Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 11:52 PM GMT

PDB ID	:	1YXA
Title	:	Serpina3n, a murine orthologue of human antichymotrypsin
Authors	:	Horvath, A.J.; Irving, J.A.; Law, R.H.; Rossjohn, J.; Bottomley, S.P.; Quinsey, N.S.; Pike, R.N.; Coughlin, P.B.; Whisstock, J.C.
Deposited on	:	2005-02-20
Resolution	:	2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

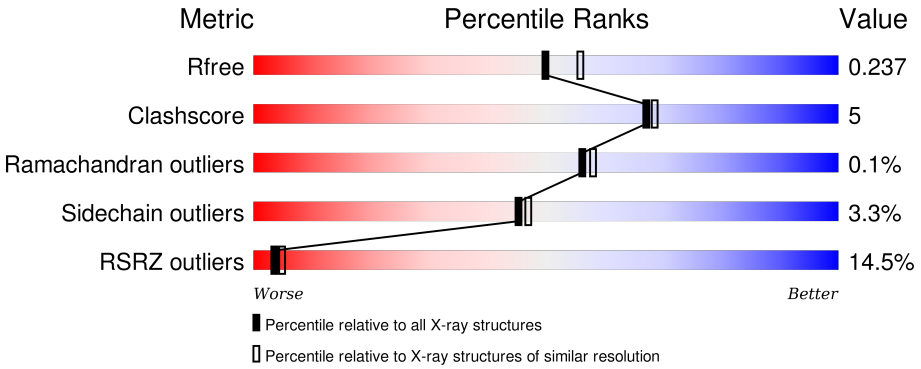
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.10 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}	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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 >=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 <=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	398	15% 83% 10% • 7%
1	B	398	12% 78% 14% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6137 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 serine (or cysteine) proteinase inhibitor, clade A, member 3N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	372	Total	C	N	O	S	0	1	0
			2874	1855	461	548	10			
1	B	368	Total	C	N	O	S	0	1	0
			2867	1849	461	548	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	69	THR	LYS	CONFLICT	UNP Q91WP6
A	84	LEU	VAL	CONFLICT	UNP Q91WP6
A	85	VAL	MET	CONFLICT	UNP Q91WP6
B	69	THR	LYS	CONFLICT	UNP Q91WP6
B	84	LEU	VAL	CONFLICT	UNP Q91WP6
B	85	VAL	MET	CONFLICT	UNP Q91WP6

- Molecule 2 is water.

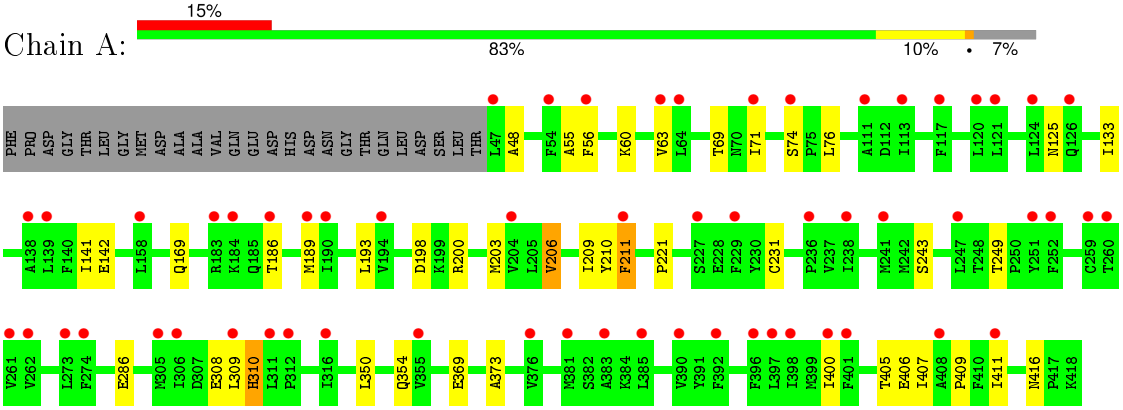
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	185	Total	O	0	0
			185	185		
2	B	211	Total	O	0	0
			211	211		

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 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.

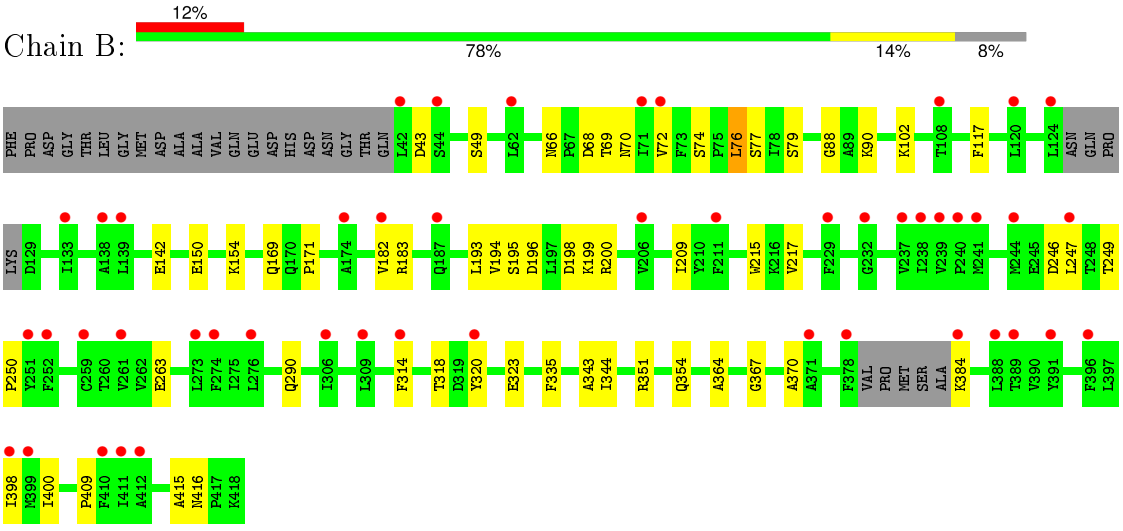
- Molecule 1: serine (or cysteine) proteinase inhibitor, clade A, member 3N

Chain A:



- Molecule 1: serine (or cysteine) proteinase inhibitor, clade A, member 3N

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	83.34Å 92.62Å 118.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.01 – 2.10 72.92 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.1 (30.01-2.10) 95.1 (72.92-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.37 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.193 , 0.239 0.194 , 0.237	Depositor DCC
R_{free} test set	2083 reflections (4.05%)	DCC
Wilson B-factor (Å^2)	28.6	Xtriage
Anisotropy	0.750	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 54117 reflections	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	6137	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.72% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.49	0/2931	0.59	0/3979
1	B	0.49	0/2919	0.59	1/3958 (0.0%)
All	All	0.49	0/5850	0.59	1/7937 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	76	LEU	CA-CB-CG	5.19	127.23	115.30

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	2874	0	2835	29	0
1	B	2867	0	2860	26	0
2	A	185	0	0	2	0
2	B	211	0	0	3	0
All	All	6137	0	5695	55	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 5.

All (55) 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:55:ALA:HB1	1:A:411:ILE:HD12	1.61	0.83
1:B:351:ARG:HD2	2:B:506:HOH:O	1.77	0.83
1:A:48:ALA:HA	1:A:407:ILE:HD13	1.64	0.80
1:A:141:ILE:HD11	1:A:203:MET:HE2	1.64	0.79
1:A:141:ILE:HD11	1:A:203:MET:CE	2.15	0.76
1:A:308:GLU:OE2	1:A:310:HIS:HE1	1.74	0.70
1:A:56:PHE:O	1:A:60:LYS:HG2	1.91	0.70
1:B:217:VAL:HG11	1:B:247:LEU:CD1	2.28	0.64
1:B:198:ASP:OD1	1:B:200:ARG:HD3	1.99	0.63
1:A:141:ILE:HD13	1:A:203:MET:HG3	1.80	0.63
1:A:198:ASP:OD1	1:A:200:ARG:HD3	1.98	0.63
1:A:141:ILE:HD12	1:A:350:LEU:HD21	1.83	0.60
1:B:70:ASN:HD21	1:B:416:ASN:H	1.47	0.60
1:A:74:SER:HB2	1:A:209:ILE:HG21	1.84	0.59
1:A:231:CYS:SG	1:A:416:ASN:HB2	2.42	0.59
1:B:400:ILE:O	1:B:409:PRO:HD2	2.03	0.58
1:A:221:PRO:HB2	1:A:373:ALA:HB2	1.86	0.57
1:A:141:ILE:CD1	1:A:203:MET:CE	2.84	0.56
1:A:405:THR:O	1:A:406:GLU:HB2	2.06	0.55
1:A:55:ALA:HB1	1:A:411:ILE:CD1	2.35	0.53
1:A:308:GLU:OE2	1:A:310:HIS:CE1	2.58	0.53
1:B:323[A]:GLU:HG3	1:B:335:PHE:CD1	2.44	0.53
1:A:63:VAL:HG22	1:A:71:ILE:CD1	2.40	0.52
1:A:354:GLN:NE2	2:A:458:HOH:O	2.43	0.51
1:A:141:ILE:HD11	1:A:203:MET:HE3	1.93	0.50
1:A:60:LYS:HE2	1:A:286:GLU:O	2.12	0.49
1:B:74:SER:HB2	1:B:209:ILE:HG21	1.94	0.48
1:B:66:ASN:O	1:B:69:THR:HG22	2.13	0.48
1:B:246:ASP:OD1	1:B:384:LYS:NZ	2.43	0.47
1:B:70:ASN:HD22	1:B:415:ALA:H	1.63	0.46
1:B:217:VAL:HG11	1:B:247:LEU:HD11	1.99	0.45
1:B:215:TRP:CD1	1:B:367:GLY:HA2	2.51	0.45
1:A:141:ILE:CD1	1:A:203:MET:HE2	2.39	0.45
1:B:150:GLU:HG2	1:B:154:LYS:HD2	1.98	0.45
1:A:189:MET:HG2	1:A:210:TYR:CD2	2.53	0.44
1:B:384:LYS:N	2:B:435:HOH:O	2.49	0.44
1:B:72:VAL:O	1:B:320:TYR:HE2	2.00	0.44
1:B:194:VAL:HG13	1:B:195:SER:O	2.18	0.43
1:A:133:ILE:HD13	1:A:211:PHE:HB2	2.01	0.43
1:A:243:SER:HA	1:A:309:LEU:O	2.19	0.43
1:B:88:GLY:HA3	1:B:344:ILE:HG13	1.99	0.43
1:A:193:LEU:HD21	1:A:206:VAL:HG13	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:364:ALA:HB1	1:B:370:ALA:HA	2.01	0.42
1:B:90:LYS:HE3	1:B:343:ALA:HB2	2.01	0.42
1:A:400:ILE:O	1:A:409:PRO:HD2	2.20	0.42
1:B:354:GLN:NE2	2:B:497:HOH:O	2.52	0.42
1:B:171:PRO:HG3	1:B:199:LYS:HG2	2.01	0.42
1:B:314:PHE:HZ	1:B:398:ILE:HD11	1.84	0.42
1:B:182:VAL:HG21	1:B:193:LEU:CD2	2.50	0.41
1:B:74:SER:HB3	1:B:77:SER:HB3	2.01	0.41
1:A:193:LEU:CD2	1:A:206:VAL:HG13	2.50	0.41
1:B:79:SER:HB3	1:B:117:PHE:HZ	1.86	0.41
1:A:186:THR:HB	1:A:189:MET:HB3	2.02	0.40
1:A:125:ASN:N	2:A:597:HOH:O	2.53	0.40
1:B:250:PRO:HB2	1:B:263:GLU:HB3	2.03	0.40

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	371/398 (93%)	363 (98%)	8 (2%)	0	100 100
1	B	363/398 (91%)	359 (99%)	3 (1%)	1 (0%)	46 45
All	All	734/796 (92%)	722 (98%)	11 (2%)	1 (0%)	56 58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	102	LYS

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	305/350 (87%)	296 (97%)	9 (3%)	48 51
1	B	309/350 (88%)	298 (96%)	11 (4%)	42 43
All	All	614/700 (88%)	594 (97%)	20 (3%)	45 47

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	69	THR
1	A	76	LEU
1	A	142	GLU
1	A	169	GLN
1	A	206	VAL
1	A	211	PHE
1	A	249	THR
1	A	310	HIS
1	A	369	GLU
1	B	43	ASP
1	B	49	SER
1	B	68	ASP
1	B	76	LEU
1	B	142	GLU
1	B	169	GLN
1	B	183	ARG
1	B	196	ASP
1	B	249	THR
1	B	290	GLN
1	B	318	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	ASN
1	A	146	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	169	GLN
1	A	310	HIS
1	A	338	GLN
1	A	354	GLN
1	B	70	ASN
1	B	122	GLN
1	B	338	GLN
1	B	354	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 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^{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	372/398 (93%)	0.99	59 (15%) 3 4	18, 32, 60, 75	0
1	B	368/398 (92%)	0.90	48 (13%) 5 6	19, 31, 53, 66	0
All	All	740/796 (92%)	0.95	107 (14%) 3 5	18, 32, 58, 75	0

All (107) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	383	ALA	11.0
1	A	385	LEU	5.8
1	B	378	PHE	4.5
1	A	190	ILE	4.0
1	A	124	LEU	3.4
1	B	124	LEU	3.3
1	B	261	VAL	3.2
1	A	398	ILE	3.1
1	A	400	ILE	3.1
1	A	117	PHE	3.1
1	B	108	THR	3.0
1	A	247	LEU	3.0
1	A	273	LEU	2.8
1	B	71	ILE	2.8
1	A	194	VAL	2.8
1	B	239	VAL	2.8
1	A	138	ALA	2.8
1	A	64	LEU	2.7
1	A	311	LEU	2.7
1	A	204	VAL	2.7
1	A	186	THR	2.7
1	B	396	PHE	2.7
1	A	305	MET	2.6
1	A	274	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	252	PHE	2.6
1	A	111	ALA	2.6
1	B	259	CYS	2.6
1	B	398	ILE	2.6
1	B	251	TYR	2.6
1	A	259	CYS	2.6
1	A	121	LEU	2.6
1	B	273	LEU	2.6
1	A	113	ILE	2.6
1	B	244	MET	2.6
1	B	62	LEU	2.5
1	B	388	LEU	2.5
1	B	241	MET	2.5
1	A	261	VAL	2.5
1	A	309	LEU	2.5
1	A	397	LEU	2.5
1	B	133	ILE	2.5
1	A	252	PHE	2.5
1	A	355	VAL	2.5
1	B	120	LEU	2.4
1	A	262	VAL	2.4
1	A	47	LEU	2.4
1	A	54	PHE	2.4
1	A	63	VAL	2.4
1	A	183	ARG	2.4
1	A	241	MET	2.4
1	B	391	TYR	2.4
1	B	237	VAL	2.3
1	A	56	PHE	2.3
1	B	72	VAL	2.3
1	B	44	SER	2.3
1	B	232	GLY	2.3
1	B	211	PHE	2.3
1	A	392	PHE	2.3
1	B	182	VAL	2.3
1	B	247	LEU	2.3
1	A	396	PHE	2.3
1	B	206	VAL	2.3
1	A	71	ILE	2.3
1	B	238	ILE	2.3
1	B	306	ILE	2.2
1	A	251	TYR	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	314	PHE	2.2
1	A	236	PRO	2.2
1	B	399	MET	2.2
1	B	411	ILE	2.2
1	A	184	LYS	2.2
1	B	139	LEU	2.2
1	B	274	PHE	2.2
1	B	174	ALA	2.2
1	B	276	LEU	2.2
1	A	390	VAL	2.2
1	B	229	PHE	2.2
1	A	376	VAL	2.2
1	B	412	ALA	2.2
1	A	306	ILE	2.2
1	A	381	MET	2.1
1	B	187	GLN	2.1
1	A	158	LEU	2.1
1	A	74	SER	2.1
1	B	389	THR	2.1
1	A	126	GLN	2.1
1	A	411	ILE	2.1
1	A	120	LEU	2.1
1	A	139	LEU	2.1
1	B	42	LEU	2.1
1	A	401	PHE	2.1
1	B	371	ALA	2.1
1	B	240	PRO	2.1
1	B	138	ALA	2.1
1	A	211	PHE	2.1
1	B	309	LEU	2.1
1	A	408	ALA	2.1
1	B	384	LYS	2.1
1	A	227	SER	2.1
1	A	229	PHE	2.1
1	B	410	PHE	2.1
1	A	312	PRO	2.0
1	B	320	TYR	2.0
1	A	316	ILE	2.0
1	A	260	THR	2.0
1	A	238	ILE	2.0
1	A	189	MET	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.