BSc Degree Examinations 2019-20

Department: BIOLOGY

Title of Exam: Molecular Recognition

Time Allowed: 24 Hours (PLEASE NOTE: Late papers will not be marked)

Time Recommended: 2 hours and 30 minutes

Word Limit: Please answer the questions within the line and word limits stated. Content beyond the line limit will not be marked.

Allocation of Marks: Total marks available for this paper: 100

Total marks available:
Section A: 50, Section B: 50
The marks available for each question are indicated on the paper.

Instructions for Candidates:
All questions should be answered on this question paper using minimum font size Arial 11.
Each question should be answered within the stated line or word limit.
Do not adjust the margin width.

Section A: Answer all questions
Section B: Answer either question A, B or C.
A note on Academic Integrity

We are treating this online examination as a time-limited open assessment, and you are therefore permitted to refer to written and online materials to aid you in your answers.

However, you must ensure that the work you submit is entirely your own, and for the whole time the assessment is live you must not:

- communicate with departmental staff on the topic of the assessment
- communicate with other students on the topic of this assessment.
- seek assistance with the assignment from the academic and/or disability support services, such as the Writing and Language Skills Centre, Maths Skills Centre and/or Disability Services. (The only exception to this will be for those students who have been recommended an exam support worker in a Student Support Plan. If this applies to you, you are advised to contact Disability Services as soon as possible to discuss the necessary arrangements.)
- seek advice or contribution from any third party, including proofreaders, friends, or family members.

We expect, and trust, that all our students will seek to maintain the integrity of the assessment, and of their award, through ensuring that these instructions are strictly followed. Failure to adhere to these requirements will be considered a breach of the Academic Misconduct regulations, where the offences of plagiarism, breach/cheating, collusion and commissioning are relevant - see AM.1.2.1” (Note this supersedes section 7.3 of the Guide to Assessment).
Section A:
Mark total for this section: 50

1. FadR is a dimeric transcriptional repressor from the archaeon *Sulfolobus acidocaldarius* that binds to the *Saci-1123* operator. Electrophoretic mobility shift assays (EMSAs) were performed incubating the protein with wild type and a mutant operator, in which one base of the motif had been changed.

![Electrophoretic mobility shift assays (EMSAs)](image)

<table>
<thead>
<tr>
<th>FadR&lt;sub&gt;Sa&lt;/sub&gt; WT (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
</tr>
<tr>
<td>B1</td>
</tr>
</tbody>
</table>

**Saci_1123** wild type operator

**Saci_1123** mutant operator

a) Discuss the patterns obtained with the two operators and rationalise the differences. (7 marks, 14 lines)

FadR binds the wild type operator forming a single complex (1 mark). The binding is very efficient as the free DNA disappears from the bottom of the gel in most of the reactions (1 mark). FadR binds to the mutant operator forming two complexes, B1 and B2, that migrate to different positions in the gel as they have different molecular weights (1 mark). The binding to the mutant operator is less efficient, as free DNA is visible at the bottom of the gel at all protein concentrations (1 mark) and the binding starts at a FadR concentration of 1.5 micromolar as compared to 0.53 micromolar with wild type operator (1 mark). In the B1 complex FadR binds as a dimer (1 mark), whereas in the B2 complex FadR binds as two dimers (1 mark).
b) The residue N37 in FadR is replaced with alanine and the mutant protein is incubated with the mutant Saci-1123 operator in an EMSA experiment. What do the results tell us about the role of N37? (3 marks, 6 lines)

![EMSA experiment image](image_url)

FadR binds the DNA forming more B1 and less B2 complex (1 mark), indicating that FadR binds mostly as a dimer rather than a dimer-of-dimer (1 mark), suggesting that N37 is important for stabilising dimer-dimer interaction (1 mark).

FEEDBACK: most students answered this question correctly. I think that having gone through the paper together during the lecture has helped.

Module Learning Outcome 3: describe common biochemical and biophysical techniques for the study of molecular recognition in vitro and in vivo, interpret data obtained using these techniques, and design experimental strategies using an appropriate combination of these techniques.

2. An anti-lysozyme antibody fragment (Fab) binds to lysozyme (L) to form a 1:1 complex.

   a) Equal volumes of 50 µM Fab and 40 µM L are mixed and allowed to equilibrate. The equilibrium concentration of L is found to be 2 µM. Showing your working, calculate K_d, the equilibrium dissociation constant of the complex. (3 marks, 3 lines)

   \[
   [\text{Fab}]_0 = 25 \text{ µM}; \quad [\text{L}]_0 = 20 \text{ µM}; \quad [\text{L}]_{\text{eq}} = 2 \text{ µM}; \quad [\text{Fab-L}]_{\text{eq}} = 18 \text{ µM}; \quad [\text{Fab}]_{\text{eq}} = 7 \text{ µM}
   
   K_d = [\text{Fab}]_{\text{eq}} x [\text{L}]_{\text{eq}} / [\text{Fab-L}]_{\text{eq}} = 7 \times 2 / 18 = 14/18 \text{ µM} = 0.8 \text{ µM}
   \]
Solvent accessible surface areas derived from crystal structures of the Fab alone, lysozyme alone and the Fab-lysozyme complex are

- Fab: 19,500 Å²
- Lysozyme: 6,500 Å²
- Fab-Lysozyme Complex: 24,600 Å²

b) Stating any assumptions, calculate the buried surface area when Fab and lysozyme interact. (2 marks, 4 lines)

\[
BSA = ASA(Fab) + ASA(Lysozyme) - ASA(Fab-Lysozyme) = 19500 + 6500 - 24600 = 1400 \text{ Å}^2
\]

Assumption is that there is no conformational change accompanying complex formation.

FEEDBACK: Very few students realised that mixing equal volumes of two solutions will halve the concentration in the resulting mixture and as a result there were very few full marks for Part a). Contrastingly almost all students correctly calculated the BSA and stated the assumption correctly.

3. The figure below shows data from three SEC-MALLS experiments on samples of a repressor protein R (A: blue trace), and its antagonist I (B: green trace) and an equimolar mixture of the two proteins (C: red trace). The continuous lines represent the differential refractive index of the column eluate, the thicker horizontal lines represent the molecular mass. The subunit molecular mass of R is 13 kDa and that of I is 7 kDa.
Giving your reasoning, deduce the following: (3 marks, 1 line per answer)

a) The oligomeric state of R
   R is a tetramer as evidenced by the mass of ~55 kDa (4 x 13 kDa)

b) The oligomeric state of I
   I is a dimer as evidenced by the mass of ~14 kDa (2 x 7 kDa)

c) The stoichiometry of the R:I complex
   The R:I complex is a heterodimer as evidenced by the mass of 20 kDa (13 kDa + 7 kDa)

FEEDBACK: This question presented few problems and many answers gained full marks.

4. A small GTPase CHOC1 is activated by a guanine nucleotide exchange factor CHOC1-GEF. The central signalling protein kinase PKB is also involved in CHOC1 activation, though CHOC1-GEF itself is not a direct target of PKB.

Figure 1. Activation of CHOC1. A) Binding of CHOC1-GEF to artificial liposomes (vesicles) containing different phosphinositides was measured and the amount of bound CHOC1 recorded. The data are normalized to the amount of CHOC1 bound to purified cellular membranes. B) GTP binding of CHOC1 was measured in the presence of CHOC1-GEF and either PI(3,4)P₂, active PKB, or both.
a) Based on the data presented in Figure 1, formulate a hypothesis for CHOC1 activation. (3 marks, 6 lines)

The hypothesis would have to include either co-recruitment of the GEF and GTPase to the membrane by PI3,4P2 or a conformational change in the GEF CHOC1 due to membrane recruitment that activates it, allowing GTPase activation. (1 mark) A further aspect should be on PKB involvement, phosphorylation by PKB of the GTPase should make it either more open ready for exchange, or the phosphate could repel the GDP, or it could allow stronger GEF binding. (1 mark)

Students should point out that the effects of PI3,4P2 binding and PKB mediated phosphorylation are synergistic, their combined effect being stronger than the addition of individual ones. (1 mark)

FEEDBACK: Many said that PKB would be used to phosphorylate PIPs, which is not correct. PKB is a protein kinase as it says in the question text. Some linking of PIP and PKB involvement was needed for full marks.

b) Describe an experiment that tests the involvement of PKB in CHOC1 activation. (2 marks, 6 lines)

Dependent on the hypothesis presented for the action of the phosphate group on the GTPase, either binding studies with the GEF, a molecular dynamics simulation of conformational flexibility, or binding studies with the membrane vesicles should be described, mentioning the technique used and at least one control condition.

FEEDBACK: Many suggested in vivo methods based on PKB knock-out or inhibition in cells rather than an in vitro method. These were given marks as long as the CHOC1 activity readout was well described, and a proper control was included.

c) Explain why the following reagents would be useful to test CHOC1 activity in vivo:

i) An inhibitor of PI-4-kinase (2 marks, 4 lines)

PI3,4P2 is an essential component of the membranes for full CHOC1 activation (1 mark). PI-4-kinase is required for the phosphorylation of PIP at the 4-position, so its inhibition will reduce PI3,4P2 levels, and therefore DU1 activity (1 mark).

FEEDBACK: Answered generally well. It was important to know that it was phosphorylation of PIP at position 4 that was inhibited by this reagent.
ii) A plasmid for the expression of the CHOC1-Gln67Ala mutant in HeLa cells. Gln67 in CHOC1 is equivalent to Gln61 in the homologous Ras GTPase.

(2 marks, 4 lines)

While the equivalent Ras mutant is dominant active (1 mark), this property does not always translate to homologous GTPases, so this mutant is not useful for testing CHOC1 activity without further tests (1 mark).

FEEDBACK: This question was referring to material taught in the lecture, and needed the student to synthesize knowledge from that lecture to get full marks.

5. Lin28 is a multi-domain RNA-binding protein, which has the following domain structure:

```
N         CSD         ZnF     ZnF         C
```

CSD: cold shock domain
ZnF: CCHC-type Zinc-finger domain

A series of EMSA experiments were performed to evaluate how the CSD and ZnF regions contributed to the recognition of a precursor microRNA target by Lin28. The data are shown below:

The same range of protein concentrations was used in each EMSA (from left to right): 8 nM, 33 nM, 130 nM, 520 nM, 2.1 µM, 8.3 µM, 33.3 µM
a) Giving your reasons, explain why the EMSA data suggest that the linker connecting CSD and ZnF regions of Lin28 is flexible. 

The individual CSD and ZnF constructs bind more weakly (8.3 uM and > 33 uM, respectively) than FL, 0.5 uM (1 mark). Combining both domains on the same construct (CSD-CCHC) shows WT-like binding affinity (1 mark) indicating that the two domains are necessary and sufficient to reconstitute WT binding (1 mark). The increase in binding on combining the two domains is less than anticipated by the product of their individual binding affinities (1 mark) which suggests that the linker between the two domains is flexible (1 mark).

Evidence of higher order binding at higher protein concentration is not relevant to the question.

FEEDBACK: Generally done quite well though only one person got full marks. Most got the first 3 marks but less got the second 2. The main feedback was that the question said “explain why the EMSA data” but several answers used knowledge about Lin28 to answer the question.

The cold-shock domain (such as that found in Lin28) and the PAZ domain (such as that found in Dicer) both contain an oligonucleotide-binding (OB) fold.

b) Contrast how the PAZ domain of Dicer and the cold shock domain of Lin28 recognise precursor microRNAs

CSD of Lin28 recognises a specific sequence motif in pre-miR ligands (1 mark, required).

Additional mark from: CSD binds to the pre-E loop of pre-miRs (1 mark) via exposed residues on a flat surface of the domain (1 mark). Surface exposed residues make sequence specific interactions (1 mark)

PAZ domain of Dicer does not recognise a specific sequence motif; or PAZ domain of Dicer recognises structure/chemistry rather than sequence (1 mark, required)

Additional mark from: PAZ domain binds to 3’ single stranded motifs in pre-miRs (1 mark). PAZ domain has a pocket that binds 3’ overhang (1 mark). Residues in binding pocket interact with sugar-phosphate backbone rather than bases (1 mark)

FEEDBACK: Very well done on the whole. I was strict on not awarding if people said that an interaction was specific/non-specific. You had to say sequence-specific or sequence independent. Both domains are specific in their interactions.
6.

a) Describe a biophysical technique that could be used to directly observe DNA loop formation. 

A single-molecule tethered particle assay could be used to directly measure loop formation (1 mark, required). This would require a linear DNA template labelled with biotin and digoxigenin labels at opposite ends, which contains binding sites with the appropriate spacing (~300 bp) (1 mark, required). DNA template would be tethered between anti-DIG IgG coated surface and streptavidin-coated sub-micron (1 mark). Real-time loop formation would be observed as a reduction in the Brownian motion of the sub-micron bead due to shortening of the DNA tether (1 mark, required).

FEEDBACK: Most students provided a subset of the experimental detail required to describe this biophysical technique. Details of the differentially end-labelled linear DNA template were frequently omitted. Some students mistakenly included details about the magnetic tweezers technique.

b) Outline an in vitro experimental strategy using this technique with appropriate controls that could be used to investigate the role of lac repressor protein tetramerisation in DNA loop formation. 

Investigate ability of wild-type repressor protein to mediate DNA loop formation using DNA templates with one or two binding (operator) sites (1 mark, required). Protein tetramerisation could be disrupted through site-directed mutagenesis or deletion of tetramerisation domains (1 mark) and the ability of the mutant protein to form DNA loops quantified on the template with two binding sites (1 mark, required). It would be important here to establish that the protein mutations only affect tetramerisation, but not dimerisation of the repressor and operator site binding (1 mark, required).

FEEDBACK: Many students provided an incomplete experimental strategy which did not properly test the role of repressor tetramerisation and/or did not include the necessary controls.
7. An electron microscopy study of a newly discovered virus revealed icosahedral structures, with \( n \) subunits per capsid, as shown schematically below.

![](image)

**a)** Locations of potential symmetry axes of the icosahedron are circled and labelled with a letter. Which letter corresponds to:

i) the 2-fold symmetry axis

\( e \)  

(1 mark, 1 line)

ii) the 3-fold symmetry axis

\( b \)  

(1 mark, 1 line)

iii) the 5-fold symmetry axis

\( d \)  

(1 mark, 1 line)

**b)** What is the \( T \) number for the capsid shown above?  

\[
h = 1, \ k = 1 \quad \text{(1 mark), hence } T = h^2 + h \times k + k^2 = 3 \quad \text{(1 mark)}
\]

(2 marks, 2 lines)

**c)** Showing any calculations, how many subunits are there in this capsid?  

\[
n = 60 \times T = 60 \times 3 = 180 \text{ subunits} \quad \text{(1 mark)}
\]

(1 mark, 1 lines)

**FEEDBACK:** Most students were able to identify correct positions for two out of three symmetry axes. However, only very few received full marks, having problems with
either locating all symmetry axes or with working out the correct h,k numbers and hence the correct T number and the overall number of subunits.

Section B

Answer either question A, B OR C. Maximum 1000 words.

Mark total for this section: 50

A) A new DNA-binding protein has been identified in *Escherichia coli* that regulates the expression of multiple genes involved in amino acid biosynthesis. Discuss how you would identify where the protein binds to chromosomal DNA and the *in vitro* methods that could be employed to characterise the DNA-binding properties of the protein.

The FadR paper by Wang et al. discussed in one of the lectures provides numerous ideas for this essay, as the paper reports the characterization of a transcriptional repressor from discovery of the gene to solution of the crystal structure.

The essay would start with the discussion of how ChIP-seq would identify the chromosomal sites with which the DNA-binding protein associates in the cell. A brief description of the method would be appropriate. Once the regions to which the protein binds are determined *in vivo*, a number of different approaches can be discussed to investigate the DNA-binding features of the protein *in vitro*. Examples of experimental approaches include, but are not limited to:

- EMSA with DNA fragments corresponding to the regions identified with ChIP-seq
- DNase I or other DNA footprinting to define the sequence of the protein binding site
- Size-exclusion chromatography to determine the stoichiometry of the protein-DNA interaction
- SPR or other quantitative method to determine the $K_D$ of the protein-DNA interaction
- Solution of the protein structure.

The essay is quite open-ended and can include methods that have been presented in the lectures of other colleagues.

FEEDBACK: I would point the students to the individual feedback that was annotated on each essay. This was the most popular essay having been chosen by 16 students. Overall, all the students provided a reasonable overview of the area, spanning from 2:2 to 1st class level. The material was mostly based on the module lectures.
Module Learning Outcome 3: describe common biochemical and biophysical techniques for the study of molecular recognition in vitro and in vivo, interpret data obtained using these techniques, and design experimental strategies using an appropriate combination of these techniques.

B) Discuss how protein phosphorylation and dephosphorylation are used to control the interactions of proteins that determine cell fate in *Bacillus subtilis*.

This question most obviously would be answered in relation to two lectures on the system of proteins controlling the activation of sigma F in the forespore during spore formation in *Bacillus subtilis*. The system comprises the sigma factor itself whose capacity to transcribe cell-type specific genes is determined by an anti-sigma factor. The anti-sigma factor is also a Ser/Thr protein kinase that phosphorylates a third protein which has a role as an anti-anti-sigma factor. Phosphorylation of the latter inactivates it. Reactivation by a membrane localising phosphatase is required for the sigma to become active. A good answer here requires clarity of thought and organisation of the essay. The scene needs to be set with a description of the underlying biological process and the alternate cell fates. An authoritative answer would include a description of phosphorylation in terms of its simple chemistry. Then the determinants of switching can be brought in all the keeping a focus on phosphorylation and dephosphorylation.

C) How do co-operativity and post-translational modification affect the affinity of the KIX domain of p300/CBP for its binding partners?

This question relates to how the KIX domain recognises its peptide binding partners.

An acceptable answer will cover material from lecture 12 dealing with two examples: 1) phosphorylated pKID, which binds to the c-Myb site on KIX; and 2) co-operativity between MLL and cMyb peptides. Answers should cover the parent proteins of each motif measured. Better answers will look at similarities of the motifs.

Answers relating to KIX/pKID should state that: the KID motif is found in CREB and that it is phosphorylated by PKA; that phosphorylation increases affinity 100-fold (from what to what would be good); the structure of the KIX/pKID complex; that pKID is unstructured in the absence of KIX; what we know about how pKID is recognised by KIX, i.e. the binding pathway as described by Wright and colleagues in 2007 (this was covered in detail in the paper). Better answers would include information about the experiments that led to our understanding and the techniques used would be welcome. pKID binds the c-Myb site. Good answers might explore the differences in...
how these two peptide ligands bind KIX: pKID via a mechanism that is more induced fit; c-Myb via a mixture of induced fit and conformational selection. These terms would be properly explained in better answers.

The second part of the question refers to what happens when peptide ligands bind at both c-Myb and MLL sites. A basic answer will state that co-operative binding means that occupancy at one site enhances binding at the other. Better answers will give values of enhancement and potentially show figures for routes to a double-bound KIX domain. Stronger answers would discuss mutations that explore binding site specificity, i.e. that MLL peptide can be seen to bind at the c-Myb site in an appropriate KIX mutant (L628A). The best answers will consider studies that explore mechanisms of allostery and the role of aliphatic residues in transmitting binding state between sites.

FEEDBACK: Only 5 chose this essay and so difficult to draw confident conclusions. Answers ranged from mid-50s to mid 70s. Most dealt with material from lectures. Lower scoring essays didn’t provide much mechanistic detail and/or left out important parts of the answer (e.g. dealing only with PTMs or modification of affinity and not linking both; not providing any details about how we know these features of KIX domain activity, etc). The best answer really dealt with what is known about the molecular basis for KIX co-operativity, i.e. that binding at one site causes a change in side chain conformation in a network of hydrophobic residues that causes residues in the other binding site to adopt their bound state conformation.

LO1: Demonstrate understanding of the basic structural features of protein-target complexes, the role that solvent plays in recognition, and how conformational changes, modularity, symmetry, intrinsic disorder and structural plasticity can influence binding affinity and specificity.