CONTRIBUTORS
Daniel J. Pallin, M.D.
Senior Author
Judene Wright, M.S., M.A.Ed.
Senior Author

TPR MCAT Biology and Biochemistry Development Team:
Jessica Adams, Ph.D.
Andrew D. Snyder, M.D.
Jenkang Tao, B.S., B.A.
Judene Wright, M.S., M.A.Ed., Senior Editor, Lead Developer
Sarah Woodruff, B.S., B.A.

Edited for Production by:
Judene Wright, M.S., M.A.Ed.
National Content Director, MCAT Program, The Princeton Review

The TPR MCAT Biology and Biochemistry Team and Judene would like to thank the following people for their contributions to this book:
### Periodic Table of the Elements

<table>
<thead>
<tr>
<th>Element</th>
<th>Atomic Number</th>
<th>Mass Number</th>
<th>Mass (amu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>Li</td>
<td>3</td>
<td>6</td>
<td>6.9</td>
</tr>
<tr>
<td>Be</td>
<td>4</td>
<td>9</td>
<td>9.0</td>
</tr>
<tr>
<td>Na</td>
<td>11</td>
<td>23</td>
<td>23.0</td>
</tr>
<tr>
<td>Mg</td>
<td>12</td>
<td>24.3</td>
<td>24.3</td>
</tr>
<tr>
<td>K</td>
<td>19</td>
<td>39</td>
<td>39.1</td>
</tr>
<tr>
<td>Ca</td>
<td>20</td>
<td>40.1</td>
<td>40.1</td>
</tr>
<tr>
<td>Sc</td>
<td>21</td>
<td>45</td>
<td>45.0</td>
</tr>
<tr>
<td>Ti</td>
<td>22</td>
<td>47</td>
<td>47.9</td>
</tr>
<tr>
<td>V</td>
<td>23</td>
<td>50</td>
<td>50.9</td>
</tr>
<tr>
<td>Cr</td>
<td>24</td>
<td>52</td>
<td>52.0</td>
</tr>
<tr>
<td>Mn</td>
<td>25</td>
<td>54</td>
<td>54.9</td>
</tr>
<tr>
<td>Fe</td>
<td>26</td>
<td>56</td>
<td>55.8</td>
</tr>
<tr>
<td>Co</td>
<td>27</td>
<td>58</td>
<td>58.9</td>
</tr>
<tr>
<td>Ni</td>
<td>28</td>
<td>60</td>
<td>58.7</td>
</tr>
<tr>
<td>Cu</td>
<td>29</td>
<td>63</td>
<td>63.5</td>
</tr>
<tr>
<td>Zn</td>
<td>30</td>
<td>65</td>
<td>65.4</td>
</tr>
<tr>
<td>Ga</td>
<td>31</td>
<td>69</td>
<td>69.7</td>
</tr>
<tr>
<td>As</td>
<td>33</td>
<td>75</td>
<td>74.9</td>
</tr>
<tr>
<td>Se</td>
<td>34</td>
<td>79</td>
<td>79.0</td>
</tr>
<tr>
<td>Br</td>
<td>35</td>
<td>87</td>
<td>79.9</td>
</tr>
<tr>
<td>Kr</td>
<td>36</td>
<td>88</td>
<td>83.8</td>
</tr>
<tr>
<td>Rb</td>
<td>37</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Sr</td>
<td>38</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zr</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nb</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mo</td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tc</td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ru</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rh</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pd</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ag</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cd</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sn</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sb</td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Br</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kr</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cs</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ba</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hf</td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ta</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Os</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ir</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pt</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Au</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hg</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tl</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pb</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bi</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Po</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rn</td>
<td>86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fr</td>
<td>87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ra</td>
<td>88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ac</td>
<td>89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rf</td>
<td>90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Db</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sg</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bh</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hs</td>
<td>94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mt</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ds</td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rg</td>
<td>97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cn</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uu</td>
<td>102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lv</td>
<td>103</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Lanthanide Series:*

<table>
<thead>
<tr>
<th>Element</th>
<th>Atomic Number</th>
<th>Mass Number</th>
<th>Mass (amu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ce</td>
<td>58</td>
<td></td>
<td>140.1</td>
</tr>
<tr>
<td>Pr</td>
<td>59</td>
<td></td>
<td>140.9</td>
</tr>
<tr>
<td>Nd</td>
<td>60</td>
<td></td>
<td>144.2</td>
</tr>
<tr>
<td>Sm</td>
<td>62</td>
<td></td>
<td>150.4</td>
</tr>
<tr>
<td>Eu</td>
<td>63</td>
<td></td>
<td>152.0</td>
</tr>
<tr>
<td>Gd</td>
<td>64</td>
<td></td>
<td>157.3</td>
</tr>
<tr>
<td>Tb</td>
<td>65</td>
<td></td>
<td>158.9</td>
</tr>
<tr>
<td>Dy</td>
<td>66</td>
<td></td>
<td>162.5</td>
</tr>
<tr>
<td>Ho</td>
<td>67</td>
<td></td>
<td>164.9</td>
</tr>
<tr>
<td>Er</td>
<td>68</td>
<td></td>
<td>167.3</td>
</tr>
<tr>
<td>Tm</td>
<td>69</td>
<td></td>
<td>168.9</td>
</tr>
<tr>
<td>Yb</td>
<td>70</td>
<td></td>
<td>173.0</td>
</tr>
<tr>
<td>Lu</td>
<td>71</td>
<td></td>
<td>175.0</td>
</tr>
</tbody>
</table>

†Actinide Series:*

<table>
<thead>
<tr>
<th>Element</th>
<th>Atomic Number</th>
<th>Mass Number</th>
<th>Mass (amu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th</td>
<td>90</td>
<td></td>
<td>232.0</td>
</tr>
<tr>
<td>Pa</td>
<td>91</td>
<td></td>
<td>231</td>
</tr>
<tr>
<td>U</td>
<td>92</td>
<td></td>
<td>238.0</td>
</tr>
<tr>
<td>Np</td>
<td>93</td>
<td></td>
<td>237</td>
</tr>
<tr>
<td>Pu</td>
<td>94</td>
<td></td>
<td>244</td>
</tr>
<tr>
<td>Am</td>
<td>95</td>
<td></td>
<td>243</td>
</tr>
<tr>
<td>Cm</td>
<td>96</td>
<td></td>
<td>247</td>
</tr>
<tr>
<td>Bk</td>
<td>97</td>
<td></td>
<td>247</td>
</tr>
<tr>
<td>Cf</td>
<td>98</td>
<td></td>
<td>251</td>
</tr>
<tr>
<td>Es</td>
<td>99</td>
<td></td>
<td>252</td>
</tr>
<tr>
<td>Fm</td>
<td>100</td>
<td></td>
<td>257</td>
</tr>
<tr>
<td>Md</td>
<td>101</td>
<td></td>
<td>258</td>
</tr>
<tr>
<td>No</td>
<td>102</td>
<td></td>
<td>259</td>
</tr>
<tr>
<td>Lr</td>
<td>103</td>
<td></td>
<td>260</td>
</tr>
</tbody>
</table>
# CONTENTS

Register Your Book Online! ........................................................................................................................................... x

## CHAPTER 1: MCAT BASICS ........................................................................................................................................ 1

## CHAPTER 2: BIOLOGY STRATEGY FOR THE MCAT .................................................................................................. 15

2.1 Science Sections Overview ................................................................................................................................. 16
2.2 Science Passage Types ........................................................................................................................................... 17
2.3 Science Question Types ........................................................................................................................................ 21
2.4 Biology on the MCAT ............................................................................................................................................ 23
2.5 Tackling a Biology Passage ................................................................................................................................... 23
2.6 Tackling the Questions ......................................................................................................................................... 32
2.7 Summary of the Approach to Biology .................................................................................................................. 36

## CHAPTER 3: BIOLOGICALLY IMPORTANT MOLECULES ............................................................................................ 39

3.1 Protein Building Blocks ......................................................................................................................................... 41
3.2 Protein Structure ..................................................................................................................................................... 42
3.3 Carbohydrates ........................................................................................................................................................ 46
3.4 Lipids ......................................................................................................................................................................... 48
3.5 Phosphorus-Containing Compounds .................................................................................................................... 54
Chapter 3 Summary ................................................................................................................................................... 56
Chapter 3 Freestanding Practice Questions ............................................................................................................. 58
Chapter 3 Practice Passage Questions ..................................................................................................................... 59
Solutions ..................................................................................................................................................................... 61

## CHAPTER 4: MOLECULAR BIOLOGY ........................................................................................................................ 63

4.1 DNA Structure .......................................................................................................................................................... 65
4.2 Genome Structure and Genomic Variations ............................................................................................................ 72
4.3 The Role of DNA ...................................................................................................................................................... 73
4.4 DNA Replication ...................................................................................................................................................... 75
4.5 Genetic Mutation ...................................................................................................................................................... 83
4.6 DNA Repair ............................................................................................................................................................. 92
4.7 Gene Expression: Transcription ............................................................................................................................. 95
4.8 Gene Expression: Translation ................................................................................................................................ 103
4.9 Controlling Gene Expression ................................................................................................................................ 112
4.10 Return to Gene Structure: A Summary .................................................................................................................... 123
Chapter 4 Summary ................................................................................................................................................... 125
Chapter 4 Freestanding Practice Questions ............................................................................................................. 127
Chapter 4 Practice Passage Questions ..................................................................................................................... 128
Solutions ..................................................................................................................................................................... 130
## CONTENTS

### CHAPTER 8: THE NERVOUS AND ENDOCRINE SYSTEMS

- 8.1 Neuronal Structure and Function .................................................. 275
- 8.2 Synaptic Transmission ................................................................. 283
- 8.3 Functional Organization of the Human Nervous System ............... 286
- 8.4 Anatomical Organization of the Nervous System ......................... 289
- 8.5 Sensation and Perception ............................................................ 297
- 8.6 The Endocrine System ............................................................... 310
- Chapter 8 Summary ......................................................................... 316
- Chapter 8 Freestanding Practice Questions ...................................... 320
- Chapter 8 Practice Passage ............................................................. 322
- Solutions ......................................................................................... 324

### CHAPTER 9: THE CIRCULATORY, LYMPHATIC, AND IMMUNE SYSTEMS

- 9.1 Overview of the Circulatory System .............................................. 328
- 9.2 The Heart .................................................................................. 330
- 9.3 Hemodynamics .......................................................................... 339
- 9.4 Components of Blood ................................................................. 342
- 9.5 Transport of Gases ..................................................................... 346
- 9.6 The Lymphatic System ................................................................ 351
- 9.7 The Immune System .................................................................. 352
- 9.8 Autoimmunity ............................................................................ 356
- Chapter 9 Summary ......................................................................... 358
- Chapter 9 Freestanding Practice Questions ...................................... 360
- Chapter 9 Practice Passage ............................................................. 362
- Solutions ......................................................................................... 364

### CHAPTER 10: THE EXCRETORY AND DIGESTIVE SYSTEMS

- 10.1 Overview of the Excretory System .............................................. 370
- 10.2 Anatomy and Function of the Urinary System ............................. 372
- 10.3 Renal Regulation of Blood Pressure and pH ............................... 379
- 10.4 Endocrine Role of the Kidney .................................................... 381
- 10.5 Overview of the Digestive System ............................................. 382
- 10.6 The Gastrointestinal Tract .......................................................... 386
- 10.7 The GI Accessory Organs ........................................................... 393
- 10.8 A Day in the Life of Food ............................................................ 396
- 10.9 Vitamins .................................................................................... 400
- Chapter 10 Summary ...................................................................... 401
- Chapter 10 Freestanding Practice Questions .................................... 403
- Chapter 10 Practice Passage .......................................................... 404
- Solutions ......................................................................................... 406
**CHAPTER 11: THE MUSCULAR AND SKELETAL SYSTEMS** .......................................................... 409  
11.1 Overview of Muscle Tissue ......................................................................................... 410  
11.2 Skeletal Muscle ......................................................................................................... 410  
11.3 Cardiac Muscle Compared to Skeletal Muscle ....................................................... 419  
11.4 Smooth Muscle Compared to Skeletal Muscle ......................................................... 421  
11.5 Overview of the Skeletal System ............................................................................. 425  
11.6 Connective Tissue .................................................................................................... 426  
11.7 Bone Structure ......................................................................................................... 426  
11.8 Tissues Found at Joints ............................................................................................ 428  
11.9 Bone Growth and Remodeling: the Cells of Bone ................................................. 429  
Chapter 11 Summary .................................................................................................. 431  
Chapter 11 Freestanding Practice Questions ............................................................... 433  
Chapter 11 Practice Passage ....................................................................................... 434  
Solutions ..................................................................................................................... 436  

**CHAPTER 12: THE RESPIRATORY SYSTEM AND THE SKIN** .......................................... 439  
12.1 Functions of the Respiratory System .................................................................. 440  
12.2 Anatomy of the Respiratory System .................................................................. 441  
12.3 Pulmonary Ventilation ......................................................................................... 445  
12.4 Gas Exchange ........................................................................................................ 448  
12.5 Regulation of Ventilation Rate ........................................................................... 452  
12.6 Structure and Layers of the Skin .......................................................................... 454  
12.7 Temperature Regulation by the Skin .................................................................. 455  
Chapter 12 Summary .................................................................................................. 456  
Chapter 12 Freestanding Practice Questions ............................................................... 458  
Chapter 12 Practice Passage ....................................................................................... 460  
Solutions ..................................................................................................................... 462  

**CHAPTER 13: THE REPRODUCTIVE SYSTEMS** .......................................................... 465  
13.1 The Male Reproductive System ........................................................................... 466  
13.2 Spermatogenesis .................................................................................................... 468  
13.3 Development of the Male Reproductive System ................................................... 472  
13.4 Androgens and Estrogens .................................................................................... 473  
13.5 The Female Reproductive System ...................................................................... 475  
13.6 Oogenesis and Ovulation .................................................................................... 477  
13.7 The Menstrual Cycle .............................................................................................. 479  
13.8 Hormonal Changes During Pregnancy .................................................................. 483  
13.9 Fertilization and Cleavage .................................................................................... 484  
13.10 Implantation and the Placenta ............................................................................. 487  
13.11 Post-Implantation Development ......................................................................... 488  
13.12 Differentiation ....................................................................................................... 490
1. Go to PrincetonReview.com/cracking

2. You’ll see a welcome page where you should register your book or boxed set of books using the ISBN. If you have a book, the ISBN can be found above the bar code on the back cover. If you have a boxed set, the ISBN can be found on the back of the box above the bar code.

3. After placing this free order, you’ll either be asked to log in or to answer a few simple questions in order to set up a new Princeton Review account.

4. Finally, click on the “Student Tools” tab located at the top of the screen. It may take an hour or two for your registration to go through, but after that, you’re good to go.

NOTE: If you are experiencing book problems (potential content errors), please contact EditorialSupport@review.com with the full title of the book, its ISBN number, and the page number of the error.

Experiencing technical issues? Please email TPRStudentTech@review.com with the following information:

- your full name
- e-mail address used to register the book
- full book title and ISBN
- your computer OS (Mac or PC) and Internet browser (Firefox, Safari, Chrome, etc.)
- description of technical issue
Once you’ve registered, you can...

- Take 3 full-length practice MCAT exams
- Find useful information about taking the MCAT and applying to medical school
- Check to see if there have been any updates to this edition

Offline Resources

If you are looking for more review or medical school advice, please feel free to pick up these books in stores right now!

- Medical School Essays That Made a Difference
- The Best 167 Medical Schools
- The Princeton Review Complete MCAT
Chapter 1
MCAT Basics
SO YOU WANT TO BE A DOCTOR

So...you want to be a doctor. If you’re like most premeds, you’ve wanted to be a doctor since you were pretty young. When people asked you what you wanted to be when you grew up, you always answered “a doctor.” You had toy medical kits, bandaged up your dog or cat, and played “hospital.” You probably read your parents’ home medical guides for fun.

When you got to high school you took the honors and AP classes. You studied hard, got straight A’s [or at least really good grades!], and participated in extracurricular activities so you could get into a good college. And you succeeded!

At college you knew exactly what to do. You took your classes seriously, studied hard, and got a great GPA. You talked to your professors and hung out at office hours to get good letters of recommendation. You were a member of the premed society on campus, volunteered at hospitals, and shadowed doctors. All that’s left to do now is get a good MCAT score.

Just the MCAT.

Just the most confidence-shattering, most demoralizing, longest, most brutal entrance exam for any graduate program. At about 7.5 hours [including breaks], the MCAT tops the list...even the closest runners up, the LSAT and GMAT, are only about 4 hours long. The MCAT tests significant science content knowledge along with the ability to think quickly, reason logically, and read comprehensively, all under the pressure of a timed exam.

The path to a good MCAT score is not as easy to see as the path to a good GPA or the path to a good letter of recommendation. The MCAT is less about what you know, and more about how to apply what you know...and how to apply it quickly to new situations. Because the path might not be so clear, you might be worried. That’s why you picked up this book.

We promise to demystify the MCAT for you, with clear descriptions of the different sections, how the test is scored, and what the test experience is like. We will help you understand general test-taking techniques as well as provide you with specific techniques for each section. We will review the science content you need to know as well as give you strategies for the Critical Analysis and Reasoning Skills [CARS] section. We’ll show you the path to a good MCAT score and help you walk the path.

After all...you want to be a doctor. And we want you to succeed.
WHAT IS THE MCAT...REALLY?
Most test-takers approach the MCAT as though it were a typical college science test, one in which facts and knowledge simply need to be regurgitated in order to do well. They study for the MCAT the same way they did for their college tests, by memorizing facts and details, formulas and equations. And when they get to the MCAT they are surprised…and disappointed.

It’s a myth that the MCAT is purely a content-knowledge test. If medical-school admission committees want to see what you know, all they have to do is look at your transcripts. What they really want to see is how you think, especially how you think under pressure. That’s what your MCAT score will tell them.

The MCAT is really a test of your ability to apply basic knowledge to different, possibly new, situations. It’s a test of your ability to reason out and evaluate arguments. Do you still need to know your science content? Absolutely. But not at the level that most test-takers think they need to know it. Furthermore, your science knowledge won’t help you on the Critical Analysis and Reasoning Skills (CARS) section. So how do you study for a test like this?

You study for the science sections by reviewing the basics and then applying them to MCAT practice questions. You study for the CARS section by learning how to adapt your existing reading and analytical skills to the nature of the test. (More information about the CARS section can be found in MCAT Critical Analysis and Reasoning Skills Review.)

The book you are holding will review all the relevant MCAT Biology content you will need for the test, and a little bit more. It includes hundreds of questions designed to make you think about the material in a deeper way, along with full explanations to clarify the logical thought process needed to get to the answer. It also comes with access to three full-length online practice exams to further hone your skills. For more information on accessing those online exams, please refer to the “Register Your Book Online!” spread on page x.
MCAT NUTS AND BOLTS

Overview
The MCAT is a computer-based test (CBT) that is not adaptive. Adaptive tests base your next question on whether or not you’ve answered the current question correctly. The MCAT is linear, or fixed-form, meaning that the questions are in a predetermined order and do not change based on your answers. However, there are many versions of the test, so that on a given test day, different people will see different versions. The following table highlights the features of the MCAT exam.

<table>
<thead>
<tr>
<th>Registration</th>
<th>Online via <a href="http://www.aamc.org">www.aamc.org</a>. Begins as early as six months prior to test date; available up until week of test (subject to seat availability).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing Centers</td>
<td>Administered at small, secure, climate-controlled computer testing rooms.</td>
</tr>
<tr>
<td>Security</td>
<td>Photo ID with signature, electronic fingerprint, electronic signature verification, assigned seat.</td>
</tr>
<tr>
<td>Proctoring</td>
<td>None. Test administrator checks examinee in and assigns seat at computer. All testing instructions are given on the computer.</td>
</tr>
<tr>
<td>Frequency of Test</td>
<td>Many times per year distributed over January, April, May, June, July, August, and September.</td>
</tr>
<tr>
<td>Format</td>
<td>Exclusively computer-based. NOT an adaptive test.</td>
</tr>
<tr>
<td>Length of Test Day</td>
<td>7.5 hours</td>
</tr>
<tr>
<td>Breaks</td>
<td>Optional 10-minute breaks between sections, with a 30-minute break for lunch.</td>
</tr>
</tbody>
</table>
| Section Names      | 1. Chemical and Physical Foundations of Biological Systems (Chem/Phys)  
2. Critical Analysis and Reasoning Skills (CARS)  
3. Biological and Biochemical Foundations of Living Systems (Bio/Biochem)  
4. Psychological, Social, and Biological Foundations of Behavior (Psych/Soc)  |
| Number of Questions and Timing | 59 Chem/Phys questions, 95 minutes  
53 CARS questions, 90 minutes  
59 Bio/Biochem questions, 95 minutes  
59 Psych/Soc questions, 95 minutes  |
| Scoring            | Test is scaled. Several forms per administration.                                                                                                                                           |
| Allowed/Not Allowed| No timers/watches. Noise reduction headphones available. Unopened package of foam earplugs is allowed. Scratch paper and pencils given at start of test and taken at end of test. Locker or secure area provided for personal items. |
| Results: Timing and Delivery | Approximately 30 days. Electronic scores only, available online through AAMC login. Examinees can print official score reports. |
| Maximum Number of Retakes | As of April 2015, the MCAT can be taken a maximum of three times in one year, four times over two years, and seven times over the lifetime of the examinee. An examinee can be registered for only one date at a time. |
Registration
Registration for the exam is completed online at www.aamc.org/students/applying/mcat/reserving. The AAMC opens registration for a given test date at least two months in advance of the date, often earlier. It’s a good idea to register well in advance of your desired test date to make sure that you get a seat.

Sections
There are four sections on the MCAT exam: Chemical and Physical Foundations of Biological Systems (Chem/Phys), Critical Analysis and Reasoning Skills (CARS), Biological and Biochemical Foundations of Living Systems (Bio/Biochem), and Psychological, Social, and Biological Foundations of Behavior (Psych/Soc). All sections consist of multiple-choice questions.

<table>
<thead>
<tr>
<th>Section</th>
<th>Concepts Tested</th>
<th>Number of Questions and Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical and Physical Foundations of Biological Systems</td>
<td>Basic concepts in chemistry and physics, including biochemistry; scientific inquiry; reasoning; research and statistics skills.</td>
<td>59 questions in 95 minutes</td>
</tr>
<tr>
<td>Critical Analysis and Reasoning Skills</td>
<td>Critical analysis of information drawn from a wide range of social science and humanities disciplines.</td>
<td>53 questions in 90 minutes</td>
</tr>
<tr>
<td>Biological and Biochemical Foundations of Living Systems</td>
<td>Basic concepts in biology and biochemistry, scientific inquiry, reasoning, research and statistics skills.</td>
<td>59 questions in 95 minutes</td>
</tr>
<tr>
<td>Psychological, Social, and Biological Foundations of Behavior</td>
<td>Basic concepts in psychology, sociology, and biology, research methods and statistics.</td>
<td>59 questions in 95 minutes</td>
</tr>
</tbody>
</table>

Most questions on the MCAT (44 in the science sections, all 53 in the CARS section) are passage-based; the science sections have 10 passages each and the CARS section has 9. A passage consists of a few paragraphs of information on which several following questions are based. In the science sections, passages often include equations or reactions, tables, graphs, figures, and experiments to analyze. CARS passages come from literature in social sciences, humanities, ethics, philosophy, cultural studies, and population health, and do not test content knowledge in any way.

Some questions in the science sections are freestanding questions (FSQs). These questions are independent of any passage information and appear in several groups of about four to five questions, interspersed throughout the passages. 15 of the questions in the science sections are freestanding, and the remainder are passage-based.
Each section on the MCAT is separated by either a 10-minute break or a 30-minute lunch break.

<table>
<thead>
<tr>
<th>Section</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Center Check-In</td>
<td>Variable, can take up to 40 minutes if center is busy.</td>
</tr>
<tr>
<td>Tutorial</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Chemical and Physical Foundations of Biological Systems</td>
<td>95 minutes</td>
</tr>
<tr>
<td>Break</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Critical Analysis and Reasoning Skills</td>
<td>90 minutes</td>
</tr>
<tr>
<td>Lunch Break</td>
<td>30 minutes</td>
</tr>
<tr>
<td>Biological and Biochemical Foundations of Living Systems</td>
<td>95 minutes</td>
</tr>
<tr>
<td>Break</td>
<td>10 minutes</td>
</tr>
<tr>
<td>Psychological, Social, and Biological Foundations of Behavior</td>
<td>95 minutes</td>
</tr>
<tr>
<td>Void Option</td>
<td>5 minutes</td>
</tr>
<tr>
<td>Survey</td>
<td>5 minutes</td>
</tr>
</tbody>
</table>

The survey includes questions about your satisfaction with the overall MCAT experience, including registration, check-in, etc., as well as questions about how you prepared for the test.

**Scoring**

The MCAT is a scaled exam, meaning that your raw score will be converted into a scaled score that takes into account the difficulty of the questions. There is no guessing penalty. All sections are scored from 118–132, with a total scaled score range of 472–528. Because different versions of the test have varying levels of difficulty, the scale will be different from one exam to the next. Thus, there is no “magic number” of questions to get right in order to get a particular score. Plus, some of the questions on the test are considered “experimental” and do not count toward your score; they are just there to be evaluated for possible future inclusion in a test.

At the end of the test (after you complete the Psychological, Social, and Biological Foundations of Behavior section), you will be asked to choose one of the following two options, “I wish to have my MCAT exam scored” or “I wish to VOID my MCAT exam.” You have five minutes to make a decision, and if you do not select one of the options in that time, the test will automatically be scored. If you choose the VOID option, your test will not be scored (you will not now, or ever, get a numerical score for this test), medical schools will not know you took the test, and no refunds will be granted. You cannot “unvoid” your scores at a later time.

So, what’s a good score? The AAMC is centering the scale at 500 (i.e., 500 will be the 50th percentile), and recommends that application committees consider applicants near the center of the range. To be on the safe side, aim for a total score of around 510. Remember that if your GPA is on the low side, you’ll need higher MCAT scores to compensate, and if you have a strong GPA, you can get away with lower MCAT scores. But the reality is that your chances of acceptance depend on a lot more than just your MCAT scores. It’s a combination of your GPA, your MCAT scores, your undergraduate coursework, letters of recommendation, experience related to the medical field (such as volunteer work or research), extracurricular activities, your personal statement, etc. Medical schools are looking for a complete package, not just good scores and a good GPA.
GENERAL LAYOUT AND TEST-TAKING STRATEGIES

Layout of the Test
In each section of the test, the computer screen is divided vertically, with the passage on the left and the range of questions for that passage indicated above (e.g., “Passage 1 Questions 1–5”). The scroll bar for the passage text appears in the middle of the screen. Each question appears on the right, and you need to click “Next” to move to each subsequent question.

In the science sections, the freestanding questions are found in groups of 4–5, interspersed with the passages. The screen is still divided vertically; on the left is the statement “Questions [X–XX] do not refer to a passage and are independent of each other,” and each question appears on the right as described above.

CBT Tools
There are a number of tools available on the test, including highlighting, strike-outs, the Mark button, the Review button, the Periodic Table button, and of course, scratch paper. The following is a brief description of each tool.

1) **Highlighting**: This is done in the passage text (including table entries and some equations, but excluding figures and molecular structures) and in the question stems by left-clicking and dragging the mouse across the words you wish to highlight; the selected words will then be highlighted in blue. When you release the mouse, a highlighting icon will appear; clicking on the icon will highlight the selected text in yellow. To remove the highlighting, left-click on the highlighted text.

2) **Strike-outs**: Right-clicking on an answer choice causes the entire text of that choice to be crossed out. The strike-out can be removed by right-clicking again. Left-clicking selects an answer choice; note that an answer choice that is selected cannot be struck out. When you strike out a figure or molecular structure, instead of being crossed out, the image turns grey.

3) **Mark button**: This allows you to flag the question for later review. When clicked, the flag on the “Mark” button turns red and says “Marked.”

4) **Review button**: Clicking this button brings up a new screen showing all questions and their status (either “completed,” “incomplete,” or “marked”). You can choose to: “review all,” “review incomplete,” or “review marked.” You can also double-click any question number to quickly return to that specific question. You can only review questions in the section of the MCAT you are currently taking, but the Review button can be clicked at any time during the allotted time for that section; you do NOT have to wait until the end of the section to click it.

5) **Periodic Table button**: Clicking this button will open a periodic table. Note that the periodic table is large, however it can be resized to see the questions and a portion of the periodic table at the same time.

6) **Scratch paper**: You will be given four pages (8 faces) of scratch paper at the start of the test. You can ask for more at any point during the test, and your first set of paper will be collected before you receive fresh paper. Scratch paper is only useful if it is kept organized; do not give in to the tendency to write on the first available open space! Good organization will be very helpful when/if you wish to review a question. Indicate the passage number and the range of questions for that passage in a box near the top of your scratch work, and indicate the question you are working on in a circle to the left of the notes for that question. Draw a line under your scratch work when you change passages to keep the work separate. Do not erase or scribble over any previous work. If you do not think it is correct, draw one line through the work and start again. You may have already done some useful work without realizing it.
General Strategy for the Science Sections

Passages vs. FSQs in the Science Sections: What to Start With
Since the questions are displayed on separate screens, it is awkward and time consuming to click through all of the questions up front to find the FSQs. Therefore, go through the section on a first pass and decide whether to do the passage now or to save it for later, basing your decision on the passage text and the first question. Tackle the FSQs as you come upon them. More details are below.

Here is an outline of the procedure:

1) For each passage, write a heading on your scratch paper with the passage number, the general topic, and its range of questions (e.g., “Passage 1, thermodynamics, Q 1–5” or “Passage 2, enzymes, Q 6–9”). The passage numbers do not currently appear in the Review screen, thus having the question numbers on your scratch paper will allow you to move through the section more efficiently.

2) Skim the text and rank the passage. If a passage is a “Now,” complete it before moving on to the next passage (also see Attacking the Questions below). If it is a “Later” passage, first write “SKIPPED” in block letters under the passage heading on your scratch paper and leave room for your work when you come back to complete that passage. (Note that the specific passages you skip will be unique to you; in the Bio/Biochem section, you might choose to do all Biology passages first, then come back for Biochemistry. Or in Chem/Phys you might choose to skip experiment-based or analytical passage. Know ahead of time what type of passage you are going to skip and follow your plan.)

3) Next, click on the “Review” button at the bottom to get to the review screen. Double-click on the first question of the next passage; you’ll be able to identify it because you know the range of questions from the passage you just skipped. This will take you to the next passage, where you will repeat steps 1–3.

4) Once you have completed the “Now” passages, go to the review screen and double-click the first question for the first passage you skipped. Answer the questions, and continue going back to the review screen and repeating this procedure for other passages you have skipped.

Attacking the Questions
As you work through the questions, if you encounter a particularly lengthy question, or a question that requires a lot of analysis, you may choose to skip it. This is a wise strategy because it ensures you will tackle all the easier questions first, the ones you are more likely to get right. If you choose to skip the question (or if you attempt it but get stuck), write down the question number on your scratch paper, click the Mark button to flag the question in the Review screen, and move on to the next question. At the end of the passage, click back through the set of questions to complete any that you skipped over the first time through, and make sure that you have filled in an answer for every question.
General Strategy for the CARS Section

Ranking and Ordering the Passages: What to Start With

**Ranking:** Since the questions are displayed on separate screens, it is awkward and time consuming to click through all of the questions before making a “Now,” “Later,” or “Killer” decision. Therefore, rank the passage and decide whether or not to do it on the first pass through the section based on the passage text, skimming the first 2–3 sentences.

**Ordering:** Because of the additional clicking through screens (or, use of the Review screen) that is required to navigate through the section, the “Two-Pass” system (completing the “Now” passages as you find them) is likely to be your most efficient approach. However, if you find that you are continuously making a lot of bad ranking decisions, it is still valid to experiment with the “Three-Pass” approach (ranking all nine passages up front before attempting your first “Now” passage).

Here is an outline of the basic Ranking and Ordering procedure to follow.

1) For each passage, write a heading on your scratch paper with the passage number and its range of questions (e.g., “Passage 1 Q 1–7”). The passage numbers do not currently appear in the Review screen, thus having the question numbers on your scratch paper will allow you to move through the section more efficiently.

2) Skim the first 2–3 sentences and rank the passage. If the passage is a “Now,” complete it before moving on to the next. If it is a “Later” or “Killer,” first write either “Later” or “Killer” and “SKIPPED” in block letters under the passage heading on your scratch paper and leave room for your work if you decide to come back and complete that passage. Then click through each question, marking each one and filling in random guesses, until you get to the next passage.

3) Once you have completed the “Now” passages, come back for your second pass and complete the “Later” passages, leaving your random guesses in place for any “Killer” passages that you choose not to complete. Go to the Review screen and use your scratch paper notes on the question numbers; double-click on the number of the first question for that passage to go back to that question and proceed from there. Alternatively, if you have consistently marked all the questions for passages you skipped in your first pass you can use “Review Marked” from the Review screen to find and complete your “Later” passages.

4) Regardless of how you choose to find your second pass passages, unmark each question after you complete it, so that you can continue to rely on the Review screen (and the “Review Marked” function) to identify questions that you have not yet attempted.

Previewing the Questions

The formatting and functioning of the tools facilitates effective previewing. Having each question on a separate screen will encourage you to really focus on that question. Even more importantly, you can now highlight in the question stem (but not in the answer choices).

Here is the basic procedure for previewing the questions:

1) Start with the first question, and if it has lead words referencing passage content, highlight them. You may also choose to jot them down on your scratch paper. Once you reach and preview the last question for the set on that passage, THEN stay on that screen and work the passage (your highlighting appears and stays on every passage screen, and persists through the whole 90 minutes).
2) Once you have worked the passage and defined the Bottom Line, work backward from the last question to the first. If you skip over any questions as you go (see Attacking the Questions below), write down the question number on your scratch paper. Then click forward through the set of questions, completing any that you skipped over the first time through. Once you reach and complete the last question for that passage, clicking “Next” will send you to the first question of the next passage. Working the questions from last to first the first time through the set will eliminate the need to click back through multiple screens to get to the first question immediately after previewing, and will also make it easier and more efficient to do the hardest questions last (see Attacking the Questions below).

Attacking the Questions
The question types and the procedure for actually attacking each type will be discussed later. However, it is still important not to attempt the hardest questions first (potentially getting stuck, wasting time, and discouraging yourself).

So, as you work the questions from last to first (see Previewing the Questions above), if you encounter a particularly difficult and/or lengthy question (or if you attempt a question but get stuck) write down the question number on your scratch paper (you may also choose to mark it) and move on backward to the next question. Then click forward through the set and complete any that you skipped over the first time through the set, unmarking any questions that you marked that first time through and making sure that you have filled in an answer for every question.

Pacing Strategy for the MCAT
Since the MCAT is a timed test, you must keep an eye on the timer and adjust your pacing as necessary. It would be terrible to run out of time at the end only to discover that the last few questions could have been easily answered in just a few seconds each.

In the science sections you will have about one minute and thirty-five seconds (1:35) per question, and in the CARS section you will have about one minute and forty seconds (1:40) per question (not taking into account time reading the passage before answering the questions).

<table>
<thead>
<tr>
<th>Section</th>
<th># of Questions in Passage</th>
<th>Approximate Time (including reading the passage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chem/Phys, Bio/Biochem, and Psych/Soc</td>
<td>4</td>
<td>6.5 minutes</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8 minutes</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>9.5 minutes</td>
</tr>
<tr>
<td>CARS</td>
<td>5</td>
<td>8.5 minutes</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>10 minutes</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>11.5 minutes</td>
</tr>
</tbody>
</table>
When starting a passage in the science sections, make note of how much time you will allot for it and the starting time on the timer. Jot down on your scratch paper what the timer should say at the end of the passage. Then just keep an eye on it as you work through the questions. If you are near the end of the time for that passage, guess on any remaining questions, make some notes on your scratch paper, Mark the questions, and move on. Come back to those questions if you have time.

For the CARS section, keep in mind that many people will maximize their score by not trying to complete every question or every passage in the section. A good strategy for test takers who cannot achieve a high level of accuracy on all nine passages is to randomly guess on at least one passage in the section, and spend your time getting a high percentage of the other questions right. To complete all nine CARS passages, you have about ten minutes per passage. To complete eight of the nine, you have about 11 minutes per passage.

To help maximize your number of correct answer choices in any section, do the questions and passages within that section in the order you want to do them in. See General Strategy.

**Process of Elimination**

Process of Elimination (POE) is probably the most useful technique you have to tackle MCAT questions. Since there is no guessing penalty, POE allows you to increase your probability of choosing the correct answer by eliminating those you are sure are wrong.

1) Strike out any choices that you are sure are incorrect or that do not address the issue raised in the question.
2) Jot down some notes to help clarify your thoughts if you return to the question.
3) Use the “Mark” button to flag the question for review. (Note, however, that in the CARS section, you generally should not be returning to rethink questions once you have moved on to a new passage.)
4) Do not leave it blank! For the sciences, if you are not sure and you have already spent more than 60 seconds on that question, just pick one of the remaining choices. If you have time to review it at the end, you can always debate the remaining choices based on your previous notes. For CARS, if you have been through the choices two or three times, have re-read the question stem and gone back to the passage and you are still stuck, move on. Do the remaining questions for that passage, take one more look at the question you were stuck on, then pick an answer and move on for good.
5) Special Note: If three of the four answer choices have been eliminated, the remaining choice must be the correct answer. Don’t waste time pondering why it is correct, just click it and move on. The MCAT doesn’t care if you truly understand why it’s the right answer, only that you have the right answer selected.
6) More subject-specific information on techniques will be presented in the next chapter.

**Guessing**

Remember, there is NO guessing penalty on the MCAT. NEVER leave a question blank!
McAT Biology Review

Question Types

In the science sections of the MCAT, the questions fall into one of three main categories.

1) Memory questions: These questions can be answered directly from prior knowledge and represent about 25 percent of the total number of questions.

2) Explicit questions: These questions are those for which the answer is explicitly stated in the passage. To answer them correctly, for example, may just require finding a definition, reading a graph, or making a simple connection. Explicit questions represent about 35 percent of the total number of questions.

3) Implicit questions: These questions require you to apply knowledge to a new situation; the answer is typically implied by the information in the passage. These questions often start “if...then...” (For example, “If we modify the experiment in the passage like this, then what result would we expect?”) Implicit style questions make up about 40 percent of the total number of questions.

In the CARS section, the questions fall into four main categories:

1) Specific questions: These either ask you for facts from the passage (Retrieval questions) or require you to deduce what is most likely to be true based on the passage (Inference questions).

2) General questions: These ask you to summarize themes (Main Idea and Primary Purpose questions) or evaluate an author’s opinion (Tone/Attitude questions).

3) Reasoning questions: These ask you to describe the purpose of, or the support provided for, a statement made in the passage (Structure questions) or to judge how well the author supports his or her argument (Evaluate questions).

4) Application questions: These ask you to apply new information from either the question stem itself (New Information questions) or from the answer choices (Strengthen, Weaken, and Analogy questions) to the passage.

More detail on question types and strategies can be found in Chapter 2.

Testing Tips

Before Test Day

- Take a trip to the test center at least a day or two before your actual test date so that you can easily find the building and room on test day. This will also allow you to gauge traffic and see if you need money for parking or anything like that. Knowing this type of information ahead of time will greatly reduce your stress on the day of your test.
- During the week before the test, adjust your sleeping schedule so that you are going to bed and getting up in the morning at the same times as on the day before and morning of the MCAT. Prioritize getting a reasonable amount of sleep during the last few nights before the test.
- Don’t do any heavy studying the day before the test. This is not a test you can cram for! Your goal at this point is to rest and relax so that you can go into test day in a good physical and mental condition.
• Eat well. Try to avoid excessive caffeine and sugar. Ideally, in the weeks leading up to the actual test you should experiment a little bit with foods and practice tests to see which foods give you the most endurance. Aim for steady blood sugar levels during the test: sports drinks, peanut-butter crackers, trail mix, etc. make good snacks for your breaks and lunch.

General Test Day Info and Tips
• On the day of the test, arrive at the test center at least a half hour prior to the start time of your test.
• Examinees will be checked in to the center in the order in which they arrive.
• You will be assigned a locker or secure area in which to put your personal items. Textbooks and study notes are not allowed, so there is no need to bring them with you to the test center.
• Your ID will be checked, a digital image of your fingerprint will be taken, and you will be asked to sign in.
• You will be given scratch paper and a couple of pencils, and the test center administrator will take you to the computer on which you will complete the test. You may not choose a computer; you must use the computer assigned to you.
• Nothing, not even your watch, is allowed at the computer station except your photo ID, your locker key (if provided), and a factory sealed packet of ear plugs.
• If you choose to leave the testing room at the breaks, you will have your fingerprint checked again, and you will have to sign in and out.
• You are allowed to access the items in your locker, except for notes and cell phones. (Check your test center’s policy on cell phones ahead of time; some centers do not even allow them to be kept in your locker.)
• Don’t forget to bring the snack foods and lunch you experimented with during your practice tests.
• At the end of the test, the test administrator will collect your scratch paper and shred it.
• Definitely take the breaks! Get up and walk around. It’s a good way to clear your head between sections and get the blood (and oxygen!) flowing to your brain.
• Ask for new scratch paper at the breaks if you use it all up.
Chapter 2
Biology Strategy for the MCAT
2.1 SCIENCE SECTIONS OVERVIEW
There are three science sections on the MCAT:

- Chemical and Physical Foundations of Biological Systems
- Biological and Biochemical Foundations of Living System
- Psychological, Social, and Biological Foundations of Behavior

The Chemical and Physical Foundations of Biological Systems section (Chem/Phys) is the first section on the test. It includes questions from General Chemistry (about 30%), Physics (about 25%), Organic Chemistry (about 15%), Biochemistry (about 25%), and Biology (about 5%). Further, the questions often test chemical and physical concepts within a biological setting: for example, pressure and fluid flow in blood vessels. A solid grasp of math fundamentals is required (arithmetic, algebra, graphs, trigonometry, vectors, proportions, and logarithms); however, there are no calculus-based questions.

The Biological and Biochemical Foundations of Living Systems section (Bio/Biochem) is the third section on the test. Approximately 65% of the questions in this section come from biology, approximately 25% come from biochemistry, and approximately 10% come from Organic and General Chemistry. Math calculations are generally not required on this section of the test; however, a basic understanding of statistics as used in biological research is helpful.

The Psychological, Social, and Biological Foundations of Behavior section (Psych/Soc) is the fourth and final section on the test. About 60% of the questions will be drawn from Psychology, about 30% from Sociology, and about 10% from Biology. As with the Bio/Biochem section, calculations are generally not required, however a basic understanding of statistics as used in research is helpful.

Most of the questions in the science sections (44 of the 59) are passage-based, and each section has ten passages. Passages consist of a few paragraphs of information and include equations, reactions, graphs, figures, tables, experiments, and data. Four to six questions will be associated with each passage.

The remaining 25% of the questions (15 of 59) in each science section are freestanding questions (FSQs). These questions appear in approximately four groups interspersed between the passages. Each group contains four to five questions.

95 minutes are allotted to each of the science sections. This breaks down to approximately one minute and 35 seconds per question.
2.2 SCIENCE PASSAGE TYPES

The passages in the science sections fall into one of three main categories: Information and/or Situation Presentation, Experiment/Research Presentation, or Persuasive (or Scientific) Reasoning.

Information and/or Situation Presentation

These passages either present straightforward scientific information or they describe a particular event or occurrence. Generally, questions associated with these passages test basic science facts or ask you to predict outcomes given new variables or new information. Here is an example of an Information/Situation Presentation passage:

Figure 1 shows a portion of the inner mechanism of a typical home smoke detector. It consists of a pair of capacitor plates which are charged by a 9-volt battery (not shown). The capacitor plates (electrodes) are connected to a sensor device, D; the resistor R denotes the internal resistance of the sensor. Normally, air acts as an insulator and no current would flow in the circuit shown. However, inside the smoke detector is a small sample of an artificially produced radioactive element, americium-241, which decays primarily by emitting alpha particles, with a half-life of approximately 430 years. The daughter nucleus of the decay has a half-life in excess of two million years and therefore poses virtually no biohazard.

![Diagram of a smoke detector mechanism](image)

**Figure 1** Smoke detector mechanism

The decay products (alpha particles and gamma rays) from the 241Am sample ionize air molecules between the plates and thus provide a conducting pathway which allows current to flow in the circuit shown in Figure 1. A steady-state current is quickly established and remains as long as the battery continues to maintain a 9-volt potential difference between its terminals. However, if smoke particles enter the space between the capacitor plates and thereby interrupt the flow, the current is reduced, and the sensor responds to this change by triggering
the alarm. (Furthermore, as the battery starts to "die out," the resulting drop in current is also detected to alert the homeowner to replace the battery.)

\[ C = \varepsilon_0 \frac{A}{d} \]

Equation 1

where \( \varepsilon_0 \) is the universal permittivity constant, equal to \( 8.85 \times 10^{-12} \text{ C}^2/(\text{N m}^2) \). Since the area \( A \) of each capacitor plate in the smoke detector is 20 cm\(^2\) and the plates are separated by a distance \( d \) of 5 mm, the capacitance is \( 3.5 \times 10^{-12} \text{ F} = 3.5 \text{ pF} \).

**Experiment/Research Presentation**

These passages present the details of experiments and research procedures. They often include data tables and graphs. Generally, questions associated with these passages ask you to interpret data, draw conclusions, and make inferences. Here is an example of an Experiment/Research Presentation passage:

The development of sexual characteristics depends upon various factors, the most important of which are hormonal control, environmental stimuli, and the genetic makeup of the individual. The hormones that contribute to the development include the steroid hormones estrogen, progesterone, and testosterone, as well as the pituitary hormones FSH (follicle-stimulating hormone) and LH (luteinizing hormone).

To study the mechanism by which estrogen exerts its effects, a researcher performed the following experiments using cell culture assays.

**Experiment 1:**

Human embryonic placental mesenchyme (HEPM) cells were grown for 48 hours in Dulbecco’s Modified Eagle Medium (DMEM), with media change every 12 hours. Upon confluent growth, cells were exposed to a 10 mg per mL solution of green fluorescent-labeled estrogen for 1 hour. Cells were rinsed with DMEM and observed under confocal fluorescent microscopy.

**Experiment 2:**

HEPM cells were grown to confluence as in Experiment 1. Cells were exposed to Pesticide A for 1 hour, followed by the 10 mg/mL solution of labeled estrogen, rinsed as in Experiment 1, and observed under confocal fluorescent microscopy.
Experiment 3:

Experiment 1 was repeated with Chinese Hamster Ovary (CHO) cells instead of HEPM cells.

Experiment 4:

CHO cells injected with cytoplasmic extracts of HEPM cells were grown to confluence, exposed to the 10 mg/mL solution of labeled estrogen for 1 hour, and observed under confocal fluorescent microscopy.

The results of these experiments are given in Table 1.

Table 1 Detection of Estrogen (+ indicates presence of Estrogen)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Cytoplasm</th>
<th>Nucleus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

After observing the cells in each experiment, the researcher bathed the cells in a solution containing 10 mg per mL of a red fluorescent probe that binds specifically to the estrogen receptor only when its active site is occupied. After 1 hour, the cells were rinsed with DMEM and observed under confocal fluorescent microscopy. The results are presented in Table 2.

The researcher also repeated Experiment 2 using Pesticide B, an estrogen analog, instead of Pesticide A. Results from other researchers had shown that Pesticide B binds to the active site of the cytosolic estrogen receptor (with an affinity 10,000 times greater than that of estrogen) and causes increased transcription of mRNA.

Table 2 Observed Fluorescence and Estrogen Effects (G = green, R = red)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Cytoplasm</th>
<th>Nucleus</th>
<th>Estrogen effects observed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G only</td>
<td>G and R</td>
<td>G and R</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>G only</td>
<td>G only</td>
<td>G only</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>G only</td>
<td>G only</td>
<td>G only</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>G only</td>
<td>G and R</td>
<td>G and R</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Based on these results, the researcher determined that estrogen had no effect when not bound to a cytosolic, estrogen-specific receptor.
2.2

**Persuasive (Scientific) Reasoning**

These passages typically present a scientific phenomenon, along with a hypothesis that explains the phenomenon, and may include counterarguments as well. Questions associated with these passages ask you to evaluate the hypothesis or arguments. Persuasive Reasoning passages in the science sections of the MCAT tend to be less common than Information Presentation or Experiment-based passages. Here is an example of a Persuasive Reasoning passage:

Two theoretical chemists attempted to explain the observed trends of acidity by applying two interpretations of molecular orbital theory. Consider the \( pK_a \) values of some common acids listed along with the conjugate base:

<table>
<thead>
<tr>
<th>acid</th>
<th>( pK_a )</th>
<th>conjugate base</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{H}_2\text{SO}_4 )</td>
<td>&lt; 0</td>
<td>( \text{HSO}_4^- )</td>
</tr>
<tr>
<td>( \text{H}_2\text{CrO}_4 )</td>
<td>5.0</td>
<td>( \text{HCrO}_4^- )</td>
</tr>
<tr>
<td>( \text{H}_2\text{PO}_4 )</td>
<td>2.1</td>
<td>( \text{H}_2\text{PO}_4^- )</td>
</tr>
<tr>
<td>HF</td>
<td>3.9</td>
<td>( \text{F}^- )</td>
</tr>
<tr>
<td>( \text{HOCl} )</td>
<td>7.8</td>
<td>( \text{ClO}^- )</td>
</tr>
<tr>
<td>HCN</td>
<td>9.5</td>
<td>( \text{CN}^- )</td>
</tr>
<tr>
<td>( \text{HIO}_3 )</td>
<td>1.2</td>
<td>( \text{IO}_3^- )</td>
</tr>
</tbody>
</table>

Recall that acids with a \( pK_a < 0 \) are called strong acids, and those with a \( pK_a > 0 \) are called weak acids. The arguments of the chemists are given below.

**Chemist #1:**

“The acidity of a compound is proportional to the polarization of the H—X bond, where X is some nonmetal element. Complex acids, such as \( \text{H}_2\text{SO}_4 \), \( \text{HClO}_4 \), and \( \text{HNO}_3 \) are strong acids because the H—O bonding electrons are strongly drawn towards the oxygen. It is generally true that a covalent bond weakens as its polarization increases. Therefore, one can conclude that the strength of an acid is proportional to the number of electronegative atoms in that acid.”

**Chemist #2:**

“The acidity of a compound is proportional to the number of stable resonance structures of that acid’s conjugate base. \( \text{H}_2\text{SO}_4 \), \( \text{HClO}_4 \), and \( \text{HNO}_3 \) are all strong acids because their respective conjugate bases exhibit a high degree of resonance stabilization.”
Mapping a Passage
“Mapping a passage” refers to the combination of on-screen highlighting and scratch paper notes that you take while working through a passage. Typically, good things to highlight include the overall topic of a paragraph, unfamiliar terms, italicized terms, unusual terms, numerical values, hypothesis, and results. Scratch paper notes can be used to summarize the paragraphs and to jot down important facts and connections that are made when reading the passage. More details on passage mapping will be presented in Section 2.5.

2.3 SCIENCE QUESTION TYPES
Questions in the science sections are generally one of three main types: Memory, Explicit, or Implicit.

Memory Questions
These questions can be answered directly from prior knowledge, with no need to reference the passage or question text. Memory questions represent approximately 25 percent of the science questions on the MCAT. Usually, Memory questions are found as FSQs, but they can also be tucked into a passage. Here’s an example of a Memory question:

Which of the following acetylation conditions will convert diethylamine into an amide at the fastest rate?

A) Acetic acid / HCl
B) Acetic anhydride
C) Acetyl chloride
D) Ethyl acetate
Explicit Questions
Explicit questions can be answered primarily with information from the passage, along with prior knowledge. They may require data retrieval, graph analysis, or making a simple connection. Explicit questions make up approximately 35–40 percent of the science questions on the MCAT; here’s an example (taken from the Information/Situation Presentation passage):

The sensor device $D$ shown in Figure 1 performs its function by acting as:

A) an ohmmeter.
B) a voltmeter.
C) a potentiometer.
D) an ammeter.

Implicit Questions
These questions require you to take information from the passage, combine it with your prior knowledge, apply it to a new situation, and come to some logical conclusion. They typically require more complex connections than do Explicit questions, and they may also require data retrieval, graph analysis, etc. Implicit questions usually require a solid understanding of the passage information. They make up approximately 35–40 percent of the science questions on the MCAT; here’s an example (taken from the Experiment/Research Presentation passage):

If Experiment 2 were repeated, but this time exposing the cells first to Pesticide A and then to Pesticide B before exposing them to the green fluorescent-labeled estrogen and the red fluorescent probe, which of the following statements will most likely be true?

A) Pesticide A and Pesticide B bind to the same site on the estrogen receptor.
B) Estrogen effects would be observed.
C) Only green fluorescence would be observed.
D) Both green and red fluorescence would be observed.

The Rod of Asclepius
You may notice this Rod of Asclepius icon as you read through the book. In Greek mythology, the Rod of Asclepius is associated with healing and medicine; the symbol continues to be used today to represent medicine and healthcare. You won’t see this on the actual MCAT, but we’ve used it here to call attention to medically related examples and questions.
2.4 BIOLOGY ON THE MCAT

Biology is by far the most information-dense section on the MCAT. MCAT Biology topics span six different semester-length courses (molecular biology, cell biology, microbiology, genetics, anatomy, and physiology). Further, the application of this material is potentially vast; passages can discuss anything from the details of some viral life cycle to the complexities of genetic studies, to the nuances of an unusual disease. Fortunately, biology is the subject that MCAT students typically find the most interesting, and the one they have the most background in. People who want to go to medical school have an inherent interest in biology; thus this subject, although vast, seems more manageable than all the others on the MCAT.

The science sections of the MCAT have 10 passages and 15 freestanding questions (FSQs). The Biological and Biochemical Foundations of Living Systems section (Bio/Biochem) is primarily biology (65%) and biochemistry (25%). The remaining 10% are General and Organic Chemistry questions. Further, Biology questions can show up in the Psychological, Social, and Biological Foundations of Behavior section (about 10%) and in the Chemical and Physical Foundations of Biological Systems section (about 5%). Note also that about 25% of the Chem/Phys section is Biochemistry, and frequently the passages and questions are biology-based.

2.5 TACKLING A BIOLOGY PASSAGE

Generally speaking, time is not an issue in the Bio/Biochem section of the MCAT. Because students have a stronger background in biology than in other subjects, the passages seem more understandable; in fact, readers sometimes find themselves getting caught up and interested in the passage. Often, students report having about 5 to 10 minutes “left over” after completing the section. This means that an additional minute or so can potentially be spent on each passage, thinking and understanding.

Passage Types as They Apply to Biology

Experiment/Research Presentation: Biology
This is the most common type of Biology passage. It typically presents the details behind an experiment along with data tables, graphs, and figures. Often these are the most difficult passages to deal with because they require an understanding of the reasoning behind the experiment, the logic to each step, and the ability to analyze the results and form conclusions. A basic understanding of biometry (basic statistics as they apply to biology and biology research) is necessary.

Information/Situation Presentation: Biology
This is the second most common type of Biology passage on the MCAT. These passages generally appear as one of two variants: either a basic concept with additional levels of detail included (for example, all the detail you ever wanted to know about the electron transport chain), or a novel concept with ties to
basic information (for example, a rare demyelinating disease). Either way, Biology passages are notorious for testing concepts in unusual contexts. The key to dealing with these passages is to, first, not become anxious about all the stuff you might not know, and second, figure out how the basics you do know apply to the new situation. For example, you might be presented with a passage that introduces hormones you never heard of or novel drugs to combat diseases you didn’t know existed. First, don’t panic. Second, look for how these new things fit into familiar categories: for example, “peptide vs. steroid” or “sympathetic antagonist.” Then answer the questions with these basics in mind.

That said, you have to know your basics. This will increase your confidence in answering freestanding questions, as well as increase the speed with which to find the information in the passage. The astute MCAT student will never waste time staring at a question thinking, “Should I know this?” Instead, because she has a solid understanding of the necessary core knowledge, she’ll say, “No, I am NOT expected to know this, and I am going to look for it in the passage.”

**Persuasive Reasoning: Biology**

This is the least common passage type in Biology. It typically describes some biological phenomenon and then offers one or more theories to explain it. Questions in Persuasive Reasoning passages ask you to determine support for one of the theories, or present new evidence and ask which theory is now contradicted.

One last thought about Biology passages in general: Because the array of topics is so vast, Biology passages often pull questions from multiple areas of biology into a single, general topic. Consider, for example, a passage on renal function. Question topics could include basics about the kidney, transmembrane transport, autonomic control, blood pressure, hormones, biochemical energy needs, or a genetics question about a rare kidney disease.

**Reading a Biology Passage**

Although tempting, try not to get bogged down reading all the little details in a passage. Again, because most premeds have an inherent interest in biology and the mechanisms behind disease, it’s very easy to get lost in the science behind the passages. In spite of having that “extra” time, you don’t want to use it all up reading what isn’t necessary. Each passage type requires a slightly different style of reading.

Information/Situation Presentation passages require the least reading. These should be skimmed to get an idea of the location of information within the passage. These passages include a fair amount of detail that you might not need, so save the reading of these details until a question comes up about them. Then go back and read for the finer nuances.

Experiment/Research Presentation passages require the most reading. You are practically guaranteed to get questions that ask you about the details of the experiment, why a particular step was carried out, why the results are what they are, how to interpret the data, or how the results might change if a particular variable is altered. It’s worth spending a little more time reading to understand the experiment. However, because there will be a fair number of questions unrelated to the experiment, you might consider answering these first and then going back for the experiment details.
Persuasive Argument passages are somewhere in the middle. You can skim them for location of information, but you also want to spend a little time reading the details of and thinking about the arguments presented. It is extremely likely that you will be asked a question about them.

**Advanced Reading Skills**

To improve your ability to read and glean information from a passage, you need to practice. Be critical when you read the content; watch for vague areas or holes in the passage that aren’t explained clearly. Remember that information about new topics will be woven throughout the passage; you may need to piece together information from several paragraphs and a figure to get the whole picture.

After you’ve read, highlighted, and mapped a passage (more on this in a bit), stop and ask yourself the following questions:

- What was this passage about? What was the conclusion or main point?
- Was there a paragraph that was mostly background?
- Were there paragraphs or figures that seemed useless?
- What information was found in each paragraph? Why was that paragraph there?
- Are there any holes in the story?
- What extra information could I have pulled out of the passage? What inferences or conclusions could I make?
- If something unique was explained or mentioned, what might be its purpose?
- What am I not being told?
- Can I summarize the purpose and/or results of the experiment in a few sentences?
- Were there any comparisons in the passage?

This takes a while at first, but eventually it will become second nature and you’ll start doing it as you read the passage. If you have a study group you are working with, consider doing this as an exercise with your study partners. Take turns asking and answering the questions above. Having to explain something to someone else not only solidifies your own knowledge, but helps you see where you might be weak.

**Mapping a Biology Passage**

Mapping a Biology passage is a combination of highlighting and scratch paper notes that can help you organize and understand the passage information.

Resist the temptation to highlight everything! (Everyone has done this: You’re reading a biology textbook with a highlighter, and then look back and realize that the whole page is yellow!) Restrict your highlighting to a few things:
Scratch paper should be organized. Make sure the passage number appears at the top of your scratch paper notes. For each paragraph, note “P1,” “P2,” etc., on the scratch paper, and jot down a few notes about that paragraph. Try to translate biology jargon into your own words using everyday language (this is particularly useful for experiments). Also, make sure to note down simple relationships (e.g., the relationship between two variables).

Pay attention to equations, figures, and the like to see what type of information they deal with. Don’t spend a lot of time analyzing at this point, but do jot down on your scratch paper “Fig 1” and a brief summary of the data. Also, if you’ve discovered a list in the passage, note its topic and location down on your scratch paper.

Let’s take a look at how we might highlight and map a passage. Below is a passage on eye physiology.

The wall of the human eye is composed of three layers of tissue, an outer layer of tough connective tissue, a middle layer of darkly pigmented vascular tissue, and an inner layer of neural tissue. The outer layer is subdivided into the sclera, the white portion, and the cornea, the clear portion. The inner layer is more commonly known as the retina and contains several types of cells.

![Retina Structure](image)

**Figure 1** Retina Structure
The photoreceptors of the retina include rods and cones which respond to light under different circumstances. Rods are more sensitive to light but cannot distinguish color; cones are less sensitive to light overall, but can respond to different wavelengths. Response to light involves visual pigments, which in all cases consist of a light-absorbing molecule called retinal (derived from vitamin A) bound to a protein called opsin. The type of opsin in the visual pigment determines the wavelength specificity of the retinal. The specific visual pigment in rod cells is called rhodopsin.

![Figure 2 The Two Forms of Retinal](image)

In the absence of light, Na⁺ channels in the membranes of rod cells are kept open by cGMP. The conformational change in retinal upon light absorption causes changes in opsin as well; this triggers a pathway by which phosphodiesterase (PDE) is activated. Active PDE converts cGMP to GMP, causing it to dissociate from the Na⁺ channel and the channel to close. Until retinal regains its bent shape (helped by enzymes), the rod is unable to respond further to light.
Visual defects can be caused by abnormal visual pigments or by misshapen eyeballs; for example, myopia (nearsightedness) is due to an eyeball that is too long, causing light rays from distant objects to focus in front of the retina so the image appears blurry.

Analysis and Passage Map

This passage is an Information Presentation passage and starts out with a paragraph about the structure of the eye and its layers. This is primarily a background paragraph and can be skimmed quickly, with a few words highlighted. Figure 1 shows the detail of the retina.

The second paragraph goes into more detail about the photoreceptors, and specifically compares the functions of rods and cones. There are few more italicized terms; this paragraph is presenting information that is beyond what you are expected to know about the eye for the MCAT. Figure 2 shows the conversion between the two forms of retinal.

The third paragraph presents details about rod cells, and in particular points out a unique feature of rod cells: that their Na⁺ channels are typically open in light. On stimulation by light, they close. This is unusual behavior in the nervous system, since it is the opposite of what typically occurs. Figure 3 confirms this, as the cell in darkness appears to be resting at −40 mV, 30 mV more positive than typical neurons rest at.
The final paragraph is a brief description of visual defects. Like paragraph 1, it only needs to be skimmed briefly. Here’s what your passage map might look like:

- **P1** – 3 layers of eyeball, Fig 1 retina detail
- **P2** – photoreceptors
  - rods no color, more sensitive
  - cones less sensitive, respond to different colors
details on vis pigments. Fig 2 convert retinal
- **P3** – rod function. WEIRD Na+ channels open in dark, close in light depol in light, hyperpol in dark
- **P4** – visual defects

Let’s take a look at a different passage. Below is an Experiment/Research Presentation passage.

The development of sexual characteristics depends upon various factors, the most important of which are hormonal control, environmental stimuli, and the genetic makeup of the individual. The hormones that contribute to the development include the steroid hormones estrogen, progesterone, and testosterone, as well as the pituitary hormones FSH (follicle-stimulating hormone) and LH (luteinizing hormone).

To study the mechanism by which estrogen exerts its effects, a researcher performed the following experiments using cell culture assays.

**Experiment 1:**

Human embryonic placental mesenchyme (HEPM) cells were grown for 48 hours in Dulbecco’s Modified Eagle Medium (DMEM), with media change every 12 hours. Upon confluent growth, cells were exposed to a 10 mg per mL solution of green fluorescent-labeled estrogen for 1 hour. Cells were rinsed with DMEM and observed under confocal fluorescent microscopy.

**Experiment 2:**

HEPM cells were grown to confluence as in Experiment 1. Cells were exposed to Pesticide A for 1 hour, followed by the 10 mg/mL solution of labeled estrogen, rinsed as in Experiment 1, and observed under confocal fluorescent microscopy.

**Experiment 3:**

Experiment 1 was repeated with Chinese Hamster Ovary (CHO) cells instead of HEPM cells.
Experiment 4:

**CHO cells injected with cytoplasmic extracts of HEPM cells** were grown to confluence, exposed to the 10 mg/mL solution of labeled estrogen for 1 hour, and observed under confocal fluorescent microscopy.

The results of these experiments are given in Table 1.

**Table 1** Detection of Estrogen
(+ indicates presence of estrogen)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Cytoplasm</th>
<th>Nucleus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

After observing the cells in each experiment, the researcher bathed the cells in a solution containing 10 mg/mL of a red fluorescent probe that binds specifically to the estrogen receptor only when its active site is occupied. After 1 hour, the cells were rinsed with DMEM and observed under confocal fluorescent microscopy. The results are presented in Table 2.

The researcher also repeated Experiment 2 using Pesticide B, an estrogen analog, instead of Pesticide A. Results from other researchers had shown that Pesticide B binds to the active site of the cytosolic estrogen receptor (with an affinity 10,000 times greater than that of estrogen) and causes increased transcription of mRNA.

**Table 2** Observed Fluorescence and Estrogen Effects
(G = green, R = red)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Media</th>
<th>Cytoplasm</th>
<th>Nucleus</th>
<th>Estrogen effects observed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G only</td>
<td>G and R</td>
<td>G and R</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>G only</td>
<td>G only</td>
<td>G only</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>G only</td>
<td>G only</td>
<td>G only</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>G only</td>
<td>G and R</td>
<td>G and R</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Based on these results, the researcher determined that estrogen had no effect when not bound to a cytosolic, estrogen-specific receptor.
Analysis and Passage Map
This passage starts out with a very general background paragraph. Not much to do here, but it does tell us that estrogen is going to be the hormone of focus.

The next few paragraphs are short descriptions of four different experiments. These should be read to understand not only what’s happening in each experiment but also what the differences in the experiments are. Note this on your scratch paper.

Table 1 shows the results of the four experiments. It should jump out at you that estrogen is found everywhere; in other words, it is not restricted from any area of the cell.

After Table 1, the passage describes two modifications to the experiments. As with the original experiments, it’s worth taking a little time to read and understand what’s going on. The first big difference is that the researchers aren’t just looking for the presence of estrogen, but also want to know when it’s bound to its receptor. The second big difference is the testing of an estrogen analog, Pesticide B.

Table 2 shows the results of when estrogen is bound and when it isn’t. These results could be combined with the experiment description results on your map:

- P1 – hormones that contribute to development, estrogen
- E1 – HEPM cells exposed to estrogen, green + red = estrogen effects
- E2 – Pesticide A, green only, must inhibit binding of estrogen to recept.
- E3 – CHO cells, green only, no recept.
- E4 – CHO cells + HEPM cytoplasm, green + red, recept is in cytoplasm

Table 1 – estrogen is not restricted from anywhere in the cell
Further exp’ts – red probe for bound active site, and Pesticide B (estrogen analog with higher affinity)

One last thought about passages: Remember that, as with all sections on the MCAT, you can do the passages in the order you want to. There are no extra points for taking the test in order. Generally, passages will fall into one of four main subject groups:

- biochemistry
- other non-physiology
- physiology
- organic/general chemistry

Figure out which group you are most comfortable with, and do those passages first. (See Chapter 1 for general strategies for moving around in the sections efficiently.)
2.6 TACKLING THE QUESTIONS

Questions in the Biology section mimic the three typical questions of the science sections in general: Memory, Explicit, and Implicit.

Question Types as They Apply to Biology

Biology Memory Questions

Memory questions are exactly what they sound like: They test your knowledge of some specific fact or concept. While Memory questions are typically found as freestanding questions, they can also be tucked into a passage. These questions, aside from requiring memorization, do not generally cause problems for students because they are similar to the types of questions that appear on a typical college biology exam. Below is an example of a freestanding Memory question:

Regarding embryogenesis, which of the following sequence of events is in correct order?

A) Implantation—cleavage—gastrulation—neurulation—blastulation
B) Blastulation—implantation—cleavage—neurulation—gastrulation
C) Implantation—blastulation—gastrulation—cleavage—neurulation
D) Cleavage—blastulation—implantation—gastrulation—neurulation

The correct answer to the question above is choice D. Here’s another example. This question is from a passage:

The genital organs of the *guevedoche* that develop at puberty are derivatives of the mesodermal germ layer. Which of the following is/are also derivatives of the mesodermal germ layer?

I. Skeletal muscle
II. Liver
III. Kidney

A) I only
B) II only
C) I and III only
D) II and III only

Note that this question includes an additional, unnecessary sentence at the beginning, but it is a Memory question all the same. You don’t need to know anything about the *guevedoche* to answer the question, and the information in that first sentence does not help you in any way. The correct answer is choice C.
There is no specific “trick” to answering Memory questions; either you know the answer or you don’t.

If you find that you are missing a fair number of Memory questions, it is a sure sign that you don’t know the content well enough. Go back and review.

**Biology Explicit Questions**

True, pure Explicit questions are rare in the Biology section. A purely Explicit question can be answered only with information in the passage. Below is an example of a pure Explicit question taken from the eye passage above:

> The middle layer of the eyeball wall most likely contains:
>  
> A) bipolar cells.  
> B) photoreceptors.  
> C) blood vessels.  
> D) collagen fibers.

Referring back to the map for this passage, it indicates that information about the layers of the eyeball are in paragraph 1. It states that the middle layer is a “darkly pigmented vascular layer,” meaning that it contains blood vessels. The correct answer is choice C.

However, more often in the biology section, Explicit questions are more of a blend of Explicit and Memory; they require not only retrieval of passage information, but also recall of some relevant fact. They usually do not require a lot of analysis or connections. Here’s an example of the more common type of Explicit question:

> Pesticide A most likely functions as:
>  
> A) an agonist.  
> B) an inhibitor.  
> C) a lipase.  
> D) a receptor.

To answer this question, you first need to retrieve information from the passage about the effects of Pesticide A. From Table 2 we know that it prevents estrogen from binding to its receptor (and we noted this on our passage map). You also need to remember the definitions of the terms in the answer choices (agonists cause similar effects, inhibitors prevent effects, lipases break down lipids, and receptors bind ligands to cause effects). Based on our known definitions, choices A and D can be eliminated, and while Pesticide A could be functioning as a lipase that breaks down estrogen, “inhibitor” is a more accurate term (choice B is better than choice C and is the correct answer).

A final subgroup in the Explicit question category are graph interpretation questions. These fall into one of two types: those that ask you to take graphical information from the passage and convert it to a text answer, or those that take text from the passage and ask you to convert it to a graph. On the following page is an example of the latter type:
Which of the following graphs would best illustrate the binding of estrogen (E) to its receptor in the presence of its analog, Pesticide B?

A) % E bound to receptor
   Pesticide B
B) % E bound to receptor
   Pesticide B
C) % E bound to receptor
   Pesticide B
D) % E bound to receptor
   Pesticide B

From our passage map, we know that information about Pesticide B is found near the end of the passage, where it describes “further experiments.” The passage states that Pesticide B functions as an estrogen analog that binds to the estrogen receptor with a much higher affinity than does estrogen. In other words, if Pesticide B is around, the receptor will preferentially bind it, and not estrogen. So as the concentration of Pesticide B rises, the amount of estrogen bound to the receptor should fall. This is shown in choice B.

If you find that you are missing Explicit questions, practice your passage mapping. Make sure you aren’t missing the critical items in the passage that lead you to the right answer. Slow down a little; take an extra 15 to 30 seconds per passage to read or think about it more carefully.

### Biology Implicit Questions

Implicit questions require the most thought. These require recall of not only biology information but also information gleaned from the passage and a more in-depth analysis of how the two relate. Implicit questions require more analysis and connections to be made than Explicit questions. Often they take the form “If...then...” Below is an example of a classic Implicit question, taken from the Experiment passage shown earlier.

If Experiment 2 were repeated, but this time exposing the cells first to Pesticide A and then to Pesticide B before exposing them to the green fluorescent-labeled estrogen and the red fluorescent probe, which of the following statements will most likely be true?

A) Pesticide A and Pesticide B bind to the same site on the estrogen receptor.
B) Estrogen effects would be observed.
C) Only green fluorescence would be observed.
D) Both green and red fluorescence would be observed.
To answer this question, conclusions have to be drawn from the experiments described in the passage, and new conclusions have to be predicted based on the new circumstance. Many more connections need to be made than when answering an Explicit question. From the passage, we need to figure out that Pesticide A is an inhibitor. We also have to figure out that it does not bind at the active site of the receptor (data from Table 2). We have to know what green fluorescence and red fluorescence imply. We have to draw on the information provided about Pesticide B to know that it is an analog and that it binds to the active site of the estrogen receptor. We have to combine all of this together and come to a logical conclusion: Since Pesticide A is an inhibitor, it would prevent the binding of Pesticide B and thus prevent estrogen effects (choice B can be eliminated). If Pesticide B cannot bind, we would only see green fluorescence (choice D can be eliminated, and choice C is probably correct). Since Pesticide A by itself does not produce red fluorescence, it must not be binding at the active site, which is where Pesticide B binds, (choice A can be eliminated, and choice C is definitely correct).

Here’s another example of an Implicit question, drawn from the same passage:

When the researcher performed Experiment 2 using Pesticide B instead of Pesticide A, which of the following fluorescence and estrogen effects did the researcher most likely observe?

A) Media: green and red  
Cytoplasm: green and red  
Nucleus: green and red  
Estrogen effects: no

B) Media: green only  
Cytoplasm: green and red  
Nucleus: green and red  
Estrogen effects: no

C) Media: green only  
Cytoplasm: green and red  
Nucleus: green and red  
Estrogen effects: yes

D) Media: green only  
Cytoplasm: green and red  
Nucleus: green only  
Estrogen effects: no

To answer this question, we again must combine passage information with logical inference and working memory. Since red fluorescence indicates binding of the receptor, and since the receptor is never in the media, there can never be red fluorescence in the media (choice A can be eliminated). We know from the passage that Pesticide B binds at the active site of the receptor, and we know that the receptor is found in the cytoplasm. We also know from the passage that Pesticide B causes increased mRNA transcription, and we know from memory that to induce mRNA transcription, the receptor must move into the nucleus. Thus, red fluorescence must be observed in the nucleus as well (choice D can be eliminated). Since Pesticide B is defined as an “estrogen analog,” and since we know from memory that analogs cause similar effects, it is likely that estrogen effects will be observed. The fact that increased mRNA transcription occurs supports this idea (choice B can be eliminated, and choice C is correct). Again, many more connections need to be made to answer Implicit questions; Process of Elimination is typically the best approach.
If you find that you are missing a lot of Implicit questions, first make sure that you are using POE aggressively. Second, go back and review the explanations for the correct answer to figure out where your logic went awry. Did you miss an important fact in the passage? Did you forget the relevant Biology content? Did you follow the logical train of thought to the right answer? Once you figure out where you made your mistake, you will know how to correct it.

### 2.7 SUMMARY OF THE APPROACH TO BIOLOGY

#### How to Map the Passage and Use Scratch Paper

1. The passage should not be read like textbook material, with the intent of learning something from every sentence (science majors especially will be tempted to read this way). Passages should be read to get a feel for the type of questions that will follow and to get a general idea of the location of information within the passage.

2. Highlighting—Use this tool sparingly, or you will end up with a passage that is completely covered in yellow highlighter! Highlighting in a Biology passage should be used to draw attention to a few words that demonstrate one of the following:
   - the main theme of a paragraph
   - an unusual or unfamiliar term that is defined specifically for that passage (e.g., something that is italicized)
   - statements that either support the main theme or counteract the main theme
   - list topics (see below)
   - relationships

3. Pay brief attention to equations, figures, and experiments, noting only what information they deal with. Do not spend a lot of time analyzing at this point.

4. For each passage, start by noting the passage number, the general topic, and the range of questions on your scratch paper. You can then work between your scratch paper and the review screen to easily get to the questions you want to (see Chapter 1).

5. For each paragraph, note “P1,” “P2,” etc. on the scratch paper and jot down a few notes about that paragraph. Try to translate biology-speak into your own words using everyday language. Especially note down simple relationships (e.g., the relationship between two variables).

6. Lists—Whenever a list appears in paragraph form, jot down on the scratch paper the paragraph and the general topic of the list. It will make returning to the passage more efficient and help to organize your thoughts.

7. Scratch paper is only useful if it is kept organized! Make sure that your notes for each passage are clearly delineated and marked with the passage number and question range. This will allow you to easily read your notes when you come back to review a marked question. Resist the temptation to write in the first available blank space as this makes it much more difficult to refer back to your work.
Biology Question Strategies

1) Remember that the content in Biology is vast, so don’t panic if something seems completely unfamiliar. Understand the basic content well, find the basics in the unfamiliar topic, and apply them to the question.

2) Process of Elimination is paramount! The strikeout tool allows you to eliminate answer choices; this will improve your chances of guessing the correct answer if you are unable to narrow it down to one choice.

3) Answer the straightforward questions first (typically the memory questions). Leave questions that require analysis of experiments and graphs for later. Take the test in the order YOU want. Make sure to use your scratch paper to indicate questions you skipped.

4) Make sure that the answer you choose actually answers the question and isn’t just a true statement.

5) Try to avoid answer choices with extreme words such as “always,” “never,” etc. In biology, there is almost always an exception and answers are rarely black-and-white.

6) I-II-III questions: Always work between the I-II-III statements and the answer choices. Unfortunately, it is not possible to strike out the Roman numerals, but this is a great use for scratch paper notes. Once a statement is determined to be true (or false), strike out answer choices which do not contain (or do contain) that statement.

7) LEAST/EXCEPT/NOT questions: Don’t get tricked by these questions that ask you to pick that answer that doesn’t fit (the incorrect or false statement). Make sure to highlight the words “LEAST,” “EXCEPT,” or “NOT” in the question stem. It’s often good to use your scratch paper and write a T or F next to answer choices A–D. The one that stands out as different is the correct answer!

8) Again, don’t leave any question blank.

A Note About Flashcards

For most of the exams you’ve taken previously, flashcards were likely very helpful. This was because those exams mostly required you to regurgitate information, and flashcards are pretty good at helping you memorize facts. However, the most challenging aspect of the MCAT is not that it requires you to memorize the fine details of content knowledge, but that it requires you to apply your basic scientific knowledge to unfamiliar situations: flashcards alone may not help you there.

Flashcards can be beneficial if your basic content knowledge is deficient in some area. For example, if you don’t know the hormones and their effects in the body, flashcards can certainly help you memorize these facts. Or, maybe you are unsure of the functions of the different brain regions. You might find that flashcards can help you memorize these. But unless you are trying to memorize basic facts in your personal weak areas, you are better off doing and analyzing practice passages than carrying around a stack of flashcards.
Chapter 3

Biologically Important Molecules
The biological macromolecules are grouped into four classes of molecules that play important roles in cells and in organisms as a whole. All of them are polymers, strings of repeated units (monomers).

This chapter discusses the biomolecules from a biological perspective: what they are made of, how they are put together, and what their roles are in the body. These molecules are also discussed in *MCAT Biochemistry Review* in more detail, as well as in *MCAT Organic Chemistry Review* from an organic chemistry perspective: nomenclature, chirality, etc.
3.1 PROTEIN BUILDING BLOCKS

Proteins are biological macromolecules that act as enzymes, hormones, receptors, channels, transporters, antibodies, and support structures inside and outside cells. Proteins are composed of twenty different amino acids linked together in polymers. The composition and sequence of amino acids in the polypeptide chain is what makes each protein unique and able to fulfill its special role in the cell.

Amino Acid Structure and Nomenclature

Understanding the structure of amino acids is key to understanding both their chemistry and the chemistry of proteins. The generic formula for all twenty amino acids is shown below.

![Generic Amino Acid Structure]

All twenty amino acids share the same nitrogen-carbon-carbon backbone. The unique feature of each amino acid is its side chain (variable R-group), which gives it the physical and chemical properties that distinguish it from the other nineteen. Much more detail about amino acid structure, including their chemical properties, can be found in MCAT Biochemistry Review.
3.2 PROTEIN STRUCTURE

There are two common types of covalent bonds between amino acids in proteins: the peptide bonds that link amino acids together into polypeptide chains and disulfide bridges between cysteine R-groups.

The Peptide Bond

Polypeptides are formed by linking amino acids together in peptide bonds. A peptide bond is formed between the carboxyl group of one amino acid and the α-amino group of another amino acid with the loss of water. The figure below shows the formation of a dipeptide from the amino acids glycine and alanine.

![Figure 2 Peptide Bond (Amide Bond) Formation](image)

In a polypeptide chain, the N–C–C–N–C–C pattern formed from the amino acids is known as the backbone of the polypeptide. An individual amino acid is termed a residue when it is part of a polypeptide chain. The amino terminus is the first end made during polypeptide synthesis, and the carboxy terminus is made last. Hence, by convention, the amino-terminal residue is also always written first.

- In the oligopeptide Phe-Glu-Gly-Ser-Ala, which residue has a free α-amino group, and which residue has a free α-carboxyl group? (Refer to the beginning of the chapter for structures.)

---

1 As stated above, the amino end is always written first. Therefore, the oligopeptide begins with an exposed Phe amino group and ends with an exposed Ala carboxyl; all the other backbone groups are hitched together in peptide bonds.
Hydrolysis of a protein by another protein is called **proteolysis** or **proteolytic cleavage**, and the protein that does the cutting is known as a **proteolytic enzyme** or **protease**. Proteolytic cleavage is a specific means of cleaving peptide bonds. Many enzymes only cleave the peptide bond adjacent to a specific amino acid. For example, the protease trypsin cleaves on the carboxyl side of the residues arginine and lysine, while chymotrypsin cleaves adjacent to hydrophobic residues such as phenylalanine. (Do *not* memorize these examples.)

![Figure 3](image)

- Based on the above, if the following peptide is cleaved by trypsin, what amino acid will be on the new N-terminus and how many fragments will result: Ala-Gly-Glu-Lys-Phe-Phe-Lys?²

**The Disulfide Bond**

Cysteine is an amino acid with a reactive thiol (sulphydryl, SH) in its side chain. The thiol of one cysteine can react with the thiol of another cysteine to produce a covalent sulfur-sulfur bond known as a disulfide bond, as illustrated below. The cysteines forming a disulfide bond may be located in the same or different polypeptide chain(s). The disulfide bridge plays an important role in stabilizing tertiary protein structure; this will be discussed in the section on protein folding. Once a cysteine residue becomes disulfide-bonded to another cysteine residue, it is called **cystine** instead of cysteine.

![Figure 4](image)

² Trypsin will cleave on the carboxyl side of the Lys residue, with Phe on the N-terminus of the new Phe-Phe-Lys fragment. There will be two fragments after trypsin cleavage: Phe-Phe-Lys and Ala-Gly-Glu-Lys.
Protein Structure in Three Dimensions
Each protein folds into a unique three-dimensional structure that is required for that protein to function properly. Improperly folded, or denatured, proteins are non-functional. There are four levels of protein folding that contribute to their final three-dimensional structure. Each level of structure is dependent upon a particular type of bond, as discussed in the following sections.

Denaturation is an important concept. It refers to the disruption of a protein’s shape without breaking peptide bonds. Proteins are denatured by urea (which disrupts hydrogen bonding interactions), by extremes of pH, by extremes of temperature, and by changes in salt concentration (tonicity).

Primary [1°] Structure: The Amino Acid Sequence
The simplest level of protein structure is the order of amino acids bonded to each other in the polypeptide chain. This linear ordering of amino acid residues is known as primary structure. Primary structure is the same as sequence. The bond which determines 1° structure is the peptide bond, simply because this is the bond that links one amino acid to the next in a polypeptide.

Secondary structure refers to the initial folding of a polypeptide chain into shapes stabilized by hydrogen bonds between backbone NH and CO groups. Certain motifs of secondary structure are found in most proteins. The two most common are the α-helix and the β-pleated sheet.

There are two types of β-sheets, one with adjacent polypeptide strands running in the same direction (parallel β-pleated sheet) and another in which the polypeptide strands run in opposite directions (antiparallel β-pleated sheet), as shown in Figure 6.
If a single polypeptide folds once and forms a β-pleated sheet with itself, would this be a parallel or antiparallel β-pleated sheet?3


The next level of protein folding, tertiary structure, concerns interactions between amino acid residues located more distantly from each other in the polypeptide chain. The folding of secondary structures such as α-helices into higher order tertiary structures is driven by interactions of R-groups with each other and with the solvent (water). Hydrophobic R-groups tend to fold into the interior of the protein, away from the solvent, and hydrophilic R-groups tend to be exposed to water on the surface of the protein (shown for the generic globular protein).

It would be antiparallel because one participant in the β-pleated sheet would have a C to N direction, while the other would be running N to C.
Under the right conditions, the forces driving hydrophobic avoidance of water and hydrogen bonding will fold a polypeptide spontaneously into the correct conformation, the lowest energy conformation.

- Which of the following may be considered an example of tertiary protein structure?
  I. van der Waals interactions between two Phe R-groups located far apart on a polypeptide
  II. Hydrogen bonds between backbone amino and carboxyl groups
  III. Covalent disulfide bonds between cysteine residues located far apart on a polypeptide


The highest level of protein structure, quaternary structure, describes interactions between polypeptide subunits. A subunit is a single polypeptide chain that is part of a large complex containing many subunits (a multisubunit complex). The arrangement of subunits in a multisubunit complex is what we mean by quaternary structure. For example, mammalian RNA polymerase II contains twelve different subunits. The interactions between subunits are instrumental in protein function, as in the cooperative binding of oxygen by each of the four subunits of hemoglobin.

The forces stabilizing quaternary structure are generally the same as those involved in tertiary structure—non-covalent interactions, van der Waals forces, hydrogen bonds, disulfide bonds, and electrostatic interactions. It is key to understand, however, that there is one bond that may not be involved in quaternary structure—the peptide bond—because this bond defines sequence (1° structure).

- What is the difference between a disulfide bridge involved in quaternary structure and one involved in tertiary structure?

**3.3 CARBOHYDRATES**

Carbohydrates can be broken down to CO₂ in a process called oxidation, which is also known as burning or combustion. Because this process releases large amounts of energy, carbohydrates generally serve as the principle energy source for cellular metabolism. Glucose in the form of the polymer cellulose is also the building block of wood and cotton.

**Monosaccharides and Disaccharides**

A single carbohydrate molecule is called a monosaccharide (meaning “single sweet unit”), also known as a simple sugar. Monosaccharides have the general chemical formula CₙH₂₀₂ₙOₙ.

---

4 This is a simple question provided to clarify the classification of the disulfide bridge. Item I is a good example of 3° structure. Item II is describes 2°, not 3°, structure. Item III describes the disulfide, which is considered to be tertiary because of when it is formed, despite the fact that it is a covalent bond.

5 Quaternary disulfides are bonds that form between chains that aren’t linked by peptide bonds. Tertiary disulfides are bonds that form between residues in the same polypeptide.
BIOLOGICALLY IMPORTANT MOLECULES

Two monosaccharides bonded together form a disaccharide, a few form an oligosaccharide, and many form a polysaccharide. The bond between two sugar molecules is called a glycosidic linkage. This is a covalent bond, formed in a dehydration reaction that requires enzymatic catalysis.

Some common disaccharides you might see on the MCAT are sucrose (Glc-α-1,2-Fru), lactose (Gal-β-1,4-Glc), maltose (Glc-α-1,4-Glc), and cellobiose (Glc-β-1,4-Glc). However, you should NOT try to memorize these linkages.
Polymers made from these disaccharides form important biological macromolecules. Glycogen serves as an energy storage carbohydrate in animals and is composed of thousands of glucose units. Starch is the same as glycogen (except that the branches are a little different), and serves the same purpose in plants. Cellulose is a polymer of cellobiose; but note that cellobiose does not exist freely in nature. It exists only in its polymerized, cellulose form. The glycosidic bonds allow the polymer to assume a long, straight, fibrous shape. Wood and cotton are made of cellulose.

![Diagram of the Polysaccharide Glycogen]

**Figure 10** The Polysaccharide Glycogen

### 3.4 LIPIDS

Lipids are oily or fatty substances that play three physiological roles, summarized here and discussed below.

1. In adipose cells, triglycerides (fats) store energy.
2. In cellular membranes, phospholipids constitute a barrier between intracellular and extracellular environments.
3. Cholesterol is a special lipid that serves as the building block for the hydrophobic steroid hormones.

The cardinal characteristic of the lipid is its **hydrophobicity**. *Hydrophobic* means “water-fearing.” It is important to understand the significance of this. Since water is very polar, polar substances dissolve well in water; these are known as “water-loving,” or **hydrophilic** substances. Carbon-carbon bonds and carbon-hydrogen bonds are nonpolar. Therefore, substances that contain only carbon and hydrogen will not dissolve well in water. Some examples: Table sugar dissolves well in water, but cooking oil floats in a layer above water or forms many tiny oil droplets when mixed with water. Cotton T-shirts become wet when exposed to water because they are made of glucose polymerized into cellulose; a nylon jacket does not become wet because it is composed of atoms covalently bound together in a nonpolar fashion. A synonym for hydrophobic is **lipophilic** (which means “lipid-loving”); a synonym for hydrophilic is **lipophobic**. We return to these concepts below.
Fatty Acid Structure

Fatty acids are composed of long unsubstituted alkanes that end in a carboxylic acid. The chain is typically 14 to 18 carbons long, and because they are synthesized two carbons at a time from acetate, only even-numbered fatty acids are made in human cells. A fatty acid with no carbon-carbon double bonds is said to be saturated with hydrogen because every carbon atom in the chain is covalently bound to the maximum number of hydrogens. Unsaturated fatty acids have one or more double bonds in the tail. These double bonds are almost always (Z) (or cis).

![Saturated fatty acid](image1)

![Unsaturated fatty acid](image2)

Figure 11 Fatty Acid Structure

- How does the shape of an unsaturated fatty acid differ from that of a saturated fatty acid?  
- If fatty acids are mixed into water, how are they likely to associate with each other?

Figure 12 illustrates how free fatty acids interact in an aqueous solution; they form a structure called a micelle. The force that drives the tails into the center of the micelle is called the hydrophobic interaction.

![A Fatty Acid Micelle](image3)

6 An unsaturated fatty acid is bent, or “kinked,” at the cis double bond.

7 The long hydrophobic chains will interact with each other to minimize contact with water, exposing the charged carboxyl group to the aqueous environment.
Triacylglycerols [TG]
The storage form of the fatty acid is fat. The technical name for fat is triacylglycerol or triglyceride (shown below). The triglyceride is composed of three fatty acids esterified to a glycerol molecule. Glycerol is a three-carbon triol with the formula HOCH₂–CHOH–CH₂OH. As you can see, it has three hydroxyl groups that can be esterified to fatty acids. It is necessary to store fatty acids in the relatively inert form of fat because free fatty acids are reactive chemicals.

\[
\begin{align*}
H₂C–O–C–R₁ & \\
H₂C–O–C–R₂ & \\
H₂C–O–C–R₃ & \\
\end{align*}
\]

R-groups may be the same or different.

Figure 13  A Triglyceride (Fat)

Lipases are enzymes that hydrolyze fats. Triacylglycerols are stored in fat cells as an energy source. Fats are more efficient energy storage molecules than carbohydrates for two reasons: packing and energy content.

1) **Packing:** Their hydrophobicity allows fats to pack together much more closely than carbohydrates. Carbohydrates carry a great amount of water-of-solvation (water molecules hydrogen bonded to their hydroxyl groups).

2) **Energy content:** All packing considerations aside, fat molecules store much more energy than carbohydrates. In other words, regardless of what you dissolve it in, a fat has more energy carbon-for-carbon than a carbohydrate.

Introduction to Lipid Bilayer Membranes
Membrane lipids are phospholipids derived from diacylglycerol phosphate or DG-P.

\[
\begin{align*}
H₂C–O–C–R₁ & \\
H₂C–O–C–R₂ & \\
H₂C–O–P–O⁻ & \\
\end{align*}
\]

Figure 14  A Phosphoglyceride (Diacylglycerol Phosphate, or DGP)
We saw above how fatty acids spontaneously form micelles. Phospholipids also minimize their interactions with water by forming an orderly structure—in this case, it is a lipid bilayer (below). Hydrophobic interactions drive the formation of the bilayer, and once formed, it is stabilized by van der Waals forces between the long tails.

Figure 15  A Small Section of a Lipid Bilayer Membrane

- Would a saturated or an unsaturated fatty acid residue have more van der Waals interactions with neighboring alkyl chains in a bilayer membrane? 

A more precise way to give the answer to the question above is to say that double bonds (unsaturation) in phospholipid fatty acids tend to increase membrane fluidity. Unsaturation prevents the membrane from solidifying by disrupting the orderly packing of the hydrophobic lipid tails. The right amount of fluidity is essential for function. Decreasing the length of fatty acid tails also increases fluidity. The steroid cholesterol (discussed a bit later) is a third important modulator of membrane fluidity. At low temperatures, it increases fluidity in the same way as kinks in fatty acid tails; therefore, it is known as membrane antifreeze. At high temperatures, however, cholesterol attenuates (reduces) membrane fluidity. Don’t ponder this paradox too long; just remember that cholesterol keeps fluidity at an optimum level. Remember, the structural determinants of membrane fluidity are: degree of saturation, tail length, and amount of cholesterol.

Terpenes

A terpene is a member of a broad class of compounds built from isoprene units (C₅H₈) with a general formula (C₅H₈)ₙ.

Figure 16  Isoprene Unit

---

8 The bent shape of the unsaturated fatty acid means that it doesn’t fit in as well and has less contact with neighboring groups to form van der Waals interactions. Phospholipids composed of saturated fatty acids make the membrane more solid.
Terpenes may be linear or cyclic, and are classified by the number of isoprene units they contain. For example, monoterpenes consist of two isoprene units, sesquiterpenes consist of three, and diterpenes contain four.

**Figure 17  Terpene Structures**

Squalene is a triterpene (made of six isoprene units) and a particularly important compound, as it is biosynthetically utilized in the manufacture of steroids. Squalene is also a component of earwax.

**Figure 18  Squalene**

Whereas a terpene is formally a simple hydrocarbon, there are a number of natural and synthetically derived species that are built from an isoprene skeleton and functionalized with other elements (O, N, S, etc.). These functionalized-terpenes are known as terpenoids. Vitamin A ($C_{20}H_{30}O$) is an example of a terpenoid.

**Figure 19  Vitamin A**
Steroids
Steroids are included here because of their hydrophobicity, and, hence, similarity to fats. Their structure is otherwise unique. All steroids have the basic tetracyclic ring system (see below), based on the structure of cholesterol.

As discussed earlier, the steroid cholesterol is an important component of the lipid bilayer. It is both obtained from the diet and synthesized in the liver. It is carried in the blood packaged with fats and proteins into lipoproteins. One type of lipoprotein has been implicated as the cause of atherosclerotic vascular disease, which refers to the build-up of cholesterol “plaques” on the inside of blood vessels.

Steroid hormones are made from cholesterol. Two examples are testosterone (an androgen or male sex hormone) and estradiol (an estrogen or female sex hormone).
3.5 PHOSPHORUS-CONTAINING COMPOUNDS

Phosphoric acid is an inorganic acid (it does not contain carbon) with the potential to donate three protons. The $K_a$ values for the three acid dissociation equilibria are 2.1, 7.2, and 12.4. Therefore, at physiological pH, phosphoric acid is significantly dissociated, existing largely in anionic form.

![Figure 21 Phosphoric Acid Dissociation](image)

Phosphate is also known as orthophosphate. Two orthophosphates bound together via an anhydride linkage form pyrophosphate. The P–O–P bond in pyrophosphate is an example of a high-energy phosphate bond. This name is derived from the fact that the hydrolysis of pyrophosphate is extremely favorable.

There are three reasons that phosphate anhydride bonds store so much energy:

1) When phosphates are linked together, their negative charges repel each other strongly.
2) Orthophosphate has more resonance forms and thus a lower free energy than linked phosphates.
3) Orthophosphate has a more favorable interaction with the biological solvent (water) than linked phosphates.

The details are not crucial. What is essential is that you fix the image in your mind of linked phosphates acting like compressed springs, just waiting to fly open and provide energy for an enzyme to catalyze a reaction.

![Figure 22 The Hydrolysis of Pyrophosphate](image)
Nucleotides
Nucleotides are the building blocks of nucleic acids (RNA and DNA). Each nucleotide contains a ribose (or deoxyribose) sugar group; a purine or pyrimidine base joined to carbon number one of the ribose ring; and one, two, or three phosphate units joined to carbon five of the ribose ring. The nucleotide adenosine triphosphate (ATP) plays a central role in cellular metabolism in addition to being an RNA precursor. Significantly more information about the structure of the nucleic acids can be found in *MCAT Biochemistry Review*. More information on the function of the nucleic acids will be provided in Chapter 4 of this book.

*ATP is the universal short-term energy storage molecule.* Energy extracted from the oxidation of foodstuffs is immediately stored in the phosphoanhydride bonds of ATP. This energy will later be used to power cellular processes; it may also be used to synthesize glucose or fats, which are longer-term energy storage molecules. This applies to *all* living organisms, from bacteria to humans. Even some viruses carry ATP with them outside the host cell, though viruses cannot make their own ATP.

![Figure 23 Adenosine Triphosphate (ATP)](image-url)
Chapter 3 Summary

- Amino acids [AAs] consist of a tetrahedral $\alpha$-carbon connected to an amino group, a carboxyl group, and a variable R group, which determines the AA’s properties.

- Proteins consist of amino acids linked by peptide bonds, which are very stable. The primary structure of a protein consists of its amino acid sequence.

- The secondary structure of proteins [$\alpha$-helices and $\beta$-sheets] is formed through hydrogen-bonding interactions between atoms in the backbone of the molecule.

- The most stable tertiary protein structure generally places polar AA’s on the exterior and nonpolar AA’s on the interior of the protein. This minimizes interactions between nonpolar AA’s and water, while optimizing interactions between side chains inside the protein.

- Proteins have a variety of functions in the body including [but not limited to] enzymes, structural roles, hormones, receptors, channels, antibodies, transporters, etc.

- The monomer for a carbohydrate is a monosaccharide [simple sugar], with the molecular formula $C_nH_{2n}O_n$. The common monosaccharides are glucose, fructose, galactose, ribose, and deoxyribose.

- Two monosaccharides joined with a glycosidic linkage form a disaccharide. The common disaccharides are maltose, sucrose, and lactose. Mammals can digest $\alpha$ glycosidic linkages, but generally not $\beta$ linkages.

- Polysaccharides consist of many monosaccharides linked together. Glycogen [animals] and starch [plants] are storage units for glucose and can be broken down for energy. Cellulose is also a glucose polymer, but the beta linkage prevents digestion. It forms wood and cotton.

- Lipids are found in several forms in the body, including triglycerides, phospholipids, cholesterol and steroids, and terpenes. Triglycerides and phospholipids are linear, while cholesterol and steroids have a ring structure.
BIOLOGICALLY IMPORTANT MOLECULES

- Lipids are hydrophobic. Triglycerides are used for energy storage, phospholipids form membranes, and cholesterol is the precursor to the steroid hormones.

- The building blocks of nucleic acids (DNA and RNA) are nucleotides, which are comprised of a pentose sugar, a purine or pyrimidine base, and 2-3 phosphate units.
CHAPTER 3 FREESTANDING PRACTICE QUESTIONS

1. Why is ATP known as a “high energy” structure at neutral pH?
   A) It exhibits a large decrease in free energy when it undergoes hydrolytic reactions.
   B) The phosphate ion released from ATP hydrolysis is very reactive.
   C) It causes cellular processes to proceed at faster rates.
   D) Adenine is the best energy storage molecule of all the nitrogenous bases.

2. Which of the following best describes the secondary structure of a protein?
   A) Various folded polypeptide chains joining together to form a larger unit
   B) The amino acid sequence of the chain
   C) The polypeptide chain folding upon itself due to hydrophobic/hydrophilic interactions
   D) Peptide bonds hydrogen-bonding to one another to create a sheet-like structure

3. Phenylketonuria (PKU) is an autosomal recessive disorder that results from a deficiency of the enzyme phenylalanine hydroxylase. This enzyme normally converts phenylalanine into tyrosine. PKU results in intellectual disability, growth retardation, fair skin, eczema and a distinct musty body odor. Which of the following is most likely true?
   A) Treatment should include a decrease in tyrosine in the diet.
   B) The musty body odor is likely caused by a disorder in aromatic amino acid metabolism.
   C) Patients with PKU should increase the amount of phenylalanine in their diet.
   D) PKU can be acquired by consuming too much aspartame (an artificial sweetener that contains high levels of phenylalanine).

4. A genetic regulator is found to contain a lysine residue that is important for its binding to DNA. If a mutation were to occur such that a different amino acid replaces the lysine at that location, which of the following resulting amino acids would likely be the least harmful to its ability to bind DNA?
   A) Glycine
   B) Glutamate
   C) Aspartate
   D) Arginine

5. Increasing the amount of cholesterol in a plasma membrane would lead to an increase in:
   A) membrane permeability.
   B) atherosclerotic plaques.
   C) membrane fluidity at low temperatures but a decrease in membrane fluidity at high temperatures.
   D) membrane fluidity at high and low temperatures.

6. A human space explorer crash-lands on a planet where the native inhabitants are entirely unable to digest glycogen, but are able to digest cellulose. Consequently, they make their clothing out of glycogen-based material. The starving space explorer eats one of the native inhabitants’ shirts and the natives are amazed. Based on this information, which of the following is/are true?
   I. The explorer can digest α-glycosidic linkages.
   II. The native inhabitants can digest α-glycosidic linkages.
   III. The native inhabitants can digest starch.
   A) I only
   B) I and III only
   C) II and III only
   D) I, II, and III
CHAPTER 3 PRACTICE PASSAGE

Photosynthesis is the process plants use to derive energy from sunlight and is associated with a cell’s chloroplasts. The energy is used to produce carbohydrates from carbon dioxide and water. Photosynthesis involves light and dark phases. Figure 1 represents two initial steps associated with the light phase.

The light phase supplies the dark phase with NADPH and a high-energy substrate.

A researcher attempted to produce a photosynthetic system outside the living organism according to the following protocols:

- Chloroplasts were extracted from green leaves and ruptured, and their membranes were thereby exposed, then a solution of hexachloroplatinate ions carrying a charge of $-2$ was added.
- The structure of the composite was analyzed, and the amount of oxygen produced by the system was measured.

The researcher concluded that the ions were bound to the membrane’s Photosystem 1 site by the attraction of opposite charges. The resulting composite is shown in Figure 2. It was found that the hexachloroplatinate-membrane composite was photosynthetically active.

1. In concluding that the hexachloroplatinum ions were bound to Photosystem 1 due to the attraction of opposite charges, the researchers apparently assumed that the structure of the membrane was:

   A) determined solely by hydrophobic bonding.
   B) positively charged.
   C) covalently bound to the platinate.
   D) negatively charged.
2. Figure 1 indicates that:
   A) photoactivation of the chloroplast membrane results in the reduction of the anhydride-containing molecule NADP⁺.
   B) electrons are lost from Photosystem 1 through the conversion of NADPH to NADP⁺, and are replaced by electrons from Photosystem 2.
   C) there is a net gain of electrons by the system.
   D) electrons are lost from Photosystem 1 through the conversion of NADP⁺ to NADPH, but are not replaced by electrons from Photosystem 2.

3. In addition to NADPH, the photosynthetic light phase must supply the dark phase with another molecule which stores energy for biosynthesis. Among the following, the molecule would most likely be:
   A) ADP.
   B) CO₂.
   C) inorganic phosphate.
   D) ATP.

4. If NADP⁺ is fully hydrolyzed to its component bases, phosphates, and sugars, what type of monosaccharide would result?
   A) A three-carbon triose
   B) A hexose
   C) A pentose
   D) An α-D-glucose

5. If in a given cell the photosynthetic dark phase were artificially arrested while the light phase proceeded, the cell would most likely experience:
   A) decreased levels of NADPH.
   B) increased levels of NADPH.
   C) increased levels of carbohydrate.
   D) increased photoactivation of the chloroplast.

6. To determine the primary structure of the protein portion of Photosystem 1, a series of cleavage reactions was undertaken. To break apart the protein, the most logical action to take would be to:
   A) decarboxylate free carboxyl groups.
   B) hydrolyze peptide bonds.
   C) repolymerize peptide bonds.
   D) hydrolyze amide branch points.

7. A researcher examined a sample of the principal substance produced by the photosynthetic dark phase and concluded that he was working with a racemic mixture of glucose isomers. Which of the following experimental findings would be inconsistent with such a conclusion?
   A) The sample is composed of carbon, hydrogen, and oxygen only.
   B) The sample consists of an aldohexose.
   C) The sample rotates the plane of polarized light to the left.
   D) The sample is optically inactive.
SOLUTIONS TO CHAPTER 3 FREESTANDING PRACTICE QUESTIONS

1. **A** Choice A is the best answer because it directly addresses the energetics of ATP hydrolysis. Choice B discusses the reactivity of the released phosphate ion and not the structure of ATP itself, so it can be eliminated. Choice C can be eliminated because it describes the rate of cellular processes not the energy of ATP. Choice D can be eliminated because the structure of adenine is not related to why ATP is a good energy storage molecule.

2. **D** The secondary structure of proteins is the initial folding of the polypeptide chain into \( \alpha \)-helices or \( \beta \)-pleated sheets. Choice A can be eliminated because it describes the formation of a quaternary protein, choice B can be eliminated because it describes the primary protein structure, and choice C can be eliminated because it describes the tertiary protein structure.

3. **B** A defect in phenylalanine hydroxylase (or the THB cofactor) would result in a build-up of phenylalanine. This would lead to an excess of phenylalanine byproducts such as phenylacetate, phenyllactate and phenylpyruvate, and a decrease in tyrosine. Therefore, patients with PKU should increase the amount of tyrosine in their diet (it becomes an essential amino acid in this condition; choice A is wrong), as well as eliminate phenylalanine from their diet (choice C is wrong). PKU is a genetically acquired disorder (autosomal recessive), as mentioned in the passage, and thus it is not acquired by consuming too much phenylalanine (choice D is wrong). It is true that phenylalanine and its derivatives are aromatic amino acids and that the high levels of these compounds lead to the distinct musty body odor (choice B is correct). Process of Elimination (POE) is probably the best method to use in answering this question since it is unclear (without prior knowledge of PKU) what the underlying mechanism of the body odor would be.

4. **D** While knowing the structures of the different amino acids is unlikely to be important for the MCAT, knowing which of the amino acids are basic (histidine, arginine, lysine) and which are acidic (glutamate, aspartate) is likely to be relevant. In this case, since lysine is basic (and therefore best at binding the negatively charged DNA), one can assume that a mutation resulting in another basic amino acid would cause the least change in its ability to bind DNA. Therefore, a mutation from lysine to arginine would cause the least harm (choice D is correct). A mutation from lysine to glutamine or aspartate (both acidic) would likely cause the most harm to its ability to bind DNA (choices B and C are wrong). Glycine is a neutral amino acid (choice A is wrong).

5. **C** Plasma membranes can be up to 50% composed of sterols. Sterols help stabilize the membrane at both spectrums of the temperature. At low temperatures, cholesterol increases fluidity because the ring structure of cholesterol does not allow for tight phospholipid tail packing. At high temperatures, cholesterol decreases membrane fluidity (the OH group of cholesterol prevents phospholipid dispersion; choice C is correct, and choice D is wrong). Cholesterol decreases the permeability of membranes by filling in the “holes” between the fatty acid tails (choice A is wrong). The formation of atherosclerotic plaques, while related to cholesterol, is due to high levels of blood cholesterol, not membrane cholesterol (choice B is wrong).
6. A Item I is true: Humans can digest $\alpha$-glycosidic linkages, such as those found in glycogen. If the natives' shirts are made of glycogen, our space explorer should have no trouble consuming and digesting them (choice C can be eliminated). Item II is false: Cellulose contains $\beta$-glycosidic linkages. If the natives can digest cellulose, but not glycogen, then they cannot digest $\alpha$-glycosidic linkages (choice D can be eliminated). Item III is false: Starch also contains $\alpha$-glycosidic linkages. If the natives cannot digest glycogen, then they likely cannot digest starch either (choice B can be eliminated, and choice A is true).

**SOLUTIONS TO CHAPTER 3 PRACTICE PASSAGE**

1. B The passage states that the ion is attracted to Photosystem 1 by the attraction of opposite charges (positively charged photosystem and negatively charged hexachloroplatinate ion).

2. A The main result of the light phase, as depicted in Figure 1, is the reduction of NADP$^+$ to make NADPH (choice A is correct). Choice B is wrong since NADP$^+$ is converted into NADPH, not vice versa. Choice C is incorrect since in any system, mass and charge are conserved. Electrons move from one molecule to another, but they are not created or destroyed in a chemical reaction. Choice D can be eliminated since Figure 1 depicts electrons moving from Photosystem 2 to Photosystem 1.

3. D The passage states that the light reactions supply the dark reactions with a “high-energy substrate.” The most likely candidate among the choices is ATP, which is choice D.

4. C NADPH contains ribose, a pentose. Choice C is correct.

5. B The light phase makes NADPH, and the dark phase consumes it. In the absence of the dark phase, NADPH will continue to be produced, but none will be consumed, making NADPH levels rise (choice B is correct). Choice C is wrong since the dark phase is responsible for biosynthesis, such as carbohydrate production, so this will decrease, not increase. Choice D can be eliminated since the amount of light and photoactivation should remain the same.

6. B Proteins are composed of amino acid residues which are joined together by peptide bonds during the translation process. To split the protein into smaller pieces, proteases and chemical reagents act to hydrolyze the peptide bond, reversing the biosynthetic process.

7. C A racemic mixture is one which contains equal quantities of two stereoisomers that rotate plane-polarized light in opposite directions. Since there are equal quantities of both, racemic mixtures are optically inactive. Thus, choice C, which states that the sample rotates light, is inconsistent with the conclusion that the sample is racemic and is the correct answer choice. All other choices are consistent with the conclusion that the sample is a racemic mixture of glucose. Carbohydrates, of which glucose is one, are made of only carbon, hydrogen, and oxygen (choice A is consistent and can be eliminated). Glucose, with six carbons and a carbonyl group on the 6th carbon, is an aldohexose (choice B is consistent and can be eliminated), and racemic mixtures do not rotate light (choice D is consistent and can be eliminated).
Chapter 4
Molecular Biology
It was once thought that simple living organisms were generated spontaneously from nonliving matter. When a steak went bad and became infested with larvae, it was because the decomposing meat actually became squirming worms. Most religions have traditional explanations for the origin of human life, too. Children are derived from adults due to the will of a deity; the original adults were placed on the earth by that deity. But as empiricism developed during the Enlightenment, rigorous experiments were used to explain life, resulting in “scientific” models that are gradually replacing more traditional explanations.

One early conclusion was that simple organisms were derived not from decomposing matter but from parental organisms. Subsequently, it was found that some organisms are too small to be seen with the naked eye. These “germs” were eventually implicated as the cause of most major diseases. Gradually the scientific community came to the conclusion that all life was derived from other life. The patterns of inheritance and evolution were elucidated by a chain of scientists, from Mendel through Darwin. But the mechanism remained a mystery. Finally, cellular biology advanced to the point that scientists were aware of two substances found in cells that seemed appropriate vehicles for the transmission of inherited information: DNA and protein. The extreme length and orderly arrangement of repeating units in DNA and protein made it seem very likely that they could contain information. Researchers had waded through a chemical ocean of alphabet soup and suddenly come upon long strings of what looked like letters.

This is where biology stood in the early 1940s. In the ’40s and ’50s, two monumental achievements in microbiology finally clarified the gears in the clock of evolution and how they turn. One was the elucidation of the structure of DNA by Watson and Crick. The other was the proof by Avery, Herriott, Hershey, Chase, and their coworkers that DNA was the fundamental unit of genetic inheritance in microorganisms. In the following discussion, we will summarize the wealth of information that has been built upon these two prescient cornerstones.
4.1 DNA STRUCTURE

General Overview
Significant detail about the structure of DNA and RNA is provided in *MCAT Biochemistry Review*; here we will give a brief overview of this topic. DNA is short for deoxyribonucleic acid. DNA and RNA (ribonucleic acid) are called nucleic acids because they are found in the nucleus and possess many acidic phosphate groups.

The building block of DNA is the deoxyribonucleoside 5’ triphosphate (dNTP, where N represents one of the four basic nucleosides). Deoxyadenosine 5’ triphosphate (dATP) is shown in Figure 1; the other bases are shown in Figure 2. Nucleotides are built from three components: a sugar (deoxyribose for DNA, ribose for RNA), an aromatic, nitrogenous base, and 1-3 phosphate groups. The bases G and A are derived from a precursor called purine, so they are referred to as the purines. C, T, and U are the pyrimidines.1

A nucleoside is ribose or deoxyribose with a purine or pyrimidine linked to the 1’ carbon; nucleotides are phosphate esters of nucleosides, with one, two, or three phosphate groups joined to the ribose ring by the 5’ hydroxy group. When nucleotides contain three phosphate residues, they are also referred to as nucleoside triphosphates, abbreviated NTP (if the sugar is deoxyribose, they are abbreviated dNTP). In individual nucleotides, N is replaced by A, G, C, T, or U.

![Figure 1: Deoxyadenosine Triphosphate (dATP)](image)

1 A mnemonic for this is: Pyramids (pyrimidines) have sharp edges, so they CUT. Another mnemonic is CUT the Py.
The sugar + phosphate portion of the nucleotide is referred to as the **backbone** of DNA, because it is invariant. The base is the variable portion of the building block.

**PYRIMIDINE BASES**

- Cytosine
- Thymine (DNA only)
- Uracil (RNA only)

**PURINE BASES**

- Adenine
- Guanine

**Figure 2** Aromatic Bases of DNA and RNA

**Polynucleotides**

Nucleotides in the DNA chain are covalently linked by **phosphodiester bonds** between the 3’ hydroxy group of one deoxyribose and the 5’ phosphate group of the next deoxyribose (Figure 3). A polymer of several nucleotides linked together is termed an oligonucleotide, and a polymer of many nucleotides is a polynucleotide. Since the only unique part of the nucleotide is the base, the sequence of a polynucleotide can be abbreviated by simply listing the bases attached to each nucleotide in the chain. The end of the chain with a free 5’ phosphate group is written first in a polynucleotide, with other nucleotides in the chain indicated in the 5’ to 3’ direction.
**The Watson–Crick Model of DNA Structure**

James Watson and Francis Crick (with the help of Maurice Wilkins and Rosalind Franklin) developed a model of the structure of DNA in the cell. According to the **Watson–Crick model**, cellular DNA is a right-handed double helix held together by hydrogen bonds between bases. It is important to understand each facet of this model.

In the cell, DNA does not exist in the form of a single long polynucleotide. Instead, the DNA found in the nucleus is double-stranded (ds). In ds-DNA, two very long polynucleotide chains are hydrogen-bonded together in an **antiparallel** orientation. Antiparallel means the 5' end of one chain is paired with the 3' end of the other. The H-bonds in ds-DNA are between the bases on adjacent chains. This H-bonding is very specific: A is always H-bonded to T, and G is always H-bonded to C (Figure 4). Note that this

---

**Figure 3**  The Polymerization of Nucleotides
means an H-bonded pair always consists of a purine plus a pyrimidine. Thus both types of base pairs (AT or GC) take up the same amount of room in the DNA double helix. The GC pair is held together by three hydrogen bonds, the AT pair by two. Two chains of DNA are said to be complementary if the bases in each strand can hydrogen bond when the strands are oriented in an antiparallel fashion. If we are talking about ds-DNA 100 nucleotides long, we would say it is 100 base pairs (bp) long. A kbp (kilobase pair) is ds-DNA 1000 nucleotides long.

The binding of two complementary strands of DNA into a double-stranded structure is termed annealing, or hybridization. The separation of strands is termed melting, or denaturation.

Which of the following is/are true about ds-DNA?

I. If the amount of G in a double helix is known, the amount of C can be calculated.
II. If the fraction of purine nucleotides and the total molecular weight of a double helix are known, the amount of cytosine can be calculated.
III. The two chains in a piece of ds-DNA containing mostly purines will be bonded together more tightly than the two chains in a piece of ds-DNA containing mostly pyrimidines.
IV. The oligonucleotide ATGTAT is complementary to the oligonucleotide ATACAT.

This fact has a fringe benefit: We can calculate the number of purines if we know the number of pyrimidines. We can actually calculate several variables. Chargoff’s rule states that \([A] = [T]\) and \([G] = [C]\); and \([A] + [G] = [T] + [C]\).

Item I: True. For every G, there is a C; and for every A there is a T. Item II: False. The ratio of purines to pyrimidines is always the same (50:50) since each purine is paired with a pyrimidine. In order to calculate the amount of any one base, you have to know the ratio of AT to GC pairs. Item III: False. Again, the ratio of purines to pyrimidines is always the same; 50:50. However, two chains containing mostly GC pairs will bond more tightly than two chains containing mostly AT pairs, since GC pairs are held together by 3 H-bonds while AT pairs have only 2. Item IV: True. Remember: the strands are antiparallel, A and T pair, G and C pair, and the 5' end is always written first.
There is another important detail about DNA structure: Not only is it double stranded, it is also coiled. In ds-DNA, the two hydrogen-bonded antiparallel DNA strands form a **right-handed double helix** (meaning it corkscrews in a clockwise motion) with the bases on the interior and the ribose/phosphate backbone on the exterior. The double helix is stabilized by van der Waals interactions between the bases, which are stacked upon each other.

![A Small Section of a DNA Double Helix](image)

**Figure 5** A Small Section of a DNA Double Helix

**Chromosome Structure and Packing**

The sum total of an organism’s genetic information is called its *genome*. Eukaryotic genomes are composed of several large pieces of linear ds-DNA; each piece of ds-DNA is called a *chromosome*. Humans have 46 chromosomes, 23 of which are inherited from each parent. Prokaryotic (bacterial) genomes are composed of a **single circular chromosome**. Viral genomes may be linear or circular DNA or RNA. The human genome consists of over $10^9$ base pairs while bacterial genomes contain only $10^6$ base pairs. But there is no direct correlation between genome size and evolutionary sophistication, since the organisms with the largest known genomes are amphibians. Much of the size difference in higher eukaryotic genomes is the result of repetitive DNA that has no known function.

If the DNA remained as a simple double helix floating free in the cell, it would be very bulky and fragile. Prokaryotes have a distinctive mechanism for making their single circular chromosome more compact and sturdy. An enzyme called *DNA gyrase* uses the energy of ATP to twist the gigantic circular molecule. Gyrase functions by breaking the DNA and twisting the two sides of the circle around each other. The resulting structure is a twisted circle that is composed of ds-DNA. As discussed earlier, the two strands are already coiled, forming a helix. The twists created by DNA gyrase are called *supercoils*, since they are coils of a structure that is already coiled.
Since eukaryotes have even more DNA in their genome than prokaryotes, the eukaryotic genome requires denser packaging to fit within the cell (Figure 6). To accomplish this, eukaryotic DNA is wrapped around globular proteins called histones. After being wrapped around histones, but before being completely packed away, DNA has the microscopic appearance of beads on a string. The beads are called nucleosomes; they are composed of DNA wrapped around an octamer of histones (a group of eight). The string between the beads is a length of double-helical DNA called linker DNA and is bound by a single linker histone. Fully packed DNA is called chromatin; it is composed of closely stacked nucleosomes.

To look for patterns and morphology, chromosomes can be stained with chemicals. Usually, condensed metaphase chromosomes are used, as they are compact and easier to see. When chromosomes are treated, distinct light and dark regions become visible. The darker regions are denser and are called heterochromatin. Heterochromatin is rich in repeats; the lighter regions are less dense and are called euchromatin. Density gives a sense of DNA coiling or compactness, and these patterns are constant and heritable. It’s now known that the lighter regions have higher transcription rates and therefore higher gene activity. The looser packing makes DNA accessible to enzymes and proteins.
**Centromeres**

A centromere is the region of the chromosome to which spindle fibers attach during cell division. The fibers attach via **kinetochores**, multiprotein complexes that act as anchor attachment sites for spindle fibers. Other protein complexes also bind the centromere after DNA replication to keep sister chromatids attached to each other. Centromeres are made of heterochromatin, and repetitive DNA sequences. Chromosomes have p (short) and q (long) arms, and centromere position defines the ratio between the two (Figure 7).

![Figure 7: Centromere Positions](image)

**Telomeres**

The ends of linear chromosomes are called **telomeres**. At the DNA level, these regions are distinguished by the presence of distinct nucleotide sequences repeated 50 to several hundred times. The repeated unit is usually 6-8 base pairs long and guanine-rich. Many vertebrates (including humans and mice) have the same repeat: 5′-TTAGGG-3′. Telomeres are composed of both single- and double-stranded DNA. Single stranded DNA is found at the very end of the chromosome and is about 300 base pairs in length. It loops around to form a knot, held together by many telomere-associated proteins. This stabilizes the end of the chromosome; specialized telomere cap proteins distinguish telomeres from double-stranded breaks (Section 4.4), and this prevents activation of repair pathways.

Telomeres function to prevent chromosome deterioration and also prevent fusion with neighboring chromosomes. They function as disposable buffers, blocking the ends of chromosomes. DNA replication of telomeres represents a special challenge to cellular machinery (see Section 4.4). Since most prokaryotes have circular genomes, their DNA does not contain telomeres.
4.2 GENOME STRUCTURE AND GENOMIC VARIATIONS

The human genome contains 24 different chromosomes (22 autosomes, plus two different sex chromosomes), 3.2 billion base pairs, and codes for about 21,000 genes. The sequence of the human genome was reported by two independent groups in 2001 (the publicly funded Human Genome Project lead by Dr. Francis Collins, and Dr. J. Craig Venter and his firm Celera Genomics).

The human genome has numerous regions with high transcription rates, separated by long stretches of intergenic space. Intergenic regions are composed of noncoding DNA; they may direct the assembly of specific chromatin structures and can contribute to the regulation of nearby genes, but many have no known function. Tandem repeats and transposons are major components of intergenic regions.

Genomic regions with high transcription rates are rich in genes. A gene is a DNA sequence that encodes a gene product. It includes both regulatory regions (such as promoters and transcription stop sites), and a region that codes for either a protein or a non-coding RNA.

Nucleotide Variation

Small-scale and large-scale variation across a genome is common. For example, one person could have the sequence CCCGGG, while another has CCTGGG. It’s been predicted that there are single nucleotide changes once in every 1,000 base pairs in the human genome. These variations are called single nucleotide polymorphisms (SNPs, pronounced “snips”) and are essentially mutations. [If the size of the human genome is just over 3 billion base pairs, approximately how many human SNPs are there?] These SNPs occur most frequently in noncoding regions of the genome, however some SNPs can lead to specific traits and phenotypes. For example, about 70% of people taste phenylthiocarbamide (PTC) as very bitter, and the remaining 30% don’t taste PTC at all. You may have done this test yourself, as PTC response is commonly used as an example in genetics classes. This ability is a dominant genetic trait and is determined by a gene on chromosome 7. Three SNPs in this gene determine PTC taste sensitivity.

Copy-Number Variation

Copy-number variations (CNVs) are structural variations in the genome that lead to different copies of DNA sections. Large regions of the genome (10^5 to 10^6 base pairs) can be duplicated (increasing copy number) or deleted (decreasing copy number). The specific mechanism by which this occurs is not clear, but it may be due to misalignment of repetitive DNA sequences during synapsis of homologous chromosomes in meiosis. These changes therefore apply to much larger regions of the genome compared to SNPs. They are a normal part of our genome (0.4% of the genome can have CNV), but have also been associated with cancer and other diseases. Genes involved in immune system function, as well as brain development and activity, are often enriched in CNVs.

\[ 3 \times 10^9 \text{ base pairs} \times 1 \text{ SNP/1000 base pairs} = 3 \times 10^6 \text{ SNPs, or approximately 3 million human SNPs.} \]
Repeated Sequences: Tandem Repeats

Much of our genome is single copy, meaning there is one copy of the gene in a haploid set of the genome. This is true for most eukaryotic genes that code for proteins. However, genomes also have regions of tandem repeats, where short sequences of nucleotides are repeated one right after the other, from as little as three to over 100 times. The human genome has over a thousand regions of tandem repeats. Repeats can be unstable, when the repeating unit is short (such as di- or trinucleotides) or when the repeat itself is very long. Unstable tandem repeats can lead to chromosome breaks and some have been implicated in disease. Tandem repeats often show variations in length between individuals, which can be useful in DNA fingerprinting. Heterochromatin, centromeres, and telomeres are all rich in repeats.

4.3 THE ROLE OF DNA

DNA encodes and transmits the genetic information passed down from parents to offspring. Before 1944 it was generally believed that protein, rather than DNA, carried genetic information, since proteins have an “alphabet” of 20 letters (the amino acids), while DNA’s “alphabet” has only 4 letters (the four nucleotides). But in that year, Oswald Avery showed that DNA was the active agent in bacterial transformation. In short, this means he proved that pure DNA from one type of E. coli bacteria could transform E. coli of another type, causing it to acquire the genetic nature of the first type. Later Hershey and Chase proved that DNA was the active chemical in the infection of E. coli bacteria by bacteriophage T2. These experiments will be discussed in more detail in Chapter 7.

The Genetic Code

DNA does not directly exert its influence on cells, but merely contains sequences of nucleotides known as genes that serve as templates for the production of another nucleic acid known as RNA. The process of reading DNA and writing the information as RNA is termed transcription. This can generate either a final gene product (as in the case of all non-coding RNAs, discussed below), or a messenger molecule. The messenger RNA (mRNA) is then read, and the information is used to construct protein. The synthesis of proteins using RNA as a template is termed translation and is accomplished by the ribosome, which is a massive enzyme composed of many proteins and pieces of RNA (known as ribosomal RNA or rRNA).

The overall process looks like this: DNA → RNA → protein. This unidirectional flow equation represents the Central Dogma (fundamental law) of molecular biology. This is the mechanism whereby inherited information is used to create actual objects, namely enzymes and structural proteins.

This language used by DNA and mRNA to specify the building blocks of proteins is known as the Genetic Code. The alphabet of the genetic code contains only four letters (A, T, G, C). How can four letters specify the ingredients of the multitude of proteins in every cell? [What is the smallest “word” size that
would allow this four-letter alphabet to encode twenty different amino acids? A number of experiments confirmed that the genetic code is written in three-letter words, each of which codes for a particular amino acid. A nucleic acid word (3 nucleotide letters) is referred to as a **codon**.

The genetic code is represented in Figure 8. The first nucleotide in a codon is given at the left, the second on top, and the third on the right. At the intersection of these three nucleotides is the amino acid called for by that codon. [Why is uracil (U) shown in the chart, and why is thymine (T) absent? The codon GTG in DNA is transcribed in RNA as __, which the ribosome translates into what amino acid?]

![Figure 8 The Genetic Code](image)

- The genetic code was studied by experimenters using a cell-free protein synthesis system. All of the materials necessary for protein synthesis (ribosomes, amino acids, tRNA, GTP, ATP) were purified and placed in a beaker. Then synthetic RNA was added, and protein was translated from this template. For example, when synthetic RNA containing only cytosine (CCCCC…) was added, polypeptides containing only proline (polyproline) resulted. What kind of synthetic RNA would give rise to a mixture of polyproline, polyhistidine, and polythreonine?

---

7 With four nucleotides, if a “word” (codon) is two nucleotides long, there are $4^2 = 16$ possible codons; too few to specify 20 unique amino acids. However, there are $4^3 = 64$ possible 3-letter “words,” and 64 is more than enough different codons to specify 20 unique amino acids. Thus, three nucleotides is the minimum codon size.

8 RNA is the nucleic acid that actually encodes protein during translation. RNA has U instead of T.

9 The RNA codon transcribed from the DNA will be CAC, coding for histidine.

10 The RNA would have to be CCACCACCACCACCACCACCAC…. This would yield polyproline if read as CCA, CCA, CCA. But if it were read as CAC, CAC, CAC, it would give rise to polyhistidine. If it were read ACC, ACC, ACC, it would encode polythreonine.
There are 64 codons. Sixty-one of them specify amino acids; the remaining three are called stop codons. Their function is to notify the ribosome that the protein is complete and cause it to stop reading the mRNA (see Section 4.5). Stop codons are also called nonsense codons, since they don’t code for any amino acid. Note that most of the twenty amino acids can be coded for by more than one codon. Often, all four of the codons with the same first two nucleotides (e.g., CU_) encode the same amino acid. [If the last nucleotide in the codon CUU is changed in a gene that codes for a protein, will the protein be affected?11] Two or more codons coding for the same amino acid are known as synonyms. Because it has such synonyms, the genetic code is said to be degenerate. However, it is very important to realize that though an amino acid may be specified by several codons, each codon specifies only a single amino acid. This means that each piece of DNA can be interpreted only one way: The code has no ambiguity.

The code in Figure 8 is the standard genetic code and is used by most organisms. However, some protists use an alternate genetic code, and the mitochondrial genome (see Section 4.10) of many organisms (including humans and many other vertebrates) uses a slightly different code.

**Beyond the Central Dogma**

There are several aspects of molecular biology that aren’t explicitly stated in the Central Dogma.

- Some viruses (retroviruses) make DNA from RNA using the enzyme reverse transcriptase (see Section A.6 in the Appendix).
- Information can also be transferred in other ways. For example, DNA methylation and post-translational modification of proteins can alter gene expression and convey information, despite the fact that neither is directly included in the Central Dogma.
- Many final gene products are not proteins but are RNAs instead.

**4.4 DNA REPLICATION**

The DNA genome is the control center of the cell. When mitosis produces two identical daughter cells from one parental cell, each daughter must have the same genome as the parent. Therefore, cell division requires duplication of the DNA, known as replication. This is an enzymatic process, just as the Krebs cycle and glycolysis are enzymatic processes. It occurs during S (synthesis) phase in interphase of the cell cycle (Chapter 6). Let’s go through the process of replication, stopping to add essential facts to a list of things to memorize. But before we get bogged down with details, we should have a look at the big picture.

There is only one logical way to make a new piece of DNA that is identical to the old one: copy it. The old DNA is called parental DNA, and the new is called daughter DNA. What is the relationship between parental and daughter DNA after replication? There are several possibilities (Figure 9). In other words, where do the atoms from the parent go when the daughters are made?

---

11 No, since CUN codes for leucine, regardless of what N is. Notice that switching the third nucleotide in the majority of codons will have no effect.
Experiments done by Meselson and Stahl in 1958 aimed to determine if DNA replication is semiconservative, conservative, or dispersive (Figure 9). In conservative replication, the parental ds-DNA would remain as-is while an entirely new double-stranded genome was created. The dispersive theory said that both copies of the genomes were composed of scattered pieces of new and old DNA. Meselson and Stahl showed that replication is semiconservative; after replication, one strand of the new double helix is parental (old) and one strand is newly synthesized daughter DNA.

Let’s begin the list of things to memorize here:

1) **DNA replication is semiconservative.**

   Individual strands of the double-stranded parent are pulled apart, and then a new daughter strand is synthesized using the parental DNA as a template to copy from.\(^\text{12}\) Each new daughter chain is perfectly \(^\text{13}\) to its template or parent.

Now we’ll look at replication at the molecular level. When it is not being replicated, DNA is tightly coiled. The replication process cannot begin unless the double helix is uncoiled and separated into two single strands. The enzyme that unwinds the double helix and separates the strands is called helicase. \([\text{Would you expect helicase to use the energy of ATP hydrolysis to do its job?}^\text{14}]\) The place where the helicase begins to unwind is not random. It is a specific location (sequence of nucleotides) on the chromosome called the **origin of replication** (abbreviated ORI). This sequence is found by proteins with tertiary structures to specifically recognize a particular pattern of nucleotides. They scan along the chromosome

---

\(^{12}\) A template is something that is copied. The metal plates used in printing presses are an example.

\(^{13}\) complementary

\(^{14}\) Yes. Separating the strands requires the breaking of many H-bonds.
like a train on a track) until they find the right spot; then they call in helicase and other enzymes to initiate DNA replication. In prokaryotes, a protein called DnaA finds the ORI to initiate DNA replication. In eukaryotes, three proteins cooperate to find the ORI, two of which are synthesized during M and G1 phases of the cell cycle (see Chapter 6) but rapidly destroyed once the S phase begins. This means these two proteins link DNA replication to the cell cycle, ensuring DNA replication doesn’t initiate during other phases of the cell cycle.

When helicase unwinds the helix at the origin of replication, the helix gets wound more tightly upstream and downstream from this point. The chromosome would get tangled and eventually break, except that enzymes called topoisomerases cut one or both of the strands and unwrap the helix, releasing the excess tension created by the helicases. Another potential problem is that single-stranded DNA is much less stable than ds-DNA. Single-strand binding proteins (SSBPs) protect DNA that has been unpackaged in preparation for replication and help keep the strands separated. The separated strands are referred to as an open complex. Replication may now begin.

An RNA primer must be synthesized for each template strand. This is accomplished by a set of proteins called the primosome, of which the central component is an RNA polymerase called primase. Primer synthesis is important because the next enzyme, DNA polymerase, cannot start a new DNA chain from scratch. It can only add nucleotides to an existing nucleotide chain. The RNA primer is usually 8–12 nucleotides long, and is later replaced by DNA.

Daughter DNA is created as a growing polymer. DNA polymerase (DNA pol) catalyzes the elongation of the daughter strand using the parental template, and elongates the primer by adding dNTPs to its 3’ end. In fact, the 3’ hydroxyl group acts as a nucleophile in the polymerization reaction to displace 5’ pyrophosphate from the dNTP to be added. [The template strand is read in what direction?] DNA pol is part of a large complex of proteins called the replisome. Other accessory proteins in this complex help DNA polymerase and allow it to polymerize DNA quickly. The prokaryotic replisome contains 13 components, and the eukaryotic replisome contains 27 proteins; additional complexity in the eukaryotic system is required because replication machinery must also unwind DNA from histone proteins.

---

15 Imagine two long ropes wound around each other. What happens if you pull them apart in the middle?
16 If the daughter is made 5’ to 3’, and the two strands have to end up antiparallel, the template must be read 3’ to 5’.
Rapid elongation of the daughter strands follows. Since the two template strands are antiparallel, the two primers will elongate toward opposite ends of the chromosome. After a while it looks like this:

![Diagram of DNA replication](image)

DNA polymerase checks each new nucleotide to make sure it forms a correct base-pair before it is incorporated in the growing polymer. The thermodynamic driving force for the polymerization reaction is the removal and hydrolysis of pyrophosphate ($P_2O_7^{4-}$) from each dNTP added to the chain. Here are some more replication rules to memorize:

2) **Polymerization occurs in the 5’ to 3’ direction, without exception.** This means the existing chain is always lengthened by the addition of a nucleotide to the 3’ end of the chain. There is never 3’ to 5’ polymerase activity.

3) **DNA pol requires a template.** It cannot make a DNA chain from scratch but must copy an old chain. This makes sense because it would be pretty useless if DNA pol just made a strand of DNA randomly, without copying a template.

4) **DNA pol requires a primer.** It cannot start a new nucleotide chain.

- Can DNA polymerase make the following partially double-stranded structure completely double stranded in the presence of excess nucleotides, using the top strand as a primer?17

![Nucleotides](image)

---

17 No. The DNA strands are antiparallel, meaning that the upper strand would have to be extended in a 3’ to 5’ direction, which is impossible. Note that the phrase “in the presence of excess nucleotides” is extraneous. It just means there are plenty of building blocks around. Typical MCAT smokescreen.
Replication proceeds along in both directions away from the origin of replication. Both template strands are read 3' to 5' while daughter strands are elongated 5' to 3'. The areas where the parental double helix continues to unwind are called the replication forks. Let’s split Figure 11 and look at an enlargement of the right side:

![Leading Strand Diagram](image1)

**Figure 12**  Leading Strand

See how it looks like a big fork? In examining these pictures, you have probably become aware of a problem. It seems like only half of each template strand will be replicated (in Figure 11, the right half of the bottom strand and the left half of the top strand). The problem is that chain elongation can only proceed in one direction, 5' to 3', but in order to replicate the right half of the top chain and the left half of the bottom one continuously, we would have to go in the opposite direction. Here's the solution:

![Leading and Lagging Strands Diagram](image2)

**Figure 13**  Leading and Lagging Strands

The solution to this problem involves building strands of DNA on opposite sides of the ORI using different methods. As the bottom chain on the right is elongated continuously, the replication fork widens. After a good bit of the top template chain becomes exposed, primase comes in and lays down a primer, which DNA pol can elongate. Then, when the replication fork widens again and more of the top template becomes exposed, these events are repeated. The bottom daughter on the right side, and the top daughter on the left side are called the leading strands because they elongate continuously right into the widening replication fork. The top daughter on the right, and the bottom daughter on the left are called the lagging strands because they must wait until the replication fork widens before beginning to polymerize. The
small chunks of DNA comprising the lagging strand are called **Okazaki fragments**, after their discoverer.

[As the replication forks grow, does helicase have to continue to unwind the double helix and separate the strands?**18**] Let’s continue our memory-list:

5) **Replication forks grow away from the origin in both directions.** Each replication fork contains a **leading strand** and a **lagging strand**.

6) Replication of the leading strand is **continuous** and leads into the replication fork, while replication of the lagging strand is **discontinuous**, resulting in Okazaki fragments.

7) Eventually **all RNA primers are replaced by DNA**, and the **fragments are joined by an enzyme called DNA ligase**.

---

**DNA Polymerase**

DNA polymerase can rapidly build DNA and is able to add tens of thousands of nucleotides before falling off the template. It is therefore said to be **processive**.

Eukaryotes have several different DNA polymerase enzymes, and their mechanisms of action are complex. You do not need to worry about this complexity.

Prokaryotes on the other hand have five types of DNA polymerases, called DNA polymerase I, II, III, IV and V. You should definitely know the functions of DNA pol III and DNA pol I:

1) **DNA pol III** is responsible for the super-fast, super-accurate elongation of the leading strand. In other words, it has high processivity. It has 5' to 3' polymerase activity as well as 3' to 5' exonuclease**19** activity. This is when the enzyme moves backward to chop off the nucleotide it just added, if it was incorrect; the ability to correct mistakes in this way is known as **proofreading function**. It has no known function in repair, and so is considered a replicative enzyme.

2) **DNA pol I** starts adding nucleotides at the RNA primer; this is 5' to 3' polymerase activity. Because of its poor processivity (it can only add 15-20 nucleotides per second), DNA pol III usually takes over about 400 base pairs downstream from the ORI. DNA pol I is also capable of 3' to 5' exonuclease activity (proofreading). DNA pol I removes the RNA primer via 5' to 3' exonuclease activity, while simultaneously leaving behind new DNA in **20** activity. Finally, DNA pol I is important for excision repair.

---

**18 Yes.**

**19 Exonuclease** means “cutting a nucleic acid chain at the end.” An **endonuclease** will cut a polynucleotide acid chain in the middle of the chain, usually at a particular sequence. Two important types of endonucleases are: **repair enzymes** that remove chemically damaged DNA from the chain, and **restriction enzymes**, which are endonucleases found in bacteria. Their role is to destroy the DNA of infecting viruses, thus restricting the host range of the virus.

**20 5’ to 3’ polymerase; remember, all polymerization is 5’ to 3’**.
The functions of DNA pol II, IV, and V are less important to know for the MCAT:

3) DNA pol II has 5' to 3' polymerase activity, and 3' to 5' exonuclease proofreading function. It participates in DNA repair pathways and is used as a backup for DNA pol III.

4) DNA pol IV and DNA pol V have similar characteristics. They are error prone in 5' to 3' polymerase activity, but function to stall other polymerase enzymes at replication forks when DNA repair pathways have been activated. This is an important part of the prokaryotic checkpoint pathway.

If a bacterium possesses a mutation in the gene for DNA polymerase III, resulting in an enzyme without the 3' to 5' exonuclease activity, will mutations occur more often than in bacteria with a normal DNA polymerase gene?21

Prokaryotic vs. Eukaryotic Replication

Prokaryotes have only one chromosome, and this one chromosome has only one origin. Because the chromosome is circular, as replication proceeds the partially duplicated genome begins to look like the Greek letter θ (theta). Hence the replication of prokaryotes is said to proceed by the theta mechanism and is referred to as theta replication (see Figure 14).

Yes. The 3' to 5' exonuclease activity is the polymerase’s way of editing its work. Without this editing function, many more point mutations would occur due to the incorporation of wrong nucleotides. The normal polymerase is remarkably adept at sensing correct base pairing and removing bases that don’t belong.
In eukaryotic replication, each chromosome has several origins. This is necessary because eukaryotic chromosomes are so huge that replicating them from a single origin would be too slow. As the many replication forks continue to widen, they create an appearance of bubbles along the DNA strand, so they are referred to as “replication bubbles.” Eventually the replication forks meet, and the many daughter strands are ligated together.

![Replication Bubbles](image1)

**Figure 15  Eukaryotic Replication**

### Replicating Telomeres in Eukaryotes

DNA polymerase can only build DNA in one direction (5' to 3'), and requires both a template and a primer. These requirements lead to a roadblock at chromosome ends. Eventually there will be no place on the lagging strand to lay down a primer, and primers close to the end of DNA cannot be replaced with DNA because there is nothing on the other side (DNA polymerase usually uses a previous length of upstream DNA to replace the primer, but this isn’t available at the end of a chromosome). This means that DNA replication machinery is unable to replicate sequences at the very ends of chromosomes, and after each round of the cell cycle and DNA replication, the ends of chromosomes shorten.

**Telomeres** are disposable repeats at the end of chromosomes. They are consumed and shorten during cell division, becoming 50 and 200 base pairs shorter.

When telomeres become too short, they reach a critical length where the chromosome can no longer replicate. As a consequence, cells can activate DNA repair pathways, enter a senescent state (where they are alive but not dividing), or activate apoptosis (pre-programmed cell death). The **Hayflick limit** is the number of times a normal human cell type can divide until telomere length stops cell division. Many age-related diseases are linked to telomere shortening.

**Telomerase** is an enzyme that adds repetitive nucleotide sequences to the ends of chromosomes and therefore lengthens telomeres. Telomerase is a ribonucleoprotein complex, containing an RNA primer and reverse transcriptase enzyme. Reverse transcriptases read RNA templates and generate DNA. In humans, the RNA template is 3'-CCCAATCCC-5', and this allows for chromosome extension, one DNA repeat (5'-TTAGGG-3') at a time (Figure 16). The telomerase complex continuously polymerizes, then translocates, allowing extension of six-nucleotide telomere repeats.
In most organisms, telomerase is only expressed in the germ line, embryonic stem cells, and some white blood cells. However, cancer cells can also express telomerase, which can help the cells immortalize. Telomere extension allows the cells to bypass senescence and apoptosis, and can therefore contribute to their transformation to a pre-cancerous state.

4.5 GENETIC MUTATION

Genetic mutation refers to any alteration of the DNA sequence of an organism’s genome. These can be inherited or acquired throughout life. Mutations that can be passed onto offspring are called germline mutations, since they occur in the germ cells (which give rise to gametes). Somatic mutations occur in somatic (non-gametic) cells and are not passed onto offspring. In other words, somatic mutations can have a major effect on an individual, but will not be passed on to future individuals in that population. Our cells have evolved elaborate repair pathways to help deal with mutations, and these will be discussed in the next section.

Causes of Mutation

There are many causes of mutation. Most are induced by an environmental factor or chemical; however, they can also occur spontaneously.

Physical Mutagens

Ionizing radiation (such as X-rays, alpha particles, and gamma rays) can cause DNA breaks. If these only occur on one strand (Figure 17, left), they can be easily patched up because the DNA helix is still held together in one piece. However, if both backbones are broken close to each other on a segment of DNA, a double-strand break (DSB) occurs (Figure 17, right). Here, the chromosome has been split into two pieces and it’s much more difficult to piece them back together.
UV light causes photochemical damage to DNA. For example, if two pyrimidines (two Cs or two Ts) are beside each other on a DNA backbone, UV light can cause them to become covalently linked. These pyrimidine dimers distort the DNA backbone (Figure 18) and can cause mutations during DNA replication if they are not repaired.
Reactive Chemicals
Many chemicals interact directly with DNA, and many others turn into damaging agents as they’re being processed by a cell. Chemicals can covalently alter bases or can cause cross-linking or strand breaks. Cross-links are abnormal covalent bonds between different parts of DNA. Any compound that can cause mutations is called a mutagen.

Compounds that look like purines and pyrimidines (with large flat aromatic ring structures) cause mutations by inserting themselves between base pairs, or intercalating, thereby causing errors in DNA replication. Ethidium bromide is often used to visualize nucleic acids during gel electrophoresis in molecular biology labs. This chemical is used because it is planar (and therefore intercalates with the DNA ladder), and glows orange when exposed to UV light (meaning nucleic acids in a gel can be easily visualized). However, because it intercalates with DNA, it also distorts the structure and can therefore disrupt DNA replication and transcription. Thus, ethidium bromide is a mutagen.

Biological Agents
Biological agents can also cause mutations. For example, although DNA polymerase has proofreading and correction abilities, it can still make a mistake. An incorrect base pair may be repaired (see Section 4.6), but if not, it will be passed on to all daughter cells. In this case, there is no mutagen. The mistake is spontaneous. Viruses can also affect DNA. Lysogenic viruses insert into the genome of the host cell (see Chapter 5), and this can cause mutations and disrupt genetic function. Some viruses can cause cancer because of this function. And finally, transposons can induce mutations. This will be described in the next section.

Types of Mutations
Based on structure, there are seven kinds of mutations:

1) Point mutations
2) Insertions
3) Deletions
4) Inversions
5) Amplifications
6) Translocations and rearrangements
7) Loss of heterozygosity

Point mutations are single base pair substitutions (A in place of G, for example). Point mutations can be transitions (substitution of a pyrimidine for another pyrimidine or substitution of a purine for another purine) or transversions (substitution of a purine for a pyrimidine or vice versa). There are three types of point mutations:

1) Missense mutation: This causes one amino acid to be replaced with a different amino acid. This may not be serious if the amino acids are similar. [How can this occur?]

---

22 For example, substituting a small hydrophobe such as valine for another small hydrophobe like leucine will probably cause little disruption of protein structure. Another way of defining conservative mutations is that they cause changes in primary structure but do not affect secondary, tertiary, or quaternary structure.
2) **Nonsense mutation**: A stop codon replaces a regular codon and prematurely shortens the protein.

3) **Silent mutation**: A codon is changed into a new codon for the same amino acid, so there is no change in the protein’s amino acid sequence.

Insertion refers to the addition of one or more extra nucleotides into the DNA sequence, and deletion is the removal of nucleotides from the sequence. Both of these mutations can cause a shift in the reading frame. For example, AAACCCACC is read as AAA, CCC, ACC. It would code for Lys-Pro-Thr. Inserting an extra G into the first codon could produce this: AGAACCCACC. This would be read AGA, ACC, CAC, C. It now codes for Arg-Thr-His (plus there’s an extra C). Not only has the first codon and amino acid changed, the whole gene will be read differently and all amino acids in the protein from that point on will change. Mutations that cause a change in the reading frame are called **frameshift mutations**. Generally speaking, frameshift mutations are very serious. Note that a frameshift can lead to premature termination of translation (yielding an incomplete polypeptide) if it results in the presence of an abnormal stop codon. [Are all insertions and deletions frameshift mutations?] If the following oligonucleotide is mutated by inserting a G between the fifth and sixth codons, what effect will this have on the oligopeptide it encodes: AUG AAG GGG CCC UUU AAA UGA CCC? For each type of mutation, does it involve a change in the genotype, the phenotype, or both?

In addition to mutations at individual nucleotides, larger-scale mutations are also common. Insertions and deletions can involve thousands of bases. An **inversion** is when a segment of a chromosome is reversed end to end. The chromosome undergoes breakage and rearrangement within itself (Figure 19). Insertions, deletions, and inversions can be caused by transposons (see below).

---

23 No. If you insert or delete one whole codon or several whole codons, you add or remove amino acids to the polypeptide without changing the reading frame.

24 The original RNA codes for Met-Lys-Gly-Pro-Phe-Lys. After the insertion, the oligonucleotide will code for Met-Lys-Gly-Pro-Phe-Glu-Met-Thr. Note that this contains different amino acids and it’s longer. The extra length is due to the fact that a stop codon, UGA, changed by the frameshift.

25 By definition, all mutations involve a change in the genotype. Most mutations also cause a change in the phenotype, but in the case of conservative mutations it is a very subtle change that would be hard to detect.
Chromosome amplification is when a segment of a chromosome is duplicated. This is similar to copy number variations discussed above. Translocations result when recombination occurs between nonhomologous chromosomes (Figure 20). This can create a gene fusion, where a new gene product is made from parts of two genes that were not previously connected. This is a common occurrence in many types of cancer. Translocations can be balanced (where no genetic information is lost), or unbalanced (where genetic information is lost or gained).

![Figure 20](image)

**Before translocation**

Chromosome 4

Chromosome 20

**After translocation**

Derivative Chromosome 4

Derivative Chromosome 20

**Figure 20** Chromosome Amplification

Transposons

Both prokaryotes and eukaryotes have mobile genetic elements in their genomes, called transposable elements or transposons. It is thought that many eukaryotic transposons are degenerate (old and defective) retroviruses. “Genetic mobility” means that these short segments can jump around the genome. Transposons can cause mutations and chromosome changes such as inversions, deletions, and rearrangements.

There are three common types of transposons, each with a different structure. The first type is the simplest and is called an IS element (Figure 21, top). It is composed of a transposase gene (discussed below), flanked by inverted repeat sequences. The structure of an example inverted repeat is shown in Figure 22. Some transposons are more complex, in that they also contain additional genes (Figure 21, middle). For example, some transposons contain genes for antibiotic resistance. Finally, composite transposons have two similar or identical IS elements with a central region in between (Figure 21, bottom).
All transposons contain a gene that codes for a protein called transposase. This enzyme has “cut and paste” activity, where it catalyzes mobilization of the transposon (excision from the donor site) and integration into a new genetic location (the acceptor site). Sometimes the transposon sequence is completely excised and moved, and sometimes it is duplicated and moved, while still maintained at the original location (Figure 23). The inverted repeats are important for this mobilization.
(1) The Transposase gene codes for a "cut and paste" Transposase enzyme.

(2) Transposase cuts a copy of the transposon out of the original or donor site (Chromosome 7 in this example).

(3) The transposon mobilizes to a new recipient site (Chromosome 3 in this case), but a copy also stays at the donor site.

Figure 23  The Mechanism of Transposon Mobilization

Many mobilizations have no effect because the transposon inserts into a relatively unimportant part of the genome. However, transposons can cause mutations if they jump into an important part of the genome.

When transposons are mobilized, they can insert in any part of the genome, and this can affect gene expression or cause mutations. They can jump into a promoter and turn gene expression off. They can jump into a protein-coding region and disrupt (or mutate) the sequence. They can also jump into regulatory parts of the genome and ramp up gene expression at a nearby site.

In addition to jumping around the genome, transposons can cause structural changes to chromosomes when they work in pairs. Directionality of the transposon is important here, as it determines what happens to the chromosome. If a chromosome has two transposons with the same direction (Figure 24), the transposons can line up beside each other, so they are parallel. This causes the chromosomal segment between them to loop around. Recombination occurs between the transposons, and this causes deletion of the DNA between the two transposons. The original chromosome therefore completely loses the DNA segment between the transposons (a deletion). The segment of DNA that is lost takes one transposon with it, meaning it can actually jump back into the genome somewhere else, causing chromosome rearrangement: one chunk of a chromosome has moved to a new location in the genome.
Inverted transposons pair

Recombination

DNA segment between direct transposons is deleted

DNA segment can integrate into a different genomic site, leading to chromosomal rearrangement

**Figure 24** Deletion and Chromosomal Rearrangements via Transposons

If a chromosome has two transposons with inverted orientations (Figure 25), they can again pair and align with each other. After recombination, the sequence of DNA between the two transposons ends up inverted.

**Figure 25** Chromosome Inversion via Transposons
Diploid organisms have two copies of each gene, and generally a mutation in one copy is tolerated as long as the other copy of the gene is normal. However, if a deletion removes the normal copy of the gene, the only remaining copy is the mutated version. This is referred to as **loss of heterozygosity**. This makes the locus **hemizygous**: there is only one gene copy in the diploid organism. If the remaining allele is mutant or defective, all gene expression of the normal gene product is lost. For example, hereditary retinoblastoma is a type of retinal cancer common in young children. It occurs when a child receives a flawed copy of the tumor suppressor Rb1 from one parent, and a loss of heterozygosity event leads to loss of the normal allele (from the other parent). With no functional Rb protein (due to having only one copy of Rb1, and it being a flawed or mutant copy), the child almost invariably develops retinoblastoma.

**Effects of Mutations**

There are many mechanisms by which mutations can exert their effects on the cell. A single amino acid change can affect protein activity, localization, degradation, half-life, or interactions, or, it may have no effect at all. The outcome of a mutation on a protein depends on where the mutation occurs. Mutations on sex chromosomes typically have a greater effect than mutations on autosomes since autosomes are present in double copies. Males have only one X chromosome and one Y chromosome, with no back-up copy of either. Similarly, most females only express one of their X chromosomes (see Section 4.9), and so they, too, often don't have a back-up copy. Haploid expression in a diploid organism is **hemizygosity**, and this can lead to an increased effect of mutations on these chromosomes.

Gain-of-function mutations increase the activity of a certain gene product, or change it such that it gains a new and abnormal function. Loss-of-function mutations are the opposite; they result in the gene product having less or no function. In **haploinsufficiency**, a diploid organism has only a single functional copy of a gene, and this single copy is not enough to support a normal state. Haploinsufficiency highlights the importance of gene dose: many times, just expressing a gene is not enough. You must express enough of the gene to maintain good health.

**Good and Bad Mutations**

Despite the bad reputation they have, not all mutations are bad. Many mutations are neutral, and have no effect. Evolution is based on mutations and selection, and some mutations are beneficial. Those that confer a survival advantage will be selected for in a population.

There are examples of beneficial mutations in humans:

- Sickle-cell anemia is caused by mutations in the gene for hemoglobin (Hb). One of the most common mutations allows deoxygenated Hb to dimerize and form long chains, which distorts the red blood cell shape, causing it to sickle. These deformed cells cannot function properly and are prematurely destroyed, leading to anemia. However, people who carry this gene also have an advantage in that they are more resistant to malaria. In areas where malaria is common, this is an important benefit.
• Some humans are missing 32 base pairs in a gene called CCR5. This deletion confers HIV resistance to homozygotes and delays AIDS onset in heterozygotes. This mutation may have also conferred resistance to diseases in the past (such as the bubonic plague or smallpox), explaining its prevalence in populations of European descent, where these diseases were prevalent.

Mutations can also be disease causing. In some cases, one mutation is sufficient to induce a diseased state. In other cases, many mutations have to cooperate and occur together to cause a disease.

Inborn errors of metabolism are a huge group of genetic diseases that involve disorders of metabolism. Most of these are due to a single mutation in a single gene that codes for some sort of metabolic enzyme. Symptoms are caused by either the build-up of a toxic compound that can’t be broken down or by the deficiency of an essential molecule that cannot be synthesized. Because cellular metabolism is crucial, many symptoms are possible and a wide range of systems can be affected. Inborn errors of metabolism are typically organized into groups of disorders, depending on what type of metabolic pathways they affect: carbohydrate, amino acid, urea cycle, organic acids, fatty acid oxidation, mitochondrial, porphyrin, purine or pyrimidine, steroid, peroxisomal function, or lysosomal storage.

Cancer is driven by mutation accumulation. These mutations can either be inherited, or can be caused by carcinogen exposure. A carcinogen is a mutagen that is directly involved in causing cancer. Tumors typically have hundreds of mutations, ranging from point mutations to massive chromosomal changes. These mutations are often in oncogenes and tumor suppressors. An oncogene is a gene that can cause cancer when it is mutated or expressed at high levels. Tumor suppressors are the opposite in that their deletion (or expression at decreased levels) can cause cancer. Some mutations will drive tumor growth and are highly selected for. These mutations are the most promising targets for developing cancer treatments, as the cancer cells rely on these mutations for growth.

4.6 DNA REPAIR

Cells have developed several mechanisms to deal with DNA damage. First, cell cycle checkpoints are activated, and arrest cell cycle progression. In eukaryotes, checkpoint pathways function at phase boundaries (such as the G1/S transition, and the G2/M transition), and can also be activated within some phases. Extensive DNA damage can induce apoptosis in eukaryotes, but before this happens, cells try to repair the DNA damage. This is important so that defective DNA isn’t passed on to daughter cells. There are several types of DNA repair.

Direct Reversal

Many types of DNA damage are irreversible and require repair pathways to fix the damage. However, a few can be directly reversed. For example, some enzymes can repair UV-induced pyrimidine photodimers using visible light. This process is called photoreactivation, and directly repairs the UV damage to DNA. This is commonly performed by bacteria and many plants. If pyrimidine dimers are not directly repaired, nucleotide excision repair can be used instead. This is the main mechanism of repair in humans, but can introduce a mutation when trying to complete the repair. If left un repaired, pyrimidine dimers in humans may lead to melanoma, a type of very dangerous and malignant skin tumor.
**Homology-Dependent Repair**

One of the benefits of DNA structure is the presence of a back-up copy; because DNA is double stranded, mutations on one strand of DNA can be repaired using the undamaged, complementary information on the other strand. Repair pathways that rely on this characteristic of DNA are called **homology-dependent repair pathways**. These can be divided into repair that happens before DNA replication (**excision repair**), or repair that happens during and after DNA replication (**post-replication repair**).

**Excision Repair**

Excision repair involves removing defective bases or nucleotides and replacing them. If these bases are not repaired, they can induce mutations during DNA replication, since replication machinery cannot pair them properly.

**Post-Replication Repair**

The **mismatch repair pathway** (MMR) targets mismatched Watson-Crick base pairs that were not repaired by DNA polymerase proofreading during replication. To do this, mispaired bases must be identified and fixed, but the crucial question is: which base is the correct one and which is the mistake? For example, if DNA contains an AC base pair, is the adenine correct and C should be removed and replaced with T? Or is the cytosine correct and A should be removed and replaced with G?

Some bacteria use genome methylation to help differentiate between the older DNA template strand and the newly synthesized daughter strand. Methylation takes a while to complete, which means that shortly after DNA synthesis, the parental template strand will be labeled with methylated bases and the new daughter strand will not. Bacterial machinery can read these methyl tags and know which base is the correct one (the one on the older strand) and which needs to be replaced (the newer one).

Other prokaryotes and most eukaryotes use a different system, where the newly synthesized strand is recognized by the free 3’-terminus on the leading strand, or by the presence of gaps between Okazaki fragments on the lagging strand.

**Double-Strand Break Repair**

DNA double-strand breaks (DSBs) can be caused by reactive oxygen species, ionizing radiation, UV light or chemical agents. Cells have two pathways to help in DSB repair: homologous recombination and nonhomologous end-joining. The goal of both is to reattach and fuse chromosomes that have come apart because of DSB. If done incorrectly, this can lead to deletions (where genetic information is lost) or translocations (where chromosome segments move to other chromosomes).

**Homologous Recombination**

After DNA replication, the genome contains identical sister chromatids. Homologous recombination is a process where one sister chromatid can help repair a DSB in the other. First, the DSB is identified and trimmed at 5’ ends to generate single-stranded DNA (Figure 26). This is done by nucleases (which break phosphodiester bonds) and helicase (to unwind the DNA). Many proteins bind these ends and start a search of the genome to find a sister chromatid region that is complementary to the single-stranded
DNA. Once found, the complementary sequences are used as a template to repair and connect the broken chromatid. This requires a “joint molecule,” where damaged and undamaged sister chromatids cross over. DNA polymerase and ligase build a corrected DNA strand.

**Figure 26** Homologous Recombination to Repair Double-Strand Breaks

**Nonhomologous End Joining**

Cells that aren’t actively growing or cycling through the cell cycle don’t have the option of using sister chromatids to repair DSBs in an error-free way. Since DNA replication isn’t happening, there is no chromosome backup to use. In this case, even a poorly repaired chromosome is better than one with a DSB, since chromosome breaks can lead to rearrangements.

Nonhomologous end joining is used to accomplish repair in this case. This process is common in eukaryotes but relatively uncommon in prokaryotes. First, broken ends are stabilized and processed, then DNA ligase connects the fragments. Nothing about this process requires specificity; the goal is just to reconnect broken chromosomes. Often this can result in base pairs being lost or chromosomes being connected in an abnormal way.
4.7 GENE EXPRESSION: TRANSCRIPTION

Gene expression refers to the process whereby the information contained in genes begins to have effects in the cell. The Central Dogma tells us that genetic information must be written in the form of RNA (i.e., it must be transcribed); and then it must be expressed as protein (i.e., it must be translated). Therefore, the logical place to begin our discussion of gene expression is with the nature of RNA and transcription.

Characteristics of RNA

RNA is chemically distinct from DNA in three important ways:

1) RNA is single-stranded, except in some viruses.
2) RNA contains uracil instead of thymine.
3) The pentose ring in RNA is ribose rather than 2’ deoxyribose.

There are several different types of RNA, each with a unique role.

Coding RNA

You are already familiar with messenger RNA (mRNA), the only type of coding RNA. This molecule carries genetic information to the ribosome, where it can be translated into protein; each unique polypeptide is created according to the sequence of codons on a particular piece of mRNA, which was transcribed from a particular gene. To allow for this, each mRNA has several regions. The 5’ region is not translated into protein (so is called the 5’ untranslated region, or 5’UTR), but is important in initiation and regulation. Following the 5’UTR is the region that codes for a protein. This starts at a start codon and ends at a stop codon, and is called the open reading frame (ORF). The 3’ end of the mRNA (after the stop codon) isn’t translated into protein, but often contains regulatory regions that influence post-transcriptional gene expression (see Section 4.9).

Eukaryotic mRNA is usually monocistronic and obeys the “one gene, one protein” principle. This means that each piece of mRNA encodes only one polypeptide (and so contains one ORF). Hence, there are as many different mRNAs as there are proteins. Because each mRNA can be read many times, each transcript can be used to make many copies of its polypeptide. There are a few exceptions to the “one gene, one protein” principle; recently, some polycistronic eukaryotic mRNAs have been discovered, and these will be discussed below.

In contrast, prokaryotic mRNA often codes for more than one polypeptide and is termed polycistronic. Different open reading frames on the same polycistronic mRNA are generally related in function. Translation termination and initiation sequences are found between the ORFs. The termination information helps finish the previous peptide chain, and initiation information helps start translation of the next open reading frame on the transcript.

For instance, if five enzymes are necessary for the synthesis of a particular molecule, then all five enzymes might be encoded on a single piece of mRNA.
Messenger RNA is constantly produced and degraded, according to the cell’s need for the protein encoded by each piece of mRNA. In fact, this is the principal means whereby cells regulate the amount of each particular protein they synthesize. This is an important point that will be emphasized later. Note that in eukaryotes, the first RNA transcribed from DNA is an immature or precursor to mRNA called heterogeneous nuclear RNA (hnRNA). Processing events (such as addition of a cap and tail, and splicing) are required for hnRNA to become mature mRNA. Since prokaryotes do not process their primary transcripts, hnRNA is only found in eukaryotes.

Non-Coding RNA
Non-coding RNA (ncRNA) is a functional RNA that is not translated into a protein. The human genome codes for thousands of ncRNAs, and there are several types. The two major types to know for the MCAT are transfer RNA (tRNA) and ribosomal RNA (rRNA).

Transfer RNA (tRNA) is responsible for translating the genetic code. Transfer RNA carries amino acids from the cytoplasm to the ribosome to be added to a growing protein. The structure of tRNA and how it does its job will be discussed in Section 4.8. [Estimate how many different tRNAs there are. 27]

Ribosomal RNA (rRNA) is the major component of the ribosome. Humans have only four different types of rRNA molecules (18S, 5.8S, 28S and 5S), although almost all the RNA made in a given cell is rRNA. All rRNAs serve as components of the ribosome, along with many polypeptide chains. One rRNA provides the catalytic function of the ribosome, which is a little odd. In most other cases, enzymes are made from polypeptides. Catalytic RNAs are also called ribozymes (or ribonucleic acid enzymes), since they are capable of performing specific biochemical reactions, similar to protein enzymes. There are additional examples of ribozymes, including snRNA (discussed below) and some introns that are self-splicing.

Some other interesting non-coding RNAs are:

- **Small nuclear RNA** (snRNA) molecules (150 nucleotides) associate with proteins to form snRNP (small nuclear ribonucleic particles) complexes in the spliceosome.
- **MicroRNA** (miRNA) and **small interfering RNA** (siRNA) function in RNA interference (RNAi), a form of post-transcriptional regulation of gene expression. Both can bind specific mRNA molecules to either increase or decrease translation. This will be discussed more in Section 4.9.
- **PIWI-interacting RNAs** (piRNAs) are single stranded and short (typically between 21 and 31 nucleotides in length). They work with a class of regulatory proteins called PIWI proteins to prevent transposons from mobilizing.
- **Long ncRNAs** are longer than 200 nucleotides. They help control the basal transcription level in a cell by regulating initiation complex assembly on promoters. They also contribute to many types of post-transcriptional regulation by controlling splicing and translation, and they function in imprinting and X-chromosome inactivation (see Section 4.9).

---

27 Each tRNA must recognize a codon on mRNA and respond by delivering the appropriate amino acid to the ribosome. There are 20 different amino acids, so there at least 20 different tRNAs. However, there are 61 possible codons, so there could be as many as 61 different tRNAs. The actual number is between 20 and 61, because the third nucleotide of the codon is often not needed for specificity of the amino acid.
Replication vs. Transcription

Transcription is the synthesis of RNA (usually mRNA, tRNA, or rRNA) using DNA as the template. The word transcription indicates that in the process of reading and writing information, the language does not change. Information is transferred from one polynucleotide to another. This should lead you to expect transcription to be fairly similar to replication. And it is.

Both replication and transcription involve template-driven polymerization. [Because of this, the RNA transcript produced in transcription is 28 complementary to the DNA template, just as the daughter strand produced in replication was.] The driving force for both processes is the removal and subsequent hydrolysis of pyrophosphate from each nucleotide added to the chain, with the existing chain acting as nucleophile. [Transcription, like replication, can occur only in the 29 5’ to 3’ direction. Do the polymerase enzymes in both replication and transcription require a primer? 30] Another important difference between transcription and DNA replication is that RNA polymerase has not been shown to possess the ability to remove mismatched nucleotides (it lacks exonuclease activity); in other words, it cannot correct its errors. Thus, transcription is a lower fidelity process than replication. [A virus possessing an RNA genome relies on RNA polymerase rather than DNA polymerase to replicate its genome. Will this virus have a higher or a lower rate of spontaneous mutation than organisms with ds-DNA genomes? 31]

Another similarity is that transcription, like replication, begins at a specific spot on the chromosome. The name of the site where transcription starts (the start site) is different from the name of the place where replication begins, 32 the origin. The sequence of nucleotides on a chromosome that activates RNA polymerase to begin the process of transcription is called the promoter, and the point where RNA polymerization actually starts is called the start site. In fact, from this point forward, just about every event in transcription is given a different name from the events in replication.

Reference Points in Transcription

Before we discuss the mechanics of transcription, we need to clarify a few reference points (see Figure 27). We noted previously that the chromosome is referred to as the template, not parent. What about the individual strands of the chromosome? Are they both templates for the same mRNA? Let’s answer with a thought experiment: Say there is a strand of DNA which has the sequence AAAAAAAAA. If we transcribe this strand, the resulting mRNA will look like: UUUUUUUUU. When it is translated, this mRNA will result in an oligopeptide with this primary structure: Phe-Phe-Phe. (Refer to the genetic code table in Section 4.3.) Now, what if we transcribe the other strand of the chromosome? What is its DNA sequence? What will the transcript look like? And the oligopeptide? 33 Our conclusion is that only one of the strands of the DNA template encodes a particular mRNA molecule. But it makes sense: paired DNA

28 complementary
29 5’ to 3’
30 No, RNA pol does not require a primer. Remember, the primer in replication is a piece of RNA, made by an RNA polymerase.
31 The virus will have a very high rate of mutation. It is a general law that most mutations are harmful. Hence, individual viruses will be far less likely to survive than organisms with DNA genomes. However, the high mutation rate will allow the entire species of virus to evolve very rapidly, making it very successful as a parasite (since it will evade host defense systems).
32 the origin
33 The DNA strand must be complementary to the first strand we discussed. So the sequence must be TTTTTTTTTT. Thus the transcript will have to be AAAAAAAAA. Because AAA codes for lysine, the oligopeptide would be Lys-Lys-Lys.

| 97 |
strands are complementary, not identical. The strand which is actually transcribed is called the template, non-coding, transcribed, or antisense strand; it is complementary to the transcript. The other DNA strand is called the coding or sense strand; it has the same sequence as the transcript (except it has T in place of U). It is customary to say that transcription starts at a point and proceeds downstream, which means toward the 3’ end of the coding strand and transcript. Upstream means toward the 5’ end of the coding strand, beyond the 5’ end of the transcript. Upstream nucleotide sequences are referred to using negative numbers, and downstream sequences are referred to using positive numbers. The first nucleotide on the template strand which is actually transcribed is called the start site. The corresponding nucleotide on the coding strand is given the number +1. As we’ll see below, regulatory sequences on the chromosome are referred to by where they occur on the coding strand.

Figure 27  Reference Points in Transcription

- The figure above labels the transcript “mRNA.” Is this accurate in all life forms? (Hint: In eukaryotes, is the initial transcript mature mRNA, ready to be translated?)

Prokaryotic Transcription

It is important to understand all the vocabulary and general principles presented above. In this section and the next, we will present some more detailed information.

In bacteria (prokaryotes), all types of RNA are made by the same RNA polymerase. Prokaryotic RNA polymerase is a large enzyme complex consisting of five subunits: two alpha subunits, a beta subunit, a beta’ subunit, and an omega subunit ($\alpha_2\beta\beta'\omega$). This is the core enzyme responsible for rapid elongation of the transcript. However, the core enzyme alone cannot initiate transcription. An additional subunit termed the sigma factor ($\sigma$) is required to form what is sometimes referred to as the holoenzyme (holo = complete), which is responsible for initiation.

Transcription occurs in three stages: initiation, elongation, and termination. Initiation occurs when RNA polymerase holoenzyme binds to a promoter. The typical bacterial promoter contains two primary sequences: the Pribnow box at –10 and the –35 sequence. Holoenzyme scans along the chromosome like a train on a railroad track until it recognizes a promoter and then stops, forming a closed complex. The RNA polymerase must unwind a portion of the DNA double helix before it can begin to synthesize RNA.

---

34 No, it is accurate for prokaryotes only. In eukaryotes, the RNA transcript must be processed (spliced) and transported out of the nucleus before it can be translated. We will discuss this in depth later in the chapter.
The RNA polymerase bound at the promoter with a region of single-stranded DNA is termed the open complex. Once the open complex has formed, transcription can begin.

The sigma factor plays two roles in helping the polymerase find promoters. The first is to greatly increase the ability of RNA polymerase to recognize promoters. The second is to decrease the nonspecific affinity of holoenzyme for DNA. Once the open complex and several phosphodiester bonds have been formed, the sigma factor is no longer necessary and leaves the RNA polymerase complex.

The core enzyme elongates the RNA chain processively, with one polymerase complex synthesizing an entire RNA molecule. As the core enzyme elongates the RNA, it moves along the DNA downstream in a transcription bubble in which a region of the DNA double helix is unwound to allow the polymerase to access the complementary DNA template. When a termination signal is detected, in some cases with the help of a protein called rho, the polymerase falls off of the DNA, releases the RNA, and the transcription bubble closes.

### Comparing Prokaryotic and Eukaryotic Transcription

Eukaryotic and prokaryotic transcription are similar, but you need to be aware of four major differences. Differences in location, RNA polymerases and primary transcripts are discussed here. Regulation of transcription is another major difference and is discussed in Section 4.9.

#### Location

Eukaryotic means “true-kernelled.” Prokaryotic means “before-the-kernel.” The karyon (kernel) is, of course, the nucleus. The fact that prokaryotes have no nucleus means transcription occurs free in the cytoplasm, in the same compartment where translation occurs, and transcription and translation can occur simultaneously. Eukaryotes must transcribe their mRNA in the nucleus, modify it (see below), and then transport it across the nuclear membrane to the cytoplasm where it can be translated. Transcription and translation in eukaryotes do not occur simultaneously.

Another important difference between prokaryotic and eukaryotic gene expression is that the primary transcript in prokaryotes is mRNA. In other words, the product of transcription by prokaryotic RNA polymerase is ready to be translated. In fact, translation of prokaryotic mRNA begins before transcription is completed!

In contrast, the eukaryotic primary transcript (hnRNA made by RNA pol II, see below for info on eukaryotic RNA polymerases) is modified extensively before translation (Figure 30). The most important example is splicing. Eukaryotic DNA has non-coding sequences intervening between the segments that actually code for proteins. Sometimes these intervening sequences contain enhancers or other regulatory sequences and they can be quite long. The average size of a mammalian intron, for example, is about 2000 nucleotides. Intervening sequences in the RNA are called introns. Note that introns are intragenic regions (and not intergenic space, discussed in Section 4.2). Protein-coding regions of the RNA are termed exons because they actually get expressed. Before the RNA can be translated, introns must be removed and exons joined together; this is accomplished via splicing.
Splicing is mediated by the spliceosome, a complex that contains over 100 proteins and 5 small nuclear RNA (snRNA) molecules. About half the proteins stably bind snRNAs, and these form three small nuclear ribonucleic particles (snRNPs). Each snRNP is therefore made of proteins and snRNAs. The spliceosome is not a pre-assembled complex, but rather assembles around each intron that needs to be removed. This happens in a series of steps, where different snRNP components are recruited and released as the reaction proceeds. The complex undergoes many conformational changes to attain catalytic activity.

To catalyze the splicing reaction, snRNPs recognize and hydrogen bond to conserved nucleotides in the intron: typically GU at the 5’ end, AG at the 3’ end, and an adenine 15-45 bases upstream of the 3’ splice site. This aligns the hnRNA such that the splicing mechanism can take place (Figure 28). Two splicing reactions are catalyzed by the spliceosome. The first reaction attaches one end of the intron to the conserved adenine. This causes the intron to form a looped structure, then the second reaction joins the two exons (Figure 28) and releases the loop. The five conserved nucleotides necessary for this reaction (GU, A and AG) are found in all genes and across all eukaryotic species.

For a given gene, there are often different options or patterns of splicing, a phenomenon called alternative splicing. There are many different common patterns. One gene could have different promoters in the 5’ region, which can change where/how the RNA begins. There can be alternative 5’ exons or 3’ exons, which can affect either end of the RNA. In the middle, too, some exons can be included or skipped. Finally, there could be mutually exclusive exons, where sometimes one is included and sometimes the other is kept. All these patterns lead to different mRNAs being made from one DNA gene sequence; the mRNAs can be different in length and sequence. Shuffling exons in this way is one way to increase the complexity of gene expression (Figure 29).
Alternative splicing is mediated by introns and exons, as well as by the proteins that can bind to these sequences. There are almost 200,000 introns in the human genome, with an average of about seven per gene. It was initially thought that introns were unimportant and had no function. While it's true that a lot of intron sequences are probably junk, the current picture of introns is a little more complicated than first believed.

Eukaryotic hnRNA must be modified in two other ways before translation can occur. A tag is added to each end of the molecule: a 5' cap and a 3' poly-A tail. The 5' cap is a methylated guanine nucleotide stuck on the 5' end [which is the end made \(5'\) or \(3'\)]\(^9\). The poly-A tail is a string of several hundred adenine nucleotides. The cap is essential for translation, while both the cap and the poly-A tail are important in preventing digestion of the mRNA by exonucleases that are free in the cell.

---

\(^9\) It is made first, since transcription proceeds from 5' to 3'.
• Why would active exonucleases be floating free in the cell?36

Figure 30  Comparison of Prokaryotic and Eukaryotic Gene Expression

• One piece of RNA isolated from a human cell is found to produce two different polypeptides when added to a cell-free protein synthesis system containing all the enzymes necessary for eukaryotic gene expression. When the two polypeptides are separated and digested with trypsin, they produce fragments of the following molecular weights:
  Polypeptide 1: 5 kD, 8 kD, 12 kD, and 14 kD
  Polypeptide 2: 3 kD, 5 kD, 8 kD, 10 kD, 12 kD, and 14 kD
How can we explain the synthesis of two different polypeptides from one piece of RNA?37

RNA Polymerase
In prokaryotes, all RNA is made by the $\alpha\beta\beta'\sigma$ RNA polymerase complex. In eukaryotes, there are many different RNA polymerases:

• RNA polymerase I transcribes most rRNA
• RNA polymerase II transcribes hnRNA (so ultimately mRNA), most snRNA, and some miRNA
• RNA polymerase III transcribes tRNA, long ncRNA, siRNA, some miRNA, and a subset of rRNA

Please note: In our discussion of replication you learned about many prokaryotic DNA polymerases. In contrast, here you learned about many eukaryotic RNA polymerases. Don’t get mixed up!

36 Two conceivable reasons: 1) mRNA has a very short lifespan; it is degraded rapidly, and more must be made if the protein is still needed. Note that this is consistent with the idea that regulation of gene expression occurs primarily at the transcriptional level since this is more efficient. 2) Viruses may inject RNA into the cell. If it does not have the correct cap and tail modifications, exonucleases will destroy it.

37 Here is an example of the use of splicing for the regulation of gene expression. The piece of RNA must have been hnRNA. In the cell-free system it underwent differential splicing to produce one of two different mRNA molecules. Apparently, Polypeptide 1 came from an mRNA which had more material spliced out than the mRNA coding for Polypeptide 2.
4.8 GENE EXPRESSION: TRANSLATION

Translation is the synthesis of polypeptides according to the amino acid sequence dictated by the sequence of codons in mRNA. During translation, an mRNA molecule attaches to a ribosome at a specific codon, and the appropriate amino acid is delivered by a tRNA molecule. Then the second amino acid is delivered by another tRNA. Then the ribosome binds the two amino acids together, creating a dipeptide. This process is repeated until the polypeptide is complete, at which point the ribosome drops the mRNA and the new polypeptide departs.

Transfer RNA (tRNA)

Each tRNA is composed of a single transcript produced by RNA polymerase III. The tertiary structure of every tRNA molecule is similar. tRNAs have a stem-and-loop structure stabilized by hydrogen bonds between bases on neighboring segments of the RNA chain (Figures 31 and 32). Several modified nucleotides are found in tRNA (e.g., dihydrouridine). One end of the structure is responsible for recognizing the mRNA codon to be translated. This is the anticodon, a sequence of three ribonucleotides which is complementary to the mRNA codon the tRNA translates. A key step in translation is specific base pairing between the tRNA anticodon and the mRNA codon. It is this specificity that dictates which amino acid of the twenty will be added to a growing polypeptide chain by the ribosome. [Is it likely that the three nucleotides of the anticodon contribute to the tertiary structure of tRNA by base-pairing with other nucleotides in the chain?] The other end of the tRNA molecule has the amino acid acceptor site, which is where the amino acid is attached to the tRNA. [If you analyzed a thousand tRNA molecules, which region would you expect to vary the most?] Since there is a tRNA for each codon, each tRNA is specific for one amino acid, while each amino acid may have several tRNAs. Each tRNA can be named according to the amino acid it’s specific for. For example, a tRNA for valine would be written tRNA_{Val}. When the amino acid is attached, the tRNA is written this way: Val-tRNA_{Val}.

---

38 No. They must be available for base pairing with the codon.

39 The anticodon is different for each of the different tRNA molecules. Part of the rest of the molecule varies from one tRNA to the next, but about 60 percent is constant. The amino acid binding site is always the same: CCA (at the 3' end of the tRNA molecule).
tRNA molecules often contain nitrogenous bases in many positions that have been covalently modified. Base methylation is particularly common. Some specific examples are inosine (derived from adenine), pseudouridine (derived from uracil), or lysidine (derived from cytosine). Inosine in particular plays an important role in wobble base pairing.

**The Wobble Hypothesis**

Using the standard genetic code, you would guess that organisms have 61 distinct tRNA molecules to recognize the 61 amino acid-coding codons possible in mRNA. In actual fact, most organisms have fewer than 45 different types of tRNAs, meaning some anticodons must pair with more than one codon. Francis Crick’s Wobble Hypothesis explains this and states that the first two codon-anticodon pairs obey normal base pairing rules, but the third position is more flexible (Figure 33). This allows for non-traditional pairing and explains why a smaller number of tRNAs are possible.

![Figure 33 Wobble Base Pairing Between a tRNA Anticodon and an mRNA Codon](Image)

A modified inosine base (I) at the 5’ end of the anticodon is particularly wobbly, as it can bond to three different codon bases (A, U or C). Some common wobble pairing combinations are:

<table>
<thead>
<tr>
<th>5’ Base in Anticodon [tRNA]</th>
<th>3’ Base in Codon [mRNA]</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>C (Watson-Crick base) or U (wobble base)</td>
</tr>
<tr>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>A</td>
<td>U</td>
</tr>
<tr>
<td>U</td>
<td>A (Watson-Crick base) or G (wobble base)</td>
</tr>
<tr>
<td>I</td>
<td>A, U or C (all wobble bases)</td>
</tr>
</tbody>
</table>
In other words, the most common wobble base pairs are guanine-uracil, inosine-uracil, inosine-adenine, and inosine-cytosine (G-U, I-U, I-A and I-C). Both the wobble base pair and the normal Watson-Crick base pair have similar thermodynamic stabilities.

**Amino Acid Activation**

Peptide bond formation during protein synthesis is a process that requires a lot of energy because the peptide bond has unfavorable thermodynamics ($\Delta G > 0$) and slow kinetics (high activation energy). Reaction coupling is used to power the process: two high-energy phosphate bonds are hydrolyzed to provide the energy to attach an amino acid to its tRNA molecule. This process is called tRNA loading or amino acid activation, and is useful because breaking the aminoacyl-tRNA bond will drive peptide bond formation forward. Amino acid activation occurs in several steps:

1) An amino acid is attached to AMP to form **aminoacyl** AMP. In this reaction, the nucleophile is the acidic oxygen of the amino acid, and the leaving group is PP$_i$.
2) The pyrophosphate leaving group is hydrolyzed to 2 orthophosphates. This reaction is highly favorable ($\Delta G << 0$).
3) tRNA loading, an unfavorable reaction, is driven forward by the destruction of the high-energy aminoacyl—AMP bond created in Step 1.

![Figure 34](image_url)  
*Figure 34*  Amino Acid Activation as an Example of Reaction Coupling  
*Note: water as a reactant has been left out of all reactions in this figure.*

Overall, amino acid activation requires 2 ATP equivalents because it uses two high-energy bonds. An ATP equivalent is a single high-energy phosphate bond. You can get 2 ATP equivalents by hydrolyzing 2 ATP to 2 ADP + 2 P$_i$ or by hydrolyzing 1 ATP to AMP + 2 P$_i$.
Eventually, the bond between the amino acid and the tRNA molecule will be broken. This hydrolysis will power peptide bond formation: the nitrogen of another amino acid will nucleophilically attack the carbonyl carbon of this amino acid, and tRNA will be the leaving group.

**Aminoacyl-tRNA Synthetases**

We have stated that incorporation of the appropriate amino acid in a growing polypeptide depends on the delivery of the correct amino acid by a specific tRNA. But we also noted that the amino acid acceptor sites of all tRNA molecules are the same. How is the attachment of the appropriate amino acid to each tRNA molecule accomplished? **Aminoacyl-tRNA synthetase enzymes** are specific to each amino acid, and there is at least one aminoacyl-tRNA synthetase for every amino acid. This family of enzymes recognizes both the tRNA and the amino acid based on their three-dimensional structures. They are highly specific, which is important because joining the wrong amino acid to a tRNA would result in the wrong amino acid being incorporated into a polypeptide. Given that some amino acids differ only by a single methyl group, this specificity is quite amazing. Aminoacyl-tRNA synthetases also function with a very low error rate. If there is a 1/1000 error rate in amino acid incorporation, what percentage of polypeptides that are 500 amino acid residues long will not contain any errors? 40

Overall then, amino acid activation serves two functions. One is specific and accurate amino acid delivery, and the other is thermodynamic activation of the amino acid.

- A bacterial strain with a point mutation in the gene for hexokinase is not able to metabolize glucose. The mutation causes a substitution of arginine for serine. These bacteria are used to test whether chemicals are mutagenic. The chemical is added to a culture of bacteria with glucose as the only carbon source. Any bacteria that grow must have undergone a mutation which remedied the problem (this is called suppression of the original mutation). When a particular hair spray ingredient is tested, several colonies grow on the glucose-only medium. Which one of the following might act as a suppressor of the first mutation? 41
  - A) A point mutation during replication of a tRNA gene
  - B) A mutation in RNA polymerase that increases the rate of promoter recognition
  - C) A base pair deletion in the hexokinase gene
  - D) A point mutation during transcription of a tRNA molecule

40 The easiest way to calculate this is to figure out the probability of getting all amino acids in the protein correct, in other words, we must use the non-error rate for our calculation, not the error rate. If the error rate is 1/1000, then the non-error rate is 999/1000. The probability of having no errors is .999^n, where n = the number of amino acid residues. In other words, a single amino acid has .999 probability, or 99.9% probability of being correct. Two amino acids correct in a row have a .999 x .999 probability (.999^2), or .998, or 99.8% probability of happening. Continuing in this manner, a 500-amino acid protein has a .999^500 probability of being entirely correct, or .606, approximately a 60% probability. Longer proteins have a higher chance of containing errors.

41 A single base change in the anticodon of the tRNA for arginine could cause it to recognize the codon for serine. If that happened in the mutant bacteria, problems might ensue, but one good result would be that the correct amino acid would be incorporated at the mutated site in hexokinase (choice A is correct; note that point mutations in tRNA genes are actually a common means of suppression in bacteria). Increasing the rate at which RNA polymerase recognizes the promoter might increase the rate of transcription, but would not fix a mutant enzyme (choice B is wrong), and a base pair deletion in the hexokinase gene would cause a frameshift mutation and a serious significant change in protein structure and function (choice C is wrong). A point mutation during transcription of a tRNA molecule might have a temporary effect on a single bacterium, but would not be passed on to its progeny; remember than only DNA mutations have lasting effects and errors made during transcription are generally insignificant (choice D is wrong).
The Ribosome

The ribosome is composed of many polypeptides and rRNA chains held together in a massive quaternary structure. Ribosomes float around in the cytoplasm, and each has a small subunit and a large subunit. The unit of measurement is the Svedberg, or S, unit. Svedbergs are a sedimentation rate, that is, how quickly something will sink in a gradient during centrifugation, and the units are not additive.

The prokaryotic ribosome sediments in a gradient at a rate of 70S, so it is referred to as the 70S ribosome (Figure 35). It is composed of a 30S small subunit and a 50S large subunit. The small subunit is made of a 16S rRNA and 21 peptides. Two rRNA molecules (23S and 5S) and 31 peptides make up the large subunit.

Eukaryotes have an 80S ribosome. It also has a small and large subunit. The large subunit has three rRNA molecules (5S, 5.8S, and 28S) and 46 peptides, and sediments in a gradient at a rate of 60S. The small subunit has 33 peptides and one rRNA (18S) and sediments in a gradient at a rate of 40S.

The 23S rRNA in prokaryotes and the 28S rRNA in eukaryotes have ribozyme function. They help link amino acids during protein synthesis via peptidyl transferase activity. This contributes to peptide bond formation. Notice how the ribozymic activity of the ribosome is found in the large subunit of both prokaryotic and eukaryotic ribosomes.

**Figure 35** Ribosome Components
In both prokaryotes and eukaryotes, the complete ribosome (both subunits together) has three special binding sites. The A site (aminoacyl-tRNA site) is where each new tRNA delivers its amino acid. The P site (peptidyl-tRNA site) is where the growing polypeptide chain, still attached to a tRNA, is located during translation. The E site (exit-tRNA site) is where a now-empty tRNA sits prior to its release from the ribosome. [During translation, the next codon to be translated is exposed in the __ 42.] tRNAs move through the sites from A → P → E.

**Figure 36**  The Ribosome

### Prokaryotic Translation

In prokaryotes, translation occurs in the same compartment and at the same time as transcription. In other words, *while the mRNA is being made* ribosomes attach and begin translating it. [Does this mean that the first end of the mRNA to be translated is 5' or 3'?43] Note that it says ribosomes above. Several ribosomes attach to the mRNA and translate it simultaneously (see Figure 37; you may hear the term *polyribosome* used to describe this arrangement; polyribosomes are seen in both prokaryotes and eukaryotes). [You figured out the direction of translation on the mRNA from what you already know. Do you have any previous knowledge that would help you answer this: Does translation always begin at the 5' end of the mRNA, or somewhere up the chain?44]

![Prokaryotic Polyribosome](image)

**Figure 37**  A Prokaryotic Polyribosome

---

42 A site, since this is where the next amino acid to be added must bind.
43 5' first, since the mRNA is made 5' end first. Transcription and translation go in the same direction on mRNA.
44 It does not always occur at the very end. You can deduce this from the fact that mRNA is polycistronic. If there are more than one translation start site on the mRNA, they can’t all be at the 5' end.
Because prokaryotes often have polycistronic mRNAs, their ribosomes can also start translation in the middle of the chain. This means termination and initiation sequences are found between each ORF. Even for the first open reading frame on a transcript, translation doesn’t begin right at the 5’ end. An upstream regulatory sequence is essential for initiation, just as in transcription. Here, instead of a promoter, we have a ribosome binding site, also known as the Shine-Dalgarno sequence, located at –10 (ten ribonucleotides upstream, or on the 5’ side of the start codon). The Shine-Dalgarno sequence is complementary to a pyrimidine-rich region on the small subunit, and thus helps position the initiation machinery on the transcript.

Like transcription, translation has three distinct stages: initiation, elongation, and termination. Many antibiotics function by inhibiting a particular stage.

**Initiation** starts with the small ribosomal subunit (30S) binding two initiation proteins called IF1 and IF3. This complex then binds the mRNA transcript. Next, the first aminoacyl-tRNA joins, along with a third initiation factor called IF2, which is also bound to one GTP. Finally, the 50S subunit completes the complex. This process is powered by the hydrolysis of one GTP molecule. The first aminoacyl-tRNA is special; it is called the initiator tRNA, abbreviated fMet-tRNA$_{fMet}$. The “fMet” stands for formylmethionine, which is a modified methionine used as the first amino acid in all prokaryotic proteins. The initiator tRNA sits in the P site of the 70S ribosome, hydrogen-bonded with the start codon. Before elongation, all initiation factors dissociate from the complex.

**Elongation**, a three-step cycle, may now begin. In the first step, the second aminoacyl-tRNA enters the A site and hydrogen bonds with the second codon. This process requires the hydrolysis of one phosphate from GTP. This is done by an elongation factor protein called Tu (EF-Tu), which is a GTPase. A second elongation factor called EF-Ts removes the remaining GDP from EF-Tu, thus helping it reset. In the second step, the peptidyl transferase activity of the large ribosomal subunit (the 23S rRNA) catalyzes the formation of a peptide bond between fMet and the second amino acid. The amino group of amino acid #2 acts as nucleophile, and tRNA$_{fMet}$ is the leaving group; it dissociates from the ribosome. A new dipeptide is now attached to tRNA #2. Now you can figure out the direction of translation from the point of view of the polypeptide; you won’t have to memorize it. The third step is translocation, in which tRNA #1 (now empty) moves into the E site, tRNA #2 (holding the growing peptide) moves into the P site, and the next codon to be translated moves into the A site. Elongation factor G (EF-G) helps with translocation, and this process costs one GTP. EF-G is sometimes called a translase because of its function in this step. The new dipeptide is still attached to tRNA #2, and tRNA #2 is still H-bonded to codon #2. The presence of tRNA #1 in the E site (still H-bonded to codon #1), is thought to help maintain the reading frame of the mRNA (disruption of tRNA binding to the E site results in an increase in the number of frameshift mutations in the resulting protein). EF-Tu eventually helps remove this tRNA from the E site.

---

45 For example, streptomycin and tetracycline bind to the 30S subunit of the prokaryotic ribosome. Chloramphenicol and erythromycin bind to the 50S subunit.
46 This may seem odd, as ATP is normally the energy molecule. But a high-energy phosphate is a high-energy phosphate. Another example is the GTP produced in the Krebs cycle.
47 In fact, cells of our immune system release cytotoxins when they sniff out fMet, because this chemical is a sure sign that bacteria are busily translating.
48 Refer to the genetic code table. The codon for methionine is AUG; that’s the start codon. It only initiates translation when it is preceded by a Shine-Dalgarno sequence (prokaryotes).
49 The direction of synthesis is N → C, since the N of amino acid #2 binds to the C of #1. As the polypeptide elongates, its N terminus will come snaking out of the ribosome.
Does the ribosome move relative to the mRNA during translocation? These three steps repeat over and over again, connecting amino acids in the order their codons appear along the mRNA strand (and thus appear in the A site).

**Termination** occurs when a stop codon appears in the A site. Instead of a tRNA, a *release factor* now enters the A site. This causes the peptidyl transferase to hydrolyze the bond between the last tRNA and the completed polypeptide. Prokaryotes have three release factor proteins, which mediate translation termination by recognizing stop codons. RF1 recognizes termination codons UAA and UAG, and RF2 recognizes UAA and UGA. RF3 is a GTP-binding protein that doesn’t recognize a stop codon, but instead leads to the dissociation of RF1/RF2 after peptide release. Finally, the ribosome separates into its subunits and releases both mRNA and polypeptide.

![Figure 38 Translation Elongation](image)

Let’s focus for a moment on the energetics of translation. Why doesn’t peptide bond formation require GTP hydrolysis, like the other steps in translation? You should be able to answer questions like this: How many high-energy phosphate bonds are required to make a 50 amino acid polypeptide chain, including the energy used to activate amino acids to aminoacyl-tRNAs?

**Eukaryotic Translation**

There are several differences between eukaryotic and prokaryotic translation. Many of these have already been mentioned: The ribosome is larger (80S) and has different components than the prokaryotic ribosome, the mRNA must be processed before it can be translated (spliced, with cap and tail added), and the N-terminal amino acid is different (Met instead of fMet). Also remember that eukaryotic mRNA must...

---

50 It must, if the tRNA remains H-bonded to the mRNA while moving to another spot in the ribosome.
51 Because the bond between each amino acid and its tRNA is a high-energy bond whose hydrolysis drives peptide bond formation. Remember that the aminoacyl-tRNA bond was formed using the energy of two phosphate bonds from ATP.
52 There are two phosphate bonds hydrolyzed per amino acid to make the aminoacyl-tRNAs, or 100 for the 50 amino acid polypeptide. Two phosphate bonds are required for each elongation step, one for the entrance of each new aminoacyl-tRNA into the ribosomal A site and the other for translocation. Since there are 49 elongation steps for a 50-amino acid protein, 98 high-energy bonds are hydrolyzed during elongation. Finally, one GTP is hydrolyzed during initiation to position the first tRNA and mRNA on the ribosome, and one GTP is hydrolyzed in termination. Thus, a total of 200 high-energy bonds are required for the translation of a 50-amino acid protein. In other words, it costs 4n high-energy bonds to make a peptide chain, where n is the number of amino acids in the chain.
not only be spliced, capped, and tailed, but it also requires transport from nucleus to cytoplasm, thus transcription and translation cannot proceed simultaneously.

Eukaryotes do not use the Shine-Dalgarno sequence to initiate translation. There are 5’ UTR sequences in eukaryotes that function in starting translation; a common one is the Kozak sequence, which is a consensus sequence typically located a few nucleotides before the start codon.

Eukaryotic translation begins with formation of the initiation complex. First, a 43S pre-initiation complex forms, composed of the 40S small ribosomal subunit, Met-tRNA\textsubscript{Met}, and several proteins called eukaryotic initiation factors (or eIFs). Next, this assembled complex is recruited to the 5’ capped end of the transcript, by an initiation complex of proteins (including other eIF proteins). Additional proteins are recruited (such as a polyA tail binding protein) and the initiation complex starts scanning the mRNA from the 5’ end, looking for a start codon. Once the start codon has been found, the large ribosomal subunit (60S) is recruited and translation can begin.

Some eIF proteins are essential to initiate translation and others help regulate the process. For example, eIF3 binds the small ribosomal subunit and prevents it from prematurely associating with the 60S subunit. The amount of eIF proteins in the cell is closely controlled, and this affects the amount of translation occurring. eIF4A is a helicase and unwinds mRNA, eIF4E binds the 5’ cap of the mRNA and eIF4G is a scaffold protein. Each of these three function in the initiation complex, and their levels are a rate-limiting step for translation. Higher amounts of these three proteins means the cell can perform more translation, while a lower amount decreases translation. Activity of eIF proteins is controlled by post-translational modification, such as phosphorylation. This couples translation to upstream cell signaling pathways.

Eukaryotes have two elongation factors. eEF-1 has two subunits, one that helps with entry of an aminoacyl tRNA into the A site and one that is a guanine nucleotide exchange factor, catalyzing the release of GDP. The eukaryotic translocase is called eEF-2. Additional elongation factors are required to facilitate peptide bond formation.

The order in which the initiation complex is formed is different in eukaryotes. \[\text{Are the nascent (newly formed) polypeptide chains emerging from a polyribosome in a eukaryote all the same?} \]

Eukaryotic translation termination involves two release factors. eRF1 recognizes all three termination codons, and eRF3 is a ribosome-dependent GTPase that helps eRF1 release the completed polypeptide.

- Which one of the following pairs of processes may occur simultaneously on the same RNA molecule in a eukaryotic cell?\[35\]
  A) Translation and transcription
  B) Transcription and splicing
  C) Splicing and translation
  D) Messenger RNA degradation and transcription

\[35\] In eukaryotes, the answer is: yes, always, because eukaryotic mRNA is monocistronic. In prokaryotes, however, different polypeptides may be translated from a single piece of mRNA, since prokaryotic mRNA is polycistronic.

\[34\] In order for processes in eukaryotes to occur simultaneously, they must occur in the same compartment. Transcription and splicing both occur in the nucleus and could therefore occur simultaneously (choice B is correct). Translation occurs in the cytoplasm, while transcription and splicing occur in the nucleus, thus translation cannot occur at the same time as either of these processes (choices A and C are wrong). mRNA degradation and transcription cannot occur at the same time; if this were true, no mRNA molecules would survive to be translated (choice D is wrong).
**Cap-Independent Translation**

It was long thought that all eukaryotic translation started at the 5' end of an mRNA. In other words, all eukaryotic transcripts were assumed to be monocistronic, and coded for only one polypeptide chain. It is true that this mechanism is by far the major one in eukaryotic cells. Because of the important role of 5' mRNA cap recognition, it's called *cap-dependent translation*.

However, it’s recently been discovered that eukaryotes are sometimes capable of starting translation in the middle of an mRNA molecule, a process called *cap-independent translation* (because the beginning of translation doesn’t require the 5’ cap of the mRNA). To do this, the transcript must have an internal ribosome entry site, or IRES. This is a specialized nucleotide sequence, and was first discovered in viruses. Since then, IRESs have been found in a number of eukaryotic transcripts. Most code for proteins that help the cell deal with stress, or help activate apoptosis. In other words, the IRESs found so far make sure the cell can make essential proteins when under sub-optimal growth conditions. Cells under stress generally inhibit translation (via inhibiting translation initiation), and cap-independent translation allows the cell to make proteins when doing so is crucial for survival or programmed cell death. Activation of translation using an IRES requires different proteins than normal initiation.

Additional nucleotide sequences have been identified, which allow cap-independent translation in eukaryotes. While some of these are used in molecular biology labs, it’s unclear how or if they function in normal eukaryotic cells.

---

**4.9 CONTROLLING GENE EXPRESSION**

Adult humans have over 220 different types of cells, all with the same genome, but with different attributes such as morphology, lifespan, function, ability to secrete, response to signaling molecules, mobility, etc. These changes are due to differences in gene expression and protein function. In each cell type, some genes are expressed and others are silenced, further, genes that are expressed can have different levels of expression, where in one cell type the gene is expressed at a high level (to produce lots of ncRNA or protein), and in a different cell type the same gene is expressed at a low level. They can also have varying activity, stability and half-life. These variations in gene expression can be altered using many different mechanisms:

![Figure 39](image)

*Figure 39  Mechanisms of Controlling Gene Expression in Eukaryotes*
Transcription is the principle site of the regulation of gene expression in both eukaryotes and prokaryotes. This means that the amount of each protein made in every cell is affected by the amount of mRNA that gets transcribed. Gene expression can also be controlled epigenetically. Broadly speaking, epigenetics focuses on changes in gene expression that are not due to changes in DNA sequences, but are either heritable or have a long-term effect. The three most commonly studied areas in this field are DNA methylation, chromatin remodeling, and RNA interference. Let’s look at the regulation of gene expression, starting with DNA and working down toward proteins.

**Controlling Gene Expression at the DNA Level**

**DNA Methylation and Chromatin Remodeling**

Both prokaryotic and eukaryotic DNA can be covalently modified by adding a methyl group. Bacteria methylate new DNA shortly after synthesis, and the brief delay is useful in mismatch repair pathways (see above). Methylation can also control gene expression in prokaryotes, either by promoting or inhibiting transcription.

Eukaryotic DNA methylation has been found in every vertebrate genome studied so far. Broadly speaking, it plays an important role in controlling gene expression (especially during embryonic development), and has also been implicated in several diseases. DNA methylation turns off eukaryotic gene expression two ways:

1) Methylation physically blocks the gene from transcriptional proteins.
2) Certain proteins bind methylated CpG groups and recruit chromatin remodeling proteins that change the winding of DNA around histones.

- Regulation of a gene is examined *in vitro* in the presence and absence of chromatin assembly, and in the presence and absence of a sequence-specific regulator of transcription. Transcription is quantitated after the experiment and the following results are obtained:

<table>
<thead>
<tr>
<th>Sequence-Specific Factor</th>
<th>DNA</th>
<th>Relative Amount of Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. None</td>
<td>unpackaged</td>
<td>0.74</td>
</tr>
<tr>
<td>2. None</td>
<td>packaged</td>
<td>0.07</td>
</tr>
<tr>
<td>3. Present</td>
<td>unpackaged</td>
<td>1.0</td>
</tr>
<tr>
<td>4. Present</td>
<td>packaged</td>
<td>0.59</td>
</tr>
</tbody>
</table>
Which one of the following conclusions can be drawn from this experiment?\(^5\)

A) The degree of activation by the sequence-specific factor is greater in the presence of chromatin assembly than in its absence.
B) The sequence-specific factor acts to repress transcription.
C) The histones increase the rate of transcription.
D) The sequence-specific factor increases the rate of transition from a closed complex to an open complex.

**Gene Dose**
One way to increase gene expression is to increase the copy number of a gene by amplification. Increasing gene dose will allow a cell to make large quantities of the corresponding protein. Similarly, gene deletion causes a decrease in gene expression. Both are examples of copy number variation, discussed earlier in Section 4.2.

**Imprinting**
Genomic imprinting is when only one allele of a gene is expressed. In some situations, the maternal allele is expressed, and in others the paternal allele is expressed. Imprinted genes tend to be clustered together on chromosomes. Imprinting is a dynamic process and can change from generation to generation. In other words, a gene that is imprinted in an adult may be "unimprinted" and expressed in that adult’s offspring. This observation led to the notion that imprinting is an epigenetic process. Silencing of a certain gene involves DNA methylation, histone modification, and binding of long ncRNAs. These epigenetic marks are established in the germline and are maintained throughout life and mitotic divisions.

**X Chromosome Inactivation**
Female mammals have two X chromosomes, one of which is active (called Xa) and one of which is silenced, or inactive (and is called Xi). In humans, X-inactivation occurs early in development, at the blastocyst stage (Chapter 13). Each cell in the inner cell mass randomly inactivates an X chromosome, and this decision is irreversible. This means every cell derived from each cell in the inner cell mass will have the same X chromosome inactivated, however, because each cell makes its own decision, an adult can have different X chromosomes inactivated in different tissues and cells. Because of X-inactivation, all humans have the same number of gene products for the X chromosome; males have only one X chromosome, and females have only one active X chromosome. Not all animals behave the same when it comes to X-inactivation. Some animals (such as marsupials) consistently silence one X chromosome; in the case of marsupials, the paternally derived X chromosome is inactivated and the maternal X chromosome is active. Xi is very condensed, and packaged in heterochromatin. It has high levels of DNA methylation.

\(^5\) A quick glance at the data indicates that transcription is increased in the presence of the sequence-specific factor (compare lines 1 and 2 with lines 3 and 4, choice B is wrong), and that histones decrease the rate of transcription (packaged DNA has a lower rate of transcription than unpackaged, choice C is wrong). Looking closer, it appears that the sequence specific factor causes an approximate 8-fold increase in the transcription rate of packaged DNA (compare lines 2 and 4), but doesn’t even double the rate of transcription of unpackaged DNA (compare lines 1 and 3). It might be that this occurs because the factor increases the rate of transition to an open complex, but there is no data to support this (choice A is a better answer than choice D). Don’t confuse “open complex” (which means separated DNA strands) with “unpackaged” (which means not wrapped around histones).
Controlling Gene Expression at the RNA Level: Regulation of Transcription in Prokaryotes

Regulation of transcription is the primary method of regulation of gene expression in prokaryotes. One simple mechanism of transcriptional regulation in bacteria is that some promoters are simply stronger than others. The problem with this mechanism of regulation is that it is “pre-set” and cannot respond to changing conditions within the cell. Bacteria also possess far more complex regulatory mechanisms, which activate or suppress transcription depending on current needs for specific gene products. For example, bacteria only produce the enzyme β-galactosidase and other proteins required for lactose catabolism when lactose is present. [Assuming these protein products do not have a harmful effect on the cell, what advantage might there be in turning off the genes when the protein products are not required?]

- Are the terms polypeptide enzyme and gene product synonymous? Or are there gene products that are not polypeptide enzymes? Are there polypeptides which are not enzymes?

Enzymes involved in anabolism (biosynthesis) should be produced when the item they help make (their product) is scarce. Enzymes involved in catabolism (degradative metabolism) should be produced when the item they help break down (their substrate) is abundant, such as food. Thus there are two basic ways we can imagine how transcription is regulated. The transcription of enzymes involved in biosynthetic pathways should be inhibited by their product. The transcription of enzymes involved in catabolic pathways should be automatically inhibited whenever the substrate is not around, and activated when it is. That is in fact exactly what happens. Anabolic enzymes whose transcription is inhibited in the presence of excess amounts of product are repressible. Catabolic enzymes whose transcription can be stimulated by the abundance of a substrate are called inducible enzymes.

There are two common examples of this. The lac operon is inducible, since the enzymes it codes for are part of lactose catabolism, and the trp operon is repressible, since the enzymes it codes for mediate tryptophan biosynthesis or anabolism. An operon has two components, a coding sequence for enzymes, and upstream regulatory sequences or control sites. Operons may also include genes for regulatory proteins, such as repressors or activators, but don’t have to. These genes can be located elsewhere in the genome and typically have their own promoters.

The Lac Operon

The lac operon contains several components:

1) \(P\) region: the promoter site on DNA to which RNA polymerase binds to initiate transcription of \(Y, Z,\) and \(A\) genes
2) \(O\) region: the operator site to which the Lac repressor binds
3) \(Z\) gene: codes for the enzyme β-galactosidase, which cleaves lactose into glucose and galactose
4) \(Y\) gene: codes for permease, a protein which transports lactose into the cell

---

56 It takes a great deal of ATP to synthesize RNA and protein, so it’s more energy-efficient to transcribe and translate only the proteins that are needed.

57 They are not synonymous. All polypeptides are gene products, but some gene products are not polypeptides and some polypeptides are not enzymes. Transfer RNA and rRNA are gene products, but not polypeptides. Microfilaments and other elements of the cytoskeleton, as well as collagen and many other polypeptides, are not enzymes.

58 So note: The default for repressible systems is “ON”; for inducible systems the default is “OFF.”
5) A gene: codes for transacetylase, an enzyme which transfers an acetyl group from acetyl-CoA to β-galactosides (note that this function is not required for lactose metabolism)

Additionally, there are two genes, each with their own promoter, that code for proteins important in the regulation of the lac operon:

1) crp gene: located at a distant site, this gene codes for a catabolite activator protein (CAP) and helps couple the lac operon to glucose levels in the cell
2) I gene: located at a distant site, this gene codes for the Lac repressor protein

So overall, there are five protein coding genes and two regulatory sequences. Both crp and I have their own promoters. The protein products of these two genes control gene expression of Z, Y and A.

Bacterial cells preferentially use glucose as an energy source. This means that in the presence of glucose, the lac operon will be off, or expressed at low amounts (see Figures 40 and 41). This is mediated by the CAP and repressor proteins. Glucose levels control a protein called adenylyl cyclase, which converts ATP to cAMP. In high glucose conditions, adenylyl cyclase is inactivated and cAMP levels are very low. In low glucose conditions, the opposite is true: adenylyl cyclase is activated and cAMP levels are high. CAP binds cAMP and this complex binds the promoter of the lac operon (Figure 42). This helps activate RNA polymerase at the lac operon and contributes to the operon being turned on when glucose levels are low.

The I gene codes for a repressor protein, which binds the operator of the lac operon. This prevents RNA polymerase from binding the promoter and transcribing Z, Y, and A genes, thereby blocking transcription of the operon when lactose is absent (Figure 40). The repressor protein can also bind lactose, and this blocks its activity on the operator. This binding is allosteric, meaning it happens at a distant site from operator binding. It causes a conformational change in the tertiary structure of the repressor protein, such that it is no longer capable of binding to the operator. As a consequence, it falls off the DNA (Figures 41 and 42).

High transcription of Z, Y, and A genes occurs when glucose is absent and lactose is present (Figure 42). Low glucose results in an increased amount of cAMP, which binds to CAP and helps activate RNA polymerase activity at the lac operon. Lactose presence means the Lac repressor protein is unable to bind the lac operator and negatively regulate transcription; thus the polycistronic mRNA is transcribed at high levels. When the supply of lactose becomes very scarce, there isn’t enough to bind to the repressors, and most of the repressor proteins return to their original structure. They now rebind to the operator, decreasing transcription of Z, Y, and A genes.
Figure 40  The Lac Operon in the Presence of Glucose and Absence of Lactose

Figure 41  The Lac Operon in the Presence of both Glucose and Lactose
If the operator is mutated so that the lac repressor can no longer bind, what effect will this have on transcription?59

A) Transcription of Gene Z will be activated, and Genes Y and A will not be affected.
B) None of the genes will be transcribed, regardless of the presence or absence of lac repressor.
C) Transcription will still be activated by lactose.
D) All three genes will be expressed constitutively, regardless of the presence of lactose.

The Trp Operon

Bacteria use a five enzyme synthetic pathway to make the amino acid tryptophan from chorismic acid. In the presence of tryptophan, there is little point in making these enzymes, which are also co-localized in an operon.

The repressor protein is coded by the trpR gene (Figure 43). The repressor binds tryptophan when it is present, and the two together then bind the operator, to turn off transcription of the other five trp genes.

59 If the repressor cannot bind to the operator, nothing will prevent RNA polymerase from transcribing all the genes on the operon in an unregulated, constitutive (or continuous) fashion (choice D is true and choice B is false). All genes on the operon are expressed or repressed together (choice A is false), and lactose will no longer have any effect (the expression of the genes is unregulated, so choice C is false).
In the absence of tryptophan, the bacterial cell must make its own. With no tryptophan present, the repressor protein cannot bind the operator. Without this block, RNA polymerase transcribes the five genes in the trp operon, and the five gene products allow the cell to make tryptophan. This is an example of anabolic repressible transcription.

**Figure 43** The Trp Operon in the Presence of Tryptophan

### Control of Gene Expression at the RNA Level: Regulation of Transcription in Eukaryotes

Given the complexity of eukaryotes compared to prokaryotes, it is not surprising that the regulation of eukaryotic transcription is also more complex. Most of this regulation happens at initiation.

For protein-coding genes, there are upstream control elements (UCEs), usually about 200 bases upstream of the initiation site, a core promoter containing binding sites for the basal transcription complex and RNA polymerase II (about 50 bases upstream of the transcription start site), and a TATA box at –25. The TATA box is a highly conserved DNA recognition sequence for the TATA box binding protein (TBP).

Binding of TBP to the TATA box initiates transcription complex assembly at the promoter.

Enhancer sequences in DNA are bound by activator proteins, and this is another kind of transcriptional regulation. The enhancer may be located many thousands of base pairs away from a promoter (either upstream or downstream) and still regulate transcription. This is likely done by DNA looping so enhancers and their activator proteins can get close to transcriptional machinery.

Eukaryotes also have gene repressor proteins, which inhibit transcription; this can also be done by modifying chromatin structure. Transcription factors have DNA-binding domains and are crucial in transcription regulation. They can bind promoters or other regulatory sequences. In fact, in many cases,
transcription levels in eukaryotes are controlled by huge committees of proteins. This produces a combinatorial effect, where each protein contributes to regulation, and can itself be regulated. These complex networks help link transcription to cell signaling and status. The binding of transcriptional machinery to DNA is often regulated by extracellular signals. For example, steroid hormones bind to receptors in the cell, and this sends the receptor to the nucleus. The complex binds DNA to regulate transcription. [If a mutation in a eukaryotic fat cell reduces the level of several proteins related to fat metabolism, does this mean the proteins are encoded by the same mRNA?]60

Beyond regulating the initiation of transcription, eukaryotes employ several other methods of transcriptional regulation, including:

- **RNA Translocation**: mRNA transcripts must be exported from the nucleus to the cytoplasm and can also be transported to different areas of the cell. They are translationally silent while this is happening. This system is especially important in cells that have a high level of polarity, where one area or end of the cell is distinctly different from the other. For example, neurons have polarity, and some transcripts are transported to the dendrites, while others stay in the soma. This is a way of controlling gene expression: mRNA transcripts aren’t translated into proteins until they are localized properly in the cell.

- **mRNA Surveillance**: Cells closely monitor mRNA molecules to ensure that only high-quality mRNA transcripts are read by the ribosome. Defective transcripts (such as those with premature stop codons, or those without stop codons at all) and stalled transcripts (where the ribosome is stalled in translation) are degraded.

- **RNA Interference**: RNA interference (RNAi) is a way to silence gene expression after a transcript has been made. It is mediated by miRNA and siRNA (Section 4.7). Molecular biology labs often use the RNAi system experimentally, as a way to decrease protein expression (see Appendix). Generally speaking, the siRNAs bind complementary sequences on mRNAs, and this ds-RNA is then degraded. The amount of transcript in the cell decreases, and gene expression is thus negatively regulated.

### Control of Gene Expression at the Protein Level:

#### Translation Initiation

We’ve already discussed the complex process of assembling translational machinery. In both prokaryotes and eukaryotes, this is a highly regulated process that links protein synthesis with upstream signaling pathways. Otherwise there is little control at the level of translation.

---

60 No, it does not. Eukaryotic mRNA is monocistronic. A more likely explanation is that a number of different genes located throughout the genome have related regulatory sequences that bind the same sequence-specific transcription factors. This is the means used by eukaryotes to achieve coordinated expression of genes. Related proteins are clumped together on the same piece of mRNA in prokaryotes only.
Post-Translational Modification

Newly synthesized proteins released from the ribosome are rarely able to function. They need to be correctly folded, modified or processed, and transported to where they function in the cell. These modifications are called post-translational events, since they occur after protein synthesis.

Protein Folding

First, the newly synthesized nascent protein is folded into its correct three-dimensional shape. This is accomplished by a family of proteins called chaperones. If folded correctly, the protein is said to be in its native conformation. If the protein is unfolded or misfolded, it’s said to be in its non-native state. Chaperone proteins are found across all types of organisms (from bacteria to plants to mammals), and also function in assembly or folding of other macromolecular structures. For example, chaperone proteins assist in nucleosome assembly from folded histones and DNA. In eukaryotic cells, chaperones are found in many subcellular compartments.

Covalent Modification

Many proteins are covalently modified. Some have hydrophobic groups added to facilitate membrane localization. For example, the addition of a fatty acid can target a protein to a membrane (either the plasma membrane or an organelle membrane).

Smaller chemical groups can also be added. For example, proteins can be:

- Acetylated: addition of an acetyl group (−C(O)CH₃), usually at the N-terminus of a protein, or at a lysine amino acid
- Formylated: addition of a formyl group (−C(O)H)
- Alkylated: addition of an alkyl group (such as methyl, ethyl, etc). Methylation is a common post-translational modification, and is usually done to lysine or arginine amino acids
- Glycosylated: addition of a glycosyl group to arginine, asparagine, cysteine, serine, threonine, tyrosine, or tryptophan amino acids. A glycosyl group is the substituent form of a cyclic mono-, di-, or oligosaccharide. This results in a glycoprotein.
- Phosphorylated: addition of a phosphate group (PO₄³⁻) to a serine, threonine, tyrosine, or histidine amino acid.
- Sulphated: addition of a sulphate group (SO₄²⁻) to a tyrosine amino acid.

Proteins can also be linked to other proteins. For example, in ubiquitination, proteins are covalently linked to ubiquitin.

There are many other examples of protein covalent modification. Overall, these modifications can have many effects on a protein and its function. They can change protein subcellular localization, target a protein for degradation, change interactions between proteins and other molecules, activate or inhibit enzyme activity, or change enzyme affinity for substrates. These modifications are typically studied in the lab using mass spectrometry (see MCAT Organic Chemistry Review), western blotting, or eastern blotting (see Appendix).
**Processing**

Many proteins require cleavage of some sort to become mature or functional. Cleavage can occur at either end of a peptide chain, or in the middle. Protein precursors are often used when the mature protein may be dangerous to the organism. Because the precursor is already made, it allows large quantities of mature protein to be available on short notice. Enzyme precursors are called *zymogens* or *proenzymes*.

A well-known example of post-translational processing is insulin. Insulin is made from a prohormone (Figure 44); preproinsulin is the primary translational product of the human *INS* gene. This peptide is 110 amino acids in length. To form proinsulin, an N-terminus signal peptide is removed and disulphide bonds form, in the endoplasmic reticulum. Three cleavage events are necessary to process proinsulin: the C peptide is removed by a family of enzymes called proprotein convertases, and a dipeptide fragment is removed from the C-terminus of the B chain peptide by a carboxypeptidase. These cleavage events occur in a secretory vesicle. The biological effects of insulin are well known, but it’s recently been shown that peptide C also has signaling properties.

![Figure 44 Insulin Processing: An Example of Post-Translational Modification](image-url)
4.10 RETURN TO GENE STRUCTURE: A SUMMARY

Now that we have been through all the processes that a cell uses to turn a gene into a protein, and control this process, let’s review the components once again (Figure 45). Transcription begins at a start site, but needs a promoter upstream of this. It ends at a termination signal. The RNA transcript contains the open reading frame (which goes from start codon to stop codon), as well as both 5’ and 3’ regulatory regions.
DNA replication, transcription, and translation have many similarities and some differences, and these are summarized in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>DNA Replication</th>
<th>Transcription</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal to get ready</td>
<td>ORI</td>
<td>Promoter</td>
<td>• Shine-Dalgarno (prok)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Kozak sequence (euk)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• these are found in 5'UTR (untranslated region)</td>
</tr>
<tr>
<td>Signal to start</td>
<td>ORI</td>
<td>Start site</td>
<td>AUG start codon</td>
</tr>
<tr>
<td>Key synthesis enzyme</td>
<td>DNA polymerase</td>
<td>RNA polymerase</td>
<td>Ribosome (made of rRNA and peptides)</td>
</tr>
<tr>
<td>Other important enzymes</td>
<td>Helicase</td>
<td>Spliceosome machinery</td>
<td>Aminoacyl tRNA synthetases</td>
</tr>
<tr>
<td></td>
<td>Topoisomerase</td>
<td></td>
<td>Initiation factors</td>
</tr>
<tr>
<td></td>
<td>SSBPs</td>
<td></td>
<td>Elongation factors</td>
</tr>
<tr>
<td></td>
<td>Primase</td>
<td></td>
<td>Release factors</td>
</tr>
<tr>
<td></td>
<td>Ligase</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Telomerase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Template molecule</td>
<td>DNA</td>
<td>DNA</td>
<td>mRNA</td>
</tr>
<tr>
<td>Read direction</td>
<td>3' to 5' on the DNA template</td>
<td>3' to 5' on the DNA template</td>
<td>5' to 3' on the RNA template</td>
</tr>
<tr>
<td>Molecule synthesized</td>
<td>DNA</td>
<td>RNA (mRNA in prok, hnRNA in euk)</td>
<td>Peptides</td>
</tr>
<tr>
<td>Build direction</td>
<td>5' to 3'</td>
<td>5' to 3'</td>
<td>N-terminus to C-terminus</td>
</tr>
<tr>
<td>Prokaryotic location</td>
<td>Cytoplasm</td>
<td>Cytoplasm</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>Eukaryotic location</td>
<td>Nucleus</td>
<td>Nucleus</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>Signal to stop</td>
<td>When the replication bubbles or newly synthesized strands meet and are ligated together</td>
<td>Transcription stop sequence or poly-A sequence</td>
<td>Stop codon (UAG, UGA, UAA)</td>
</tr>
</tbody>
</table>

Table 1  A Review of Molecular Biology Processes
Chapter 4 Summary

• DNA is the fundamental unit of inheritance in cells.

• DNA and RNA are polymers, made of nucleotide monomers. A nucleotide contains phosphate group(s), a sugar (either deoxyribose for DNA or ribose for RNA), and a nitrogenous base, either a purine (adenine or guanine) or a pyrimidine (thymine, cytosine, or uracil).

• In DNA, adenine always pairs with thymine via two hydrogen bonds, and cytosine always pairs with guanine via three hydrogen bonds.

• Uracil replaces thymine in RNA, and the ribose in RNA has an OH group on carbon 2.

• DNA is supercoiled in prokaryotes and packaged around histone proteins in eukaryotes.

• Eukaryotic DNA is divided into several linear chromosomes which have unique structures, including the long (q) and short (p) arms, centromere and telomeres on the ends.

• Genomes have extensive variation, including single nucleotide polymorphisms and copy number variation; transposons are mobile genetic elements which also contribute to genomic variation.

• Mutations can occur spontaneously, or can be caused by environmental factors [such as ionizing radiation], chemicals, or biological agents.

• Point mutations are classified based on their effect on the DNA (transition/transversion) or their effect on the amino acid sequence (missense, nonsense, or silent).

• Frameshift mutations are caused by insertions or deletions in the DNA base sequence that affect the reading frame of a gene. These are generally very serious mutations because they affect every amino acid codon from the point of the mutation on.

• Other types of mutations include inversions, translocations, and rearrangements.
• Transposons are mobile genetic elements that can “jump” from chromosome to chromosome. This can lead to deletions, insertions, and mutations. They consist of two inverted repeats flanking the gene for transposase and may include other genes as well.

• DNA replication occurs in the S-phase of the cell cycle and is semiconservative in nature.

• Several enzymes are involved in DNA replication. Helicases unwind the parental DNA at the origin of replication. Primases synthesize an RNA primer. DNA polymerase synthesizes new DNA, proofreads, and replaces the RNA primer. DNA ligase attaches the Okazaki fragments in the lagging strand.

• Cells have developed several ways to fix mutations, including: direct reversal, homology-dependent repair pathways (such as excision repair and post-replication repair), double-strand break repair (such as homologous recombination and nonhomologous end joining), and SOS repair.

• Transcription is the first part of protein synthesis; it is the creation of an RNA transcript by an RNA polymerase that reads the DNA template. Translation is the second part of protein synthesis; it is the creation of a polypeptide chain by ribosomes that read the mRNA transcript.

• There are several types of RNA that do not encode proteins. Some are directly involved in translation [rRNA and tRNA], while others play a role in gene expression [snRNA, miRNA, siRNA].

• Key info about prokaryotes: theta replication, genome is a single circular piece of DNA, three different DNA polymerases, one RNA polymerase, no mRNA processing, polycistronic mRNA, simultaneous transcription/translation, smaller ribosomes.

• Key info about eukaryotes: replication bubbles, genome is several linear pieces of DNA, one DNA polymerase, three RNA polymerases, capping, tailing, and splicing of mRNA prior to translation, monocistronic mRNA, transcription in nucleus, translation in cytosol, larger ribosomes.
CHAPTER 4 FREESTANDING PRACTICE QUESTIONS

1. A competitive inhibitor of eukaryotic RNA polymerase III would have the greatest effect on:
   A) replication.
   B) reverse transcription.
   C) translation.
   D) mutation.

2. In the lac operon, transcription is regulated by a repressor protein and only takes place in the presence of lactose. Which of the following statements is correct?
   A) The repressor protein binds to the promoter site to inhibit transcription.
   B) Lactose binds to the promoter site to initiate transcription.
   C) Lactose binds to the repressor protein to inhibit transcription.
   D) The repressor protein binds the operator site to inhibit transcription.

3. Which of the following could not be caused by a single point mutation in the DNA?
   A) Ala-Gln-Cys-Asp-Leu → Ala-Gln
   B) Ala-Gln-Cys-Asp-Leu → Ala-Gln-Cys-Asp-Leu
   C) Ala-Gln-Cys-Asp-Leu → Ala-Gln-Cys-His-Lys
   D) Ala-Gln-Cys-Asp-Leu → Ala-Gln-Cys-His-Leu

4. Which of the following is/are true with respect to eukaryotic mRNA?
   I. Monocistronic
   II. Transcription stops at the stop codon
   III. Has the same sequence as the template DNA that it was transcribed from
   A) I only
   B) I and II
   C) II and III
   D) I, II, and III

5. Which of the following is NOT a similarity between replication and transcription?
   A) Both processes occur with the same fidelity.
   B) Polymerization in both processes is based on reading a template.
   C) A pyrophosphate is removed from every nucleotide as polymerization occurs.
   D) Both processes occur in the 5’ to 3’ direction.

6. Which of the following functions is NOT typically attributed to small nuclear RNA (snRNA)?
   A) Processing of pre-mRNA
   B) Regulation of transcription factors
   C) Coordinating amino acid addition in translation
   D) Maintaining telomeres

7. Organisms with a higher degree of complexity do not necessarily have more diverse genomes than less complex organisms, in spite of the need for a greater diversity of proteins. Post-translational modification is one method used by more complex organisms to produce proteins that serve a wider variety of distinct functions. Which of the following explains this phenomenon?
   A) The genome itself is manipulated by the agents responsible for post-translational modification in order to yield an increase in the number of transcriptional products.
   B) Post-translational modification alters the structures and functions of proteins produced from a relatively smaller number of genes.
   C) hnRNA is modified by the actions of post-transcriptional agents to provide increased variety in the mRNA used for translation.
   D) Post-translational modifications enhance the ability of the ribosome to produce distinct protein products.
**CHAPTER 4 PRACTICE PASSAGE**

Protein synthesis involves a number of complex steps, from transcription of the gene through to translation and post-translational modification. After mRNA is transcribed in eukaryotes, it must be processed (capped, poly-A tailed, and spliced) before it can be translated. Prokaryotes do not need to process their mRNA.

Due to the exonuclease activity of DNA polymerase, DNA replication is generally a high-fidelity process. Random errors occasionally occur and these mutations are classified as **frameshift mutations** (insertions or deletions in the base sequence) or **point mutations** (a single base pair change). Any mutation is subject to natural selection, with advantageous mutations preserved and the most deleterious mutations eliminated quickly. Thus, areas of the genome that appear to evolve very slowly (i.e., have a slower rate of mutation than other areas) do not actually have a slower rate; rather, that area is highly critical to normal functioning of the organism involved.

Point mutations can be further classified by their final effect on the mature protein. Because of the redundancy of the genetic code, some mutations do not alter the final amino acid sequence of the protein and are referred to as **silent mutations**. However, it was discovered that all redundant codons are not equal; some are used preferentially to enhance the speed or accuracy of protein translation. tRNAs corresponding to redundant codons are not found equally in the cell; some tRNAs are more common than others. Silent mutations can cause phenotypic changes by altering mRNA stem-and-loop folding, half-life, and splicing sites. Thus, mutations formerly considered “silent” have now been implicated in a number of different disorders, such as Marfan syndrome, phenylketonuria, Seckel syndrome, and increased pain sensitivity.

### Figure 1  The Genetic Code

<table>
<thead>
<tr>
<th>1st Position (5' End)</th>
<th>2nd Position</th>
<th>3rd Position (3' End)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U</strong></td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>Ile</td>
<td>Thr</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td>Met</td>
<td>Thr</td>
</tr>
<tr>
<td><strong>G</strong></td>
<td>Val</td>
<td>Ala</td>
</tr>
</tbody>
</table>

1. Based on information in the passage, genes coding for particularly abundant proteins in a cell would have all of the following EXCEPT:

A) codons corresponding to abundant tRNAs.
B) equal use of redundant codons.
C) greater use of preferential codons.
D) high-fidelity replication.
2. Which of the following could account for the changes brought on by silent mutations in both eukaryotes and prokaryotes? 
   I. Decrease in mRNA half-life
   II. Disruption of splicing sites
   III. Changes in mRNA folding

A) I and II only
B) II and III only
C) I and III only
D) I, II, and III

3. Researchers studying a gene associated with breast cancer found that regions where silent mutations occur (“silent sites”) in this gene evolve very slowly compared to other regions within this gene. Comparisons were made between mice and humans. Which of the following is most likely true about this gene?

A) Mutations at other sites are more detrimental to the health of the organism than mutations at the silent sites.
B) Mutations at the silent sites increase the accuracy of mRNA splicing.
C) Mutations within the silent sites often lead to the death of the organism.
D) The silent sites are less critical to overall function than the other sites.

4. Researchers studying the DNA polymerase activity in several different organisms discovered a mutant E. coli polymerase that retained almost 100% of wild-type activity. Which of the following active-site missense mutations is LEAST likely to affect enzyme activity?

A) A → D
B) V → L
C) R → G
D) L → W

5. Point mutations are found in three subclasses: nonsense mutations, missense mutations, and silent mutations. Which of the following represents a silent mutation?

A) UGC to UGA
B) UUA to CUA
C) CAC to CAA
D) CAU to CUU

6. How could changing the half-life of an mRNA lead to phenotypic changes?

A) A shorter mRNA half-life would lead to a truncated protein.
B) A longer mRNA half-life would increase the amount of time the mRNA stays bound to the template strand of DNA, and reduce the amount of protein translated.
C) Differences in mRNA folding could alter the rates of translation.
D) More or less of the protein encoded by that mRNA would be translated.
SOLUTIONS TO CHAPTER 4 FREESTANDING PRACTICE QUESTIONS

1. C RNA polymerase III transcribes transfer RNA (tRNA), which then carries amino acids to ribosomes for use in translation. This polymerase plays no role in replication (choice A is wrong), and reverse transcription uses a DNA polymerase (in any case, it is not carried out by eukaryotes; choice B is wrong). Blocking the action of this enzyme would not alter the base sequence, so mutation would not be affected (choice D is wrong).

2. D The lac operon includes an operator site to which a repressor protein binds (choice A is wrong). The operator site is located between the promoter region and the start transcription site. When the repressor is bound, RNA polymerase (which binds to the promoter site; choice B is wrong) cannot move forward to the start site; thus transcription is inhibited (choice D is correct). Lactose binds to the repressor protein at an allosteric site, causing a conformational change so that the repressor protein can no longer bind to the operator. When this happens, RNA polymerase can move forward to the start site and transcription will occur (choice C is wrong).

3. C A point mutation is a single base pair substitution. There are few possibilities that can result if a single base is substituted. If the new codon is now a stop codon, then the polypeptide will be truncated (choice A could result from a point mutation and can be eliminated). If the new codon codes for the same amino acid as before the mutation, then a silent point mutation has occurred and no change will be seen in the amino acid sequence (choice B could result from a point mutation and can be eliminated). If the mutation leads to a single new amino acid, then a missense point mutation has occurred (choice D could result from a point mutation and can be eliminated). However, if more than one base was changed, or bases were added/deleted (a frameshift mutation), this would lead to multiple new amino acids (choice C could not result from a point mutation and is the correct answer choice).

4. A Item I is true: Eukaryotic mRNA is monocistronic, meaning that only one protein is transcribed from each mRNA (choice C can be eliminated). Item II is false: Translation does not stop at a stop codon; translation stops at a stop codon (choices B and D can be eliminated, and choice A is the correct answer). Translation stops when a termination signal is reached. Item III is also false: When mRNA is transcribed, it is complementary to the template strand, not identical to it.

5. A Fidelity refers to accuracy. Because RNA polymerases do not proofread, transcription is less accurate (i.e., a lower-fidelity process; choice A is not a similarity and is the correct answer choice). Both replication and transcription use DNA as a template (choice B is a similarity and can be eliminated). In both cases, the removal of pyrophosphate provides the energy for polymerization to occur (choice C is a similarity and can be eliminated). Lastly, although RNA polymerase (in transcription) and DNA polymerase (in replication) move along the parent chain in the 3' → 5' direction, the new chain is made in the 5' → 3' direction.

6. C Transfer RNA (tRNA), not snRNA, is typically involved in the process of coordinating the amino acids that are added to a growing protein during translation (choice C is not a function of snRNA and is the correct answer choice). It should also be noted that translation is NOT taking place in the nucleus, which is the location of snRNA.
pre-mRNA, regulation of transcription factors, and maintenance of telomeres are all functions typically attributed to snRNA and take place in the nucleus (choices A, B, and D are all functions that include snRNAs and can be eliminated).

7. **B** Post-translational modification is, by definition, the manipulation of protein products after translation; the primary purpose of these modifications is to allow for a large increase in the number of possible protein products from a relatively small genome (choice B is correct). The genome itself is not changed during post-translational modification (choice A is wrong). hnRNA is processed in the nucleus after transcription to yield a variety of mRNA transcripts (and this is the other primary way that protein diversity can be achieved), but the question specifically asks about post-translational modification, not post transcriptional effects (choice C is wrong). The ribosome is essentially a factory that reads the code on an mRNA and links amino acids together, and that’s it. Ribosomes create the primary protein structure; this IS translation. Post-translational modification happens after this step (choice D is wrong).

**SOLUTIONS TO CHAPTER 4 PRACTICE PASSAGE**

1. **B** Proteins that are abundant require speed and accuracy during translation, and the passage states that this can be accomplished by using preferential codons (choice C is true and can be eliminated; choice B is false and the correct answer choice). Likewise, codons corresponding to abundant tRNAs would be used instead of those corresponding to the more rare tRNAs (choice A is true and can be eliminated). Choice D is true of all genes, abundant proteins or not (choice D can be eliminated).

2. **C** The passage states that silent mutations can lead to all three of the Roman numeral items listed; however, prokaryotes do not undergo mRNA splicing. Thus Item I is true for both eukaryotes and prokaryotes (choice B can be eliminated), Item II is only true for eukaryotes (choices A and D can be eliminated), and Item III is true for both.

3. **C** According to the passage, areas of the genome that appear to evolve very slowly are highly critical to normal functioning of the organism. Thus, mutations in these areas most likely disrupt function in a major way, leading to the death of the organism and thus the loss of the mutation (hence the reason it appears to evolve very slowly; choice C is correct and choice D is wrong). If other sites appear to evolve more quickly, mutations at those sites must be less detrimental (choice A is wrong). If mutations at the silent sites increase the accuracy of mRNA splicing, this would be beneficial and thus preserved (choice B is wrong). Note that the information on breast cancer and humans versus mice is not necessary to answer the question and is there solely to distract you. Focus on what the question is asking you.

4. **B** In order for a mutation to have a minimal effect on enzyme activity, it must be a relatively conservative mutation. Both valine (V) and leucine (L) are nonpolar amino acids; the substitution of L for V is not likely to have a significant effect on enzyme activity (choice B is correct). Alanine (A) is nonpolar, and aspartic acid (D) is acidic (and polar). The substitution
of D for A would likely be disruptive (choice A is wrong). Arginine (R) is basic (polar) and glycine (G) is nonpolar; the substitution of G for R would likely be disruptive (choice C is wrong). Leucine (L) is nonpolar, as is tryptophan (W); however, the structure of tryptophan is significantly different than leucine. Tryptophan has a large double-ring side chain, while the side chain of leucine is a short hydrocarbon, \(-\text{CH}_2\text{CH}(\text{CH}_3)_2\), very similar to valine, \(-\text{CH}(\text{CH}_3)_2\). The substitution of leucine for valine is likely to be much better tolerated than the substitution of tryptophan for leucine (choice B is better than choice D).

5. B  Nonsense mutations convert a codon for an amino acid into a stop codon, missense mutations lead to amino acid substitutions, and silent mutations do not affect the amino acid sequence of a protein. To answer this question, you must use the genetic code in Figure 1. UGC codes for cysteine and UGA is a STOP codon, making this a nonsense mutation (choice A is wrong). The codons UUA and CUA both code for leucine, making this a silent mutation (choice B is correct). CAC codes for histidine and CAA codes for glutamine; this is a missense mutation (choice C is wrong). CAU codes for histidine and CUU codes for leucine; this is also a missense mutation (choice D is wrong).

6. D  If the half-life of an mRNA is increased, it will stay in the cell longer and more of the protein would be translated. Likewise, if the mRNA’s half-life is decreased, it will be eliminated from the cell more quickly, and less of the protein would be translated. The mRNA half-life has nothing to do with the length of the protein; protein size is dictated by the length of the open reading frame on the mRNA molecule and the number of codons in the translated region (choice A is wrong). mRNA does not stay bound to the DNA template strand for any length of time, regardless of half-life. As mRNA is transcribed, the DNA helix reforms immediately behind it, releasing the mRNA from the transcription bubble as it is synthesized (choice B is wrong). Choice C is a true statement but does not address the question of half-life (choice C is wrong).
Chapter 5
Microbiology
A milestone in microbiology was the demonstration by Louis Pasteur in 1861 that microbes do not spontaneously arise in boiled broth; they must arrive there by contamination. This put the last nail in the coffin of the idea of spontaneous generation of life. Another major contribution to the golden age of microbiology was the isolation of the bacteria responsible for anthrax in 1876 by a physician named Robert Koch. This and other experiments led to the germ theory of disease, the idea that disease was not caused by bad air (“malaria”), but by microorganisms. In this chapter we will examine three major groups of disease-causing organisms, beginning with the smallest (viruses) and ending with the largest (fungi—which are often not microscopic at all but visible with the naked eye).
5.1 VIRUSES

With the identification of bacteria as the cause of anthrax and other diseases, medical science appeared in the late 1800s to be headed toward explanation of all infectious disease. Researchers soon found, however, that some infectious agents could not be trapped by passage through filters in the same manner as bacteria. These agents also proved invisible to the light microscope, unlike bacteria. With the advent of electron microscopy, the tiny infectious agents known as viruses were finally visualized. Today, molecular biology has shed great light onto viruses, down to the nucleotide sequence of entire viral genomes. However, viruses such as HIV still remain one of the most serious threats to health, indicating there is still much to learn.

Viruses infect all life forms on earth, including plants, animals, protists, and bacteria. A virus is an obligate intracellular parasite. As such, they are only able (obligated) to reproduce within (intra) cells. While within cells, viruses have some of the attributes of living organisms, such as the ability to reproduce; but outside cells, viruses are without activity. Viruses on their own are unable to perform any of the chemical reactions characteristic of life, such as synthesis of ATP and macromolecules. Viruses are not cells or even living organisms. To reproduce, they commandeer the cellular machinery of the host they infect and use it to manufacture copies of themselves. In the final analysis, a virus is nothing more than a package of nucleic acid that says: “Pick me up and reproduce me.” Remember this crucial definition: A virus is an obligate intracellular parasite that relies on host machinery whenever possible. In the following sections we will look at some of the variations on this basic theme.

- Cyanide (an inhibitor of the electron transport chain) is added to a culture of virus-infected mammalian cells. The virus has none of the components of electron transport or any other proteins that are inhibited by cyanide. Which one of the following best describes the effect of cyanide?2
  A) The mammalian cells will die, and all viruses will be destroyed as well, regardless of their stage of development.
  B) Mammalian cells are killed, and viral replication halted, but the culture remains infectious.
  C) Mammalian cells stop growing, and viral replication is unaffected.
  D) Mammalian cells continue to grow, but viral replication is halted.

Viral Structure and Function

The structure of viruses reflects their life cycle. In general, all viruses possess a nucleic acid genome packaged in a protein shell. The exterior protein packaging helps to convey the genome from one cell to infect other cells. Once in a cell, the viral genome directs the production of new copies of the genome and of the

1 Note, however, that some viruses store some ATP in their capsids. They acquired this ATP from the previous host and typically use it to power penetration (see below).

2 The mammalian cells are directly dependent on the ATP generated by the electron transport chain, so if cyanide inhibits the electron transport chain, the mammalian cells will die (choice D is wrong). The viruses are dependent on the mammalian cells for the ATP and enzymes needed for replication, so if the mammalian cells die, viral replication will stop (choice C is wrong). However, any viruses that had already completed the replication process when the cyanide was added will not be affected, and will remain infectious (choice B is correct and choice A is wrong).
protein packaging needed to produce more virus. However, the nature of the genome, the protein packaging, and the viral life cycle vary tremendously between different viruses.

A viral genome may consist of either DNA or RNA that is either single-stranded or double-stranded and is either linear or circular. Viruses utilize virtually every conceivable form of nucleic acid as their genome. However, a given type of virus can have only one type of nucleic acid as its genome, and a mature virus does not contain nucleic acid other than its genome. 3

If the ratio of adenine to thymine in a DNA virus is not one to one, what can be said about the genome of this virus? 4 A disease agent that is isolated from a human cannot reproduce on its own in cell-free broth but can reproduce in a culture of human cells. In its pure form it possesses both RNA and DNA. Is it possible that the disease agent is a virus? 5

A factor that influences all viral genomes, regardless of the form of the nucleic acid used as genome, is size as a limiting factor. Viruses are much smaller than the hosts they infect, both prokaryotic and eukaryotic. Figure 1 depicts the relative size of a bacteriophage (a virus that infects bacteria) and its host.

Not only are viruses small, but the exterior protein shell of a virus is typically a rigid structure of fixed size that cannot expand to accommodate a larger genome. [What is the likely result if a viral genome is tripled in size? 6] To adapt to this size constraint, viral genomes have evolved to be extremely economical. One adaptation is for the viral genome to carry very few genes and for the virus to rely on host-encoded proteins for transcription, translation, and replication. [How do ribosomes used to translate viral proteins compare to host ribosomes? 7] Another adaptation found in viral genomes is the ability to encode more than one protein in a given length of genome. A virus can accomplish this feat by utilizing more than one reading frame within a piece of DNA so that genes may overlap with each other.

3 There are exceptions. For example, it has recently been discovered that the Hepatitis B virus has a circular DNA genome which is part single-stranded and part double-stranded. The take-home point here is that when a virus is not inside a host cell, it contains only its genome, which is always the same (except in special situations such as when a piece of host genome accidentally becomes incorporated in the viral genome). In contrast, a true cell contains not only its genome, but also mRNA, rRNA, and tRNA.

4 Adenine base pairs with thymine in double-stranded DNA. Thus, for every A there should be one T for a one to one ratio of A to T. If the ratio differs from this, the genome must be single-stranded DNA, or RNA, which has no T.

5 No, it cannot be a virus. Viruses possess only one kind of nucleic acid. The disease agent is another kind of obligate intracellular parasite (certain bacteria can only reproduce inside host cells, e.g., Chlamydia).

6 The viral genome will probably no longer fit within the normal viral structure, and the genome will therefore not be packaged into infectious viral particles.

7 Viruses use host ribosomes. Viral and host proteins are translated by the same ribosomes.
• A 1,000 base pair region of viral genome is found to encode two polypeptides unrelated in amino acid sequence during infection of eukaryotic cells. If one of these polypeptides is 250 amino acids in length and the other is 300, what is the best explanation for this?
  A) A missense mutation
  B) Viruses use a different genetic code than eukaryotes do
  C) Overlapping multiple reading frames
  D) The polypeptides are splicing variants

Surrounding the viral nucleic acid genome is a protein coat called the capsid. The capsid provides the external morphology that is used to classify viruses. It is made from a repeating pattern of only a few protein building blocks. Helical capsids are rod-shaped, while polyhedral capsids are multiple-sided geometric figures with regular surfaces. Complex viruses may contain a mixture of shapes. For example, the T4 bacteriophage has a helical sheath and a polyhedral head (Figure 2). This virus is commonly used in research; its host is the bacterium E. coli. The genome is located within the capsid head. Other parts of the capsid are used during infection of the host. The tail fibers attach to the surface of the host cell, as does the base plate. The sheath contracts using the energy of stored ATP, injecting the genome into the host. [Why might a bacteriophage inject its DNA, while animal viruses do not?]

8 The problem is that the virus must contain at least 750 bp (250 amino acids) and 900 bp (300 amino acids) of genetic information for unrelated polypeptides in 1000 bp of DNA. The only way to do this is overlapping multiple reading frames (choice C).

9 Phage must puncture the bacterial cell wall, while animal viruses can be internalized whole into animal cells (since they do not have a cell wall).
The most important thing to understand is that the entire viral capsid is composed of protein, while the viral genome is composed of nucleic acid (DNA or RNA). Most viruses are not as structurally complex as the bacteriophage shown in Figure 2. See Figure 3 for more examples.

Many animal viruses also possess an envelope that surrounds the capsid. This is a membrane on the exterior of the virus derived from the membrane of the host cell. It contains phospholipids, proteins, and carbohydrates from the host membrane, in addition to proteins encoded by the viral genome. Enveloped viruses acquire this covering by budding through the host cell membrane. To infect a new host, some enveloped viruses fuse their envelope with the host’s plasma membrane, which leaves the de-enveloped capsid inside the host cell. Viruses which do not have envelopes are called naked viruses. All phages and plant viruses are naked. [Can you imagine why this might be true?]

---

10 Remember: Viruses acquire envelopes by budding through host membranes. Phages and plant viruses infect hosts that possess cell walls. When viruses begin to exit the cell, the cell wall is destroyed, and host membranes rupture. Thus there is no membrane through which the remaining viruses must bud; they simply escape in a lytic explosion.
Whether enveloped or naked, the surface of a virus determines what host cells it can infect. Viral infection is not a random process, but highly specific. A virus binds to a specific receptor on the cell surface as the first step in infection. After binding, the virus will be internalized, either by fusion with the plasma membrane or by receptor-mediated endocytosis. Only cells with a receptor that matches the virus will become infected, explaining why only specific species or specific cell types are susceptible to infection. The viral surface is also important for recognition by our immune system. [If antibodies to a viral capsid protein are ineffective in blocking infection, what might this indicate about the virus?]

**Bacteriophage Life Cycles**

Since viruses lack the ability to produce energy and replicate on their own, they use the machinery of the cell they infect to carry out these processes. The viral genome contains genes that redirect the infected cell to produce viral products. The first step is binding to the exterior of a bacterial cell in a process termed attachment or adsorption. The next step is injection of the viral genome into the host cell in a process termed penetration or eclipse. It is called “eclipse” because the capsid remains on the outer surface of the bacterium while the genome disappears into the cell, removing infectious virus from the media. From this point forward a phage follows one of two different paths: It enters either the lytic cycle or the lysogenic cycle.

**The Lytic Cycle of Phages**

As soon as the phage genome has entered the host cell, host polymerases and/or ribosomes begin to rapidly transcribe and translate it. One of the first viral gene products made is sometimes an enzyme called hydrolase, a hydrolytic enzyme that degrades the entire host genome. (Hydrolase is an example of an early gene; one of a group of genes that are expressed immediately after infection and which includes any special enzymes required to express viral genes.) Then multiple copies of the phage genome are produced (using the dNTPs resulting from degradation of the host genome), as well as an abundance of capsid proteins. Next, each new capsid automatically assembles itself around a new genome. Finally, an enzyme called lysozyme is produced. An example of a late gene, lysozyme is also present in human tears and saliva. It destroys the bacterial cell wall. Because osmotic pressure is no longer counteracted by the protection of the cell wall, the host bacterium bursts (“lyses,” hence the name lytic), releasing about 100 progeny viruses, which can begin another round of the cycle (see Figure 4). [If lysozyme were an early gene, would this be advantageous to the virus?]

---

11 It suggests that the virus is enveloped, so the antibody cannot reach its epitope on the capsid surface in infectious virus.

12 No. The host cell would lyse before the phage had time to replicate and assemble.
1. Attachment of phage to *E. coli* and injection of phage chromosome

2. Breakdown of bacterial chromosome by phage-specific enzyme

3. Replication of phage chromosome using bacterial materials and phage enzymes

4. Expression of phage genes to produce phage structural components

5. Assembly of progeny phage particles

6. Release of progeny phage by lysis of bacterial wall

---

**Figure 4** The Lytic Cycle

- When phage are first added to a bacterial culture, the number of infective viruses initially decreases before it later increases. Why does this occur? 

- Bacteria cultured in the presence of \(^{35}\)S-labeled cysteine and \(^{32}\)P-labeled phosphates are infected with phage T4. When phage from this culture are used to infect a new nonradiolabeled bacterial culture, which of the isotopes will be found in the interior of the newly infected bacteria? 

- A bacteriophage with an important capsid gene deleted infects the same cell as another virus with a normal copy of the same gene. At the time of host-cell lysis:

  A) all released viruses will be capable of infecting new hosts, but only some of these new infections will give rise to phage capable of infecting new hosts.
  
  B) no infective viruses will be released.
  
  C) each individual virus that is released will produce a mixture of infective and noninfective viruses in subsequent infections.
  
  D) only normal viruses will be released.

---

\(^{13}\) The initial decrease is due to the simple fact that many phage have injected their genomes into hosts and are no longer infectious.

\(^{14}\) The \(^{35}\)S cysteine will be incorporated into viral coat proteins and the \(^{32}\)P phosphate will be incorporated into the viral nucleic acid genome in newly released viral particles. (Proteins contain no P and nucleic acids contain no S.) When these viruses infect bacteria, their nucleic acids are injected into the bacteria while the capsid proteins remain on the exterior, which means that only the \(^{32}\)P will be found in the interior of the newly infected cells.

\(^{15}\) When two viruses infect the same cell, it is called co-infection. Some normal viruses will result, and some genomes from defective viruses will get packaged into capsids made from proteins encoded by the normal virus. The latter will be capable of infecting new hosts, but when they do their progeny will not survive due to the capsid abnormality. Choice A is correct, and choices B and D are wrong. Think about it: Where did the phage with the deleted capsid gene come from? The deficient virus must have come from a co-infection such as this. The deficient phage can only infect host cells and reproduce with the help of normal viruses. Note that because a single virus carries only a single genome, it can produce only one type of progeny (choice C is wrong).
The Lysogenic Cycle of Phages

The lytic cycle is an efficient way for a virus to rapidly increase its numbers. It presents a problem though: All host cells are destroyed. This is an evolutionary disadvantage. Some viruses are cleverer: They enter the lysogenic cycle. Upon infection, the phage genome is incorporated into the bacterial genome and is now referred to as a prophage; the host is now called a lysogen (Figure 5). The prophage is silent; its genes are not expressed, and viral progeny are not produced. This dormancy is due to the fact that transcription of phage genes is blocked by a phage-encoded repressor protein that binds to specific DNA elements in phage promoters (operators). The cleverness of the lysogenic cycle lies in the fact that every time the host cell reproduces itself, the prophage is reproduced too. Eventually, the prophage becomes activated. It now removes itself from the host genome (in a process called excision) and enters the lytic cycle.

One potential consequence of the lysogenic cycle is that when the viral genome activates, excising itself from the host genome, it may take part of the host genome along with it. When the virus replicates, the small piece of host genome will be replicated and packaged with the viral genome. In subsequent infections, the virus will integrate the "stolen" host DNA along with its own genome into the new host's genome. The presence of the new DNA will become evident if it codes for a trait that the newly infected host did not previously possess, such as the ability to metabolize galactose. This process is called transduction. [Why would a bacterial gene, carried with a virus and integrated with viral genes into a new bacterial genome, not be repressed along with the viral genes during lysogeny?]

---

5. In rare cases, the prophage may separate and the cell will be induced to lyse.
3. The phage DNA integrates and becomes a noninfective prophage.
2. The phage DNA enters the host cell.
1. The phage binds to the bacterium.

Bacterial Chromosome
Prophage

Figure 5 The Lysogenic Cycle

---

16 Prophage latency results from a viral repressor protein binding to viral DNA in a sequence-specific manner. The specific DNA sequence to which the repressor binds is present in the viral genes but not in the bacterial genes, so the bacterial gene can be expressed while the viral genes are repressed.
Replication of Animal Viruses

There are a number of differences between phages and viruses which infect animal cells. (Animal viruses don't have a special name like "phage.") The general outline of the viral life cycle, however, remains the same. The virus must specifically bind to a proper host cell, release its genetic material into the host, take over host machinery, replicate its genome, synthesize capsid components, assemble itself, and finally escape to infect a new cell.

Animal cells have proteins on the surface of their plasma membranes that serve as specific receptors for viruses. These receptors play a role in normal cellular function; they do not exist simply for the benefit of the virus. Part of the tissue-specificity of animal viruses is due to the distribution of receptors necessary for adsorption. For example, the binding of the HIV virus protein gp120 to a T cell membrane protein termed CD4 is one of the first steps in HIV infection.

- Would treatment of an HIV-infected person with a soluble form of CD4 protein affect the infectivity of the virus?\(^\text{17}\)
- Mutation of the cell-surface receptor that viruses attach to would be a means for an organism to become resistant to viral infection. Why is this mechanism not common?\(^\text{18}\)
- Treatment of an enveloped animal virus with a mild detergent solubilizes several proteins from the virus, although the genome does not become accessible. Which one of the following is consistent with this scenario?\(^\text{19}\)
  - A) Some of the proteins that are released by detergent may be encoded by the genome of the infected cell.
  - B) The infectivity of the virus is not affected by detergent treatment.
  - C) The proteins released by detergent are capsid proteins.
  - D) All the proteins released by the detergent are encoded by the viral genome.

The next step in the infection of an animal cell is penetration into the cell, just as in bacterial infection by a phage. Many animal-viruses enter cells by endocytosis (a process whereby the host cell engulfs the virus and internalizes it). [Why don’t phages enter their hosts by endocytosis?\(^\text{20}\)] Once inside the host, the viral genome is uncoated, meaning it is released from the capsid. Alternatively, some viruses fuse with the plasma membrane to release virus into the cytoplasm. From this point, an animal virus may enter either a lytic cycle, a lytic-like cycle called the productive cycle, or a lysogenic cycle.

The lytic cycle in animal viruses is the same as in phages. The productive cycle is similar to the lytic cycle but does not destroy the host cell. It is possible because enveloped viruses exit the host cell by budding through the host’s cell membrane, becoming coated with this membrane in the process. Budding does not necessarily destroy a cell since the lipid bilayer membrane can reseal as the virus leaves. Finally, in the animal virus lysogenic cycle the dormant form of the viral genome is called a provirus (analogous to a prophage). For example, Herpes simplex I is the virus that causes oral herpes. After infection, it may remain

\(^{17}\) Yes, it would. The soluble CD4 protein would bind to the virus’s CD4 receptor (gp120) and block attachment of the virus to the T cells.

\(^{18}\) Two reasons: 1) The receptor has a specific role in the normal physiology of the host, which a mutation might compromise. 2) Viruses generally evolve so rapidly that they can keep up with any changes in the host, but this is not an absolute rule. Cells of our immune system keep us alive by keeping up with most microorganisms’ tricks.

\(^{19}\) The detergent solubilized the viral envelope (choice C is wrong). As stated in the text, some envelope proteins are encoded by the virus and some are derived from the host’s membranes during budding (choice A is correct, and choice D is wrong). Removal of envelope proteins will impair viral adsorption and reduce infectivity (choice B is wrong).

\(^{20}\) Bacteria do not perform endocytosis, in part because they have a rigid cell wall that does not permit them to.
dormant as a provirus for an indefinite period of time. Then one day, usually when the host encounters stress (e.g., lack of sleep, upcoming professional school entrance exams), the virus reactivates.

**Viral Genomes**

Many factors determine the uniqueness of each virus. The type of genome, possession or lack of an envelope, nature of cell-surface proteins, and type of life cycle are examples. All of these parameters are used in the classification of viruses, and all are potential targets for therapeutic intervention. The nature of the genome is perhaps the most important of these and has important consequences for how infection by each virus proceeds. In the following discussion we will look at a few viral genomes with an eye to *what proteins the virus must encode or actually carry in its capsid based on its genome type*. Our purpose is not to provide new information, but rather to demonstrate what conclusions can be drawn from what you already know (typical MCAT passage material). Do not memorize, but rather read for comprehension. We will not discuss ds-RNA or ss-DNA genomes, but by the end of this section you should be able to imagine components they might require.

**[+] RNA Viruses**

—must *encode* RNA-dependent RNA pol (and do not have to carry it).

A (+) RNA virus, with a single-stranded RNA genome, is the simplest imaginable type of viral genome. (A piece of single-stranded viral RNA which serves as mRNA is called (+) RNA.) As soon as the (+) RNA genome is in the host cell, host ribosomes begin to translate it, creating viral proteins. The viral genome acts directly as mRNA. The technical way to describe this scenario is to say the genome is *infective*, meaning injecting an isolated genome into the host cell will result in virus production. In order for the virus to replicate itself, one of the proteins it encodes must be an RNA-dependent RNA polymerase, the role of which is __? 21 (+) RNA viruses cause the common cold, polio, and rubella. [Will an infectious virus be produced if the genome of an enveloped (+) strand RNA virus is added to an extract prepared from the cytoplasm of eukaryotic cells that retains translational activity but lacks DNA replication or transcription of host genes? 22 If a viral genome is (+) strand RNA, what is used as a template by the RNA-dependent RNA polymerase? 23]

**[–] RNA Viruses**

—must *carry* RNA-dependent RNA pol (and, of course, encode it too).

The genome of a (–) RNA virus is *complementary* to the piece of RNA that encodes viral proteins. In other words, the genome of a (–) RNA virus is the template for viral mRNA production. If host ribosomes translate (–) RNA, useless polypeptides will be made. Hence, the virus must not only encode an RNA-dependent RNA polymerase, it must actually carry one with it in the capsid. When the virus enters the

---

21 to copy the RNA genome for viral replication; the host never makes RNA from RNA.

22 No. The (+) strand RNA virus will be able to produce viral genome and proteins, but progeny will not be able to acquire the envelope they need to be infectious.

23 To make (+) strand copies of the genome, the virus needs the complementary strand as a template: the (–) strand RNA. Thus, the RNA-dependent RNA polymerase produces a (–) strand intermediate before generating new (+) strand genomes.
host cell, this enzyme will create a (+) strand from the (–) genome. Then the viral life cycle can proceed. (–) RNA viruses cause rabies, measles, mumps, and influenza. [Do (–) strand RNA viruses use host enzymes to catalyze RNA production in transcription or in replication of the genome?]

**Retroviruses**
—must encode reverse transcriptase.

HIV, the virus that causes AIDS, and HTLV (Human T-cell Leukemia Virus) are examples of retroviruses. These are (+) RNA viruses that undergo lysogeny. In other words, they integrate into the host genome as proviruses. In order to integrate into our double-stranded DNA genome, a viral genome must also be composed of double-stranded DNA. Since these viral genomes enter the cell in an RNA form, they must undergo reverse transcription to make DNA from an RNA template. This snubbing of the central dogma is accomplished by an RNA-dependent DNA polymerase (reverse transcriptase) encoded by the viral genome. Retroviruses are theoretically not required to carry this enzyme, only to encode it. [Why?]
The three main retroviral genes are gag (codes for viral capsid proteins), pol (polymerase codes for reverse transcriptase) and env (envelope codes for viral envelope proteins). [After integration of a retrovirus into the cellular genome, a reverse transcriptase inhibitor is added to the cell. Will the production of new viruses be blocked?]

**Double-Stranded DNA Viruses**
—often encode enzymes required for dNTP synthesis and DNA replication.

These viruses often have large genomes that include genes for enzymes involved in deoxyribonucleotide synthesis (which we do whenever we make DNA) and DNA replication. [Given the limited information that viruses may contain in their genomes, why carry around genes for an enzyme possessed by the host? Why don’t RNA viruses do this? What is a factor likely to limit the size of RNA genomes? Some DNA viruses induce infected host cells to enter mitosis and may even override cellular inhibition of cell division so strongly that the cell becomes cancerous; what is the advantage to the virus of inducing host-cell division?]

---

24 Neither. Viral RNA-dependent RNA polymerase first makes (+) strand as mRNA and then uses the (+) strand as the template to replicate new (–) strand genomes.

25 Because the viral RNA genome can be translated by host ribosomes; thus reverse transcriptase may be made after the viral genome enters the host. It just so happens that HIV does carry its reverse transcriptase within its capsid. You should understand why this is not a theoretical necessity.

26 No, it will not. Reverse transcriptase is required for only one phase of the retrovirus life cycle: the copying of the viral RNA genome into DNA so that it can integrate into the host genome and be transcribed. Once the viral genome has integrated, transcription to produce viral mRNA and new viral RNA genomes does not involve reverse transcriptase. It can proceed with the normal host-cell enzymes.

27 The host cell will only make dNTPs in preparation for replication. If the virus wants to reproduce without waiting for the host to do so, it must encode its own enzymes for the synthesis of DNA building blocks.

28 Transcription is always occurring in all cells, so NTPs (not dNTPs) are always present.

29 The error rate in RNA synthesis is much higher than in DNA synthesis, in part because there are mechanisms to proofread and correct errors in DNA synthesis (but not in RNA synthesis). If an RNA genome were too large, every copy of the viral genome synthesized would suffer from so many errors that no infectious virus would be produced.

30 To replicate, the DNA virus must either provide all of the necessary components (such as dNTPs) itself, infect a cell that is already dividing, or induce the cell it infects to enter mitosis and produce the ingredients for DNA synthesis.
• Adenoviruses have a single linear ds-DNA genome, which contains a number of different promoters that are regulated during infection. Although transcription is carried out by cellular RNA polymerase, the viral E1A gene product is required for transcription of most viral genes. If the E1A gene is deleted from the virus or if the gene product is inactivated, viral infection is unable to proceed. Adenoviruses also encode much of their own replication machinery, including DNA polymerase. If two different adenoviruses infect the same cell, one with a deleted E1A gene and another with a deleted DNA polymerase gene, will successful infection of the cell result?31

5.2 SUBVIRAL PARTICLES

Some infectious agents are even smaller and simpler than viruses and are termed subviral particles. These include prions and viroids.

Prions

As infectious agents, prions do not strictly follow the Central Dogma because they are self-replicating proteins [Why does this violate the Central Dogma?32]. The prion itself is a misfolded version of a protein that already exists (see Figure 6).

![Figure 6](image-url)  
*Figure 6  Comparison of the PrPC structure to the PrPSc structure*

31 Yes, thanks to complementation. The mutant viruses will complement each other, one providing the E1A protein and the other providing DNA polymerase. Note that this had to have happened before; how else could a defective virus such as these exist? One virus which complements another is called a helper virus.

32 The Central Dogma states that information flows in its nucleotide form from DNA to RNA (transcription), and then in its amino acid form from RNA to protein (translation). Prions take both transcription and translation out of the process and have proteins being shaped based on other proteins, hence the term “self-replicating.”
When the normally folded protein (designated PrP\(^\text{C}\)) comes into contact with the prion (designated PrP\(^\text{Sc}\)), the prion acts as a template; the shape of the normal protein is altered and it too becomes infectious. Prions are responsible for a class of diseases in mammals referred to as the **transmissible spongiform encephalopathies** (TSEs). These diseases cause degeneration in the nervous system, especially the brain where characteristic holes develop, and are always fatal. The misfolded proteins are found in the nervous tissues and are very resistant to degradation by chemicals or heat, making them hard to destroy. Bovine spongiform encephalopathy (BSE, commonly called *mad cow disease*) is the prion disease found in cows; this was originally transmitted to cows from sheep because all types of tissue from sheep, including the brain, is used as a supplement in the feed for other farm animals. Though much less common, the disease *kuru* follows a similar transmission path in humans; it is only found in a limited number of tribes where consumption of the body, particularly of the brain, is part of honoring the dead (since identification of the transmission route, this practice has stopped and kuru has virtually disappeared).

However, prion diseases can also be genetically linked, through mutations in the gene that codes for the prion protein. For example, fatal familial insomnia (FFI) is an autosomal dominant condition inherited on chromosome 20, and Creutzfeldt Jakob disease (CJD) is also inherited. It is also possible for these diseases to arise spontaneously (through mutation) in someone with no prior family history. In general, however, prion diseases are very rare, striking only 1–2 people per million.

Whether transmitted, inherited, or spontaneously arising, prion diseases are characterized by their very long incubation periods, which can be several months to years in animals and several years to decades in humans. The misfolded proteins cause the destruction of neurons, particularly in the central nervous system, leading to loss of coordination, dementia, and death. Diagnosis is difficult, in part because of the long incubation periods and in part because the symptoms can be indicative of other conditions.

**Viroids**

Viroids consist of a short piece of circular, single-stranded RNA (200–400 bases long) with extensive self-complementarity (i.e., it can base-pair with itself to create some regions that are double-stranded; see Figure 7). Generally they do not code for proteins and they lack capsids. Some viroids are catalytic ribozymes, while others, when replicated, produce siRNAs that can silence normal gene expression.

![Figure 7](structure_of_a_viroid_showing_double-stranded_regions.png)

Reproduction of some viroids shares similarity to the replication of RNA viruses. A viroid RNA-dependent RNA polymerase synthesizes a (–) strand, which is circularized by an RNA ligase derived from the host; this is then used as the round, rolling template to make more (+) copies that match the original RNA viroid sequence. An alternative to this mechanism leaves the (–) stand in a more linear state where it
can still act as a template for (+) strand creation and then become circularized. In other cases, viroids somehow hijack the cell’s DNA dependent RNA polymerase and direct it to read RNA templates. This mechanism is not well understood.

Most of the diseases caused by viroids are found in plants. The only human disease linked to viroids is Hepatitis D. The Hepatitis D viroid can only enter hepatocytes (liver cells) if it is contained in a capsid with a binding protein; since viroids do not have capsids, successful Hepatitis D infection required coinfection with Hepatitis B, from which it derives its capsid.

5.3 PROKARYOTES (DOMAIN BACTERIA)

Cell Theory
Advances in microbiology have been made possible by advancing technologies in magnification. Once humans were able to utilize basic, if crude, microscopy, the cell as the monomer of tissues and organs could be studied. In 1655, this led the English scientist, Robert Hooke, to define the Cell Theory based on his studies of cork. Its tenets are as follows:

1) All living organisms are composed of one or more cells and their products.
2) Cells are the monomer for any organism.
3) New cells arise from pre-existing, living cells.

Though these basic principles are still true, more modern extensions of Cell Theory also include the idea that no matter what the species, the chemical composition of cells is similar, that DNA is the source of hereditary programming information passed from cell to cell, that an organism’s activity is determined by the total activity of its cells, and that biochemical energy flow occurs within cells. These additional principles have been explored and verified due to vast improvements both in microscopy as well as biochemical and genetic testing.

All living organisms (which does not include viruses) can be classified as either prokaryotes or eukaryotes. The classification of organisms into these groups is based on examination of their internal cellular structure. Representatives from both groups are able to carry out the basic biochemical processes of photosynthesis, the Krebs cycle, and oxidative phosphorylation to produce ATP. The primary feature of prokaryotes that distinguishes them from eukaryotes is that they do not contain membrane-bound organelles (nucleus, mitochondria, lysosomes, etc.). Prokaryote means “before the nucleus,” and the lack of a nucleus indicates that prokaryotes are evolutionarily the oldest domains. Unlike viruses, however, prokaryotes possess all of the machinery required for life. They are true cells; true living organisms. The prokaryotes include bacteria, archa (extremophiles), and blue-green algae (cyanobacteria).

The classification of living organisms, taxonomy, is an important part of biology because it is used to determine the evolutionary relationship of organisms to one another. The largest taxonomic division is the domain. There are three recognized domains: Bacteria, Archa, and Eukarya. Domains Bacteria and
Archea include prokaryotic organisms, and Domain Eukarya includes eukaryotic organisms. Each domain can be further subdivided into kingdoms. Currently there are three well-recognized eukaryotic kingdoms (Animalia, Plantae, and Fungi), and great debate over the number of kingdoms that should be present in the other prokaryotic domains and in the single-celled eukaryotes (protists).

In this section, we will begin to study the most basic and ancient of organisms, the prokaryotes.

**Bacterial Structure and Classification**

**Contents of the Cytoplasm**

In this section we will tour the bacterial cell from the inside out. Unlike a eukaryotic cell, there are no membrane-bound organelles in prokaryotic cells (note that ribosomes, which are not membrane-bound, are found in bacteria). The prokaryotic genome is a single double-stranded circular DNA chromosome. It is not located in a nucleus, and it is not associated with histone proteins as the eukaryotic genome is. In bacteria, transcription and translation occur in the same place at the same time. Ribosomes begin to translate mRNA before it is completely transcribed. Many ribosomes translating a single piece of mRNA form a structure known as a polyribosome.

[In Figure 9 on the next page, is the free end of the mRNA the 3' or the 5' end? Which end of the nascent polypeptides is the free end?] Remember that the bacterial ribosome is structurally different from the eukaryotic ribosome, though both function the same way. The differences allow us to prescribe various antibiotics which interfere with bacterial translation without disrupting our own. (Examples are streptomycin and tetracycline, which bind only to bacterial ribosomes.)

---

33 There are a few exceptions to this (e.g., bacteria with more than one chromosome and/or linear chromosomes), but you do not have to know them for the MCAT.

34 The 5' end of the mRNA polymer is free, since elongation of mRNA proceeds 5' to 3'. Proteins are made N to C, so the free end of the polypeptides is the N terminus.
One last genetic element that can be found in prokaryotic cells is the plasmid. This is a circular piece of double-stranded DNA which is much smaller than the genome. Plasmids are referred to as extrachromosomal genetic elements. They often encode gene products which may confer an advantage upon a bacterium carrying the plasmid. For example, plasmids frequently carry antibiotic-resistance genes (genes that encode proteins which can break down antibiotics). Many plasmids are capable of autonomous replication, which means that a single plasmid molecule within a bacterial cell may cause itself to be replicated into many copies. Plasmids are important not only because they may encode advantageous gene products, but also because they orchestrate bacterial exchange of genetic information, or conjugation, which is discussed below.

**Bacterial Shape**

Bacteria are often classified according to their shape. The three shapes and their proper names are organized in the following table:

<table>
<thead>
<tr>
<th>Shape</th>
<th>Proper name (plural)</th>
<th>Proper name (singular)</th>
</tr>
</thead>
<tbody>
<tr>
<td>round</td>
<td>cocci,</td>
<td>coccus</td>
</tr>
<tr>
<td>rod-shaped</td>
<td>bacilli</td>
<td>bacillus</td>
</tr>
<tr>
<td>spiral-shaped</td>
<td>spirochetes or spirilla</td>
<td>spirochete, spirillum</td>
</tr>
</tbody>
</table>

Table 1  Bacterial Classification by Shape

**The Cell Membrane and the Cell Wall**

The bacterial cytoplasm is bounded by a lipid bilayer which is similar to our own plasma membrane. Outside the lipid bilayer is a rigid cell wall. It provides support for the cell, preventing lysis due to osmotic pressure. (As we will discuss in Chapter 6, animal cells lack a cell wall. They deal with the problem of osmotic pressure by continuously pumping ions across the cell membrane.) The bacterial cell wall is composed of peptidoglycan, a complex polymer unique to prokaryotes. It contains cross-linked chains made of sugars and amino acids, including D-alanine, which is not found in animal cells (our amino acids have the L configuration). The bacterial cell wall is the target of many antibiotics, such as penicillin. The enzyme lysozyme, which is found in tears and saliva and made by lytic viruses, destroys the peptidoglycan.
in the bacterial cell wall, resulting in an osmotically fragile structure called a protoplast. [Would a protoplast moved from salt water to fresh water shrivel or burst?]35

**Gram Staining of the Cell Wall**

As part of our tour of the bacterial cell, we will say a word about classification of bacteria according to two different types of cell wall. The method of classification is derived from the extent to which bacteria turn color in a procedure termed Gram staining. The two groupings are Gram-positive, which stain strongly (a dark purple color) and Gram-negative bacteria, which stain weakly (a light pink color).

Gram-positive bacteria have a thick peptidoglycan layer outside of the cell membrane and no other layer beyond this. Gram-negative bacteria have a thinner layer of peptidoglycan in the cell wall but have an additional outer layer containing lipopolysaccharide. The intermediate space in Gram-negative bacteria between the cell membrane and the outer layer is termed the periplasmic space, in which are sometimes found enzymes that degrade antibiotics (see Figure 10). The increased protection of Gram-negative bacteria from the environment is reflected in their weak staining, as well as in their increased resistance to antibiotics. [Which bacteria would be more susceptible to lysis when treated with lysozyme: Gram-positive or Gram-negative?]36

---

![Figure 10](https://example.com/fig10.png)

**Figure 10** Gram-positive vs. Gram-negative Bacteria

35 It would burst, since water would flow into the cell by osmosis.

36 Lysozyme hydrolyzes linkages in peptidoglycan to weaken the cell wall. The peptidoglycan in Gram-positive cells is more accessible, since these cells do not possess an additional outer layer; therefore, Gram-positive cells will lyse more easily when treated with lysozyme.
Endotoxins vs. Exotoxins

Endotoxins are normal components of the outer membrane of Gram-negative bacteria that aren’t inherently poisonous. However, they cause our immune system to have such an extreme reaction that we may die as a result. Endotoxins cause the most trouble when many bacteria die and their disintegrated outer membranes are released into the circulation. When this occurs, cells of the immune system release so many chemicals that the patient goes into what is called septic shock, in which much of the aqueous portion of the blood is leaked into the tissues causing a drop in blood pressure, and other problems, which may be fatal. Endotoxins can have various chemical structures, including lipopolysaccharide, which contains sugars bound to lipids.

Exotoxins are very toxic substances secreted by both Gram-negative and Gram-positive bacteria into the surrounding medium. Exotoxins help the bacterium compete with other bacterial species, such as normal inhabitants of the mammalian gut. Some diseases that are caused by exotoxins are botulism, diptheria, tetanus, and toxic shock syndrome.

The Capsule

Another attribute which only some bacteria have is the capsule or glycocalyx. This is a sticky layer of polysaccharide “goo” surrounding the bacterial cell and often surrounding an entire colony of bacteria. It makes bacteria more difficult for immune system cells to eradicate. It also enables bacteria to adhere to smooth surfaces such as rocks in a stream or the lining of the human respiratory tract.

Flagella

Another item only some bacteria have are long, whip-like filaments known as flagella, which are involved in bacterial motility. [Can viruses move via flagellar propulsion to find host cells?] A bacterium which possesses one or more flagella is said to be motile, because flagella are the only means of bacterial locomotion. Bacteria may be monotrichous (meaning they have a flagellum located at only one end), amphitrichous (meaning they have a flagellum located at both ends), or peritrichous (meaning that they have multiple flagella). The following is which?

The structure of the flagellum is fairly complicated, with components encoded by over 35 genes, but it can be broken down into a few major components: the filament, the hook, and the basal structure (Figure 11). The basal structure contains a number of rings that anchor the flagellum to the inner and outer membrane (for a Gram-negative bacterium) and serve to rotate the rod and the rest of the attached flagellum in either a clockwise or counterclockwise manner. The most important thing to remember about the prokaryotic flagellum is that its structure is different from the eukaryotic one (which contains a “9 + 2” arrangement of microtubules, discussed in Chapter 6).

---

37 No. Viruses lack any means of energy production on their own and any means of active movement. They rely on diffusion to find host cells.

38 Monotrichous
The rotation of the rod is powered by the diffusion of H\(^+\) down the proton gradient generated across the inner membrane by electron transport. Bacterial motion can be directed toward attractants, such as food, or away from toxins, such as acid, in a process termed **chemotaxis**. The connection between chemotaxis and flagellar propulsion is dependent upon **chemoreceptors** on the cell surface that bind attractants or repellents and transmit a signal that influences the direction of flagellar rotation. A good analogy would be the blind man’s bluff game played by children, in which a person is blindfolded and moves randomly but selects among favorable or unfavorable movements toward the goal based on the responses “warmer” or “colder” (like chemoreceptors binding attractant or repellent and sending a signal to the bacteria to tumble or not to tumble). The response of flagellar rotation to chemical attractants (or repellents) is not dependent on an **absolute concentration**, but to a **change** in the concentration over time. Thus, as the bacterium moves through the solution it is able to detect whether it is moving toward or away from the highest concentration and respond accordingly.

**Pili**

Pili are long projections on the bacterial surface involved in attaching to different surfaces. The **sex pilus** is a special pilus attaching F\(^+\) (male) and F\(^-\) (female) bacteria which facilitates the formation of **conjugation bridges** (discussed below). **Fimbriae** are smaller structures that are not involved in locomotion or conjugation but are involved in adhering to surfaces. [What other bacterial structure is involved in adhering to surfaces? Is it possible that the fimbriae play a role in infection by pathogenic organisms?]

---

\[^{39}\] The capsule, or glyocalyx is also involved in adherence. And yes, fimbriae do play a role in infection, by facilitating adhesion to cells so that the bacteria can colonize a tissue.
Bacterial Growth Requirements and Classification

Temperature
Another characteristic of bacteria used to categorize them is their ability to tolerate environmental variables, such as temperature. Though bacteria as a group can grow at a wide range of temperatures, each species has an optimal growth temperature. If the temperature is too high or too low, bacteria fail to grow and may be killed, hence the use of boiling to kill bacteria and refrigeration to slow bacterial growth and prevent food spoilage. Most bacteria favor mild temperatures similar to the ones that humans and other organisms favor (30°C); they are called mesophiles (moderate temperature lovers). Thermophiles (heat lovers) can survive at temperatures up to 100°C in boiling hot springs or near geothermal vents in the ocean floor. Bacteria that thrive at very low temperatures (near 0°C) are termed psychrophiles (cold lovers). [How might a decrease in temperature increase the bacterial growth rate?]

Nutrition
Bacteria can be classified according to their carbon source and their energy source. "Troph" is a Latin root meaning “eat.” Autotrophs utilize CO₂ as their carbon source. Heterotrophs rely on organic nutrients (glucose, for example) created by other organisms. Chemotrophs get their energy from chemicals. Phototrophs get their energy from light; not only plants but also some bacteria do this. Each bacterium is either a chemotroph or a phototroph and is either an autotroph or a heterotroph. There are thus four types of bacteria:

1) **Chemoautotrophs** build organic macromolecules from CO₂ using the energy of chemicals. They obtain energy by oxidizing inorganic molecules like H₂S.
2) **Chemoheterotrophs** require organic molecules such as glucose made by other organisms as their carbon source and for energy. (We are chemoheterotrophs.)
3) **Photoautotrophs** use only CO₂ as a carbon source and obtain their energy from the Sun. (Plants are photoautotrophs.)
4) **Photoheterotrophs** are odd in that they get their energy from the Sun, like plants, but require an organic molecule made by another organism as their carbon source.

- A bacterium that causes an infection in the bloodstream of humans is most likely to be classified as which one of the following? 
  A) Chemoautotroph  
  B) Photoautotroph  
  C) Chemoheterotroph  
  D) Photoheterotroph

40 Normally, you expect decreasing temperature to decrease the rate of all chemical, biochemical, and biological processes, since reactions accelerate when kinetic energy increases. However, bacteria that have evolved to live at low temperature (psychrophiles) possess enzymes that may be optimally active at low temperature, leading to better growth.

41 Since there’s no sunlight in the bloodstream, B and D are out. If it’s a parasite, it most likely uses some of our chemicals, so it must be a heterotroph, which eliminates A. The answer is C.
• Which one of the following categories best describes an organism which uses sunlight to drive ATP production but cannot incorporate carbon dioxide into sugars?
  A) Chemoautotroph
  B) Photoautotroph
  C) Chemoheterotroph
  D) Photoheterotroph

**Growth Media**

The environment in which bacteria grow is the medium (plural: media). In the lab, the most common solid medium is agar, a firm transparent gel made from seaweed. Bacteria live in the agar but do not metabolize it. The agar is usually kept in a clear plastic plate called a Petri dish, and the process of putting bacteria on such a plate is called plating. When one bacterium is plated onto a dish, if it grows, it will eventually give rise to many progeny in an isolated spot called a colony. Minimal medium contains nothing but glucose (in addition to the agar). More key terms: A wild-type bacterium (or a wild-type strain) is one which possesses all the characteristics normal to that particular species. The dense growth of bacteria seen in laboratory Petri dishes is known as a bacterial lawn. A plaque is a clear area in the lawn. Plaques result from death of bacteria and are caused by lytic viruses or toxins.

Bacteria can reproduce very rapidly, provided that the conditions of their environment are favorable and nutrients are abundant. The doubling time is the amount of time required for a population of bacteria to double its number. It ranges from a minimum of 20 minutes for *E. coli* to a day or more for slow growers, such as the bacteria responsible for tuberculosis and leprosy. The doubling time of a bacterial species will vary, depending upon the availability of nutrients and other environmental factors.

One other important term in bacterial nutrition is auxotroph (don’t confuse this term with autotroph). This is a bacterium which cannot survive on minimal medium because it can’t synthesize a molecule it needs to live. Therefore, it requires an auxiliary trophic substance to live. For instance, a bacterium which is auxotrophic for arginine won’t form a colony when plated onto minimal medium, but if the medium is supplemented with arginine, a colony will form. This arginine auxotrophy is denoted arg⁻. Auxotrophy results from a mutation in a gene coding for an enzyme in a synthetic pathway.

Bacteria can be differentiated not only by what substances they require, but also by what substances they are capable of metabolizing for energy. For instance, a strain of bacteria may be capable of surviving on minimal medium that has the disaccharide lactose as the only carbon source (no glucose). This would be denoted lac⁺. Mutation in a gene for the enzyme lactase would impair the bacterium’s ability to survive on lactose-only medium. A bacterial strain incapable of growing with lactose as its only carbon source would be denoted lac⁻. Genetic exchange between bacteria by means of conjugation, transduction, or transformation (discussed below) can remedy these disabilities.

**Oxygen Utilization and Tolerance**

Oxygen metabolism is aerobic metabolism. Bacteria which require oxygen are called obligate aerobes. Bacteria which do not require oxygen are called anaerobes. There are three subcategories: facultative anaerobes will use oxygen when it’s around, but they don’t need it. [How much more ATP can they make

---

42 The ability to use sunlight indicates that the organism is a phototroph, and the inability to use carbon dioxide as a carbon source indicates that it is a heterotroph—it must use organic molecules as a carbon source. The answer is D.
per glucose molecule when \( \text{O}_2 \) is present?\(^{24}\) Tolerant anaerobes can grow in the presence or absence of oxygen but do not use it in their metabolism. Obligate anaerobes are poisoned by oxygen. This is because they lack certain enzymes necessary for the detoxification of free radicals which form spontaneously whenever oxygen is around.\(^{44}\) Obligate anaerobes commonly infect wounds.

- If a bacterium cannot use oxygen as an electron acceptor, is it an obligate anaerobe, a tolerant anaerobe, or a facultative anaerobe, or is it not possible to distinguish based on the information given?\(^{25}\)

- A sample of bacteria is evenly mixed into a cool liquid agar nutrient mix in the absence of oxygen and then poured into a glass-walled tube that is open to the atmosphere on top. When the agar mix cools, it solidifies, and bacterial growth is observed as shown below. How would you classify the bacteria in terms of oxygen utilization and tolerance? (Note: Agar is practically impermeable to oxygen.\(^{26}\)

\[
\text{Fermentation vs. Respiration}
\]

Respiration is glucose catabolism with use of an inorganic electron acceptor such as oxygen. In contrast, fermentation is glucose catabolism which does not use an electron acceptor such as \( \text{O}_2 \); instead, a reduced by-product of glucose catabolism such as lactate or ethanol is given off as waste. [Why is fermentation necessary whenever an external electron acceptor is not used?\(^{27}\)

\[
\text{Anaerobic Respiration}
\]

This is not a contradiction in terms! It refers to glucose metabolism with electron transport and oxidative phosphorylation relying on an external electron acceptor other than \( \text{O}_2 \). For example, instead of reducing \( \text{O}_2 \) to \( \text{H}_2\text{O} \), some anaerobic bacteria reduce \( \text{SO}_4^{2-} \) to \( \text{H}_2\text{S} \), or \( \text{CO}_2 \) to \( \text{CH}_4 \). Nitrate (\( \text{NO}_3^- \)) is another possible electron acceptor.

---

\(^{24}\) Sixteen times as much. For more on this topic, see MCAT Biochemistry Review.

\(^{25}\) The enzymes include superoxide dismutase (converts \( \text{O}_2^- \) to \( \text{H}_2\text{O}_2 \)) and catalase (converts \( \text{H}_2\text{O}_2 \) to \( \text{H}_2\text{O} + \text{O}_2 \)). An example of a harmful \( \text{O}_2 \) by-product is superoxide anion, \( \text{O}_2^- \).

\(^{26}\) The bacterium cannot be a facultative anaerobe, since the question states it cannot use \( \text{O}_2 \). It could be either an obligate or a tolerant anaerobe depending on its ability to neutralize harmful oxygen-free radicals.

\(^{27}\) Since the bacteria grew only at the bottom of the tube, farthest away from any oxygen, this indicates that they could only grow in the absence of oxygen. Thus, they are obligate anaerobes.

\(^{28}\) Because NAD\(^+\) must be regenerated from NADH for glycolysis to continue. In fermentation, the electrons are passed from NADH to a molecule other than \( \text{O}_2 \), such as pyruvic acid.
• In an experiment, facultative anaerobic bacteria that are growing on glucose in air are shifted to anaerobic conditions. If they continue to grow at the same rate while producing lactic acid, then the rate of glucose consumption will:
  
  A) increase 16 fold.  
  B) decrease 16 fold.  
  C) decrease 2 fold.  
  D) not change.

Bacterial Life Cycle

Bacteria reproduce asexually. In asexual reproduction, there is no meiosis, no meiotic generation of haploid gametes, and no fusion of gametes to form a new individual organism. Instead, each bacterium grows in size until it has synthesized enough cellular components for two cells rather than one, replicates its genome, then divides in two. This process in bacteria is also known as binary fission (fission means "to split"). [In prokaryotes, does reproduction increase genetic diversity? 49 If a eukaryote reproduces strictly by asexual reproduction, how will this affect the genetic diversity of a population? 50 How is asexual reproduction in a eukaryote different from asexual reproduction in a prokaryote? 51] Although bacteria do not reproduce sexually, they do possess a mechanism, termed conjugation, for exchanging genetic information (more on this later).

Growth of bacterial populations is described in stages (see Figure 12). Under ideal conditions, bacterial population growth is exponential, meaning that the number of bacterial cells increases exponentially with time. This also means the log of the population size grows linearly with time, hence the name log phase. [If 10 bacteria in log phase are placed in ideal growth conditions and the doubling time is 20 minutes, how many bacteria will there be after four hours? 52]
Prior to achieving exponential growth, bacteria that were not previously growing undergo a lag phase, during which cell division does not occur even if the growth conditions are ideal.

- If growth conditions are ideal, why wouldn’t cell division occur immediately?  
- Will bacteria that are transferred from a culture that is in log phase to a fresh new culture show a lag phase?

As metabolites in the growth medium are depleted, and metabolic waste products accumulate, the bacterial population passes from log phase to stationary phase, in which cells cease to divide for lack of nutrients. The maximum population at the stationary phase is referred to as the carrying capacity for that environment. In the last stages of the stationary phase, cell death may occur as a result of the medium’s inability to support growth. [If bacteria are grown in a medium with glucose as the main source of energy, when will the glycolytic pathway be more active: during the lag phase or during the stationary phase?]

**Endospore Formation**

Some types of Gram-positive bacteria, such as the bacteria responsible for botulism, form endospores under unfavorable growth conditions. Endospores have tough, thick external shells comprised of peptidoglycan. Within the endospore are found the genome, ribosomes, and RNA which are required for the spore to become metabolically active when conditions become favorable. Endospores are able to survive temperatures above 100°C, which is why autoclaves or pressure cookers are required to completely sterilize liquids and substances that cannot be heated sufficiently in a dry oven. The metabolic reactivation of an endospore is termed germination. A single bacterium is able to form only one spore per cell. Thus, bacteria cannot increase their population through spore formation. [When are bacteria most likely to

---

53 Cells that are not growing are not actively producing components that are needed for cell division, such as dNTPs. The lag period is a time when biosynthetic pathways are very actively producing new cellular components so that cells can then begin to divide.

54 No, since they will have all the gear necessary for population growth at the ready.

55 The bacteria will use glucose during the lag phase to produce ATP and cellular machinery. During this period, glucose is abundant, and the cell is actively performing biosynthesis, so glycolysis is very active. During the stationary phase, however, the glucose will be depleted, and the rate of metabolism will have slowed dramatically, so the rate of glycolysis will decrease as well.
form endospores: during lag phase, log phase, or stationary phase? Is endospore formation a means for bacteria to reproduce?\[56\]

**Genetic Exchange Between Bacteria**

Bacteria reproduce asexually, but genetic exchange is evolutionarily favorable because it fosters genetic diversity. Bacteria have three mechanisms of acquiring new genetic material: transduction, transformation, and conjugation. Note that none of these has anything to do with reproduction! Transduction was discussed in Section 5.1; it is the transfer of genomic DNA from one bacterium to another by a lysogenic phage. Transformation refers to a peculiar phenomenon: If pure DNA is added to a bacterial culture, the bacteria internalize the DNA in certain conditions and gain any genetic information in the DNA. Conjugation appears most likely to be related to normal bacterial function, however.

**Conjugation**

Although bacteria reproduce asexually, they have developed conjugation to exchange genetic information. In conjugation, bacteria make physical contact and form a bridge between the cells. One cell copies DNA, and this copy is transferred through the bridge to the other cell. A key to bacterial conjugation is an extrachromosomal element known as the F (fertility) factor. Bacteria that have the F factor are male, or F+, and will transfer the F factor to female cells. Bacteria that do not contain the F factor are female, F−, and will receive the F factor from male cells to become male. [If all cells in a population are F+ will conjugation occur?\[57\]]

The F factor is a single circular DNA molecule. Although much smaller than the bacterial chromosome, the F factor contains several genes, many of which are involved in conjugation itself. [Which cell will produce sex pili: the male cell or the female cell?\[58\]] After the male cell produces sex pili and the pili contact a female cell, a conjugation bridge forms. The F factor is replicated and transferred from the F+ to the F− cell. DNA transfer between F+ and F− cells is unidirectional; it occurs in one direction only (see Figure 13).

Although the F factor is an extrachromosomal element, it does sometimes become integrated into the bacterial chromosomes through recombination. A cell with the F factor integrated into its genome is called an Hfr (high frequency of recombination) cell. [Will an Hfr cell undergo conjugation with an F− cell?\[59\]] When an Hfr cell performs conjugation, replication of the F factor DNA occurs as in F+ cells with the extra chromosomal F factor. Since the F factor DNA is integrated in the bacterial genome in Hfr cells, replication of F factor DNA continues into bacterial genes, and these too can be transferred into the F− cell (see Figure 13).

---

56 Stationary. Forming an endospore is like hibernating, not reproducing. Bacteria do it in order to sleep through the bad times.
57 No. Conjugation occurs only between F+ (male) and F− (female).
58 The male cell contains the F factor that encodes the genes for pili production and will produce pili.
59 Yes. All of the genes of the F factor are still present and expressed normally in the Hfr cell.
If bacteria contain only one copy of the bacterial genome, how can recombination occur?\textsuperscript{260}
If the F factor in an Hfr strain integrates near a gene required for lactose metabolism, is it likely that other genes involved in lactose metabolism will be transferred during conjugation at the same time?\textsuperscript{261}

\begin{center}
\includegraphics[width=0.8\textwidth]{conjugation.png}
\end{center}

\textbf{Conjugation Mapping}
Hfr bacteria provide a mechanism of mapping the bacterial genome. By allowing Hfr cells to conjugate in the lab and stopping the conjugation process after different time intervals, researchers can figure out the order of the genes on the bacterial chromosome by analyzing recipient cells to see what genes were transferred.

For example, you have two strains of \textit{E. coli}. One is a normal Hfr bacterium. The other is F\textsuperscript{−} and auxotrophic for arginine, leucine, and histidine (F\textsuperscript{−} Arg\textsuperscript{−} Leu\textsuperscript{−} His\textsuperscript{−}). You allow conjugation to begin and stop it after 2 minutes. You find that all the recipients are now F\textsuperscript{−} Arg\textsuperscript{−} Leu\textsuperscript{−} His\textsuperscript{+}. Then you take another bunch

\textsuperscript{260} When an Hfr cell conjugates with an F\textsuperscript{−} cell and transfers a portion of the bacterial chromosomes, the F\textsuperscript{−} cell will have two copies of some genes, and recombination can occur between the two copies.

\textsuperscript{261} Yes. Genes for proteins of related functions are often adjacent to each other in prokaryotes (in operons) and so will transfer to an F\textsuperscript{−} cell together.
of bacteria and allow conjugation to proceed for 5 minutes. Now all the recipients are F' Arg' Leu' His'. You do the experiment a third and final time, allowing 8 minutes of conjugation, and find the recipients to be F' Arg' Leu' His'.

- What is the arrangement on the genome of the enzymes responsible for synthesis of each amino acid, relative to the site of F plasmid integration?62

**Domain Archaea**

Though all bacteria are prokaryotes, not all prokaryotes are equal. Certain prokaryotes belong to the domain *Archaea*, to be distinguished from the more “typical” bacteria (or eubacteria) which we have just discussed. The Archaea are the organisms that live in the world’s most extreme environments, including hot springs, thermal vents, and hypersaline environments (although they can also be found in less extreme environments, such as soil, water, the human colon, etc.). Structurally, they differ from other bacteria because their cell wall lacks peptidoglycan. Genetically, they share traits with eukaryotes including the presence of introns and the use of many similar mRNA sequences. However, since they are single celled, they do reproduce via fission or budding. [What does this mean for their ability to increase their genetic diversity?63]

Since Archaea have to produce enzymes that can function in extreme environments, they are of great use in industrial applications, such as food processing and sewage treatment. The development of applications for products from these cells is an ongoing area of research.

**Parasitic Bacteria**

Parasitic bacteria can either be *obligate*, meaning that they must be inside a host cell to replicate, or *facultative*, meaning that they can live and replicate inside or outside of a host cell. In either case, the designation as a *parasite* means that damage is being done to the host cell. However, in order to ensure a continued supply of energy and cellular materials needed to survive and reproduce, parasitic bacteria need to modulate the course of that damage. [How is this model similar to viruses? 64]

T cells (lymphocytes involved in immunity; see Chapter 9) are responsible for monitoring cellular contents; people who are T cell deficient have a hard time fighting off these types of bacterial infections, just as they would also struggle with viral infections. *Mycobacteria*, the genus of bacteria which encompasses the cause of tuberculosis as well as other diseases, has members which are obligate and others which are facultative, whereas the sexually transmitted disease chlamydia is caused solely by an obligate parasitic bacteria.

---

62 The experiments showed that the ability to make histidine was transferred in a short time. After a slightly longer time, the ability to make both histidine and arginine was transferred. Lastly, the ability to make leucine were transferred. So the arrangement on the genome (the map) must be: His-Arg-Leu-plasmid integration site.

63 Archaea would need to use separate strategies to increase their genetic diversity, just like eubacteria. The ability to become more genetically diverse would not be built into reproduction as it is in humans, in part because meiosis is not occurring.

64 Viruses are obligate intracellular parasites. They do not have the option to replicate outside of a host cell, but must also balance the damage that is done to the host cell against what is needed for more virus to be made.
Symbiotic Bacteria
Symbiotic bacteria coexist with a host, where both the bacterial cell and the host cell derive a benefit. An example of this would be the *Rhizobia* genus, which is responsible for the fixing of nitrogen in the nodules that exist on the roots of legumes. Without these bacteria the legume plants would not be able to grow, as they would be unable to derive the necessary nitrogen from the soil on their own. Similarly, *Cyanobacteria* are responsible for nitrogen fixing in marine environments. Some of the bacterial flora in the human gut is also composed of symbionts which aid the human body in defending against other pathogenic strains. [What other functions do the bacteria in the gut have?]

Due to their close relationship with their host cells, these bacteria often have smaller genomes with a more limited number of cellular products that are made, since the host cells can provide some of what the bacteria need. This can often mean that the symbiotic bacteria do not survive long outside of the host environment.

---

65 The gut flora is responsible for the production of vitamin K, which is necessary for blood clotting and in feeding off of undigested material from what humans have consumed; they are one of the final stages in processing our solid waste for excretion.
Chapter 5 Summary

- All viruses are made up of nucleic acids (either RNA or DNA) surrounded by a protein coat (capsid). They are obligate intracellular parasites and must rely on other cells to reproduce.

- Animal viruses may also have an envelope [lipid bilayer] surrounding the capsid. The envelope is derived from the host cell and is acquired by budding through the host cell membrane.

- Viral infection is specific; molecules on the viral surface determine which type of host cell it will infect.

- Viruses replicate via two major life cycles, the lytic cycle [in which more virus is made very quickly] and the lysogenic cycle [in which the virus goes dormant by integrating into the host cell genome]. Viruses in the lysogenic cycle can excise from the genome and enter the lytic cycle.

- Animal viruses can also participate in a third life cycle, the productive cycle. This is very similar to the lytic cycle, but the new viruses escape by budding instead of by lysing the host.

- Lysogenic viruses can take pieces of the host DNA with them when they excise and transfer it to the next host. This is called transduction.

- RNA viruses require special virus-derived enzymes [RNA dependent RNA polymerases] in order to replicate their genomes.

- Prions and viroids are subviral particles that can cause disease and infection. They are unique in that prions are simply abnormal proteins [with no genetic material] and viroids are small pieces of RNA with no associated capsid, that do not code for proteins.

- The primary difference between prokaryotes and eukaryotes is that prokaryotes have no membrane-bound organelles [e.g., nucleus, mitochondria, etc.], thus all cellular processes occur in the cytosol.

- The shapes of bacteria can be used to classify them [round = coccus, rod = bacillus, spiral = spirochete].
Bacteria have cell walls made out of peptidoglycan that can bind crystal violet (a purple stain used in Gram stain). Gram-positive bacteria have thick cell walls and stain a dark purple. Gram-negative bacteria have thinner cell walls and an outer membrane; they stain a light pink.

Some bacteria can be classified by the presence or absence of flagella. Bacterial flagella are used for motility and are distinct from eukaryotic flagella in structure.

Preferred growth temperature, nutrition, and oxygen use/tolerance are means of characterizing bacteria and can be used to select for growth of a particular bacteria.

Binary fission is a means of asexual bacterial reproduction that increases the population size exponentially, but does not increase the genetic diversity of the population.

Conjugation is a means of increasing genetic diversity in a bacterial population by exchanging DNA (plasmid or genomic) via a conjugation bridge.

Bacteria in Domain Archaea are sometimes classified as extremophiles because they can live in harsh, extreme environments, like hot springs, thermal vents, extremes acids/bases, and hypersaline environments.

Parasitic bacteria can live inside or outside of host cells and harm the host cells. Symbiotic bacteria coexist with host cells, but provide the host cells with a benefit; for example, the gut bacteria provide us with vitamin K.
CHAPTER 5 FREESTANDING PRACTICE QUESTIONS

1. A researcher has an agar plate covered with a lawn of *E. coli*. She adds a drop of a substance, and the next day there is a clear spot on the plate where the substance was added. This substance could be:

I. a virus undergoing the lytic cycle.
II. a virus undergoing the productive cycle.
III. a chemical that is toxic to prokaryotes.

A) I only
B) III only
C) I and III
D) I, II, and III

2. A lab technician grows a liquid bacterial culture overnight, in media without any antibiotics. The next morning, the culture is cloudy. She takes a small amount of this culture and puts it into new media containing tetracycline. The next day, she checks the culture and the media is not cloudy. What happened?

A) The bacterial culture grew the first night but not the second night.
B) The bacteria were resistant to the antibiotic tetracycline.
C) The bacteria were in the lag phase after the first night of growth.
D) The bacteria were in the stationary phase after the second night of growth.

3. Which of the following is associated with prokaryotes and does NOT introduce new genetic material?

A) Mitosis
B) Binary fission
C) Transformation
D) Transduction

4. Which of the following statements concerning viruses is true?

A) The productive cycle is the most efficient infective cycle for phages.
B) Viruses that infect human cells must have an envelope.
C) Genetic information can be transferred between hosts via transfection.
D) A virus with an RNA genome must code for an RNA-dependent RNA polymerase.

5. A researcher is trying to characterize a novel prokaryotic organism that has been found in the Indian Ocean. When Gram stained, the cells are a light pink color under the microscope. When exposed to antibiotics commonly used in the lab, the bacteria are able to enter the log growth phase in a manner similar to *E. coli* grown in media lacking ampicillin. A reasonable explanation is that:

A) this is a Gram-positive bacterium with an additional lipopolysaccharide layer that increases their resistance to antibiotics.
B) this is a Gram-positive bacterium with a cell membrane outside its peptidoglycan layer that increases their resistance to antibiotics.
C) this is a Gram-negative bacterium with an additional lipopolysaccharide layer that increases their resistance to antibiotics.
D) this is a Gram-negative bacterium with a peptidoglycan layer outside the cell membrane that increases their resistance to antibiotics.

6. Which of the following is true regarding prokaryotic flagella?

A) It is the predominant form of bacterial locomotion.
B) It is made of microtubules connected by dynein proteins.
C) It allows viruses to maneuver between host cells.
D) It can only be located on one end of a bacterium and this defines the polarity of the cell.

7. In prion diseases like Creutzfeldt-Jakob disease (CJD), the characteristic misfolded proteins are notoriously resistant to degradation. As a result, abnormal proteins accumulate in the endosomes and lysosomes of the cell, eventually leading to cellular dysfunction and death through a chronic, neurodegenerative process. Which of the following is the most likely explanation for the unusual resistance of these proteins to breakdown?

A) Accelerated rate of protein biosynthesis leads to early cell lysis
B) Excessive amount of protein accumulation results in cellular dysfunction
C) Aberrant protein function leads to disruption of normal cellular processes
D) Abnormal protein secondary structure results in poor binding with innate lysosomal proteases
Laboratory tests are a useful diagnostic tool for determining the cause of illness. One such test is the complete blood count or, CBC, as it is commonly called. In an infected individual, the white blood cell count can increase from a normal range of 4000–10,000 cells/μL to 15,000 to 20,000 cells/μL. Circulating neutrophils have a short lifespan upon release from the bone marrow (generally about ten hours); however, the demand for phagocytic cells during an infection increases markedly. The result is the release of immature neutrophils called band cells. In a differential white blood cell count, the presence of band cells is referred to as a shift to the left. A decrease in the number of neutrophils (neutropenia) can also occur as a result of inflammation or severe infection, when the removal of the neutrophils from the circulation outpaces their production. Neutropenia is also seen in certain blood cancers, such as leukemia and lymphomas, as the neutrophil precursor cells are crowded out by the cancerous cells.

The type of microorganism responsible for an infection can often be determined by changes identified in the population of white blood cells. For example, an increase in neutrophils (neutropenia) can also occur as a result of inflammation or severe infection, when the removal of the neutrophils from the circulation outpaces their production. Neutropenia is also seen in certain blood cancers, such as leukemia and lymphomas, as the neutrophil precursor cells are crowded out by the cancerous cells.

One busy spring Saturday evening in a hospital emergency room, several patients are presented with respiratory complaints. The patients had either a productive (mucus-producing) or nonproductive cough. All patients presented with some form of fever, either mild or severe. A CBC was ordered on each patient. In addition, a blood sample was obtained from each patient for the purpose of culturing and identifying any infection-causing bacteria. Standard growth media (growth media that contains glucose, amino acids, and some vitamins) providing sufficient nutrients for a wide range of bacteria was used for this purpose. The following results were obtained:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Growth in culture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Table 1** Bacterial Culture Results in Four Different Patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Elevated eosinophils</th>
<th>Elevated neutrophils</th>
<th>Elevated lymphocytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

**Table 2** WBC Counts in Four Different Patients

1. If placed on a course of antibiotic therapy, which of the following patients would feel significantly improved after approximately 1–2 days?
   A) Patient 1
   B) Patient 2
   C) Patient 3
   D) Patient 4

2. Why did the culture performed on the sample obtained from Patient 3 not yield any growth?
   A) The bacteria causing the infection in Patient 3 is a uracil auxotroph.
   B) Patient 3 has a bacterial infection.
   C) Patient 3 is suffering from allergic symptoms.
   D) A different growth medium was required.

3. Patient 4 was prescribed a broad-spectrum antibiotic and released. The patient returned to the emergency room after a two-week period complaining of worsening symptoms. The patient admitted to discontinuing use of the antibiotic after four days of therapy because they felt much improved. Which of the following are possible explanations for the patient’s symptoms?
   I. The susceptible bacterial population was not fully eradicated.
   II. The patient was resistant to the antibiotic.
   III. A resistant population of bacteria has begun to proliferate.
   A) I only
   B) II only
   C) I and III only
   D) I, II, and III
4. To determine the most appropriate type of antibiotic to prescribe, which of the following additional tests could be performed on a patient sample for classification purposes?
   A) Phage-typing
   B) Gram-staining
   C) Fermentation
   D) Transduction

5. Which of the following bacterial types are LEAST likely to cause a respiratory infection?
   A) Tolerant anaerobe
   B) Obligate aerobe
   C) Facultative anaerobe
   D) Obligate anaerobe

6. If a patient’s symptoms included neutropenia and elevated lymphocyte counts, which of the following diagnoses could be possible?
   I. Allergies
   II. Leukemia/lymphoma
   III. Viral infection
   A) I only
   B) II only
   C) I and III only
   D) II and III only

7. An experimental therapy to treat patients with multiple antibiotic-resistant bacteria involves introduction of a highly specific bacteriophage to the infected patient’s bloodstream. Which of the following bacteriophage types would be the LEAST useful for this type of therapy?
   A) A lytic bacteriophage
   B) A lysogenic bacteriophage
   C) An RNA virus
   D) An enveloped virus
SOLUTIONS TO CHAPTER 5 FREESTANDING PRACTICE QUESTIONS

1. **C** A clear spot on a plate (known as a plaque) indicates that the *E. coli* are dead. This could be due to the addition of a lytic virus (Item I is true, and choice B can be eliminated) or toxin (Item III is true, and choice A can be eliminated). However, only animal viruses can go through the productive cycle because viruses cannot bud out of a cell with a cell wall, such as bacteria (Item II is false; choice D can be eliminated, and choice C is correct).

2. **A** Cloudy cultures are usually in the stationary phase and clear cultures are either not growing or still in the lag phase. Since the culture was cloudy on the first morning, bacteria had grown overnight and were most likely in stationary phase (choice A is correct, and choice C is wrong). The culture on the second morning was clear, indicating minimal growth (choice D is wrong). Since the first overnight culture did not contain tetracycline and the second overnight culture did, it is possible that the strain was sensitive to tetracycline, not resistant (choice B is wrong).

3. **B** Binary fission is the means by which bacteria divide and reproduce. It produces two progeny cells that are genetically identical to the parent; no new genetic information is introduced (choice B is correct). Although mitosis also does not introduce new genetic information, it is a process undergone by eukaryotic cells, not prokaryotes (choice A is wrong). Both transformation and transduction are associated with prokaryotes, but both involve the introduction of new genetic material. Transformation is the uptake of genetic material (plasmids or chromosomal DNA) from the extracellular environment (choice C is wrong) and transduction is the transfer of genetic information from one bacteria to another via a lysogenic phage (choice D is wrong).

4. **D** In order to replicate its genome, an RNA virus must code for an RNA-dependent RNA polymerase; this enzyme will create a new strand of RNA by reading a template strand of RNA. Viral host cells will not express these enzymes naturally; they have no need to make RNA by reading RNA. Host cells normally produce RNA using DNA as a template (choice D is correct). Phages only infect bacteria, and can only undergo the lytic and lysogenic cycles; the productive cycle involves budding through cell membrane and cannot occur in hosts with cell walls, such as bacteria (choice A is wrong). Although viruses with an envelope (lipid bilayer coating) are restricted to infecting animal cells, the outer membrane is not required (choice B is wrong). Genetic information can indeed be transferred between hosts, but this process is called *transduction*, not *transfection* (choice C is wrong).

5. **C** The answer options all start with Gram positive or Gram negative, and the light pink staining in the question stem indicates that this bacteria is Gram negative. Gram-positive bacteria have a peptidoglycan layer outside the cell membrane and therefore stain a dark purple (choices A and B are wrong). Gram-negative bacteria have a lipopolysaccharide layer outside the peptidoglycan layer. This additional outer layer prevents dark staining (thus the light pink color) and increases resistance to antibiotics (choice C is correct, and choice D is wrong).
6. A  Prokaryotic flagella are the predominant means of bacterial locomotion (choice A is correct). Only eukaryotic flagella are made of microtubules and dynein; bacterial flagella have a different structure and are made of the protein flagellin (choice B is wrong). Viruses rely on diffusion to maneuver between host cells, not flagella (choice C is wrong). Bacteria can have flagella on one end (monotrichous), both ends (amphitrichous), or in multiple places (peritrichous; choice D is wrong).

7. D  Typically in a cell, abnormal proteins are targeted to proteasomes in the cytosol for degradation or to the lysosomes for digestion. The accumulation of the abnormal prions in the lysosomes suggests that they are somehow resistant to the lysosomal enzymes, and this could be due to their abnormal structure that prevents efficient binding at the active sites of the lysosomal digestive enzymes (choice D is correct). CJD and other prion diseases do not involve accelerated protein biosynthesis; normal prion proteins are produced at their typical rate, but are converted after translation into the abnormal prion version (choice A is wrong). While the accumulation of misfolded proteins likely contributes to cellular dysfunction, and while prions could (and likely do) lead to disruption of normal cellular processes, neither of these explain why the abnormal proteins are resistant to breakdown (choice D is a better choice than either B or C).

SOLUTIONS TO CHAPTER 5 PRACTICE PASSAGE

1. D  Antibiotics only treat bacterial infections. Table 1 shows that Patients 2 and 4 have positive cultures, confirming a bacterial infection (choices A and C can be eliminated). However, from Table 2, only Patient 4 has elevated neutrophils, which the passage states is indicative of a bacterial infection, making this patient a candidate for antibiotic therapy. Patient 2 has elevated lymphocytes (not neutrophils), which indicates a concomitant viral infection; the lack of elevated neutrophils in this patient is most likely the result of overwhelming infection. While the bacterial infection of Patient 2 would begin to subside by 1–2 days, the viral infection would take longer to eradicate, and therefore this patient would still be feeling poorly at this point (choice D is better than choice B).

2. D  From Table 2, Patient 3 has elevated lymphocytes, indicating a viral infection (choices A, B, and C can be eliminated); no growth would occur on standard growth media. Viruses are obligate intracellular parasites and require special growth media containing live cells for reproduction. According to the passage, standard growth media for bacterial culturing was used. Note that if the patient had elevated neutrophils and normal lymphocyte levels (indicating a bacterial infection), choice A would be the best answer, since uracil was not added to the culture media.

3. C  Item I is true: If the symptoms return and/or worsen, it is possible that the infection was not completely eradicated (choice B can be eliminated). Item II is false: Bacteria are resistant, the patient is not (this is a common misconception; choices B and D can be eliminated). Item III is true: It is also possible that a strain resistant to the antibiotic being used remained alive and is now proliferating (choice C is correct).
4. B Determination of Gram status can aid in antibiotic selection. Phage typing applies to bacteriophage, not bacteria (choice A can be eliminated). Fermentation refers to metabolic activity (choice C can be eliminated). Transduction is not a classification method (choice D can be eliminated).

5. D The lungs are a high oxygen environment that is unfavorable to obligate anaerobes. Tolerant anaerobes, obligate aerobes, and facultative anaerobes can all survive in the presence of oxygen (choices A, B, and C can be eliminated).

6. D Item I is false: The passage states that elevated eosinophil counts (not lymphocytes) accompany allergic reactions (choices A and C can be eliminated). Note that both remaining answer choices include Item II, so it must be true: The passage states that neutropenia and elevated lymphocyte counts are seen in lymphoma and leukemia (blood cancers). Item III is true: Neutropenia and lymphocytosis are seen in viral infections (choice B can be eliminated, and choice D is correct).

7. D Enveloped viruses infect only animal cells and would not be useful in eliminating bacteria from a human patient. The most useful type of virus would attack the infecting bacteria and cause the bacteria to lyse, eradicating the patient’s infection (choice A would be useful and can be eliminated), and this virus could have either an RNA or a DNA genome (choice C could be useful and can be eliminated). A virus that incorporates itself into the bacterial genome and then goes dormant (a lysogenic virus) would not be as helpful as a lytic virus in eradicating an infection; however, it would still be more useful than a virus that cannot infect bacteria at all (choice B would be more useful than choice D).