



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 03:48 am GMT

PDB ID : 5ELP
Title : Ketosynthase from module 1 of the bacillaene synthase from *Bacillus amyloliquefaciens* FZB42
Authors : Wagner, D.T.; Gay, D.C.; Keatinge-Clay, A.T.; Zogzas, C.E.
Deposited on : 2015-11-04
Resolution : 2.93 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

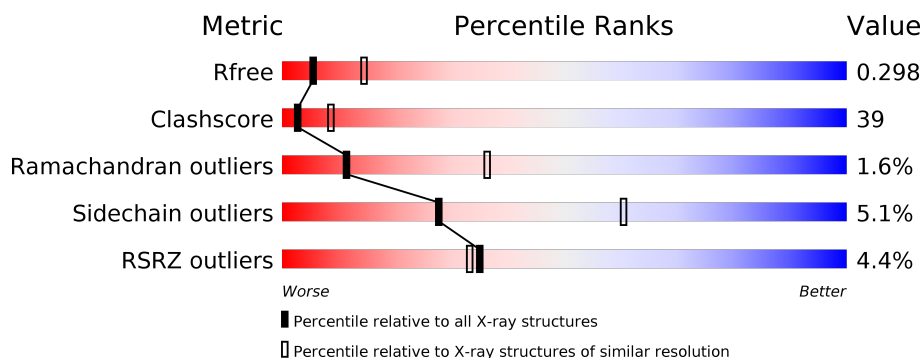
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.93 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}	100719	2289 (2.98-2.90)
Clashscore	112137	2543 (2.98-2.90)
Ramachandran outliers	110173	2475 (2.98-2.90)
Sidechain outliers	110143	2477 (2.98-2.90)
RSRZ outliers	101464	2301 (2.98-2.90)

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 $\leq 5\%$. 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	622	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 47%, green 36%, grey 13%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> % 47% 36% • • 13% </div> </div>
1	B	622	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 4%, orange 1%, yellow 43%, green 41%, grey 12%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 4% 43% 41% • 12% </div> </div>
1	C	622	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 4%, orange 1%, yellow 47%, green 35%, grey 10%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 4% 47% 35% 5% • 10% </div> </div>
1	D	622	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 6%, orange 1%, yellow 40%, green 38%, grey 16%);"></div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> 6% 40% 38% 6% • 16% </div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16749 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 NRPS/PKS protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	557	Total	C	N	O	S	0	0	0
			4317	2757	722	821	17			
1	A	540	Total	C	N	O	S	0	0	0
			4169	2656	704	792	17			
1	B	547	Total	C	N	O	S	0	0	0
			4213	2691	699	806	17			
1	D	523	Total	C	N	O	S	0	0	0
			4050	2591	677	766	16			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-16	MET	-	initiating methionine	UNP Q1RS73
C	-15	GLY	-	expression tag	UNP Q1RS73
C	-14	SER	-	expression tag	UNP Q1RS73
C	-13	SER	-	expression tag	UNP Q1RS73
C	-12	HIS	-	expression tag	UNP Q1RS73
C	-11	HIS	-	expression tag	UNP Q1RS73
C	-10	HIS	-	expression tag	UNP Q1RS73
C	-9	HIS	-	expression tag	UNP Q1RS73
C	-8	HIS	-	expression tag	UNP Q1RS73
C	-7	HIS	-	expression tag	UNP Q1RS73
C	-6	SER	-	expression tag	UNP Q1RS73
C	-5	SER	-	expression tag	UNP Q1RS73
C	-4	GLY	-	expression tag	UNP Q1RS73
C	-3	LEU	-	expression tag	UNP Q1RS73
C	-2	VAL	-	expression tag	UNP Q1RS73
C	-1	PRO	-	expression tag	UNP Q1RS73
C	0	ARG	-	expression tag	UNP Q1RS73
C	1	GLY	-	expression tag	UNP Q1RS73
C	2	SER	-	expression tag	UNP Q1RS73
C	3	SER	-	expression tag	UNP Q1RS73
A	-16	MET	-	initiating methionine	UNP Q1RS73

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	GLY	-	expression tag	UNP Q1RS73
A	-14	SER	-	expression tag	UNP Q1RS73
A	-13	SER	-	expression tag	UNP Q1RS73
A	-12	HIS	-	expression tag	UNP Q1RS73
A	-11	HIS	-	expression tag	UNP Q1RS73
A	-10	HIS	-	expression tag	UNP Q1RS73
A	-9	HIS	-	expression tag	UNP Q1RS73
A	-8	HIS	-	expression tag	UNP Q1RS73
A	-7	HIS	-	expression tag	UNP Q1RS73
A	-6	SER	-	expression tag	UNP Q1RS73
A	-5	SER	-	expression tag	UNP Q1RS73
A	-4	GLY	-	expression tag	UNP Q1RS73
A	-3	LEU	-	expression tag	UNP Q1RS73
A	-2	VAL	-	expression tag	UNP Q1RS73
A	-1	PRO	-	expression tag	UNP Q1RS73
A	0	ARG	-	expression tag	UNP Q1RS73
A	1	GLY	-	expression tag	UNP Q1RS73
A	2	SER	-	expression tag	UNP Q1RS73
A	3	SER	-	expression tag	UNP Q1RS73
B	-16	MET	-	initiating methionine	UNP Q1RS73
B	-15	GLY	-	expression tag	UNP Q1RS73
B	-14	SER	-	expression tag	UNP Q1RS73
B	-13	SER	-	expression tag	UNP Q1RS73
B	-12	HIS	-	expression tag	UNP Q1RS73
B	-11	HIS	-	expression tag	UNP Q1RS73
B	-10	HIS	-	expression tag	UNP Q1RS73
B	-9	HIS	-	expression tag	UNP Q1RS73
B	-8	HIS	-	expression tag	UNP Q1RS73
B	-7	HIS	-	expression tag	UNP Q1RS73
B	-6	SER	-	expression tag	UNP Q1RS73
B	-5	SER	-	expression tag	UNP Q1RS73
B	-4	GLY	-	expression tag	UNP Q1RS73
B	-3	LEU	-	expression tag	UNP Q1RS73
B	-2	VAL	-	expression tag	UNP Q1RS73
B	-1	PRO	-	expression tag	UNP Q1RS73
B	0	ARG	-	expression tag	UNP Q1RS73
B	1	GLY	-	expression tag	UNP Q1RS73
B	2	SER	-	expression tag	UNP Q1RS73
B	3	SER	-	expression tag	UNP Q1RS73
D	-16	MET	-	initiating methionine	UNP Q1RS73
D	-15	GLY	-	expression tag	UNP Q1RS73
D	-14	SER	-	expression tag	UNP Q1RS73

Continued on next page...

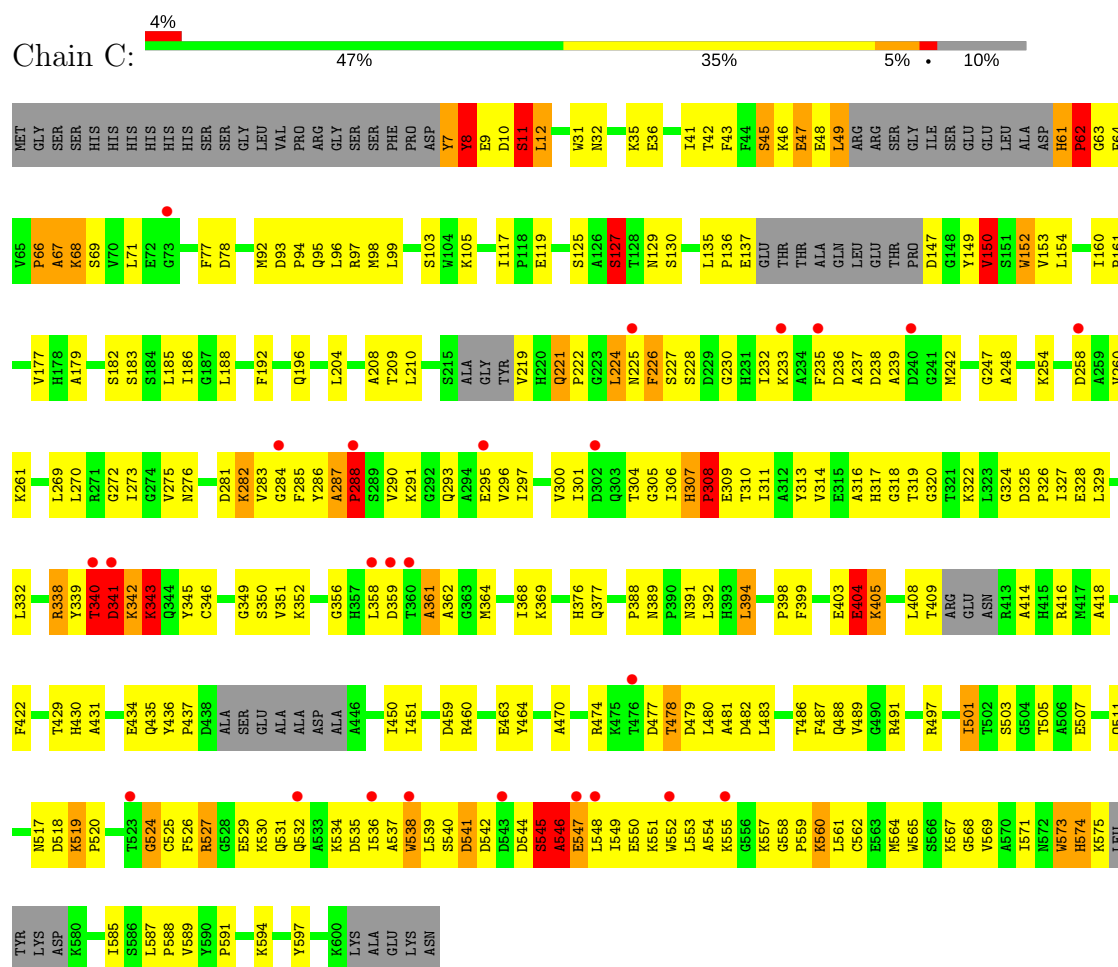
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-13	SER	-	expression tag	UNP Q1RS73
D	-12	HIS	-	expression tag	UNP Q1RS73
D	-11	HIS	-	expression tag	UNP Q1RS73
D	-10	HIS	-	expression tag	UNP Q1RS73
D	-9	HIS	-	expression tag	UNP Q1RS73
D	-8	HIS	-	expression tag	UNP Q1RS73
D	-7	HIS	-	expression tag	UNP Q1RS73
D	-6	SER	-	expression tag	UNP Q1RS73
D	-5	SER	-	expression tag	UNP Q1RS73
D	-4	GLY	-	expression tag	UNP Q1RS73
D	-3	LEU	-	expression tag	UNP Q1RS73
D	-2	VAL	-	expression tag	UNP Q1RS73
D	-1	PRO	-	expression tag	UNP Q1RS73
D	0	ARG	-	expression tag	UNP Q1RS73
D	1	GLY	-	expression tag	UNP Q1RS73
D	2	SER	-	expression tag	UNP Q1RS73
D	3	SER	-	expression tag	UNP Q1RS73

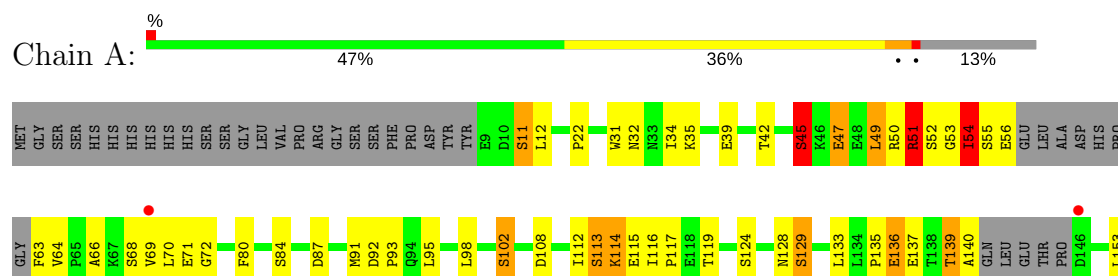
3 Residue-property plots

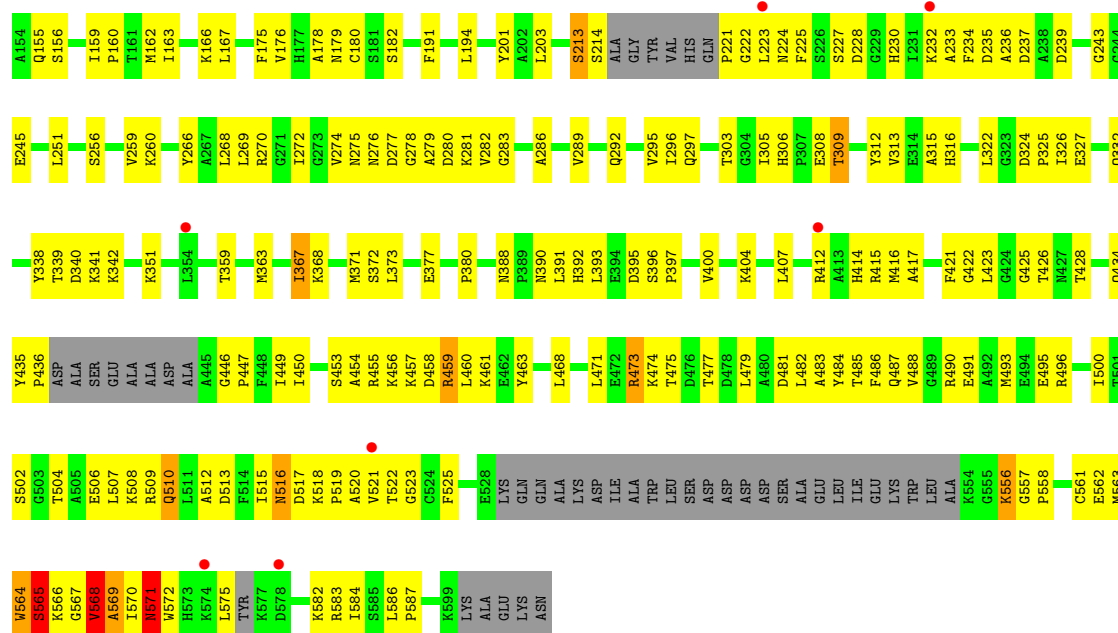
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: NRPS/PKS protein

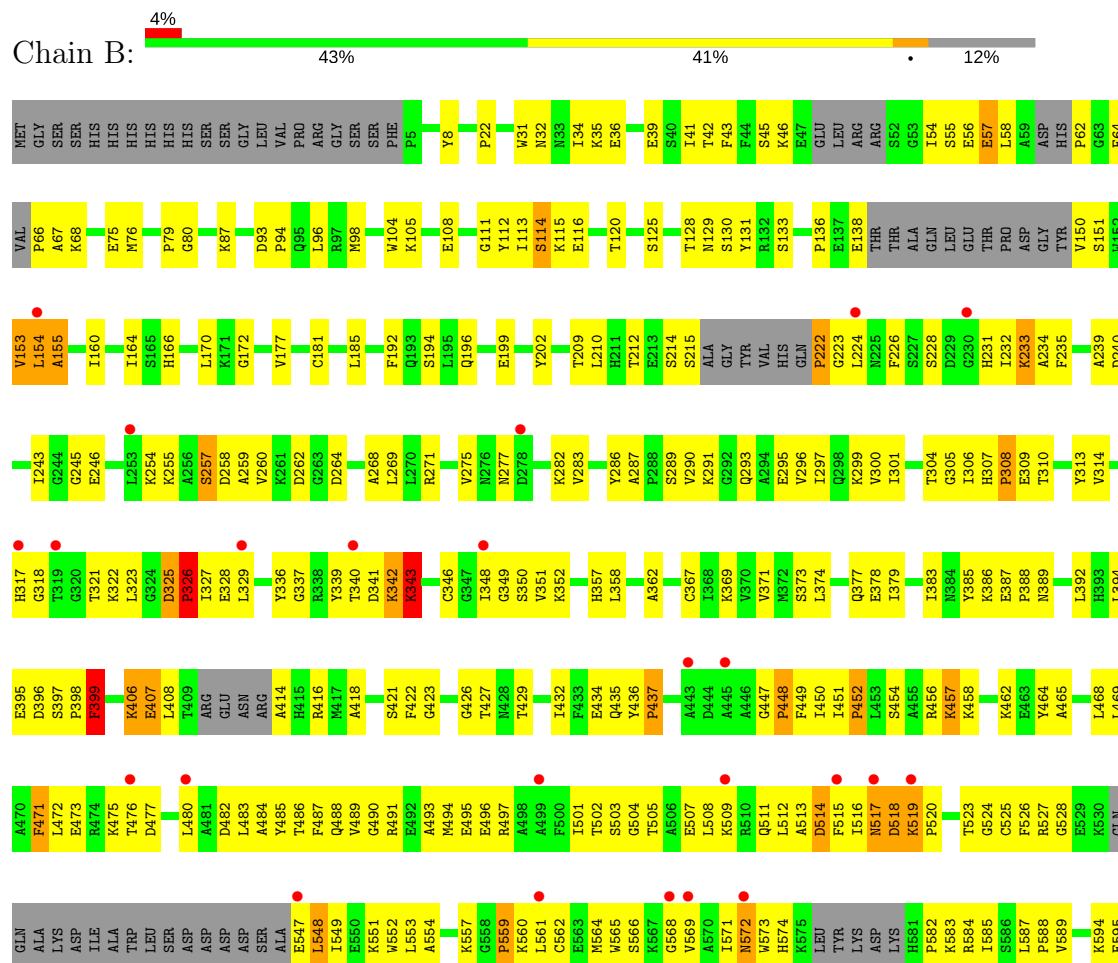


• Molecule 1: NRPS/PKS protein





• Molecule 1: NRPS/PKS protein



P596
Y597
W598
P599
X600
LVS
ALA
GLU
LVS
ASN

● Molecule 1: NRPS/PKS protein



RET	LEU	V150	K233	A316	Y385	S454	F515	H581	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
ALA	ASP	S151	A234	H317	P388	A455	I516	P582	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
HIS	HIS	L154	D236	T319	N389	K457	D518	K583	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
P62	G63	I160	A237	T321	P390	R460	LVS	I585	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
F64	V65	P161	D238	K322	N391	L461	PRO	I587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
G73	G74	M163	G245	P326	L392	K462	VAL	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
HIS	HIS	M163	G246	I327	L393	E463	GLY	I587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
SER	SER	I164	G247	T327	L394	A464	CYS	I587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
GLY	GLY	K167	G248	E328	P398	Q467	F526	P582	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
VAL	VAL	L168	A249	L329	F399	L468	G528	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
PRO	PRO	G172	L253	A331	E404	L469	LVS	I587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
ARG	ARG	H178	K254	Y334	K405	F471	GLN	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
GLY	GLY	A179	K255	Y335	K406	L472	GLN	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
SER	SER	M180	A256	Y336	E407	R474	ALA	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
PHE	PHE	C181	V260	G337	T409	K475	ASP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
P5	P5	S182	Y267	Y339	E411	D477	TRP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
D6	D6	S183	A268	Y340	ASN	T478	ALA	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
Y7	Y7	S186	L269	D341	ARG	D479	LEU	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
Y8	Y8	F192	L270	K342	H415	A481	SER	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
E9	E9	I186	R271	K343	H416	R416	ASP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
L12	L12	Q196	V275	Y345	M417	L483	ASP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
I17	I17	L195	D278	C346	A418	A484	ASP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
S18	S18	K201	K282	S350	S421	Q488	LEU	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
G19	G19	A203	G284	K352	G423	V489	ILE	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
E20	E20	L204	A287	K354	G426	R491	LVS	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
P22	P22	T120		L355		E492	TRP	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
E29	E29	T209	V290	L356	T429	A493	LEU	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
N32	N32	L210	K291	L357		E494	LEU	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
I34	I34	E213	G292	L358	E434	E495		K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
K35	K35	SER	Q293	D359	Q435	E496	K560	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
E36	E36	ALA			Y436	R497	L561	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
I41	I41	GLY	V296	A362	P437	A498	C562	K587	Y590	K594	E595	P596	Y597	W598	P599	K600	LVS	ALA	GLU	ASN	P5	D6	Y7	Y8	E9	L12	I17	S18	G19	E20	F21	P22	E29	N32	N33	I34	K35	E36	I41	T42	F43	K46	GLU	GLU	LEU	ARG	ARG	SER	GLY	ILE	SER	GLU	GLU
T42	T42																																																				

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	61.62Å 99.95Å 100.84Å 91.93° 88.18° 96.04°	Depositor
Resolution (Å)	99.35 – 2.93 53.01 – 2.93	Depositor EDS
% Data completeness (in resolution range)	94.9 (99.35-2.93) 68.8 (53.01-2.93)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.80 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, R_{free}	0.270 , 0.303 0.269 , 0.298	Depositor DCC
R_{free} test set	2426 reflections (5.24%)	DCC
Wilson B-factor (Å ²)	64.1	Xtriage
Anisotropy	0.098	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 43.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	16749	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.63	1/4264 (0.0%)	0.97	17/5749 (0.3%)
1	B	0.56	1/4311 (0.0%)	0.86	15/5813 (0.3%)
1	C	0.68	3/4421 (0.1%)	1.09	58/5968 (1.0%)
1	D	0.62	6/4148 (0.1%)	1.01	35/5595 (0.6%)
All	All	0.62	11/17144 (0.1%)	0.98	125/23125 (0.5%)

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	3
1	B	0	3
1	C	0	4
1	D	0	1
All	All	0	11

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	308	PRO	N-CD	13.48	1.66	1.47
1	A	309	THR	CB-OG1	12.87	1.69	1.43
1	C	288	PRO	N-CD	10.09	1.61	1.47
1	D	414	ALA	CA-C	6.99	1.71	1.52
1	D	414	ALA	N-CA	-6.51	1.33	1.46

The worst 5 of 125 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	309	THR	CA-CB-OG1	24.69	160.85	109.00
1	A	565	SER	CB-CA-C	16.56	141.57	110.10
1	B	57	GLU	CB-CA-C	16.45	143.30	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	45	SER	C-N-CA	16.14	162.05	121.70
1	C	477	ASP	CB-CA-C	-15.74	78.93	110.40

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	45	SER	Peptide
1	C	11	SER	Mainchain
1	C	546	ALA	Peptide
1	C	568	GLY	Peptide
1	C	8	TYR	Mainchain

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	4169	0	4098	248	1
1	B	4213	0	4124	331	6
1	C	4317	0	4213	330	4
1	D	4050	0	3958	391	3
All	All	16749	0	16393	1294	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 39.

The worst 5 of 1294 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:309:THR:OG1	1:A:309:THR:CB	1.69	1.37
1:D:308:PRO:O	1:D:345:TYR:OH	1.54	1.23
1:D:345:TYR:N	1:D:398:PRO:O	1.78	1.15
1:D:340:THR:HA	1:D:341:ASP:HB2	1.18	1.14
1:D:342:LYS:HD3	1:D:342:LYS:H	1.13	1.13

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:A:155:GLN:OE1	1:D:128:THR:OG1[1_556]	1.71	0.49
1:C:284:GLY:O	1:B:166:HIS:CD2[1_565]	1.78	0.42
1:B:477:ASP:OD2	1:D:223:GLY:O[1_655]	1.95	0.25
1:C:284:GLY:C	1:B:166:HIS:NE2[1_565]	2.03	0.17
1:C:284:GLY:O	1:B:166:HIS:CE1[1_565]	2.06	0.14

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	526/622 (85%)	483 (92%)	34 (6%)	9 (2%)	11	35
1	B	529/622 (85%)	473 (89%)	48 (9%)	8 (2%)	12	38
1	C	543/622 (87%)	485 (89%)	48 (9%)	10 (2%)	10	33
1	D	505/622 (81%)	461 (91%)	38 (8%)	6 (1%)	15	45
All	All	2103/2488 (84%)	1902 (90%)	168 (8%)	33 (2%)	11	36

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	226	PHE
1	C	288	PRO
1	C	554	ALA
1	A	447	PRO
1	A	571	ASN

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	439/506 (87%)	417 (95%)	22 (5%)	28	61
1	B	442/506 (87%)	426 (96%)	16 (4%)	40	73
1	C	453/506 (90%)	427 (94%)	26 (6%)	24	55
1	D	425/506 (84%)	399 (94%)	26 (6%)	22	52
All	All	1759/2024 (87%)	1669 (95%)	90 (5%)	28	60

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	436	PRO
1	B	114	SER
1	D	464	TYR
1	A	459	ARG
1	A	565	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	178	HIS
1	B	265	HIS
1	D	393	HIS
1	B	196	GLN
1	B	231	HIS

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 [i](#)

There are no ligands in this entry.

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](#)

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(Å ²)	Q<0.9
1	A	540/622 (86%)	0.07	9 (1%) 70 70	23, 48, 87, 109	0
1	B	547/622 (87%)	0.30	24 (4%) 35 33	24, 62, 100, 121	0
1	C	557/622 (89%)	0.32	25 (4%) 34 32	23, 55, 104, 162	0
1	D	523/622 (84%)	0.44	38 (7%) 16 13	24, 63, 108, 133	0
All	All	2167/2488 (87%)	0.28	96 (4%) 35 33	23, 57, 101, 162	0

The worst 5 of 96 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	548	LEU	8.3
1	D	503	SER	6.4
1	B	572	ASN	5.3
1	D	240	ASP	5.3
1	D	471	PHE	5.0

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.