wwPDB X-ray Structure Validation Summary Report (i)

Oct 26, 2024 – 12:39 PM EDT

PDB ID	:	2QO3
Title	:	Crystal Structure of [KS3][AT3] didomain from module 3 of 6-deoxyerthronolide B synthase
Authors	:	Khosla, C.; Cane, E.D.; Tang, Y.; Chen, Y.A.; Kim, C.Y.
Deposited on	:	2007-07-19
Resolution	:	2.59 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:

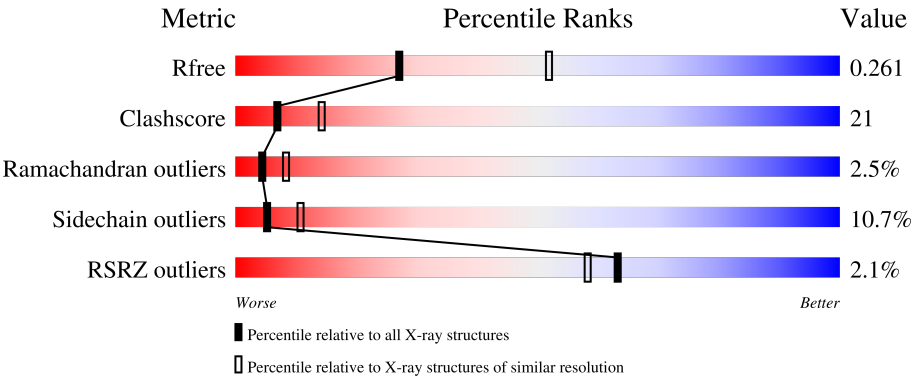
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.59 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}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (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 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	915	2% 60% 30% 5% • 5%
1	B	915	2% 61% 27% 6% • 5%

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	ACT	A	950	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13148 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 EryAII Erythromycin polyketide synthase modules 3 and 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	869	Total	C	N	O	S	0	0	0
			6449	4003	1178	1250	18			
1	B	873	Total	C	N	O	S	0	0	0
			6479	4020	1185	1256	18			

There are 38 discrepancies between the modelled and reference sequences:

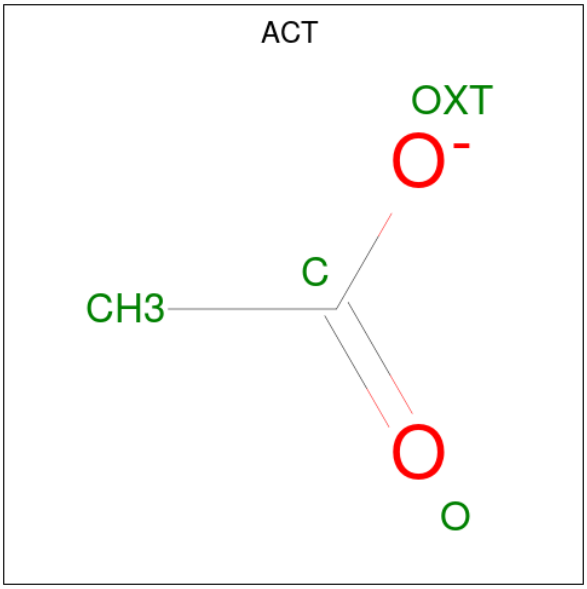
Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	-	initiating methionine	UNP A4F7P0
A	923	SER	-	expression tag	UNP A4F7P0
A	924	SER	-	expression tag	UNP A4F7P0
A	925	SER	-	expression tag	UNP A4F7P0
A	926	VAL	-	expression tag	UNP A4F7P0
A	927	ASP	-	expression tag	UNP A4F7P0
A	928	LYS	-	expression tag	UNP A4F7P0
A	929	LEU	-	expression tag	UNP A4F7P0
A	930	ALA	-	expression tag	UNP A4F7P0
A	931	ALA	-	expression tag	UNP A4F7P0
A	932	ALA	-	expression tag	UNP A4F7P0
A	933	LEU	-	expression tag	UNP A4F7P0
A	934	GLU	-	expression tag	UNP A4F7P0
A	935	HIS	-	expression tag	UNP A4F7P0
A	936	HIS	-	expression tag	UNP A4F7P0
A	937	HIS	-	expression tag	UNP A4F7P0
A	938	HIS	-	expression tag	UNP A4F7P0
A	939	HIS	-	expression tag	UNP A4F7P0
A	940	HIS	-	expression tag	UNP A4F7P0
B	26	MET	-	initiating methionine	UNP A4F7P0
B	923	SER	-	expression tag	UNP A4F7P0
B	924	SER	-	expression tag	UNP A4F7P0
B	925	SER	-	expression tag	UNP A4F7P0
B	926	VAL	-	expression tag	UNP A4F7P0
B	927	ASP	-	expression tag	UNP A4F7P0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	928	LYS	-	expression tag	UNP A4F7P0
B	929	LEU	-	expression tag	UNP A4F7P0
B	930	ALA	-	expression tag	UNP A4F7P0
B	931	ALA	-	expression tag	UNP A4F7P0
B	932	ALA	-	expression tag	UNP A4F7P0
B	933	LEU	-	expression tag	UNP A4F7P0
B	934	GLU	-	expression tag	UNP A4F7P0
B	935	HIS	-	expression tag	UNP A4F7P0
B	936	HIS	-	expression tag	UNP A4F7P0
B	937	HIS	-	expression tag	UNP A4F7P0
B	938	HIS	-	expression tag	UNP A4F7P0
B	939	HIS	-	expression tag	UNP A4F7P0
B	940	HIS	-	expression tag	UNP A4F7P0

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

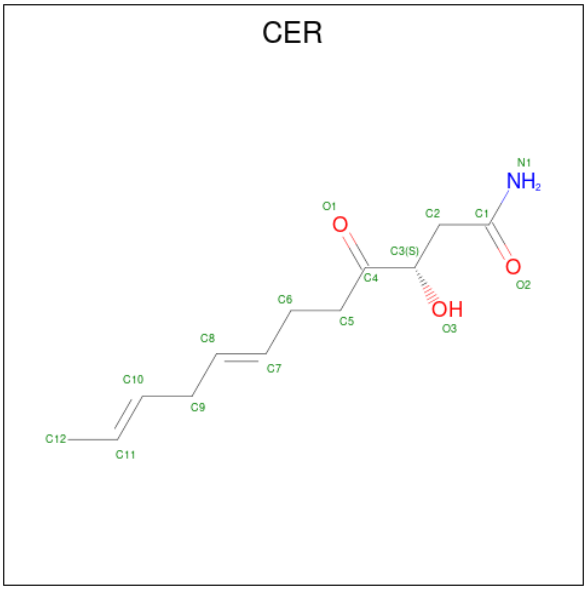


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0

- Molecule 4 is (2S, 3R)-3-HYDROXY-4-OXO-7,10-TRANS,TRANS-DODECADIENAMID E (three-letter code: CER) (formula: $C_{12}H_{19}NO_3$).



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

- Molecule 5 is water.

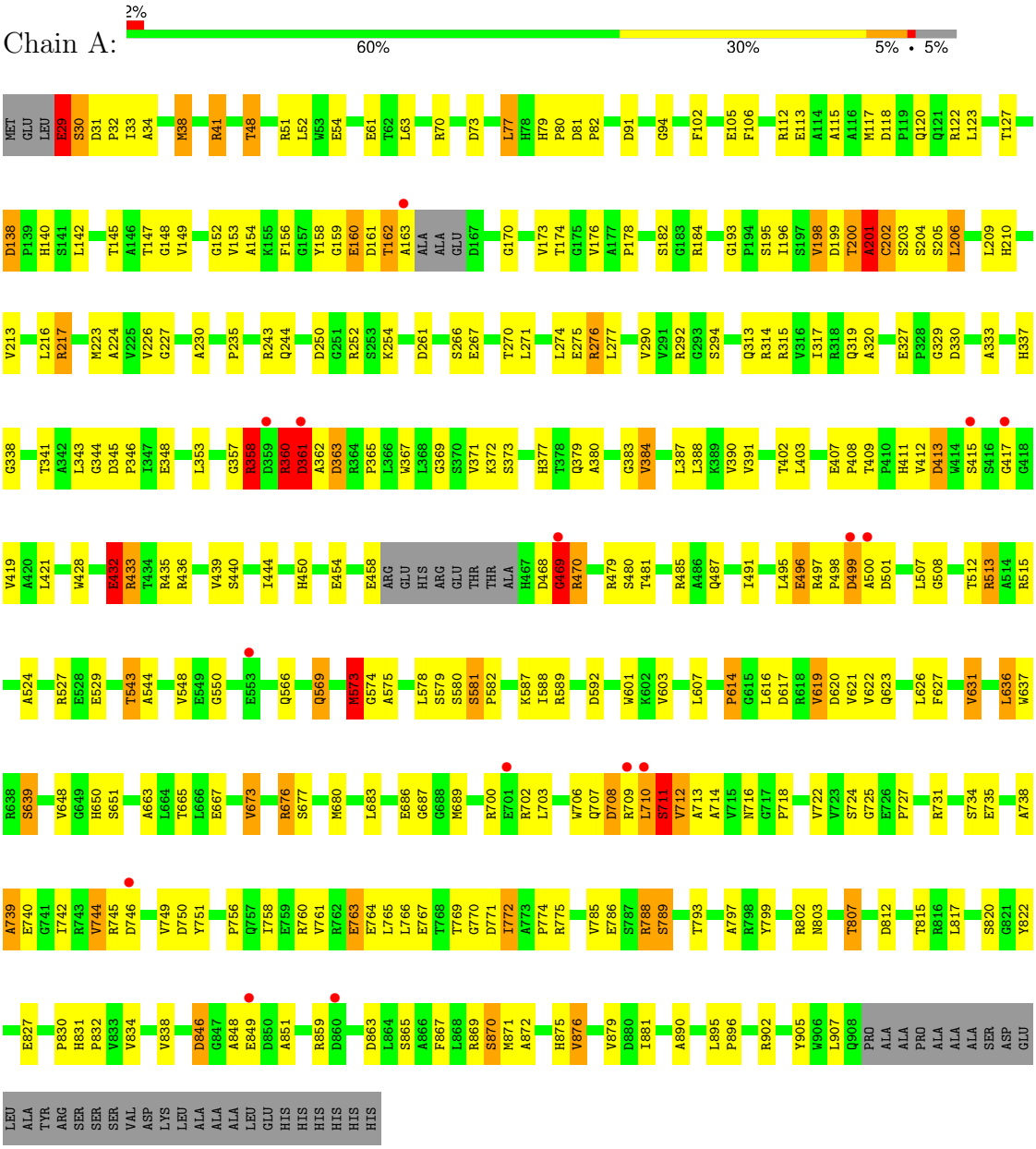
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	95	Total O 95 95	0	0
5	B	99	Total O 99 99	0	0

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.

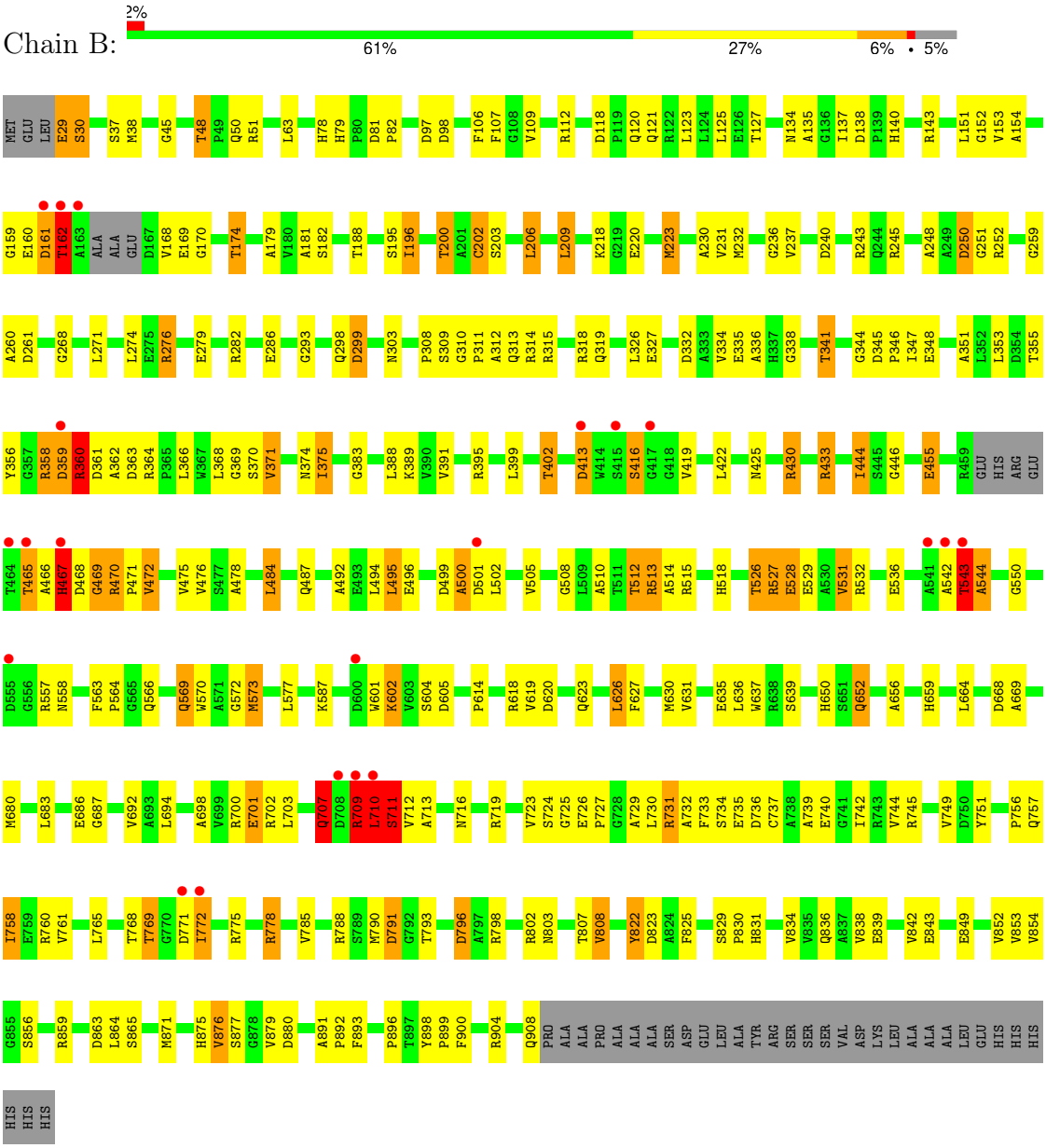
• Molecule 1: EryAII Erythromycin polyketide synthase modules 3 and 4

Chain A:



• Molecule 1: EryAII Erythromycin polyketide synthase modules 3 and 4

Chain B:


2%
61%
27%
6%
5%
MET GLU LEU E29 S30 S37 M38 G45 T48 P49 Q50 R51 L63 H78 H79 P80 D81 P82 D97 D98 F106 F107 G108 V109 R112 D118 P119 Q120 Q121 R122 L123 L124 L125 E126 T127 N134 A135 G136 I137 D138 P139 H140 R143 L151 G152 V153 A154
G159 E160 D161 T162 A163 ALA ALA GLU D167 V168 E169 G170 T174 A179 V180 A181 S182 T188 S195 I196 T200 A201 C202 S203 L206 L209 K218 G219 E220 M223 A230 V231 M232 G236 V237 D240 R243 Q244 R245 A248 A249 D250 G251 R252 G259
A260 D261 G268 L271 L274 E275 R276 E279 R282 E286 G293 Q298 D299 N303 P308 S309 G310 P311 A312 Q313 R314 R315 R318 Q319 L326 E327 D332 A333 V334 E335 A336 H337 G338 T341 G344 D345 P346 I347 E348 A351 L352 L353 D354 T355
Y356 G357 R358 D359 R360 D361 A362 D363 R364 P365 L366 W367 L368 G369 S370 V371 N374 I375 G383 L388 K389 V390 V391 R395 L399 T402 D413 W414 S415 S416 G417 G418 V419 L422 N425 R430 R433 I444 S445 G446 E455 R459 GLU HIS ARG GLU
T464 T465 A466 H467 D468 G469 R470 P471 V472 V475 V476 S477 A478 L484 Q487 A492 E493 L494 L495 E496 D499 A500 D501 L502 V505 G508 L509 A510 T511 T512 R513 A514 R515 H518 T526 R527 E528 E529 A530 V531 R532 E536 A541 A542 T543 A544 G550
D555 G556 R557 N558 F563 P564 G565 Q566 Q569 W570 A571 G572 M573 L577 K587 D600 W601 K602 V603 S604 D605 P614 R618 V619 D620 Q623 L626 F627 M630 V631 E635 L636 W637 R638 S639 H650 S651 Q652 A656 H659 L664 D668 A669
M680 L683 E686 G687 V692 A693 L694 A698 V699 R700 E701 R702 L703 Q707 D708 R709 L710 S711 V712 A713 N716 R719 V723 S724 G725 E726 P727 G728 A729 L730 R731 A732 F733 S734 E735 D736 C737 A738 A739 E740 G741 I742 R743 V744 R745 V749 D750 Y751 P756 Q757
I758 E759 R760 V761 L765 T768 T769 G770 D771 I772 R775 R778 V785 R788 S789 M790 D791 G792 T793 D796 A797 R798 R802 N803 T807 V808 Y822 D823 A824 F825 S829 P830 H831 V834 V835 Q836 A837 V838 E839 V842 E843 E849 V852 V853 V854
G855 S856 R859 D863 L864 S865 M871 H875 V876 S877 G878 V879 D880 A891 P892 F893 P896 T897 Y898 P899 F900 R904 Q908 PRO ALA ALA PRO ALA ALA ALA SER ASP GLU LEU ALA TYR ARG SER SER SER VAL ASP LYS LEU ALA ALA ALA 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	75.20Å 139.00Å 102.34Å 90.00° 106.14° 90.00°	Depositor
Resolution (Å)	41.52 – 2.59 41.52 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.5 (41.52-2.59) 99.4 (41.52-2.60)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.12 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.214 , 0.268 0.210 , 0.261	Depositor DCC
R_{free} test set	3145 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 34.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	13148	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.39% 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: CL, ACT, CER

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	0.83	0/6573	0.97	18/8942 (0.2%)
1	B	0.86	1/6603 (0.0%)	0.95	13/8983 (0.1%)
All	All	0.84	1/13176 (0.0%)	0.96	31/17925 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	11
1	B	0	9
All	All	0	20

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	327	GLU	CG-CD	5.68	1.60	1.51

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	201	ALA	O-C-N	-10.90	105.26	122.70
1	A	201	ALA	C-N-CA	10.14	147.05	121.70
1	B	202	CYS	O-C-N	-9.46	107.57	122.70
1	A	360	ARG	NE-CZ-NH1	8.20	124.40	120.30
1	A	202	CYS	C-N-CA	-7.67	102.53	121.70

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	160	GLU	Peptide
1	A	201	ALA	Mainchain
1	A	202	CYS	Mainchain
1	A	29	GLU	Peptide
1	A	361	ASP	Peptide

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	6449	0	6307	269	0
1	B	6479	0	6340	270	0
2	A	4	0	3	2	0
2	B	4	0	3	0	0
3	A	1	0	0	1	0
3	B	1	0	0	0	0
4	A	8	0	5	1	0
4	B	8	0	5	2	0
5	A	95	0	0	22	0
5	B	99	0	0	21	0
All	All	13148	0	12663	532	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 21.

The worst 5 of 532 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:B:711:SER:HB3	1:B:724:SER:N	1.40	1.35
1:B:711:SER:CB	1:B:724:SER:H	1.56	1.19
1:A:711:SER:HB2	1:A:724:SER:H	0.98	1.12
1:A:709:ARG:NH1	1:A:709:ARG:HB2	1.64	1.12
1:B:37:SER:HB2	1:B:135:ALA:HB2	1.34	1.07

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	863/915 (94%)	779 (90%)	64 (7%)	20 (2%)	5 10
1	B	867/915 (95%)	781 (90%)	62 (7%)	24 (3%)	4 7
All	All	1730/1830 (94%)	1560 (90%)	126 (7%)	44 (2%)	4 8

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	SER
1	A	358	ARG
1	A	415	SER
1	A	432	GLU
1	A	470	ARG

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	654/688 (95%)	592 (90%)	62 (10%)	7 14
1	B	657/688 (96%)	579 (88%)	78 (12%)	4 8
All	All	1311/1376 (95%)	1171 (89%)	140 (11%)	5 11

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	709	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	733	PHE
1	B	808	VAL
1	A	744	VAL
1	A	710	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	210	HIS
1	B	467	HIS
1	B	404	HIS
1	B	490	GLN
1	A	516	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
2	ACT	B	950	-	3,3,3	0.77	0	3,3,3	1.38	0
4	CER	B	960	1	6,7,15	0.59	0	5,8,17	0.55	0
4	CER	A	960	1	6,7,15	0.60	0	5,8,17	0.59	0
2	ACT	A	950	-	3,3,3	0.90	0	3,3,3	1.09	0

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	CER	A	960	1	-	0/5/6/16	-
4	CER	B	960	1	-	2/5/6/16	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	960	CER	C2-C3-C4-O1
4	B	960	CER	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	960	CER	2	0
4	A	960	CER	1	0
2	A	950	ACT	2	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

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	869/915 (94%)	-0.09	15 (1%)	69	64	7, 20, 38, 50	0
1	B	873/915 (95%)	-0.06	21 (2%)	59	54	7, 19, 40, 61	0
All	All	1742/1830 (95%)	-0.08	36 (2%)	63	58	7, 20, 39, 61	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	710	LEU	5.1
1	A	499	ASP	4.3
1	B	771	ASP	4.2
1	B	710	LEU	3.8
1	A	860	ASP	3.5

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
4	CER	A	960	8/16	0.81	0.16	38,40,42,43	0
2	ACT	B	950	4/4	0.84	0.20	37,37,37,37	0
2	ACT	A	950	4/4	0.84	0.15	30,30,31,31	0
4	CER	B	960	8/16	0.91	0.15	40,42,47,49	0
3	CL	A	2	1/1	0.95	0.08	36,36,36,36	0
3	CL	B	1	1/1	0.95	0.08	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.