	D	61	ASP	CB-CG-OD2	-5.44	113.41	118.30
2	B	134	ARG	NE-CZ-NH1	5.32	122.96	120.30

There are no chirality outliers.

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	2120	0	2146	51	0
1	D	2232	0	2264	79	0
2	B	1954	0	1948	103	0
2	C	1954	0	1948	83	0
3	A	14	0	9	1	0
3	D	14	0	9	0	0
4	A	9	0	0	0	0
5	A	2	0	0	0	0
5	D	2	0	0	0	0
6	C	4	0	6	0	0
7	D	14	0	10	3	0
8	A	268	0	0	14	1
8	B	243	0	0	13	1
8	C	294	0	0	15	1
8	D	317	0	0	15	1
All	All	9441	0	8340	307	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 18.

The worst 5 of 307 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)
2:B:2:GLN:HB2	2:B:3:PRO:HD3	1.21	1.20
2:B:229:GLU:HG3	8:B:289:HOH:O	1.39	1.17
2:C:250:ASN:HB3	8:C:1039:HOH:O	1.57	1.03
2:C:230:ASN:O	2:C:234:LYS:HG3	1.58	1.02
2:C:28:VAL:HG22	1:D:157:VAL:CG2	1.96	0.94

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)
8:A:843:HOH:O	8:C:805:HOH:O[4_555]	1.99	0.21

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:B:589:HOH:O	8:D:588:HOH:O[3_555]	2.07	0.13

## 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	275/295 (93%)	263 (96%)	12 (4%)	0	100	100
1	D	293/295 (99%)	280 (96%)	12 (4%)	1 (0%)	44	33
2	B	257/274 (94%)	237 (92%)	14 (5%)	6 (2%)	7	1
2	C	257/274 (94%)	230 (90%)	18 (7%)	9 (4%)	4	0
All	All	1082/1138 (95%)	1010 (93%)	56 (5%)	16 (2%)	12	3

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	227	ILE
2	C	91	VAL
2	C	163	ASP
2	C	227	ILE
2	B	2	GLN

### 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	222/234 (95%)	218 (98%)	4 (2%)	64	57
1	D	234/234 (100%)	231 (99%)	3 (1%)	73	69
2	B	201/214 (94%)	193 (96%)	8 (4%)	36	21
2	C	201/214 (94%)	197 (98%)	4 (2%)	60	52
All	All	858/896 (96%)	839 (98%)	19 (2%)	57	49

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	227	ILE
2	B	232	ILE
2	C	199	GLU
2	B	134	ARG
1	D	174	GLU

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

Mol	Chain	Res	Type
2	C	225	GLN
2	C	230	ASN
1	D	186	GLN
2	C	183	HIS
1	D	156	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 carbohydrates in this entry.

## 5.6 Ligand geometry

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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
3	IPE	A	2003	-	12,13,13	3.31	4 (33%)	12,19,19	1.30	2 (16%)
4	PPV	A	2004	5	8,8,8	1.44	1 (12%)	8,13,13	0.36	0
6	EDO	C	2005	-	3,3,3	0.70	0	2,2,2	0.46	0
3	IPE	D	2001	-	12,13,13	3.24	3 (25%)	12,19,19	1.11	1 (8%)
7	DST	D	2002	5	10,13,13	1.84	2 (20%)	9,19,19	2.56	3 (33%)

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
3	IPE	A	2003	-	-	0/13/13/13	0/0/0/0
4	PPV	A	2004	5	-	0/6/6/6	0/0/0/0
6	EDO	C	2005	-	-	0/1/1/1	0/0/0/0
3	IPE	D	2001	-	-	0/13/13/13	0/0/0/0
7	DST	D	2002	5	-	0/9/13/13	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	2001	IPE	C2-C3	-7.35	1.36	1.51
3	A	2003	IPE	C2-C3	-6.96	1.36	1.51
7	D	2002	DST	P1-O2	-4.46	1.52	1.60
3	D	2001	IPE	PB-O3A	-3.38	1.54	1.60
3	A	2003	IPE	PB-O3A	-2.72	1.55	1.60

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	2002	DST	O4-P1-O5	-3.26	97.75	110.50
3	A	2003	IPE	O3B-PB-O1B	-2.41	97.90	107.61
3	D	2001	IPE	C5-C3-C4	-2.23	116.71	121.88
7	D	2002	DST	C13-C12-C11	2.24	125.96	111.50
3	A	2003	IPE	O1-C1-C2	2.28	118.42	108.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2003	IPE	1	0
7	D	2002	DST	3	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	279/295 (94%)	0.28	20 (7%) 16 25	27, 39, 68, 99	0
1	D	295/295 (100%)	0.45	24 (8%) 13 20	26, 41, 91, 109	0
2	B	259/274 (94%)	0.90	48 (18%) 1 1	29, 46, 116, 126	0
2	C	259/274 (94%)	0.82	37 (14%) 3 5	27, 46, 96, 115	0
All	All	1092/1138 (95%)	0.60	129 (11%) 5 8	26, 42, 98, 126	0

The worst 5 of 129 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	251	LEU	12.5
2	B	248	GLY	10.4
2	B	257	LEU	9.2
2	B	246	PHE	8.7
2	B	250	ASN	8.3

### 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<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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	EDO	C	2005	4/4	0.96	0.19	2.43	36,37,37,38	0
3	IPE	A	2003	14/14	0.89	0.20	1.15	72,81,88,88	0
7	DST	D	2002	14/14	0.96	0.15	-0.08	49,55,62,62	0
3	IPE	D	2001	14/14	0.98	0.12	-0.21	39,40,45,48	0
5	MG	D	3002	1/1	0.95	0.06	-0.86	49,49,49,49	0
4	PPV	A	2004	9/9	0.93	0.10	-1.09	76,76,77,78	0
5	MG	D	3001	1/1	0.93	0.06	-1.16	53,53,53,53	0
5	MG	A	3003	1/1	0.94	0.11	-	50,50,50,50	0
5	MG	A	3004	1/1	0.83	0.07	-	58,58,58,58	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.