



# wwPDB X-ray Structure Validation Summary Report

Nov 21, 2017 – 02:57 AM EST

PDB ID : 2PCD  
Title : STRUCTURE OF PROTOCATECHUATE 3,4-DIOXYGENASE FROM  
PSEUDOMONAS AERUGINOSA AT 2.15 ANGSTROMS RESOLUTION  
Authors : Ohlendorf, D.H.; Orville, A.M.; Lipscomb, J.D.  
Deposited on : unknown  
Resolution : 2.15 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

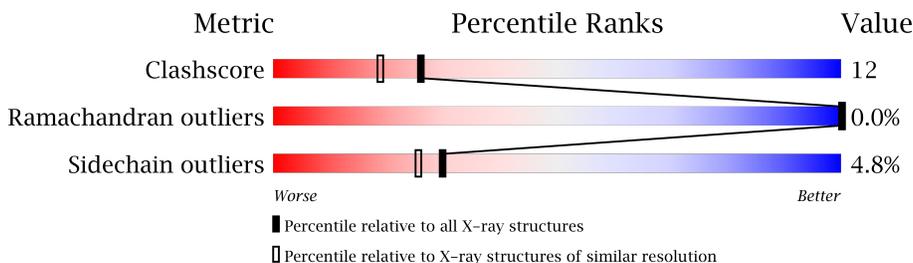
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	200	
1	B	200	
1	C	200	
1	D	200	
1	E	200	
1	F	200	
2	M	238	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	N	238	 78% 15% 5% .
2	O	238	 73% 20% . .
2	P	238	 74% 19% . .
2	Q	238	 69% 24% 5% .
2	R	238	 66% 27% . . .

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 21906 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 PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	200	1571	993	276	299	3	0	0	0
1	B	200	1571	993	276	299	3	0	0	0
1	C	200	1571	993	276	299	3	0	0	0
1	D	200	1571	993	276	299	3	0	0	0
1	E	200	1571	993	276	299	3	0	0	0
1	F	200	1571	993	276	299	3	0	0	0

- Molecule 2 is a protein called PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	M	233	1840	1171	334	328	7	0	0	0
2	N	233	1840	1171	334	328	7	0	0	0
2	O	233	1840	1171	334	328	7	0	0	0
2	P	233	1840	1171	334	328	7	0	0	0
2	Q	233	1840	1171	334	328	7	0	0	0
2	R	233	1840	1171	334	328	7	0	0	0

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	P	1	Total Fe 1 1	0	0
3	Q	1	Total Fe 1 1	0	0
3	N	1	Total Fe 1 1	0	0
3	O	1	Total Fe 1 1	0	0
3	R	1	Total Fe 1 1	0	0
3	M	1	Total Fe 1 1	0	0

- Molecule 4 is water.

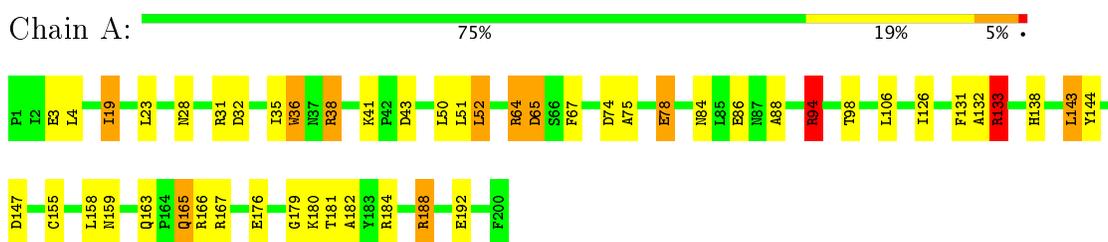
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	83	Total O 83 83	0	0
4	M	154	Total O 154 154	0	0
4	B	79	Total O 79 79	0	0
4	N	163	Total O 163 163	0	0
4	C	80	Total O 80 80	0	0
4	O	158	Total O 158 158	0	0
4	D	77	Total O 77 77	0	0
4	P	159	Total O 159 159	0	0
4	E	77	Total O 77 77	0	0
4	Q	163	Total O 163 163	0	0
4	F	83	Total O 83 83	0	0
4	R	158	Total O 158 158	0	0

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

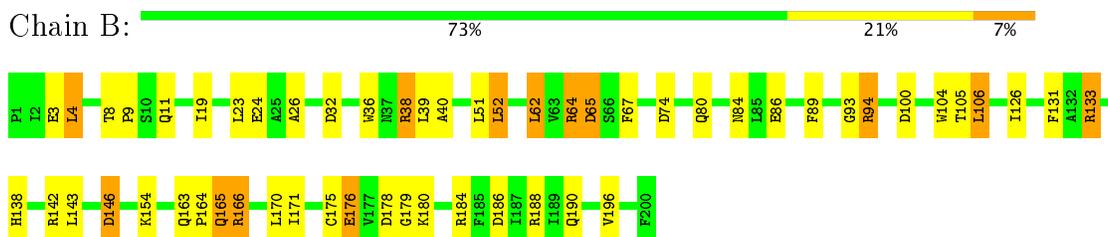
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

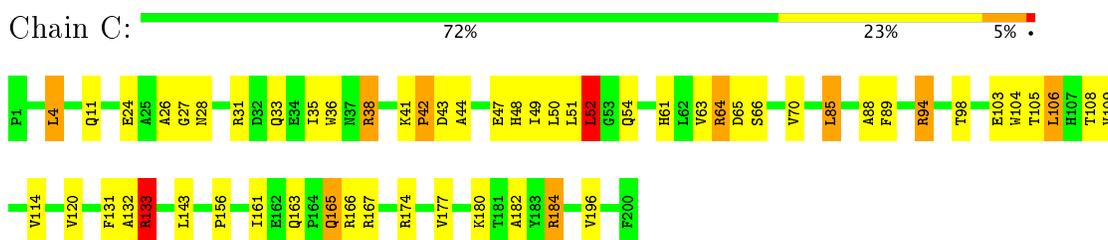
- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)



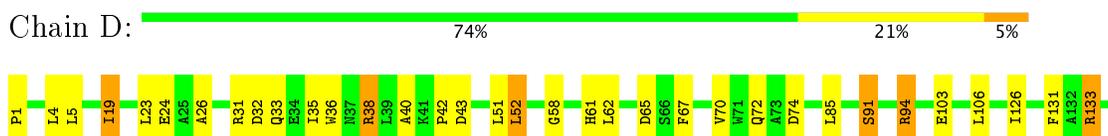
- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)



- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)



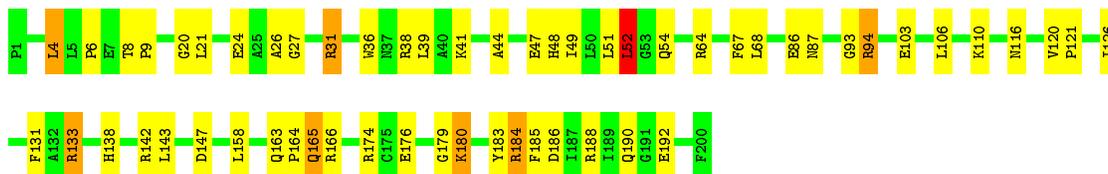
- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)





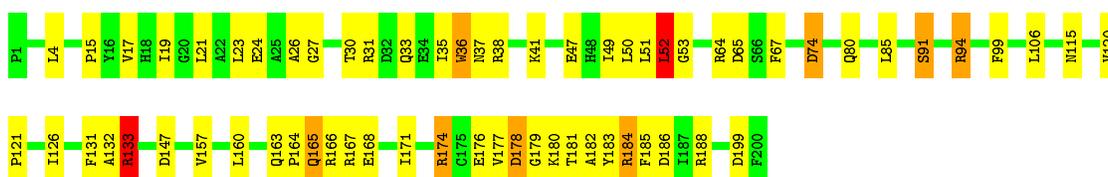
- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)

Chain E: 72% 25% ..



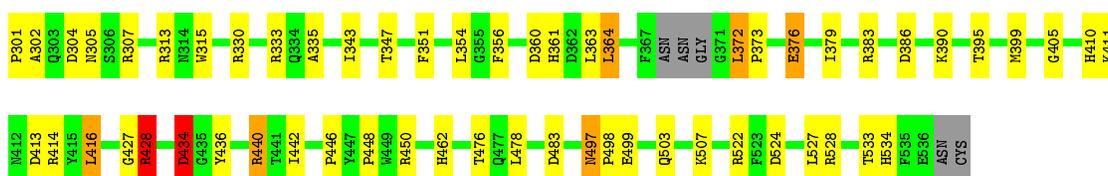
- Molecule 1: PROTOCATECHUATE 3,4-DIOXYGENASE (ALPHA CHAIN)

Chain F: 68% 27% ..



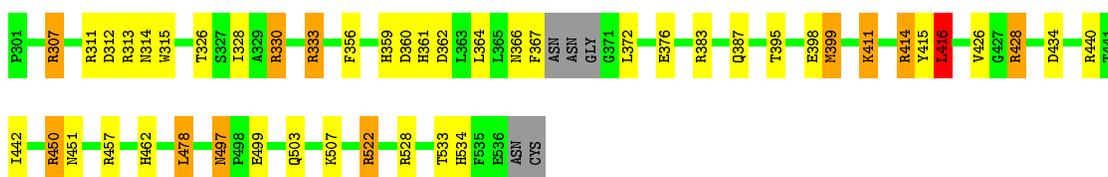
- Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

Chain M: 74% 21% ..



- Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

Chain N: 78% 15% 5% .



- Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

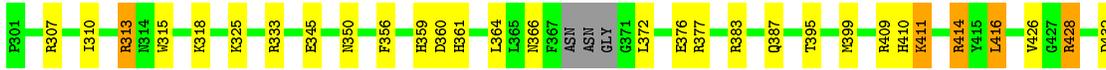
Chain O: 73% 20% ..





• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

Chain P:     ..



• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

Chain Q:     ..



• Molecule 2: PROTOCATECHUATE 3,4-DIOXYGENASE (BETA CHAIN)

Chain R:     ..



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	197.17Å 127.03Å 134.18Å 90.00° 97.64° 90.00°	Depositor
Resolution (Å)	5.00 – 2.15	Depositor
% Data completeness (in resolution range)	(Not available) (5.00-2.15)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, $R_{free}$	0.172 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	21906	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

## 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: FE

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	0.99	1/1611 (0.1%)	1.72	32/2195 (1.5%)
1	B	0.98	1/1611 (0.1%)	1.61	29/2195 (1.3%)
1	C	1.00	0/1611	1.64	22/2195 (1.0%)
1	D	0.99	1/1611 (0.1%)	1.73	32/2195 (1.5%)
1	E	1.02	0/1611	1.72	27/2195 (1.2%)
1	F	1.06	0/1611	1.72	28/2195 (1.3%)
2	M	1.00	1/1895 (0.1%)	1.63	26/2580 (1.0%)
2	N	1.02	1/1895 (0.1%)	1.66	33/2580 (1.3%)
2	O	1.01	0/1895	1.60	34/2580 (1.3%)
2	P	1.01	2/1895 (0.1%)	1.57	25/2580 (1.0%)
2	Q	1.08	1/1895 (0.1%)	1.63	35/2580 (1.4%)
2	R	1.07	0/1895	1.65	32/2580 (1.2%)
All	All	1.02	8/21036 (0.0%)	1.66	355/28650 (1.2%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	94	ARG	CD-NE	-5.99	1.36	1.46
2	P	428	ARG	CD-NE	-5.71	1.36	1.46
2	N	440	ARG	CD-NE	-5.67	1.36	1.46
1	B	94	ARG	CD-NE	-5.65	1.36	1.46
2	P	345	GLU	CD-OE1	-5.59	1.19	1.25

The worst 5 of 355 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	N	440	ARG	NE-CZ-NH2	-24.05	108.28	120.30
2	M	440	ARG	NE-CZ-NH2	-23.06	108.77	120.30
2	Q	440	ARG	NE-CZ-NH2	-22.31	109.14	120.30
2	R	440	ARG	NE-CZ-NH2	-20.05	110.28	120.30
1	F	38	ARG	CD-NE-CZ	19.67	151.13	123.60

There are no chirality outliers.

There are no planarity outliers.

## 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	A	1571	0	1499	34	0
1	B	1571	0	1499	45	0
1	C	1571	0	1499	43	0
1	D	1571	0	1499	29	0
1	E	1571	0	1499	42	0
1	F	1571	0	1499	49	0
2	M	1840	0	1793	51	0
2	N	1840	0	1793	32	0
2	O	1840	0	1793	34	0
2	P	1840	0	1793	48	0
2	Q	1840	0	1793	50	0
2	R	1840	0	1793	63	0
3	M	1	0	0	0	0
3	N	1	0	0	0	0
3	O	1	0	0	0	0
3	P	1	0	0	0	0
3	Q	1	0	0	0	0
3	R	1	0	0	0	0
4	A	83	0	0	2	0
4	B	79	0	0	1	0
4	C	80	0	0	2	0
4	D	77	0	0	0	0
4	E	77	0	0	0	0
4	F	83	0	0	1	0
4	M	154	0	0	4	0
4	N	163	0	0	2	0
4	O	158	0	0	5	0
4	P	159	0	0	1	0
4	Q	163	0	0	6	0
4	R	158	0	0	4	0
All	All	21906	0	19752	475	0

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 12.

The worst 5 of 475 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:E:165:GLN:NE2	1:E:165:GLN:H	1.34	1.23
1:E:165:GLN:N	1:E:165:GLN:HE21	1.39	1.18
1:C:163:GLN:HB3	1:C:165:GLN:NE2	1.68	1.07
1:B:165:GLN:NE2	1:B:165:GLN:H	1.64	0.95
1:E:176:GLU:OE2	1:E:179:GLY:HA2	1.68	0.94

There are no symmetry-related clashes.

## 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	198/200 (99%)	193 (98%)	5 (2%)	0	100	100
1	B	198/200 (99%)	195 (98%)	3 (2%)	0	100	100
1	C	198/200 (99%)	193 (98%)	4 (2%)	1 (0%)	32	25
1	D	198/200 (99%)	187 (94%)	11 (6%)	0	100	100
1	E	198/200 (99%)	187 (94%)	11 (6%)	0	100	100
1	F	198/200 (99%)	185 (93%)	13 (7%)	0	100	100
2	M	229/238 (96%)	224 (98%)	5 (2%)	0	100	100
2	N	229/238 (96%)	220 (96%)	9 (4%)	0	100	100
2	O	229/238 (96%)	221 (96%)	8 (4%)	0	100	100
2	P	229/238 (96%)	224 (98%)	5 (2%)	0	100	100
2	Q	229/238 (96%)	223 (97%)	6 (3%)	0	100	100
2	R	229/238 (96%)	220 (96%)	9 (4%)	0	100	100
All	All	2562/2628 (98%)	2472 (96%)	89 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	42	PRO

### 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	162/163 (99%)	151 (93%)	11 (7%)	18	12
1	B	162/163 (99%)	155 (96%)	7 (4%)	33	30
1	C	162/163 (99%)	157 (97%)	5 (3%)	45	44
1	D	162/163 (99%)	152 (94%)	10 (6%)	21	15
1	E	162/163 (99%)	158 (98%)	4 (2%)	53	55
1	F	162/163 (99%)	155 (96%)	7 (4%)	33	30
2	M	196/202 (97%)	187 (95%)	9 (5%)	31	27
2	N	196/202 (97%)	186 (95%)	10 (5%)	28	22
2	O	196/202 (97%)	185 (94%)	11 (6%)	25	19
2	P	196/202 (97%)	185 (94%)	11 (6%)	25	19
2	Q	196/202 (97%)	186 (95%)	10 (5%)	28	22
2	R	196/202 (97%)	187 (95%)	9 (5%)	31	27
All	All	2148/2190 (98%)	2044 (95%)	104 (5%)	30	25

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	O	428	ARG
1	D	91	SER
2	R	395	THR
2	O	434	ASP
2	O	534	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	O	503	GLN
2	P	412	ASN
2	R	497	ASN
1	D	163	GLN
1	D	165	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 [i](#)

Of 6 ligands modelled in this entry, 6 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.