



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 03:54 PM GMT

PDB ID : 4DTT
Title : Crystal structure of human insulin degrading enzyme (ide) in complex with
compund 41367
Authors : Guo, Q.; Deprez-Poulain, R.; Deprez, B.; Tang, W.J.
Deposited on : 2012-02-21
Resolution : 3.22 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

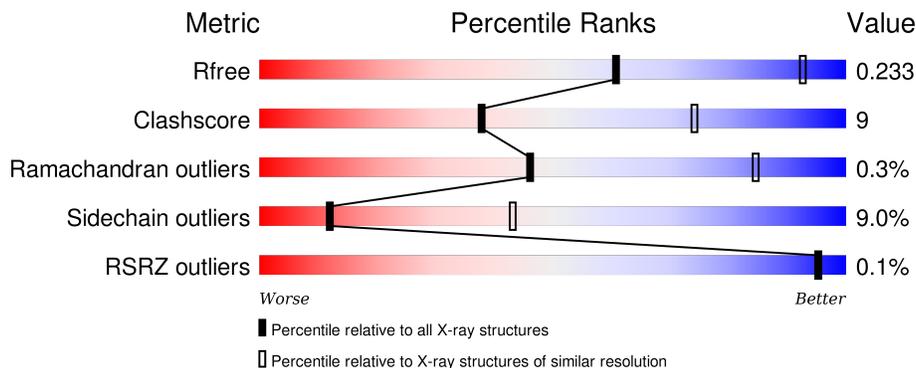
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.22 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}	91344	1095 (3.26-3.18)
Clashscore	102246	1046 (3.24-3.20)
Ramachandran outliers	100387	1026 (3.24-3.20)
Sidechain outliers	100360	1025 (3.24-3.20)
RSRZ outliers	91569	1100 (3.26-3.18)

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 ≥ 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	990	 73% 21% . .
1	B	990	 71% 23% . .

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
3	I41	A	1102	-	-	-	X
3	I41	A	1103	-	-	-	X
3	I41	B	1102	-	-	-	X
3	I41	B	1103	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 15713 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 Insulin-degrading enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	956	7768	5008	1306	1432	22	0	0	0
1	B	954	7758	5002	1304	1430	22	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	EXPRESSION TAG	UNP P14735
A	31	HIS	-	EXPRESSION TAG	UNP P14735
A	32	HIS	-	EXPRESSION TAG	UNP P14735
A	33	HIS	-	EXPRESSION TAG	UNP P14735
A	34	HIS	-	EXPRESSION TAG	UNP P14735
A	35	HIS	-	EXPRESSION TAG	UNP P14735
A	36	HIS	-	EXPRESSION TAG	UNP P14735
A	37	ALA	-	EXPRESSION TAG	UNP P14735
A	38	ALA	-	EXPRESSION TAG	UNP P14735
A	39	GLY	-	EXPRESSION TAG	UNP P14735
A	40	ILE	-	EXPRESSION TAG	UNP P14735
A	41	PRO	-	EXPRESSION TAG	UNP P14735
A	110	LEU	CYS	ENGINEERED MUTATION	UNP P14735
A	171	SER	CYS	ENGINEERED MUTATION	UNP P14735
A	178	ALA	CYS	ENGINEERED MUTATION	UNP P14735
A	257	VAL	CYS	ENGINEERED MUTATION	UNP P14735
A	414	LEU	CYS	ENGINEERED MUTATION	UNP P14735
A	573	ASN	CYS	ENGINEERED MUTATION	UNP P14735
A	590	SER	CYS	ENGINEERED MUTATION	UNP P14735
A	789	SER	CYS	ENGINEERED MUTATION	UNP P14735
A	812	ALA	CYS	ENGINEERED MUTATION	UNP P14735
A	819	ALA	CYS	ENGINEERED MUTATION	UNP P14735
A	904	SER	CYS	ENGINEERED MUTATION	UNP P14735
A	966	ASN	CYS	ENGINEERED MUTATION	UNP P14735
A	974	ALA	CYS	ENGINEERED MUTATION	UNP P14735

Continued on next page...

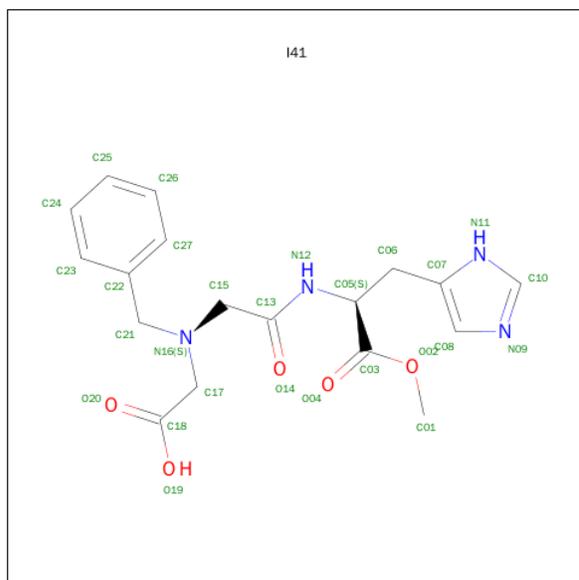
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	30	MET	-	EXPRESSION TAG	UNP P14735
B	31	HIS	-	EXPRESSION TAG	UNP P14735
B	32	HIS	-	EXPRESSION TAG	UNP P14735
B	33	HIS	-	EXPRESSION TAG	UNP P14735
B	34	HIS	-	EXPRESSION TAG	UNP P14735
B	35	HIS	-	EXPRESSION TAG	UNP P14735
B	36	HIS	-	EXPRESSION TAG	UNP P14735
B	37	ALA	-	EXPRESSION TAG	UNP P14735
B	38	ALA	-	EXPRESSION TAG	UNP P14735
B	39	GLY	-	EXPRESSION TAG	UNP P14735
B	40	ILE	-	EXPRESSION TAG	UNP P14735
B	41	PRO	-	EXPRESSION TAG	UNP P14735
B	110	LEU	CYS	ENGINEERED MUTATION	UNP P14735
B	171	SER	CYS	ENGINEERED MUTATION	UNP P14735
B	178	ALA	CYS	ENGINEERED MUTATION	UNP P14735
B	257	VAL	CYS	ENGINEERED MUTATION	UNP P14735
B	414	LEU	CYS	ENGINEERED MUTATION	UNP P14735
B	573	ASN	CYS	ENGINEERED MUTATION	UNP P14735
B	590	SER	CYS	ENGINEERED MUTATION	UNP P14735
B	789	SER	CYS	ENGINEERED MUTATION	UNP P14735
B	812	ALA	CYS	ENGINEERED MUTATION	UNP P14735
B	819	ALA	CYS	ENGINEERED MUTATION	UNP P14735
B	904	SER	CYS	ENGINEERED MUTATION	UNP P14735
B	966	ASN	CYS	ENGINEERED MUTATION	UNP P14735
B	974	ALA	CYS	ENGINEERED MUTATION	UNP P14735

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0

- Molecule 3 is 2-[[2-[[[(2S)-3-(3H-IMIDAZOL-4-YL)-1-METHOXY-1-OXO-PROPAN-2-YL] AMINO]-2-OXO-ETHYL]-(PHENYLMETHYL)AMINO]ETHANOIC ACID (three-letter code: I41) (formula: C₁₈H₂₂N₄O₅).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			27	18	4	5		
3	A	1	Total	C	N	O	0	0
			27	18	4	5		
3	B	1	Total	C	N	O	0	0
			27	18	4	5		
3	B	1	Total	C	N	O	0	0
			27	18	4	5		

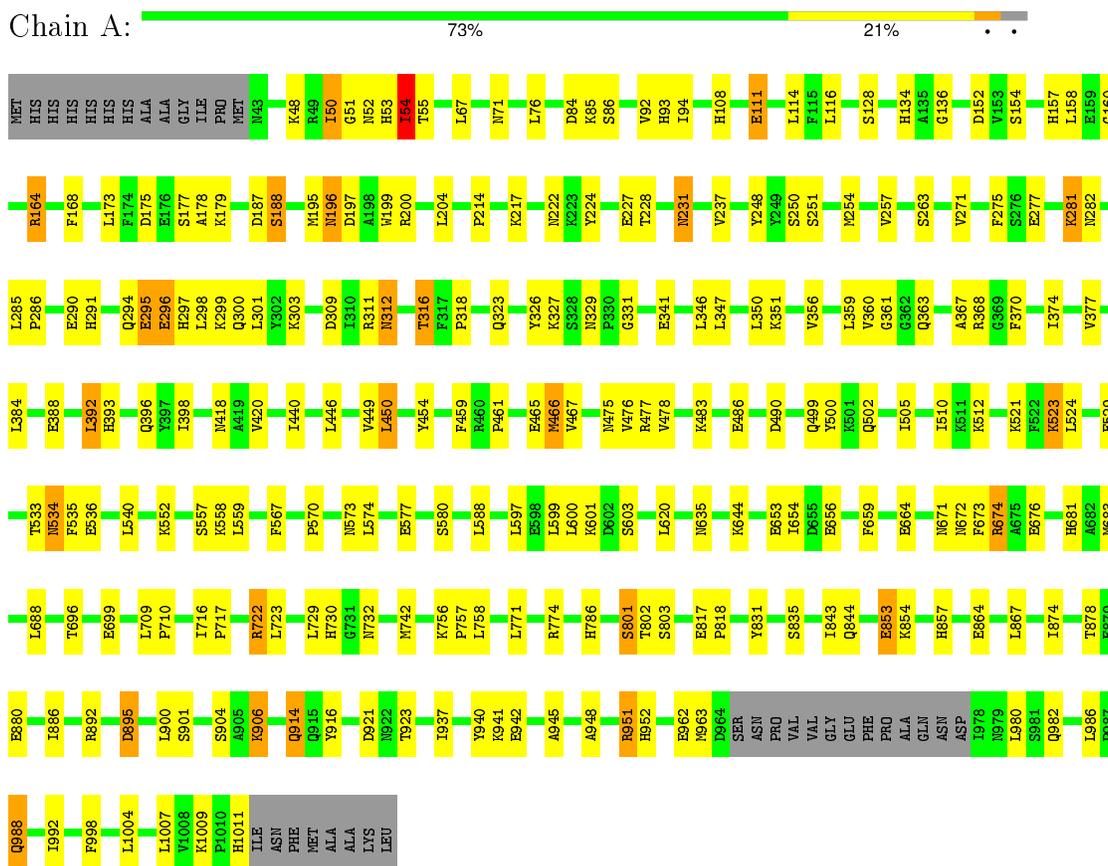
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total	O	0	0
			47	47		
4	B	30	Total	O	0	0
			30	30		

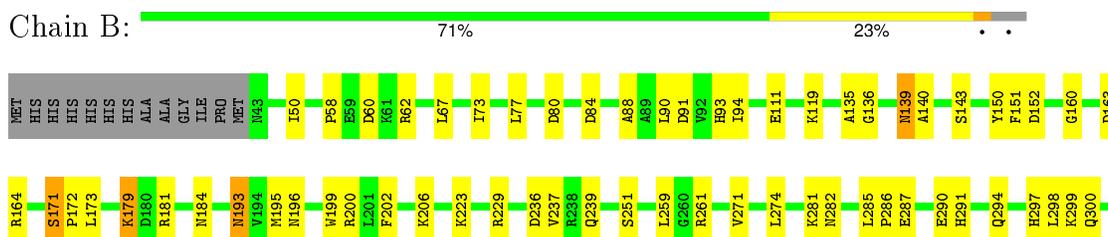
3 Residue-property plots [i](#)

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



- Molecule 1: Insulin-degrading enzyme



D809	R445	L686	R824	L959
Y314	L446	N573	R825	A960
P318	E447	F578	Q828	R961
Y325	L450	D586	W695	E962
Y326	T451	P587	R689	N963
K327	A452	L588	L700	D964
	E453	L597	K701	SER
	Y454	E598	L709	ASN
R336	E458	L599	Q844	PRO
L337	F459	L600	Q845	VAL
	R460	L604	R711	VAL
	M466	E612	L604	GLY
	E349	L616	I716	GLU
	R472	L620	Q851	PHE
S352	M475	T623	S852	PRO
K353	V476	Y624	E853	ALA
G354	A479	Y625	R859	GLN
Y356	E494	G626	L859	ASN
L359	T498	M627	E864	ASP
	R368	N635	L867	ILE
	F370	Q638	L874	ASN
R371	Y500	L641	R874	L980
F372	Y501	L645	E875	P384
	Q502	E647	D876	A985
L379	I510	R653	R877	L986
	K511	L645	E880	I992
L391	K512	L646	A881	Q993
R394	W513	E653	R884	R1000
Q396	N514	L646	L891	L1004
	A516	R653	R892	P1010
E413	K523	L662	R896	H1011
	D416	E676	P897	ILE
	L417	Q677	R898	ASN
	N418	P678	R899	PHE
	A419	R679	L900	MET
	V420	Q680	S901	ALA
	T545	M683	Q814	ALA
F422	I550		Q915	LYS
R423	T554		R915	LEU
F424	L559		R920	
K425	M560		L670	
E428	D564		L810	
H442	Y443		Q813	
Y444	D665		S816	
			E817	
			P818	
			A819	
			T822	
			L823	

4 Data and refinement statistics i

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	263.52Å 263.52Å 90.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.22 49.80 – 3.22	Depositor EDS
% Data completeness (in resolution range)	99.2 (50.00-3.22) 99.2 (49.80-3.22)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.66 (at 3.19Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.175 , 0.242 0.172 , 0.233	Depositor DCC
R_{free} test set	2938 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	60.1	Xtrriage
Anisotropy	0.179	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 48.4	EDS
Estimated twinning fraction	0.025 for h,-h-k,-l	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Outliers	0 of 58137 reflections	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15713	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.71% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: ZN, I41

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	0.82	4/7963 (0.1%)	0.83	7/10779 (0.1%)
1	B	0.80	3/7953 (0.0%)	0.80	0/10765
All	All	0.81	7/15916 (0.0%)	0.82	7/21544 (0.0%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	656	GLU	CG-CD	5.93	1.60	1.51
1	B	447	GLU	CB-CG	5.64	1.62	1.52
1	A	676	GLU	CG-CD	5.14	1.59	1.51
1	B	699	GLU	CG-CD	5.14	1.59	1.51
1	A	853	GLU	CG-CD	5.12	1.59	1.51

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	674	ARG	NE-CZ-NH1	5.86	123.23	120.30
1	A	895	ASP	CB-CG-OD1	5.50	123.25	118.30
1	A	94	ILE	CB-CA-C	-5.47	100.65	111.60
1	A	674	ARG	NE-CZ-NH2	-5.46	117.57	120.30
1	A	173	LEU	CB-CG-CD1	-5.17	102.22	111.00

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	7768	0	7674	142	0
1	B	7758	0	7670	140	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	54	0	42	7	0
3	B	54	0	42	8	0
4	A	47	0	0	7	0
4	B	30	0	0	2	0
All	All	15713	0	15428	286	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 9.

The worst 5 of 286 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:285:LEU:HD12	1:A:286:PRO:HD2	1.47	0.94
1:A:50:ILE:O	1:A:50:ILE:HG22	1.72	0.90
1:B:309:ASP:H	1:B:672:ASN:HD21	1.18	0.90
1:B:962:GLU:O	1:B:963:MET:HB2	1.73	0.89
1:A:300:GLN:HE21	1:A:502:GLN:HE21	1.15	0.89

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	952/990 (96%)	880 (92%)	71 (8%)	1 (0%)	56 91
1	B	950/990 (96%)	867 (91%)	78 (8%)	5 (0%)	34 77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1902/1980 (96%)	1747 (92%)	149 (8%)	6 (0%)	46 84

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	961	ARG
1	B	963	MET
1	B	353	LYS
1	B	1010	PRO
1	A	54	ILE

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	837/879 (95%)	769 (92%)	68 (8%)	15 51
1	B	837/879 (95%)	755 (90%)	82 (10%)	10 38
All	All	1674/1758 (95%)	1524 (91%)	150 (9%)	12 43

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1011	HIS
1	B	282	ASN
1	B	898	LYS
1	B	84	ASP
1	B	173	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	730	HIS
1	B	139	ASN
1	B	914	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	914	GLN
1	A	988	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 [i](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	I41	A	1102	2	21,28,28	2.02	3 (14%)	27,36,36	3.03	10 (37%)
3	I41	A	1103	-	21,28,28	2.23	5 (23%)	27,36,36	2.87	14 (51%)
3	I41	B	1102	2	21,28,28	2.40	8 (38%)	27,36,36	2.57	9 (33%)
3	I41	B	1103	-	21,28,28	2.08	3 (14%)	27,36,36	2.11	8 (29%)

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	I41	A	1102	2	-	0/24/26/26	0/2/2/2
3	I41	A	1103	-	-	0/24/26/26	0/2/2/2
3	I41	B	1102	2	-	0/24/26/26	0/2/2/2
3	I41	B	1103	-	-	0/24/26/26	0/2/2/2

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1102	I41	C23-C22	2.03	1.43	1.38
3	B	1102	I41	C21-C22	2.14	1.55	1.51
3	A	1103	I41	C26-C27	2.15	1.43	1.38
3	B	1102	I41	C21-N16	2.28	1.51	1.47
3	B	1102	I41	C15-N16	2.38	1.51	1.46

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1103	I41	C13-C15-N16	-5.17	100.14	113.20
3	A	1102	I41	C21-N16-C15	-4.96	101.55	112.19
3	A	1102	I41	C25-C24-C23	-4.91	113.00	120.19
3	B	1102	I41	C25-C24-C23	-4.56	113.51	120.19
3	A	1103	I41	O02-C03-O04	-4.09	115.34	123.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1102	I41	1	0
3	A	1103	I41	6	0
3	B	1102	I41	3	0
3	B	1103	I41	5	0

5.7 Other polymers [i](#)

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 [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, 95th 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 ²)	Q<0.9
1	A	956/990 (96%)	-0.25	0 100 100	32, 48, 65, 84	0
1	B	954/990 (96%)	-0.18	1 (0%) 95 95	37, 55, 71, 90	0
All	All	1910/1980 (96%)	-0.22	1 (0%) 95 95	32, 52, 69, 90	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	984	PRO	2.2

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, 95th 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(Å ²)	Q<0.9
3	I41	A	1102	27/27	0.82	0.35	9.97	98,107,108,108	0
3	I41	B	1103	27/27	0.76	0.37	7.80	82,95,106,106	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	I41	A	1103	27/27	0.79	0.36	6.04	85,99,109,110	0
3	I41	B	1102	27/27	0.85	0.31	3.19	98,106,108,110	0
2	ZN	B	1101	1/1	0.99	0.12	-	63,63,63,63	0
2	ZN	A	1101	1/1	0.99	0.13	-	57,57,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.