



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2018 – 01:12 am GMT

PDB ID : 2A1D
Title : Staphylocoagulase bound to bovine thrombin
Authors : Friedrich, R.; Panizzi, P.; Kawabata, S.; Bode, W.; Bock, P.E.; Fuentes-Prior, P.
Deposited on : 2005-06-20
Resolution : 3.50 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30686

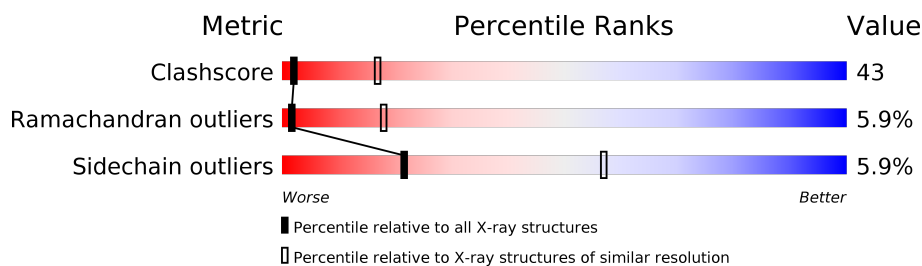
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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))
Clashscore	122078	1485 (3.60-3.40)
Ramachandran outliers	120005	1446 (3.60-3.40)
Sidechain outliers	119972	1447 (3.60-3.40)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	41	 37% 51% 12%
1	E	41	 34% 54% 12%
2	B	259	 33% 58% 9%
2	F	259	 33% 58% 9%
3	D	329	 31% 50% 5% 14%
3	H	329	 31% 50% . . 14%

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
4	0G6	B	500	X	-	-	-
4	0G6	F	600	X	-	-	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9634 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 thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	41	Total	C	N	O	S	0	0	0
			338	214	55	68	1			
1	E	41	Total	C	N	O	S	0	0	0
			338	214	55	68	1			

- Molecule 2 is a protein called thrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	259	Total	C	N	O	S	3	0	0
			2094	1337	376	369	12			
2	F	259	Total	C	N	O	S	0	0	0
			2094	1337	376	369	12			

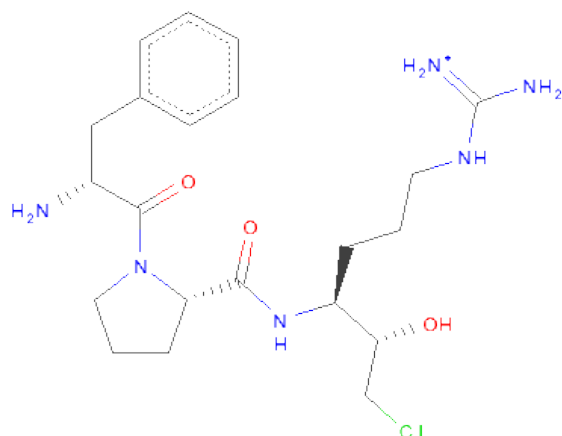
- Molecule 3 is a protein called Staphylocoagulase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	282	Total	C	N	O	S	0	0	0
			2347	1483	387	472	5			
3	H	282	Total	C	N	O	S	0	0	0
			2347	1483	387	472	5			

There are 4 discrepancies between the modelled and reference sequences:

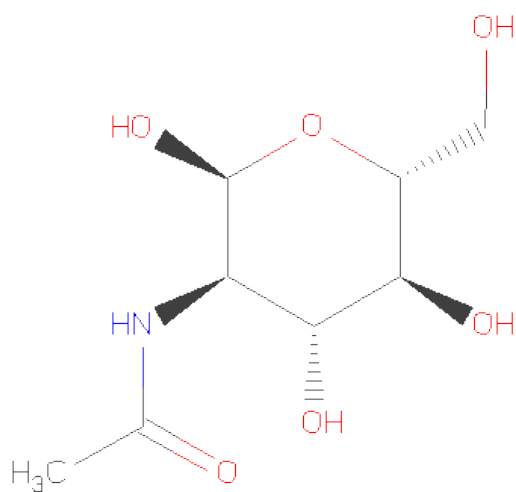
Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	INITIATING METHIONINE	UNP P17855
D	?	-	GLY	SEE REMARK 999	UNP P17855
H	0	MET	-	INITIATING METHIONINE	UNP P17855
H	?	-	GLY	SEE REMARK 999	UNP P17855

- Molecule 4 is D-phenylalanyl-N-[(2S,3S)-6-{[amino(iminio)methyl]amino}-1-chloro-2-hydroxyhexan-3-yl]-L-prolinamide (three-letter code: 0G6) (formula: C₂₁H₃₄ClN₆O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			30	21	6	3		
4	F	1	Total	C	N	O	0	0
			30	21	6	3		

- Molecule 5 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Na 1	0	0
6	F	1	Total 1	Na 1	0	0

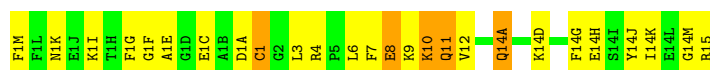
3 Residue-property plots

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.

Note EDS was not executed.

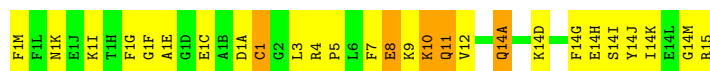
• Molecule 1: thrombin

Chain A: 



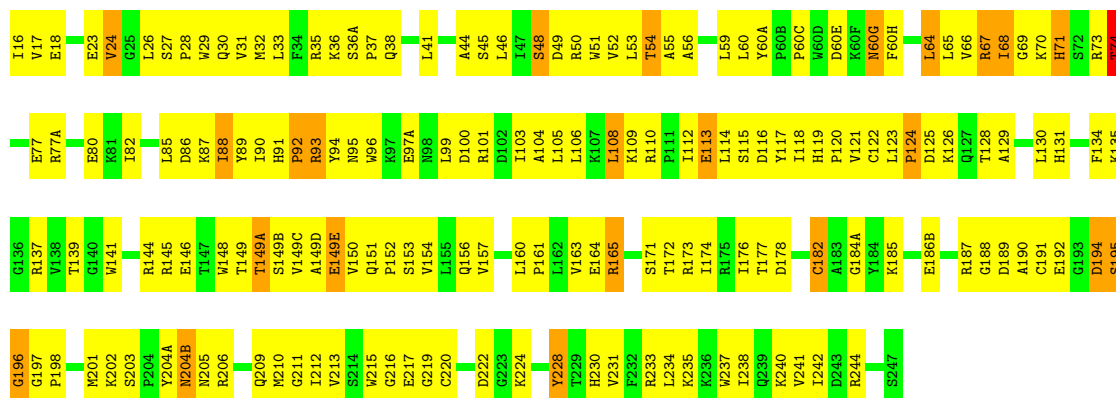
• Molecule 1: thrombin

Chain E: 

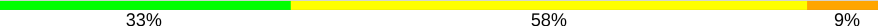


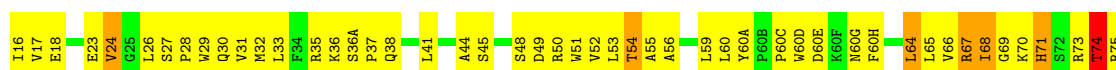
• Molecule 2: thrombin

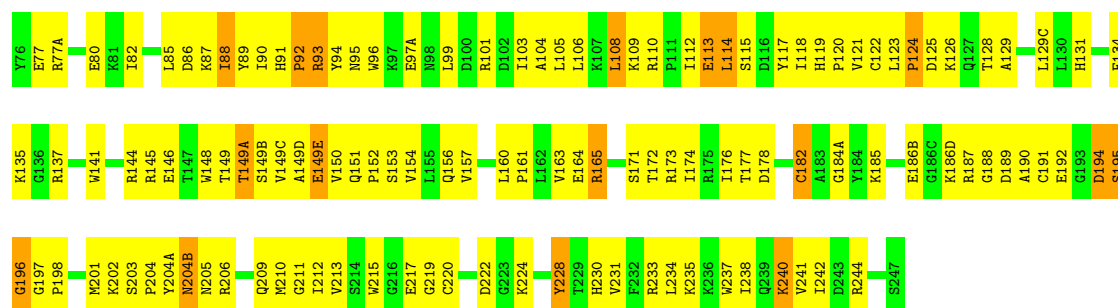
Chain B: 



• Molecule 2: thrombin

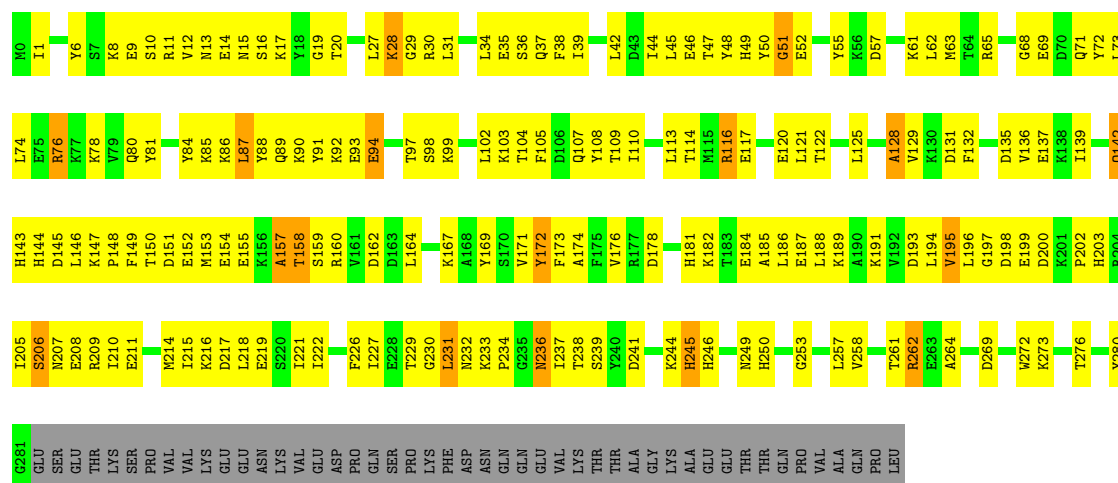
Chain F: 





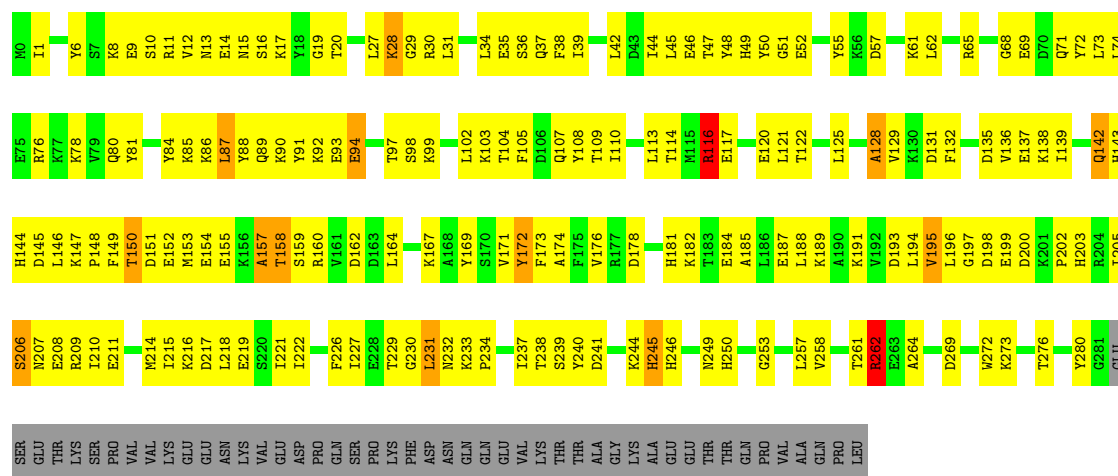
• Molecule 3: Staphylocoagulase

Chain D: 31% 50% 5% 14%



• Molecule 3: Staphylocoagulase

Chain H: 31% 50% 14%



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	183.78Å 102.54Å 134.04Å 90.00° 129.14° 90.00°	Depositor
Resolution (Å)	31.69 – 3.50	Depositor
% Data completeness (in resolution range)	86.7 (31.69-3.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.233 , 0.306	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9634	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

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: 0G6, NA, NDG

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	0.42	0/344	0.59	0/456
1	E	0.43	0/344	0.59	0/456
2	B	0.40	0/2148	0.69	0/2905
2	F	0.39	0/2148	0.68	0/2905
3	D	0.39	0/2389	0.79	6/3206 (0.2%)
3	H	0.39	0/2389	0.78	6/3206 (0.2%)
All	All	0.40	0/9762	0.73	12/13134 (0.1%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	262	ARG	NE-CZ-NH1	-15.01	112.80	120.30
3	H	262	ARG	NE-CZ-NH2	-14.67	112.96	120.30
3	D	262	ARG	NE-CZ-NH2	14.57	127.58	120.30
3	H	262	ARG	NE-CZ-NH1	13.87	127.23	120.30
3	D	116	ARG	NE-CZ-NH2	-12.50	114.05	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	338	0	319	38	0
1	E	338	0	319	38	0
2	B	2094	0	2095	219	0
2	F	2094	0	2096	209	0
3	D	2347	0	2311	183	0
3	H	2347	0	2311	178	0
4	B	30	0	30	3	0
4	F	30	0	30	3	0
5	B	14	0	13	4	0
6	B	1	0	0	0	0
6	F	1	0	0	0	0
All	All	9634	0	9524	817	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 43.

The worst 5 of 817 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:F:172:THR:HG22	2:F:174:ILE:H	1.14	1.06
2:B:172:THR:HG22	2:B:174:ILE:H	1.19	1.04
2:F:31:VAL:HB	2:F:44:ALA:HB3	1.46	0.97
2:F:211:GLY:HA2	2:F:231:VAL:HG23	1.46	0.97
2:B:31:VAL:HB	2:B:44:ALA:HB3	1.47	0.93

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	39/41 (95%)	28 (72%)	6 (15%)	5 (13%)	0 4
1	E	39/41 (95%)	30 (77%)	4 (10%)	5 (13%)	0 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	257/259 (99%)	195 (76%)	47 (18%)	15 (6%)	2	18
2	F	257/259 (99%)	193 (75%)	48 (19%)	16 (6%)	1	17
3	D	280/329 (85%)	201 (72%)	65 (23%)	14 (5%)	2	22
3	H	280/329 (85%)	204 (73%)	63 (22%)	13 (5%)	2	24
All	All	1152/1258 (92%)	851 (74%)	233 (20%)	68 (6%)	2	18

5 of 68 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	CYS
1	A	10	LYS
2	B	24	VAL
2	B	74	THR
2	B	149(E)	GLU

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	35/35 (100%)	33 (94%)	2 (6%)	23	57
1	E	35/35 (100%)	33 (94%)	2 (6%)	23	57
2	B	226/226 (100%)	211 (93%)	15 (7%)	18	53
2	F	226/226 (100%)	210 (93%)	16 (7%)	16	50
3	D	260/303 (86%)	248 (95%)	12 (5%)	29	64
3	H	260/303 (86%)	246 (95%)	14 (5%)	24	59
All	All	1042/1128 (92%)	981 (94%)	61 (6%)	21	57

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	237	ILE
2	F	67	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	H	194	LEU
3	D	245	HIS
1	E	11	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	203	HIS
1	E	11	GLN
3	H	181	HIS
3	D	232	ASN
1	E	1(K)	ASN

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 5 ligands modelled in this entry, 2 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	$\# Z > 2$
4	0G6	B	500	2	30,31,32	1.92	8 (26%)	37,41,42	1.20	3 (8%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	NDG	B	700	2	14,14,15	0.84	1 (7%)	17,19,21	0.82	1 (5%)
4	0G6	F	600	2	30,31,32	1.94	8 (26%)	37,41,42	1.14	2 (5%)

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
4	0G6	B	500	2	1/1/8/10	0/31/41/43	0/2/2/2
5	NDG	B	700	2	-	0/6/23/26	0/1/1/1
4	0G6	F	600	2	1/1/8/10	0/31/41/43	0/2/2/2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	600	0G6	O2-C2	-6.97	1.23	1.43
4	B	500	0G6	O2-C2	-6.77	1.23	1.43
4	F	600	0G6	CZ-CE1	2.19	1.43	1.38
5	B	700	NDG	C1-C2	2.30	1.55	1.52
4	B	500	0G6	C-N1	2.33	1.39	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	500	0G6	CG2-CB2-CA2	-2.13	109.56	113.96
5	B	700	NDG	C2-N2-C7	-2.07	119.92	122.94
4	F	600	0G6	CB1-CA1-N1	2.24	106.22	103.01
4	B	500	0G6	CB1-CA1-N1	2.39	106.44	103.01
4	B	500	0G6	O2-C2-C3	4.09	122.40	109.71

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	F	600	0G6	C2
4	B	500	0G6	C2

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	500	0G6	3	0
5	B	700	NDG	4	0
4	F	600	0G6	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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.