W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 04:39 am GMT

PDB ID	:	4C65
Title	:	Crystal structure of A. niger ochratoxinase
Authors	:	Dobritzsch, D.; Wang, H.; Schneider, G.; Yu, S.
Deposited on	:	2013-09-17
Resolution	:	2.20 Å(reported)

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
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)	:	2.36

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 2.20 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	Percentile Ranks	Value
Rfree		0.250
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		3.2%
RSRZ outliers		5.5%

Worse   Better

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
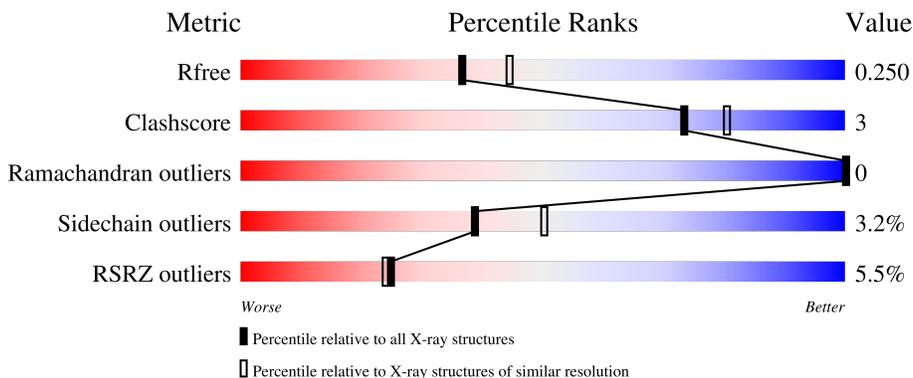
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	438	4% / 89% / 11%
1	B	438	3% / 89% / 8% .
1	C	438	4% / 89% / 9% ..
1	D	438	5% / 90% / 10% .
1	E	438	4% / 89% / 10% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	438	4% 91% 8% ·
1	G	438	9% 88% 9% ··
1	H	438	9% 88% 10% ·
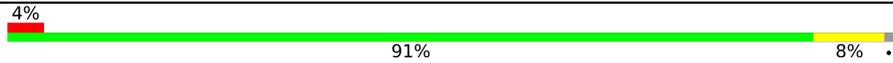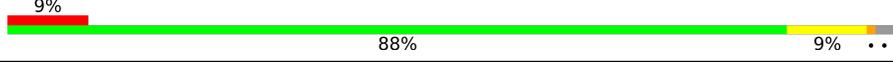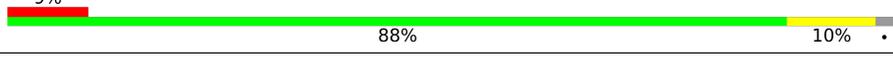
2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 27300 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 OCHRATOXINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	436	Total 3281	C 2078	N 569	O 621	S 13	0	2	0
1	B	430	Total 3250	C 2060	N 565	O 612	S 13	0	4	0
1	C	433	Total 3274	C 2075	N 568	O 618	S 13	0	5	0
1	D	435	Total 3264	C 2067	N 567	O 617	S 13	0	1	0
1	E	435	Total 3267	C 2069	N 567	O 618	S 13	0	1	0
1	F	434	Total 3260	C 2065	N 569	O 613	S 13	0	1	0
1	G	431	Total 3236	C 2050	N 563	O 610	S 13	0	0	0
1	H	429	Total 3228	C 2045	N 564	O 606	S 13	0	1	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	223	Total 225	O 225	0	2
2	B	175	Total 175	O 175	0	0
2	C	138	Total 138	O 138	0	0
2	D	128	Total 129	O 129	0	1
2	E	187	Total 187	O 187	0	0
2	F	154	Total 154	O 154	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	137	Total 137	O 137	0	0
2	H	95	Total 95	O 95	0	0

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: OCHRATOXINASE

Chain A:

- Molecule 1: OCHRATOXINASE

Chain B:

- Molecule 1: OCHRATOXINASE

Chain C:

- Molecule 1: OCHRATOXINASE

Chain D:

V332 I333 F336 S339 N340 G341 E342 G343 L344 V345 K346 E347 S348 W349 Q353 K361 G365 K368 G376 A435 E438 K451 H455 V456 W457 L462 R476 N477 L480

• Molecule 1: OCHRATOXINASE

Chain E:

4%

89%   10%

SER THR ASP E46 A47 I58 R66 K88 L91 P105 H131 E145 Q148 Y154 E163 S181 L197 V202 L221 C222 I223 V244 M254 S255 R256 Q264 F265 R279 K299 V308 E313 W316 G323 T329 V332

I333 F336 S339 N340 G341 E342 G343 L344 V345 K346 E347 S348 W349 A350 K351 L352 G365 G376 R426 A435 H455 V456 W457 L462 D474 A475 R476 N477 L480

• Molecule 1: OCHRATOXINASE

Chain F:

4%

91%   8%

SER THR ASP GLU A47 K48 I58 P105 D119 S126 H131 R138 E145 Y154 E163 S181 G192 V202 Y206 L221 C222 I223 R239 V244 M254 S255 R256 R279 K299 V308 E313 W316 T329 V332 I333

F336 S339 N340 G341 E342 G343 L344 V345 K346 E347 S348 W349 A350 A354 G365 G376 A435 P450 K451 H455 L462 I468 N477 L480

• Molecule 1: OCHRATOXINASE

Chain G:

9%

88%   9%

SER THR ASP E46 A47 K48 Y53 I58 L65 R66 N67 V71 F79 V80 G81 S82 E83 A84 D85 I86 P87 K88 K89 Y90 L91 R92 S93 T94 Q95 S96 T97 H98 R99 P105 D119 H131 L147 Q148 N149 Y154 E163 V174 S181 V202 I223

R239 V244 V253 M254 S255 R256 Q264 F265 R279 K299 E306 H307 V308 E313 W316 T329 V332 I333 F336 L337 A338 S339 N340 GLU GLY GLY LEU V345 K346 E347 S348 W349 A350 G365 L375 G376 R395 A406 A407 R426 A435

E442 H455 V456 W457 L462 P470 R476 N477 L480

• Molecule 1: OCHRATOXINASE

Chain H:

9%

88%   10%

SER THR ASP GLU A47 K48 Y53 A54 I58 L65 R66 N67 A68 A78 F79 S82 E83 A84 K88 K89 Y90 L91 R92 S93 T94 Q95 S96 R99 P105 D119 Y120 Y121 N122 H131 R138 E145 Y154 E163 K166 S181 L197 V244

M254 S255 R256 Q264 F265 R279 V308 E313 W316 T329 V332 I333 F336 L337 A338 S339 N340 GLY GLU GLY LEU VAL K346 E347 S348 W349 A350 K351 L352 Q353 A354 G365 G376 G423 R426 Y429 E430 A435 L436 E437 E438 P450 H455

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	210.85Å 78.92Å 217.00Å 90.00° 105.05° 90.00°	Depositor
Resolution (Å)	37.79 – 2.20 37.76 – 2.20	Depositor EDS
% Data completeness (in resolution range)	87.6 (37.79-2.20) 87.6 (37.76-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.213 , 0.245 0.217 , 0.250	Depositor DCC
R_{free} test set	7696 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	20.1	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 38.5	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.000 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	27300	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.80 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.4166e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.56	0/3359	0.75	6/4557 (0.1%)
1	B	0.54	0/3333	0.78	8/4521 (0.2%)
1	C	0.55	0/3360	0.75	8/4557 (0.2%)
1	D	0.55	0/3339	0.75	8/4531 (0.2%)
1	E	0.56	0/3342	0.74	5/4536 (0.1%)
1	F	0.55	0/3335	0.81	12/4525 (0.3%)
1	G	0.54	1/3307 (0.0%)	0.77	11/4487 (0.2%)
1	H	0.54	0/3302	0.84	12/4479 (0.3%)
All	All	0.55	1/26677 (0.0%)	0.77	70/36193 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	306	GLU	CD-OE1	5.46	1.31	1.25

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	138	ARG	NE-CZ-NH2	-14.20	113.20	120.30
1	H	138[A]	ARG	NE-CZ-NH2	-13.90	113.35	120.30
1	H	138[B]	ARG	NE-CZ-NH2	-13.90	113.35	120.30
1	F	138	ARG	NE-CZ-NH1	13.75	127.18	120.30
1	H	138[A]	ARG	NE-CZ-NH1	12.84	126.72	120.30

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	3281	0	3261	29	1
1	B	3250	0	3238	22	0
1	C	3274	0	3263	23	1
1	D	3264	0	3243	19	0
1	E	3267	0	3244	27	1
1	F	3260	0	3245	24	1
1	G	3236	0	3214	20	1
1	H	3228	0	3212	22	1
2	A	225	0	0	2	0
2	B	175	0	0	1	0
2	C	138	0	0	2	0
2	D	129	0	0	1	0
2	E	187	0	0	3	0
2	F	154	0	0	3	0
2	G	137	0	0	1	0
2	H	95	0	0	2	0
All	All	27300	0	25920	172	3

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 3.

The worst 5 of 172 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:H:65:LEU:HG	1:H:68:ALA:HB2	1.55	0.88
1:C:341:GLY:HA2	1:C:344:LEU:HD12	1.60	0.82
1:A:166:LYS:HE2	2:A:2077:HOH:O	1.81	0.79
1:H:462:LEU:HD23	1:H:468:ILE:HD11	1.69	0.75
1:A:206:TYR:CD2	1:A:221:LEU:HD23	2.21	0.75

All (3) 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:F:239[A]:ARG:NH2	1:H:197:LEU:O[2_554]	1.77	0.43
1:A:197:LEU:O	1:C:239[A]:ARG:NH2[2_555]	1.83	0.37
1:E:197:LEU:O	1:G:239:ARG:NH2[2_554]	2.02	0.18

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	436/438 (100%)	421 (97%)	15 (3%)	0	100	100
1	B	430/438 (98%)	415 (96%)	15 (4%)	0	100	100
1	C	434/438 (99%)	419 (96%)	15 (4%)	0	100	100
1	D	434/438 (99%)	419 (96%)	15 (4%)	0	100	100
1	E	434/438 (99%)	416 (96%)	18 (4%)	0	100	100
1	F	433/438 (99%)	414 (96%)	19 (4%)	0	100	100
1	G	427/438 (98%)	412 (96%)	15 (4%)	0	100	100
1	H	426/438 (97%)	411 (96%)	15 (4%)	0	100	100
All	All	3454/3504 (99%)	3327 (96%)	127 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	333/333 (100%)	325 (98%)	8 (2%)	49	62
1	B	331/333 (99%)	324 (98%)	7 (2%)	53	67
1	C	334/333 (100%)	323 (97%)	11 (3%)	38	49
1	D	331/333 (99%)	320 (97%)	11 (3%)	38	49
1	E	331/333 (99%)	317 (96%)	14 (4%)	30	38
1	F	330/333 (99%)	320 (97%)	10 (3%)	41	53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	328/333 (98%)	312 (95%)	16 (5%)	25	31
1	H	327/333 (98%)	317 (97%)	10 (3%)	40	51
All	All	2645/2664 (99%)	2558 (97%)	87 (3%)	39	49

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	332	VAL
1	G	336	PHE
1	F	339	SER
1	G	239	ARG
1	G	395	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	113	HIS
1	B	169	ASN
1	E	148	GLN
1	F	280	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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, 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(Å²)	Q<0.9
1	A	436/438 (99%)	0.01	19 (4%)	34	32	12, 20, 48, 86	0
1	B	430/438 (98%)	-0.01	15 (3%)	44	42	13, 23, 48, 80	0
1	C	433/438 (98%)	0.05	19 (4%)	34	32	10, 22, 50, 82	0
1	D	435/438 (99%)	0.17	22 (5%)	28	26	12, 25, 49, 97	0
1	E	435/438 (99%)	0.01	18 (4%)	37	35	13, 21, 49, 92	0
1	F	434/438 (99%)	0.03	18 (4%)	37	35	10, 24, 51, 98	0
1	G	431/438 (98%)	0.26	40 (9%)	8	7	14, 28, 59, 89	0
1	H	429/438 (97%)	0.28	38 (8%)	9	8	13, 30, 62, 80	0
All	All	3463/3504 (98%)	0.10	189 (5%)	25	24	10, 24, 56, 98	0

The worst 5 of 189 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	343	GLY	9.0
1	E	344	LEU	8.4
1	A	343	GLY	8.2
1	F	343	GLY	7.3
1	D	341	GLY	6.9

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.