W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Mar 13, 2024 – 06:23 PM JST

|		
PDB ID	:	5J1M
Title	:	Crystal structure of Csd1-Csd2 dimer II
Authors	:	An, D.R.; Suh, S.W.
Deposited on	:	2016-03-29
Resolution	:	2.35 Å(reported)

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 2.35 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.
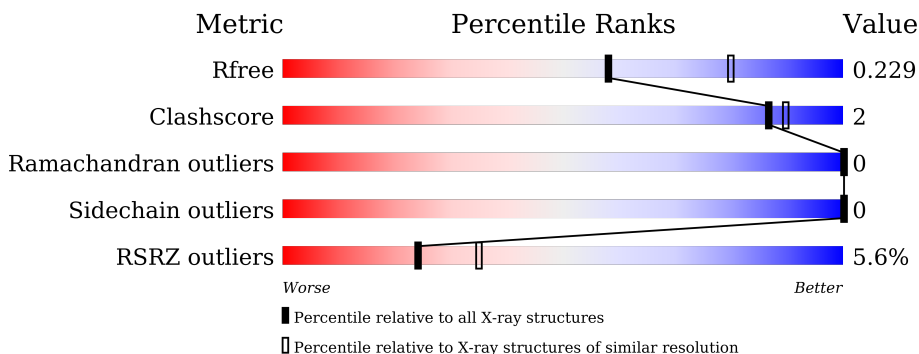
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	216	6% ... 74% · 23%
1	C	216	7% ... 69% · 26%
2	B	189	2% ... 87% 6% 7%
2	D	189	3% ... 86% 7% 6%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 5560 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 ToxR-activated gene (TagE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	166	Total	C	N	O	S	0	0	0
			1328	860	231	232	5			
1	C	159	Total	C	N	O	S	0	0	0
			1271	822	220	224	5			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	105	MET	-	initiating methionine	UNP O26068
A	106	GLY	-	expression tag	UNP O26068
A	107	SER	-	expression tag	UNP O26068
A	108	SER	-	expression tag	UNP O26068
A	109	HIS	-	expression tag	UNP O26068
A	110	HIS	-	expression tag	UNP O26068
A	111	HIS	-	expression tag	UNP O26068
A	112	HIS	-	expression tag	UNP O26068
A	113	HIS	-	expression tag	UNP O26068
A	114	HIS	-	expression tag	UNP O26068
A	115	SER	-	expression tag	UNP O26068
A	116	SER	-	expression tag	UNP O26068
A	117	GLY	-	expression tag	UNP O26068
A	118	LEU	-	expression tag	UNP O26068
A	119	VAL	-	expression tag	UNP O26068
A	120	PRO	-	expression tag	UNP O26068
A	121	ARG	-	expression tag	UNP O26068
A	122	GLY	-	expression tag	UNP O26068
A	123	SER	-	expression tag	UNP O26068
A	124	HIS	-	expression tag	UNP O26068
A	313	LEU	-	expression tag	UNP O26068
A	314	GLU	-	expression tag	UNP O26068
A	315	HIS	-	expression tag	UNP O26068
A	316	HIS	-	expression tag	UNP O26068
A	317	HIS	-	expression tag	UNP O26068

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	318	HIS	-	expression tag	UNP O26068
A	319	HIS	-	expression tag	UNP O26068
A	320	HIS	-	expression tag	UNP O26068
C	105	MET	-	initiating methionine	UNP O26068
C	106	GLY	-	expression tag	UNP O26068
C	107	SER	-	expression tag	UNP O26068
C	108	SER	-	expression tag	UNP O26068
C	109	HIS	-	expression tag	UNP O26068
C	110	HIS	-	expression tag	UNP O26068
C	111	HIS	-	expression tag	UNP O26068
C	112	HIS	-	expression tag	UNP O26068
C	113	HIS	-	expression tag	UNP O26068
C	114	HIS	-	expression tag	UNP O26068
C	115	SER	-	expression tag	UNP O26068
C	116	SER	-	expression tag	UNP O26068
C	117	GLY	-	expression tag	UNP O26068
C	118	LEU	-	expression tag	UNP O26068
C	119	VAL	-	expression tag	UNP O26068
C	120	PRO	-	expression tag	UNP O26068
C	121	ARG	-	expression tag	UNP O26068
C	122	GLY	-	expression tag	UNP O26068
C	123	SER	-	expression tag	UNP O26068
C	124	HIS	-	expression tag	UNP O26068
C	313	LEU	-	expression tag	UNP O26068
C	314	GLU	-	expression tag	UNP O26068
C	315	HIS	-	expression tag	UNP O26068
C	316	HIS	-	expression tag	UNP O26068
C	317	HIS	-	expression tag	UNP O26068
C	318	HIS	-	expression tag	UNP O26068
C	319	HIS	-	expression tag	UNP O26068
C	320	HIS	-	expression tag	UNP O26068

- Molecule 2 is a protein called ToxR-activated gene (TagE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1405	911	238	254	2			
2	D	177	Total	C	N	O	S	0	0	0
			1414	916	239	257	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	120	MET	-	initiating methionine	UNP O26069
D	120	MET	-	initiating methionine	UNP O26069

• Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0

• Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	25	Total 25	O 25	0	0
4	B	47	Total 47	O 47	0	0
4	C	22	Total 22	O 22	0	0
4	D	46	Total 46	O 46	0	0

3 Residue-property plots ⓘ

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.
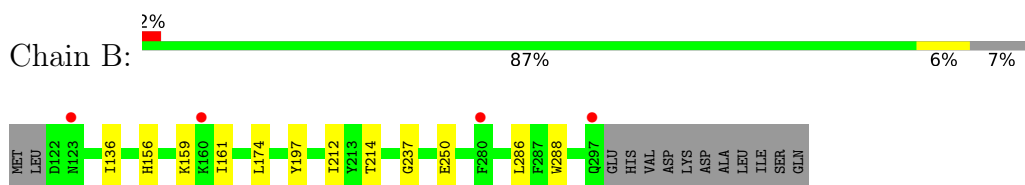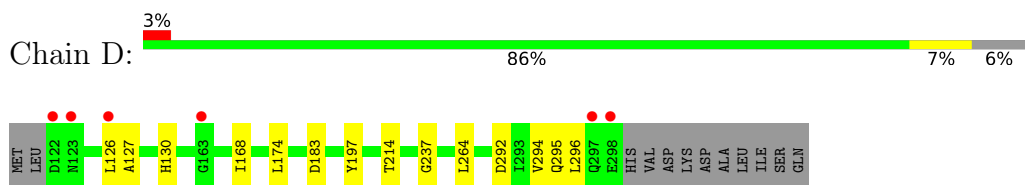
- Molecule 1: ToxR-activated gene (TagE)
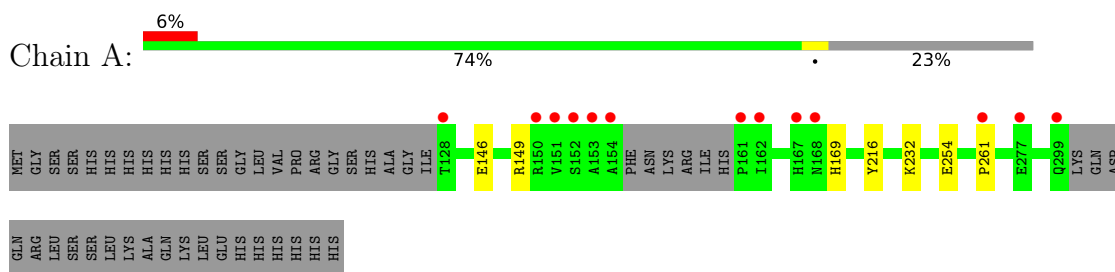
- Molecule 1: ToxR-activated gene (TagE)
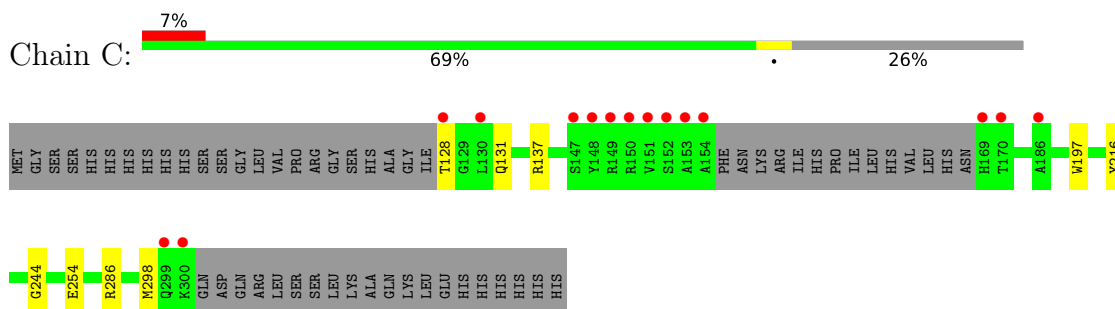
- Molecule 2: ToxR-activated gene (TagE)

- Molecule 2: ToxR-activated gene (TagE)

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	53.15Å 80.01Å 74.59Å	Depositor
a, b, c, α, β, γ	90.00° 104.43° 90.00°	
Resolution (Å)	50.01 – 2.35	Depositor
	31.59 – 2.35	EDS
% Data completeness	99.7 (50.01-2.35)	Depositor
(in resolution range)	99.7 (31.59-2.35)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.38 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.168 , 0.233	Depositor
	0.171 , 0.229	DCC
R_{free} test set	1246 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	38.6	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.3	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5560	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.15% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.54	0/1364	0.71	0/1844
1	C	0.50	0/1304	0.71	0/1760
2	B	0.55	0/1440	0.71	0/1947
2	D	0.55	0/1449	0.72	1/1959 (0.1%)
All	All	0.54	0/5557	0.72	1/7510 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	183	ASP	CB-CG-OD1	5.18	122.96	118.30

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	1328	0	1330	4	0
1	C	1271	0	1273	6	0
2	B	1405	0	1406	7	0
2	D	1414	0	1412	8	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	25	0	0	0	0
4	B	47	0	0	1	0
4	C	22	0	0	0	1
4	D	46	0	0	0	1
All	All	5560	0	5421	23	1

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 2.

The worst 5 of 23 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:B:156:HIS:HB3	2:B:159:LYS:O	2.03	0.59
2:D:197:TYR:O	2:D:214:THR:HG23	2.06	0.54
2:B:212:ILE:HG22	2:B:250:GLU:HB2	1.90	0.54
2:D:292:ASP:O	2:D:295:GLN:HG2	2.08	0.54
1:A:216:TYR:HB2	1:A:254:GLU:HB2	1.92	0.52

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:C:511:HOH:O	4:D:441:HOH:O[2_656]	1.65	0.55

5.3   Torsion angles   ⓘ

5.3.1   Protein backbone   ⓘ

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	162/216 (75%)	152 (94%)	10 (6%)	0	100	100
1	C	155/216 (72%)	145 (94%)	10 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	174/189 (92%)	169 (97%)	5 (3%)	0	100	100
2	D	175/189 (93%)	167 (95%)	8 (5%)	0	100	100
All	All	666/810 (82%)	633 (95%)	33 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	144/188 (77%)	144 (100%)	0	100	100
1	C	137/188 (73%)	137 (100%)	0	100	100
2	B	152/164 (93%)	152 (100%)	0	100	100
2	D	153/164 (93%)	153 (100%)	0	100	100
All	All	586/704 (83%)	586 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	169	HIS
2	B	285	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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 (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(Å²)	Q<0.9
1	A	166/216 (76%)	0.36	13 (7%)	13	19	28, 41, 83, 118	0
1	C	159/216 (73%)	0.47	15 (9%)	8	13	29, 45, 90, 123	0
2	B	176/189 (93%)	-0.02	4 (2%)	60	70	25, 38, 65, 100	0
2	D	177/189 (93%)	0.10	6 (3%)	45	57	25, 37, 59, 99	0
All	All	678/810 (83%)	0.22	38 (5%)	24	35	25, 39, 79, 123	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	152	SER	8.5
1	A	151	VAL	7.5
1	C	149	ARG	7.1
1	A	154	ALA	6.4
1	C	154	ALA	6.3

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(Å²)	Q<0.9
3	ZN	C	401	1/1	0.96	0.03	74,74,74,74	0
3	ZN	A	401	1/1	1.00	0.02	43,43,43,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.