ΕΡΓΑΣΙΑ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ

Ονοματεπώνυμο:Πεΐδου Ραφαέλα
Α.Ε.Μ.:8365

Πρωτεΐνη: 3LYM(P00698)

- *Fasta sequence:*

```
>sp|P00698|LYSC_CHICK Lysozyme C OS=Gallus gallus
GN=LYZ PE=1 SV=1
MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFE
SNFNTQA
TNRNTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCA
KKIVSDG
NGMNAWVAWRNRCKGTDVQAWIRGCRL
```

- *ProtParam*

User-provided sequence:

```
        10         20         30         40
50         60
MRSLLILVLC FLPLAALGKV FGRCELAAAM KRHGLDNYRG
YSLGNWVCAA KFESNFNTQA

        70         80         90        100
110        120
TNRNTDGSTD YGILQINSRW WCNDGRTPGS RNLCNIPCSA
LLSSDITASV NCAKKIVSDG

       130        140
NGMNAWVAWR NRCKGTDVQA WIRGCRL
```

References and documentation are available.

```
Number of amino acids: 147

Molecular weight: 16238.6
```

Theoretical pI: 9.36

Amino acid composition:

```
Ala (A)  14     9.5%
Arg (R)  12     8.2%
Asn (N)  14     9.5%
Asp (D)   7     4.8%
Cys (C)   9     6.1%
Gln (Q)   3     2.0%
Glu (E)   2     1.4%
Gly (G)  13     8.8%
His (H)   1     0.7%
Ile (I)   7     4.8%
Leu (L)  15    10.2%
Lys (K)   6     4.1%
Met (M)   3     2.0%
Phe (F)   4     2.7%
Pro (P)   3     2.0%
Ser (S)  11     7.5%
Thr (T)   7     4.8%
Trp (W)   6     4.1%
Tyr (Y)   3     2.0%
Val (V)   7     4.8%
Pyl (O)   0     0.0%
Sec (U)   0     0.0%

 (B)   0   0.0%
 (Z)   0   0.0%
 (X)   0   0.0%
```

Total number of negatively charged residues (Asp + Glu): 9
Total number of positively charged residues (Arg + Lys): 18

Atomic composition:

```
Carbon      C        705
Hydrogen    H       1116
Nitrogen    N        214
Oxygen      O        204
Sulfur      S         12
```

Formula: $C_{705}H_{1116}N_{214}O_{204}S_{12}$
Total number of atoms: 2251

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 37970
Abs 0.1% (=1 g/l) 2.338, assuming all pairs of Cys residues form cystines

Ext. coefficient 37470
Abs 0.1% (=1 g/l) 2.307, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 19.86
This classifies the protein as stable.

Aliphatic index: 81.70

Grand average of hydropathicity (GRAVY): -0.150

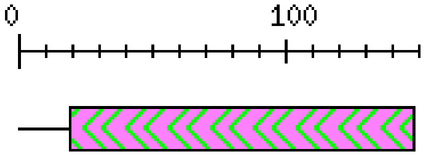
- *ProDom*

database:	multiple alignments
Program:	ncbi-blastp

Matrix:	BLOSUM62
Expect:	0.01
Filter:	seg

Graphical results and forms to other applications

The following is the graphical representation of the HSP found by BLAST. Please note that HSPs are sorted from highest to lowest scores, so that lower scoring HSPs may be hidden.



Align subsequence with ProDom domains, using Multalin

Domain ID	BEGIN	END
PD000577	19	147
PDC5E2N9	45	145
PDC1J6J2	42	134

Domain 3D modelling using Swiss-Model

Domain ID	BEGIN	END
PD000577	19	147

Domain 3D modelling using Geno3D

Domain ID	BEGIN	END
PD000577	19	147

HSP Results

Warning: Original output has been filtered to yield non-redundant similarities

```
blastp 2.2.26 [Sep-21-2011]


Reference: Altschul, Stephen F., Thomas L. Madden,
Alejandro A. Schaffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J.
Lipman (1997),
_quot;Gapped BLAST and PSI-BLAST: a new generation of
protein database search
programs_quot;,  Nucleic Acids Res. 25:3389-3402.

Query: unkwown
     (147 letters)

Database: prodom2010.1 multiple alignments
     45,292,438 sequences; 2,147,483,647 total letters



ProDom domains producing High-scoring Segment Pairs:

  Position  ProDom domain
Score E value

    19-147  #PD000577
678 3e-89
    42-134  #PDC1J6J2
180 2e-14
    45-145  #PDC5E2N9
202 1e-17


>PD000577 (Closest domain: LYSC_CHICK 19-147)
Number of domains in family: 566
Commentary (automatic):
DISULFIDE BOND SUBNAME: FULL=LYSOZYME C RECNAME: ALTNAME:
```

```
EC=3.2.1.17 GLYCOSIDASE PRECURSOR
Length = 129
Score = 678  (265.8 bits), Expect = 3e-89
Identities = 129/129 (100%), Positives = 129/129 (100%)

 Query:     19
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQ
INS 78

KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQ
INS
 Sbjct:     19
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQ
INS 78

 Query:     79
RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKG
TDV 138

RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKG
TDV
 Sbjct:     79
RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKG
TDV 138

 Query:    139 QAWIRGCRL 147
               QAWIRGCRL
 Sbjct:    139 QAWIRGCRL 147




>PDC5E2N9 (Closest domain: Q7Q6R2_ANOGA 366-501)
Number of domains in family: 2
Commentary (automatic):
SUBNAME: DISULFIDE BOND FULL=LYSOZYME FULL=AGAP005717-PA
REFERENCE C-6
Length = 136
Score = 202  (82.4 bits), Expect = 1e-17
Identities = 44/107 (41%), Positives = 59/107 (55%), Gaps
= 9/107 (8%)

 Query:     45 NWVCAAKFESNFNTQATNR-
NTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLS 103
                WVC A  ES FNT A  R N DGS D+G+ QI+  +WC+
PG+   C + C AL
 Sbjct:    366 TWVCIAYHESRFNTSAEGRLNADGSGDHGLFQISDIYWCS--
-PPGNGWACGVSCDALKD 422

 Query:    104 SDITASVNCAKKIVSD-----
```

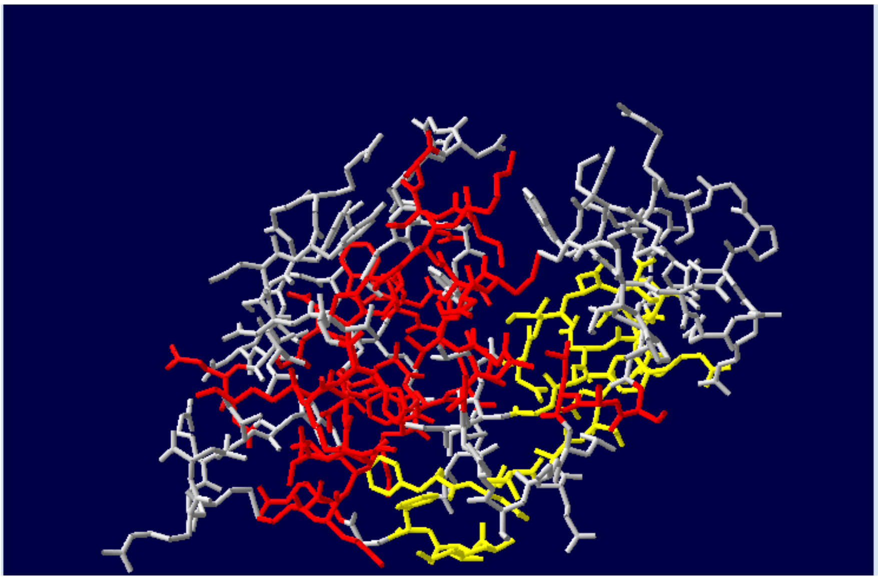
```
GNGMNAWVAWRNRCKGTDVQAWIRGC 145
               SDI+  V C K I  +     G+G NAW  ++  C+   V
++RGC
 Sbjct:    423
SDISDDVQCVKTIYEEHQRLSGDGFNAWSVYKPYCQRDAVDTFVRGC 469
```

>**PDC1J6J2** (Closest domain: B4IWP9_DROGR 949-1127)
Number of domains in family: 3
Commentary (automatic):
SUBNAME: DISULFIDE BOND FULL=GI13315 FULL=GH16162
Length = 179
Score = 180 (73.9 bits), Expect = 2e-14
Identities = 38/100 (38%), Positives = 55/100 (55%), Gaps = 10/100 (10%)

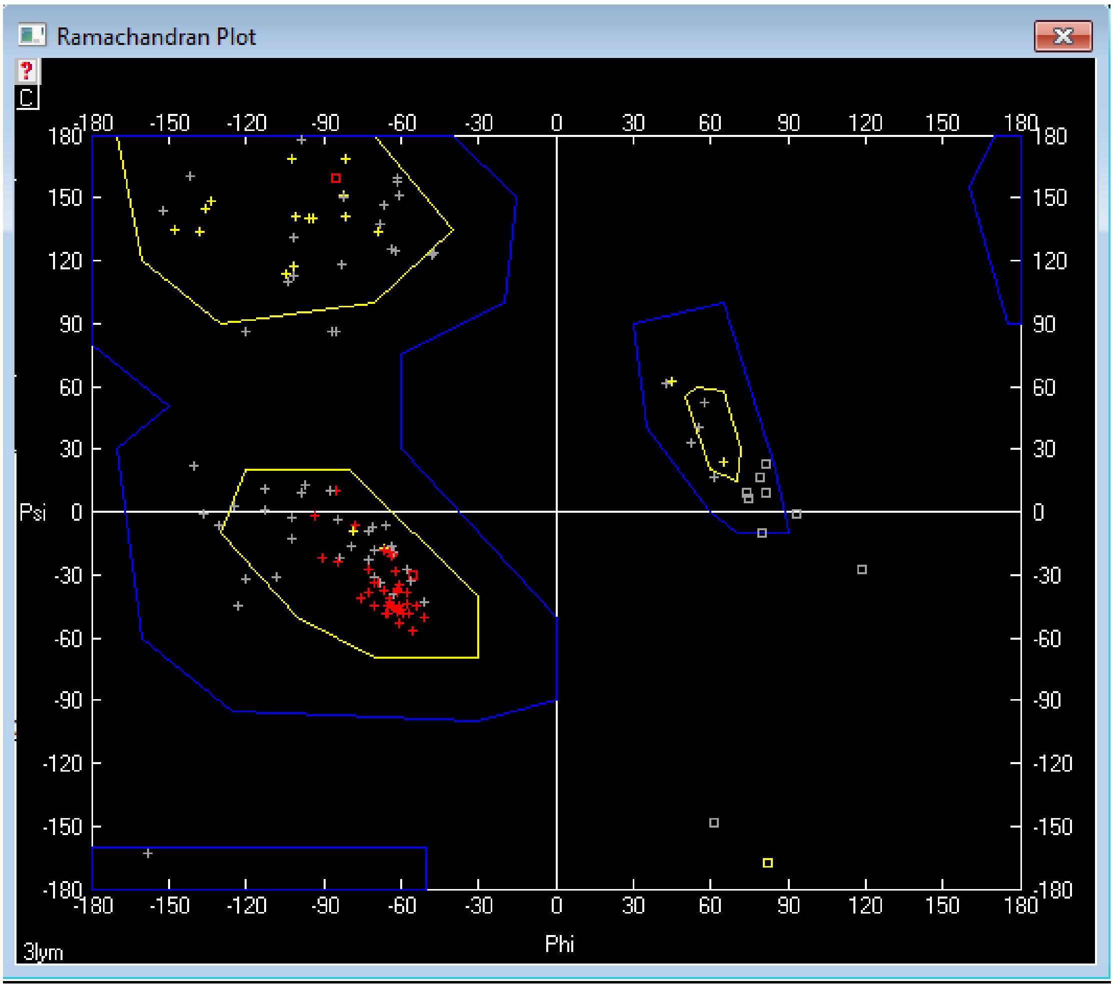
```
 Query:     42 SLGNWVCAAKFESNFNTQATNR-
NTDGSTDYGILQINSRWWC-NDGRTPGSRNLCNIPCS 99
                +  WVC A+ ES+++T A  R NTD S D+G+ QI+  +WC
+DG    S   C+I C
 Sbjct:    960
EIPTWVCIAQHESSYSTAAVGRLNTDSSEDHGLFQISDLYWCTHDG---
SSGKACHIECD 1016

 Query:    100 ALLSSDITASVNCAKKIVSD-----GNGMNAWVAWRNRCK
134
                LL SDI+  + C K I  +     G+G  AW  +   C+
 Sbjct:   1017 RLLDSDISDDIECIKTIYKEHTRISGDGFTAWTVYNGHCR
1056
```

- *Rendered by 3D/colored by secondary structure*



- *Ramachandran Plot*



Τα αμινοξέα που εμφανίζονται εκτός των επιτρεπτών στερεοτακτικών δομών είναι τα:

GLY102 GLY104 GLY 54

Επειδή η πλευρική αλυσίδα της γλυκίνης είναι μικρή (μόλις ένα άτομο υδρογόνου), η γλυκίνη μπορεί να συμμετέχει σε πολλές διαμορφώσεις μη αποδεκτές για άλλα αμινοξέα