wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 11:21 PM EDT

PDB ID	:	7N6H
Title	:	The crystal structure of the GH30 subfamily 10 enzyme, AcXbh30A from Acetivibrio clariflavus
Authors	:	Tan, K.; St John, F.J.
Deposited on	:	2021-06-08
Resolution	:	1.28 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

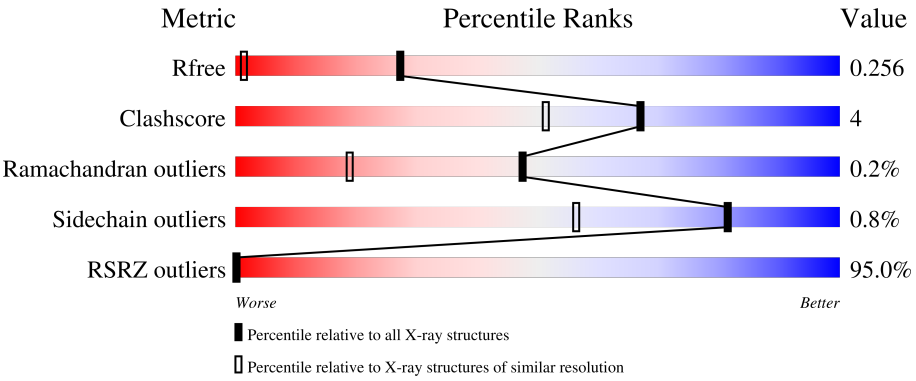
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.28 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	1850 (1.30-1.26)
Clashscore	141614	1926 (1.30-1.26)
Ramachandran outliers	138981	1860 (1.30-1.26)
Sidechain outliers	138945	1859 (1.30-1.26)
RSRZ outliers	127900	1807 (1.30-1.26)

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 >=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 <=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	444	92% 89% 9%
1	B	444	93% 88% 9%

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	CL	A	504	-	-	-	X
2	CL	A	506	-	-	-	X
2	CL	A	508	-	-	X	-
2	CL	A	511	-	-	-	X
2	CL	A	515	-	-	X	-
2	CL	B	503	-	-	X	-
2	CL	B	507	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7649 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 AcXbh30A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	433	Total	C	N	O	S	0	3	0
			3454	2194	570	679	11			
1	B	433	Total	C	N	O	S	0	1	0
			3443	2187	569	676	11			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	MET	-	initiating methionine	UNP G8LU16
A	464	LEU	-	expression tag	UNP G8LU16
A	465	GLU	-	expression tag	UNP G8LU16
A	466	HIS	-	expression tag	UNP G8LU16
A	467	HIS	-	expression tag	UNP G8LU16
A	468	HIS	-	expression tag	UNP G8LU16
A	469	HIS	-	expression tag	UNP G8LU16
A	470	HIS	-	expression tag	UNP G8LU16
A	471	HIS	-	expression tag	UNP G8LU16
B	28	MET	-	initiating methionine	UNP G8LU16
B	464	LEU	-	expression tag	UNP G8LU16
B	465	GLU	-	expression tag	UNP G8LU16
B	466	HIS	-	expression tag	UNP G8LU16
B	467	HIS	-	expression tag	UNP G8LU16
B	468	HIS	-	expression tag	UNP G8LU16
B	469	HIS	-	expression tag	UNP G8LU16
B	470	HIS	-	expression tag	UNP G8LU16
B	471	HIS	-	expression tag	UNP G8LU16

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

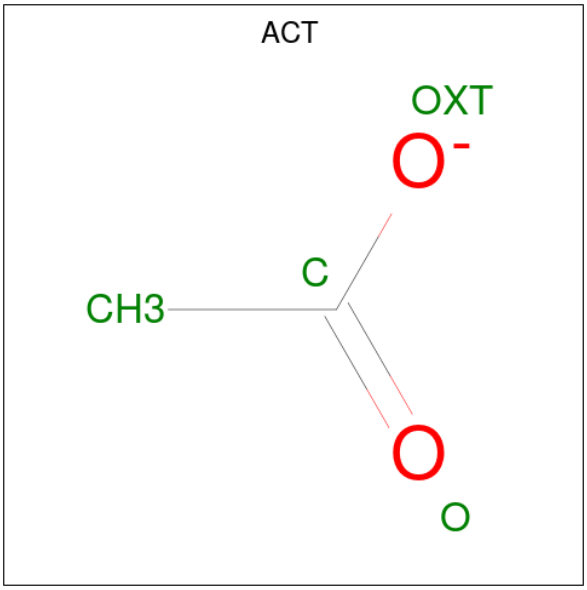
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	15	Total	Cl	0	0
			15	15		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	8	Total Cl 8 8	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0

- Molecule 4 is water.

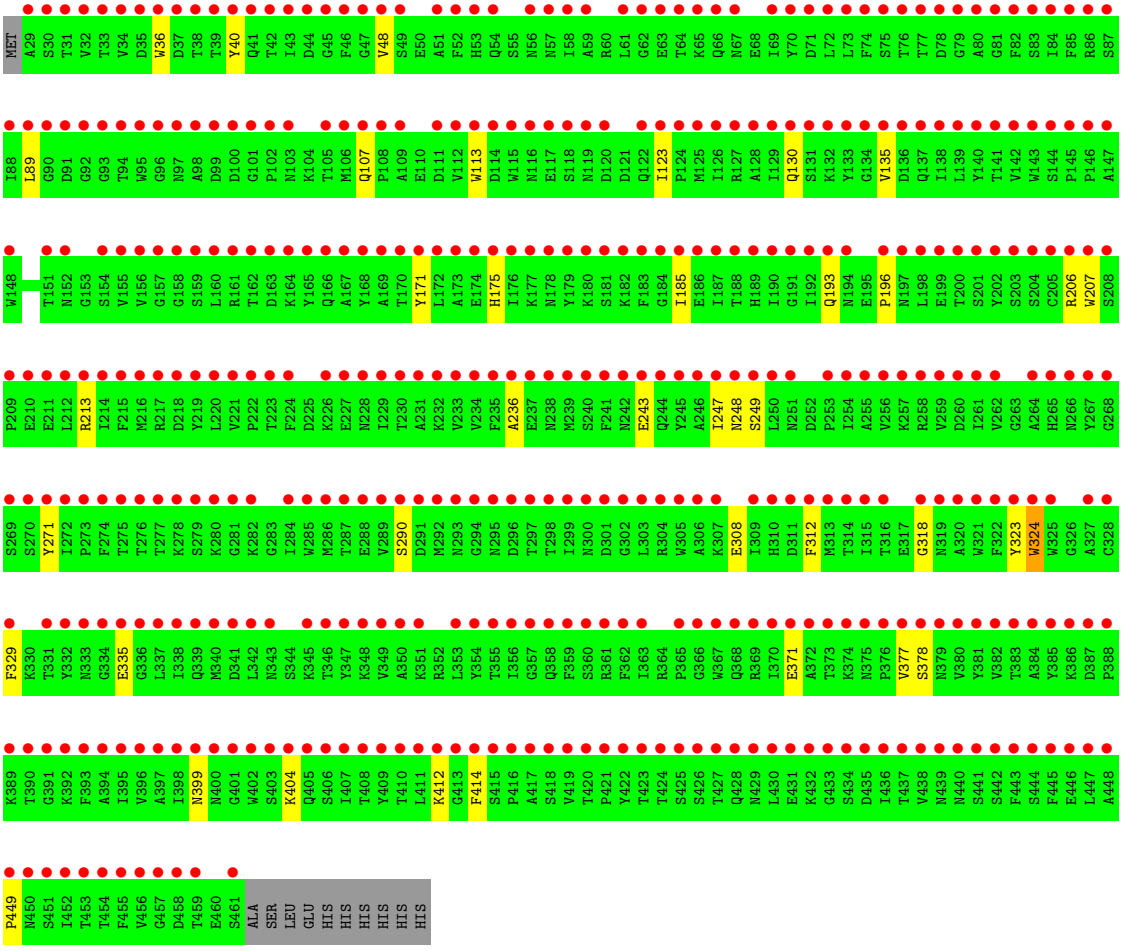
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	388	Total O 388 388	0	0
4	B	329	Total O 329 329	0	0

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.

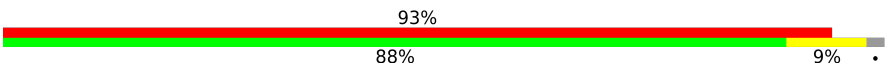
• Molecule 1: AcXbh30A

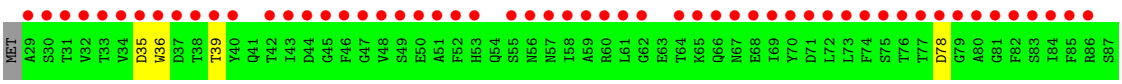
Chain A:

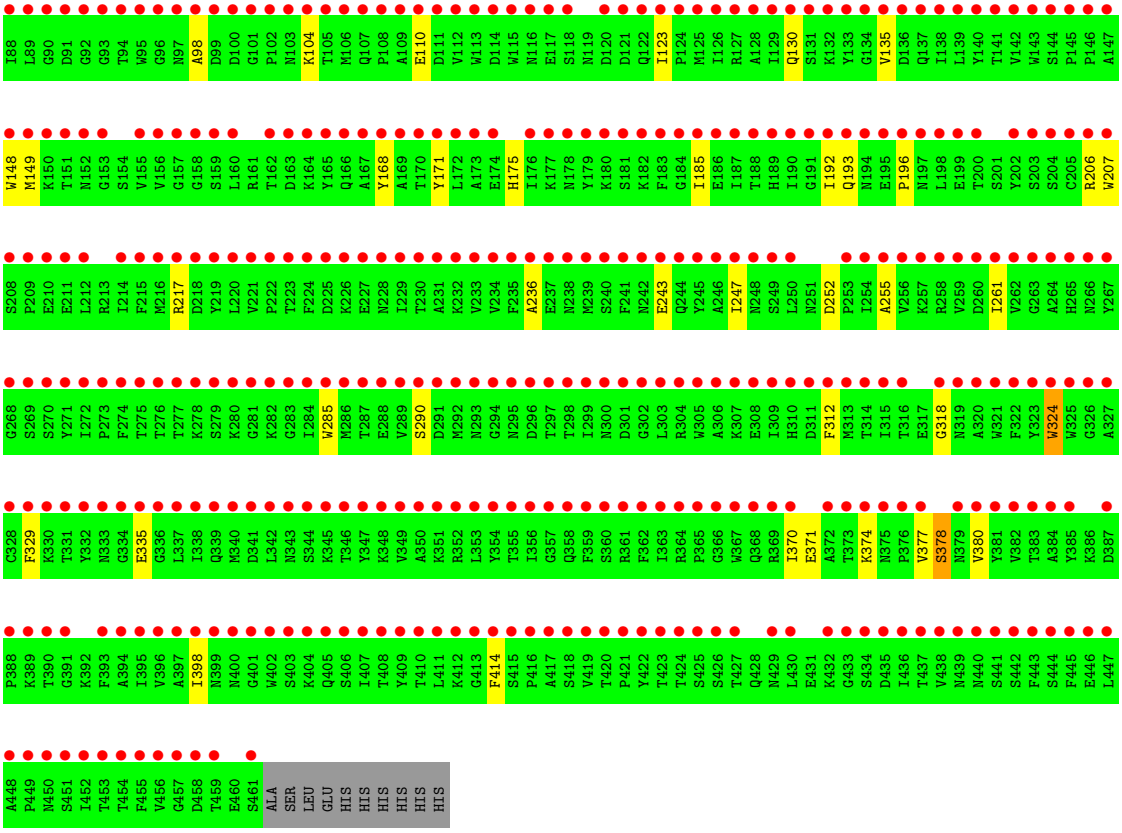


• Molecule 1: AcXbh30A

Chain B:






I88 L89 G90 D91 G92 G93 T94 W95 G96 N97 A98 D99 D100 G101 P102 N103 K104 T105 M106 Q107 P108 A109 E110 D111 V112 W113 D114 W115 N116 E117 S118 N119 D120 D121 Q122 I123 P124 M125 I126 R127 A128 I129 Q130 S131 K132 Y133 G134 V135 D136 Q137 I138 L139 Y140 T141 V142 W143 S144 P145 P146 A147
W148 M149 K150 T151 N152 G153 S154 V155 V156 G157 G158 S159 L160 R161 T162 D163 K164 Y165 Q166 A167 Y168 A169 T170 Y171 L172 A173 E174 H175 I176 K177 N178 Y179 K180 S181 K182 F183 G184 I185 E186 I187 T188 H189 I190 G191 I192 Q193 N194 E195 P196 N197 L198 E199 T200 S201 Y202 S203 S204 C205 R206 W207
S208 P209 E210 E211 L212 R213 I214 F215 M216 R217 D218 Y219 L220 V221 P222 T223 F224 D225 K226 E227 N228 I229 T230 A231 K232 V233 V234 F235 A236 E237 N238 M239 S240 F241 N242 E243 Q244 Y245 A246 I247 N248 S249 L250 N251 D252 P253 I254 A255 V256 K257 R258 V259 D260 I261 V262 G263 A264 H265 N266 Y267
G268 S269 S270 Y271 I272 P273 F274 T275 T276 T277 K278 S279 K280 G281 K282 G283 I284 W285 M286 T287 E288 V289 S290 D291 M292 N293 G294 N295 D296 T297 T298 I299 N300 D301 G302 L303 R304 W305 A306 K307 E308 I309 H310 D311 F312 N313 T314 I315 T316 E317 G318 N319 A320 W321 F322 Y323 W324 W325 G326 A327
C328 F329 K330 T331 Y332 N333 G334 E335 G336 L337 I338 Q339 M340 D341 L342 N343 S344 K345 T346 Y347 K348 V349 A350 K351 R352 L353 Y354 T355 I356 G357 Q358 F359 S360 R361 F362 I363 R364 P365 G366 W367 Q368 R369 I370 E371 A372 T373 K374 N375 P376 V377 S378 N379 V380 Y381 V382 T383 A384 Y385 K386 D387
P388 K389 T390 G391 K392 F393 A394 I395 V396 A397 I398 N399 N400 G401 W402 S403 K404 Q405 S406 I407 T408 Y409 T410 L411 K412 G413 F414 S415 P416 A417 S418 V419 T420 P421 Y422 T423 T424 S425 S426 T427 Q428 N429 L430 E431 K432 G433 S434 D435 T436 T437 V438 N439 N440 S441 S442 F443 S444 F445 E446 L447
A448 P449 N450 S451 I452 T453 T454 F455 V456 G457 D458 T459 E460 S461 ALA SER LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.87Å 108.02Å 66.84Å 90.00° 90.14° 90.00°	Depositor
Resolution (Å)	41.18 – 1.28 41.18 – 1.28	Depositor EDS
% Data completeness (in resolution range)	96.3 (41.18-1.28) 96.2 (41.18-1.28)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	13.51 (at 1.28Å)	Xtriage
Refinement program	PHENIX (1.19_4092: ???)	Depositor
R, R_{free}	0.202 , 0.254 0.203 , 0.256	Depositor DCC
R_{free} test set	10529 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	13.1	Xtriage
Anisotropy	0.270	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 26.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.387 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7649	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: ACT, 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	0.24	0/3544	0.47	0/4821
1	B	0.24	0/3533	0.46	0/4806
All	All	0.24	0/7077	0.47	0/9627

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	3454	0	3292	24	0
1	B	3443	0	3281	24	0
2	A	15	0	0	8	0
2	B	8	0	0	5	0
3	A	8	0	6	0	0
3	B	4	0	3	0	0
4	A	388	0	0	4	0
4	B	329	0	0	3	0
All	All	7649	0	6582	53	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 4.

The worst 5 of 53 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:A:510:CL:CL	4:A:896:HOH:O	2.38	0.78
1:A:213:ARG:NE	2:A:503:CL:CL	2.50	0.73
2:A:515:CL:CL	4:A:860:HOH:O	2.43	0.73
2:A:514:CL:CL	4:A:952:HOH:O	2.53	0.61
1:A:404:LYS:HG3	1:A:449:PRO:HD3	1.85	0.59

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	433/444 (98%)	418 (96%)	14 (3%)	1 (0%)	47	19
1	B	432/444 (97%)	418 (97%)	13 (3%)	1 (0%)	47	19
All	All	865/888 (97%)	836 (97%)	27 (3%)	2 (0%)	47	19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	324	TRP
1	B	324	TRP

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	376/384 (98%)	373 (99%)	3 (1%)	81	56
1	B	374/384 (97%)	371 (99%)	3 (1%)	81	56
All	All	750/768 (98%)	744 (99%)	6 (1%)	81	56

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	324	TRP
1	B	329	PHE
1	B	378	SER
1	A	324	TRP
1	A	206	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 26 ligands modelled in this entry, 23 are monoatomic - leaving 3 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	ACT	B	509	-	3,3,3	1.28	0	3,3,3	1.27	0
3	ACT	A	516	-	3,3,3	1.27	0	3,3,3	1.41	0
3	ACT	A	517	-	3,3,3	1.25	0	3,3,3	1.32	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

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

6.1 Protein, DNA and RNA chains

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^{th} 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	433/444 (97%)	4.62	410 (94%)	0	0	10, 15, 21, 38	0
1	B	433/444 (97%)	4.76	413 (95%)	0	0	9, 14, 21, 34	0
All	All	866/888 (97%)	4.69	823 (95%)	0	0	9, 15, 21, 38	0

The worst 5 of 823 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	29	ALA	36.8
1	B	306	ALA	14.5
1	B	146	PRO	12.5
1	B	246	ALA	11.7
1	B	329	PHE	11.5

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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^2)	Q<0.9
3	ACT	A	517	4/4	0.30	0.31	21,21,22,22	0
3	ACT	A	516	4/4	0.33	0.29	17,18,20,21	0
2	CL	A	511	1/1	0.37	0.49	33,33,33,33	0
2	CL	A	506	1/1	0.39	0.45	45,45,45,45	0
2	CL	A	501	1/1	0.46	0.29	39,39,39,39	0
2	CL	A	509	1/1	0.48	0.38	68,68,68,68	0
2	CL	A	502	1/1	0.49	0.15	31,31,31,31	0
2	CL	A	513	1/1	0.49	0.22	49,49,49,49	0
2	CL	A	503	1/1	0.51	0.20	52,52,52,52	0
3	ACT	B	509	4/4	0.61	0.31	24,24,26,27	0
2	CL	A	510	1/1	0.64	0.20	50,50,50,50	0
2	CL	B	507	1/1	0.66	0.54	79,79,79,79	0
2	CL	B	506	1/1	0.66	0.20	27,27,27,27	0
2	CL	A	507	1/1	0.67	0.33	23,23,23,23	0
2	CL	B	501	1/1	0.72	0.25	33,33,33,33	0
2	CL	B	505	1/1	0.73	0.25	38,38,38,38	0
2	CL	A	505	1/1	0.73	0.11	29,29,29,29	0
2	CL	A	514	1/1	0.74	0.21	45,45,45,45	0
2	CL	B	502	1/1	0.74	0.10	23,23,23,23	0
2	CL	A	512	1/1	0.75	0.22	40,40,40,40	0
2	CL	A	504	1/1	0.76	0.46	48,48,48,48	0
2	CL	A	515	1/1	0.78	0.21	26,26,26,26	0
2	CL	A	508	1/1	0.81	0.17	22,22,22,22	0
2	CL	B	504	1/1	0.82	0.24	39,39,39,39	0
2	CL	B	508	1/1	0.88	0.16	45,45,45,45	0
2	CL	B	503	1/1	0.91	0.06	24,24,24,24	0

6.5 Other polymers (i)

There are no such residues in this entry.