Full wwPDB X-ray Structure Validation Report (i)

May 16, 2020 – 12:26 am BST

PDB ID	:	1SGP
Title	:	ALA 18 VARIANT OF TURKEY OVOMUCOID INHIBITOR THIRD DOMAIN COMPLEXED WITH STREPTOMYCES GRISEUS PROTEINASE B
Authors	:	Huang, K.; James, M.N.G.
Deposited on	:	1995-05-26
Resolution	:	1.40 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

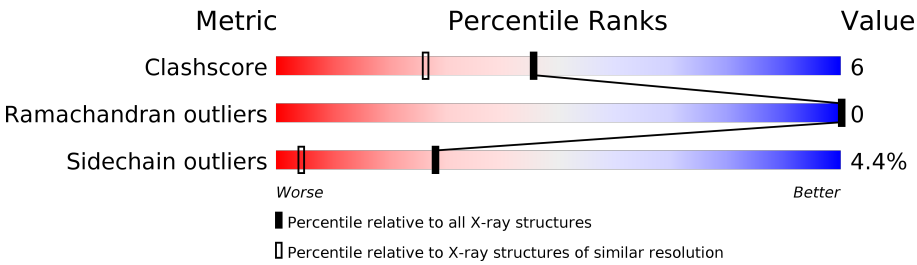
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.40 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	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)

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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	185	81% 16% • •
2	I	51	67% 20% 12% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1880 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 STREPTOMYCES GRISEUS PROTEINASE B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	185	1310	801	228	275	6	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	235A	VAL	SER	CONFLICT	UNP P00777

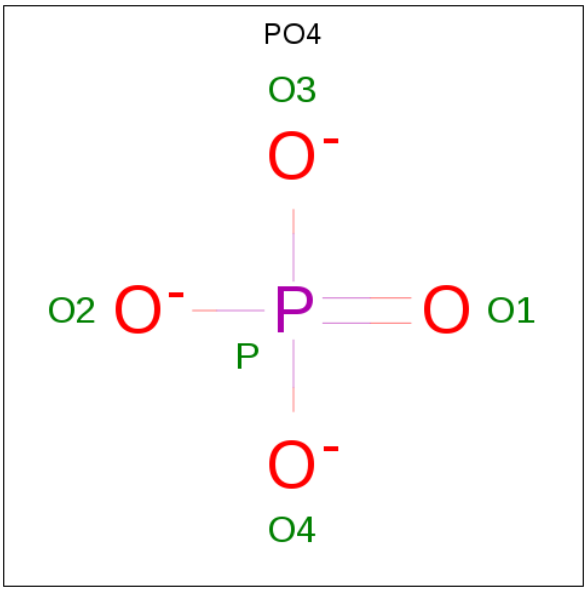
- Molecule 2 is a protein called TURKEY OVOMUCOID INHIBITOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	I	51	384	235	65	78	6	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	18	ALA	LEU	CONFLICT	UNP P68390

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total O P 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	132	Total O 132 132	0	0
4	I	49	Total O 49 49	0	0

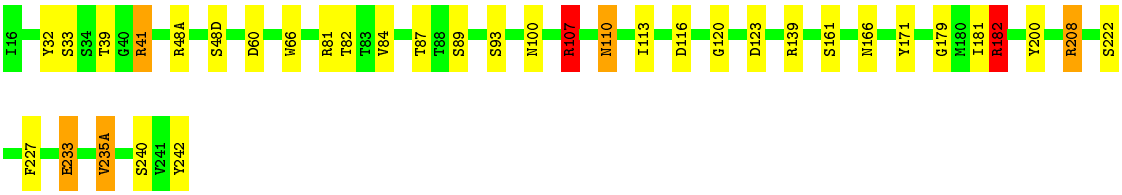
3 Residue-property plots (i)

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. 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. 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.

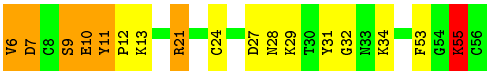
- Molecule 1: STREPTOMYCES GRISEUS PROTEINASE B

Chain E:



- Molecule 2: TURKEY OVOMUCOID INHIBITOR

Chain I:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	45.38Å 54.62Å 45.47Å 90.00° 119.20° 90.00°	Depositor
Resolution (Å)	20.00 – 1.40	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.40)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.171 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1880	wwPDB-VP
Average B, all atoms ($Å^2$)	18.0	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: PO4

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	E	1.15	1/1335 (0.1%)	1.96	39/1820 (2.1%)
2	I	1.21	1/392 (0.3%)	2.02	18/529 (3.4%)
All	All	1.17	2/1727 (0.1%)	1.97	57/2349 (2.4%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	10	GLU	CD-OE1	8.45	1.34	1.25
1	E	233	GLU	CD-OE1	8.14	1.34	1.25

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	182	ARG	NE-CZ-NH1	13.66	127.13	120.30
1	E	107	ARG	NE-CZ-NH2	-13.38	113.61	120.30
1	E	120	GLY	C-N-CA	-11.00	99.19	122.30
1	E	60	ASP	CB-CG-OD2	-10.62	108.74	118.30
1	E	41	ARG	NE-CZ-NH1	10.01	125.30	120.30
1	E	116	ASP	CB-CG-OD1	9.52	126.87	118.30
1	E	182	ARG	CD-NE-CZ	9.45	136.83	123.60
1	E	39	THR	CA-CB-CG2	-9.00	99.80	112.40
2	I	11	TYR	CB-CG-CD1	8.48	126.09	121.00
1	E	182	ARG	NE-CZ-NH2	-8.40	116.10	120.30
1	E	208	ARG	NE-CZ-NH2	8.17	124.38	120.30
2	I	7	ASP	N-CA-CB	-8.12	95.98	110.60
1	E	240	SER	CB-CA-C	-7.94	95.01	110.10
1	E	116	ASP	CB-CG-OD2	-7.81	111.27	118.30
2	I	11	TYR	CB-CG-CD2	-7.60	116.44	121.00
2	I	31	TYR	CB-CG-CD2	-7.49	116.51	121.00
1	E	242	TYR	CB-CG-CD2	-6.96	116.82	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	53	PHE	CB-CG-CD1	-6.77	116.06	120.80
1	E	222	SER	CB-CA-C	-6.77	97.24	110.10
1	E	139	ARG	NE-CZ-NH2	-6.63	116.98	120.30
1	E	81	ARG	NE-CZ-NH2	6.62	123.61	120.30
1	E	81	ARG	NE-CZ-NH1	-6.48	117.06	120.30
2	I	9	SER	CB-CA-C	-6.30	98.12	110.10
1	E	171	TYR	CG-CD2-CE2	6.27	126.31	121.30
2	I	7	ASP	CB-CG-OD1	-6.26	112.66	118.30
2	I	55	LYS	CB-CA-C	-6.13	98.15	110.40
1	E	120	GLY	O-C-N	-6.07	112.89	123.20
1	E	107	ARG	NE-CZ-NH1	6.05	123.33	120.30
2	I	27	ASP	CB-CG-OD1	6.04	123.73	118.30
1	E	100	ASN	N-CA-CB	5.96	121.33	110.60
1	E	200	TYR	CB-CG-CD2	-5.94	117.44	121.00
2	I	27	ASP	CB-CG-OD2	-5.94	112.96	118.30
1	E	171	TYR	CB-CG-CD2	5.92	124.55	121.00
2	I	11	TYR	CA-CB-CG	-5.92	102.16	113.40
1	E	93	SER	N-CA-CB	-5.89	101.67	110.50
2	I	6	VAL	CG1-CB-CG2	5.80	120.18	110.90
1	E	48(D)	SER	N-CA-CB	-5.80	101.80	110.50
1	E	82	THR	CA-CB-CG2	-5.77	104.32	112.40
1	E	48(A)	ARG	NE-CZ-NH1	5.77	123.18	120.30
2	I	31	TYR	CD1-CE1-CZ	-5.75	114.63	119.80
1	E	39	THR	N-CA-CB	-5.74	99.39	110.30
2	I	31	TYR	CG-CD2-CE2	-5.73	116.71	121.30
1	E	235(A)	VAL	CA-CB-CG2	-5.60	102.50	110.90
2	I	53	PHE	CB-CG-CD2	5.58	124.70	120.80
1	E	123	ASP	CB-CA-C	-5.56	99.28	110.40
2	I	9	SER	C-N-CA	-5.39	108.23	121.70
1	E	32	TYR	CB-CG-CD2	-5.32	117.81	121.00
1	E	89	SER	CB-CA-C	-5.26	100.10	110.10
2	I	32	GLY	CA-C-O	-5.26	111.13	120.60
1	E	235(A)	VAL	CG1-CB-CG2	-5.18	102.62	110.90
1	E	161	SER	CB-CA-C	5.16	119.91	110.10
1	E	139	ARG	NE-CZ-NH1	5.15	122.88	120.30
1	E	87	THR	OG1-CB-CG2	-5.14	98.17	110.00
1	E	242	TYR	CZ-CE2-CD2	-5.11	115.20	119.80
2	I	24	CYS	CA-CB-SG	-5.09	104.83	114.00
1	E	84	VAL	CG1-CB-CG2	-5.06	102.80	110.90
1	E	60	ASP	CB-CG-OD1	5.03	122.83	118.30

There are no chirality outliers.

There are no planarity outliers.

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	E	1310	0	1231	12	0
2	I	384	0	352	9	0
3	E	5	0	0	1	0
4	E	132	0	0	5	0
4	I	49	0	0	5	0
All	All	1880	0	1583	21	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 6.

All (21) 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:166:ASN:HD22	1:E:179:GLY:HA2	1.45	0.81
1:E:107:ARG:NH2	4:E:565:HOH:O	2.28	0.67
2:I:6:VAL:HA	4:I:105:HOH:O	1.98	0.64
2:I:28:ASN:ND2	4:I:101:HOH:O	2.26	0.62
2:I:11:TYR:HB3	2:I:12:PRO:HA	1.82	0.61
1:E:235(A):VAL:HG23	4:E:626:HOH:O	2.02	0.59
1:E:110:ASN:C	1:E:110:ASN:HD22	2.06	0.58
1:E:182:ARG:HD3	4:E:539:HOH:O	2.05	0.57
1:E:110:ASN:ND2	1:E:113:ILE:H	2.03	0.57
1:E:166:ASN:ND2	1:E:179:GLY:HA2	2.19	0.54
1:E:233:GLU:HG3	4:E:630:HOH:O	2.08	0.54
2:I:55:LYS:HB2	2:I:55:LYS:NZ	2.25	0.51
1:E:41:ARG:NE	3:E:500:PO4:O2	2.42	0.49
2:I:55:LYS:CB	2:I:55:LYS:NZ	2.79	0.46
1:E:33:SER:HB3	1:E:66:TRP:CH2	2.51	0.45
2:I:29:LYS:HE2	4:I:81:HOH:O	2.17	0.45
2:I:9:SER:N	4:I:103:HOH:O	2.51	0.43
2:I:7:ASP:N	4:I:105:HOH:O	2.46	0.43
1:E:181:ILE:O	1:E:227:PHE:HA	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:208:ARG:NH1	4:E:560:HOH:O	2.29	0.41
2:I:21:ARG:HH11	2:I:21:ARG:HB3	1.85	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	E	183/185 (99%)	174 (95%)	9 (5%)	0	100	100
2	I	49/51 (96%)	48 (98%)	1 (2%)	0	100	100
All	All	232/236 (98%)	222 (96%)	10 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	E	138/138 (100%)	135 (98%)	3 (2%)	52	19
2	I	44/44 (100%)	39 (89%)	5 (11%)	5	0
All	All	182/182 (100%)	174 (96%)	8 (4%)	28	4

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	107	ARG
1	E	110	ASN
1	E	182	ARG
2	I	10	GLU
2	I	13	LYS
2	I	21	ARG
2	I	34	LYS
2	I	55	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	110	ASN
1	E	166	ASN

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)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	PO4	E	500	-	4,4,4	1.88	2 (50%)	6,6,6	1.64	1 (16%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	500	PO4	P-O2	-2.45	1.47	1.54
3	E	500	PO4	P-O3	-2.09	1.48	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	500	PO4	O4-P-O3	-3.44	96.92	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	500	PO4	1	0

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.