



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 02:45 PM GMT

PDB ID : 4LCD  
Title : Structure of an Rsp5xUbxSna3 complex: Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3  
Authors : Kamadurai, H.B.; Miller, D.; Schulman, B.A.  
Deposited on : 2013-06-21  
Resolution : 3.10 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

---

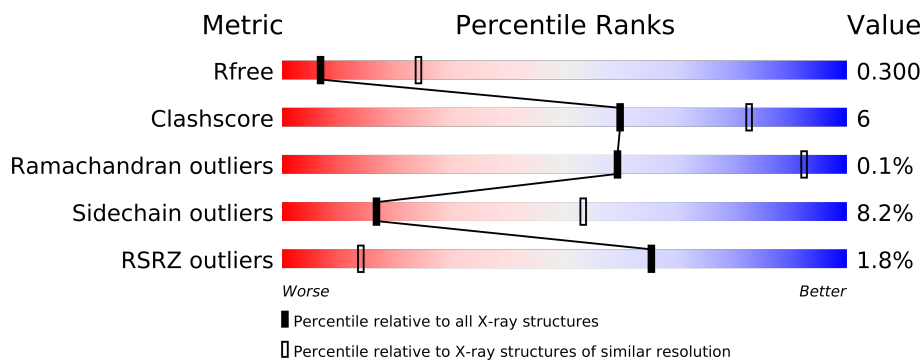
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.10 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))
$R_{free}$	66092	1007 (3.18-3.02)
Clashscore	79885	1078 (3.16-3.04)
Ramachandran outliers	78287	1044 (3.16-3.04)
Sidechain outliers	78261	1044 (3.16-3.04)
RSRZ outliers	66119	1008 (3.18-3.02)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	432	
1	B	432	
2	C	24	
2	D	24	
3	E	83	
3	F	83	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7944 atoms, of which 0 are hydrogen and 0 are deuterium.

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 E3 ubiquitin-protein ligase RSP5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	419	Total	C	N	O	S	0	0	0
			3389	2169	585	624	11			
1	B	413	Total	C	N	O	S	0	0	0
			3289	2110	565	604	10			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	378	GLY	-	EXPRESSION TAG	UNP P39940
A	379	SER	-	EXPRESSION TAG	UNP P39940
A	380	GLY	-	EXPRESSION TAG	UNP P39940
A	381	GLY	-	EXPRESSION TAG	UNP P39940
A	382	SER	-	EXPRESSION TAG	UNP P39940
A	455	LEU	CYS	ENGINEERED MUTATION	UNP P39940
A	517	GLY	CYS	ENGINEERED MUTATION	UNP P39940
A	721	ALA	CYS	ENGINEERED MUTATION	UNP P39940
B	378	GLY	-	EXPRESSION TAG	UNP P39940
B	379	SER	-	EXPRESSION TAG	UNP P39940
B	380	GLY	-	EXPRESSION TAG	UNP P39940
B	381	GLY	-	EXPRESSION TAG	UNP P39940
B	382	SER	-	EXPRESSION TAG	UNP P39940
B	455	LEU	CYS	ENGINEERED MUTATION	UNP P39940
B	517	GLY	CYS	ENGINEERED MUTATION	UNP P39940
B	721	ALA	CYS	ENGINEERED MUTATION	UNP P39940

- Molecule 2 is a protein called Protein SNA3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	8	Total	C	N	O	0	0	0
			51	34	8	9			
2	D	14	Total	C	N	O	0	0	0
			80	51	14	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	125	ALA	LYS	ENGINEERED MUTATION	UNP P14359
D	125	ALA	LYS	ENGINEERED MUTATION	UNP P14359

- Molecule 3 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	75	Total	C	N	O	S	0	0	0
			565	353	98	112	2			
3	F	75	Total	C	N	O	S	0	0	0
			570	359	96	113	2			

There are 18 discrepancies between the modelled and reference sequences:

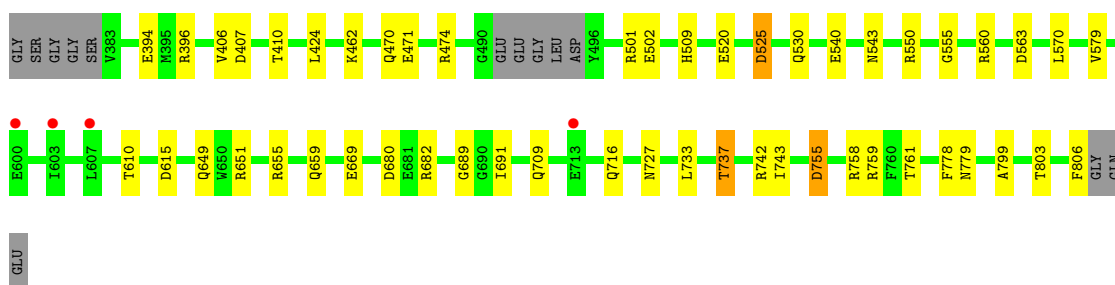
Chain	Residue	Modelled	Actual	Comment	Reference
E	-7	MET	-	EXPRESSION TAG	UNP P0CG48
E	-6	GLY	-	EXPRESSION TAG	UNP P0CG48
E	-5	HIS	-	EXPRESSION TAG	UNP P0CG48
E	-4	HIS	-	EXPRESSION TAG	UNP P0CG48
E	-3	HIS	-	EXPRESSION TAG	UNP P0CG48
E	-2	HIS	-	EXPRESSION TAG	UNP P0CG48
E	-1	HIS	-	EXPRESSION TAG	UNP P0CG48
E	0	HIS	-	EXPRESSION TAG	UNP P0CG48
E	75	CYS	GLY	ENGINEERED MUTATION	UNP P0CG48
F	-7	MET	-	EXPRESSION TAG	UNP P0CG48
F	-6	GLY	-	EXPRESSION TAG	UNP P0CG48
F	-5	HIS	-	EXPRESSION TAG	UNP P0CG48
F	-4	HIS	-	EXPRESSION TAG	UNP P0CG48
F	-3	HIS	-	EXPRESSION TAG	UNP P0CG48
F	-2	HIS	-	EXPRESSION TAG	UNP P0CG48
F	-1	HIS	-	EXPRESSION TAG	UNP P0CG48
F	0	HIS	-	EXPRESSION TAG	UNP P0CG48
F	75	CYS	GLY	ENGINEERED MUTATION	UNP P0CG48

### 3 Residue-property plots

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.

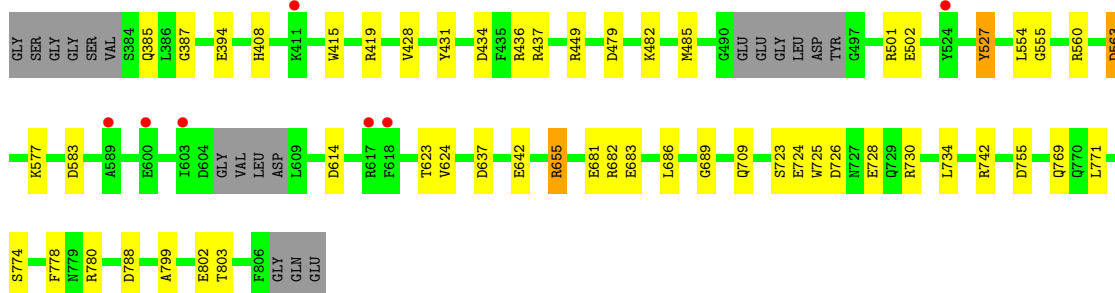
- Molecule 1: E3 ubiquitin-protein ligase RSP5

Chain A: 



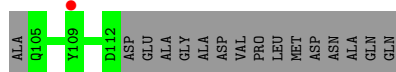
- Molecule 1: E3 ubiquitin-protein ligase RSP5

Chain B: 



- Molecule 2: Protein SNA3

Chain C: 



- Molecule 2: Protein SNA3

Chain D: 



- Molecule 3: Ubiquitin

Chain E: 



● Molecule 3: Ubiquitin

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.05Å 78.92Å 96.72Å 90.00° 101.67° 90.00°	Depositor
Resolution (Å)	29.40 – 3.10 29.40 – 3.10	Depositor EDS
% Data completeness (in resolution range)	95.0 (29.40-3.10) 91.4 (29.40-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.48 (at 3.11Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.7.3_928)	Depositor
R, $R_{free}$	0.251 , 0.299 0.250 , 0.300	Depositor DCC
$R_{free}$ test set	1051 reflections (5.13%)	DCC
Wilson B-factor (Å <sup>2</sup> )	73.2	Xtriage
Anisotropy	0.488	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 50.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 21344 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	7944	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.16% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.47	0/3472	0.57	0/4702
1	B	0.47	0/3371	0.59	0/4574
2	C	0.41	0/53	0.76	0/74
2	D	0.28	0/82	0.67	0/114
3	E	0.36	0/571	0.56	0/774
3	F	0.33	0/576	0.56	0/782
All	All	0.45	0/8125	0.58	0/11020

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3389	0	0	19	0
1	B	3289	0	0	21	0
2	C	51	0	0	0	0
2	D	80	0	0	6	0
3	E	565	0	0	3	0
3	F	570	0	0	6	0
All	All	7944	0	0	48	0



Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (48) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:774:SER:N	3:F:73:LEU:O	2.25	0.68
1:B:415:TRP:NE1	2:D:104:ALA:O	2.29	0.66
1:B:771:LEU:O	3:F:40:GLN:NE2	2.29	0.66
1:B:449:ARG:CG	1:B:449:ARG:NH1	2.59	0.65
1:B:563:ASP:OD1	1:B:742:ARG:NH1	2.32	0.63
1:B:387:GLY:O	1:B:419:ARG:NH1	2.32	0.62
1:B:502:GLU:OE2	3:F:74:ARG:NH2	2.33	0.62
3:E:42:ARG:NE	3:E:49:GLN:OE1	2.34	0.61
1:B:385:GLN:O	1:B:449:ARG:NH1	2.35	0.60
3:F:71:LEU:N	3:F:71:LEU:CD1	2.68	0.56
1:A:579:VAL:O	1:A:649:GLN:NE2	2.38	0.56
2:D:114:GLU:O	2:D:115:ALA:C	2.44	0.55
1:B:479:ASP:OD1	1:B:482:LYS:NZ	2.41	0.53
1:A:755:ASP:OD1	1:A:755:ASP:N	2.40	0.53
1:A:799:ALA:O	1:A:803:THR:OG1	2.27	0.53
1:A:520:GLU:OE2	1:A:530:GLN:NE2	2.44	0.50
1:B:655:ARG:CG	1:B:655:ARG:NH1	2.75	0.49
1:A:525:ASP:N	1:A:525:ASP:OD1	2.44	0.49
1:B:555:GLY:O	1:B:560:ARG:N	2.46	0.49
3:E:18:GLU:N	3:E:21:ASP:OD2	2.45	0.49
1:A:550:ARG:NE	1:A:669:GLU:OE1	2.46	0.49
1:B:614:ASP:OD1	1:B:623:THR:OG1	2.31	0.48
1:B:799:ALA:O	1:B:803:THR:OG1	2.32	0.48
1:B:527:TYR:CD2	1:B:527:TYR:N	2.82	0.48
1:B:408:HIS:ND1	2:D:109:TYR:OH	2.46	0.48
3:F:72:ARG:O	3:F:72:ARG:CG	2.62	0.47
1:A:502:GLU:OE1	3:E:74:ARG:NH2	2.48	0.47
1:A:501:ARG:NH2	1:A:779:ASN:OD1	2.48	0.47
1:A:659:GLN:N	1:A:659:GLN:OE1	2.48	0.46
1:A:540:GLU:OE1	1:A:543:ASN:ND2	2.48	0.46
1:B:802:GLU:CB	3:F:71:LEU:CD2	2.95	0.45
1:A:755:ASP:OD2	1:A:758:ARG:NE	2.49	0.45
2:D:109:TYR:CD1	2:D:109:TYR:O	2.70	0.45
1:A:689:GLY:O	1:A:742:ARG:NE	2.51	0.44
1:B:725:TRP:O	1:B:730:ARG:NE	2.51	0.44
1:A:555:GLY:O	1:A:560:ARG:N	2.49	0.44
2:D:110:ASP:C	2:D:112:ASP:N	2.72	0.43
1:A:733:LEU:O	1:A:737:THR:OG1	2.35	0.43
1:B:689:GLY:O	1:B:742:ARG:NE	2.52	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:683:GLU:O	1:B:686:LEU:N	2.52	0.42
1:A:471:GLU:OE2	1:A:474:ARG:NH2	2.54	0.41
1:B:577:LYS:NZ	1:B:681:GLU:OE1	2.53	0.41
2:D:109:TYR:C	2:D:111:GLU:N	2.72	0.41
1:A:394:GLU:O	1:A:406:VAL:N	2.54	0.41
1:A:691:ILE:N	1:A:743:ILE:O	2.54	0.41
1:A:615:ASP:OD1	1:A:655:ARG:NH2	2.54	0.41
1:A:407:ASP:OD2	1:A:410:THR:OG1	2.39	0.41
1:B:434:ASP:OD1	1:B:437:ARG:NH1	2.54	0.40

There are no symmetry-related clashes.

## 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	415/432 (96%)	390 (94%)	25 (6%)	0	100	100
1	B	407/432 (94%)	366 (90%)	41 (10%)	0	100	100
2	C	6/24 (25%)	2 (33%)	4 (67%)	0	100	100
2	D	12/24 (50%)	6 (50%)	5 (42%)	1 (8%)	1	9
3	E	73/83 (88%)	63 (86%)	10 (14%)	0	100	100
3	F	73/83 (88%)	65 (89%)	8 (11%)	0	100	100
All	All	986/1078 (92%)	892 (90%)	93 (9%)	1 (0%)	59	92

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	111	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	353/386 (92%)	332 (94%)	21 (6%)	28	68
1	B	333/386 (86%)	307 (92%)	26 (8%)	18	57
2	C	3/18 (17%)	3 (100%)	0	100	100
2	D	3/18 (17%)	3 (100%)	0	100	100
3	E	60/76 (79%)	50 (83%)	10 (17%)	3	11
3	F	62/76 (82%)	52 (84%)	10 (16%)	3	13
All	All	814/960 (85%)	747 (92%)	67 (8%)	17	53

All (67) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	396	ARG
1	A	424	LEU
1	A	462	LYS
1	A	470	GLN
1	A	509	HIS
1	A	525	ASP
1	A	563	ASP
1	A	570	LEU
1	A	610	THR
1	A	651	ARG
1	A	680	ASP
1	A	682	ARG
1	A	709	GLN
1	A	716	GLN
1	A	727	ASN
1	A	737	THR
1	A	755	ASP
1	A	759	ARG
1	A	761	THR
1	A	778	PHE
1	A	806	PHE
1	B	394	GLU
1	B	428	VAL
1	B	431	TYR
1	B	436	ARG
1	B	485	MET
1	B	501	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	527	TYR
1	B	554	LEU
1	B	563	ASP
1	B	583	ASP
1	B	624	VAL
1	B	637	ASP
1	B	642	GLU
1	B	655	ARG
1	B	682	ARG
1	B	709	GLN
1	B	723	SER
1	B	724	GLU
1	B	726	ASP
1	B	728	GLU
1	B	734	LEU
1	B	755	ASP
1	B	769	GLN
1	B	778	PHE
1	B	780	ARG
1	B	788	ASP
3	E	1	MET
3	E	9	THR
3	E	13	ILE
3	E	32	ASP
3	E	42	ARG
3	E	43	LEU
3	E	63	LYS
3	E	72	ARG
3	E	74	ARG
3	E	75	CYS
3	F	9	THR
3	F	11	LYS
3	F	13	ILE
3	F	14	THR
3	F	48	LYS
3	F	55	THR
3	F	56	LEU
3	F	67	LEU
3	F	71	LEU
3	F	75	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	419/432 (96%)	0.13	4 (0%) 79 23	43, 84, 147, 180	0
1	B	413/432 (95%)	0.20	7 (1%) 67 15	46, 92, 140, 206	0
2	C	8/24 (33%)	0.60	1 (12%) 5 1	107, 128, 133, 134	0
2	D	14/24 (58%)	0.51	0 100 100	113, 143, 157, 161	0
3	E	75/83 (90%)	0.32	2 (2%) 52 8	84, 124, 183, 204	0
3	F	75/83 (90%)	0.43	4 (5%) 25 4	94, 122, 143, 157	0
All	All	1004/1078 (93%)	0.20	18 (1%) 65 14	43, 96, 149, 206	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	603	ILE	6.7
1	B	589	ALA	3.5
3	F	2	GLN	3.2
3	F	75	CYS	2.9
1	A	607	LEU	2.8
3	E	2	GLN	2.8
3	F	1	MET	2.7
1	A	600	GLU	2.5
1	B	524	TYR	2.5
3	E	1	MET	2.5
1	B	411	LYS	2.3
1	B	600	GLU	2.2
3	F	3	ILE	2.2
1	B	618	PHE	2.2
2	C	109	TYR	2.2
1	B	617	ARG	2.2
1	A	603	ILE	2.0
1	A	713	GLU	2.0

## 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 carbohydrates 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.