



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 11:50 pm BST

PDB ID : 2QHF  
Title : Mycobacterium tuberculosis Chorismate synthase in complex with NCA  
Authors : Bruning, M.; Bourenkov, G.P.; Strizhov, N.I.; Bartunik, H.D.  
Deposited on : 2007-07-02  
Resolution : 1.65 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

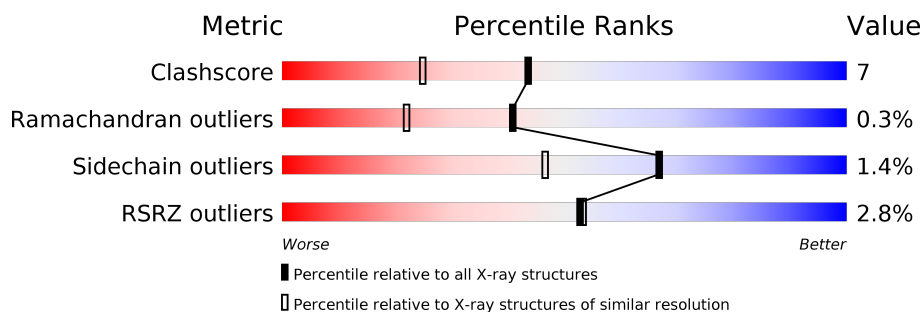
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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 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  $\leq 5\%$ . 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	407	

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	507	-	-	X	-
2	ACT	A	510	-	-	X	-
2	ACT	A	511	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 3563 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 Chorismate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	392	3022	1878	560	570	14	0	28	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	393	GLU	-	EXPRESSION TAG	UNP P63611
A	394	ALA	-	EXPRESSION TAG	UNP P63611
A	395	PRO	-	EXPRESSION TAG	UNP P63611
A	396	ALA	-	EXPRESSION TAG	UNP P63611
A	397	ALA	-	EXPRESSION TAG	UNP P63611
A	398	ARG	-	EXPRESSION TAG	UNP P63611
A	399	VAL	-	EXPRESSION TAG	UNP P63611
A	400	SER	-	EXPRESSION TAG	UNP P63611
A	401	GLY	-	EXPRESSION TAG	UNP P63611
A	402	HIS	-	EXPRESSION TAG	UNP P63611
A	403	HIS	-	EXPRESSION TAG	UNP P63611
A	404	HIS	-	EXPRESSION TAG	UNP P63611
A	405	HIS	-	EXPRESSION TAG	UNP P63611
A	406	HIS	-	EXPRESSION TAG	UNP P63611
A	407	HIS	-	EXPRESSION TAG	UNP P63611

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



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

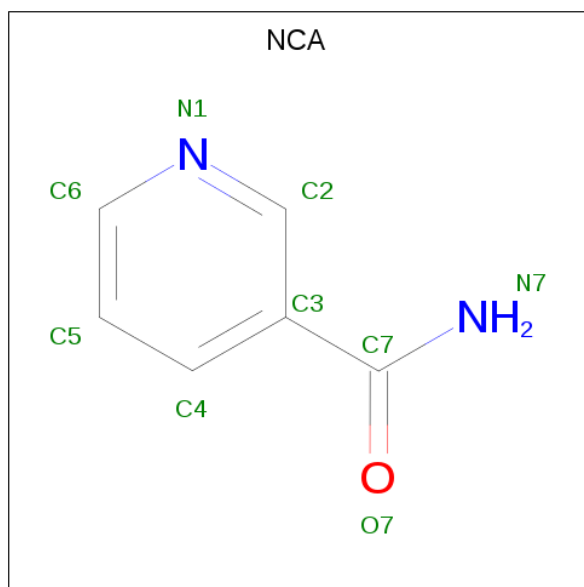
- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Na	0	0
			2	2		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is NICOTINAMIDE (three-letter code: NCA) (formula: C<sub>6</sub>H<sub>6</sub>N<sub>2</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			9	6	2	1		

- Molecule 6 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	473	Total	O	0	0
			473	473		

i

- Molecule 1: Chorismate synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.59Å 132.59Å 159.56Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.65 – 1.65 19.71 – 1.65	Depositor EDS
% Data completeness (in resolution range)	98.8 (19.65-1.65) 98.8 (19.71-1.65)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.45 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.164 , 0.186 0.177 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.8	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 60.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3563	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, NCA, ACT, TRS, CL

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	0/3155	0.71	1/4279 (0.0%)

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	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	309	ARG	NE-CZ-NH1	5.16	122.88	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	316	PRO	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	3022	0	3085	36	0
2	A	48	0	36	17	0
3	A	2	0	0	0	0
4	A	1	0	0	0	0
5	A	9	0	6	1	0
6	A	8	0	12	2	0
7	A	473	0	0	17	1
All	All	3563	0	3139	42	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 7.

All (42) 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:A:139[B]:ARG:NH1	2:A:507:ACT:O	1.59	1.34
6:A:501:TRS:H31	7:A:910:HOH:O	1.42	1.18
1:A:199:LYS:CE	7:A:1067:HOH:O	1.91	1.15
1:A:199:LYS:HE3	7:A:1067:HOH:O	1.49	1.04
1:A:5[B]:ILE:HD11	7:A:1050:HOH:O	1.77	0.85
1:A:139[B]:ARG:NH1	2:A:507:ACT:C	2.44	0.80
1:A:139[B]:ARG:CZ	2:A:507:ACT:O	2.31	0.77
1:A:46:ARG:HH22	5:A:493:NCA:HN72	1.30	0.76
1:A:199:LYS:NZ	7:A:1067:HOH:O	2.17	0.71
1:A:3[A]:ARG:HH21	2:A:510:ACT:C	2.16	0.59
1:A:338[A]:ILE:CD1	2:A:508:ACT:O	2.52	0.58
1:A:1:MET:HE2	1:A:3[B]:ARG:CZ	2.34	0.57
1:A:55:ASP:OD1	1:A:80:THR:HG23	2.07	0.54
1:A:253[B]:MET:HG2	1:A:261:VAL:CG2	2.37	0.54
1:A:253[B]:MET:SD	1:A:261:VAL:HG23	2.47	0.54
1:A:79:ASN:HB2	2:A:507:ACT:CH3	2.39	0.53
1:A:253[B]:MET:HG2	1:A:261:VAL:HG21	1.89	0.53
1:A:3[A]:ARG:NE	2:A:510:ACT:OXT	2.42	0.50
1:A:95[B]:ASP:N	1:A:95[B]:ASP:OD2	2.42	0.49
2:A:511:ACT:CH3	7:A:1076:HOH:O	0.79	0.49
1:A:334[B]:GLU:CG	7:A:623:HOH:O	2.62	0.48
1:A:139[B]:ARG:HH12	2:A:507:ACT:C	2.22	0.48
1:A:204:ASP:O	1:A:208[A]:GLN:HG2	2.13	0.47
1:A:380[B]:ARG:NE	7:A:985:HOH:O	2.00	0.47
1:A:239[A]:SER:OG	1:A:242:HIS:CD2	2.68	0.47
1:A:139[B]:ARG:NH2	2:A:507:ACT:OXT	2.47	0.46
1:A:1:MET:CE	1:A:3[B]:ARG:CZ	2.93	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:507:ACT:CH3	7:A:1055:HOH:O	0.75	0.45
2:A:511:ACT:H2	7:A:1076:HOH:O	0.73	0.45
1:A:79:ASN:HB2	2:A:507:ACT:C	2.48	0.44
1:A:199:LYS:HD2	7:A:1068:HOH:O	2.18	0.44
1:A:3[A]:ARG:HH21	2:A:510:ACT:CH3	2.31	0.43
1:A:20[A]:GLU:CG	7:A:687:HOH:O	2.67	0.43
1:A:20[A]:GLU:HG2	7:A:687:HOH:O	2.17	0.43
6:A:501:TRS:C3	7:A:910:HOH:O	2.25	0.42
1:A:334[B]:GLU:HG2	7:A:623:HOH:O	2.19	0.41
1:A:41:ARG:NH2	1:A:54:ARG:HA	2.36	0.41
2:A:507:ACT:H1	7:A:1055:HOH:O	0.73	0.41
1:A:139[B]:ARG:CZ	2:A:507:ACT:C	2.96	0.41
1:A:322:ARG:HA	1:A:336:VAL:HG11	2.02	0.41
1:A:334[B]:GLU:CD	7:A:623:HOH:O	2.59	0.40
1:A:139[B]:ARG:HD2	2:A:506:ACT:C	2.52	0.40

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)
7:A:609:HOH:O	7:A:808:HOH:O[11_655]	2.17	0.03

## 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	418/407 (103%)	415 (99%)	2 (0%)	1 (0%)	47 28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	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	311/294 (106%)	304 (98%)	7 (2%)	50 25

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20[A]	GLU
1	A	20[B]	GLU
1	A	99[A]	LEU
1	A	99[B]	LEU
1	A	338[A]	ILE
1	A	338[B]	ILE
1	A	356	THR

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

Mol	Chain	Res	Type
1	A	242	HIS
1	A	386	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

Of 17 ligands modelled in this entry, 3 are monoatomic - leaving 14 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	A	502	-	1,3,3	1.03	0	0,3,3	0.00	-
2	ACT	A	510	-	1,3,3	0.56	0	0,3,3	0.00	-
2	ACT	A	506	-	1,3,3	1.00	0	0,3,3	0.00	-
2	ACT	A	503	-	1,3,3	1.19	0	0,3,3	0.00	-
2	ACT	A	507	-	1,3,3	1.85	0	0,3,3	0.00	-
2	ACT	A	509	-	1,3,3	1.77	0	0,3,3	0.00	-
6	TRS	A	501	-	7,7,7	0.37	0	9,9,9	0.54	0
2	ACT	A	511	-	1,3,3	1.11	0	0,3,3	0.00	-
5	NCA	A	493	-	9,9,9	0.73	0	11,11,11	1.48	4 (36%)
2	ACT	A	508	-	1,3,3	2.36	1 (100%)	0,3,3	0.00	-
2	ACT	A	505	-	1,3,3	1.26	0	0,3,3	0.00	-
2	ACT	A	512	-	1,3,3	1.53	0	0,3,3	0.00	-
2	ACT	A	504	-	1,3,3	1.32	0	0,3,3	0.00	-
2	ACT	A	513	-	1,3,3	1.89	0	0,3,3	0.00	-

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
6	TRS	A	501	-	-	0/9/9/9	-
5	NCA	A	493	-	-	0/4/4/4	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	508	ACT	CH3-C	2.36	1.51	1.48

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	A	493	NCA	C6-N1-C2	2.24	120.72	116.85
5	A	493	NCA	C3-C2-N1	-2.23	120.20	123.49
5	A	493	NCA	C5-C4-C3	-2.15	117.80	120.34
5	A	493	NCA	C4-C3-C2	2.11	120.02	117.63

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	510	ACT	3	0
2	A	506	ACT	1	0
2	A	507	ACT	10	0
6	A	501	TRS	2	0
2	A	511	ACT	2	0
5	A	493	NCA	1	0
2	A	508	ACT	1	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, 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	A	392/407 (96%)	-0.46	11 (2%) 53 53	15, 22, 39, 82	14 (3%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	52	PHE	9.1
1	A	285	PRO	6.9
1	A	284	GLY	3.8
1	A	286	ASP	3.5
1	A	47	GLY	3.0
1	A	53	GLU	2.6
1	A	48	ALA	2.4
1	A	51	THR	2.3
1	A	54	ARG	2.2
1	A	97	ALA	2.2
1	A	391	ASP	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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	ACT	A	503	4/4	0.66	0.25	36,39,47,50	0
2	ACT	A	510	4/4	0.73	0.18	20,26,29,32	4
2	ACT	A	511	4/4	0.80	0.20	33,33,34,42	4
2	ACT	A	513	4/4	0.80	0.34	43,44,45,46	0
2	ACT	A	508	4/4	0.88	0.13	21,23,26,27	4
2	ACT	A	512	4/4	0.89	0.13	32,33,36,37	0
6	TRS	A	501	8/8	0.90	0.16	16,20,23,25	8
2	ACT	A	509	4/4	0.90	0.21	42,45,49,50	0
5	NCA	A	493	9/9	0.93	0.08	20,26,29,30	0
2	ACT	A	507	4/4	0.94	0.21	22,24,28,28	4
2	ACT	A	504	4/4	0.96	0.11	35,41,43,48	0
2	ACT	A	505	4/4	0.96	0.07	45,46,47,47	4
2	ACT	A	506	4/4	0.96	0.07	27,29,30,31	0
3	NA	A	602	1/1	0.97	0.04	31,31,31,31	0
4	CL	A	603	1/1	0.98	0.05	25,25,25,25	0
3	NA	A	601	1/1	0.99	0.05	23,23,23,23	0
2	ACT	A	502	4/4	0.99	0.04	22,23,25,26	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.