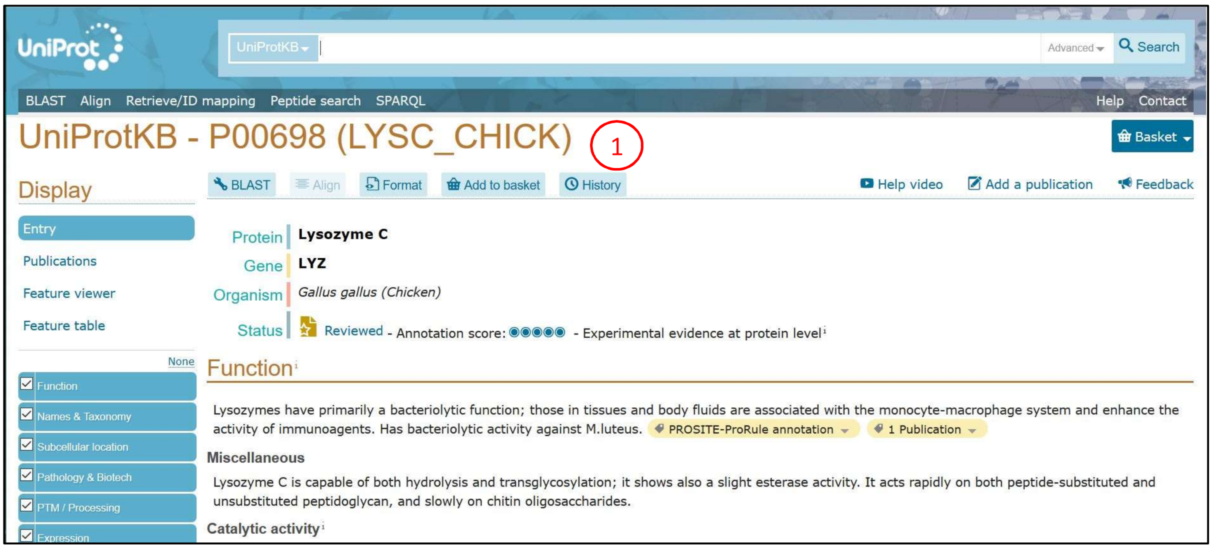
PHASING WITH THE MOLECULAR REPLACAMENT METHOD.

For phasing, crystallographic data should be used after scaling and merging, in order to have a list of unique reflections with respective intensities. To apply the Molecular Replacement method (MR), a model protein with known structure should be identified. Various resources are available to determine the best protein model, among which: (a) the UniProt online data bank [1] to obtain the primary sequence of the protein; (b) the online software Blast [2] to search for proteins with similar sequence and known structure; (c) the Protein Data Bank [3] to download the 3D structure of the selected protein; (d) the MOLREP software [4] of the crystallographic suite CCP4 [5] to solve the phase problem with MR; (e) the software Refmac [6] for the rigid body refinement and the determination of R_{work} e R_{free} values for the MR solution; (f) the Coot software [7] to visualize model and electron density obtained by Fourier transform using phases determined through MR.

Selection and preparation of the model structure.

If the primary sequence of the protein is now yet available, it can be obtained from the UniProt data bank. In the example, we search for the Hen Egg White Lysozyme (https://www.uniprot.org/uniprot/P00698) (**1**). The databank contains a lot of information, among which the sequence (**2**) and the post-translational modifications (PTM/processing, **3**).



UniProtKB - P00698 (LYSC_CHICK)

Protein	Lysozyme C
Gene	LYZ
Organism	*Gallus gallus (Chicken)*
Status	Reviewed - Annotation score: ●●●●● - Experimental evidence at protein level

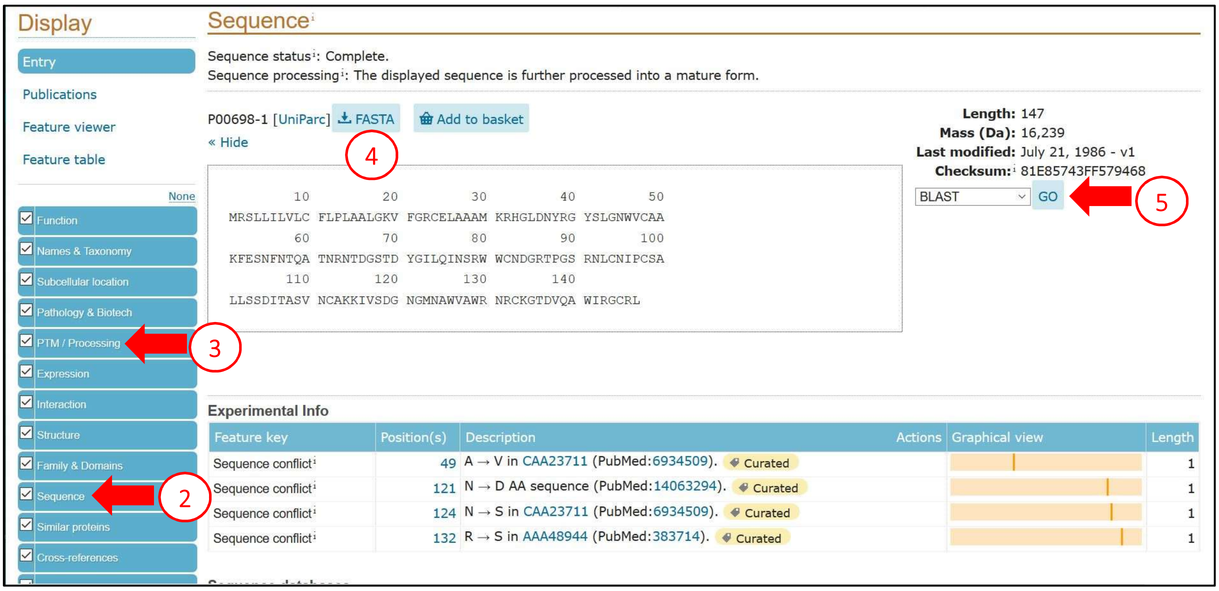
Function

Lysozymes have primarily a bacteriolytic function; those in tissues and body fluids are associated with the monocyte-macrophage system and enhance the activity of immunoagents. Has bacteriolytic activity against M.luteus. PROSITE-ProRule annotation 1 Publication

Miscellaneous

Lysozyme C is capable of both hydrolysis and transglycosylation; it shows also a slight esterase activity. It acts rapidly on both peptide-substituted and unsubstituted peptidoglycan, and slowly on chitin oligosaccharides.

Catalytic activity



Sequence

Sequence status: Complete.
Sequence processing: The displayed sequence is further processed into a mature form.

P00698-1 [UniParc]

Length: 147
Mass (Da): 16,239
Last modified: July 21, 1986 - v1
Checksum: 81E85743FF579468

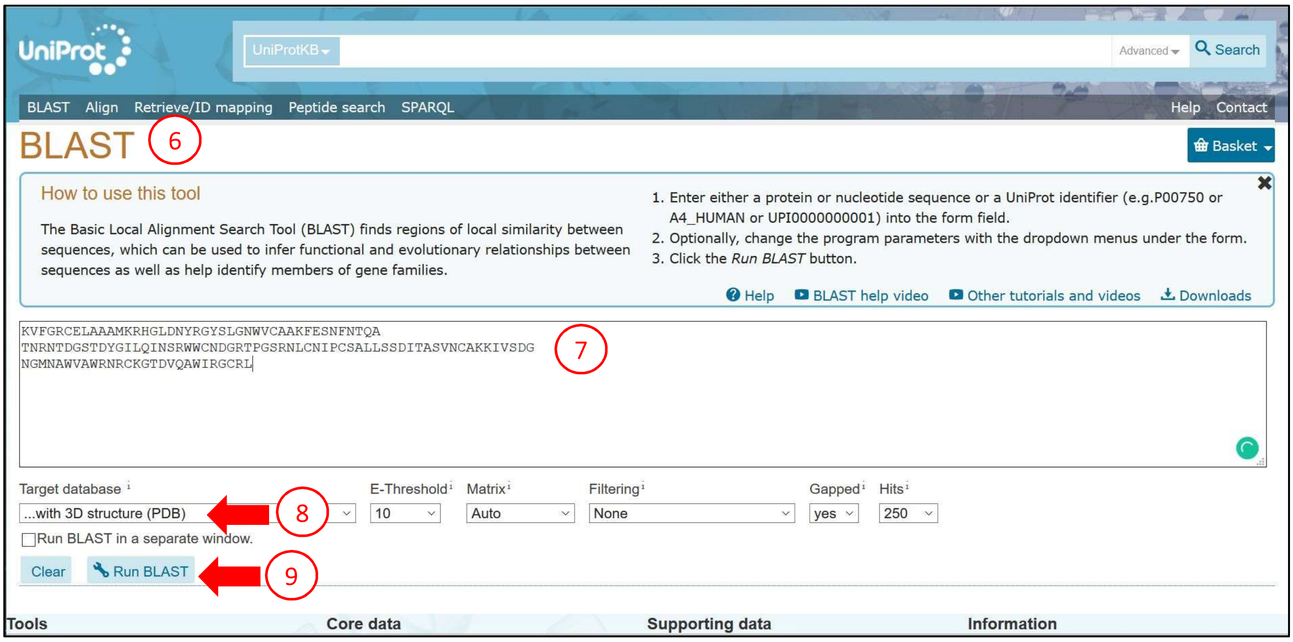
```
        10         20         30         40         50
MRSLLILVLC FLPLAALGKV FGRCELAAAM KRHGLDNYRG YSLGNWVCAA
        60         70         80         90        100
KFESNFNTQA TNRNTDGSTD YGILQINSRW WCNDGRTPGS RNLCNIPCSA
       110        120        130        140
LLSSDITASV NCAKKIVSDG NGMNAWVAWR NRCKGTDVQA WIRGCRL
```

Experimental Info

Feature key	Position(s)	Description	Actions	Graphical view	Length
Sequence conflict	49	A → V in CAA23711 (PubMed:6934509). Curated			1
Sequence conflict	121	N → D AA sequence (PubMed:14063294). Curated			1
Sequence conflict	124	N → S in CAA23711 (PubMed:6934509). Curated			1
Sequence conflict	132	R → S in AAA48944 (PubMed:383714). Curated			1

Considering the information available on the databank, the mature form of the protein lacks the first 18 residues of the sequence, which constitute the signal peptide and which are removed by proteolysis during protein maturation. The protein sequence in *Fasta* format can be obtained with the suitable button (4), it can be copied and the first 18 residues can be manually removed. From the UniProt webpage, the software Blast (5) can be opened to search for proteins with similar primary sequence.

A new Blast window opens (6) and the primary sequence of the protein in *Fasta* format can be pasted after removing the signal peptide residues (7). Among the options in the lower part of the window, it is advisable to select only the Protein Data Bank as target database (8), so that the proteins identified will have a known structure. The database search is started with the button "Run BLAST" (9).



BLAST

How to use this tool

The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences, which can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families.

1. Enter either a protein or nucleotide sequence or a UniProt identifier (e.g.P00750 or A4_HUMAN or UPI0000000001) into the form field.
2. Optionally, change the program parameters with the dropdown menus under the form.
3. Click the *Run BLAST* button.

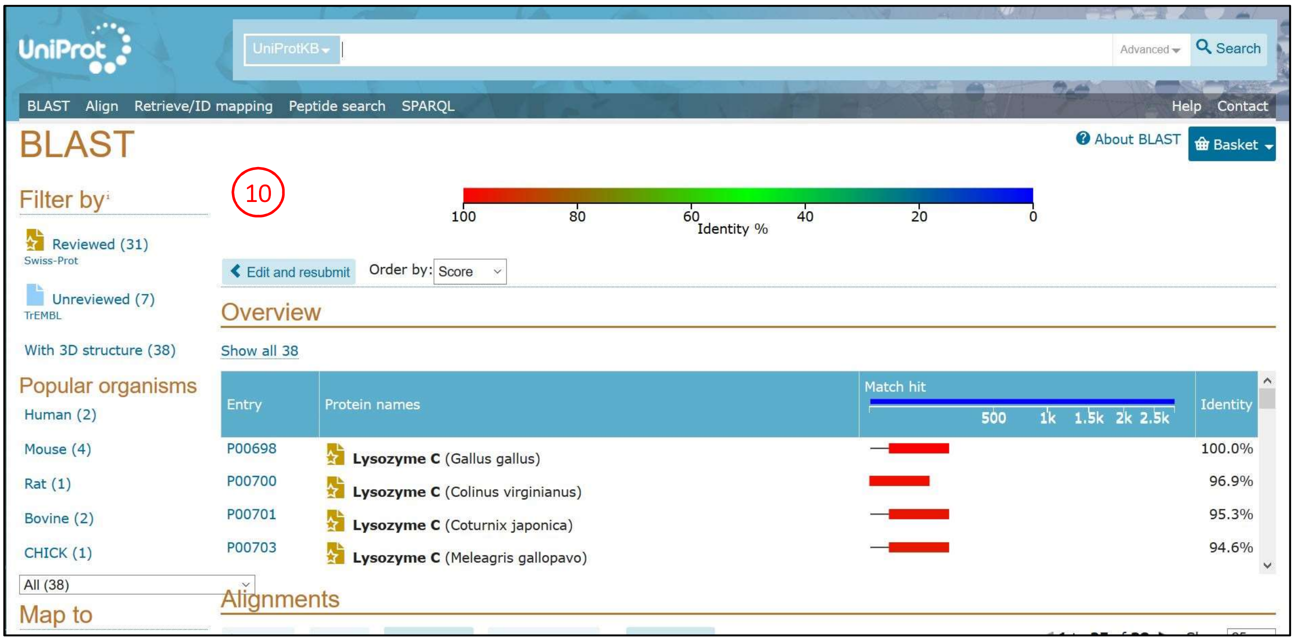
Help BLAST help video Other tutorials and videos Downloads

```
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQA
TNRNTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDG
NGMNAWVAWRNRCKGTDVQAWIRGCRL
```

Target database: ...with 3D structure (PDB)
E-Threshold: 10
Matrix: Auto
Filtering: None
Gapped: yes
Hits: 250

☐ Run BLAST in a separate window.

At the end of the search, the software shows a list of proteins with a sequence similar to the query (10), together with protein alignments (11) and identity percentage (12) between the query and the identified protein.



BLAST

Filter by

Reviewed (31) Swiss-Prot

Unreviewed (7) TrEMBL

With 3D structure (38)

Popular organisms

Human (2)

Mouse (4)

Rat (1)

Bovine (2)

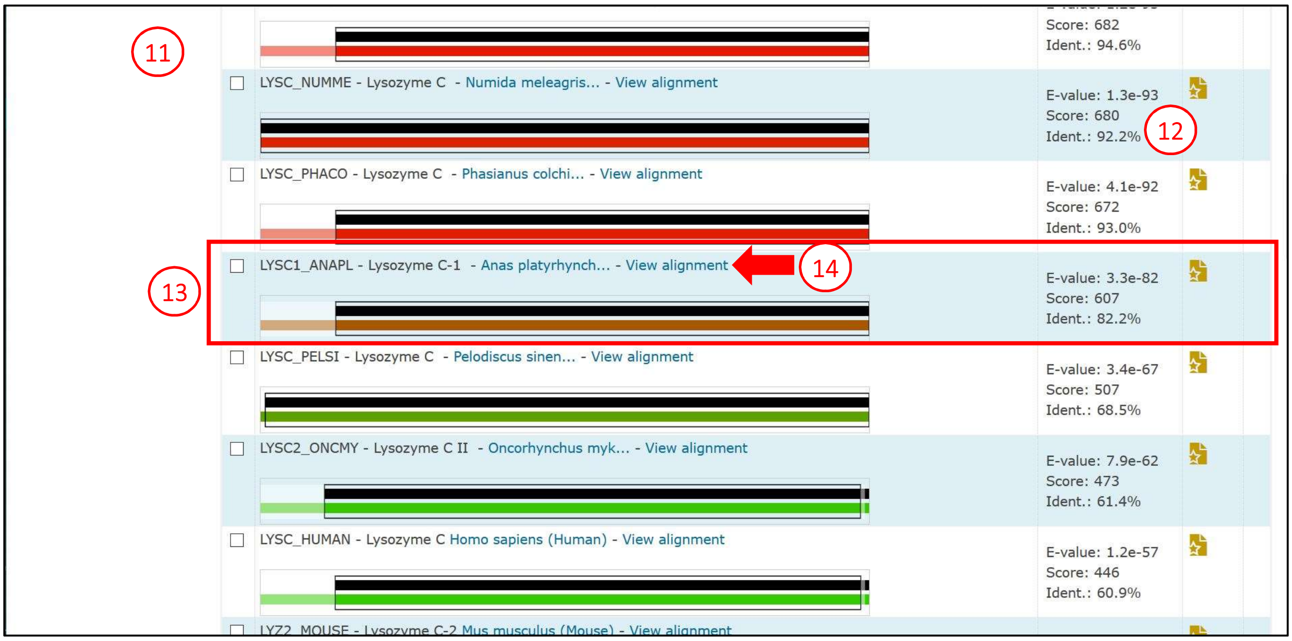
CHICK (1)

All (38)

Overview

Entry	Protein names	Match hit	Identity
P00698	**Lysozyme C** (Gallus gallus)		100.0%
P00700	**Lysozyme C** (Colinus virginianus)		96.9%
P00701	**Lysozyme C** (Coturnix japonica)		95.3%
P00703	**Lysozyme C** (Meleagris gallopavo)		94.6%

Alignments



Score: 682
Ident.: 94.6%

□ LYSC_NUMME - Lysozyme C - Numida meleagris... - View alignment
E-value: 1.3e-93
Score: 680
Ident.: 92.2%

□ LYSC_PHACO - Lysozyme C - Phasianus colchi... - View alignment
E-value: 4.1e-92
Score: 672
Ident.: 93.0%

□ LYSC1_ANAPL - Lysozyme C-1 - Anas platyrhynch... - View alignment
E-value: 3.3e-82
Score: 607
Ident.: 82.2%

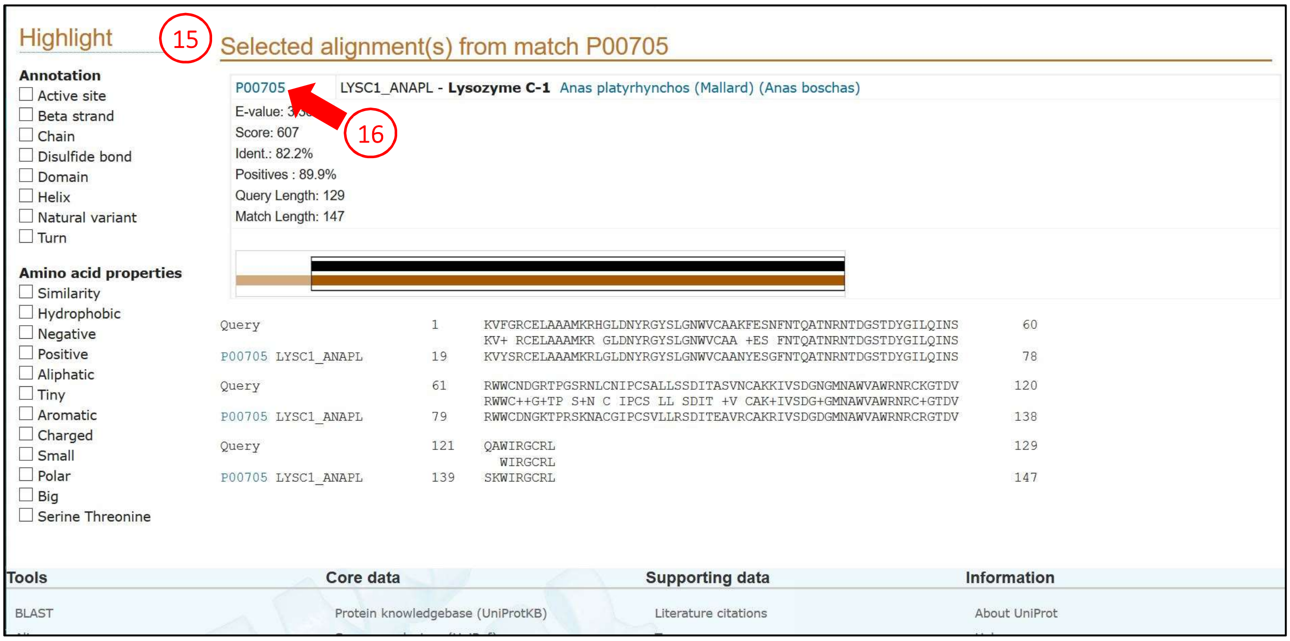
□ LYSC_PELSI - Lysozyme C - Pelodiscus sinen... - View alignment
E-value: 3.4e-67
Score: 507
Ident.: 68.5%

□ LYSC2_ONCMY - Lysozyme C II - Oncorhynchus myk... - View alignment
E-value: 7.9e-62
Score: 473
Ident.: 61.4%

□ LYSC_HUMAN - Lysozyme C Homo sapiens (Human) - View alignment
E-value: 1.2e-57
Score: 446
Ident.: 60.9%

□ LYZ2_MOUSE - Lysozyme C-2 Mus musculus (Mouse) - View alignment

Among the protein sequences, we select a protein with about 80% identity (13) (in a real case, using the sequence with the highest identity value ensures a higher success probability for the MR search). For the selected sequence, the alignment details can be analyzed in a separate browser window (14 and 15). By clicking on the identification code of the sequence (in the example P00705, 16), the user can open the UniProt page of the protein identified as probe (17, opening the link in a separate browser window).



Selected alignment(s) from match P00705

P00705 LYSC1_ANAPL - **Lysozyme C-1** Anas platyrhynchos (Mallard) (Anas boschas)

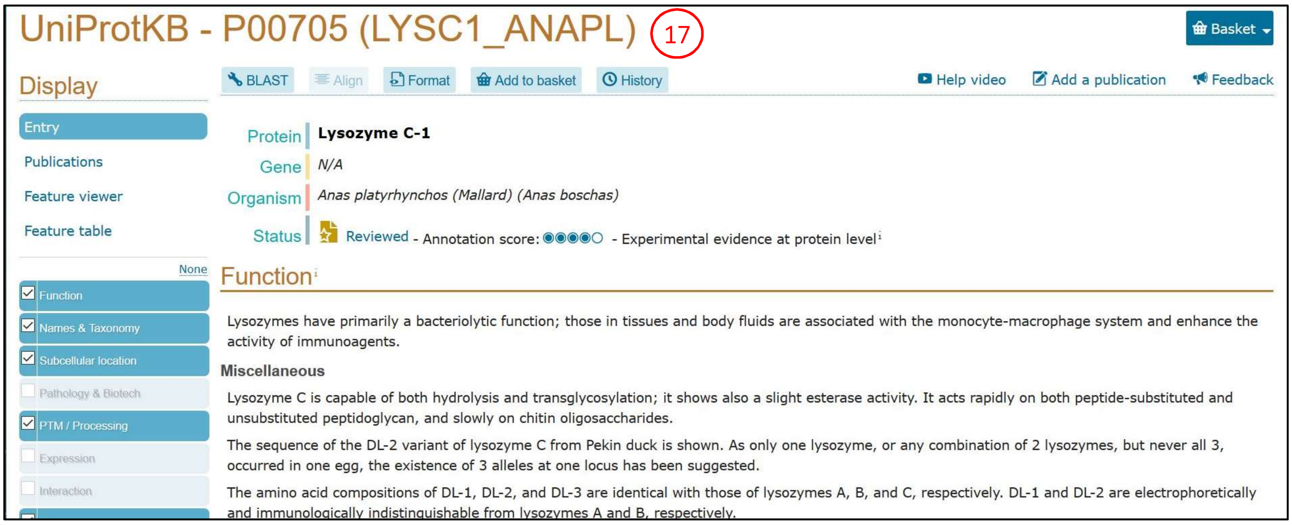
E-value: 3
Score: 607
Ident.: 82.2%
Positives : 89.9%
Query Length: 129
Match Length: 147

```
Query                 1     KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS    60
                            KV+ RCELAAAMKR GLDNYRGYSLGNWVCAA +ES FNTQATNRNTDGSTDYGILQINS
P00705 LYSC1_ANAPL   19    KVYSRCELAAAMKRLGLDNYRGYSLGNWVCAANYESGFNTQATNRNTDGSTDYGILQINS    78

Query                 61    RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDV   120
                            RWWC++G+TP S+N C IPCS LL SDIT +V CAK+IVSDG+GMNAWVAWRNRC+GTDV
P00705 LYSC1_ANAPL   79    RWWCDNGKTPRSKNACGIPCSVLLRSDITEAVRCAKRIVSDGDGMNAWVAWRNRCRGTDV   138

Query                 121   QAWIRGCRL                                                      129
                              WIRGCRL
P00705 LYSC1_ANAPL   139   SKWIRGCRL                                                      147
```

In this UniProt page, the Structure tab can be selected from the left menu (18), showing all the 3D structures corresponding to the primary sequence of the model protein, including structures determined using computational software such as AlphaFold. Among these structures, the user will choose the preferred for the MR step (for example the structure with PDB code 5V94). The second link on the right (19) opens the Protein Data Bank page of the structure (20).



UniProtKB - P00705 (LYSC1_ANAPL)

Protein	**Lysozyme C-1**
Gene	*N/A*
Organism	*Anas platyrhynchos (Mallard) (Anas boschas)*
Status	Reviewed - Annotation score: ●●●●○ - Experimental evidence at protein level

Function

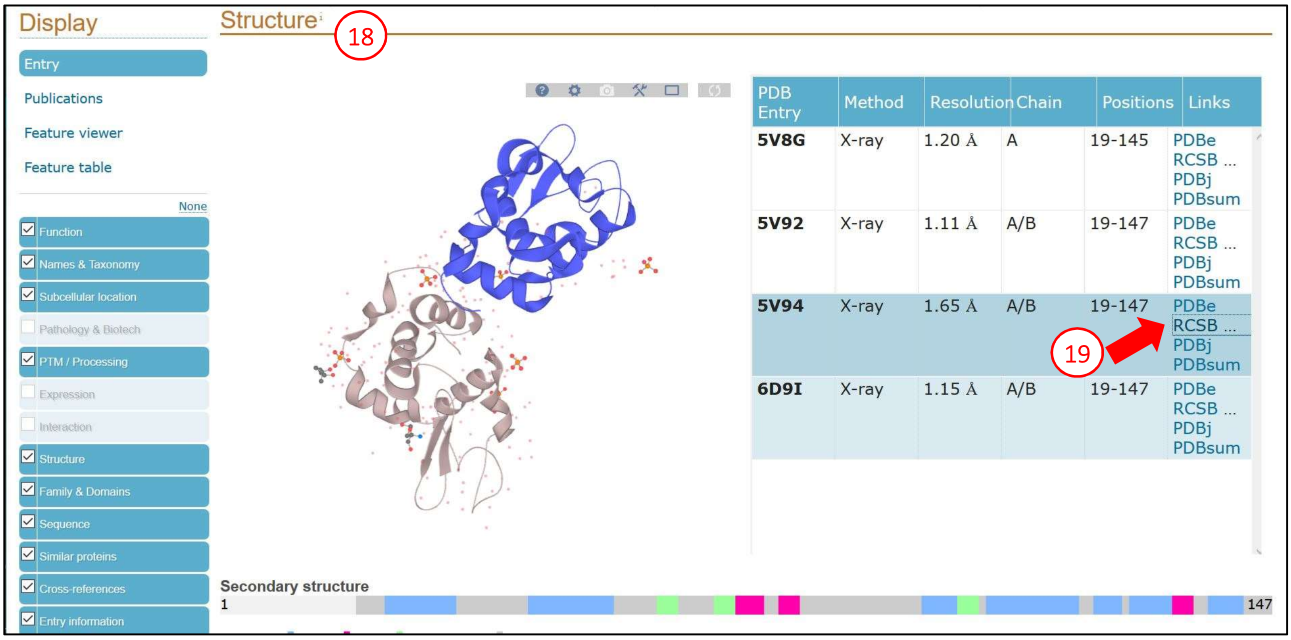
Lysozymes have primarily a bacteriolytic function; those in tissues and body fluids are associated with the monocyte-macrophage system and enhance the activity of immunoagents.

Miscellaneous

Lysozyme C is capable of both hydrolysis and transglycosylation; it shows also a slight esterase activity. It acts rapidly on both peptide-substituted and unsubstituted peptidoglycan, and slowly on chitin oligosaccharides.

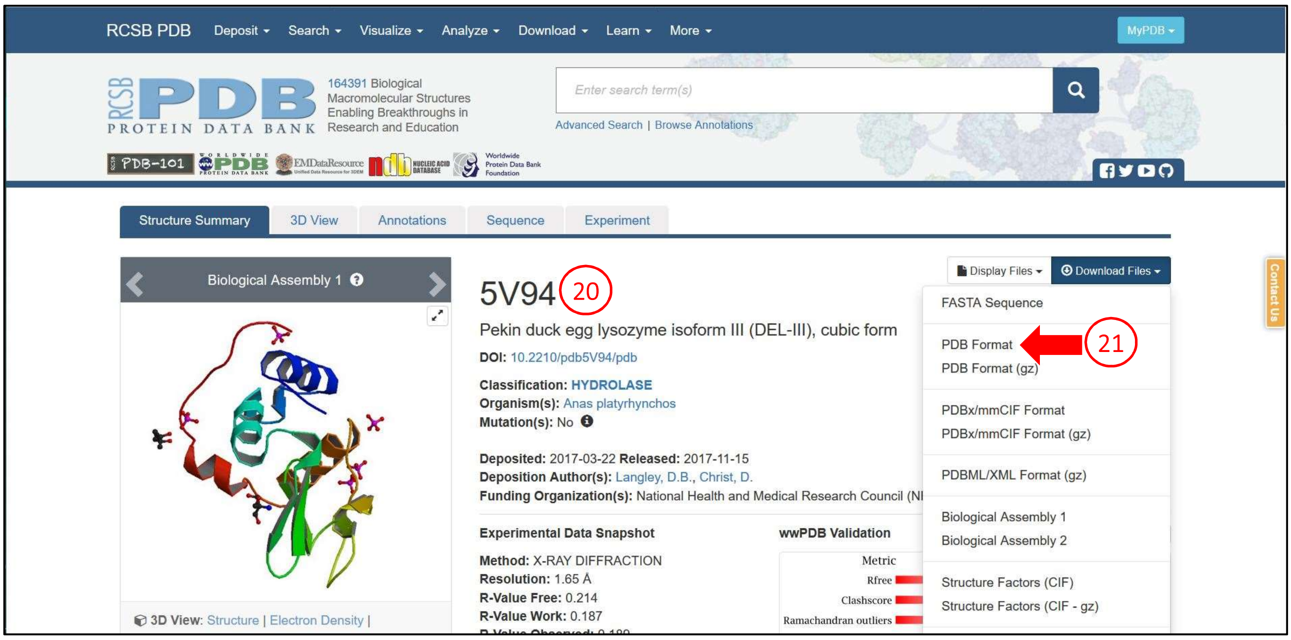
The sequence of the DL-2 variant of lysozyme C from Pekin duck is shown. As only one lysozyme, or any combination of 2 lysozymes, but never all 3, occurred in one egg, the existence of 3 alleles at one locus has been suggested.

The amino acid compositions of DL-1, DL-2, and DL-3 are identical with those of lysozymes A, B, and C, respectively. DL-1 and DL-2 are electrophoretically and immunologically indistinguishable from lysozymes A and B, respectively.



Structure

PDB Entry	Method	Resolution	Chain	Positions	Links
5V8G	X-ray	1.20 Å	A	19-145	PDBe RCSB ... PDBj PDBsum
5V92	X-ray	1.11 Å	A/B	19-147	PDBe RCSB ... PDBj PDBsum
5V94	X-ray	1.65 Å	A/B	19-147	PDBe RCSB ... PDBj PDBsum
6D9I	X-ray	1.15 Å	A/B	19-147	PDBe RCSB ... PDBj PDBsum



5V94

Pekin duck egg lysozyme isoform III (DEL-III), cubic form

DOI: 10.2210/pdb5V94/pdb

Classification: HYDROLASE
Organism(s): Anas platyrhynchos
Mutation(s): No

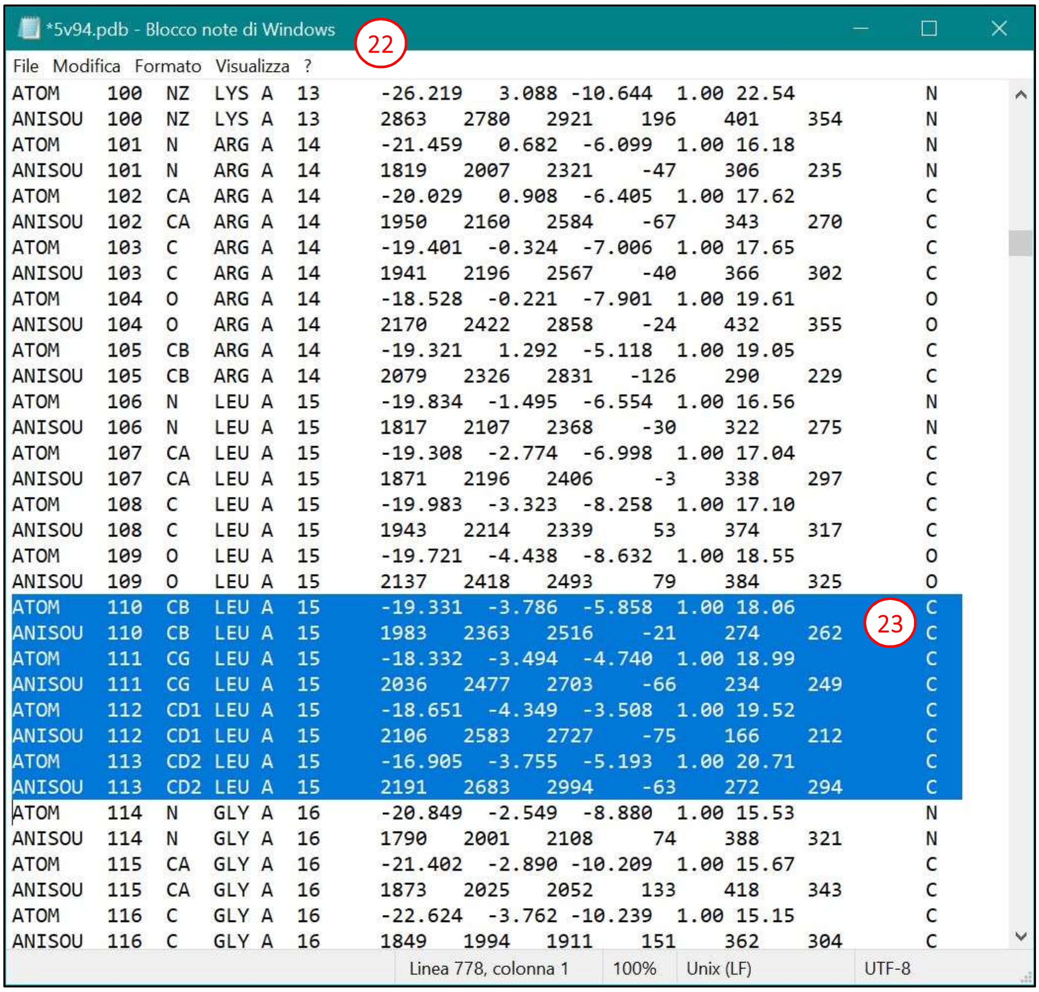
Deposited: 2017-03-22 **Released:** 2017-11-15
Deposition Author(s): Langley, D.B., Christ, D.
Funding Organization(s): National Health and Medical Research Council (NI

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 1.65 Å
R-Value Free: 0.214
R-Value Work: 0.187

wwPDB Validation

On the PDB page, the user can download the file containing atomic coordinates of the selected structure (21). Such file should be opened with a text editor (22) to manually edit the residues following the previous alignment (15). In particular, a common modification that yields good results is the removal of side chains (23) of residues that differ between the two protein sequences. In addition, only a single protein sequence should be included in the model file, while multiple chain should be removed together with water molecules, ions, ligands or other molecular species present in the 3D structure. The edited file is saved in *pdb* format. The edited structure can be visualized using a suitable software, such as PyMOL.


*5v94.pdb - Blocco note di Windows
22
File Modifica Formato Visualizza ?
23
Linea 778, colonna 1
100%
Unix (LF)
UTF-8

```
ATOM     100  NZ   LYS A  13     -26.219   3.088 -10.644  1.00 22.54           N
ANISOU   100  NZ   LYS A  13     2863   2780   2921    196    401    354       N
ATOM     101  N    ARG A  14     -21.459   0.682  -6.099  1.00 16.18           N
ANISOU   101  N    ARG A  14     1819   2007   2321    -47    306    235       N
ATOM     102  CA   ARG A  14     -20.029   0.908  -6.405  1.00 17.62           C
ANISOU   102  CA   ARG A  14     1950   2160   2584    -67    343    270       C
ATOM     103  C    ARG A  14     -19.401  -0.324  -7.006  1.00 17.65           C
ANISOU   103  C    ARG A  14     1941   2196   2567    -40    366    302       C
ATOM     104  O    ARG A  14     -18.528  -0.221  -7.901  1.00 19.61           O
ANISOU   104  O    ARG A  14     2170   2422   2858    -24    432    355       O
ATOM     105  CB   ARG A  14     -19.321   1.292  -5.118  1.00 19.05           C
ANISOU   105  CB   ARG A  14     2079   2326   2831   -126    290    229       C
ATOM     106  N    LEU A  15     -19.834  -1.495  -6.554  1.00 16.56           N
ANISOU   106  N    LEU A  15     1817   2107   2368    -30    322    275       N
ATOM     107  CA   LEU A  15     -19.308  -2.774  -6.998  1.00 17.04           C
ANISOU   107  CA   LEU A  15     1871   2196   2406     -3    338    297       C
ATOM     108  C    LEU A  15     -19.983  -3.323  -8.258  1.00 17.10           C
ANISOU   108  C    LEU A  15     1943   2214   2339     53    374    317       C
ATOM     109  O    LEU A  15     -19.721  -4.438  -8.632  1.00 18.55           O
ANISOU   109  O    LEU A  15     2137   2418   2493     79    384    325       O
ATOM     110  CB   LEU A  15     -19.331  -3.786  -5.858  1.00 18.06           C
ANISOU   110  CB   LEU A  15     1983   2363   2516    -21    274    262       C
ATOM     111  CG   LEU A  15     -18.332  -3.494  -4.740  1.00 18.99           C
ANISOU   111  CG   LEU A  15     2036   2477   2703    -66    234    249       C
ATOM     112  CD1  LEU A  15     -18.651  -4.349  -3.508  1.00 19.52           C
ANISOU   112  CD1  LEU A  15     2106   2583   2727    -75    166    212       C
ATOM     113  CD2  LEU A  15     -16.905  -3.755  -5.193  1.00 20.71           C
ANISOU   113  CD2  LEU A  15     2191   2683   2994    -63    272    294       C
ATOM     114  N    GLY A  16     -20.849  -2.549  -8.880  1.00 15.53           N
ANISOU   114  N    GLY A  16     1790   2001   2108     74    388    321       N
ATOM     115  CA   GLY A  16     -21.402  -2.890 -10.209  1.00 15.67           C
ANISOU   115  CA   GLY A  16     1873   2025   2052    133    418    343       C
ATOM     116  C    GLY A  16     -22.624  -3.762 -10.239  1.00 15.15           C
ANISOU   116  C    GLY A  16     1849   1994   1911    151    362    304       C
```

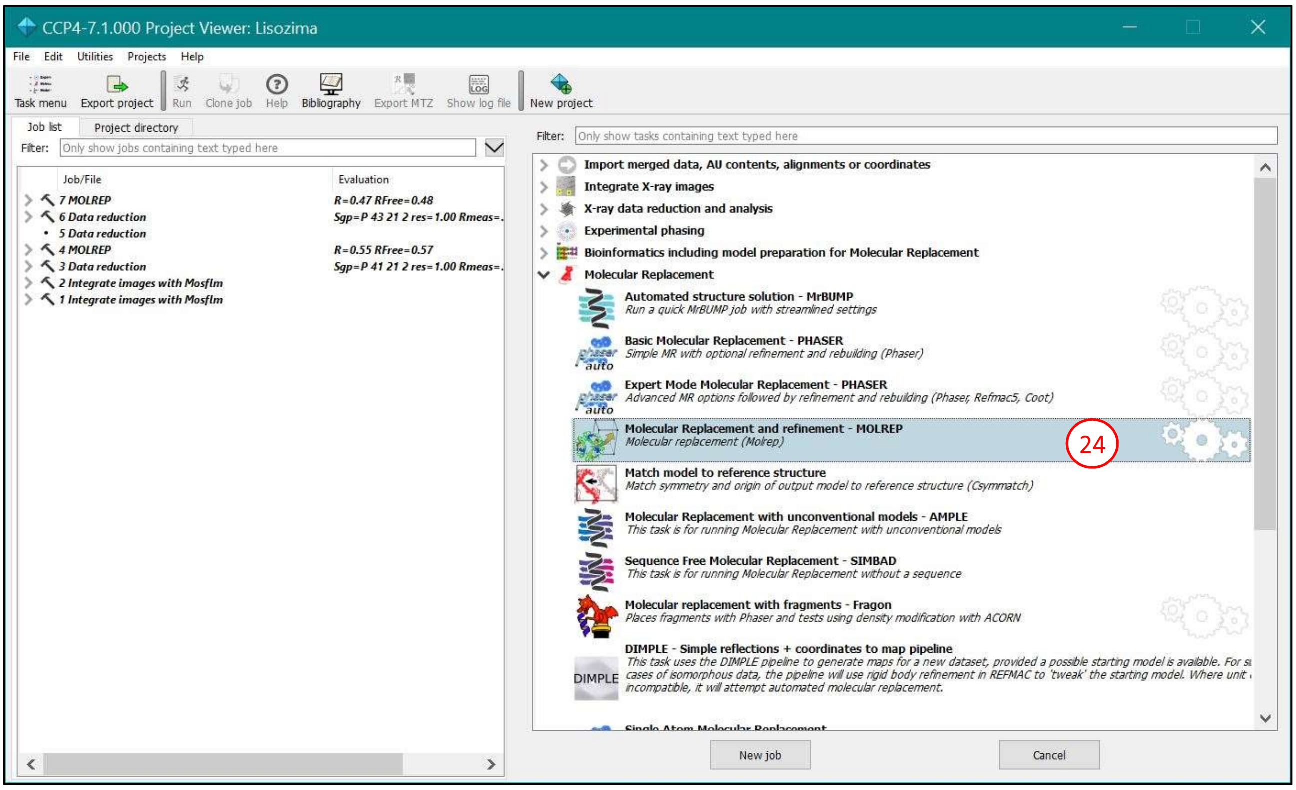
Phasing using the Molecular Replacement method with the MOLREP software.

The MOLREP software can be started from the CCP4i2 interface (24). In the window that opens after selecting the program, the input diffraction data, i.e. the *mtz* file containing scaled intensities, can be selected.

Considering the enantiomorphism of the space group, the phasing protocol should be tested for both possible solutions. The first test is conducted with intensities scaled in the *P 41 21 2* space group (25) during the previous “job 3” (26). The menu on the right reports the space group, allowing the user to check the correctness of the desired scaling procedure.

In order to perform the MR search, a second input is required, namely the model previously prepared. The edited *pdb* file is selected in the appropriate space (27).

The MR search, corresponding to a rotation matrix and a translation vector search, can be can be started with the “Run” button (28).

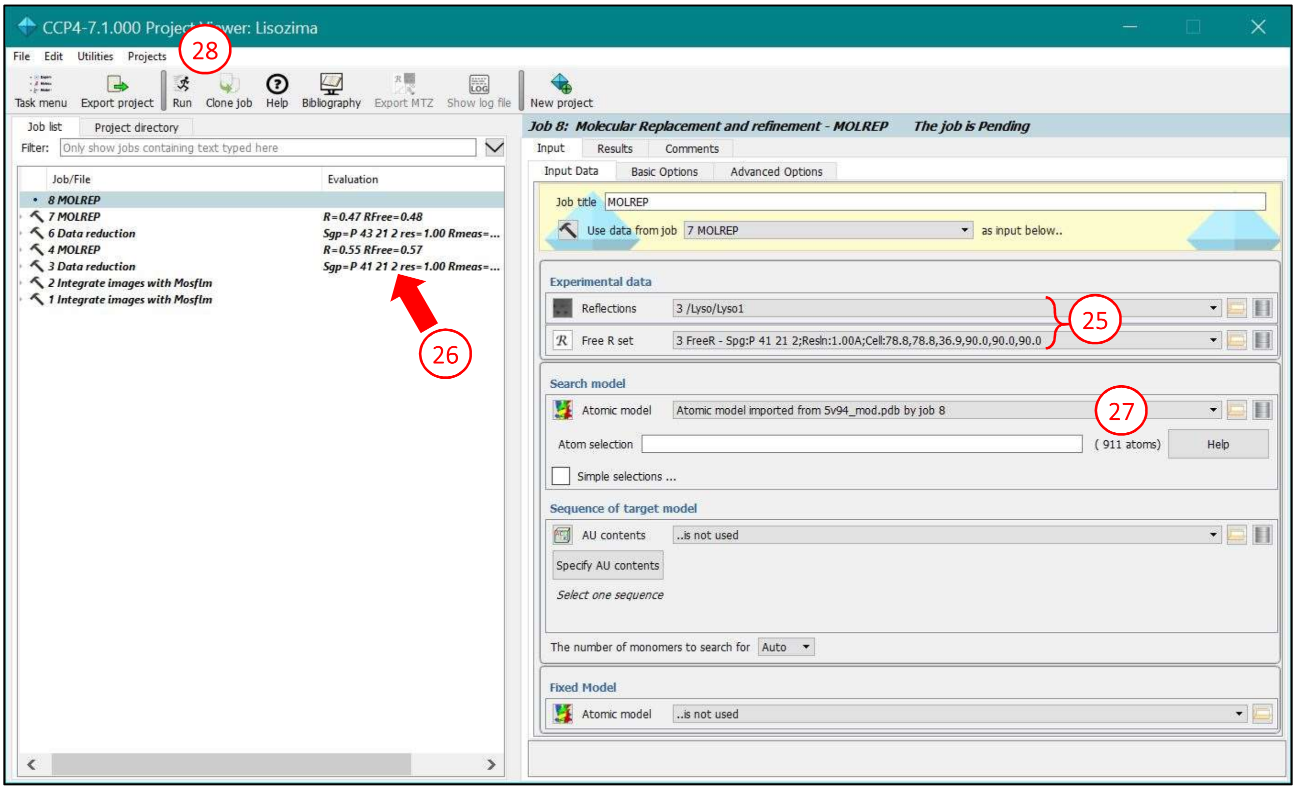


Filter: Only show jobs containing text typed here

Job/File	Evaluation
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=..
5 Data reduction	
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=..
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Filter: Only show tasks containing text typed here

- Import merged data, AU contents, alignments or coordinates
- Integrate X-ray images
- X-ray data reduction and analysis
- Experimental phasing
- Bioinformatics including model preparation for Molecular Replacement
- Molecular Replacement
  - Automated structure solution - MrBUMP
    Run a quick MrBUMP job with streamlined settings
  - Basic Molecular Replacement - PHASER
    Simple MR with optional refinement and rebuilding (Phaser)
  - Expert Mode Molecular Replacement - PHASER
    Advanced MR options followed by refinement and rebuilding (Phaser, Refmac5, Coot)
  - Molecular Replacement and refinement - MOLREP
    Molecular replacement (Molrep)
  - Match model to reference structure
    Match symmetry and origin of output model to reference structure (Csymmatch)
  - Molecular Replacement with unconventional models - AMPLE
    This task is for running Molecular Replacement with unconventional models
  - Sequence Free Molecular Replacement - SIMBAD
    This task is for running Molecular Replacement without a sequence
  - Molecular replacement with fragments - Fragon
    Places fragments with Phaser and tests using density modification with ACORN
  - DIMPLE - Simple reflections + coordinates to map pipeline
    This task uses the DIMPLE pipeline to generate maps for a new dataset, provided a possible starting model is available. For su cases of isomorphous data, the pipeline will use rigid body refinement in REFMAC to 'tweak' the starting model. Where unit incompatible, it will attempt automated molecular replacement.
  - Single Atom Molecular Replacement



Filter: Only show jobs containing text typed here

Job/File	Evaluation
8 MOLREP	
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=...
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=...
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Job 8: Molecular Replacement and refinement - MOLREP *The job is Pending*

Job title: MOLREP
Use data from job: 7 MOLREP as input below..

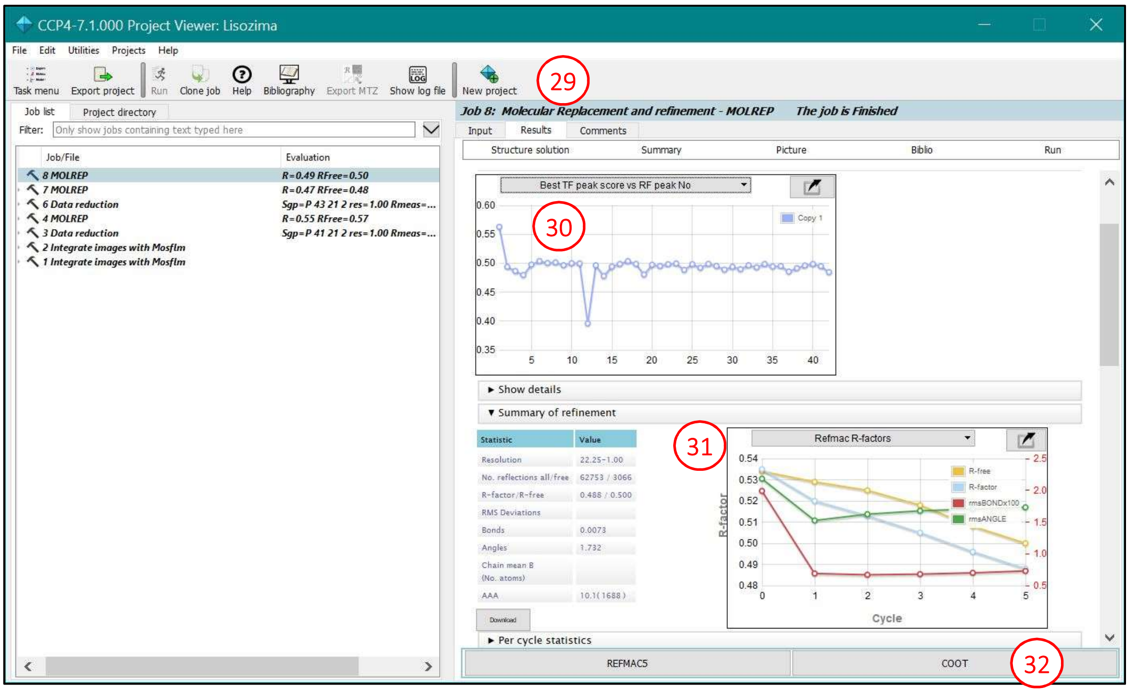
Experimental data
Reflections: 3 /Lyso/Lyso1
Free R set: 3 FreeR - Sgp:P 41 21 2;Resln:1.00A;Cell:78.8,78.8,36.9,90.0,90.0,90.0

Search model
Atomic model: Atomic model imported from 5v94_mod.pdb by job 8
Atom selection: (911 atoms)
□ Simple selections ...

Sequence of target model
AU contents: ..is not used
Select one sequence
The number of monomers to search for: Auto

Fixed Model
Atomic model: ..is not used

At the end of the calculation, the software provides the best solution identified for the model positioning in the unit cell (29). The graph on the right side of the window show the quality of the MR solution. In particular, graph 30 reports the quality of the best translational solution for each of the rotational solutions and, in the example, shows no optimal solution. The software automatically performs a rigid body refinement of the best solution (i.e., refining only the position of the whole protein structure, with no modification allowed on reciprocal positions of the atoms and residues), by recalling the Refmac software. This program yields also values for the R_{work} e R_{free} factors after refinement. The graph **31** shows the variation of these indexes in the refinement cycles. In this case, the MR solution in the *P 4₁ 2₁ 2* space group yields an R_{work} value of 0.49 and an R_{free} value of 0.50 at the end of the 5th rigid body refinement cycle. These unsatisfactory values are indicative of a possible mistake in the space group choice, but this hypothesis can be confirmed only by testing the MR solution in the other enantiomorphic space group.

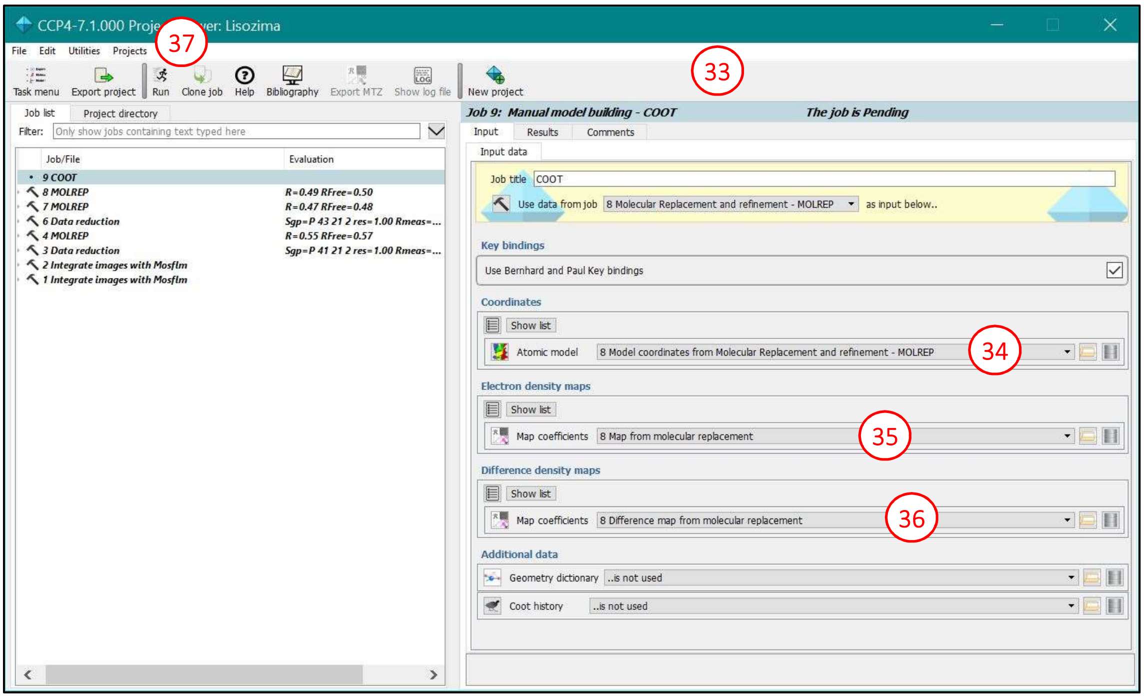


Job 8: Molecular Replacement and refinement - MOLREP *The job is Finished*

Job/File	Evaluation
8 MOLREP	R=0.49 RFree=0.50
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=...
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=...
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Statistic	Value
Resolution	22.25-1.00
No. reflections all/free	62753 / 3066
R-factor/R-free	0.488 / 0.500
RMS Deviations	
Bonds	0.0073
Angles	1.732
Chain mean B (No. atoms)	
AAA	10.1(1688)

A further indication that the solution obtained is wrong comes from the direct observation of the calculated electron density, compared with the protein model used in the MR. The Coot software, that allows for the electron density inspection, can be started from the CCP4i2 interface (**32**). In the right window (**33**), the user can select both the model protein (**34**) and the data from which electron density (**35**) and difference electron density (**36**) are obtained. The "Run" button (**37**) opens the Coot window (**38**).



Job 9: Manual model building - COOT *The job is Pending*

Job/File	Evaluation
9 COOT	
8 MOLREP	R=0.49 RFree=0.50
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=...
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=...
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Job title COOT

Use data from job 8 Molecular Replacement and refinement - MOLREP as input below..

Use Bernhard and Paul Key bindings ☑

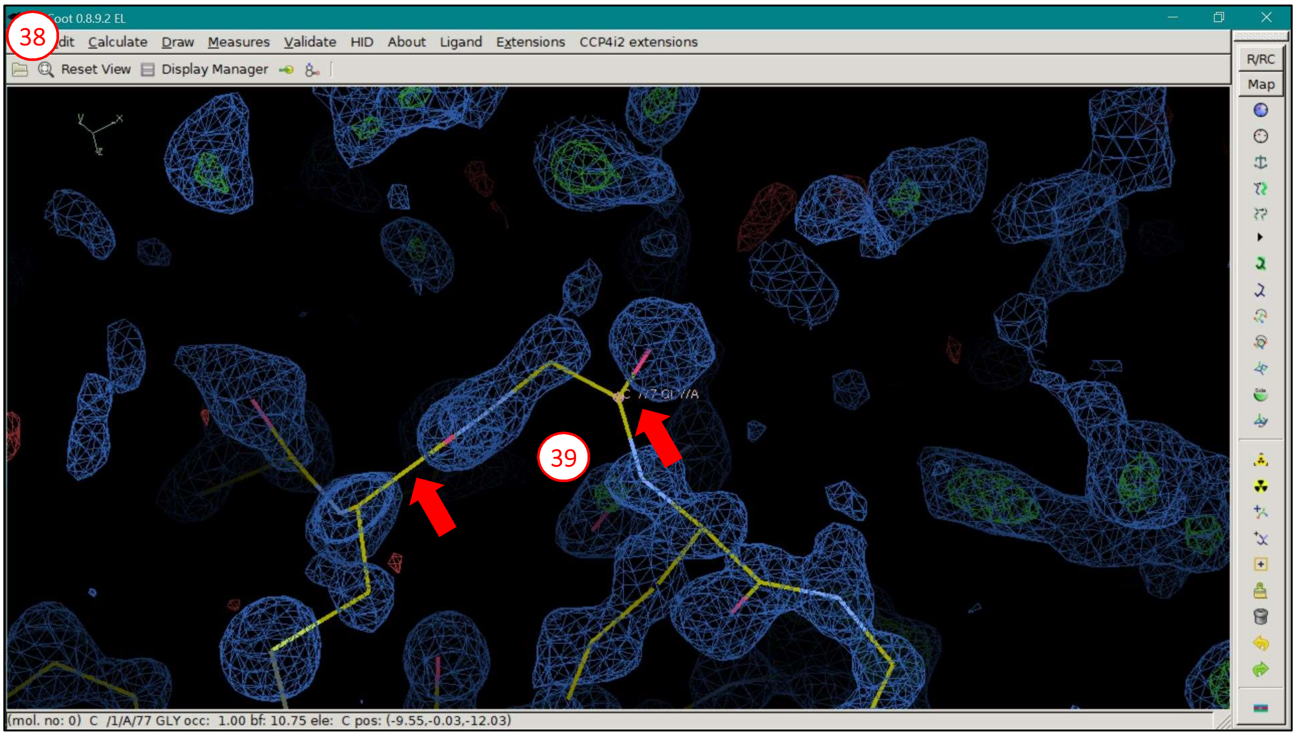
Atomic model 8 Model coordinates from Molecular Replacement and refinement - MOLREP

Map coefficients 8 Map from molecular replacement

Map coefficients 8 Difference map from molecular replacement

Geometry dictionary ..is not used

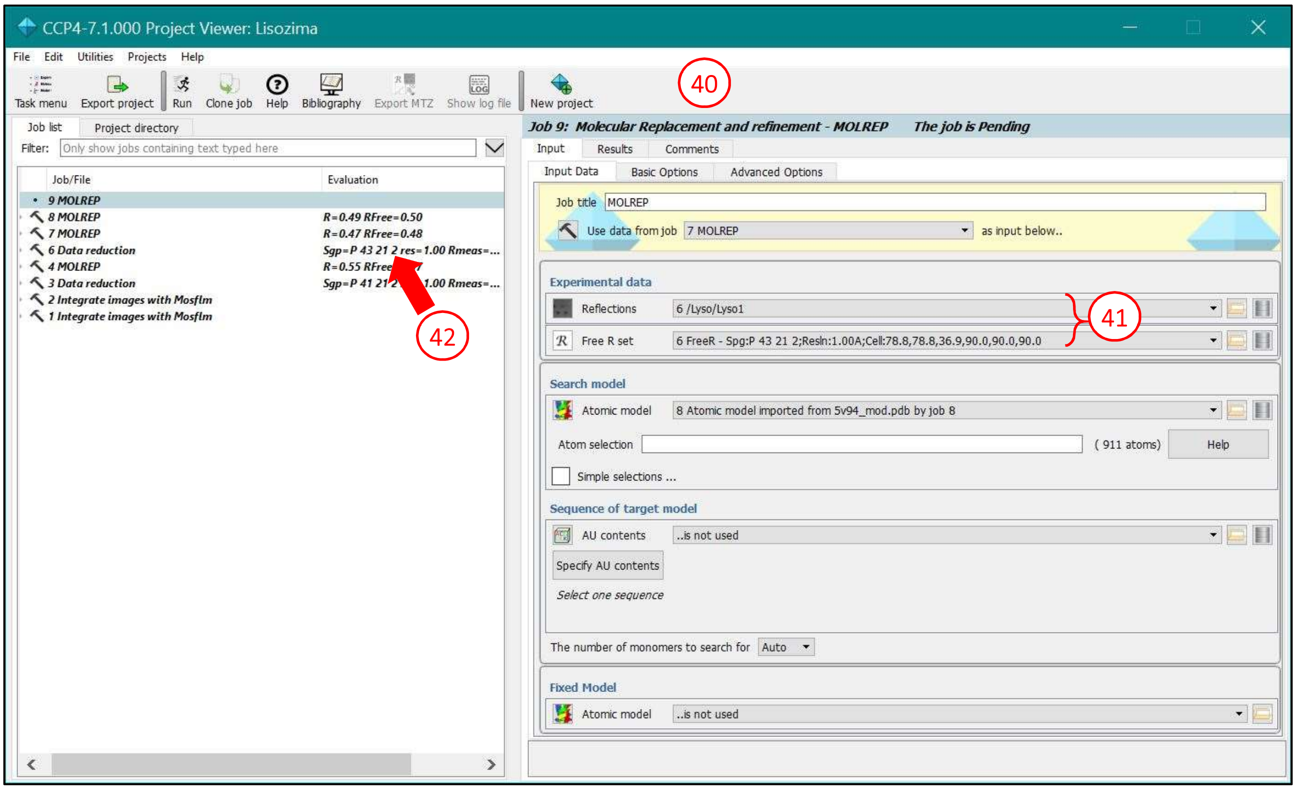
Coot history ..is not used


38
oot 0.8.9.2 EL
dit Calculate Draw Measures Validate HID About Ligand Extensions CCP4i2 extensions
Reset View Display Manager
R/RC
Map
77 GLY/A
39

(mol. no: 0) C /1/A/77 GLY occ: 1.00 bf: 10.75 ele: C pos: (-9.55,-0.03,-12.03)

Despite the fact that the *P* 4_1 2_1 *2* is the wrong space group, the user can mistakenly think that there is a similarity between the electron density and the model. This apparent similarity is due to model bias. However, a more careful inspection shows that the density is not continuous, particularly in the main chain (39).

The same steps performed for the *P* 4_1 2_1 *2* solution will be repeated using data scaled in the *P* 4_3 2_1 *2* space group (40). In this case, data selected are those obtained after scaling in "job 6" (41 e 42).



Filter: Only show jobs containing text typed here

Job/File	Evaluation
9 MOLREP	
8 MOLREP	R=0.49 RFree=0.50
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=...
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=...
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Job 9: Molecular Replacement and refinement - MOLREP *The job is Pending*

Job title: MOLREP

Use data from job 7 MOLREP as input below..

Experimental data

Reflections: 6 /Lyso/Lyso1

Free R set: 6 FreeR - Spg:P 43 21 2;Resln:1.00A;Cell:78.8,78.8,36.9,90.0,90.0,90.0

Search model

Atomic model: 8 Atomic model imported from 5v94_mod.pdb by job 8

Atom selection: (911 atoms)

□ Simple selections ...

Sequence of target model

AU contents: ..is not used

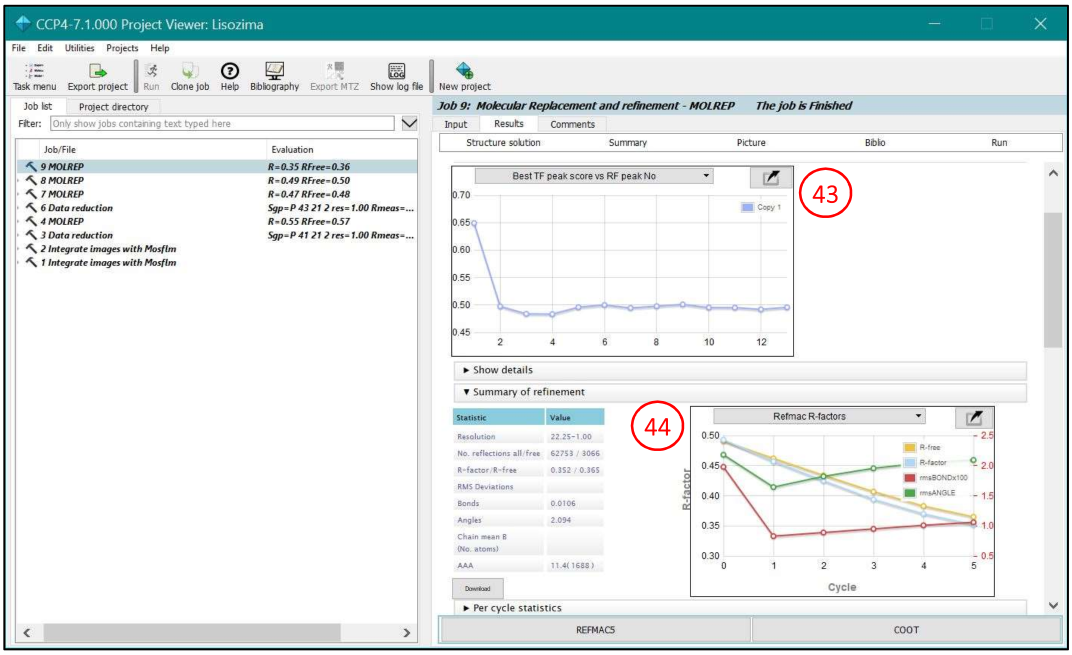
Select one sequence

The number of monomers to search for: Auto

Fixed Model

Atomic model: ..is not used

The MR solution obtained by the MOLREP software for this space group is of higher quality. The graph **43** relative to the best translational solution shows that the best solution stands out among the other, with a significant difference. The presence of a clear optimal solution is an indication that the model has been correctly positioned in the unit cell. In the graph **44**, values of R_{work} and R_{free} (0.35 and 0.36, respectively) indicate a good fitting between model and experimental data, confirming the correct space group choice.

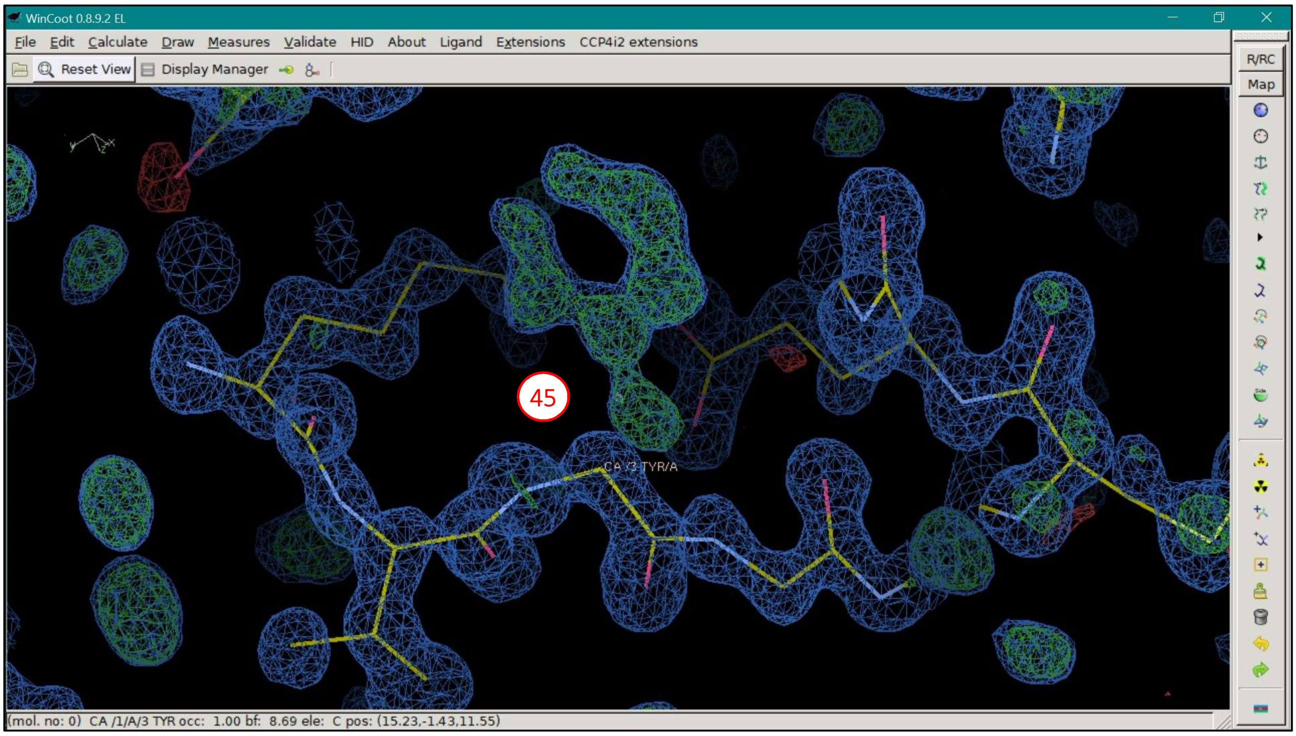


Job 9: Molecular Replacement and refinement - MOLREP *The job is Finished*

Job/File	Evaluation
9 MOLREP	R=0.35 RFree=0.36
8 MOLREP	R=0.49 RFree=0.50
7 MOLREP	R=0.47 RFree=0.48
6 Data reduction	Sgp=P 43 21 2 res=1.00 Rmeas=...
4 MOLREP	R=0.55 RFree=0.57
3 Data reduction	Sgp=P 41 21 2 res=1.00 Rmeas=...
2 Integrate images with Mosflm	
1 Integrate images with Mosflm	

Statistic	Value
Resolution	22.25-1.00
No. reflections all/free	62753 / 3066
R-factor/R-free	0.352 / 0.365
RMS Deviations	
Bonds	0.0106
Angles	2.094
Chain mean B (No. atoms)	
AAA	11.4(1688)

The analysis of the electron density with the Coot software shows a continuous electron density in the main chain. In addition, the calculated electron density predicts the mutation of some residues that differ from the model probe used in MR. For example, a tyrosine in position 3 was removed from the model, due to its mutation to phenylalanine in the analyzed protein. The electron density, **45**, shows the features of the aromatic ring, predicting the correct mutation.


WinCoot 0.8.9.2 EL
File Edit Calculate Draw Measures Validate HID About Ligand Extensions CCP4i2 extensions
Reset View
Display Manager
R/RC
Map
45
CA Y3 TYR/A

(mol. no: 0) CA /1/A/3 TYR occ: 1.00 bf: 8.69 ele: C pos: (15.23,-1.43,11.55)

References.

[1] The UniProt Consortium, *“UniProt: a worldwide hub of protein knowledge”*. Nucleic Acids Res. 2019; 47, D506-515.

[2] S. McGinnis, and T.L. Madden, *"BLAST: at the core of a powerful and diverse set of sequence analysis tools"*. Nucleic Acids Res. 2004; 32, W20-W25.

[3] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *“The Protein Data Bank”*. Nucleic Acids Res. 2000; 28, 235-242.

[4] A.Vagin and A.Teplyakov, *“MOLREP: an automated program for molecular replacement”*. J Appl Cryst. 1997; 30, 1022-1025.

[5] M. D. Winn et al., *"Overview of the CCP4 suite and current developments"*. Acta Cryst. 2011; D67, 235-242.

[6] G.N. Murshudov, A.A. Vagin, and E.J. Dodson, *"Refinement of Macromolecular Structures by the Maximum-Likelihood method"*. Acta Cryst. 1997; D53, 240-255.

[7] P. Emsley, B. Lohkamp, W.G. Scott, and K. Cowtan, *“Features and Development of Coot”*. Acta Cryst. 2010; D66(Pt 4), 486-501.