wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 01:48 AM EDT

PDB ID	:	6F8Y
Title	:	Crystal structure of P. abyssi Sua5 complexed with L-threonine
Authors	:	Pichard-Kostuch, A.; Zhang, W.; Liger, D.; Daugeron, M.C.; Letoquart, J.; Li de la Sierra-Gallay, I.; Forterre, P.; Collinet, B.; van Tilbeurgh, H.; Basta, T.
Deposited on	:	2017-12-13
Resolution	:	2.86 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:

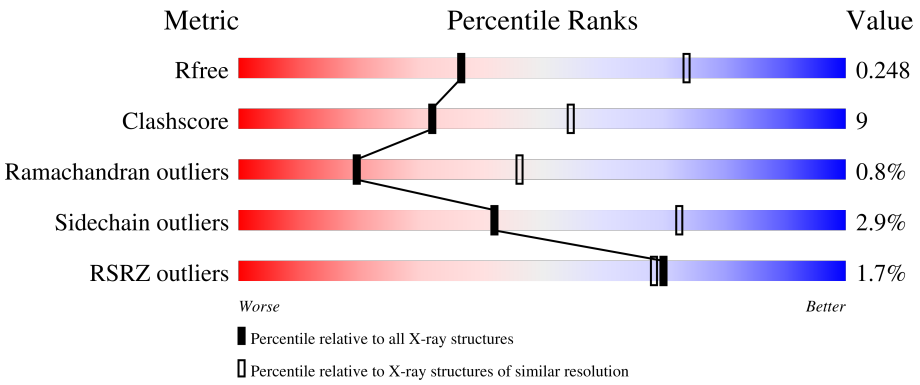
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.86 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	345	% 79% 18% ••
1	B	345	% 80% 19% •
1	C	345	80% 18% ••
1	D	345	4% 72% 26% ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10631 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 Threonylcarbamoyl-AMP synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	340	2640	1690	460	484	6	0	0	0
1	B	340	2640	1690	460	484	6	0	0	0
1	C	341	2650	1696	463	485	6	0	0	0
1	D	340	2640	1690	460	484	6	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

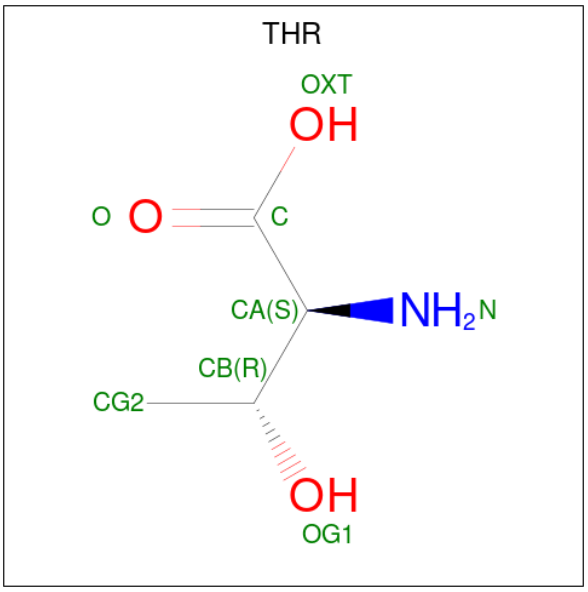
Chain	Residue	Modelled	Actual	Comment	Reference
A	341	HIS	-	expression tag	UNP Q9UYB2
A	342	HIS	-	expression tag	UNP Q9UYB2
A	343	HIS	-	expression tag	UNP Q9UYB2
A	344	HIS	-	expression tag	UNP Q9UYB2
A	345	HIS	-	expression tag	UNP Q9UYB2
A	346	HIS	-	expression tag	UNP Q9UYB2
B	341	HIS	-	expression tag	UNP Q9UYB2
B	342	HIS	-	expression tag	UNP Q9UYB2
B	343	HIS	-	expression tag	UNP Q9UYB2
B	344	HIS	-	expression tag	UNP Q9UYB2
B	345	HIS	-	expression tag	UNP Q9UYB2
B	346	HIS	-	expression tag	UNP Q9UYB2
C	341	HIS	-	expression tag	UNP Q9UYB2
C	342	HIS	-	expression tag	UNP Q9UYB2
C	343	HIS	-	expression tag	UNP Q9UYB2
C	344	HIS	-	expression tag	UNP Q9UYB2
C	345	HIS	-	expression tag	UNP Q9UYB2
C	346	HIS	-	expression tag	UNP Q9UYB2
D	341	HIS	-	expression tag	UNP Q9UYB2
D	342	HIS	-	expression tag	UNP Q9UYB2
D	343	HIS	-	expression tag	UNP Q9UYB2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	344	HIS	-	expression tag	UNP Q9UYB2
D	345	HIS	-	expression tag	UNP Q9UYB2
D	346	HIS	-	expression tag	UNP Q9UYB2

- Molecule 2 is THREONINE (three-letter code: THR) (formula: $C_4H_9NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	8	4	1	3	0	0
2	B	1	8	4	1	3	0	0
2	C	1	8	4	1	3	0	0
2	D	1	8	4	1	3	0	0

- Molecule 3 is water.

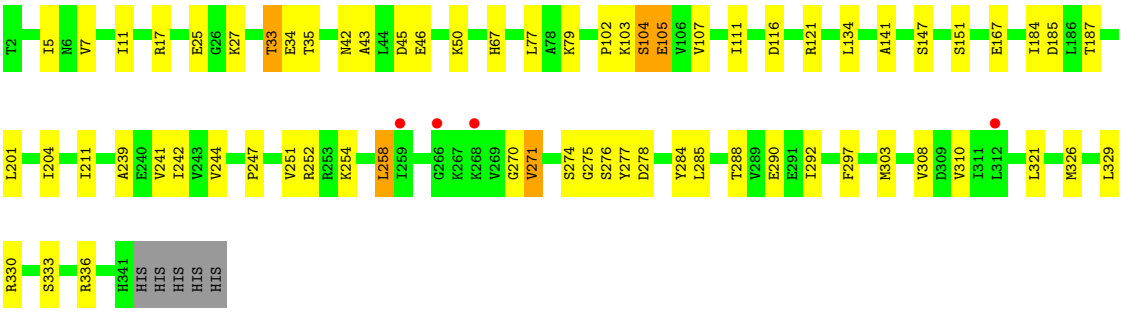
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	9	9	9	0	0
3	B	7	7	7	0	0
3	C	9	9	9	0	0
3	D	4	4	4	0	0

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.

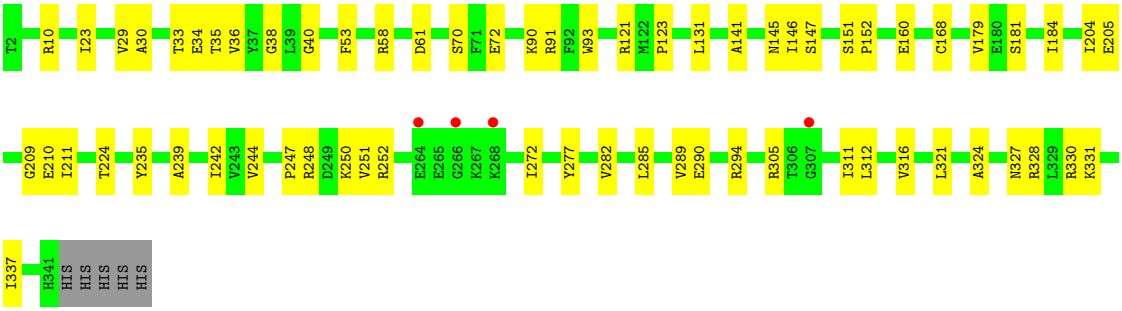
• Molecule 1: Threonylcarbamoyl-AMP synthase

Chain A:



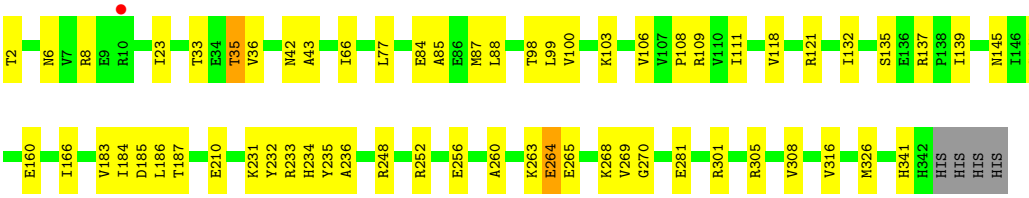
• Molecule 1: Threonylcarbamoyl-AMP synthase

Chain B:



• Molecule 1: Threonylcarbamoyl-AMP synthase

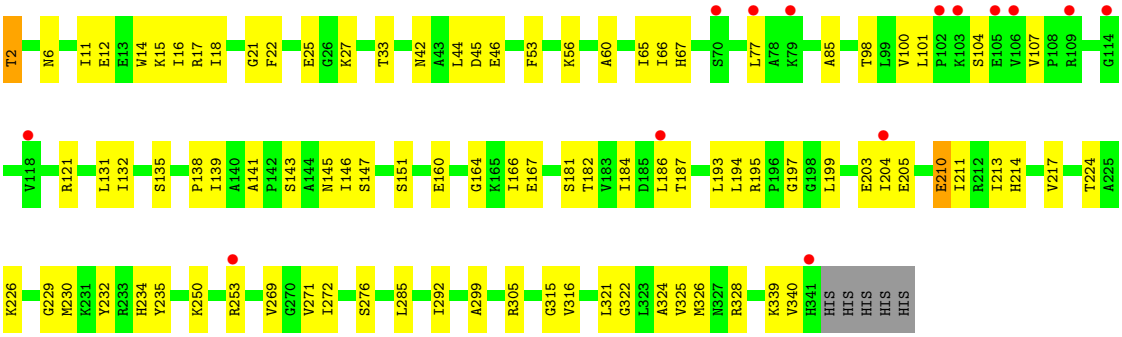
Chain C:



• Molecule 1: Threonylcarbamoyl-AMP synthase

Chain D:

4%
72%
26%


T2 N6 I11 E12 E13 W14 K15 I16 R17 I18 G21 F22 E25 G26 K27 T33 N42 A43 L44 D45 E46 F53 K56 A60 I65 I66 H67 S70 L77 A78 K79 A85 T98 L99 V100 L101 P102 K103 S104 E105 V106 V107 P108 R109 G114
V118 R121 L131 I132 S135 P138 I139 A140 A141 P142 S143 A144 N145 I146 S147 S151 E160 G164 K165 I166 E167 S181 T182 V183 I184 D185 L186 T187 L193 L194 R195 P196 G197 G198 L199 E203 I204 E205 E210 I211 R212 I213 H214 V217 T224 A225
K226 G229 M230 K231 Y232 R233 H234 Y235 K250 R253 V269 G270 V271 I272 S276 L285 I292 A299 R305 G315 V316 L321 G322 L323 A324 V325 M326 N327 R328 K339 V340 H341 HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	152.23Å 142.38Å 103.86Å 90.00° 101.80° 90.00°	Depositor
Resolution (Å)	48.78 – 2.86 48.78 – 2.86	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.78-2.86) 99.3 (48.78-2.86)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.188 , 0.252 0.192 , 0.248	Depositor DCC
R_{free} test set	2497 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	75.4	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 58.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	10631	wwPDB-VP
Average B, all atoms (Å^2)	79.0	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 47.30 % 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 1.0120e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

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	0.66	0/2691	0.83	1/3642 (0.0%)
1	B	0.67	1/2691 (0.0%)	0.87	1/3642 (0.0%)
1	C	0.67	0/2702	0.86	0/3657
1	D	0.62	0/2691	0.85	1/3642 (0.0%)
All	All	0.65	1/10775 (0.0%)	0.85	3/14583 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	61	ASP	CB-CG	-5.74	1.39	1.51

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	17	ARG	NE-CZ-NH1	5.07	122.83	120.30
1	A	17	ARG	NE-CZ-NH1	5.04	122.82	120.30
1	B	91	ARG	NE-CZ-NH2	-5.01	117.80	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	2640	0	2729	36	0
1	B	2640	0	2729	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2650	0	2736	52	0
1	D	2640	0	2729	69	0
2	A	8	0	6	1	0
2	B	8	0	6	1	0
2	C	8	0	6	3	0
2	D	8	0	6	1	0
3	A	9	0	0	0	0
3	B	7	0	0	0	0
3	C	9	0	0	0	0
3	D	4	0	0	0	0
All	All	10631	0	10947	192	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 9.

The worst 5 of 192 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:2:THR:HG22	1:C:166:ILE:O	1.27	1.28
1:C:2:THR:CG2	1:C:166:ILE:O	2.16	0.93
1:C:248:ARG:NH1	1:C:252:ARG:HD2	1.87	0.89
1:C:121:ARG:NH2	2:C:401:THR:OG1	2.06	0.89
1:D:2:THR:HG22	1:D:166:ILE:O	1.74	0.87

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	338/345 (98%)	316 (94%)	18 (5%)	4 (1%)	13 35
1	B	338/345 (98%)	316 (94%)	19 (6%)	3 (1%)	17 43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	339/345 (98%)	320 (94%)	18 (5%)	1 (0%)	41	68
1	D	338/345 (98%)	314 (93%)	21 (6%)	3 (1%)	17	43
All	All	1353/1380 (98%)	1266 (94%)	76 (6%)	11 (1%)	19	46

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	209	GLY
1	C	264	GLU
1	B	248	ARG
1	D	46	GLU
1	D	210	GLU

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	279/284 (98%)	268 (96%)	11 (4%)	32	63
1	B	279/284 (98%)	272 (98%)	7 (2%)	47	76
1	C	280/284 (99%)	273 (98%)	7 (2%)	47	76
1	D	279/284 (98%)	271 (97%)	8 (3%)	42	72
All	All	1117/1136 (98%)	1084 (97%)	33 (3%)	42	72

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	203	GLU
1	D	224	THR
1	D	276	SER
1	B	147	SER
1	B	146	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	145	ASN
1	B	157	HIS
1	C	234	HIS
1	D	145	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	THR	B	401	-	7,7,7	1.01	1 (14%)	8,9,9	1.10	1 (12%)
2	THR	A	401	-	7,7,7	0.80	1 (14%)	8,9,9	1.54	2 (25%)
2	THR	D	401	-	7,7,7	1.22	1 (14%)	8,9,9	1.21	1 (12%)
2	THR	C	401	-	7,7,7	0.92	0	8,9,9	1.55	2 (25%)

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	THR	B	401	-	-	0/8/8/8	-
2	THR	A	401	-	-	0/8/8/8	-
2	THR	D	401	-	-	0/8/8/8	-
2	THR	C	401	-	-	1/8/8/8	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	401	THR	OXT-C	-2.36	1.23	1.30
2	B	401	THR	OXT-C	-2.15	1.23	1.30
2	A	401	THR	OXT-C	-2.06	1.24	1.30

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	THR	OXT-C-O	-3.35	116.48	124.08
2	C	401	THR	OXT-C-O	-3.22	116.78	124.08
2	B	401	THR	OXT-C-O	-2.65	118.07	124.08
2	C	401	THR	OXT-C-CA	2.59	123.09	114.15
2	A	401	THR	OXT-C-CA	2.53	122.88	114.15

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	401	THR	O-C-CA-N

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	THR	1	0
2	A	401	THR	1	0
2	D	401	THR	1	0
2	C	401	THR	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 (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, 95th 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	340/345 (98%)	-0.08	4 (1%)	79	78	47, 73, 116, 136	0
1	B	340/345 (98%)	-0.07	4 (1%)	79	78	45, 70, 119, 145	0
1	C	341/345 (98%)	-0.14	1 (0%)	94	94	47, 68, 102, 158	0
1	D	340/345 (98%)	0.12	14 (4%)	37	31	55, 87, 123, 151	0
All	All	1361/1380 (98%)	-0.04	23 (1%)	70	68	45, 75, 117, 158	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	341	HIS	4.7
1	B	268	LYS	4.4
1	D	70	SER	3.3
1	D	106	VAL	3.2
1	D	118	VAL	3.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no monosaccharides 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, 95th 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
2	THR	D	401	8/8	0.92	0.16	73,82,88,88	0
2	THR	C	401	8/8	0.94	0.23	64,69,75,78	0
2	THR	A	401	8/8	0.98	0.17	59,64,72,72	0
2	THR	B	401	8/8	0.98	0.18	51,57,62,63	0

6.5 Other polymers (i)

There are no such residues in this entry.