wwPDB X-ray Structure Validation Summary Report (i)

Sep 17, 2023 – 05:28 AM EDT

PDB ID	:	4WHB
Title	:	Crystal structure of phenylurea hydrolase B
Authors	:	Sugrue, E.; Carr, P.D.; Khurana, J.L.; Jackson, C.J.
Deposited on	:	2014-09-21
Resolution	:	2.96 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 (i) 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 (i)) were used in the production of this report:

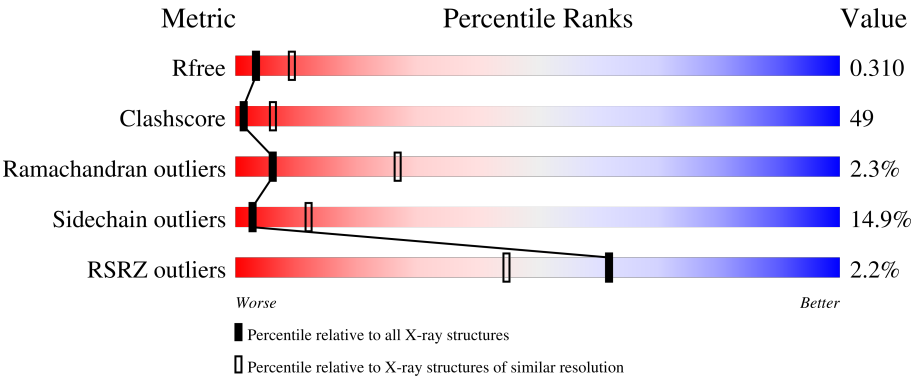
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.96 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}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	461	2% 39% 50% 10%
1	B	461	46% 43% 10%
1	C	461	% 45% 46% 8%
1	D	461	% 40% 46% 12%
1	E	461	2% 41% 48% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	461	6% 37% 53% 8%
1	G	461	3% 42% 46% 10%
1	H	461	3% 34% 52% 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 55405 atoms, of which 27459 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 Phenylurea hydrolase B.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	E	459	Total	C	H	N	O	S	0	0	0
			6918	2190	3430	618	670	10			
1	A	459	Total	C	H	N	O	S	0	0	0
			6915	2190	3427	618	670	10			
1	B	459	Total	C	H	N	O	S	0	2	0
			6944	2196	3445	622	671	10			
1	C	459	Total	C	H	N	O	S	0	2	0
			6941	2196	3442	622	671	10			
1	D	459	Total	C	H	N	O	S	0	0	0
			6917	2190	3429	618	670	10			
1	F	459	Total	C	H	N	O	S	0	0	0
			6916	2190	3428	618	670	10			
1	G	459	Total	C	H	N	O	S	0	0	0
			6917	2190	3429	618	670	10			
1	H	459	Total	C	H	N	O	S	0	0	0
			6917	2190	3429	618	670	10			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	4	Total	O	0	0
			4	4		
3	A	3	Total	O	0	0
			3	3		
3	B	3	Total	O	0	0
			3	3		

Continued on next page...

Continued from previous page...

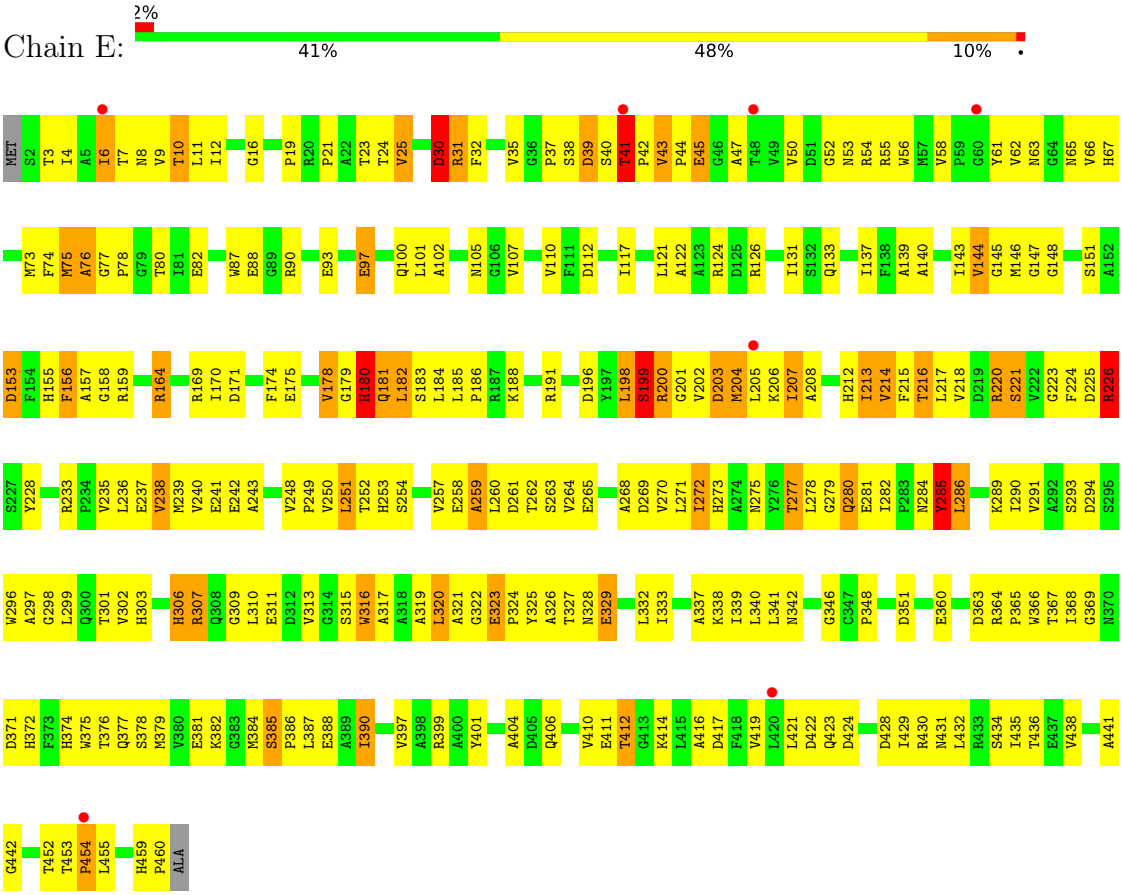
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	2	Total O 2 2	0	0
3	F	3	Total O 3 3	0	0
3	G	1	Total O 1 1	0	0
3	H	3	Total O 3 3	0	0

3 Residue-property plots (i)

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.

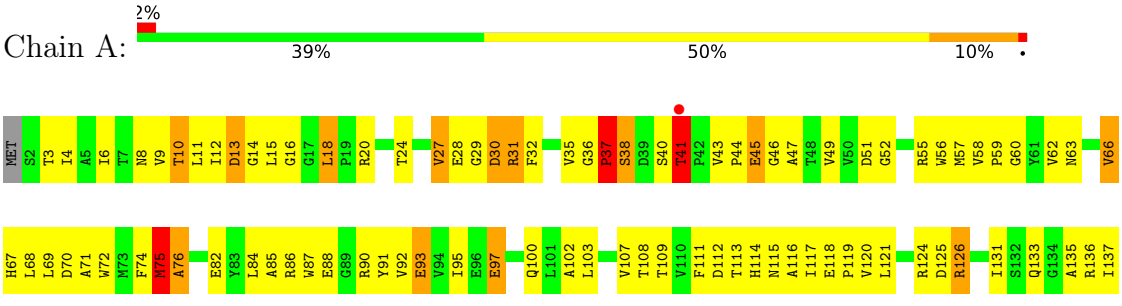
• Molecule 1: Phenylurea hydrolase B

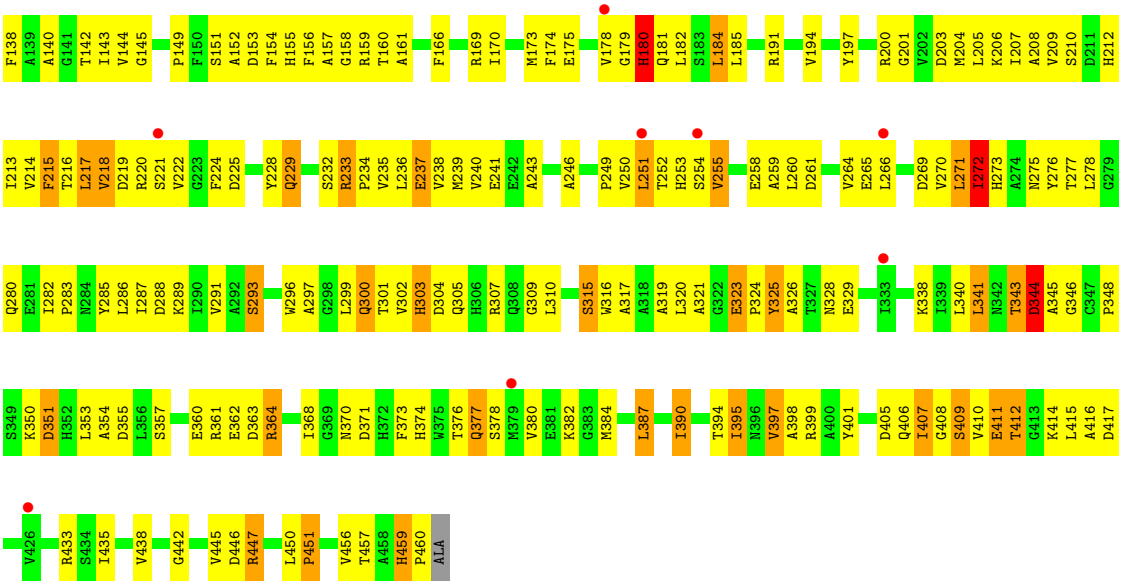
Chain E:



• Molecule 1: Phenylurea hydrolase B

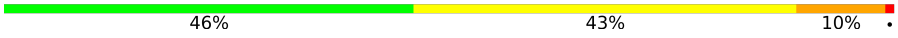
Chain A:

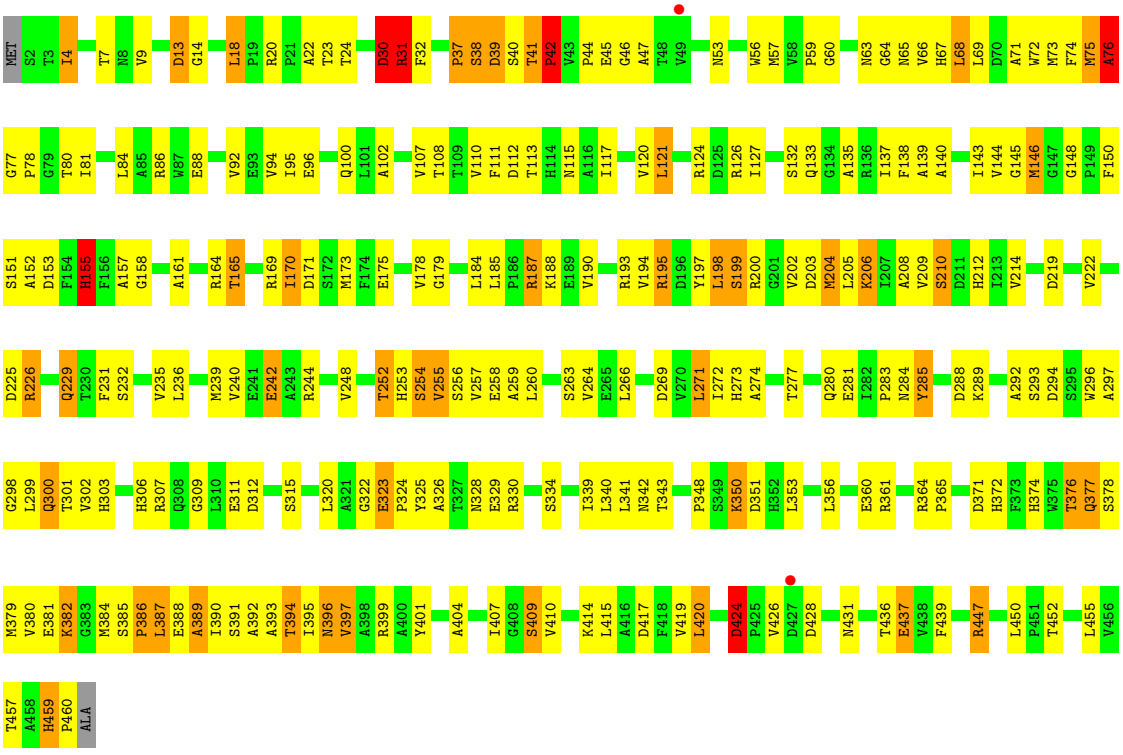




• Molecule 1: Phenylurea hydrolase B

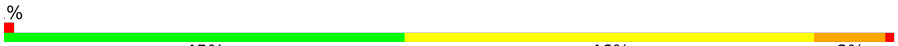
Chain B:

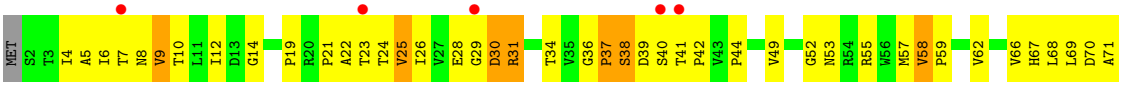


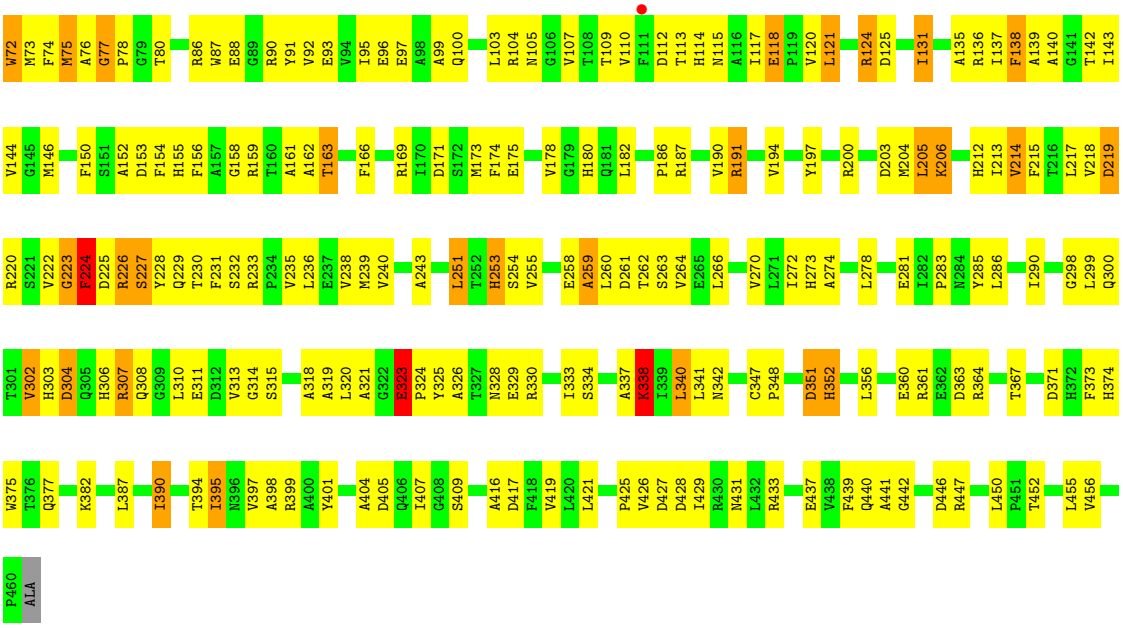


• Molecule 1: Phenylurea hydrolase B

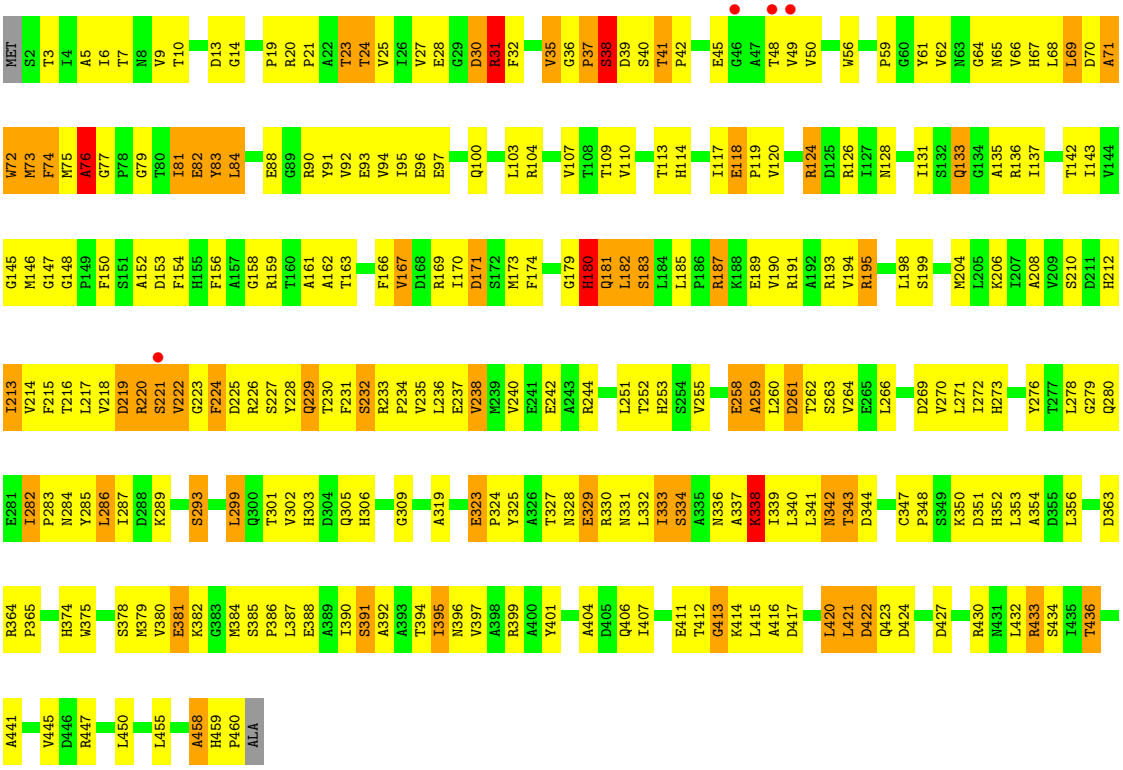
Chain C:



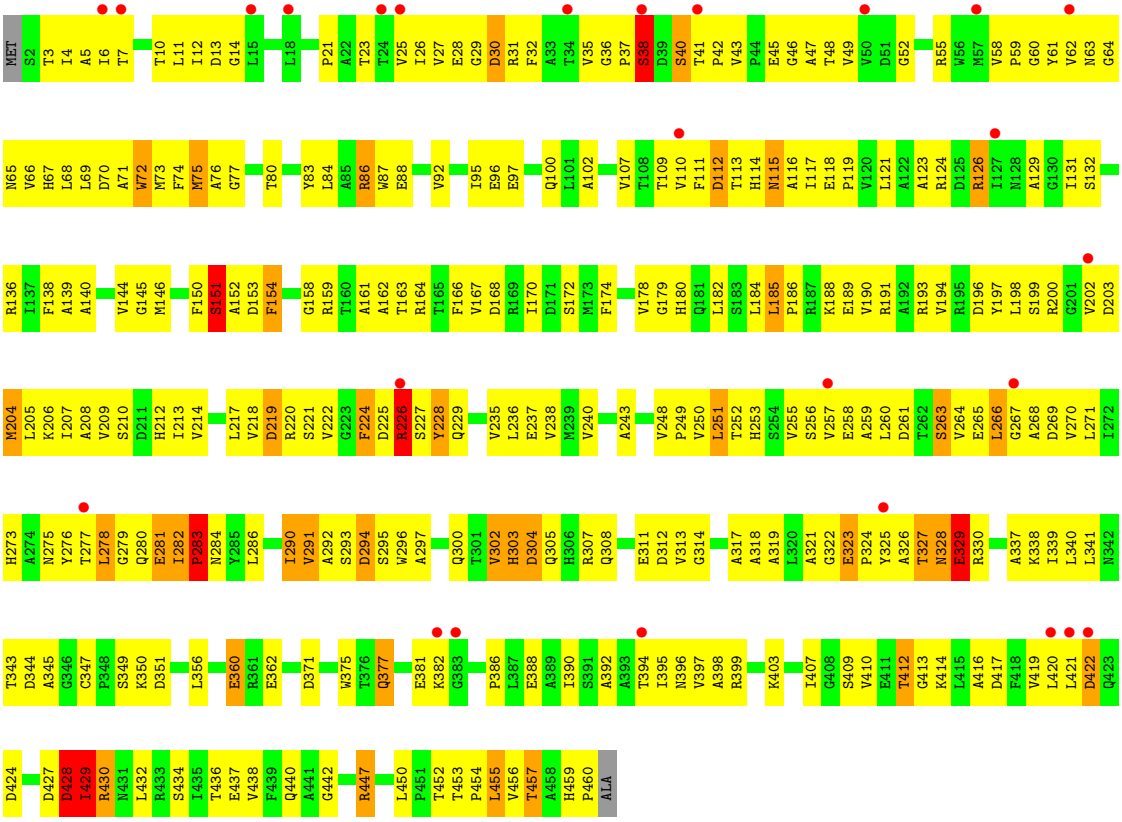




• Molecule 1: Phenylurea hydrolase B

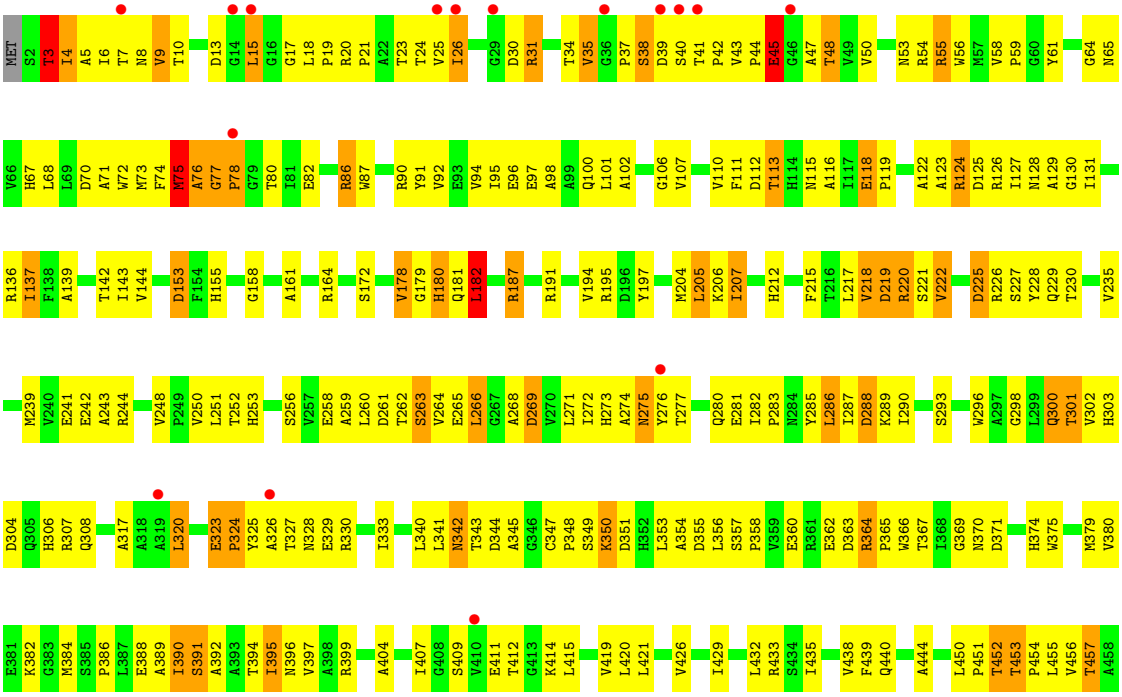


• Molecule 1: Phenylurea hydrolase B



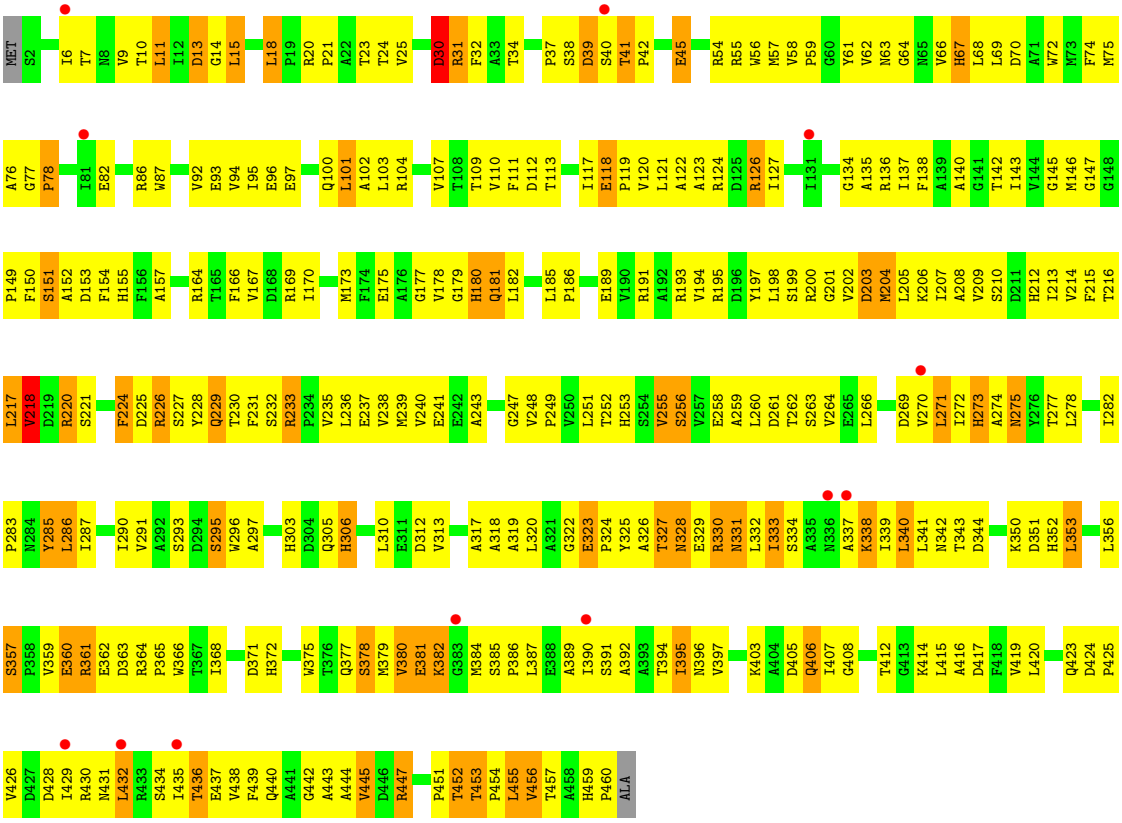
• Molecule 1: Phenylurea hydrolase B

Chain G:



• Molecule 1: Phenylurea hydrolase B

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	77.37Å 100.53Å 238.55Å 90.00° 98.37° 90.00°	Depositor
Resolution (Å)	39.59 – 2.96 39.59 – 2.96	Depositor EDS
% Data completeness (in resolution range)	98.9 (39.59-2.96) 89.7 (39.59-2.96)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 2.95Å)	Xtriage
Refinement program	PHENIX, REFMAC	Depositor
R, R_{free}	0.252 , 0.309 0.264 , 0.310	Depositor DCC
R_{free} test set	3768 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	31.9	Xtriage
Anisotropy	1.235	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 26.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.085 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	55405	wwPDB-VP
Average B, all atoms (Å^2)	41.0	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 24.56 % 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 3.7450e-03.*

[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

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.69	1/3560 (0.0%)	0.85	7/4857 (0.1%)
1	B	0.83	1/3576 (0.0%)	0.93	9/4878 (0.2%)
1	C	0.69	1/3576 (0.0%)	0.83	2/4878 (0.0%)
1	D	0.79	1/3560 (0.0%)	0.90	2/4857 (0.0%)
1	E	0.79	1/3560 (0.0%)	0.84	2/4857 (0.0%)
1	F	0.71	2/3560 (0.1%)	0.83	2/4857 (0.0%)
1	G	0.75	2/3560 (0.1%)	0.85	5/4857 (0.1%)
1	H	0.68	0/3560	0.84	0/4857
All	All	0.74	9/28512 (0.0%)	0.86	29/38898 (0.1%)

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	6
1	B	0	6
1	C	0	5
1	D	0	7
1	E	0	4
1	F	0	3
1	G	0	7
1	H	0	1
All	All	0	39

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	181	GLN	CD-OE1	6.14	1.37	1.24
1	G	300	GLN	CD-OE1	5.73	1.36	1.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	229	GLN	CD-OE1	5.53	1.36	1.24
1	F	72	TRP	CB-CG	-5.44	1.40	1.50
1	D	229	GLN	CD-OE1	5.33	1.35	1.24

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	226	ARG	NE-CZ-NH2	-8.43	116.09	120.30
1	F	283	PRO	CA-N-CD	-7.30	101.29	111.50
1	A	76	ALA	N-CA-C	-7.27	91.37	111.00
1	B	13	ASP	CB-CG-OD1	6.87	124.48	118.30
1	B	42	PRO	CA-N-CD	-6.51	102.39	111.50

There are no chirality outliers.

5 of 39 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	37	PRO	Peptide
1	E	30	ASP	Peptide
1	E	39	ASP	Peptide
1	E	41	THR	Peptide
1	E	6	ILE	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	3488	3427	3427	386	1
1	B	3499	3445	3437	299	0
1	C	3499	3442	3437	271	1
1	D	3488	3429	3429	369	5
1	E	3488	3430	3429	317	0
1	F	3488	3428	3428	382	1
1	G	3488	3429	3429	257	5
1	H	3488	3429	3429	427	1
2	A	1	0	0	0	0
3	A	3	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	3	0	0	0	0
3	D	2	0	0	2	0
3	E	4	0	0	0	0
3	F	3	0	0	1	0
3	G	1	0	0	1	0
3	H	3	0	0	1	0
All	All	27946	27459	27445	2687	7

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 49.

The worst 5 of 2687 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:6:ILE:HB	1:H:25:VAL:CG1	1.25	1.64
1:B:391:SER:CB	1:B:394:THR:HB	1.15	1.63
1:A:251:LEU:HD12	1:A:270:VAL:CB	1.46	1.45
1:E:224:PHE:CD1	1:E:278:LEU:HD11	1.51	1.43
1:F:322:GLY:HA2	1:F:326:ALA:CB	1.48	1.42

The worst 5 of 7 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:D:226:ARG:HE	1:G:180:HIS:HE2[1_455]	0.67	0.93
1:D:226:ARG:HE	1:G:180:HIS:NE2[1_455]	1.26	0.34
1:A:261:ASP:OD2	1:C:187:ARG:HH21[1_655]	1.42	0.18
1:F:219:ASP:OD2	1:H:164:ARG:NH2[1_655]	2.12	0.08
1:D:187:ARG:NH2	1:G:258:GLU:OE2[1_455]	2.13	0.07

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	457/461 (99%)	391 (86%)	57 (12%)	9 (2%)	7 30
1	B	459/461 (100%)	420 (92%)	32 (7%)	7 (2%)	10 38
1	C	459/461 (100%)	415 (90%)	35 (8%)	9 (2%)	7 30
1	D	457/461 (99%)	410 (90%)	33 (7%)	14 (3%)	4 19
1	E	457/461 (99%)	412 (90%)	31 (7%)	14 (3%)	4 19
1	F	457/461 (99%)	402 (88%)	45 (10%)	10 (2%)	6 28
1	G	457/461 (99%)	407 (89%)	40 (9%)	10 (2%)	6 28
1	H	457/461 (99%)	401 (88%)	46 (10%)	10 (2%)	6 28
All	All	3660/3688 (99%)	3258 (89%)	319 (9%)	83 (2%)	6 27

5 of 83 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	323	GLU
1	A	41	THR
1	A	255	VAL
1	B	31	ARG
1	B	41	THR

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	367/368 (100%)	313 (85%)	54 (15%)	3 12
1	B	368/368 (100%)	313 (85%)	55 (15%)	3 12
1	C	368/368 (100%)	327 (89%)	41 (11%)	6 22
1	D	367/368 (100%)	307 (84%)	60 (16%)	2 9
1	E	367/368 (100%)	315 (86%)	52 (14%)	3 13
1	F	367/368 (100%)	315 (86%)	52 (14%)	3 13
1	G	367/368 (100%)	308 (84%)	59 (16%)	2 10
1	H	367/368 (100%)	302 (82%)	65 (18%)	2 8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2938/2944 (100%)	2500 (85%)	438 (15%)	3 12

5 of 438 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	216	THR
1	F	281	GLU
1	H	229	GLN
1	D	238	VAL
1	D	433	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	100	GLN
1	G	273	HIS
1	F	212	HIS
1	A	370	ASN
1	F	253	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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

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^{th} 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	459/461 (99%)	0.16	9 (1%)	65	48	19, 40, 57, 93	0
1	B	459/461 (99%)	-0.09	2 (0%)	92	84	9, 26, 44, 70	0
1	C	459/461 (99%)	0.12	6 (1%)	77	61	16, 38, 60, 77	0
1	D	459/461 (99%)	-0.12	4 (0%)	84	71	10, 27, 48, 80	0
1	E	459/461 (99%)	0.03	7 (1%)	73	57	9, 32, 55, 80	0
1	F	459/461 (99%)	0.40	26 (5%)	23	14	23, 47, 68, 87	0
1	G	459/461 (99%)	0.11	16 (3%)	44	29	12, 34, 64, 92	0
1	H	459/461 (99%)	0.28	12 (2%)	56	39	22, 43, 58, 78	0
All	All	3672/3688 (99%)	0.11	82 (2%)	62	45	9, 37, 60, 93	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	41	THR	5.9
1	F	62	VAL	5.6
1	F	6	ILE	5.1
1	D	49	VAL	4.9
1	F	421	LEU	4.6

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

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^{th} 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^2)	Q<0.9
2	ZN	A	501	1/1	0.98	0.05	72,72,72,72	0

6.5 Other polymers

There are no such residues in this entry.