wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:32 PM GMT

PDB ID	:	1PH7
Title	:	CRYSTAL STRUCTURE OF THE OXYTRICHA NOVA TELOMERE END-BINDING PROTEIN COMPLEXED WITH NONCOGNATE SSDNA GGGGTTTTGIGG
Authors	:	Theobald, D.L.; Schultz, S.C.
Deposited on	:	2003-05-29
Resolution	:	2.90 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

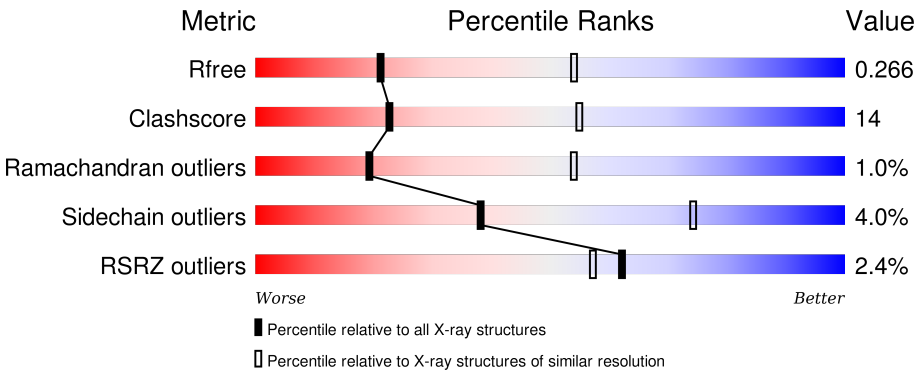
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 <=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	G	13	38% 54% 8%
1	H	13	62% 31% 8%
2	D	11	100%
3	A	460	2% 73% 25% •
4	B	216	4% 60% 38% •

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
5	NA	H	1002	-	-	-	X
5	NA	H	1003	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6681 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 DNA chain called 5'-D(*GP*GP*GP*GP*TP*TP*TP*TP*GP*GP*GP*GP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	G	12	253	120	48	74	11	0	0	0
1	H	12	253	120	48	74	11	0	0	0

- Molecule 2 is a DNA chain called 5'-D(*GP*GP*GP*GP*TP*TP*TP*TP*GP*IP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	11	230	110	42	68	10	0	0	0

- Molecule 3 is a protein called Telomere-binding protein alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	460	3706	2355	637	712	2	0	0	0

- Molecule 4 is a protein called Telomere-binding protein beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	216	1733	1112	294	326	1	0	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
5	H	3	3	3	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 7 is water.

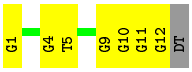
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	332	332	332	0	0
			Total	O		
7	B	75	75	75	0	0
			Total	O		
7	D	29	29	29	0	0
			Total	O		
7	G	31	31	31	0	0
			Total	O		
7	H	35	35	35	0	0

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 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.

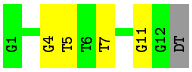
- Molecule 1: 5'-D(*GP*GP*GP*GP*TP*TP*TP*TP*GP*GP*GP*GP*T)-3'

Chain G:



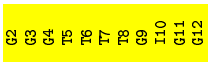
- Molecule 1: 5'-D(*GP*GP*GP*GP*TP*TP*TP*TP*GP*GP*GP*GP*T)-3'

Chain H:



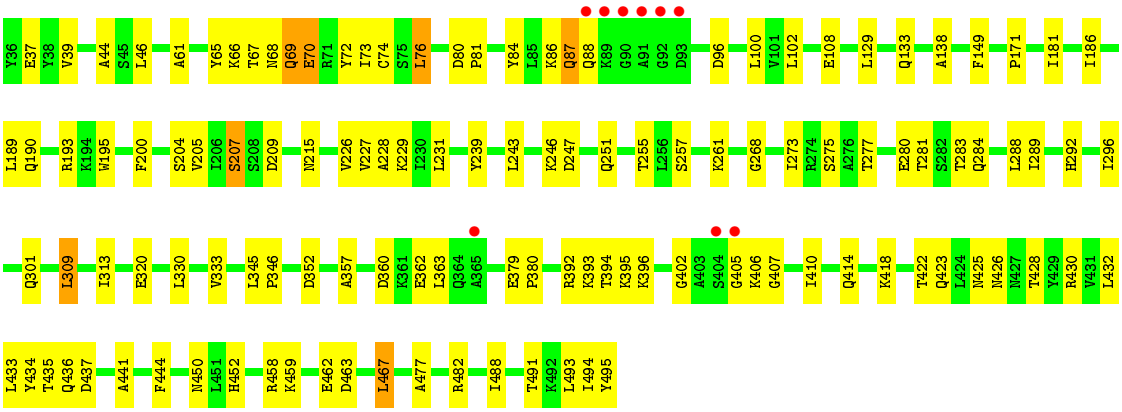
- Molecule 2: 5'-D(*GP*GP*GP*GP*TP*TP*TP*TP*GP*IP*GP*G)-3'

Chain D:



- Molecule 3: Telomere-binding protein alpha subunit

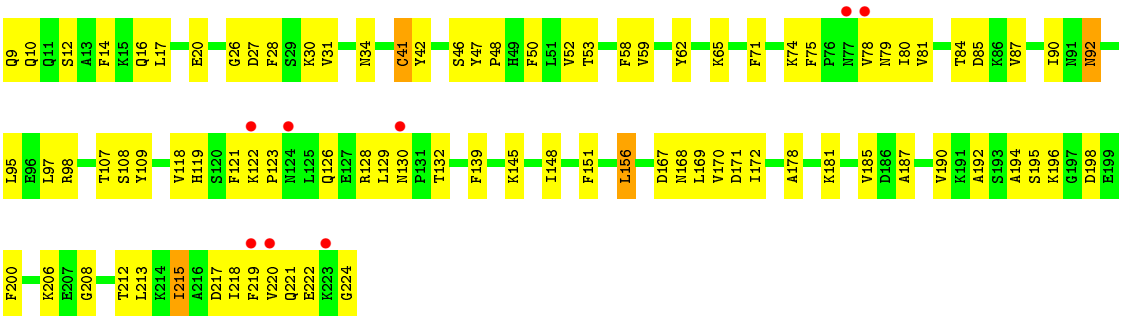
Chain A:



- Molecule 4: Telomere-binding protein beta subunit

Chain B:

4%
60%
38%


Q9 Q10 Q11 S12 A13 F14 K15 Q16 L17 E20 G26 D27 F28 S29 K30 V31 N34 G41 Y42 S46 Y47 P48 H49 F50 I51 V52 T53 F58 V59 Y62 K65 F71 K74 F75 P76 N77 V78 N79 I80 V81 T84 D85 K86 V87 I90 N91 N92
L95 E96 L97 R98 T107 S108 Y109 V118 H119 S120 F121 K122 P123 N124 L125 Q126 E127 R128 L129 N130 P131 T132 F139 K145 I148 F151 L156 D167 N168 L169 V170 D171 I172 A178 K181 V185 D186 A187 V190 K191 A192 S193 A194 S195 K196 G197 D198 E199
F200 K206 E207 G208 T212 L213 K214 I215 A216 D217 I218 F219 V220 Q221 E222 K223 G224

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	93.19Å 93.19Å 421.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.64 – 2.90 40.51 – 2.59	Depositor EDS
% Data completeness (in resolution range)	96.4 (39.64-2.90) 95.7 (40.51-2.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.34 (at 2.58Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.230 , 0.268 0.226 , 0.266	Depositor DCC
R_{free} test set	2427 reflections (11.08%)	DCC
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.336	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 40.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Outliers	0 of 33513 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6681	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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, 0.375 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: NA, 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	G	0.96	0/284	0.73	0/440
1	H	1.04	0/284	0.78	0/440
2	D	0.95	0/233	0.85	0/358
3	A	0.58	0/3777	0.61	0/5104
4	B	0.51	0/1771	0.52	0/2391
All	All	0.63	0/6349	0.62	0/8733

There are no bond length outliers.

There are no bond angle outliers.

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	G	253	0	138	7	0
1	H	253	0	138	6	0
2	D	230	0	126	13	0
3	A	3706	0	3699	86	0
4	B	1733	0	1718	67	1
5	H	3	0	0	0	0
6	A	1	0	0	0	0
7	A	332	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	75	0	0	4	0
7	D	29	0	0	1	0
7	G	31	0	0	2	0
7	H	35	0	0	4	0
All	All	6681	0	5819	163	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 14.

The worst 5 of 163 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)
4:B:10:GLN:HE21	4:B:12:SER:H	0.97	0.92
4:B:80:ILE:HD12	4:B:218:ILE:HD11	1.52	0.91
4:B:65:LYS:HD2	4:B:213:LEU:HD21	1.54	0.89
4:B:75:PHE:HZ	4:B:122:LYS:HE3	1.38	0.86
3:A:345:LEU:HD12	3:A:346:PRO:HD2	1.59	0.85

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:B:130:ASN:ND2	4:B:130:ASN:ND2[12_555]	1.97	0.23

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
3	A	458/460 (100%)	417 (91%)	36 (8%)	5 (1%)	17 51
4	B	214/216 (99%)	186 (87%)	26 (12%)	2 (1%)	21 57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	672/676 (99%)	603 (90%)	62 (9%)	7 (1%)	19 54

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	87	GLN
3	A	70	GLU
3	A	402	GLY
3	A	406	LYS
4	B	126	GLN

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
3	A	408/408 (100%)	394 (97%)	14 (3%)	44 79
4	B	189/189 (100%)	179 (95%)	10 (5%)	28 63
All	All	597/597 (100%)	573 (96%)	24 (4%)	38 74

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	362	GLU
3	A	482	ARG
4	B	206	LYS
3	A	432	LEU
3	A	467	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	372	GLN
3	A	423	GLN
4	B	92	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	356	HIS
4	B	79	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)

Of 4 ligands modelled in this entry, 4 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 (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	G	12/13 (92%)	0.20	0	100	100	30, 37, 46, 49	0
1	H	12/13 (92%)	0.77	0	100	100	35, 43, 47, 49	0
2	D	10/11 (90%)	-0.39	0	100	100	22, 31, 48, 51	0
3	A	460/460 (100%)	-0.38	9 (1%)	68	64	17, 27, 52, 80	0
4	B	216/216 (100%)	0.05	8 (3%)	45	38	22, 45, 72, 79	0
All	All	710/713 (99%)	-0.22	17 (2%)	62	57	17, 31, 66, 80	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	404	SER	7.8
3	A	405	GLY	5.7
3	A	90	GLY	5.6
3	A	91	ALA	4.5
3	A	89	LYS	4.4

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(Å^2)	Q<0.9
5	NA	H	1002	1/1	0.79	0.32	2.46	37,37,37,37	0
5	NA	H	1003	1/1	0.80	0.33	2.38	44,44,44,44	0
6	CL	A	1005	1/1	0.95	0.15	-0.47	37,37,37,37	0
5	NA	H	1004	1/1	-0.02	0.21	-0.86	41,41,41,41	0

6.5 Other polymers (i)

There are no such residues in this entry.