



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 11:27 am BST

PDB ID : 3ZS2  
Title : TyrB25,NMePheB26,LysB28,ProB29-insulin analogue crystal structure  
Authors : Antolikova, E.; Zakova, L.; Turkenburg, J.P.; Watson, C.J.; Hanclova, I.;  
Sanda, M.; Cooper, A.; Kraus, T.; Brzozowski, A.M.; Jiracek, J.A.  
Deposited on : 2011-06-21  
Resolution : 1.97 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

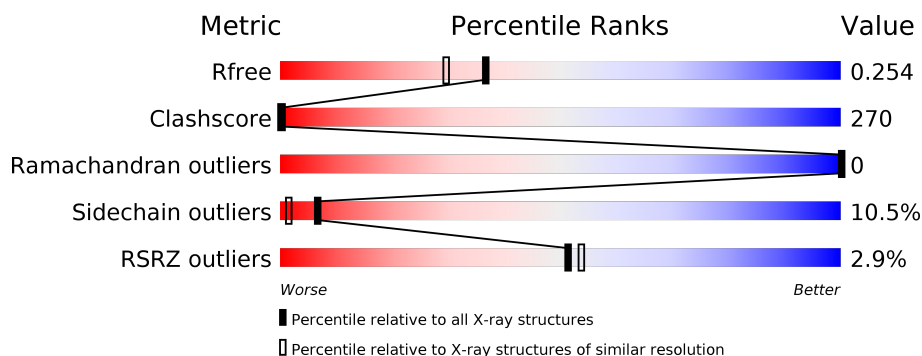
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 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.



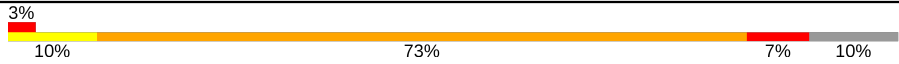

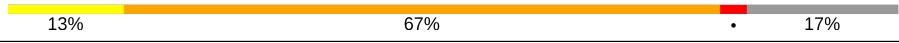
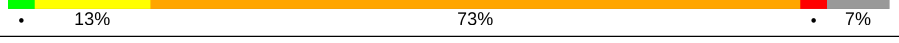


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 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	21	
1	C	21	
1	E	21	
1	G	21	
1	I	21	
1	K	21	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	30	
2	D	30	
2	F	30	
2	H	30	
2	J	30	
2	L	30	

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	MEA	B	26	-	X	-	-
2	MEA	D	26	-	X	-	-
2	MEA	F	26	-	X	X	-
2	MEA	H	26	-	X	-	-
2	MEA	J	26	-	X	-	-
2	MEA	L	26	-	X	-	-
3	IPH	A	1022	-	X	-	-
3	IPH	C	1022	-	X	-	-
3	IPH	E	1022	-	X	-	-
3	IPH	G	1022	-	X	-	-
3	IPH	I	1022	-	X	-	-
3	IPH	K	1022	-	X	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 2440 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 A CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	20	Total	C	N	O	S	0	0	0
			159	97	24	34	4			
1	C	21	Total	C	N	O	S	0	0	0
			163	99	25	35	4			
1	E	21	Total	C	N	O	S	0	0	0
			163	99	25	35	4			
1	G	20	Total	C	N	O	S	0	0	0
			159	97	24	34	4			
1	I	21	Total	C	N	O	S	0	0	0
			163	99	25	35	4			
1	K	21	Total	C	N	O	S	0	0	0
			163	99	25	35	4			

- Molecule 2 is a protein called INSULIN B CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	27	Total	C	N	O	S	0	0	0
			219	144	36	37	2			
2	D	25	Total	C	N	O	S	0	0	0
			194	125	34	33	2			
2	F	25	Total	C	N	O	S	0	0	0
			194	125	34	33	2			
2	H	28	Total	C	N	O	S	0	0	1
			220	144	37	37	2			
2	J	29	Total	C	N	O	S	0	0	0
			235	155	39	39	2			
2	L	26	Total	C	N	O	S	0	0	0
			212	140	35	35	2			

There are 24 discrepancies between the modelled and reference sequences:

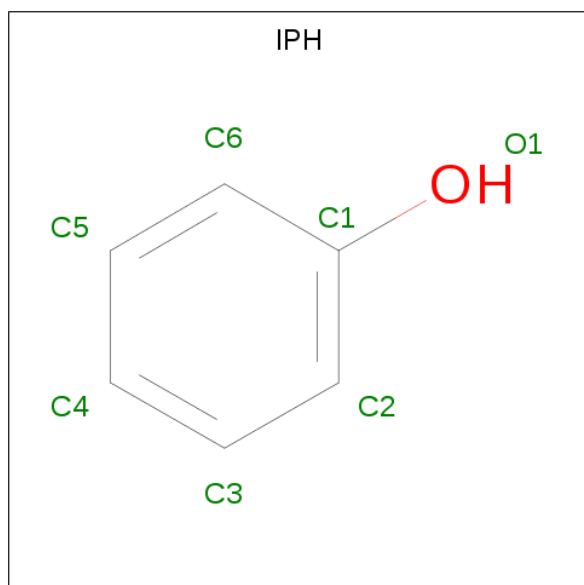
Chain	Residue	Modelled	Actual	Comment	Reference
B	25	TYR	PHE	engineered mutation	UNP P01308

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	26	MEA	TYR	engineered mutation	UNP P01308
B	28	LYS	PRO	engineered mutation	UNP P01308
B	29	PRO	LYS	engineered mutation	UNP P01308
D	25	TYR	PHE	engineered mutation	UNP P01308
D	26	MEA	TYR	engineered mutation	UNP P01308
D	28	LYS	PRO	engineered mutation	UNP P01308
D	29	PRO	LYS	engineered mutation	UNP P01308
F	25	TYR	PHE	engineered mutation	UNP P01308
F	26	MEA	TYR	engineered mutation	UNP P01308
F	28	LYS	PRO	engineered mutation	UNP P01308
F	29	PRO	LYS	engineered mutation	UNP P01308
H	25	TYR	PHE	engineered mutation	UNP P01308
H	26	MEA	TYR	engineered mutation	UNP P01308
H	28	LYS	PRO	engineered mutation	UNP P01308
H	29	PRO	LYS	engineered mutation	UNP P01308
J	25	TYR	PHE	engineered mutation	UNP P01308
J	26	MEA	TYR	engineered mutation	UNP P01308
J	28	LYS	PRO	engineered mutation	UNP P01308
J	29	PRO	LYS	engineered mutation	UNP P01308
L	25	TYR	PHE	engineered mutation	UNP P01308
L	26	MEA	TYR	engineered mutation	UNP P01308
L	28	LYS	PRO	engineered mutation	UNP P01308
L	29	PRO	LYS	engineered mutation	UNP P01308

- Molecule 3 is PHENOL (three-letter code: IPH) (formula: C<sub>6</sub>H<sub>6</sub>O).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 7 6 1	0	0
3	C	1	Total C O 7 6 1	0	0
3	E	1	Total C O 7 6 1	0	0
3	G	1	Total C O 7 6 1	0	0
3	I	1	Total C O 7 6 1	0	0
3	K	1	Total C O 7 6 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	C	1	Total Cl 1 1	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Zn 1 1	0	0
5	D	1	Total Zn 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total O 2 2	0	0
6	B	17	Total O 17 17	0	0
6	C	17	Total O 17 17	0	0
6	D	10	Total O 10 10	0	0
6	E	7	Total O 7 7	0	0

*Continued on next page...*

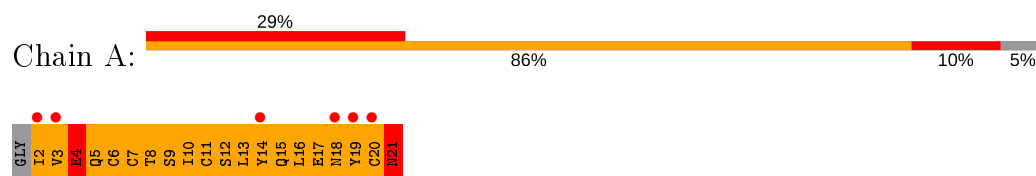
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	15	Total 15	O 15	0	0
6	G	17	Total 17	O 17	0	0
6	H	17	Total 17	O 17	0	0
6	I	19	Total 19	O 19	0	0
6	J	16	Total 16	O 16	0	0
6	K	9	Total 9	O 9	0	0
6	L	4	Total 4	O 4	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.

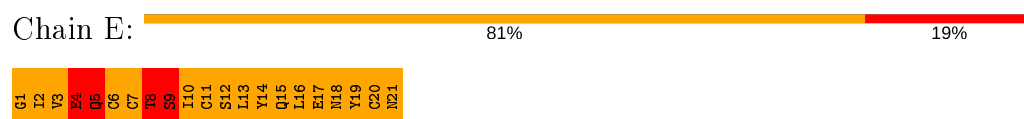
- Molecule 1: INSULIN A CHAIN



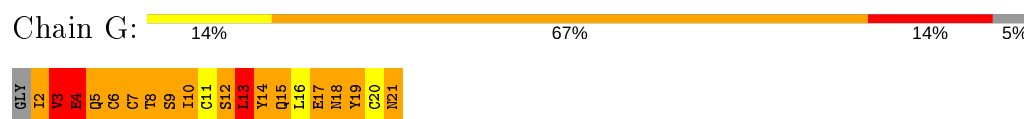
- Molecule 1: INSULIN A CHAIN



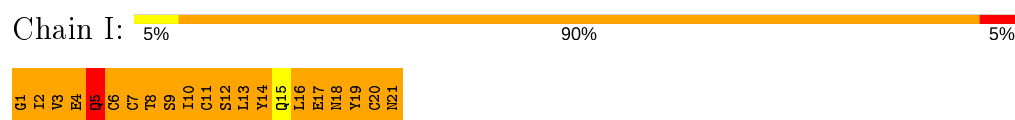
- Molecule 1: INSULIN A CHAIN



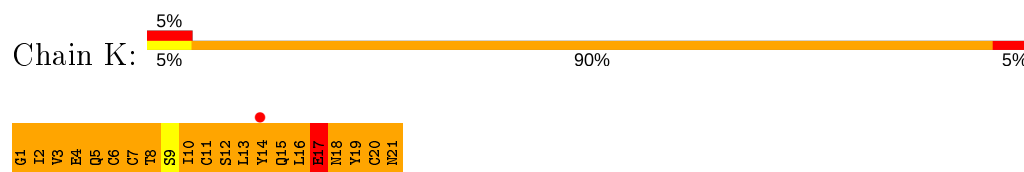
- Molecule 1: INSULIN A CHAIN



- Molecule 1: INSULIN A CHAIN



- Molecule 1: INSULIN A CHAIN

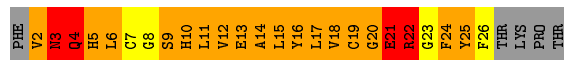
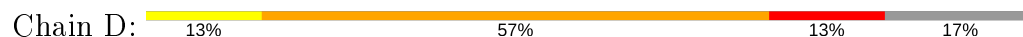




## • Molecule 2: INSULIN B CHAIN



## • Molecule 2: INSULIN B CHAIN



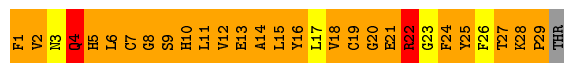
## • Molecule 2: INSULIN B CHAIN



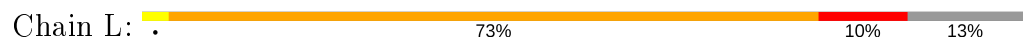
## • Molecule 2: INSULIN B CHAIN



## • Molecule 2: INSULIN B CHAIN



## • Molecule 2: INSULIN B CHAIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.66 Å   62.17 Å   46.68 Å 90.00°   111.32°   90.00°	Depositor
Resolution (Å)	53.68 – 1.97 29.03 – 1.58	Depositor EDS
% Data completeness (in resolution range)	89.9 (53.68-1.97) 29.5 (29.03-1.58)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 1.58 Å)	Xtriage
Refinement program	REFMAC 5.6.0116	Depositor
R, $R_{free}$	0.193   ,   0.252 0.201   ,   0.254	Depositor DCC
$R_{free}$ test set	617 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.8	Xtriage
Anisotropy	0.575	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33   ,   44.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.035 for -h-l,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2440	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.18% 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: ZN, MEA, IPH, CL

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	8.64	89/160 (55.6%)	6.45	74/215 (34.4%)
1	C	9.23	114/164 (69.5%)	5.50	68/220 (30.9%)
1	E	8.84	92/164 (56.1%)	5.86	98/220 (44.5%)
1	G	8.32	76/160 (47.5%)	5.40	68/215 (31.6%)
1	I	9.44	104/164 (63.4%)	5.56	77/220 (35.0%)
1	K	9.06	106/164 (64.6%)	5.30	78/220 (35.5%)
2	B	9.26	128/211 (60.7%)	6.53	115/283 (40.6%)
2	D	8.05	104/185 (56.2%)	5.47	87/249 (34.9%)
2	F	7.48	89/185 (48.1%)	5.67	111/249 (44.6%)
2	H	8.79	116/212 (54.7%)	5.73	103/285 (36.1%)
2	J	8.32	130/228 (57.0%)	5.93	102/306 (33.3%)
2	L	8.93	133/205 (64.9%)	6.40	100/276 (36.2%)
All	All	8.70	1281/2202 (58.2%)	5.85	1081/2958 (36.5%)

The worst 5 of 1281 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	7	CYS	CB-SG	29.76	2.32	1.82
1	I	12	SER	CB-OG	28.65	1.79	1.42
1	K	20	CYS	CB-SG	28.63	2.31	1.82
2	H	25	TYR	CG-CD2	25.70	1.72	1.39
1	G	6	CYS	CB-SG	25.51	2.25	1.82

The worst 5 of 1081 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	22	ARG	NE-CZ-NH1	-39.15	100.72	120.30
1	A	14	TYR	CB-CG-CD1	-33.41	100.95	121.00
2	J	16	TYR	CB-CG-CD1	-29.25	103.45	121.00
2	J	22	ARG	NE-CZ-NH1	29.24	134.92	120.30
2	B	24	PHE	CB-CG-CD2	26.18	139.13	120.80

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	159	0	142	88	0
1	C	163	0	149	108	0
1	E	163	0	148	131	0
1	G	159	0	143	73	0
1	I	163	0	149	112	0
1	K	163	0	146	122	0
2	B	219	0	205	146	0
2	D	194	0	178	94	0
2	F	194	0	180	63	0
2	H	220	0	206	96	0
2	J	235	0	223	132	0
2	L	212	0	195	117	0
3	A	7	0	5	2	0
3	C	7	0	6	0	0
3	E	7	0	5	0	0
3	G	7	0	6	2	0
3	I	7	0	5	1	0
3	K	7	0	5	2	0
4	B	1	0	0	0	0
4	C	1	0	0	1	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	A	2	0	0	0	0
6	B	17	0	0	5	0
6	C	17	0	0	3	0
6	D	10	0	0	6	0
6	E	7	0	0	1	0
6	F	15	0	0	5	0
6	G	17	0	0	6	0
6	H	17	0	0	3	0
6	I	19	0	0	6	0
6	J	16	0	0	1	0
6	K	9	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	L	4	0	0	0	0
All	All	2440	0	2096	1179	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 270.

The worst 5 of 1179 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:H:10:HIS:CB	2:H:10:HIS:CG	1.75	1.70
1:I:14:TYR:CE1	1:I:14:TYR:CD1	1.75	1.70
1:I:14:TYR:CD2	1:I:14:TYR:CE2	1.75	1.68
2:B:24:PHE:CB	2:B:24:PHE:CG	1.77	1.68
2:H:24:PHE:CD1	2:H:24:PHE:CE1	1.75	1.65

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	18/21 (86%)	17 (94%)	1 (6%)	0	100	100
1	C	19/21 (90%)	17 (90%)	2 (10%)	0	100	100
1	E	19/21 (90%)	17 (90%)	2 (10%)	0	100	100
1	G	18/21 (86%)	17 (94%)	1 (6%)	0	100	100
1	I	19/21 (90%)	19 (100%)	0	0	100	100
1	K	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
2	B	24/30 (80%)	24 (100%)	0	0	100	100
2	D	23/30 (77%)	22 (96%)	1 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	23/30 (77%)	22 (96%)	1 (4%)	0	100	100
2	H	25/30 (83%)	24 (96%)	1 (4%)	0	100	100
2	J	26/30 (87%)	24 (92%)	2 (8%)	0	100	100
2	L	24/30 (80%)	23 (96%)	1 (4%)	0	100	100
All	All	257/306 (84%)	244 (95%)	13 (5%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	20/20 (100%)	18 (90%)	2 (10%)	7	1
1	C	20/20 (100%)	19 (95%)	1 (5%)	24	12
1	E	20/20 (100%)	16 (80%)	4 (20%)	1	0
1	G	20/20 (100%)	17 (85%)	3 (15%)	3	0
1	I	20/20 (100%)	19 (95%)	1 (5%)	24	12
1	K	20/20 (100%)	19 (95%)	1 (5%)	24	12
2	B	22/25 (88%)	20 (91%)	2 (9%)	9	2
2	D	19/25 (76%)	15 (79%)	4 (21%)	1	0
2	F	19/25 (76%)	18 (95%)	1 (5%)	22	10
2	H	22/25 (88%)	20 (91%)	2 (9%)	9	2
2	J	24/25 (96%)	22 (92%)	2 (8%)	11	3
2	L	21/25 (84%)	18 (86%)	3 (14%)	3	0
All	All	247/270 (92%)	221 (90%)	26 (10%)	7	1

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	8	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	3	VAL
2	L	17	LEU
1	E	9	SER
2	F	3	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	5	GLN
1	E	21	ASN
1	I	5	GLN
2	D	4	GLN
1	I	21	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues 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	MEA	D	26	2	11,12,13	9.50	8 (72%)	13,14,16	7.91	12 (92%)
2	MEA	B	26	2	11,12,13	6.38	7 (63%)	13,14,16	5.40	10 (76%)
2	MEA	L	26	2	11,12,13	5.51	10 (90%)	13,14,16	5.77	9 (69%)
2	MEA	J	26	2	11,12,13	6.48	5 (45%)	13,14,16	4.52	10 (76%)
2	MEA	H	26	2	11,12,13	5.92	6 (54%)	13,14,16	4.54	11 (84%)
2	MEA	F	26	2	11,12,13	8.83	8 (72%)	13,14,16	7.08	11 (84%)

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	MEA	D	26	2	-	2/5/8/10	0/1/1/1
2	MEA	B	26	2	-	1/5/8/10	0/1/1/1
2	MEA	L	26	2	-	1/5/8/10	0/1/1/1
2	MEA	J	26	2	-	1/5/8/10	0/1/1/1
2	MEA	H	26	2	-	3/5/8/10	0/1/1/1
2	MEA	F	26	2	-	2/5/8/10	0/1/1/1

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	26	MEA	CA-N	-17.27	1.16	1.47
2	J	26	MEA	CB-CA	16.48	1.93	1.54
2	D	26	MEA	CA-N	-16.42	1.18	1.47
2	D	26	MEA	CD1-CG	15.23	1.71	1.38
2	D	26	MEA	CE2-CZ	13.11	1.72	1.38

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	26	MEA	CZ-CE1-CD1	-14.61	97.94	120.19
2	D	26	MEA	CE2-CD2-CG	-14.22	98.82	120.63
2	B	26	MEA	CB-CA-C	-13.46	86.14	111.65
2	F	26	MEA	CZ-CE1-CD1	-13.23	100.04	120.19
2	F	26	MEA	CE2-CD2-CG	-12.60	101.30	120.63

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	26	MEA	O-C-CA-CB
2	D	26	MEA	C-CA-CB-CG
2	B	26	MEA	O-C-CA-CB
2	L	26	MEA	O-C-CA-CB
2	J	26	MEA	O-C-CA-CB

There are no ring outliers.

6 monomers are involved in 24 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	26	MEA	4	0
2	B	26	MEA	3	0
2	L	26	MEA	3	0
2	J	26	MEA	5	0
2	H	26	MEA	2	0
2	F	26	MEA	7	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	IPH	I	1022	-	7,7,7	2.39	4 (57%)	8,8,8	3.00	5 (62%)
3	IPH	K	1022	-	7,7,7	6.55	5 (71%)	8,8,8	4.75	7 (87%)
3	IPH	E	1022	-	7,7,7	2.29	2 (28%)	8,8,8	2.54	6 (75%)
3	IPH	G	1022	-	7,7,7	7.83	7 (100%)	8,8,8	7.85	7 (87%)
3	IPH	A	1022	-	7,7,7	8.45	5 (71%)	8,8,8	7.39	8 (100%)
3	IPH	C	1022	-	7,7,7	4.67	4 (57%)	8,8,8	3.56	6 (75%)

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	IPH	I	1022	-	-	-	0/1/1/1
3	IPH	K	1022	-	-	-	0/1/1/1
3	IPH	E	1022	-	-	-	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	IPH	G	1022	-	-	-	0/1/1/1
3	IPH	A	1022	-	-	-	0/1/1/1
3	IPH	C	1022	-	-	-	0/1/1/1

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1022	IPH	C6-C1	17.90	1.73	1.38
3	G	1022	IPH	C6-C1	13.40	1.64	1.38
3	A	1022	IPH	C4-C3	12.85	1.72	1.38
3	G	1022	IPH	C4-C3	10.26	1.65	1.38
3	K	1022	IPH	C3-C2	-9.56	1.18	1.38

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1022	IPH	C3-C2-C1	-11.93	102.45	119.31
3	G	1022	IPH	C3-C2-C1	-11.76	102.69	119.31
3	A	1022	IPH	C4-C5-C6	-11.24	103.07	120.19
3	G	1022	IPH	C4-C5-C6	-10.71	103.88	120.19
3	G	1022	IPH	C5-C6-C1	10.29	133.85	119.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	1022	IPH	1	0
3	K	1022	IPH	2	0
3	G	1022	IPH	2	0
3	A	1022	IPH	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	L	11
2	J	10
2	H	9
2	B	8
1	I	5
1	A	5
2	D	4
1	K	4
2	F	4
1	G	3
1	E	3
1	C	3

The worst 5 of 69 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	25:TYR	C	26:MEA	N	1.68
1	J	12:VAL	C	13:GLU	N	1.67
1	A	11:CYS	C	12:SER	N	1.66
1	L	12:VAL	C	13:GLU	N	1.66
1	F	7:CYS	C	8:GLY	N	1.65

## 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	20/21 (95%)	0.92	6 (30%) 0 0	39, 54, 68, 70	0
1	C	21/21 (100%)	-0.59	0 100 100	27, 36, 48, 55	0
1	E	21/21 (100%)	-0.51	0 100 100	28, 39, 57, 73	0
1	G	20/21 (95%)	-0.65	0 100 100	31, 36, 57, 66	0
1	I	21/21 (100%)	-0.81	0 100 100	27, 32, 44, 59	0
1	K	21/21 (100%)	-0.52	1 (4%) 30 32	27, 37, 53, 59	0
2	B	26/30 (86%)	-0.15	1 (3%) 40 43	24, 30, 69, 94	0
2	D	24/30 (80%)	-0.12	0 100 100	21, 30, 59, 64	0
2	F	24/30 (80%)	-0.25	0 100 100	24, 30, 52, 66	0
2	H	27/30 (90%)	-0.12	0 100 100	23, 30, 62, 76	0
2	J	28/30 (93%)	-0.25	0 100 100	24, 29, 55, 73	0
2	L	25/30 (83%)	0.00	0 100 100	23, 32, 55, 67	0
All	All	278/306 (90%)	-0.25	8 (2%) 51 54	21, 35, 65, 94	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	18	ASN	3.5
1	A	14	TYR	2.8
1	A	19	TYR	2.5
1	A	2	ILE	2.3
1	A	3	VAL	2.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	MEA	L	26	12/13	0.72	0.20	58,69,80,82	0
2	MEA	B	26	12/13	0.82	0.21	83,90,94,95	0
2	MEA	F	26	12/13	0.87	0.12	44,49,53,53	0
2	MEA	D	26	12/13	0.90	0.09	53,65,71,73	0
2	MEA	H	26	12/13	0.94	0.07	36,41,45,46	0
2	MEA	J	26	12/13	0.97	0.09	28,31,37,39	0

### 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(Å <sup>2</sup> )	Q<0.9
3	IPH	E	1022	7/7	0.96	0.08	32,33,35,35	0
3	IPH	A	1022	7/7	0.96	0.07	29,31,36,36	0
3	IPH	G	1022	7/7	0.97	0.13	25,27,29,32	0
3	IPH	C	1022	7/7	0.97	0.09	27,32,33,34	0
3	IPH	K	1022	7/7	0.98	0.11	28,28,31,32	0
3	IPH	I	1022	7/7	0.99	0.06	25,26,28,30	0
4	CL	B	1028	1/1	1.00	0.11	27,27,27,27	0
4	CL	C	1028	1/1	1.00	0.10	25,25,25,25	0
5	ZN	B	1030	1/1	1.00	0.08	25,25,25,25	0
5	ZN	D	1030	1/1	1.00	0.06	23,23,23,23	0

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

There are no such residues in this entry.