

UNIVERSITY OF ESWATINI
FACULTY OF SCIENCE AND ENGINEERING
DEPARTMENT OF BIOLOGICAL SCIENCES
MAIN EXAMINATION PAPER 2019/2020

COURSE CODE: BIO451
TITLE OF PAPER: BIOINFORMATICS
TIME ALLOWED: **THREE (3) HOURS**
INSTRUCTIONS: NUMBERS IN BRACKETS DENOTE THE NUMBER OF MARKS
THIS PAPER COMPRISES OF **TWO SECTIONS**.
SECTION A COMPRISES OF EIGHT (8) QUESTIONS.
SECTION B COMPRISES OF TEN (10) QUESTIONS.
ANSWER ALL QUESTIONS IN SECTIONS A AND B.

TOTAL MARKS AVAILABLE: 70

NO ADDITIONAL MATERIAL (E.G. NOTES, BOOKS, PHONES, SMART WATCHES ETC) MAY BE TAKEN INTO THE EXAMINATION.

THE USE OF SCIENTIFIC CALCULATORS IS PERMITTED.

DO NOT OPEN THIS PAPER UNTIL PERMISSION HAS BEEN GRANTED BY THE CHIEF INVIGILATOR

Section A

Total marks available: 40

A patient has arrived at Mbabane Government Referral Hospital after having been bitten by a snake. Unfortunately, the patient did not bring the snake with them and no one knows what species of snake has envomated the patient. The doctor has given antivenom according to the national guidelines, but the patient is not responding. The doctor has taken a sample from the interstitial tissue around the bite site and also from the patient's blood. It has been sent to your lab for analysis. Knowing the venom is composed of a mixture of proteins, you have done all your usual protein identification assays. Unfortunately, none of them have helped you identify the snake or the proteins circulating in the patient. Finally, after separating the proteins using 2D PAGE, you have sent some of the spots for sequencing and have received the sequence data back for one of the spots.

AQ1: State what type of molecule and format of the data you expect to receive from the sequencer. [2]

AQ2 Describe what you would do with the sequence data to help identify the protein. In your answer, include what databases you would use. [4]

AQ3: The results have come back. Explain what the four columns inside the black oval (Query cover, E value, Per. Ident and Accession) tells you about your search. [4]

The screenshot shows a search results interface with tabs for Descriptions, Graphic Summary, Alignments, and Taxonomy. The main heading is "Sequences producing significant alignments". Below this, there are options for "Download", "Manage Columns", and "Show 100". A checkbox indicates "select all 99 sequences selected". The table has columns for "GenPept", "Graphics", "Distance tree of results", and "Multiple alignment". The table header row is circled in black and contains the following columns: "Description", "Max Score", "Total Query Score", "Query Cover", "E value", "Per. Ident", and "Accession".

Description	Max Score	Total Query Score	Query Cover	E value	Per. Ident	Accession
RecName: Full=Acidic phospholipase A2 2; Short=svPLA2; AltName: Full=Phosphatidylcholine 2-acylhydrolase; Flags: Precursor [Naja a...	304	304	97%	1e-104	100.00%	Q91131.1
RecName: Full=Acidic phospholipase A2 1; Short=svPLA2; AltName: Full=Muscarnic protein; Short=MP; AltName: Full=Phosphatidylpho...	298	298	97%	5e-102	97.85%	P00598.2
RecName: Full=Acidic phospholipase A2 2; Short=svPLA2; AltName: Full=CM-III; AltName: Full=NnkPLA-II; AltName: Full=Phosphatidylc...	293	293	97%	2e-100	98.58%	P00597.3
RecName: Full=Acidic phospholipase A2 C; Short=svPLA2; AltName: Full=NAJPLA-2C; Short=APLA; AltName: Full=Phosphatidylcholine...	287	287	97%	7e-98	93.84%	Q92068.1
RecName: Full=Neutral phospholipase A2 muscarinic inhibitor; Short=NPLA; Short=svPLA2; AltName: Full=NAJPLA-2A; AltName: Full=F...	287	287	97%	0e-98	93.84%	Q92084.1
RecName: Full=Neutral phospholipase A2 B; Short=svPLA2; AltName: Full=NAJPLA-2B; Short=NPLA; AltName: Full=Phosphatidylcholin...	286	286	97%	1e-97	93.15%	Q92085.1
RecName: Full=Acidic phospholipase A2 D; Short=svPLA2; AltName: Full=APLA; AltName: Full=Phosphatidylcholine 2-acylhydrolase; Flr...	286	286	97%	2e-97	93.84%	Q91900.1
RecName: Full=Acidic phospholipase A2 1; Short=svPLA2; AltName: Full=CM-II; AltName: Full=NnkPLA-I; AltName: Full=Phosphatidylch...	285	285	97%	3e-97	93.84%	P00599.2
phospholipase a2 INaaja.nai1	240	240	80%	1e-78	94.17%	CAA45372.1

AQ4: Based on the results above, what accession number would you pursue? [1]

AQ5: After further investigation, using GenBank, you come across this result (see Annex 1). Identify the following:

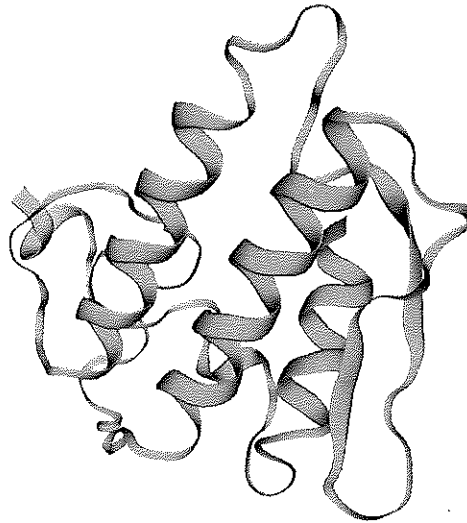
- a) The size and type of molecule [2]
- b) The full name of the GenBank Division [2]
- c) The protein name [1]
- d) Organism (common and scientific) [3]
- e) Name of the gene encoding for this protein [2]
- f) The accession number of the gene encoding for this protein [2]

Q6: Describe **how and why** you would use **bioinformatics** to determine the structure of this protein.

[5]

AQ7: Describe the key feature(s) of the protein below.

[2]



The patient was successfully treated. Your research lab has decided to further investigate the properties of this protein. Unfortunately, you do not have a lot of the protein left from the tissue sample the doctor gave and you cannot get any more snake venom from the same species. You do not have enough sample to run a couple of experiments only.

Q8: Explain how you would use **bioinformatics to inform** the biotechnology methods needed to produce more of this protein in the laboratory.

[10]

Section B

Total marks available: 30

Q1: For the following BLAST applications, state the type of molecular sequence and the type of molecular database searched:

- a) BLASTn [2]
- b) BLASTp [2]
- c) BLASTx [2]
- d) tBLASTn [2]

Q2: What is the default scoring matrix for BLASTp? [1]

Q3: Describe the difference between homologous and orthologous genes. [3]

Q4: Describe the difference between homology and similarity. [3]

Q5: What type of matrix would you use to determine the likelihood of homology between two sequences? [1]

- BQ6: Provide a brief explanation of dynamic programming algorithms and when they are most useful. [3]
- BQ7: Give two examples of dynamic programming algorithms. [2]
- BQ8: Provide a brief explanation of progressive algorithms and when they are most useful. [3]
- BQ9: Give two examples of progressive alignment algorithms [2]
- BQ10: Describe the differences between global and local alignments. [4]

END OF EXAMINATION

ANNEX 1 Sequence 1

Protein

GenPept (full)

phospholipase A2 [Naja naja]

GenBank: CAA54802.1

[Identical Proteins](#) [FASTA](#) [Graphics](#) [Item in clipboard](#)

[Go to:](#)

LOCUS CAA54802 146 aa linear VRT 27-JUN-2018

DEFINITION phospholipase A2 [Naja naja].

ACCESSION CAA54802

VERSION CAA54802.1

DBSOURCE embl accession [X77755.1](#)

KEYWORDS .

SOURCE Naja naja (Indian cobra)

ORGANISM [Naja naja](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Bifurcata; Unidentata; Episquamata; Toxicofera; Serpentes; Colubroidea; Elapidae; Elapinae; Naja.

REFERENCE 1

AUTHORS Pan, F.M., Chang, W.C. and Chiou, S.H.

TITLE cDNA and protein sequences coding for the precursor of phospholipase A2 from Taiwan cobra, *Naja naja atra*

JOURNAL Biochem. Mol. Biol. Int. 33 (1), 187-194 (1994)

PUBMED [7521702](#)

REFERENCE 2 (residues 1 to 146)

AUTHORS Chiou, S.H.H.

TITLE Direct Submission

JOURNAL Submitted (18-FEB-1994) S.H. Chiou, Inst of Biochemical Sciences, National Taiwan University & Inst of, Biochemical Chemistry, Academia Sinica, PO Box 23-106, Taipei, Taiwan 10764, TAIWAN

FEATURES

Location/Qualifiers

source 1..146
/organism="Naja naja"
/sub_species="atra"
/db_xref="taxon:35670"
/tissue_type="venom gland"

Protein 1..146
/product="phospholipase A2"

Region 28..145
/region_name="PA2c"
/note="Phospholipase A2; smart00085"
/db_xref="CDD:214568"

Site order(29,32,36,46,57,90)
/site_type="other"
/note="putative hydrophobic channel"
/db_xref="CDD:153091"

Site order(54,56,58,75)
/site_type="other"
/note="primary metal binding site"
/db_xref="CDD:153091"

Site order(56,74..75,78,94,120)
/site_type="active"
/note="catalytic network [active]"
/db_xref="CDD:153091"

CDS 1..146
/gene="PLA2"
/coded_by="X77755.1:13..453"
/db_xref="GOA:Q91133"
/db_xref="InterPro:IPR001211"
/db_xref="InterPro:IPR016090"
/db_xref="InterPro:IPR033112"
/db_xref="InterPro:IPR033113"
/db_xref="InterPro:IPR036444"
/db_xref="UniProtKB/Swiss-Prot:Q91133"

ORIGIN

1 mtpahllila avcvspigas srrpmpnlly qfknmiqctv psrswdfad ygcycgrggs
61 gtpvddldrc cqvhdyhcyne aekisgcwpy sktysyecsq gtlitckggnn acaavcdcd

121 rlaaifaga pynnnynid lkarq

//