



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:31 PM GMT

PDB ID : 1U0W
Title : An Aldol Switch Discovered in Stilbene Synthases Mediates Cyclization Specificity of Type III Polyketide Synthases: 18xCHS+resveratrol Structure
Authors : Austin, M.B.; Bowman, M.E.; Ferrer, J.-L.; Schroder, J.; Noel, J.P.
Deposited on : 2004-07-14
Resolution : 2.00 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
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

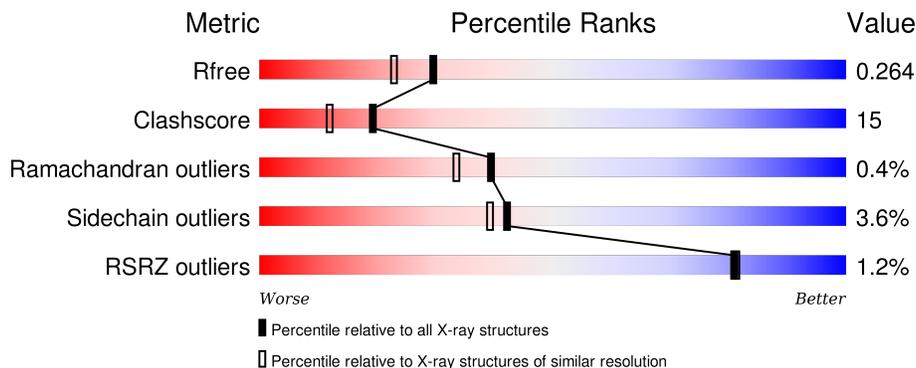
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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.

Mol	Chain	Length	Quality of chain
1	A	393	 3% 70% 26% ..
1	B	393	 % 73% 24% ..
1	C	393	 % 76% 21% ..
1	D	393	 % 65% 32% ..

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	STL	D	2003	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13154 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 Chalcone synthase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	388	2974	1892	505	560	17	0	0	0
1	B	388	2974	1892	505	560	17	0	0	0
1	C	388	2974	1892	505	560	17	0	0	0
1	D	388	2974	1892	505	560	17	0	0	0

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	CLONING ARTIFACT	UNP P30074
A	-2	SER	-	CLONING ARTIFACT	UNP P30074
A	-1	HIS	-	CLONING ARTIFACT	UNP P30074
A	0	GLY	-	CLONING ARTIFACT	UNP P30074
A	96	ALA	ASP	ENGINEERED	UNP P30074
A	98	LEU	VAL	ENGINEERED	UNP P30074
A	99	ALA	VAL	ENGINEERED	UNP P30074
A	100	MET	VAL	ENGINEERED	UNP P30074
A	131	SER	THR	ENGINEERED	UNP P30074
A	133	THR	SER	ENGINEERED	UNP P30074
A	134	THR	GLY	ENGINEERED	UNP P30074
A	135	PRO	VAL	ENGINEERED	UNP P30074
A	137	LEU	MET	ENGINEERED	UNP P30074
A	157	VAL	TYR	ENGINEERED	UNP P30074
A	158	GLY	MET	ENGINEERED	UNP P30074
A	159	VAL	MET	ENGINEERED	UNP P30074
A	160	PHE	TYR	ENGINEERED	UNP P30074
A	162	HIS	GLN	ENGINEERED	UNP P30074
A	268	LYS	LEU	ENGINEERED	UNP P30074
A	269	GLY	LYS	ENGINEERED	UNP P30074
A	270	ALA	ASP	ENGINEERED	UNP P30074

Continued on next page...

Continued from previous page...

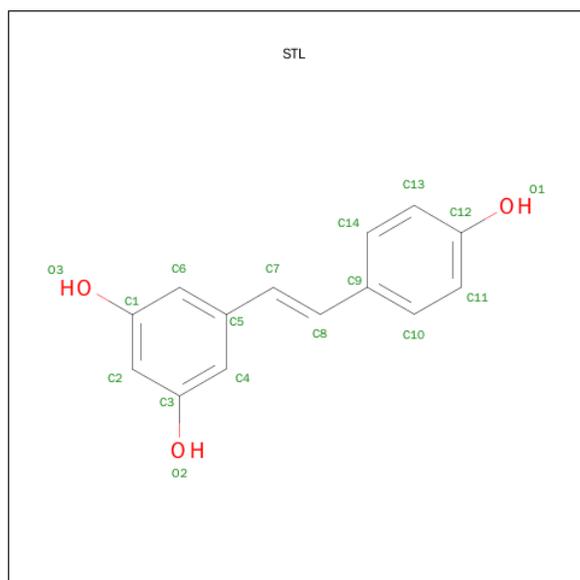
Chain	Residue	Modelled	Actual	Comment	Reference
A	273	ASP	GLY	ENGINEERED	UNP P30074
B	-3	GLY	-	CLONING ARTIFACT	UNP P30074
B	-2	SER	-	CLONING ARTIFACT	UNP P30074
B	-1	HIS	-	CLONING ARTIFACT	UNP P30074
B	0	GLY	-	CLONING ARTIFACT	UNP P30074
B	96	ALA	ASP	ENGINEERED	UNP P30074
B	98	LEU	VAL	ENGINEERED	UNP P30074
B	99	ALA	VAL	ENGINEERED	UNP P30074
B	100	MET	VAL	ENGINEERED	UNP P30074
B	131	SER	THR	ENGINEERED	UNP P30074
B	133	THR	SER	ENGINEERED	UNP P30074
B	134	THR	GLY	ENGINEERED	UNP P30074
B	135	PRO	VAL	ENGINEERED	UNP P30074
B	137	LEU	MET	ENGINEERED	UNP P30074
B	157	VAL	TYR	ENGINEERED	UNP P30074
B	158	GLY	MET	ENGINEERED	UNP P30074
B	159	VAL	MET	ENGINEERED	UNP P30074
B	160	PHE	TYR	ENGINEERED	UNP P30074
B	162	HIS	GLN	ENGINEERED	UNP P30074
B	268	LYS	LEU	ENGINEERED	UNP P30074
B	269	GLY	LYS	ENGINEERED	UNP P30074
B	270	ALA	ASP	ENGINEERED	UNP P30074
B	273	ASP	GLY	ENGINEERED	UNP P30074
C	-3	GLY	-	CLONING ARTIFACT	UNP P30074
C	-2	SER	-	CLONING ARTIFACT	UNP P30074
C	-1	HIS	-	CLONING ARTIFACT	UNP P30074
C	0	GLY	-	CLONING ARTIFACT	UNP P30074
C	96	ALA	ASP	ENGINEERED	UNP P30074
C	98	LEU	VAL	ENGINEERED	UNP P30074
C	99	ALA	VAL	ENGINEERED	UNP P30074
C	100	MET	VAL	ENGINEERED	UNP P30074
C	131	SER	THR	ENGINEERED	UNP P30074
C	133	THR	SER	ENGINEERED	UNP P30074
C	134	THR	GLY	ENGINEERED	UNP P30074
C	135	PRO	VAL	ENGINEERED	UNP P30074
C	137	LEU	MET	ENGINEERED	UNP P30074
C	157	VAL	TYR	ENGINEERED	UNP P30074
C	158	GLY	MET	ENGINEERED	UNP P30074
C	159	VAL	MET	ENGINEERED	UNP P30074
C	160	PHE	TYR	ENGINEERED	UNP P30074
C	162	HIS	GLN	ENGINEERED	UNP P30074
C	268	LYS	LEU	ENGINEERED	UNP P30074

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	269	GLY	LYS	ENGINEERED	UNP P30074
C	270	ALA	ASP	ENGINEERED	UNP P30074
C	273	ASP	GLY	ENGINEERED	UNP P30074
D	-3	GLY	-	CLONING ARTIFACT	UNP P30074
D	-2	SER	-	CLONING ARTIFACT	UNP P30074
D	-1	HIS	-	CLONING ARTIFACT	UNP P30074
D	0	GLY	-	CLONING ARTIFACT	UNP P30074
D	96	ALA	ASP	ENGINEERED	UNP P30074
D	98	LEU	VAL	ENGINEERED	UNP P30074
D	99	ALA	VAL	ENGINEERED	UNP P30074
D	100	MET	VAL	ENGINEERED	UNP P30074
D	131	SER	THR	ENGINEERED	UNP P30074
D	133	THR	SER	ENGINEERED	UNP P30074
D	134	THR	GLY	ENGINEERED	UNP P30074
D	135	PRO	VAL	ENGINEERED	UNP P30074
D	137	LEU	MET	ENGINEERED	UNP P30074
D	157	VAL	TYR	ENGINEERED	UNP P30074
D	158	GLY	MET	ENGINEERED	UNP P30074
D	159	VAL	MET	ENGINEERED	UNP P30074
D	160	PHE	TYR	ENGINEERED	UNP P30074
D	162	HIS	GLN	ENGINEERED	UNP P30074
D	268	LYS	LEU	ENGINEERED	UNP P30074
D	269	GLY	LYS	ENGINEERED	UNP P30074
D	270	ALA	ASP	ENGINEERED	UNP P30074
D	273	ASP	GLY	ENGINEERED	UNP P30074

- Molecule 2 is RESVERATROL (three-letter code: STL) (formula: C₁₄H₁₂O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			17	14	3		
2	B	1	Total	C	O	0	0
			17	14	3		
2	C	1	Total	C	O	0	0
			17	14	3		
2	D	1	Total	C	O	0	0
			17	14	3		

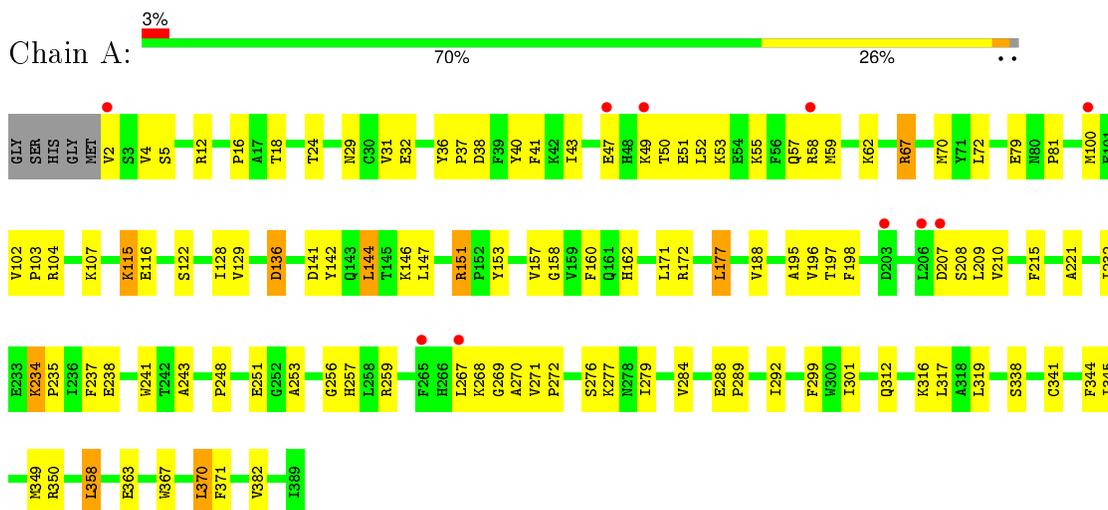
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	275	Total	O	0	0
			275	275		
3	B	310	Total	O	0	0
			310	310		
3	C	315	Total	O	0	0
			315	315		
3	D	290	Total	O	0	0
			290	290		

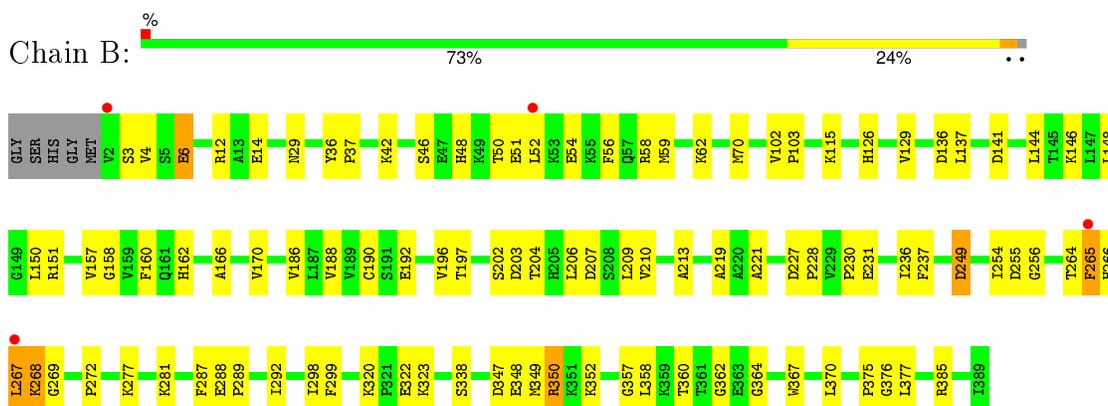
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

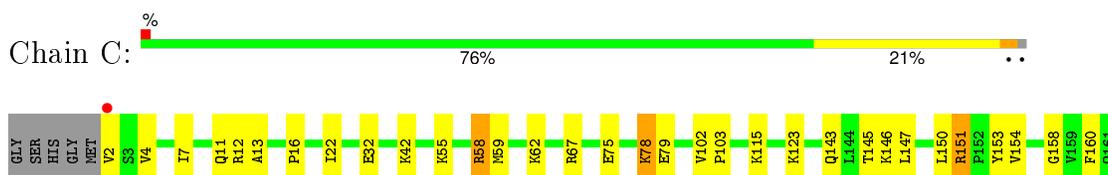
• Molecule 1: Chalcone synthase 2

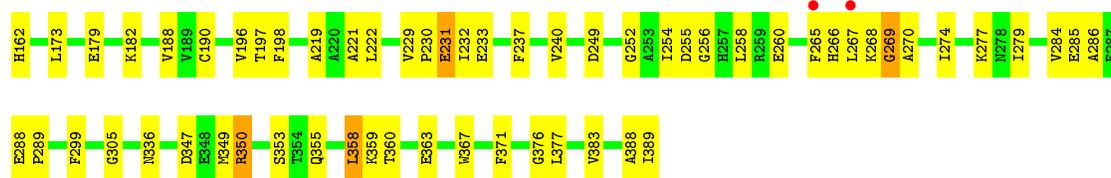


• Molecule 1: Chalcone synthase 2

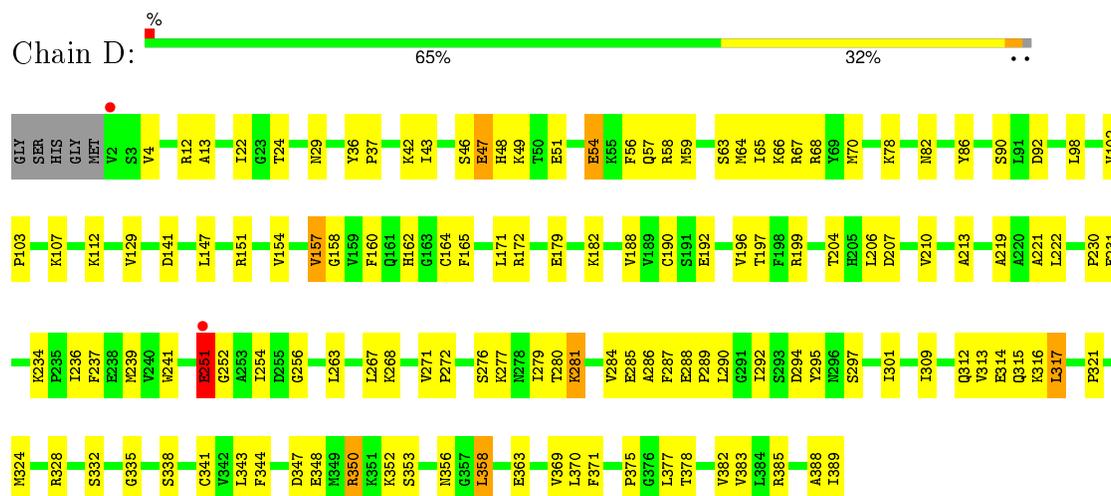


• Molecule 1: Chalcone synthase 2





• Molecule 1: Chalcone synthase 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	64.33Å 71.72Å 85.75Å 111.39° 91.61° 90.07°	Depositor
Resolution (Å)	49.34 - 2.00 49.34 - 2.00	Depositor EDS
% Data completeness (in resolution range)	97.0 (49.34-2.00) 85.2 (49.34-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.38 (at 2.00Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.203 , 0.264 0.203 , 0.264	Depositor DCC
R_{free} test set	4706 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	23.8	Xtrriage
Anisotropy	0.725	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 49.6	EDS
Estimated twinning fraction	0.119 for h,-k,-l	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Outliers	0 of 93036 reflections	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13154	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.43% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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: STL

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	0.49	0/3032	0.72	0/4106
1	B	0.49	0/3032	0.72	1/4106 (0.0%)
1	C	0.50	0/3032	0.71	0/4106
1	D	0.49	0/3032	0.72	0/4106
All	All	0.49	0/12128	0.72	1/16424 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	137	LEU	N-CA-C	-5.11	97.20	111.00

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	2974	0	3021	100	0
1	B	2974	0	3021	92	0
1	C	2974	0	3021	83	0
1	D	2974	0	3021	118	0
2	A	17	0	10	2	0
2	B	17	0	10	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	17	0	9	1	0
2	D	17	0	9	2	0
3	A	275	0	0	16	0
3	B	310	0	0	16	0
3	C	315	0	0	9	0
3	D	290	0	0	12	0
All	All	13154	0	12122	364	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 15.

The worst 5 of 364 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:58:ARG:HG2	1:B:62:LYS:HE2	1.51	0.91
1:D:171:LEU:HD22	1:D:239:MET:HE3	1.53	0.91
1:A:284:VAL:HG13	1:A:288:GLU:HG3	1.55	0.88
1:C:353:SER:HA	1:C:358:LEU:HD22	1.56	0.86
1:D:281:LYS:O	1:D:281:LYS:HE2	1.76	0.85

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	386/393 (98%)	362 (94%)	21 (5%)	3 (1%)	24 15
1	B	386/393 (98%)	369 (96%)	17 (4%)	0	100 100
1	C	386/393 (98%)	370 (96%)	15 (4%)	1 (0%)	46 41
1	D	386/393 (98%)	372 (96%)	12 (3%)	2 (0%)	34 26
All	All	1544/1572 (98%)	1473 (95%)	65 (4%)	6 (0%)	39 33

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	269	GLY
1	D	251	GLU
1	A	50	THR
1	C	269	GLY
1	D	252	GLY

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	321/324 (99%)	310 (97%)	11 (3%)	44	41
1	B	321/324 (99%)	310 (97%)	11 (3%)	44	41
1	C	321/324 (99%)	310 (97%)	11 (3%)	44	41
1	D	321/324 (99%)	308 (96%)	13 (4%)	38	33
All	All	1284/1296 (99%)	1238 (96%)	46 (4%)	42	39

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	350	ARG
1	C	115	LYS
1	D	281	LYS
1	B	358	LEU
1	C	58	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	162	HIS
1	D	325	ASN
1	D	162	HIS
1	B	162	HIS
1	D	119	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 carbohydrates 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	STL	A	2000	-	18,18,18	1.81	10 (55%)	24,24,24	1.05	1 (4%)
2	STL	B	2001	-	18,18,18	1.68	6 (33%)	24,24,24	0.94	0
2	STL	C	2002	-	18,18,18	1.67	5 (27%)	24,24,24	0.95	1 (4%)
2	STL	D	2003	-	18,18,18	1.69	5 (27%)	24,24,24	0.94	1 (4%)

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	STL	A	2000	-	-	0/5/5/5	0/2/2/2
2	STL	B	2001	-	-	0/5/5/5	0/2/2/2
2	STL	C	2002	-	-	0/5/5/5	0/2/2/2
2	STL	D	2003	-	-	0/5/5/5	0/2/2/2

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2002	STL	C6-C5	2.00	1.43	1.39
2	D	2003	STL	C14-C9	2.02	1.43	1.39
2	A	2000	STL	C13-C12	2.05	1.43	1.38
2	B	2001	STL	C14-C9	2.06	1.43	1.39
2	A	2000	STL	C11-C12	2.06	1.43	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	2002	STL	C3-C2-C1	2.03	122.08	119.17
2	A	2000	STL	C3-C2-C1	2.07	122.13	119.17
2	D	2003	STL	C3-C2-C1	2.07	122.14	119.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2000	STL	2	0
2	B	2001	STL	1	0
2	C	2002	STL	1	0
2	D	2003	STL	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 [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	388/393 (98%)	-0.32	10 (2%) 59 60	15, 27, 53, 65	0
1	B	388/393 (98%)	-0.36	4 (1%) 84 84	14, 27, 50, 61	0
1	C	388/393 (98%)	-0.39	3 (0%) 87 88	15, 28, 45, 61	0
1	D	388/393 (98%)	-0.38	2 (0%) 91 92	16, 28, 47, 61	0
All	All	1552/1572 (98%)	-0.36	19 (1%) 81 81	14, 28, 49, 65	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	267	LEU	3.5
1	C	267	LEU	3.2
1	C	2	VAL	3.0
1	B	265	PHE	2.9
1	A	58	ARG	2.9

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	STL	D	2003	17/17	0.88	0.19	2.95	24,32,49,53	0
2	STL	A	2000	17/17	0.89	0.15	1.85	32,34,43,45	0
2	STL	C	2002	17/17	0.87	0.14	1.47	22,29,41,49	0
2	STL	B	2001	17/17	0.90	0.15	1.32	22,30,47,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.