



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 01:46 am GMT

PDB ID : 2CBQ
Title : Crystal structure of the neocarzinostatin 1Tes15 mutant bound to testosterone hemisuccinate.
Authors : Drevelle, A.; Graille, M.; Heyd, B.; Sorel, I.; Ulryck, N.; Pecorari, F.; Desmadril, M.; Van Tilbeurgh, H.; Minard, P.
Deposited on : 2006-01-06
Resolution : 2.60 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)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

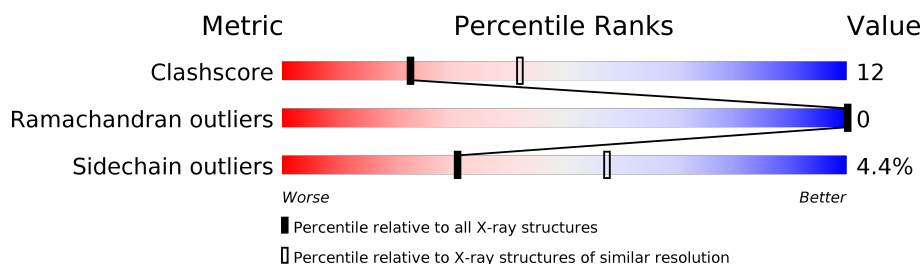
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	122126	3110 (2.60-2.60)
Ramachandran outliers	120053	3062 (2.60-2.60)
Sidechain outliers	120020	3062 (2.60-2.60)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	114	 75% 20% . .
1	B	114	 75% 22% ...
1	C	114	 75% 21% . .
1	D	114	 74% 24% . .
1	E	114	 75% 20% . . .
1	F	114	 75% 18% . .

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5016 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 NEOCARZINOSTATIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	111	Total	C	N	O	S	0	0	0
			773	475	134	162	2			
1	B	113	Total	C	N	O	S	0	0	0
			791	485	139	165	2			
1	C	112	Total	C	N	O	S	0	0	0
			781	479	136	164	2			
1	D	112	Total	C	N	O	S	0	0	0
			781	479	136	164	2			
1	E	112	Total	C	N	O	S	0	0	0
			781	479	136	164	2			
1	F	111	Total	C	N	O	S	0	0	0
			773	475	134	162	2			

There are 42 discrepancies between the modelled and reference sequences:

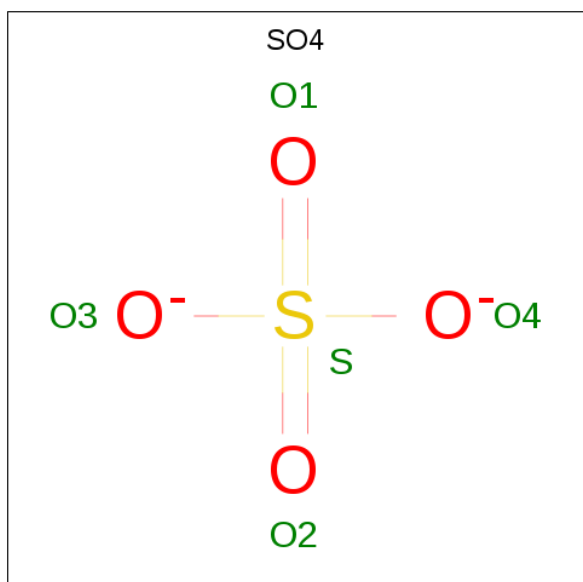
Chain	Residue	Modelled	Actual	Comment	Reference
A	33	TRP	ASP	engineered mutation	UNP P0A3R9
A	35	TYR	GLY	engineered mutation	UNP P0A3R9
A	37	ARG	CYS	engineered mutation	UNP P0A3R9
A	39	ALA	TRP	engineered mutation	UNP P0A3R9
A	45	HIS	LEU	engineered mutation	UNP P0A3R9
A	47	SER	CYS	engineered mutation	UNP P0A3R9
A	52	LEU	PHE	engineered mutation	UNP P0A3R9
B	33	TRP	ASP	engineered mutation	UNP P0A3R9
B	35	TYR	GLY	engineered mutation	UNP P0A3R9
B	37	ARG	CYS	engineered mutation	UNP P0A3R9
B	39	ALA	TRP	engineered mutation	UNP P0A3R9
B	45	HIS	LEU	engineered mutation	UNP P0A3R9
B	47	SER	CYS	engineered mutation	UNP P0A3R9
B	52	LEU	PHE	engineered mutation	UNP P0A3R9
C	33	TRP	ASP	engineered mutation	UNP P0A3R9
C	35	TYR	GLY	engineered mutation	UNP P0A3R9
C	37	ARG	CYS	engineered mutation	UNP P0A3R9

Continued on next page...

Continued from previous page...

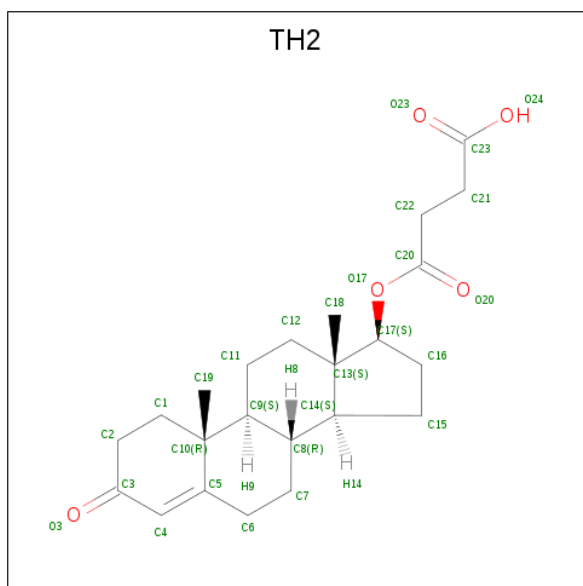
Chain	Residue	Modelled	Actual	Comment	Reference
C	39	ALA	TRP	engineered mutation	UNP P0A3R9
C	45	HIS	LEU	engineered mutation	UNP P0A3R9
C	47	SER	CYS	engineered mutation	UNP P0A3R9
C	52	LEU	PHE	engineered mutation	UNP P0A3R9
D	33	TRP	ASP	engineered mutation	UNP P0A3R9
D	35	TYR	GLY	engineered mutation	UNP P0A3R9
D	37	ARG	CYS	engineered mutation	UNP P0A3R9
D	39	ALA	TRP	engineered mutation	UNP P0A3R9
D	45	HIS	LEU	engineered mutation	UNP P0A3R9
D	47	SER	CYS	engineered mutation	UNP P0A3R9
D	52	LEU	PHE	engineered mutation	UNP P0A3R9
E	33	TRP	ASP	engineered mutation	UNP P0A3R9
E	35	TYR	GLY	engineered mutation	UNP P0A3R9
E	37	ARG	CYS	engineered mutation	UNP P0A3R9
E	39	ALA	TRP	engineered mutation	UNP P0A3R9
E	45	HIS	LEU	engineered mutation	UNP P0A3R9
E	47	SER	CYS	engineered mutation	UNP P0A3R9
E	52	LEU	PHE	engineered mutation	UNP P0A3R9
F	33	TRP	ASP	engineered mutation	UNP P0A3R9
F	35	TYR	GLY	engineered mutation	UNP P0A3R9
F	37	ARG	CYS	engineered mutation	UNP P0A3R9
F	39	ALA	TRP	engineered mutation	UNP P0A3R9
F	45	HIS	LEU	engineered mutation	UNP P0A3R9
F	47	SER	CYS	engineered mutation	UNP P0A3R9
F	52	LEU	PHE	engineered mutation	UNP P0A3R9

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0

- Molecule 3 is TESTOSTERONE HEMISUCCINATE (three-letter code: TH2) (formula: $C_{23}H_{32}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 28 23 5	0	0
3	A	1	Total C O 28 23 5	0	0
3	B	1	Total C O 28 23 5	0	0
3	B	1	Total C O 28 23 5	0	0
3	C	1	Total C O 28 23 5	0	0
3	D	1	Total C O 28 23 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			28	23	5		
3	F	1	Total	C	O	0	0
			28	23	5		

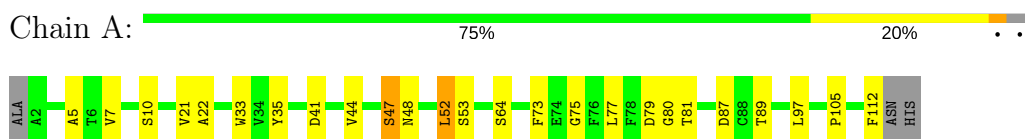
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total	O	0	0
			15	15		
4	B	17	Total	O	0	0
			17	17		
4	C	17	Total	O	0	0
			17	17		
4	D	17	Total	O	0	0
			17	17		
4	E	14	Total	O	0	0
			14	14		
4	F	7	Total	O	0	0
			7	7		

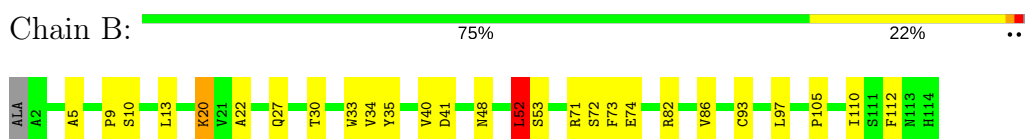
3 Residue-property plots

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. 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.

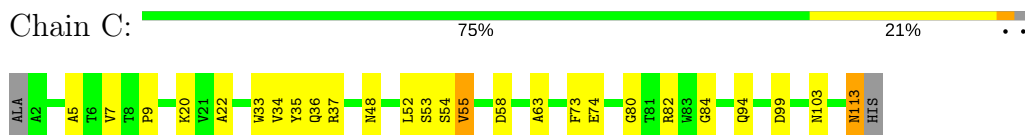
• Molecule 1: NEOCARZINOSTATIN



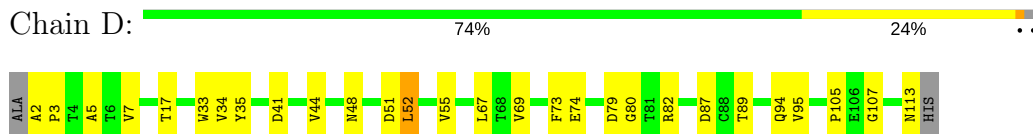
• Molecule 1: NEOCARZINOSTATIN



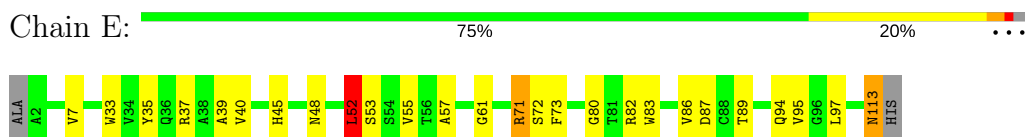
• Molecule 1: NEOCARZINOSTATIN



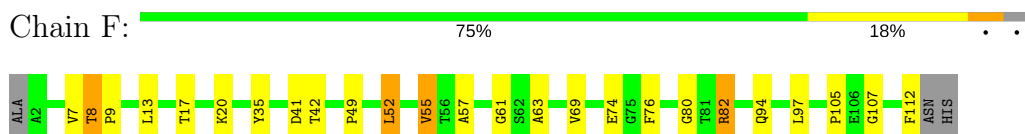
• Molecule 1: NEOCARZINOSTATIN



• Molecule 1: NEOCARZINOSTATIN



• Molecule 1: NEOCARZINOSTATIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	52.98Å 97.83Å 129.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 25.51 – 1.28	Depositor EDS
% Data completeness (in resolution range)	94.6 (20.00-2.60) 42.5 (25.51-1.28)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.74 (at 1.28Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.207 , 0.257 0.777 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	24.4	Xtriage
Anisotropy	0.650	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 19.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.67$, $\langle L^2 \rangle = 0.55$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.30	EDS
Total number of atoms	5016	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.30 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8852e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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: SO4, TH2

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	0.35	0/787	0.64	0/1078
1	B	0.38	0/806	0.64	1/1104 (0.1%)
1	C	0.43	1/795 (0.1%)	0.68	1/1089 (0.1%)
1	D	0.42	1/795 (0.1%)	0.66	1/1089 (0.1%)
1	E	0.41	1/795 (0.1%)	0.67	2/1089 (0.2%)
1	F	0.36	0/787	0.63	0/1078
All	All	0.39	3/4765 (0.1%)	0.66	5/6527 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	113	ASN	CG-OD1	5.51	1.36	1.24
1	D	113	ASN	CG-OD1	5.50	1.36	1.24
1	E	113	ASN	CG-OD1	5.47	1.35	1.24

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	113	ASN	OD1-CG-ND2	-6.88	106.08	121.90
1	D	113	ASN	OD1-CG-ND2	-6.88	106.08	121.90
1	C	113	ASN	OD1-CG-ND2	-6.87	106.11	121.90
1	E	52	LEU	CA-CB-CG	5.89	128.84	115.30
1	B	52	LEU	CA-CB-CG	5.47	127.87	115.30

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	773	0	731	21	0
1	B	791	0	744	24	0
1	C	781	0	737	23	0
1	D	781	0	737	16	0
1	E	781	0	737	22	0
1	F	773	0	731	19	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
3	A	56	0	62	3	0
3	B	56	0	62	1	0
3	C	28	0	31	0	0
3	D	28	0	31	0	0
3	E	28	0	31	0	0
3	F	28	0	31	0	0
4	A	15	0	0	0	0
4	B	17	0	0	1	0
4	C	17	0	0	2	0
4	D	17	0	0	0	0
4	E	14	0	0	0	0
4	F	7	0	0	0	0
All	All	5016	0	4665	119	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 12.

The worst 5 of 119 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:C:37:ARG:HE	1:C:94:GLN:HE21	1.17	0.92
1:D:35:TYR:HB3	1:D:52:LEU:HD23	1.53	0.90
3:A:1115:TH2:H6C2	1:F:20:LYS:HB2	1.53	0.88
3:B:1117:TH2:H6C2	1:C:20:LYS:HB2	1.64	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:37:ARG:NE	1:E:94:GLN:HE21	1.86	0.74

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	109/114 (96%)	106 (97%)	3 (3%)	0	100	100
1	B	111/114 (97%)	107 (96%)	4 (4%)	0	100	100
1	C	110/114 (96%)	104 (94%)	6 (6%)	0	100	100
1	D	110/114 (96%)	105 (96%)	5 (4%)	0	100	100
1	E	110/114 (96%)	107 (97%)	3 (3%)	0	100	100
1	F	109/114 (96%)	105 (96%)	4 (4%)	0	100	100
All	All	659/684 (96%)	634 (96%)	25 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	79/81 (98%)	75 (95%)	4 (5%)	26	50
1	B	81/81 (100%)	78 (96%)	3 (4%)	37	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	80/81 (99%)	78 (98%)	2 (2%)	50	76
1	D	80/81 (99%)	77 (96%)	3 (4%)	36	63
1	E	80/81 (99%)	75 (94%)	5 (6%)	20	40
1	F	79/81 (98%)	75 (95%)	4 (5%)	26	50
All	All	479/486 (99%)	458 (96%)	21 (4%)	31	57

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	52	LEU
1	D	79	ASP
1	F	8	THR
1	C	113	ASN
1	F	52	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	HIS
1	A	94	GLN
1	C	94	GLN
1	D	45	HIS
1	E	94	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

13 ligands 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	SO4	A	1113	-	4,4,4	0.36	0	6,6,6	0.08	0
3	TH2	A	1114	-	28,31,31	1.61	7 (25%)	45,48,48	2.00	11 (24%)
3	TH2	A	1115	-	28,31,31	1.65	8 (28%)	45,48,48	1.99	10 (22%)
2	SO4	B	1115	-	4,4,4	0.38	0	6,6,6	0.17	0
3	TH2	B	1116	-	28,31,31	1.62	6 (21%)	45,48,48	1.94	10 (22%)
3	TH2	B	1117	-	28,31,31	1.65	7 (25%)	45,48,48	2.06	11 (24%)
2	SO4	C	1114	-	4,4,4	0.35	0	6,6,6	0.14	0
3	TH2	C	1115	-	28,31,31	1.61	7 (25%)	45,48,48	1.95	10 (22%)
3	TH2	D	1114	-	28,31,31	1.64	8 (28%)	45,48,48	2.00	10 (22%)
2	SO4	E	1114	-	4,4,4	0.35	0	6,6,6	0.10	0
3	TH2	E	1115	-	28,31,31	1.59	8 (28%)	45,48,48	2.02	11 (24%)
2	SO4	F	1113	-	4,4,4	0.33	0	6,6,6	0.12	0
3	TH2	F	1114	-	28,31,31	1.54	9 (32%)	45,48,48	1.97	9 (20%)

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	SO4	A	1113	-	-	0/0/0/0	0/0/0/0
3	TH2	A	1114	-	-	0/7/67/67	0/4/4/4
3	TH2	A	1115	-	-	0/7/67/67	0/4/4/4
2	SO4	B	1115	-	-	0/0/0/0	0/0/0/0
3	TH2	B	1116	-	-	0/7/67/67	0/4/4/4
3	TH2	B	1117	-	-	0/7/67/67	0/4/4/4
2	SO4	C	1114	-	-	0/0/0/0	0/0/0/0
3	TH2	C	1115	-	-	0/7/67/67	0/4/4/4
3	TH2	D	1114	-	-	1/7/67/67	0/4/4/4
2	SO4	E	1114	-	-	0/0/0/0	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	TH2	E	1115	-	-	0/7/67/67	0/4/4/4
2	SO4	F	1113	-	-	0/0/0/0	0/0/0/0
3	TH2	F	1114	-	-	0/7/67/67	0/4/4/4

The worst 5 of 60 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1116	TH2	O17-C17	-2.74	1.41	1.46
3	A	1114	TH2	O17-C17	-2.60	1.41	1.46
3	E	1115	TH2	O17-C17	-2.49	1.42	1.46
3	F	1114	TH2	O17-C17	-2.40	1.42	1.46
3	B	1116	TH2	O17-C20	-2.32	1.27	1.34

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1115	TH2	C11-C9-C10	-4.55	106.89	113.10
3	B	1116	TH2	C11-C9-C10	-4.24	107.32	113.10
3	B	1117	TH2	C11-C9-C10	-4.20	107.37	113.10
3	D	1114	TH2	C11-C9-C10	-4.19	107.38	113.10
3	A	1114	TH2	C11-C9-C10	-4.18	107.40	113.10

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1114	TH2	C20-O17-C17-C16

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1114	TH2	2	0
3	A	1115	TH2	1	0
3	B	1117	TH2	1	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 ⓘ

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.