

MCQs FOR FINAL EXAMS PREPARATION

1. A simple way (but not the best) to score an alignment is to count 1 for each match and 0 for each mismatch .
a) multiple alignment score **b) Simple Alignment Scores** c) scoring matrix
d) match model
2. Substitution matrices are the log-odds matrices used for scoring substitutions in pairwise alignments.
a) nucleotides b) DNA **c) amino acid** d) All
3. When amino Acids are dissimilar, mutations from one into the other occur less during evolution it will be
a) positive score **b) Negative Score** c) score 1 d) score 0
4. Substitution scores can be derived from,
a) Match model b) simple matrix c) probabilistic model **d) a and c**
5. The probability of two sequences is the product of probability of each amino acid, in the case of
a) Unrelated Model b) random Model c) Match model **d) a and b**
6. Dayhoff, Schwartz and Orcutt presented their famous PAM matrix in
a) 1977 b) 1976 **c) 1978** d) 1975
7. do ~~not capture~~ true difference between short time substitutions and long term ones.
a) BLOSUM matrix **b) Dayhoff matrices** c) optimal alignment d) a and b
8. Which one can perform well only in case of closely related proteins.
a) PAM matrices b) BLOSUM matrix c) simple scoring matrix d) match model
9. BLOSUM or BLOCKS database was derived in by Henikoff&Henikoff.
a) 1990 b) 1995 c) 1998 **d) 1992**
10. BLOSUM62 is good for Alignments.
a) multiple b) substitution **c) ungapped** d) gapped

11. PAM and BLOSUM are famous matrices.

- a) addition b) deletion c) multiplication d) substitution

12. Smith-Waterman algorithm finds the best alignment between two sequences.

- a) multiple b) optimal c) local d) global

13. Dynamic programming is used to find an alignment of two sequences and its scores.

14. a) multiple b) optimal c) local d) global

15. Needleman Wunsch algorithm was developed by Saul B. Needleman and Christian D. Wunsch and published in

16. a) 1966 b) 1870 c) 1975 d) 1970

17. In needle man Wunsch algorithm while doing matrix filling is the sum of the cell to the left and the gap score

- a) vertical sequence score b) horizontal match score c) horizontal gap score d) b and c

18. was awarded by Noble prize after sequencing a virus having size 5386 nucleotides.

- a) Walter Gilbert b) Steve Fodor c) Fred Sanger d) Needle man

19. Human genome project was started in

- a) 1977 b) 1989 c) 1988 d) 1975

20. After completion of genome project approximately nucleotides were sequenced.

- a) 5 billion b) 5 million c) 2 million d) 3 billion

21. How many nucleotides could be reads in starting sequencing methods?

- a) 200 – 300 b) 400 – 500 c) 500 – 600 d) 500 - 700

22. After completion of H.G.P, Shotgun sequencing was came into being after

- a) 5 years b) 20 years c) 7 years d) 10 years

23. Shot gun sequencing is used when there will be Amount of DNA sequence.

- a) Huge b) small c) medium d) All

24. In shot gun we use for the breakdown of DNA into small pieces.

- a) PCR b) hybridization c) electrophoreses d) sonication

25. For obtaining shortest super string we use a technique

- a) super string problem b) shortest super string problem

- c) shortest string problem d) shortest super problem
26. Poly nomial algorithm is for shortest super string problem.
a) best b) likely **c) unlikely** d) a and b
27. In sequencing hybridization to identify a fragment there was a small prob of nucleotides size
a) 9 – 20 b) 12 – 17 c) 7 – 15 **d) 8 – 32**
28. Which of the following is time consuming technique?
a) sequencing hybridization b) Human genome project
c) Steve fodor **d) a and b**
29. Light directed polymer synthesis was develop by
a) sanger b) Walter Gilbert **c) steve fodor** d) albert jackob
30. Steve gave his sequencing technique in
a) 1989 **b) 1991** c) 1975 d) 1985
31. 64-kb DNA array was build in 1994 by a company of California named as
a) affymatrix b) prob matrix c) symmatrix d) array matrix
32. In an array of DNA chip each prob locate at a distance with
a) unknown location **b) known location** c) unknown sequence d) b and c
33. What we use to determine the prob which is hybridized to DNA.
a) spectrophoto meter b) micro array c) spectrometer **d) spectroscopic detector**
34. Fragment assembly is actually assemblage of fragments into
a) short array b) string c) l-mers **d) entire genome**
35. How many %age of human genome consist of repeat sequence
a) 10% **b) 20%** c) 30% d) 40%
36. How many times the 300bp ALU is present in human genome?
a) 1 billion b) 2 million **c) 1 million** d) 3 billion
37. In human T cells Locus there are 5 different closely related fragment of gene.
a) tyrosine b) tryptophan c) terasin **d) tryptosine**

38. LINE and SINE present more than times in human genome.

- a) 1 million b) 2 billion c) 1 laks d) 2 laks

39. is a special design vector in which large size of fragment can be ligated.

- a) LINE b) SINE c) BAC d) ORF

40. The size of fragments that we can use in BAC is almost

- a) 150,00 bp b) 170,00 bp c) 160,00 bp d) 150,000 bp

41. We often use 30,000 bp BAC to sequence the

- a) mouse genome b) E.coli genome c) human genome d) all organisms

42. Whole genome sequencing using BAC to BAC strategy was advocated by

- a) sanger b) james weber c) gene myers d) b and c

43. Clear biotechnology company announced 1st human draft in

- a) 2005 b) 2001 c) 1999 d) 2007

44. The problem of repeated units found in a sequence was resolved by strategy

- a) BAC to BAC b) disulfide bridge c) Mate d) Dayhoff matrix

45. In Mate strategy we actually the size of read sequence.

- a) increase b) decrease c) ignore d) count down

46. Who sequence the insulin 1st time in 1940?

- a) james weber b) sanger c) gene myers d) Edman degradation

47. We can cut the terminal amino acids for sequencing by using,

- a) Edman degradation b) BAC to BAC c) Disulfide bridge d) Dayhoff matrix

48. Sanger use enzyme for the breakdown of insulin

- a) ligase b) polymerase c) lipase d) protease

49. Which of following is sulfur containing amino acid,

- a) insulin b) cystene c) lysine d) all of these

50. purified spliceosome complex,

- a) gene myers b) Edman degradation c) Mattias Man d) james weber

51. To identify apoptosis in higher organisms a commonly used technique is
a) mass spectrometer b) gel electrophoresis c) mass spectrophotometer
d) a and c
52. A programmed cell death also known as
a) Agglutination b) myosis c) necrosis d) apoptosis
53. In the Tandem mass spectrometry most frequent C terminal ion of peptide is ion and N terminal ion is ion
a) B – C b) Y – B c) B – Y d) X – Y
54. If we have access to data base for all protein from a genome then,
a) should must use MS/MS spectrum
b) MS/MS spectrum is better than data base
c) No need of MS/MS spectrum
d) a and b is right
55. To search sequence of protein from already present data base there is an algorithm called
a) FASTA b) CHOU FASMAN c) Martinez d) SEQUEST
56. SEQUEST develop by
a) gene myers b) John Yates c) james weber d) Mattias man
57. Over a million protein sequences we know only aboutprotein structures.
a) 10,000 b) 100,000 c) 10000 d) 1000
58. Mostly proteins cannot be
a) sequenced b) functioned c) crystallized d) digested
59. is used to determine the structure of protein.
a) SEQUEST b) ESI c) MALDI d) NMR
60. In nature proteins either fold spontaneously or by
a) computationally b) enzymes c) chaperons d) b and c
61. Which can determine the protein structure
a) nucleic acid b) amino acid c) enzymes d) chaperons
62. We can obtain a large protein sequence dataset from

- a) Swiss prot b) NCBI c) Uni prot d) DDBJ
63. To predict the protein structure 1st step is to predict structure,
a) Primary b) secondary c) tertiary d) quaternary
64. if we have several in the sequence, then we can anticipate that a helix may be formed by them.
a) Proline b) Thiamin c) Alanines d) Adenine
65. To predict secondary structure of amino acids 1st algorithm that was use is
a) Needle man wunch b) smith water man c) Chou-Fasman Algorithm d) all
66. SEQUEST algorithm is a approach,
a) Recursive b) branch and bound c) statistical d) Exhaustive
67. Product of the propensity values is computed for For each amino acid.
a) individual propensity b) overall propensity c) greatest propensity d) All
68. For Alpha Helices contiguous amino acids are required.
a) 6 b) 5 c) 8 d) 4
69. Alpha-Helix propensity should be more than Once this propensity falls, Alpha-Helix stops.
a) 1.0 b) 1.5 c) 2.0 d) 2.5
70. amino acids are needed to start Alpha Helices
a) 3 b) 4 c) 5 d) 6
71. Chou Fasman algorithm help us to predict
a) Alpha Helices b) Beta Sheets c) Turns d) All .
72. In Chou fasman algorithm Alpha Helices can be finalized if there is the propensity from Beta Sheets.
a) equal b) less c) higher d) different
73. Beta sheets can be predicted from amino acid sequences
a) primary b) secondary c) tertiary d) quaternary
74. How many amino acids are needed to start a Beeta Sheets
a) 4 b) 5 c) 6 d) 3
75. Chou Fasman Algorithm is based on statistical occurrence of in known structure

- a) protein **b) Amino Acids** c) nucleotide d) a and b
76. Protein sequence determined the structure, and structure determine the
- a) sequence **b) function** c) propensity d) all
77. Protein Secondary structure form due to
- a) ionic bonds b) covalent bond **c) hydrogen bonds** d) coordinat
covalent bond
78. Alpha helix form by H⁺ bonds b/w terminal
- a) C – N** b) C – C c) N – N d) b and c
79. For the prediction of protein 2ndry structure we have to see the propensity of amino acids
- a) Terminal b) core **c) neighboring** d) all
80. Chou fasman algorithm introduced by Chou Et Al in
- a) 1975 b) 1874 c) 1956 **d) 1974**
81. We use propensity table to see
- a) which amino acid found in which 2^o structure
b) amount of amino acid found in 2^o structure
c) amount of nucleic acid in DNA
d) a and b
82. Glutamine and Alanine are more likely found in
- a) alpha helix** b) beta sheets c) loops d) turns
83. An Alpha helix will be start in which of the following condition
- a) $p(a) < 1$ b) $p(a) > 0.1$ **c) $p(a) > 1.0$** d) $p(a) < 0$
84. are very flexible in nature.
- a) loops b) a helix c) beta sheets **d) turns**
85. If we have 4 amino acids in a sequence having specific propensity then we can construct
- a) polypeptide b) tri peptide **c) tetra peptide** d) di peptide
86. F(total) means product of
- a) number amino acids b) sequences of amino acids c) number of protein

d) propensity of amino acids

87. Which condition is not including for forming a turn,

a) $f(\text{total}) < 0.000075$

b) average value for $p(\text{turn}) > 1.0$ in tetrapeptide

c) average value for tetra peptide as, $p(a - \text{helix}) < p(\text{turn}) > p(b - \text{sheets})$

d) $f(\text{total}) > 0.000075$

88. Chou fasman algorithm is based on

a) computational analysis of amino acids

b) simple sequence of amino acids

c) statistical occurrence of amino acids

d) biological role of amino acids

89. For the formation of B – sheets, 1st we need to scan the

a) particular region

b) entire sequence

c) specific amino acids

d) All

90. For scanning a sequence we have to start it from

a) center

b) most right side

c) left side

d) from end

91. If 4 out of 6 amino acids have propensity > 0.1 we call it as

a) loop

b) turn

c) B – sheets

d) a – helix

92. If 4 contiguous amino acids have $p < 0.1$ and $> \text{also} > p(a)$ it must be a..

a) turn

b) B – sheets

c) loop

d) a – helix

93. If 4 out of 6 amino acids have propensity for a – helix then

a) we have no need to check it for b – sheets

b) we must check it for b – sheet

c) we should simply extend it till the end

d) a and c are right

94. According to Chou fasman, If a sequence does not have propensity for a – helix or B – sheets then it will be computed for

a) B – turn

b) B – loop

c) bulges

d) B – sheets

95. For the improvement of Chou Fasman algorithm we must be improve its statistical numbers that we obtain from

- a) NMR b) experiments c) DDBJ d) PDB

96. For which we cannot rely on 4 amino acids bcs it has variable sizes

- a) a – helix b) B – turn c) loops d) B – sheets

97. For imrovment of Chou Fasman algorithm according to latest protein data set we must be consider

- a) Nucleation regions b) Membrane proteins c) Hydrophobic domains d) ALL

98. Functional evaluation of proteins can be performed by.....

- a) Structure visualization b) , classification c) prediction d) All

99. X-Ray Crystallography and NMR Spectroscopy are used to find the of protein.

- a) sequence b) alignment c) structure d) function

100. it is then possible to identify unknown protein structures by just examining the protein sequences

- a) heterogeneous b) homologous c) multiple d) binary

101. In homology modelling, proteins with similar are considered.

- a) 1' sequences b) 2' sequences c) 3' sequences d) 4' sequences

102. Homology modelling is used to predict structures of proteins having high sequence similarity with other proteins with!

- a) known structures b) unknown sequence c) similar sequence d) a and c

103. What is true about target protein from following

- a) protein with known sequence and structure
b) protein with unknown structure and unknown sequence
c) protein with known sequence but unknown structure
d) protein with unknown sequence and known structure

104. By visualization of a structure of protein we actually see

- a) the sequence of amino acids
b) the statistical aspect of protein
c) computational value of protein
d) biological role of protein

105. Which is not true about experimental determination of overall protein structure, that it is,

- a) difficult **b) cheap** c) expensive d) time consuming
106. Which is use to determine the amino acids sequence in protein?
a) mass spectrometry b) NMR c) Admen degradation **d) a and c**
107. For structural prediction of protein we can use
a) Admen degradation b) NMR c) X – Crystallography **d) b and c**
108. Which of the following statement is not true about Twilight Zone?
a) it can be predicted by Threading
b) it consist of lower sequence identity
c) we cannot use Homology modeling in this case
d) it require high sequence alignment
109. is 1st requirement to visualization and classification of protein according to its function.
a) sequencing b) alignment **c) structural prediction** d) folding
110. Protein classification can be done by different hierarchies like
a) SCOP b) MALDI c) CATH **d) a and c**
111. is the function unit of protein
a) Alpha helix b) beta sheets **c) Domain** d) Motifs
112. is a software for homology modelling.
a) Homolog **b) Modeller** c) iTASSER d) a and b
113. There are stable folds of protein in nature.
a) 30,000 b) 4,000 c) 40,000 **d) 5,000**
114. Several tools are available to perform homology modelling in a programmatic or automated way such as....
a) Swiss Model b) Robetta c) 3D Jigsaw **d) all**
115. Fold recognition is also called
a) scoring sequencing **c) Threading** d) modeling
116. Homology modelling fails to predict quality structures is a technique for predicting these structures.
a) Fold recognition b) Swiss Model c) Threading **d) a and c**
117. Threading Scoring typically involves using a function based on energy of a molecule
a) Z-Score b) C-score c) A-score d) Y-score
118. cannot be predicted using threading.

- a) invtro protein b) primary protein c) 2ndary protein d) Novel proteins
119. Is online tool for threading
- a) Homolog b) Modeller c) iTASSER d) CLUSTLE
120. In a structure less energy present in amino acids make it more
- a) unstable b) stable c) ionized d) a and c
121. What is the basic principle of fold recognition
- a) align and compare sequence on the base of energy
- b) align and compare sequence on the base of functionality
- c) align and compare sequence on the base of folding
- d) All of above
122. A common fold is a Bundle and trim barrel
- a) 2 helix b) 5 helix c) 3 helix d) 4 helix