



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 13, 2018 – 01:14 pm GMT

PDB ID : 2Q4M  
Title : Ensemble refinement of the crystal structure of protein from Arabidopsis thaliana At5g01750  
Authors : Levin, E.J.; Kondrashov, D.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2007-05-31  
Resolution : 1.70 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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:

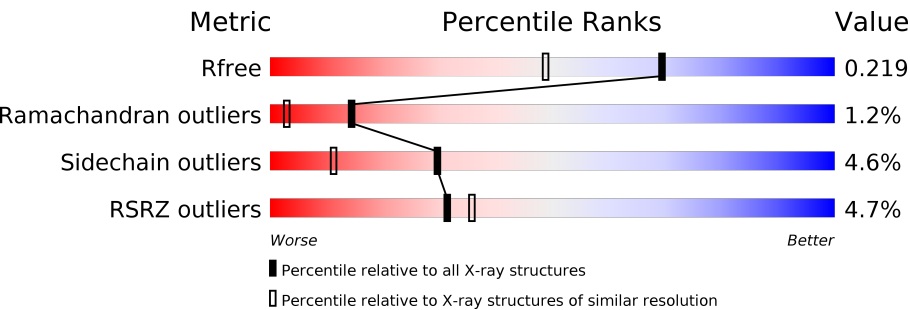
MolProbity	:	4.02b-467
Mogul	:	1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

# 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 1.70 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))
$R_{free}$	111664	3793 (1.70-1.70)
Ramachandran outliers	120053	4100 (1.70-1.70)
Sidechain outliers	120020	4100 (1.70-1.70)
RSRZ outliers	108989	3718 (1.70-1.70)

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  $\geq 3$ , 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	1-A	217	<div><div>3%</div><div>71%</div><div>•</div><div>25%</div></div>
1	10-A	217	<div><div>3%</div><div>70%</div><div>5%</div><div>25%</div></div>
1	11-A	217	<div><div>3%</div><div>71%</div><div>•</div><div>25%</div></div>
1	12-A	217	<div><div>3%</div><div>72%</div><div>•</div><div>25%</div></div>
1	13-A	217	<div><div>3%</div><div>69%</div><div>• •</div><div>25%</div></div>
1	14-A	217	<div><div>3%</div><div>71%</div><div>• •</div><div>25%</div></div>
1	15-A	217	<div><div>3%</div><div>65%</div><div>10%</div><div>25%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	16-A	217	 3% 69% 5% 25%
1	2-A	217	 3% 69% 5% 25%
1	3-A	217	 3% 71% 5% 25%
1	4-A	217	 3% 71% 5% 25%
1	5-A	217	 3% 71% 5% 25%
1	6-A	217	 3% 71% 5% 25%
1	7-A	217	 3% 73% 5% 25%
1	8-A	217	 3% 71% 5% 25%
1	9-A	217	 3% 69% 6% 25%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 22688 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.

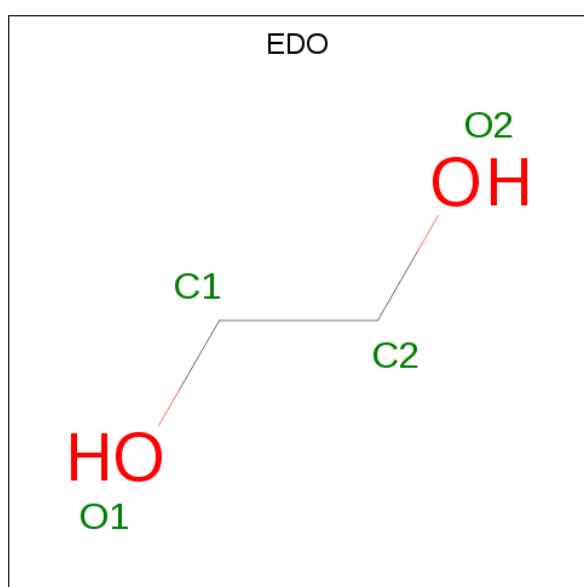
- Molecule 1 is a protein called Protein At5g01750.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	1-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	2-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	3-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	4-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	5-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	6-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	7-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	8-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	9-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	10-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	11-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	12-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	13-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	14-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	15-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			
1	16-A	162	Total	C	N	O	S	Se	0	0	0
			1287	820	228	234	3	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	40	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	46	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	47	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	95	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	98	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	124	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1
A	171	MSE	MET	MODIFIED RESIDUE	UNP Q9LZX1

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	1-A	1	Total C O 4 2 2	0	0
2	2-A	1	Total C O 4 2 2	0	0
2	3-A	1	Total C O 4 2 2	0	0
2	4-A	1	Total C O 4 2 2	0	0
2	5-A	1	Total C O 4 2 2	0	0
2	6-A	1	Total C O 4 2 2	0	0
2	7-A	1	Total C O 4 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	8-A	1	Total C O 4 2 2	0	0
2	9-A	1	Total C O 4 2 2	0	0
2	10-A	1	Total C O 4 2 2	0	0
2	11-A	1	Total C O 4 2 2	0	0
2	12-A	1	Total C O 4 2 2	0	0
2	13-A	1	Total C O 4 2 2	0	0
2	14-A	1	Total C O 4 2 2	0	0
2	15-A	1	Total C O 4 2 2	0	0
2	16-A	1	Total C O 4 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	1-A	127	Total O 127 127	0	0
3	2-A	127	Total O 127 127	0	0
3	3-A	127	Total O 127 127	0	0
3	4-A	127	Total O 127 127	0	0
3	5-A	127	Total O 127 127	0	0
3	6-A	127	Total O 127 127	0	0
3	7-A	127	Total O 127 127	0	0
3	8-A	127	Total O 127 127	0	0
3	9-A	127	Total O 127 127	0	0
3	10-A	127	Total O 127 127	0	0

*Continued on next page...*

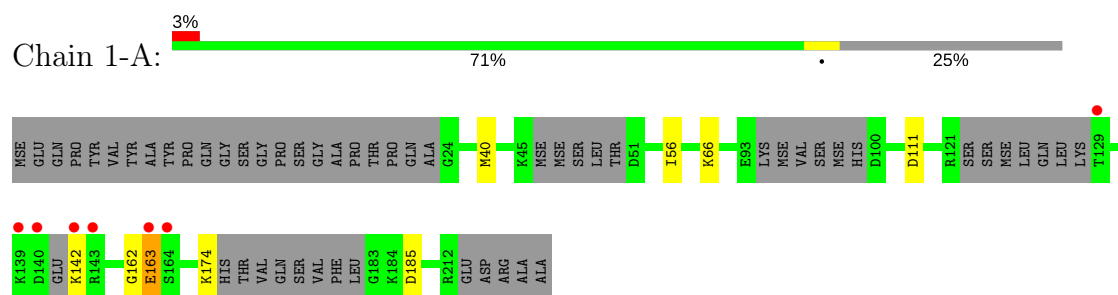
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	11-A	127	Total 127	O 127	0	0
3	12-A	127	Total 127	O 127	0	0
3	13-A	127	Total 127	O 127	0	0
3	14-A	127	Total 127	O 127	0	0
3	15-A	127	Total 127	O 127	0	0
3	16-A	127	Total 127	O 127	0	0

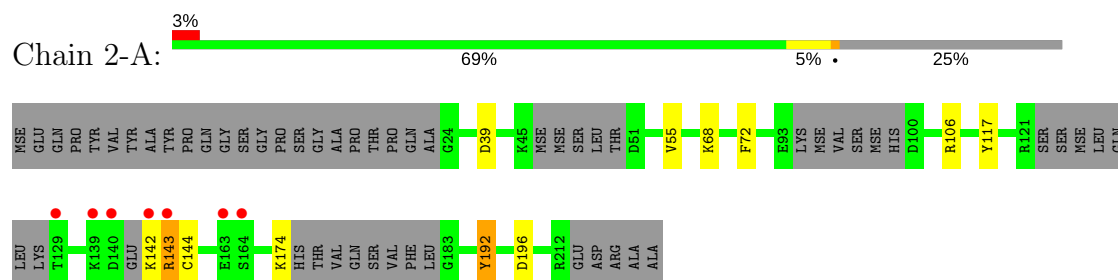
### 3 Residue-property plots [i](#)

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 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.

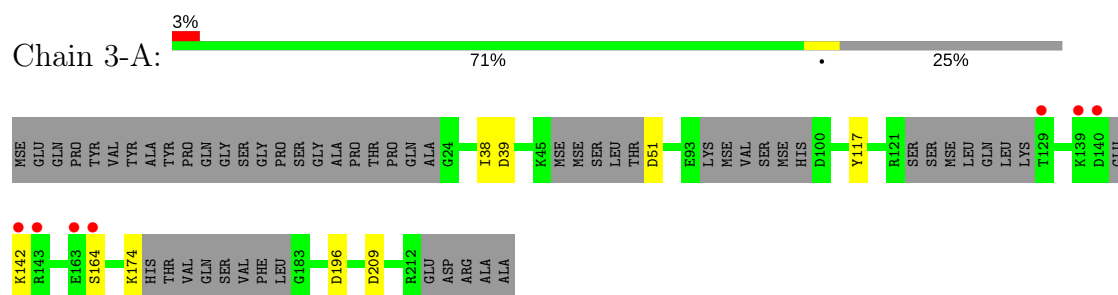
#### • Molecule 1: Protein At5g01750



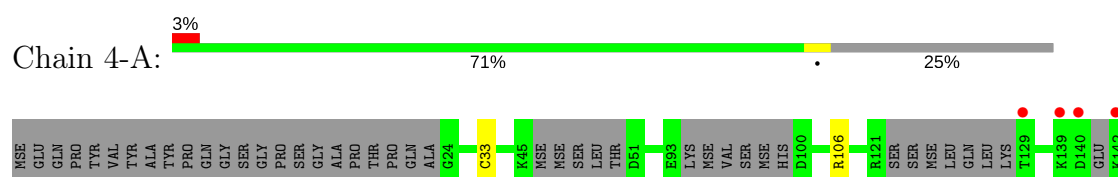
#### • Molecule 1: Protein At5g01750



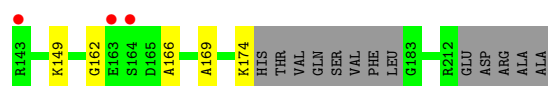
#### • Molecule 1: Protein At5g01750



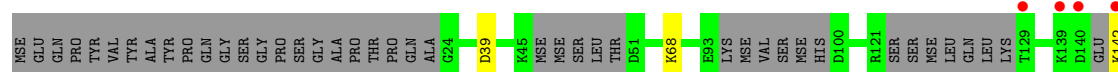
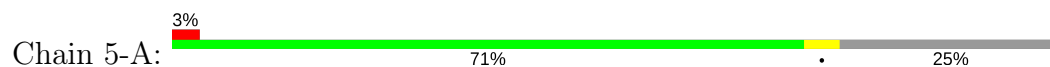
#### • Molecule 1: Protein At5g01750



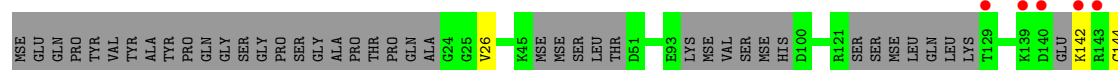
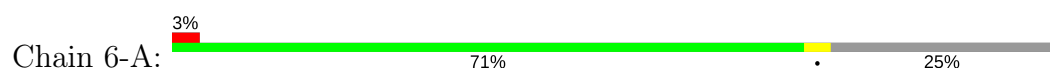




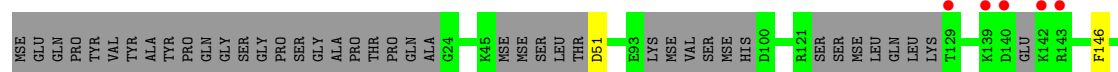
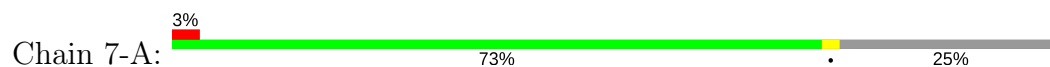
## ● Molecule 1: Protein At5g01750



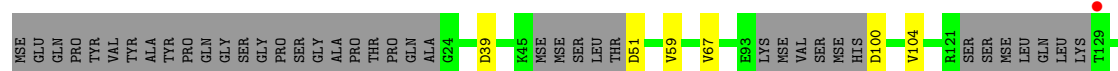
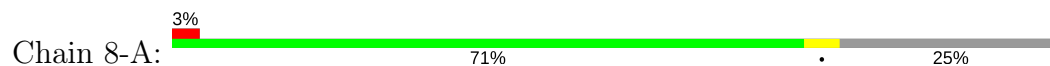
## ● Molecule 1: Protein At5g01750



## ● Molecule 1: Protein At5g01750

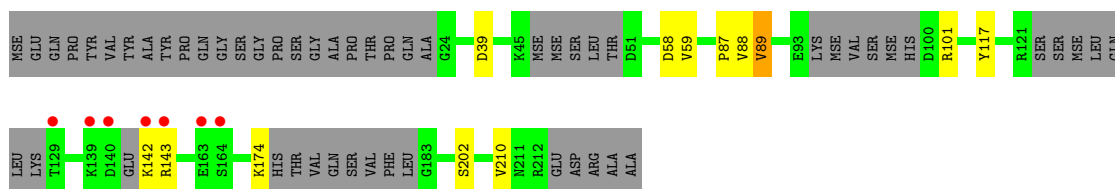


## ● Molecule 1: Protein At5g01750

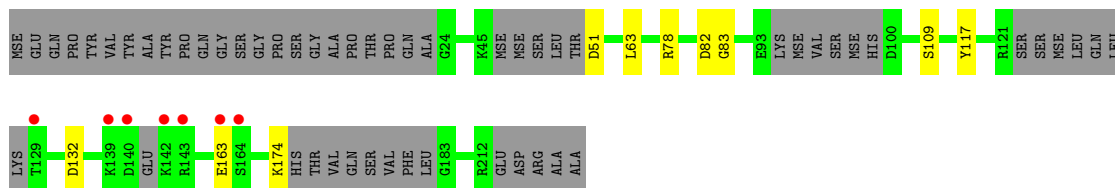
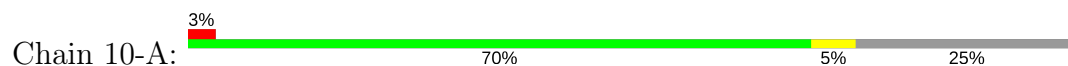


## ● Molecule 1: Protein At5g01750

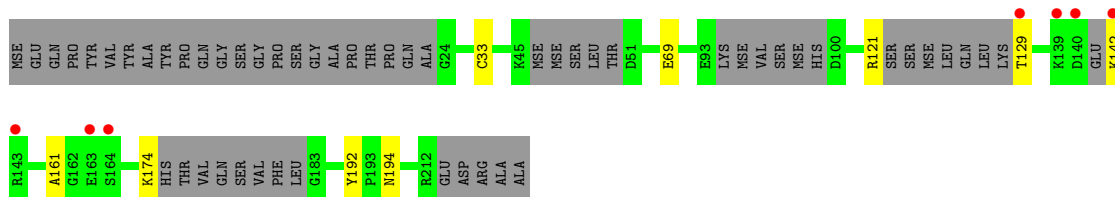




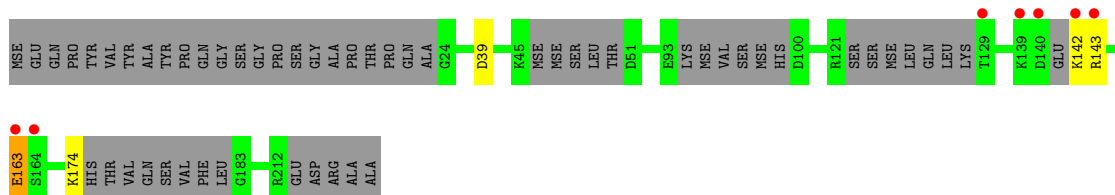
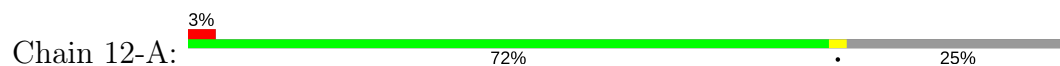
• Molecule 1: Protein At5g01750



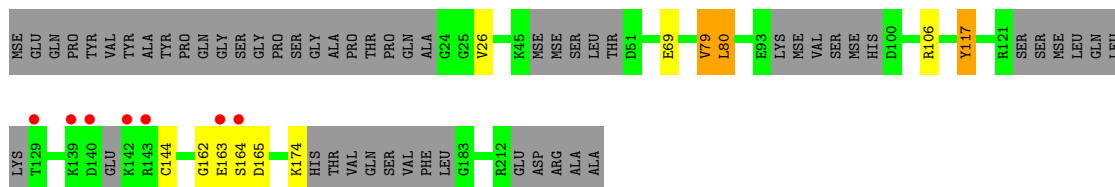
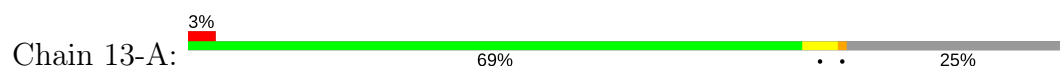
• Molecule 1: Protein At5g01750



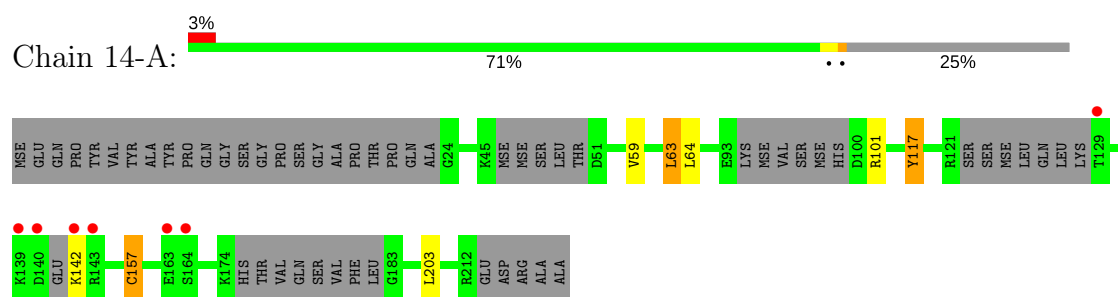
• Molecule 1: Protein At5g01750



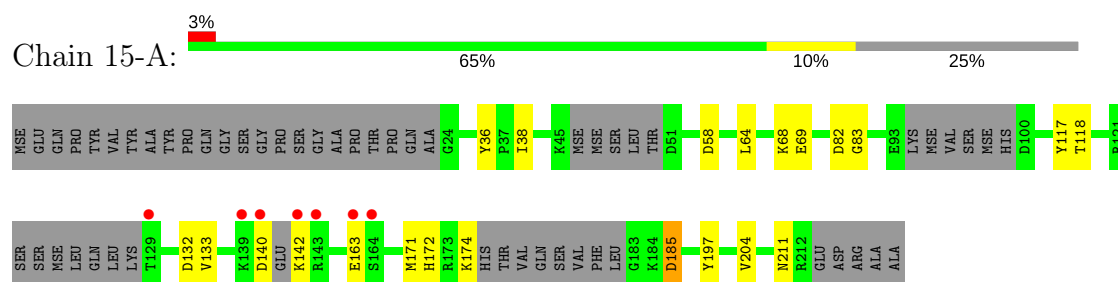
• Molecule 1: Protein At5g01750



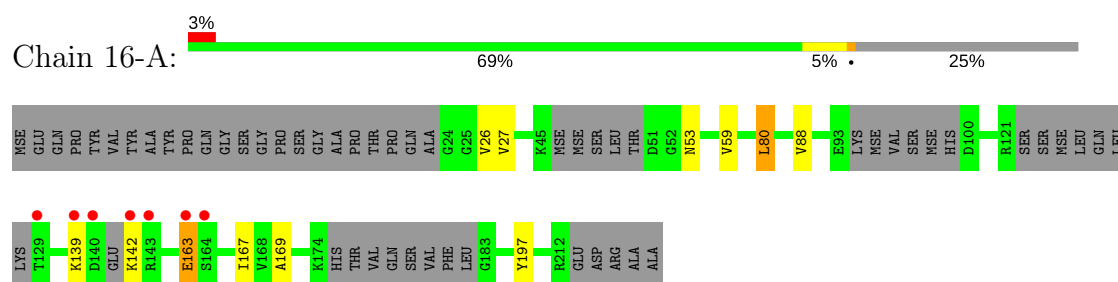
• Molecule 1: Protein At5g01750



• Molecule 1: Protein At5g01750



• Molecule 1: Protein At5g01750



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	41.07Å 57.50Å 75.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.68 – 1.70 37.68 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (37.68-1.70) 99.8 (37.68-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.66 (at 1.70Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.158 , 0.207 0.169 , 0.219	Depositor DCC
$R_{free}$ test set	1037 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.1	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 71.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	22688	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.64% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: EDO

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	1-A	0.90	0/1306	0.92	1/1756 (0.1%)
1	2-A	0.93	1/1306 (0.1%)	0.96	2/1756 (0.1%)
1	3-A	0.88	1/1306 (0.1%)	0.96	1/1756 (0.1%)
1	4-A	0.92	0/1306	0.97	0/1756
1	5-A	0.88	0/1306	0.91	0/1756
1	6-A	0.90	0/1306	0.92	0/1756
1	7-A	0.91	0/1306	0.92	0/1756
1	8-A	0.91	2/1306 (0.2%)	0.91	0/1756
1	9-A	0.94	1/1306 (0.1%)	0.97	0/1756
1	10-A	0.90	1/1306 (0.1%)	0.93	2/1756 (0.1%)
1	11-A	0.92	1/1306 (0.1%)	0.92	2/1756 (0.1%)
1	12-A	0.90	0/1306	0.94	0/1756
1	13-A	1.10	3/1306 (0.2%)	1.08	3/1756 (0.2%)
1	14-A	1.16	3/1306 (0.2%)	1.17	4/1756 (0.2%)
1	15-A	1.11	4/1306 (0.3%)	1.23	6/1756 (0.3%)
1	16-A	1.11	2/1306 (0.2%)	1.13	2/1756 (0.1%)
All	All	0.96	19/20896 (0.1%)	1.00	23/28096 (0.1%)

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	2-A	0	2
1	6-A	0	1
1	11-A	0	1
1	15-A	0	2
1	16-A	0	1
All	All	0	7

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	13-A	79	VAL	CA-CB	9.51	1.74	1.54
1	14-A	157	CYS	CB-SG	8.85	1.97	1.82
1	14-A	117	TYR	CD1-CE1	7.77	1.50	1.39
1	15-A	117	TYR	CE1-CZ	6.84	1.47	1.38
1	13-A	117	TYR	CD1-CE1	6.84	1.49	1.39
1	9-A	117	TYR	CD1-CE1	6.55	1.49	1.39
1	8-A	104	VAL	CB-CG2	6.27	1.66	1.52
1	15-A	133	VAL	CB-CG2	6.08	1.65	1.52
1	3-A	117	TYR	CD1-CE1	6.06	1.48	1.39
1	11-A	33	CYS	CB-SG	5.92	1.92	1.82
1	14-A	203	LEU	CG-CD1	5.57	1.72	1.51
1	8-A	67	VAL	CB-CG2	5.42	1.64	1.52
1	10-A	117	TYR	CD1-CE1	5.31	1.47	1.39
1	16-A	88	VAL	CB-CG2	5.26	1.63	1.52
1	15-A	117	TYR	CD2-CE2	-5.19	1.31	1.39
1	13-A	79	VAL	CA-C	5.06	1.66	1.52
1	15-A	117	TYR	CD1-CE1	5.05	1.47	1.39
1	16-A	167	ILE	CB-CG2	5.03	1.68	1.52
1	2-A	117	TYR	CD1-CE1	5.01	1.46	1.39

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	13-A	80	LEU	CA-CB-CG	-10.25	91.73	115.30
1	15-A	117	TYR	CB-CG-CD2	-9.24	115.45	121.00
1	1-A	40	MSE	CB-CG-SE	-8.26	87.93	112.70
1	15-A	117	TYR	OH-CZ-CE2	-7.01	101.18	120.10
1	16-A	80	LEU	CA-CB-CG	-6.82	99.61	115.30
1	14-A	157	CYS	CA-CB-SG	6.81	126.27	114.00
1	3-A	196	ASP	CB-CG-OD1	-6.74	112.24	118.30
1	14-A	117	TYR	CB-CG-CD2	-6.26	117.24	121.00
1	2-A	106	ARG	NE-CZ-NH1	6.06	123.33	120.30
1	2-A	143	ARG	N-CA-C	6.05	127.32	111.00
1	10-A	63	LEU	CA-CB-CG	-6.01	101.47	115.30
1	14-A	63	LEU	CA-CB-CG	-5.95	101.63	115.30
1	15-A	117	TYR	CG-CD1-CE1	-5.83	116.64	121.30
1	13-A	69	GLU	N-CA-C	-5.72	95.55	111.00
1	15-A	69	GLU	N-CA-C	-5.60	95.88	111.00
1	10-A	78	ARG	NE-CZ-NH2	-5.59	117.50	120.30
1	15-A	117	TYR	CE1-CZ-OH	5.51	134.99	120.10
1	16-A	142	LYS	CD-CE-NZ	5.29	123.86	111.70
1	15-A	171	MSE	CB-CG-SE	-5.25	96.95	112.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	13-A	117	TYR	CB-CG-CD2	-5.15	117.91	121.00
1	14-A	117	TYR	CG-CD1-CE1	-5.13	117.20	121.30
1	11-A	161	ALA	N-CA-C	-5.09	97.25	111.00
1	11-A	69	GLU	N-CA-C	-5.05	97.36	111.00

There are no chirality outliers.

All (7) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	11-A	192	TYR	Sidechain
1	15-A	197	TYR	Sidechain
1	15-A	36	TYR	Sidechain
1	16-A	197	TYR	Sidechain
1	2-A	192	TYR	Sidechain
1	2-A	72	PHE	Sidechain
1	6-A	192	TYR	Sidechain

## 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	1-A	1287	0	1289	0	0
1	2-A	1287	0	1289	0	0
1	3-A	1287	0	1289	0	0
1	4-A	1287	0	1289	0	0
1	5-A	1287	0	1289	0	0
1	6-A	1287	0	1289	0	0
1	7-A	1287	0	1289	0	0
1	8-A	1287	0	1289	0	0
1	9-A	1287	0	1289	0	0
1	10-A	1287	0	1289	0	0
1	11-A	1287	0	1289	0	0
1	12-A	1287	0	1289	0	0
1	13-A	1287	0	1289	0	0
1	14-A	1287	0	1289	0	0
1	15-A	1287	0	1289	0	0
1	16-A	1287	0	1289	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	1-A	4	0	6	0	0
2	2-A	4	0	6	0	0
2	3-A	4	0	6	0	0
2	4-A	4	0	6	0	0
2	5-A	4	0	6	0	0
2	6-A	4	0	6	0	0
2	7-A	4	0	6	0	0
2	8-A	4	0	6	0	0
2	9-A	4	0	6	0	0
2	10-A	4	0	6	0	0
2	11-A	4	0	6	0	0
2	12-A	4	0	6	0	0
2	13-A	4	0	6	0	0
2	14-A	4	0	6	0	0
2	15-A	4	0	6	0	0
2	16-A	4	0	6	0	0
3	1-A	127	0	0	0	0
3	2-A	127	0	0	0	0
3	3-A	127	0	0	0	0
3	4-A	127	0	0	0	0
3	5-A	127	0	0	0	0
3	6-A	127	0	0	0	0
3	7-A	127	0	0	0	0
3	8-A	127	0	0	0	0
3	9-A	127	0	0	0	0
3	10-A	127	0	0	0	0
3	11-A	127	0	0	0	0
3	12-A	127	0	0	0	0
3	13-A	127	0	0	0	0
3	14-A	127	0	0	0	0
3	15-A	127	0	0	0	0
3	16-A	127	0	0	0	0
All	All	22688	0	20720	0	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). Clashscore could not be calculated for this entry.

There are no clashes within the asymmetric unit.

There are no symmetry-related clashes.



## 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	1-A	150/217 (69%)	139 (93%)	9 (6%)	2 (1%)	13	2
1	2-A	150/217 (69%)	140 (93%)	9 (6%)	1 (1%)	24	9
1	3-A	150/217 (69%)	136 (91%)	12 (8%)	2 (1%)	13	2
1	4-A	150/217 (69%)	132 (88%)	15 (10%)	3 (2%)	8	1
1	5-A	150/217 (69%)	139 (93%)	10 (7%)	1 (1%)	24	9
1	6-A	150/217 (69%)	142 (95%)	6 (4%)	2 (1%)	13	2
1	7-A	150/217 (69%)	140 (93%)	9 (6%)	1 (1%)	24	9
1	8-A	150/217 (69%)	141 (94%)	9 (6%)	0	100	100
1	9-A	150/217 (69%)	133 (89%)	14 (9%)	3 (2%)	8	1
1	10-A	150/217 (69%)	141 (94%)	7 (5%)	2 (1%)	13	2
1	11-A	150/217 (69%)	145 (97%)	5 (3%)	0	100	100
1	12-A	150/217 (69%)	137 (91%)	12 (8%)	1 (1%)	24	9
1	13-A	150/217 (69%)	137 (91%)	9 (6%)	4 (3%)	5	1
1	14-A	150/217 (69%)	139 (93%)	11 (7%)	0	100	100
1	15-A	150/217 (69%)	135 (90%)	11 (7%)	4 (3%)	5	1
1	16-A	150/217 (69%)	136 (91%)	11 (7%)	3 (2%)	8	1
All	All	2400/3472 (69%)	2212 (92%)	159 (7%)	29 (1%)	14	3

All (29) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-A	163	GLU
1	4-A	169	ALA
1	13-A	163	GLU
1	1-A	162	GLY
1	4-A	162	GLY
1	6-A	164	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	7-A	185	ASP
1	10-A	109	SER
1	12-A	163	GLU
1	13-A	162	GLY
1	13-A	164	SER
1	15-A	185	ASP
1	15-A	211	ASN
1	16-A	163	GLU
1	3-A	164	SER
1	4-A	166	ALA
1	9-A	87	PRO
1	10-A	83	GLY
1	15-A	83	GLY
1	15-A	204	VAL
1	16-A	169	ALA
1	5-A	164	SER
1	6-A	144	CYS
1	9-A	88	VAL
1	2-A	192	TYR
1	3-A	38	ILE
1	9-A	89	VAL
1	16-A	27	VAL
1	13-A	79	VAL

### 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	1-A	142/181 (78%)	135 (95%)	7 (5%)	27	10
1	2-A	142/181 (78%)	134 (94%)	8 (6%)	23	7
1	3-A	142/181 (78%)	137 (96%)	5 (4%)	39	18
1	4-A	142/181 (78%)	138 (97%)	4 (3%)	47	27
1	5-A	142/181 (78%)	134 (94%)	8 (6%)	23	7
1	6-A	142/181 (78%)	138 (97%)	4 (3%)	47	27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	7-A	142/181 (78%)	139 (98%)	3 (2%)	56	38
1	8-A	142/181 (78%)	135 (95%)	7 (5%)	27	10
1	9-A	142/181 (78%)	132 (93%)	10 (7%)	16	4
1	10-A	142/181 (78%)	137 (96%)	5 (4%)	39	18
1	11-A	142/181 (78%)	137 (96%)	5 (4%)	39	18
1	12-A	142/181 (78%)	137 (96%)	5 (4%)	39	18
1	13-A	142/181 (78%)	135 (95%)	7 (5%)	27	10
1	14-A	142/181 (78%)	135 (95%)	7 (5%)	27	10
1	15-A	142/181 (78%)	129 (91%)	13 (9%)	10	2
1	16-A	142/181 (78%)	136 (96%)	6 (4%)	32	13
All	All	2272/2896 (78%)	2168 (95%)	104 (5%)	29	11

All (104) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	1-A	56	ILE
1	1-A	66	LYS
1	1-A	111	ASP
1	1-A	142	LYS
1	1-A	163	GLU
1	1-A	174	LYS
1	1-A	185	ASP
1	2-A	39	ASP
1	2-A	55	VAL
1	2-A	68	LYS
1	2-A	142	LYS
1	2-A	143	ARG
1	2-A	144	CYS
1	2-A	174	LYS
1	2-A	196	ASP
1	3-A	39	ASP
1	3-A	51	ASP
1	3-A	142	LYS
1	3-A	174	LYS
1	3-A	209	ASP
1	4-A	33	CYS
1	4-A	106	ARG
1	4-A	149	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	4-A	174	LYS
1	5-A	39	ASP
1	5-A	68	LYS
1	5-A	142	LYS
1	5-A	163	GLU
1	5-A	174	LYS
1	5-A	202	SER
1	5-A	208	ASP
1	5-A	209	ASP
1	6-A	26	VAL
1	6-A	142	LYS
1	6-A	163	GLU
1	6-A	174	LYS
1	7-A	51	ASP
1	7-A	146	PHE
1	7-A	151	SER
1	8-A	39	ASP
1	8-A	51	ASP
1	8-A	59	VAL
1	8-A	100	ASP
1	8-A	142	LYS
1	8-A	163	GLU
1	8-A	174	LYS
1	9-A	39	ASP
1	9-A	58	ASP
1	9-A	59	VAL
1	9-A	89	VAL
1	9-A	101	ARG
1	9-A	142	LYS
1	9-A	143	ARG
1	9-A	174	LYS
1	9-A	202	SER
1	9-A	210	VAL
1	10-A	51	ASP
1	10-A	82	ASP
1	10-A	132	ASP
1	10-A	163	GLU
1	10-A	174	LYS
1	11-A	121	ARG
1	11-A	129	THR
1	11-A	142	LYS
1	11-A	174	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	11-A	194	ASN
1	12-A	39	ASP
1	12-A	142	LYS
1	12-A	143	ARG
1	12-A	163	GLU
1	12-A	174	LYS
1	13-A	26	VAL
1	13-A	80	LEU
1	13-A	106	ARG
1	13-A	117	TYR
1	13-A	144	CYS
1	13-A	165	ASP
1	13-A	174	LYS
1	14-A	59	VAL
1	14-A	63	LEU
1	14-A	64	LEU
1	14-A	101	ARG
1	14-A	117	TYR
1	14-A	142	LYS
1	14-A	157	CYS
1	15-A	38	ILE
1	15-A	58	ASP
1	15-A	64	LEU
1	15-A	68	LYS
1	15-A	82	ASP
1	15-A	118	THR
1	15-A	132	ASP
1	15-A	140	ASP
1	15-A	142	LYS
1	15-A	163	GLU
1	15-A	172	HIS
1	15-A	174	LYS
1	15-A	185	ASP
1	16-A	26	VAL
1	16-A	53	ASN
1	16-A	59	VAL
1	16-A	80	LEU
1	16-A	139	LYS
1	16-A	163	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (44) such sidechains are listed below:

Mol	Chain	Res	Type
1	1-A	60	ASN
1	1-A	137	HIS
1	1-A	170	GLN
1	1-A	172	HIS
1	2-A	60	ASN
1	2-A	194	ASN
1	3-A	60	ASN
1	3-A	103	GLN
1	3-A	112	GLN
1	3-A	172	HIS
1	3-A	194	ASN
1	4-A	60	ASN
1	4-A	103	GLN
1	4-A	137	HIS
1	4-A	170	GLN
1	4-A	172	HIS
1	4-A	194	ASN
1	5-A	60	ASN
1	5-A	170	GLN
1	5-A	172	HIS
1	6-A	60	ASN
1	6-A	172	HIS
1	7-A	60	ASN
1	7-A	137	HIS
1	7-A	172	HIS
1	7-A	194	ASN
1	8-A	137	HIS
1	9-A	172	HIS
1	10-A	170	GLN
1	10-A	172	HIS
1	10-A	194	ASN
1	11-A	60	ASN
1	11-A	137	HIS
1	11-A	172	HIS
1	11-A	194	ASN
1	12-A	60	ASN
1	12-A	172	HIS
1	12-A	194	ASN
1	13-A	172	HIS
1	13-A	194	ASN
1	14-A	172	HIS
1	15-A	170	GLN
1	15-A	172	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	16-A	62	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

16 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	EDO	1-A	700	-	3,3,3	0.55	0	2,2,2	0.29	0
2	EDO	10-A	700	-	3,3,3	0.54	0	2,2,2	0.27	0
2	EDO	11-A	700	-	3,3,3	0.57	0	2,2,2	0.27	0
2	EDO	12-A	700	-	3,3,3	0.54	0	2,2,2	0.27	0
2	EDO	13-A	700	-	3,3,3	0.58	0	2,2,2	0.25	0
2	EDO	14-A	700	-	3,3,3	0.62	0	2,2,2	0.25	0
2	EDO	15-A	700	-	3,3,3	0.63	0	2,2,2	0.24	0
2	EDO	16-A	700	-	3,3,3	0.65	0	2,2,2	0.23	0
2	EDO	2-A	700	-	3,3,3	0.55	0	2,2,2	0.27	0
2	EDO	3-A	700	-	3,3,3	0.58	0	2,2,2	0.25	0
2	EDO	4-A	700	-	3,3,3	0.58	0	2,2,2	0.26	0
2	EDO	5-A	700	-	3,3,3	0.59	0	2,2,2	0.26	0
2	EDO	6-A	700	-	3,3,3	0.59	0	2,2,2	0.26	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	EDO	7-A	700	-	3,3,3	0.61	0	2,2,2	0.23	0
2	EDO	8-A	700	-	3,3,3	0.59	0	2,2,2	0.26	0
2	EDO	9-A	700	-	3,3,3	0.60	0	2,2,2	0.24	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
2	EDO	1-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	10-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	11-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	12-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	13-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	14-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	15-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	16-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	2-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	3-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	4-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	5-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	6-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	7-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	8-A	700	-	-	0/1/1/1	0/0/0/0
2	EDO	9-A	700	-	-	0/1/1/1	0/0/0/0

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 ⓘ

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	1-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	2-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	3-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	4-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	5-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	6-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	7-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	8-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	9-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	10-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	11-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	12-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	13-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	14-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	15-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
1	16-A	160/217 (73%)	0.04	7 (4%) 34 39	7, 16, 36, 50	160 (100%)
All	All	2560/3472 (73%)	0.04	112 (4%) 31 39	7, 16, 36, 50	2560 (100%)

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	163	GLU	3.5
1	2-A	163	GLU	3.5
1	3-A	163	GLU	3.5
1	4-A	163	GLU	3.5
1	5-A	163	GLU	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	6-A	163	GLU	3.5
1	7-A	163	GLU	3.5
1	8-A	163	GLU	3.5
1	9-A	163	GLU	3.5
1	10-A	163	GLU	3.5
1	11-A	163	GLU	3.5
1	12-A	163	GLU	3.5
1	13-A	163	GLU	3.5
1	14-A	163	GLU	3.5
1	15-A	163	GLU	3.5
1	16-A	163	GLU	3.5
1	1-A	129	THR	3.2
1	2-A	129	THR	3.2
1	3-A	129	THR	3.2
1	4-A	129	THR	3.2
1	5-A	129	THR	3.2
1	6-A	129	THR	3.2
1	7-A	129	THR	3.2
1	8-A	129	THR	3.2
1	9-A	129	THR	3.2
1	10-A	129	THR	3.2
1	11-A	129	THR	3.2
1	12-A	129	THR	3.2
1	13-A	129	THR	3.2
1	14-A	129	THR	3.2
1	15-A	129	THR	3.2
1	16-A	129	THR	3.2
1	1-A	143	ARG	3.1
1	2-A	143	ARG	3.1
1	3-A	143	ARG	3.1
1	4-A	143	ARG	3.1
1	5-A	143	ARG	3.1
1	6-A	143	ARG	3.1
1	7-A	143	ARG	3.1
1	8-A	143	ARG	3.1
1	9-A	143	ARG	3.1
1	10-A	143	ARG	3.1
1	11-A	143	ARG	3.1
1	12-A	143	ARG	3.1
1	13-A	143	ARG	3.1
1	14-A	143	ARG	3.1
1	15-A	143	ARG	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	16-A	143	ARG	3.1
1	1-A	142	LYS	3.1
1	2-A	142	LYS	3.1
1	3-A	142	LYS	3.1
1	4-A	142	LYS	3.1
1	5-A	142	LYS	3.1
1	6-A	142	LYS	3.1
1	7-A	142	LYS	3.1
1	8-A	142	LYS	3.1
1	9-A	142	LYS	3.1
1	10-A	142	LYS	3.1
1	11-A	142	LYS	3.1
1	12-A	142	LYS	3.1
1	13-A	142	LYS	3.1
1	14-A	142	LYS	3.1
1	15-A	142	LYS	3.1
1	16-A	142	LYS	3.1
1	1-A	164	SER	2.5
1	2-A	164	SER	2.5
1	3-A	164	SER	2.5
1	4-A	164	SER	2.5
1	5-A	164	SER	2.5
1	6-A	164	SER	2.5
1	7-A	164	SER	2.5
1	8-A	164	SER	2.5
1	9-A	164	SER	2.5
1	10-A	164	SER	2.5
1	11-A	164	SER	2.5
1	12-A	164	SER	2.5
1	13-A	164	SER	2.5
1	14-A	164	SER	2.5
1	15-A	164	SER	2.5
1	16-A	164	SER	2.5
1	1-A	140	ASP	2.1
1	2-A	140	ASP	2.1
1	3-A	140	ASP	2.1
1	4-A	140	ASP	2.1
1	5-A	140	ASP	2.1
1	6-A	140	ASP	2.1
1	7-A	140	ASP	2.1
1	8-A	140	ASP	2.1
1	9-A	140	ASP	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	10-A	140	ASP	2.1
1	11-A	140	ASP	2.1
1	12-A	140	ASP	2.1
1	13-A	140	ASP	2.1
1	14-A	140	ASP	2.1
1	15-A	140	ASP	2.1
1	16-A	140	ASP	2.1
1	1-A	139	LYS	2.1
1	2-A	139	LYS	2.1
1	3-A	139	LYS	2.1
1	4-A	139	LYS	2.1
1	5-A	139	LYS	2.1
1	6-A	139	LYS	2.1
1	7-A	139	LYS	2.1
1	8-A	139	LYS	2.1
1	9-A	139	LYS	2.1
1	10-A	139	LYS	2.1
1	11-A	139	LYS	2.1
1	12-A	139	LYS	2.1
1	13-A	139	LYS	2.1
1	14-A	139	LYS	2.1
1	15-A	139	LYS	2.1
1	16-A	139	LYS	2.1

## 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 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	EDO	13-A	700	4/4	0.58	0.18	46,48,49,51	4
2	EDO	12-A	700	4/4	0.58	0.18	47,48,49,51	4
2	EDO	4-A	700	4/4	0.58	0.18	48,48,49,50	4
2	EDO	11-A	700	4/4	0.58	0.18	48,48,49,50	4
2	EDO	16-A	700	4/4	0.58	0.18	48,49,49,50	4
2	EDO	15-A	700	4/4	0.58	0.18	47,49,49,50	4
2	EDO	9-A	700	4/4	0.58	0.18	48,49,49,50	4
2	EDO	3-A	700	4/4	0.58	0.18	48,48,49,50	4
2	EDO	10-A	700	4/4	0.58	0.18	46,48,48,50	4
2	EDO	2-A	700	4/4	0.58	0.18	47,48,49,51	4
2	EDO	7-A	700	4/4	0.58	0.18	48,49,49,50	4
2	EDO	14-A	700	4/4	0.58	0.18	47,49,49,50	4
2	EDO	1-A	700	4/4	0.58	0.18	47,48,49,51	4
2	EDO	6-A	700	4/4	0.58	0.18	48,48,49,50	4
2	EDO	8-A	700	4/4	0.58	0.18	48,48,49,50	4
2	EDO	5-A	700	4/4	0.58	0.18	48,48,49,50	4

## 6.5 Other polymers [i](#)

There are no such residues in this entry.