



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 11:54 PM UTC

PDB ID : 2CXE / pdb_00002cxe
Title : Crystal structure of octameric ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from *Pyrococcus horikoshii* OT3 (form-2 crystal)
Authors : Mizohata, E.; Mishima, C.; Akasaka, R.; Uda, H.; Terada, T.; Shirouzu, M.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2005-06-28
Resolution : 3.00 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

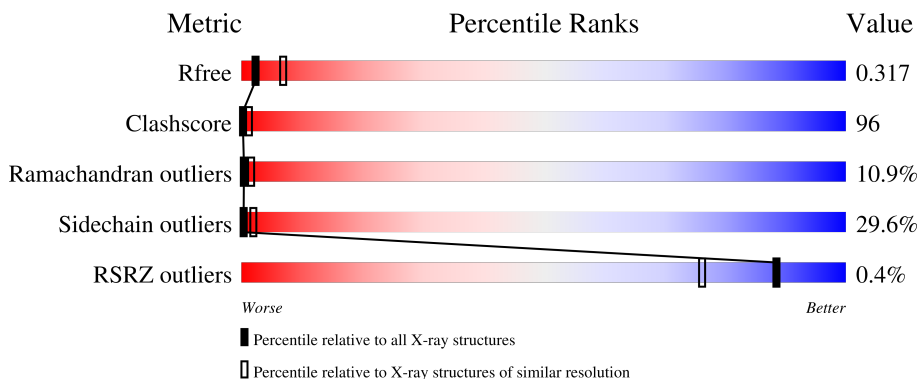
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

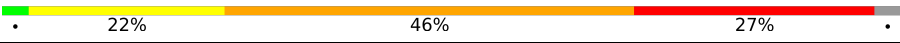
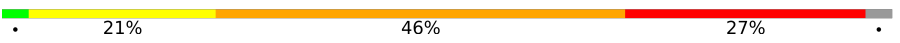
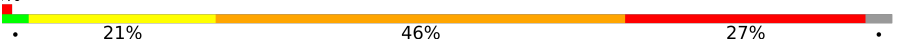
The reported resolution of this entry is 3.00 Å.

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}	180053	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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 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 $\leq 5\%$. 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	430	
1	B	430	
1	C	430	
1	D	430	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13348 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 Ribulose biphosphate carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	S	0	0	0
			3286	2107	565	599	15			
1	B	416	Total	C	N	O	S	0	0	0
			3286	2107	565	599	15			
1	C	416	Total	C	N	O	S	0	0	0
			3286	2107	565	599	15			
1	D	416	Total	C	N	O	S	0	0	0
			3286	2107	565	599	15			

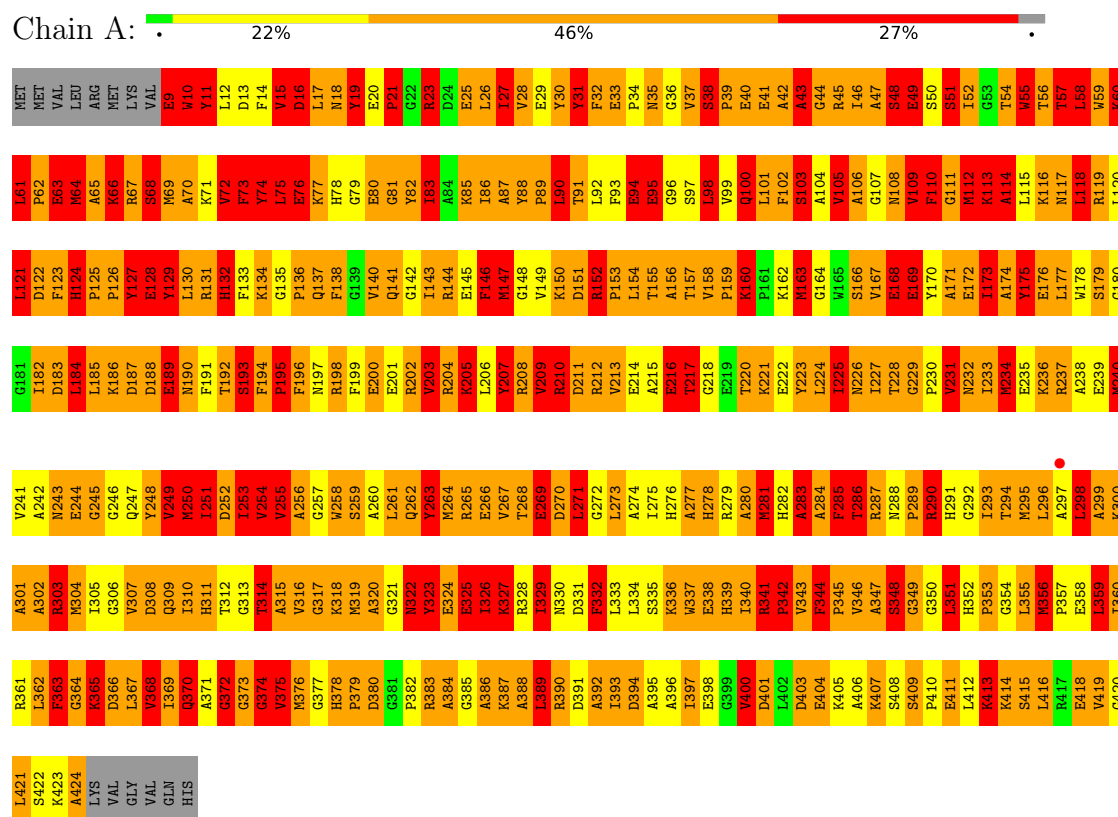
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	52	Total	O	0	0
			52	52		
2	B	51	Total	O	0	0
			51	51		
2	C	52	Total	O	0	0
			52	52		
2	D	49	Total	O	0	0
			49	49		

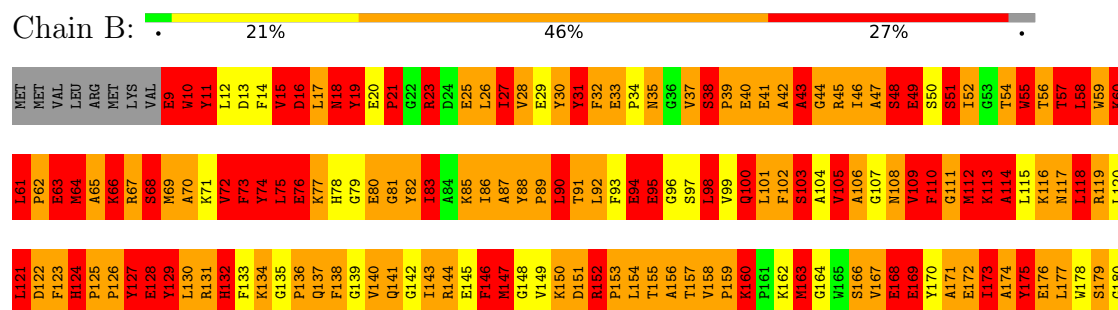
3 Residue-property plots

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: Ribulose biphosphate carboxylase



• Molecule 1: Ribulose biphosphate carboxylase



L421	L121	G181	G181	V241	A301	R361	L421
S422	D122	I182	I182	A242	A302	L362	S422
K423	F123	D183	D183	M243	R303	F363	K423
A424	L124	L184	L184	E244	R304	G364	A424
	P125	L185	L185	G245	I305	K365	LYS
	P126	K186	K186	G246	G306	D366	VAL
	Y127	D187	D187	Q247	V307	L367	GLY
	E128	D188	D188	Y248	D308	V368	VAL
Y129	L130	E189	E189	Y249	Q309	I369	GLN
L131	N190	N190	N190	M250	I310	Q370	HIS
	F191	F191	F191	I251	H311	A371	
	H132	T192	T192	D252	T312	G372	
F133	S193	S193	S193	I253	G313	G373	
K134	F194	F194	F194	V254	T314	G374	
	P195	P195	P195	V255	A315	V375	
	P196	P196	P196	A256	V316	M376	
	N197	N197	N197	G257	G317	G377	
	R198	R198	R198	W258	K318	H378	
	F199	F199	F199	S259	M319	P379	
	E200	E200	E200	A260	A320	D380	
V140	E201	E201	E201	L261	G321	G381	
Q141	R202	R202	R202	Q262	N322	P382	
G142	Y203	Y203	Y203	Y263	Y323	R383	
I143	R204	R204	R204	M264	E324	A384	
R144	K205	K205	K205	R265	E325	G385	
E145	L206	L206	L206	E266	I326	A386	
F146	Y207	Y207	Y207	V267	K327	K387	
M147	R208	R208	R208	T268	R328	A388	
G148	V209	V209	V209	E269	I329	L389	
V149	R210	R210	R210	D270	N330	R390	
K150	D211	D211	D211	L271	D331	D391	
D151	R212	R212	R212	G272	F332	A392	
R152	P213	P213	P213	L273	L333	I393	
P153	L154	L154	L154	A274	L334	D394	
L154	A215	A215	A215	I275	S335	A395	
T155	E216	E216	E216	H276	K336	A396	
A156	T217	T217	T217	A277	W337	I397	
T157	G218	G218	G218	H278	E338	E398	
V158	E219	E219	E219	R279	H339	G399	
P159	T220	T220	T220	A280	I340	V400	
K160	K221	K221	K221	H281	R341	D401	
P161	E222	E222	E222	H282	P342	L402	
K162	Y223	Y223	Y223	A283	Y343	D403	
M163	L224	L224	L224	A284	F344	E404	
G164	I225	I225	I225	F285	P345	K405	
W165	N226	N226	N226	T286	V346	A406	
S166	I227	I227	I227	R287	A347	K407	
V167	T228	T228	T228	N288	S348	S408	
E168	G229	G229	G229	P289	G349	S409	
E169	P230	P230	P230	R290	G350	P410	
Y170	V231	V231	V231	H291	L351	E411	
A171	N232	N232	N232	G292	H352	L412	
E172	I233	I233	I233	T293	P353	K413	
I173	M234	M234	M234	T294	G354	K414	
A174	E235	E235	E235	M295	L355	S415	
Y175	K236	K236	K236	L296	K356	L416	
E176	R237	R237	R237	A297	P357	R417	
L177	A238	A238	A238	L298	E358	E418	
W178	E239	E239	E239	A299	L359	V419	
S179	M240	M240	M240	K300	I360	G420	
G180							

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	171.48Å 149.21Å 107.46Å 90.00° 127.37° 90.00°	Depositor
Resolution (Å)	46.72 – 3.00 46.72 – 3.00	Depositor EDS
% Data completeness (in resolution range)	85.7 (46.72-3.00) 85.7 (46.72-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 2.91Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.295 , 0.328 0.286 , 0.317	Depositor DCC
R_{free} test set	1838 reflections (4.25%)	wwPDB-VP
Wilson B-factor (Å ²)	63.7	Xtriage
Anisotropy	0.760	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 60.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.002 for -h-2*k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	13348	wwPDB-VP
Average B, all atoms (Å ²)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	3.95	548/3365 (16.3%)	3.22	527/4546 (11.6%)
1	B	3.95	547/3365 (16.3%)	3.22	524/4546 (11.5%)
1	C	3.95	550/3365 (16.3%)	3.22	527/4546 (11.6%)
1	D	3.95	547/3365 (16.3%)	3.22	524/4546 (11.5%)
All	All	3.95	2192/13460 (16.3%)	3.22	2102/18184 (11.6%)

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	16
1	B	0	16
1	C	0	16
1	D	0	16
All	All	0	64

The worst 5 of 2192 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	64	MET	SD-CE	46.26	2.95	1.79
1	A	64	MET	SD-CE	46.22	2.95	1.79
1	B	64	MET	SD-CE	46.21	2.95	1.79
1	D	64	MET	SD-CE	46.21	2.95	1.79
1	C	281	MET	SD-CE	35.43	2.68	1.79

The worst 5 of 2102 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	289	PRO	N-CA-C	17.45	132.83	114.68
1	C	289	PRO	N-CA-C	17.44	132.81	114.68
1	B	289	PRO	N-CA-C	17.43	132.81	114.68

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	289	PRO	N-CA-C	17.42	132.79	114.68
1	C	378	HIS	CA-C-N	17.16	138.15	119.28

There are no chirality outliers.

5 of 64 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	110	PHE	Sidechain
1	A	127	TYR	Sidechain
1	A	129	TYR	Sidechain
1	A	19	TYR	Sidechain
1	A	73	PHE	Sidechain

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	3286	0	3271	612	7
1	B	3286	0	3271	677	11
1	C	3286	0	3269	696	4
1	D	3286	0	3269	628	10
2	A	52	0	0	1	1
2	B	51	0	0	2	0
2	C	52	0	0	6	0
2	D	49	0	0	2	1
All	All	13348	0	13080	2528	19

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 96.

The worst 5 of 2528 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:336:LYS:CD	1:A:336:LYS:CE	1.74	1.65
1:C:160:LYS:CD	1:C:160:LYS:CE	1.75	1.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:56:THR:CA	1:A:56:THR:CB	1.75	1.63
1:C:56:THR:CA	1:C:56:THR:CB	1.75	1.63
1:D:419:VAL:CB	1:D:419:VAL:CG2	1.76	1.63

The worst 5 of 19 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)
1:B:56:THR:O	1:D:421:LEU:CD1[3_444]	0.58	1.62
1:C:422:SER:OG	1:D:424:ALA:O[3_444]	1.07	1.13
1:C:422:SER:OG	1:D:424:ALA:C[3_444]	1.34	0.86
1:A:23:ARG:NH1	1:B:23:ARG:NH2[2_555]	1.63	0.57
1:A:23:ARG:NH2	1:B:23:ARG:NH1[2_555]	1.66	0.54

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	414/430 (96%)	298 (72%)	71 (17%)	45 (11%)	0	1
1	B	414/430 (96%)	298 (72%)	71 (17%)	45 (11%)	0	1
1	C	414/430 (96%)	298 (72%)	71 (17%)	45 (11%)	0	1
1	D	414/430 (96%)	298 (72%)	71 (17%)	45 (11%)	0	1
All	All	1656/1720 (96%)	1192 (72%)	284 (17%)	180 (11%)	0	1

5 of 180 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	TYR
1	A	50	SER
1	A	60	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	114	ALA
1	A	128	GLU

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	338/351 (96%)	238 (70%)	100 (30%)	0	2
1	B	338/351 (96%)	238 (70%)	100 (30%)	0	2
1	C	338/351 (96%)	238 (70%)	100 (30%)	0	2
1	D	338/351 (96%)	238 (70%)	100 (30%)	0	2
All	All	1352/1404 (96%)	952 (70%)	400 (30%)	0	2

5 of 400 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	95	GLU
1	C	329	ILE
1	D	418	GLU
1	C	116	LYS
1	C	231	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	35	ASN
1	C	276	HIS
1	D	282	HIS
1	C	100	GLN
1	C	226	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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	416/430 (96%)	-0.52	1 (0%) 91 83	61, 89, 123, 154	0
1	B	416/430 (96%)	-0.50	1 (0%) 91 83	61, 89, 123, 154	0
1	C	416/430 (96%)	-0.47	0 100 100	61, 89, 123, 154	0
1	D	416/430 (96%)	-0.40	4 (0%) 79 59	61, 89, 123, 154	0
All	All	1664/1720 (96%)	-0.47	6 (0%) 88 76	61, 89, 123, 154	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	422	SER	3.1
1	B	419	VAL	3.0
1	D	420	GLY	2.8
1	D	56	THR	2.4
1	D	241	VAL	2.3

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.