wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2023 – 07:56 am GMT

PDB ID	:	2IU3
Title	:	Crystal structures of transition state analogue inhibitors of inosine monophosphate cyclohydrolase
Authors	:	Xu, L.; Chong, Y.; Hwang, I.; D'Onofrio, A.; Amore, K.; Beardsley, G.P.; Li, C.; Olson, A.J.; Boger, D.L.; Wilson, I.A.
Deposited on	:	2006-05-27
Resolution	:	2.90 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 (i)) were used in the production of this report:

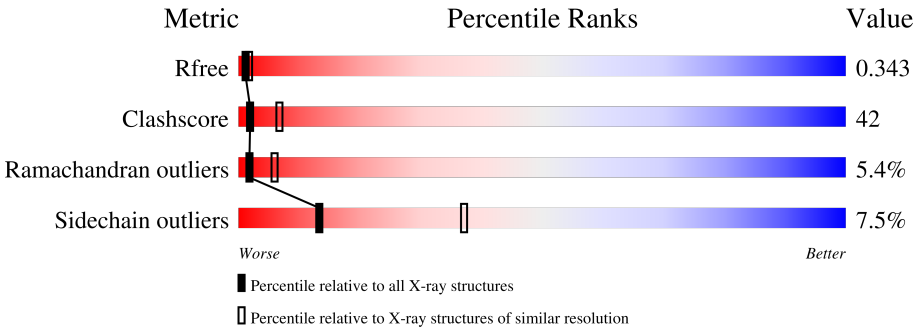
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)

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%

Mol	Chain	Length	Quality of chain
1	A	593	36% 55% 8%
1	B	593	39% 53% 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9145 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.

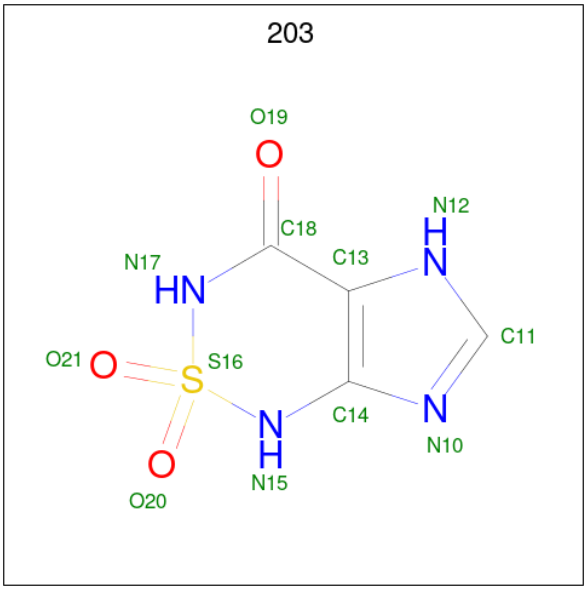
- Molecule 1 is a protein called BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	590	Total	C	N	O	S	0	1	0
			4523	2852	801	851	19			
1	B	590	Total	C	N	O	S	0	0	0
			4511	2843	800	849	19			

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

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

- Molecule 3 is 1,5-DIHYDROIMIDAZO[4,5-C][1,2,6]THIADIAZIN-4(3H)-ONE 2,2-DIOXIDE (three-letter code: 203) (formula: $C_4H_4N_4O_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	S	0	0
			12	4	4	3	1		

- Molecule 4 is water.

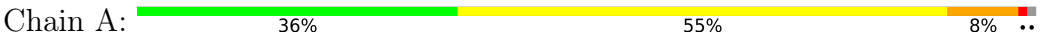
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	52	Total	O	0	0
			52	52		
4	B	45	Total	O	0	0
			45	45		

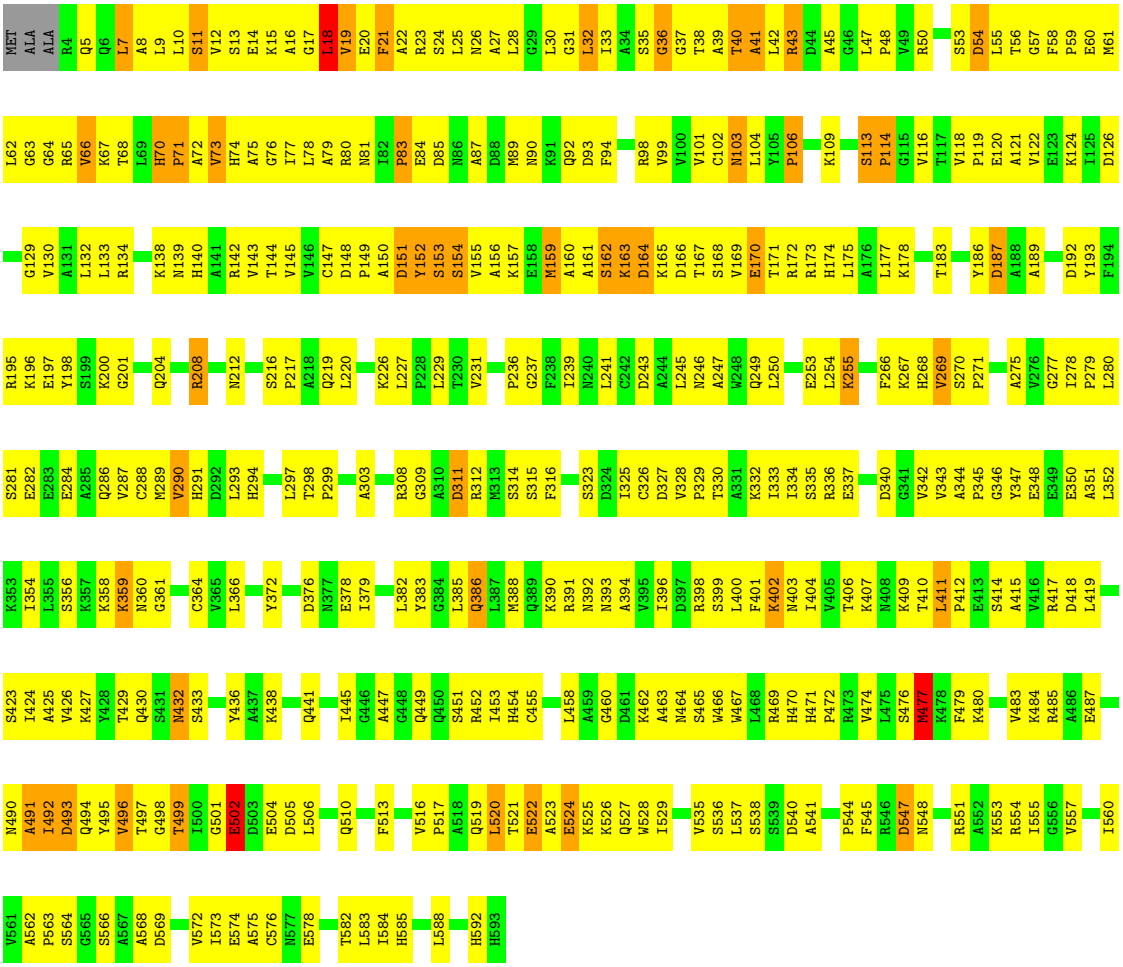
3 Residue-property plots (i)

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. 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. 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: BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH

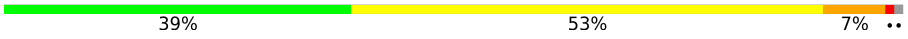
Chain A:

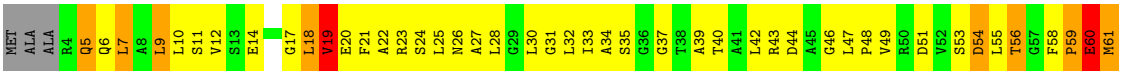


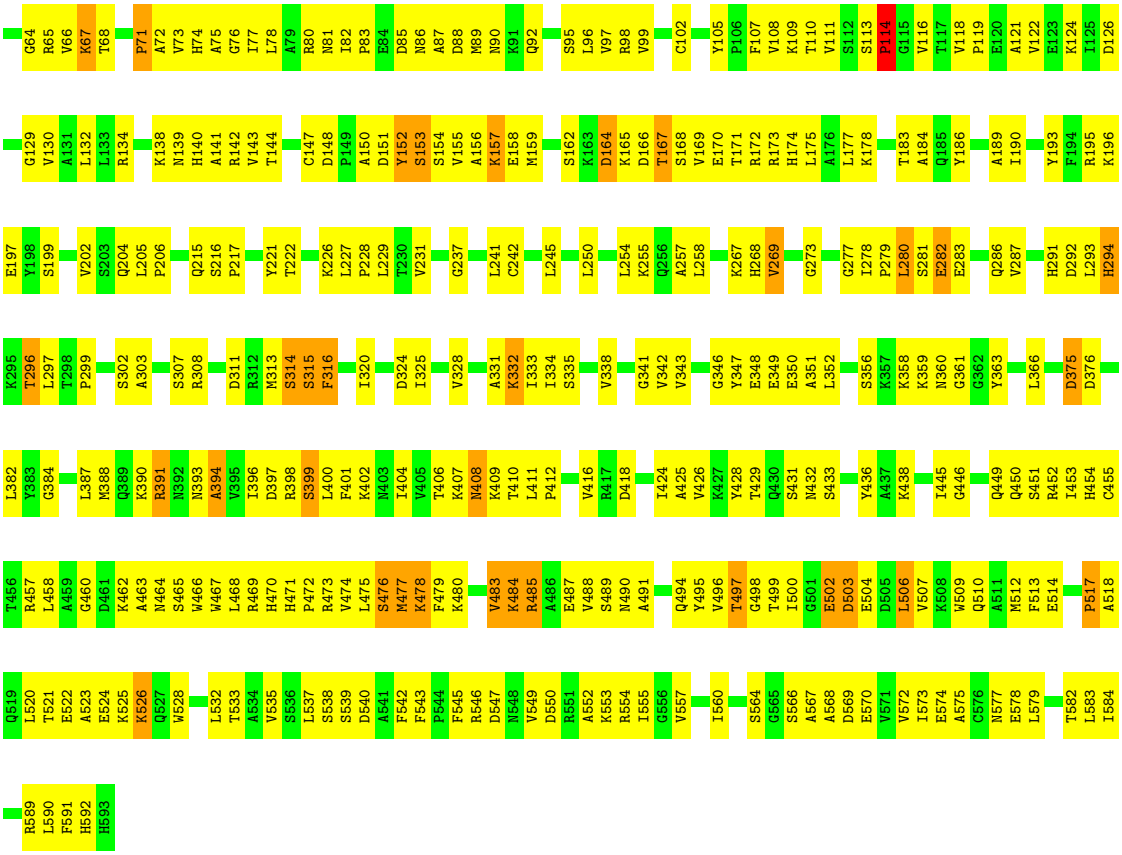


• Molecule 1: BIFUNCTIONAL PURINE BIOSYNTHESIS PROTEIN PURH

Chain B:






G64 R65 V66 K67 T68 P71 A72 V73 H74 A75 G76 I77 L78 A79 R80 N81 I82 P83 E84 D85 N86 A87 D88 M89 N90 K91 Q92 S95 L96 V97 R98 V99 C102 Y105 P106 F107 V108 K109 T110 V111 S112 S113 R114 G115 V116 T117 V118 P119 E120 A121 V122 E123 K124 I125 D126
G129 V130 A131 L132 L133 R134 K138 N139 H140 A141 R142 V143 T144 C147 D148 P149 A150 D151 Y152 S153 S154 V155 A156 K157 E158 M159 S162 K163 D164 K165 D166 T167 S168 V169 E170 T171 R172 R173 H174 L175 A176 L177 K178 T183 A184 Q185 Y186 A189 I190 Y193 F194 R195 K196
E197 Y198 S199 V202 S203 Q204 L205 P206 Q215 S216 P217 Y221 T222 K226 L227 P228 L229 T230 V231 G237 L241 C242 L245 L250 L254 K255 Q256 A257 L258 K267 H268 V269 G273 G277 I278 P279 L280 S281 E282 E283 Q286 V287 H291 D292 L293 H294
K295 T296 L297 T298 P299 S302 A303 S307 R308 D311 R312 M313 S314 S315 F316 I320 D324 I325 V328 A331 K332 I333 I334 S335 V338 G341 V342 V343 G346 Y347 E348 E349 E350 A351 L352 S356 K357 K358 K359 N360 G361 G362 Y363 L366 D375 D376
L382 Y383 G384 L387 M388 Q389 K390 R391 N392 N393 A394 V395 I396 D397 R398 S399 L400 F401 K402 M403 I404 V405 T406 K407 M408 K409 T410 L411 P412 V416 R417 D418 I424 A425 V426 K427 Y428 T429 Q430 S431 N432 S433 Y436 A437 K438 I445 G446 Q449 Q450 S451 R452 I453 H454 C455
T456 R457 L458 A459 G460 D461 K462 A463 N464 S465 W466 W467 L468 R469 H470 H471 P472 R473 V474 L475 S476 M477 K478 F479 K480 V483 K484 R485 A486 E487 V488 S489 N490 A491 Q494 Y495 V496 T497 G498 T499 I500 G501 E502 D503 E504 D505 L506 V507 K508 W509 Q510 A511 M512 F513 E514 P517 A518
Q519 L520 T521 E522 A523 E524 K525 K526 Q527 W528 L532 T533 A534 V535 S536 L537 S538 S539 D540 A541 F542 F543 P544 F545 R546 D547 N548 V549 D550 R551 A552 K553 R554 I555 G556 V557 I560 S564 G565 S566 A567 A568 D569 E570 V571 V572 I573 E574 A575 C576 N577 E578 L579 T582 L583 I584
R589 L590 F591 H592 H593

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	387.00Å 57.00Å 62.10Å 90.00° 98.90° 90.00°	Depositor
Resolution (Å)	35.20 – 2.90 35.13 – 2.87	Depositor EDS
% Data completeness (in resolution range)	91.5 (35.20-2.90) 90.9 (35.13-2.87)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 2.85Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.215 , 0.300 0.286 , 0.343	Depositor DCC
R_{free} test set	1210 reflections (3.99%)	wwPDB-VP
Wilson B-factor (Å^2)	50.6	Xtriage
Anisotropy	0.315	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 30.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.027 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	9145	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% 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: 203, K

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	0.61	0/4608	0.75	1/6249 (0.0%)
1	B	0.59	0/4595	0.75	0/6230
All	All	0.60	0/9203	0.75	1/12479 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	45	ALA	N-CA-C	-5.98	94.86	111.00

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	4523	0	4570	434	0
1	B	4511	0	4561	363	4
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	B	12	0	4	1	0
4	A	52	0	0	2	0
4	B	45	0	0	3	0
All	All	9145	0	9135	763	4

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 42.

The worst 5 of 763 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:510:GLN:HE21	1:A:516:VAL:HG21	1.19	1.04
1:B:485:ARG:HG2	1:B:485:ARG:HH11	1.22	1.04
1:A:379:ILE:HG12	1:A:388:MET:HG3	1.34	1.04
1:A:510:GLN:NE2	1:A:516:VAL:HG21	1.76	1.00
1:B:479:PHE:HA	1:B:513:PHE:HA	1.46	0.97

All (4) 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)
1:B:47:LEU:O	1:B:47:LEU:O[2_656]	1.94	0.26
1:B:46:GLY:O	1:B:48:PRO:C[2_656]	2.02	0.18
1:B:46:GLY:CA	1:B:49:VAL:O[2_656]	2.07	0.13
1:B:46:GLY:O	1:B:48:PRO:CA[2_656]	2.11	0.09

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	589/593 (99%)	464 (79%)	90 (15%)	35 (6%)	1 5
1	B	588/593 (99%)	484 (82%)	75 (13%)	29 (5%)	2 8
All	All	1177/1186 (99%)	948 (80%)	165 (14%)	64 (5%)	2 6

5 of 64 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	18	LEU
1	A	19	VAL
1	A	40	THR
1	A	66	VAL

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	485/485 (100%)	450 (93%)	35 (7%)	14	39
1	B	484/485 (100%)	446 (92%)	38 (8%)	12	34
All	All	969/970 (100%)	896 (92%)	73 (8%)	13	37

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	296	THR
1	B	506	LEU
1	B	338	VAL
1	B	408	ASN
1	A	360	ASN

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

Mol	Chain	Res	Type
1	A	510	GLN
1	B	510	GLN
1	B	70	HIS
1	B	454	HIS
1	B	5	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	203	B	1595	-	6,13,13	7.86	4 (66%)	5,20,20	8.02	5 (100%)

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	203	B	1595	-	-	-	0/1/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1595	203	O20-S16	15.37	1.65	1.43
3	B	1595	203	O21-S16	9.84	1.57	1.43
3	B	1595	203	C14-N15	4.30	1.42	1.37
3	B	1595	203	C13-C18	4.06	1.55	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1595	203	O21-S16-O20	-16.63	97.97	118.85
3	B	1595	203	C11-N12-C13	-4.78	93.89	102.99
3	B	1595	203	O19-C18-C13	-3.27	117.98	124.37
3	B	1595	203	O19-C18-N17	2.75	124.53	120.82
3	B	1595	203	C13-C18-N17	2.03	117.49	113.95

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1595	203	1	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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.