Calculations in the molecular biosciences 3

Part 1: Characterization of biological molecules

The next two chapters will explain the main types of calculations involved in studying proteins. In this chapter, we shall first of all deal with some general principles important in any calculations, and then concentrate on how to study the individual components of reactions. This includes the preparation and dilution of solutions and how spectrophotometry can be used to check concentrations in some cases. The behaviour of acids, bases, and buffers is described and this is followed by a discussion of the specific activities of proteins and the manner in which their purification is recorded. Chapter 4 will concentrate on the analysis of a number of processes and reactions of biological interest.

At the end of this chapter, there are several problems for you to check your understanding of the material covered.

If you wish to explore the topics covered here in more detail, the books mentioned in Chapter 1 can be consulted.

3.1 The golden rules for successful calculations

KEY CONCEPTS

- Working out the strategy for a calculation
- Setting out and explain the working
- Making sure the number of significant figures and the units are correct

Whenever you do a calculation, make sure that the answer to each of the following questions is 'yes'. This will not absolutely guarantee success but will go a very long way towards that goal.

- Have I worked out the correct strategy for the calculation?
- Is my working clearly set out and explained?
- Is the value obtained reasonable?
- Is it reported to the appropriate number of significant figures?
- Is it quoted in the correct units?

The aim of this chapter is to help you carry out calculations successfully in a number of topics covering the purification and characterization of proteins.

We can expand these points as follows:

Have I worked out the correct strategy for the calculation? Before starting to write down your solution, make sure that you have worked out how you will proceed from the information supplied in the problem you are trying to solve to the point where you can calculate the final answer. If the path is not immediately obvious, it can be a good idea to think about what you can deduce from the information and then try to think of what might be required immediately before the final step. Working from both ends in this way can often help to define the strategy.

Is the working clearly set out and explained? It is very important to break down the calculation into steps and use words to explain. Do not just write down numbers without explanation. If you do, you will not convince anyone else reading your work that you know what you are doing. If you make a small slip in the first step and then this is (correctly) carried through subsequent steps, you should obtain credit for this; if you have not explained what you are doing and the final result is incorrect it is highly unlikely that you will obtain any credit at all.

Is the value obtained reasonable? Sometimes it is not clear what sort of value might be expected, although it is usually possible to check that the outcome of the final calculation step is of the right order of magnitude (see Chapter 2, section 2.1). However, in many cases it is possible to think about whether the value is reasonable.

It is difficult to overestimate the importance of the need for this check.

For example, one would expect to find small integral values for the number of ligand-binding sites on a protein, e.g. one substrate binding site per polypeptide chain. Values of 0.0009 or 926 sites per polypeptide chain are very unlikely and suggest that a serious error has been made in the calculation, probably involving either the use of incorrect units or mistakes in handling powers of 10. A second example would be if you were asked to use some experimental chemical modification data to calculate the number of cysteine side chains in a protein of molecular mass 30 kDa. Before starting, one can make a useful estimate as follows. A 30 kDa protein will contain about 30 000/110 amino acids, i.e. 273 amino acids (see Chapter 2, section 2.1). There are 20 different kinds of amino acids in proteins, so if they occurred with equal frequency (which in practice they do not, but that is not important here), there would be about 14 of each type. Thus, we might expect to find about 14 cysteines in each molecule of the protein. In fact, cysteine is among the rarer amino acids so in most proteins of this size the actual value would be less than 14. The key point is that a value of, for example, 538 would clearly be incorrect. In addition, a value of 0.25 would be incorrect since the value should be either zero or an integral value such as 1, 2, etc.

The average frequency of occurrence of amino acids in proteins is given in Chapter 1, Table 1.1. Leucine and alanine occur most frequently (9.6% and 7.8%, respectively); cysteine (1.5%) and tryptophan (1.2%) occur least frequently. These are only average figures and do not necessarily apply to any particular protein, for example the small protein metallothionein that plays a role in storing a number of heavy metal ions consists of about 30% cysteine residues.

Is the appropriate number of significant figures used? This has been dealt with in Chapter 2, section 2.2; remember to use the display in the calculator with caution when writing down the final answer.

Is the answer quoted in the correct units? The subject of units of quantities is discussed in section 3.3. You should get into the habit of not only quoting the correct units of the final answer, but also quoting the units at the various stages of the calculation. This will act as a guide to the direction of the calculation process, and will also help to convince anyone else that you know what you are doing!

SELF TEST ?

The key concepts at the start of the section can be checked at the end of section 3.3.

3.2 Magnitudes of quantities

KEY CONCEPTS

- Appreciating the range of sizes of quantities used in describing biological systems
- Performing calculations involving powers of 10

Molecular biology and biochemistry are quantitative sciences; the values of quantities we discuss range from the very large (numbers of molecules, Avogadro's number, frequency of radiation, etc.) to the very small (dimensions of cells, organelles and molecular complexes, concentrations of hormones, etc.); the range of dimensions of biological objects is illustrated in Fig. 3.1. This range of magnitudes means that you need to develop confidence in handling powers of 10 and the prefixes used for quantities (section 3.3).

Fig. 3.1 shows the advantage of using a logarithmic scale. For example, if we wished to represent the range of dimensions from 10 nm to 100 m on linear scale graph paper with a grid where each 1 mm represents 10 nm, then to represent 100 m, we would need a piece of paper 10^4 km long (more than the distance from London to Los Angeles!).

SELF TEST ?

The key concepts at the start of the section can be checked at the end of section 3.3.

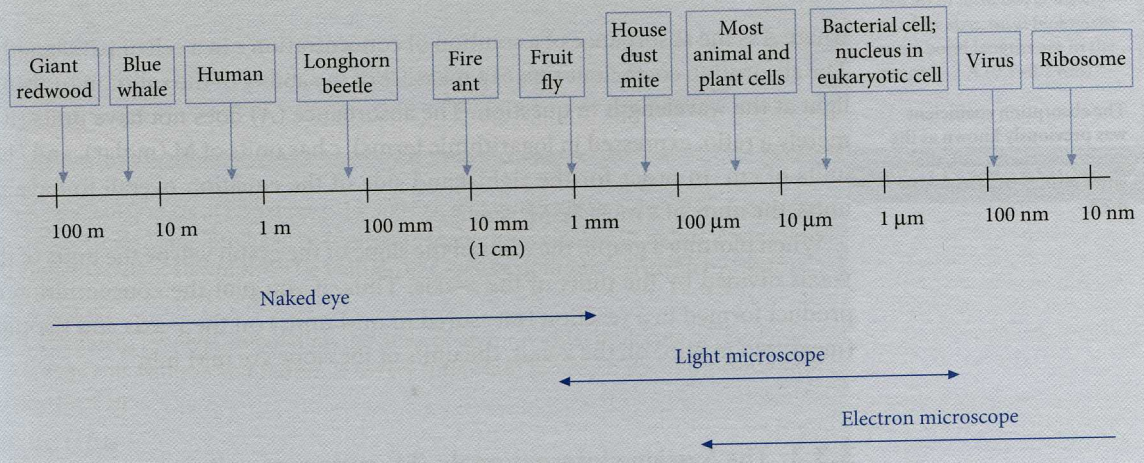

Fig. 3.1 The dimensions of biological objects. The scale is logarithmic, with each division representing a factor of 10. The dimensions associated with each object are for guidance only as there will be a range of values in each case. Individual atoms and small molecules would occupy the zone between 0.1 nm (100 pm) and 1 nm. The lower arrow bars indicate the observation ranges for the naked eye, light microscope, and electron microscope. Highly sophisticated electron microscopes can resolve objects down to about 0.2 nm (200 pm) in size.

3.3 Units of quantities

KEY CONCEPTS

- Knowing the SI prefixes for units and interconvert them as appropriate
- Using correct units in equations and graphs

Virtually all quantities we come across have units, for example mass (e.g. kg, kDa), length (e.g. mm, nm), concentration (e.g. g L^{-1}, mol L^{-1} (or M)), time (e.g. ms), etc. It is essential to state these units in any description of a system or during calculations. The only exceptions arise where the quantity represents a ratio of two values, e.g. the absorbance (A) of a solution is defined in terms of the ratios of intensities of two light beams (incident on the solution and transmitted by the solution) and therefore does not have units.

It should be pointed out that equilibrium constants (see Chapter 4, section 4.1) raise an apparent contradiction; strictly speaking they are dimensionless since all the terms in the definition refer to ratios to the standard state for that component. However, since for solutes, the values (of concentrations) are each expressed relative to a standard state of 1 M concentration (see Chapter 4, section 4.1), equilibrium constants are usually quoted with units of M to indicate this.

It is important to refer to a 1 M solution for calculation purposes, i.e. a K_d of 50 μM should be expressed as 50×10^{-6} M (or 5×10^{-5} M).

Remember that in any equation, the units must balance on the two sides. For example, in spectrophotometry the Beer-Lambert law (eqn. 3.1; see section 3.6) states that:

$$A = \varepsilon \times c \times l \qquad 3.1$$

where A is the absorbance of a solution of concentration c in a cell of path length l. The absorption coefficient ε gives a measure of the ability of the solution to absorb light at the wavelength in question. The absorbance (A) does not have units (it is merely a ratio, expressed in logarithmic terms), c has units of M (molar), and l has units of cm. In order for the right-hand side of the equation overall to have no units, the units of ε must be M^{-1} cm^{-1}.

The absorption coefficient was previously known as the extinction coefficient; the latter term is still frequently used.

When plotting a graph, the units of the slope of the graph will be the units of the y-axis divided by the units of the x-axis. Thus, if you plot the concentration of product formed in a reaction (measured in mM units) on the y-axis and the time (measured in min) on the x-axis, the units of the slope are mM min^{-1}.

3.3.1 The Système Internationale (SI) system of units

The SI system of units is based on the metre (m), kilogram (kg), second (s) quantities and units derived from these. Information on these units and units derived from them as well as the values of physical constants is included in Appendix 3.1 at the end of this chapter. It uses prefixes to bring numerical values of quantities into the 'easy to scale' range, i.e. between 0.1 and 100; thus 800 nm = 0.8 μm;

2000 μm = 2 mm, etc. The most important prefixes involve factors of 1000, so we can set up a scale as shown in Table 3.1:

Table 3.1 The prefixes used in the SI system of units

Prefix	G, giga	M, mega	k, kilo	Base unit	m, milli	μ, micro	n, nano	p, pico	f, femto
Numerical value	10^9	10^6	10^3	1	10^{-3}	10^{-6}	10^{-9}	10^{-12}	10^{-15}

We only need to go outside the range shown on very rare occasions; for interest 10^{12} is known as Tera (abbreviated 'T') and 10^{-18} is known as atto (abbreviated as 'a').

One example of the use of the prefix T would be in denoting the frequency of radiation. Thus, red light of wavelength 700 nm has a frequency of 4.3×10^{14} Hz (hertz) or 430 THz.

It is very important indeed for you to be able to convert quantities between the different prefixed units rapidly and reliably, for example to show that 5000 nm = 5 μm and that 0.012 μmol = 12 nmol.

Some intermediate prefixes are retained for convenience, e.g. c (centi, 10^{-2}), as in cm; d (deci, 10^{-1}), as in dm (note 1 litre = 1 dm^3). For measures of volume there would be an inconveniently large factor of 10^9 between 1 m^3 and 1 mm^3, hence the use of 1 dm^3 and 1 cm^3 (1 cm^3 = 1000 mm^3; 1 dm^3 = 1000 cm^3).

WORKED EXAMPLE

The length of the C–N bond in the peptide bond unit is 0.132 nm. Express this in terms of μm, pm and also in terms of Angstrom units (Å), which are commonly used to describe protein structures (1 Å = 10^{-8} cm).

STRATEGY

This is a straightforward application of the interconversion of prefixes in the SI system of units.

SOLUTION

1 nm = 10^{-3} μm, so 0.132 nm = 1.32×10^{-4} μm = 132 pm = 1.32 Å

The Ångstrom unit (Å) is not an SI unit, but is widely used by structural biologists; most bond lengths are in the range 1–2 Å (0.1–0.2 nm).

WORKED EXAMPLE

The molecular mass of lysozyme is 14 300 Da, and its density is 1.4 g mL^{-1}. Assuming that the lysozyme molecule can be regarded as spherical, calculate its radius. Avogadro's number is 6.02×10^{23} mol^{-1}. The volume of a sphere of radius r is given by $(4\pi/3)r^3$.

STRATEGY

This problem involves calculating the mass of an individual molecule of lysozyme and hence its volume and radius. It is very important to keep track of the units and powers of 10.

SOLUTION

The mass of 1 mol lysozyme is 14 300 g (see section 3.4).

Hence the mass of 1 molecule = $14\,300/(6.02 \times 10^{23})$ g = 2.375×10^{-20} g.

Since volume = mass/density, the volume of 1 molecule = 1.697×10^{-20} mL (cm^3).

So $(4\pi/3)r^3 = 1.697 \times 10^{-20}$, hence $r^3 = 4.05 \times 10^{-21}$, therefore $r = 1.59 \times 10^{-7}$ cm, or 1.59 nm (note 1 cm = 10^{-2} m; 1 nm = 10^{-9} m).

Thus, the radius of a lysozyme molecule = 1.59 nm.

In this calculation, the cube root can be readily derived by using logarithms (see Chapter 2, section 2.3).

? SELF TEST

Check that you have mastered the key concepts at the start of sections 3.1–3.3 by attempting the following questions.

ST 3.1 A pipetting device is stated to be capable of delivering 2.5 mm^3 portions of a solution. Express this volume in terms of mL and μL.

ST 3.2 A Scatchard plot is often used to analyse binding data. In this plot, $r/[L]_{free}$ is plotted against r (r is a ratio and therefore has no units; $[L]_{free}$ is the concentration of free ligand and in the case in question has units of μM). What are the units of the slope of the graph?

ST 3.3 Calculate the radius of a bacterial ribosome which can be assumed to be spherical. The molecular mass of the ribosome is 2.7 MDa and the density is 1.4 g mL^{-1}. Avogadro's number is 6.02×10^{23} mol^{-1}. The volume of a sphere of radius r is given by $(4\pi/3)r^3$.

Answers

ST 3.1 The volume delivered = 2.5×10^{-3} mL or 2.5 μL.

ST 3.2 The units of the slope of the graph are μM^{-1}.

ST 3.3 The radius of the ribosome = 9.1 nm.

Attempt Problem 3.1 at the end of the chapter.

3.4 The concepts of moles and molarity

KEY CONCEPTS

- Distinguishing clearly between moles and molarity
- Knowing the molecular mass of the solute, interconverting concentrations expressed in terms of mass/volume and in terms of molarity

It is important to note that the term mole can refer to molecules, atoms, subatomic particles, photons, etc.

One mole (abbreviated mol), which is the gram formula mass of a substance, contains Avogadro's number (6×10^{23}, or more accurately 6.022×10^{23}) of entities (which can be molecules, atoms, ions, etc.).

Substances react with one another in terms of moles, not in terms of masses. Thus, in the hydrolysis of urea:

$$CO(NH_2)_2 + H_2O \rightarrow CO_2 + 2NH_3$$

1 mol of urea (60 g) reacts with 1 mol water (18 g) to give 1 mol carbon dioxide (44 g) and 2 mol ammonia (2×17 g, i.e. 34 g).

We can also refer to 1 mol of atoms (e.g. H atoms or Cl atoms) or ions (e.g. H^+ ions or electrons); in each case this will mean the Avogadro number of that species.

3.4.1 Molarity of solutions

Virtually all biochemical reactions take place in solution. (A solution consists of one or more *solutes* dissolved in the *solvent*). A 1 molar (M or mol dm^{-3}) solution contains 1 mol of the solute in 1 litre (L or dm^3) of solution.

KEY RELATIONSHIP

number of moles = concentration (expressed in M) × volume (expressed in L)

WORKED EXAMPLE

A certain solution of NADH has a concentration of 0.6 mM. How many moles of NADH are present in 0.5 mL of this solution?

STRATEGY

This is an application of the relationship: number of moles = concentration × volume

SOLUTION

1 L of the solution contains 0.6 mmol. Hence 0.5 mL contains 0.3×10^{-6} mol, i.e. 0.3 μmol.

NADH (nicotinamide adenine dinucleotide, reduced form) is a biological reducing agent, which participates in redox reactions catalysed by dehydrogenases, such as lactate dehydrogenase (pyruvate + NADH + H^+ ⇌ lactate + NAD^+). The structure of NADH is given in Chapter 1, section 1.8).

If the molecular mass of the solute is X dalton (Da), then 1 mol = X g (e.g. for glucose $C_6H_{12}O_6$ the gram formula mass would be 180 g).

A 1 M solution therefore contains X g solute L^{-1} (for glucose, this would be 180 g L^{-1}).

With practice, you should be able to work out rapidly how many mol of a solute are present in a certain volume of solution of specified molarity. Most biochemical reactions involve small volumes of dilute solutions, so we are likely to be dealing with μmol, nmol or even pmol quantities of solute.

We can readily work out the concentrations of solutions as follows:

A solution containing 1 g solute of molecular mass X per litre would be $(1/X)$ M (for glucose this would be (1/180) M = 5.56×10^{-3} M).

For biochemistry, g and L are large quantities; mg and mL (or even μg and μL) are much more likely to represent the scale of experiments undertaken and to take account of the high costs of many of the reagents involved.

KEY RELATIONSHIP

In terms of concentration, 1 μg μL^{-1} = 1 mg mL^{-1} = 1 g L^{-1}.

Thus, we can always calculate and obtain the molar concentration of a solution by dividing the mg mL^{-1} (or g L^{-1} or μg μL^{-1}) concentration by the molecular mass (in Da).

Thus, a 0.2 mg mL^{-1} solution of glucose = 0.2/180 M = 1.11×10^{-3} M (1.11 mM).

BSA is an abundant protein in the serum of cattle blood. It can bind a number of non-polar substances and serves to transport fatty acids in the blood stream.

A 0.15 mg mL^{-1} solution of bovine serum albumin (BSA) (molecular mass 66 000 Da) = 0.15/66 000 M = 2.273 μM.

WORKED EXAMPLE

The concentration of protein required for an nuclear magnetic resonance experiment is 0.5 mM. If the molecular mass of the protein is 27.5 kDa, what concentration is required in terms of mg mL^{-1}? If the sample volume is 0.4 mL, what mass of protein is required?

STRATEGY

This is a relatively straightforward calculation, provided that you keep track of the units and powers of 10.

SOLUTION

A 1 M solution of protein contains 27 500 mg mL^{-1}. Thus a 0.5 mM solution contains $0.5 \times 10^{-3} \times 27\,500$ mg $mL^{-1} = 13.75$ mg mL^{-1}. In 0.4 mL there would be 13.75×0.4 mg, i.e. 5.5 mg.

It should be noted that for large molecules such as proteins a 1 M solution is physically unrealistic; it would be impossible for instance to dissolve 27 500 mg (27.5 g) protein in buffer to give a 1 mL solution. Nevertheless, the concept of a 1 M solution represents a useful basis for calculations.

In a number of cases, it is convenient to describe the concentration of a solute in a solution in terms of % weight/volume (abbreviated as w/v); this would represent the grams of solute per 100 mL solution. If two liquids are being mixed (e.g. ethanol and water) it can be convenient to express the concentration of ethanol in terms of percentage volume/volume (abbreviated as v/v).

WORKED EXAMPLE

The structure of acrylamide is given in Chapter 1, section 1.8. When acrylamide is polymerized and cross-linked it gives a gel which provides a convenient support medium for electrophoresis in which convection is minimized. The monomer (but not the polymer) is a powerful neurotoxin and must be handled with great care.

Acrylamide, which can be polymerized to make a gel widely used in electrophoresis, is available as an aqueous solution at 30% (w/v). The molecular mass of acrylamide is 71.1 Da. What is the concentration of this solution in molar terms?

STRATEGY

This is again a relatively straightforward calculation provided the correct units are used throughout the calculation.

SOLUTION

A 30% (w/v) solution corresponds to 300 g L^{-1}. A 1 M solution of acrylamide would be 71.1 g L^{-1}. Hence, the concentration is 300/71.1 M = 4.22 M.

A final note of caution: it is *absolutely essential* that you appreciate the distinction between the *amount* of a solute in a solution (measured in *moles*) and its *concentration* (measured in *molar*). You should be able to work out that, for example, 1 mL of a 1 mM solution contains 1 μmol and that 0.02 mL of a 0.7 μM solution contains 14 pmol.

For some calculations it is appropriate to work in terms of mol, for others it may be best to work in terms of M. You must never slip from one to the other during a calculation without being aware of what you are doing. If you label the quantities with the correct units as you go along, you should avoid making mistakes.

SELF TEST ?

Check that you have mastered the key concepts at the start of the section by attempting the following questions.

ST 3.4 How many moles of solute are present in 2.5 mL of a solution whose concentration is 1.35 mM?

ST 3.5 A solution contains 2.1 pmol of a given protein in 0.5 µL. What is the concentration of the solution?

ST 3.6 A mass spectrometer requires 0.2 ng protein in a volume of 2.5 µL. If the molecular mass of the protein is 34 kDa, what is the concentration of the solution in molar terms?

ST 3.7 The molecular mass of a small peptide is 791 Da. What would be the concentration of a 5 mM solution of the peptide in terms of mg mL^{-1}?

Answers

ST 3.4 The amount of solute = 3.375×10^{-6} mol or 3.375 µmol.

ST 3.5 The concentration = 4.2 µM.

ST 3.6 The concentration = 2.35 nM.

ST 3.7 The concentration = 3.955 mg mL^{-1}.

Attempt Problems 3.2 to 3.8 at the end of the chapter.

3.5 Preparation and dilution of stock solutions

KEY CONCEPTS

- Working out how to prepare stock solutions of the required concentration
- Calculating how to dilute these solutions appropriately for particular applications

The accurate preparation and dilution of solutions is essential for accurate work in the molecular biosciences. The following sections should give you guidance as to how to carry out these operations successfully.

3.5.1 Stock solutions

In experimental work, we often make use of 'stock solutions' of compounds which have either been prepared to a high degree of accuracy in the laboratory or purchased from a reputable supplier. These stock solutions are usually much more

concentrated than our 'working solutions' (in actual experiments) and must be diluted appropriately before use.

The preparation of stock solutions is usually carried out as follows. Suppose that you wished to prepare 5 mL of a 5 mM stock solution of NADH. You would need to know the molecular mass of NADH; inspection of the catalogue reveals that the disodium salt hydrate of NADH (which is the form you have purchased) has a molecular mass of 709.4 Da. A 1 M solution of NADH would be 709.4 g L^{-1}, i.e. 709.4 mg mL^{-1}. Thus, a 5 mM solution would be $5 \times 10^{-3} \times 709.4$ mg mL^{-1}, i.e. 3.55 mg mL^{-1}.

Many biological compounds are stored at low temperatures because they are unstable at room temperature. When taking bottles from the fridge or freezer it is important to allow them to warm up to room temperature before opening, otherwise water vapour will condense on the inside of the bottle.

If we wished to make up exactly 5 mL of this solution, we would have to weigh out 5×3.55 mg, i.e. 17.75 mg NADH on an appropriate balance. This could take some considerable time. It is much more likely that we require about 5 mL of the stock solution, in which case we can adjust the volume of solvent (water, buffer, etc.) so that the concentration is correct. Thus, if we weighed out 17.43 mg NADH, we would add 4.91 mL solvent to obtain the correct concentration. (In practice, we would probably use measurements of light absorbance to confirm the concentration of this stock solution, as explained in section 3.6. This would be particularly important in those cases where a substance may have absorbed moisture, which often happens to powders stored at low temperatures.)

The dissolution of many substances including urea and GdmCl is endothermic (i.e. takes in heat), so it is often necessary to warm solutions gently to aid dissolution.

In some cases, particularly of very concentrated stock solutions, there can be significant volume changes when the solvent is added to the solute. For example, if we wished to prepare 10 mL of an 8 M solution of the denaturant guandinium chloride (GdmCl) of molecular mass 95.5 Da, we can readily calculate that we would require 95.5×8 mg mL^{-1}, i.e. 764 mg mL^{-1} GdmCl. For 10 mL, this would be 7.64 g GdmCl. If we were to put this into a graduated tube and add 10 mL of water or buffer, the final volume would be close to 14 mL and the solution would not be of the required concentration. Instead, we would carefully add the solvent, making sure that the GdmCl was dissolved and bringing the volume of the solution up to the 10 mL mark. It should be noted that 8 M is close to the solubility limit of GdmCl, so it takes some time to ensure that it is all dissolved. (Since GdmCl does not have a convenient absorption band, the concentration of the solution can be checked by density or refractive index measurements, by reference to standard tables of values).

3.5.2 Dilution of stock solutions

Many procedures involve the preparation of working solutions from more concentrated stock solutions which have either been purchased or made up in the laboratory. It is very important to be able to do this accurately.

Two basic points are:

A. The amount of solute cannot be increased or decreased by any dilution procedure (matter cannot be created or destroyed!), and

B. The *amount* of solute = the *volume* of solution multiplied by its *concentration*.

Thus if:

C_S = concentration of stock solution

V_S = volume of stock solution taken

and

C_W = final concentration of working solution

V_W = final volume of working solution

Then we can derive the following important equations:

KEY RELATIONSHIP

concentration of stock × volume of stock = final concentration × final volume

i.e. $C_S \times V_S = C_W \times V_W$ 3.2

from which it follows that:

volume of stock = (final concentration × final volume)/concentration of stock

i.e. $V_S = \dfrac{C_W \times V_W}{C_S}$ 3.3

Note the importance of relating the volume of the stock solution to the final (total) volume of the solution (eqn. 3.2), not to the volume of buffer or solvent added to dilute the stock solution.

The application of eqn. 3.3 is illustrated in the worked examples below.

WORKED EXAMPLE

How much of a stock solution of ATP (4.0 mM) in a given buffer should be taken to give a concentration of 0.1 mM in a working solution of total volume 1.0 mL?

STRATEGY

This is an application of eqn. (3.3); it is important to ensure that the units are the same on both sides of the equation.

SOLUTION

In this case, $C_S = 4.0$ mM, $C_W = 0.1$ mM, $V_W = 1.0$ mL

Hence from eqn. (3.3) $V_S = (0.1 \times 1.0)/4.0$ mL = 0.025 mL (25 µL)

The remaining volume (0.975 mL or 975 µL) would be made up of the appropriate buffer.

WORKED EXAMPLE

Stock solutions of ATP and $MgCl_2$ are 4.0 and 20 mM, respectively, in a given buffer. What volumes of ATP, $MgCl_2$ and buffer would you add together to produce a working solution of final volume 2.0 mL, in which the concentrations of ATP and $MgCl_2$ were 0.5 and 1.5 mM, respectively?

STRATEGY

This calculation involves the application of eqn. (3.3) to each solute separately. Again it is important to ensure that the correct units are used on both sides of the equation.

SOLUTION

We must consider the components separately.

For ATP, $C_S = 4.0$ mM, $C_W = 0.5$ mM and $V_W = 2.0$ mL

Hence from eqn. (2.3) $V_S = (0.5 \times 2.0)/4.0$ mL $= 0.25$ mL

For $MgCl_2$, $C_S = 20$ mM, $C_W = 1.5$ mM and $V_W = 2.0$ mL

Hence from eqn. (2.3) $V_S = (1.5 \times 2.0)/20$ mL $= 0.15$ mL

Thus, the working solution would consist of 0.25 mL ATP, 0.15 mL $MgCl_2$ and $(2.0 - (0.25 + 0.15))$ mL, i.e. 1.6 mL buffer.

SELF TEST

Check that you have mastered the key concepts at the start of the section by attempting the following question.

ST 3.8 You are provided with stock solutions of an enzyme (2.5 mg mL^{-1}), its substrate (5 mM), and urea (8 M), all in a given buffer. You wish to make up a solution of volume 0.5 mL with the following final concentrations: enzyme, 0.1 mg mL^{-1}; substrate 0.25 mM; urea, 2 M. What volumes of the stock solutions and buffer would you require?

In ST 3.8 remember to use eqn. 3.3 to calculate the volume of each stock solution required. The volume of buffer required to make up the total volume is calculated by subtracting the stock volumes added from the required volume of the working solution.

Answer

ST 3.8 The volumes of the solutions required are: enzyme, 0.02 mL (20 µL); substrate, 0.025 mL (25 µL); urea 0.125 mL (125 µL); buffer, 0.33 mL (330 µL).

Attempt Problems 3.9 to 3.15 at the end of the chapter.

3.6 Absorbance measurements

KEY CONCEPTS

- Using the Beer–Lambert law to calculate concentrations from absorbance data
- Knowing the range of absorbance values where measurements are most reliable

Measurements of the absorption of light by a solution (spectrophotometry) can be an extremely convenient way of determining the concentration of that solution,

provided that (a) the solute has a suitable absorption band, and (b) you have access to a properly calibrated instrument (spectrophotometer).

The fundamental equation governing spectrophotometry is the Beer–Lambert law (eqn. 3.1):

$$A = \varepsilon \times c \times l \qquad 3.1$$

where A is the absorbance at a particular wavelength, ε is the absorption coefficient (a measure of the degree of absorption of light at that wavelength), and l is the path length of the cuvette (or cell) that holds the solution of concentration c. From this equation we can see that at a given wavelength and at a constant path length, A is proportional to c.

We have already seen this equation (section 3.3), where it was noted that the units of ε will depend on the units of c and l, since overall, each side of the equation has no units. This is because A is defined by eqn. 3.4 as:

$$A = \log\left(\frac{I_0}{I_t}\right) \qquad 3.4$$

where I_0 is the intensity of the incident light, i.e. the light striking the cuvette, and I_t the intensity of the transmitted light, i.e. the light leaving the cuvette.

WORKED EXAMPLE ✓

What percentage of the incident light is transmitted after passage through solutions which have absorbance values of (a) 0.05, (b) 0.3, (c) 1.0 and (d) 2.0?

STRATEGY

This is an application of eqn. 3.4 for each value of A by putting the value of A into the calculator and use the 10^x function to obtain (I_0/I_t). The reciprocal of this value gives (I_t/I_0).

SOLUTION

The value of (I_t/I_0) at each value of A is multiplied by 100 to give the per cent of the incident light that is transmitted. Thus, the per cent of transmitted light is (a) 89.1%, (b) 50.1%, (c) 10% and (d) 1%.

The logarithmic nature of A is not always readily appreciated because most modern instruments merely give a digital display of the absorbance. *However, it is particularly important to note that you should always aim to measure absorbance values in the range 0.1–1.0. Despite what the manufacturers of the instrument may claim, absorbance values significantly above 1.0 are not usually reliable.* You can bring absorbance values into the appropriate range either by an accurate dilution of a concentrated solution, or sometimes by using a cuvette of different path length. Measurements will be most accurately made when the absorbance is about 0.3.

The absorbance at a given wavelength is often denoted by a subscript; thus A_{280} is the absorbance at 280 nm. The absorption coefficient would be similarly abbreviated, e.g. ε_{280}.

If we know (or can look up) the absorption coefficient for a substance at a particular wavelength, we can use the Beer-Lambert law to determine the concentration of a solution, as illustrated below in the worked examples below.

WORKED EXAMPLE

The absorption coefficient of NADH at 340 nm is 6220 M^{-1} cm^{-1}. You have made up a 2.5 mM stock solution of NADH (see section 3.5.1). How would you check its concentration spectrophotometrically?

STRATEGY

The key point is to make sure that the absorbance to be measured is in the range for reliable measurements (0.1–1.0).

SOLUTION

A 2.5 mM solution of NADH in a cuvette of 1 cm path length would, according to the Beer-Lambert law, have an absorbance at 340 nm (denoted as A_{340}) of $6220 \times 2.5 \times 10^{-3} = 15.6$. This is far too high to measure, so we need to perform an accurate dilution. If the solution were diluted 50-fold, e.g. by adding 0.02 mL to 0.98 mL buffer, the A_{340} should be $15.6/50 = 0.312$, which is within the required range for measurement. The spectrophotometer should be adjusted so that the absorbance of buffer alone is set at zero. (Note that this dilution procedure would be appropriate for a cuvette of 1 mL capacity with a path length of 1 cm; if we were using a cuvette of 3 mL capacity, we would add 0.06 to 2.94 mL buffer.)

When performing dilutions it is important to ensure that the correct pipettes are used and the solutions are completely mixed.

WORKED EXAMPLE

A solution consists of a mixture of NADH and its oxidized form NAD^+. NADH and NAD^+ have equal absorbances at 260 nm (absorption coefficient 18 000 M^{-1} cm^{-1}). NADH absorbs at 340 nm (absorption coefficient 6220 M^{-1} cm^{-1}); NAD^+ does not absorb at this wavelength. In a cuvette of 1 cm path length, the absorbance of the solution is 0.933 at 260 nm and 0.135 at 340 nm. What are the concentrations of NADH and NAD^+ in the solution?

STRATEGY

We can use the A_{260} to determine the sum of [NADH] and [NAD^+], since both absorb equally at this wavelength. The A_{340} can be used to determine the [NADH].

SOLUTION

From the A_{260}, the sum of [NADH] and [NAD^+] is $0.933/18\,000$ M = 51.8 μM. From the A_{340}, [NADH] = 0.135/6220 M = 21.7 μM. Hence by difference [NAD^+] = 30.1 μM.

Spectrophotometry provides a convenient way of measuring the concentration of a protein solution. Virtually all proteins absorb radiation at 280 nm due to the aromatic amino acid side chains (principally tyrosine and tryptophan) they contain. The absorption coefficients of different proteins at 280 nm will vary because proteins differ in terms of the exact proportions of these amino acids they contain. For example, the absorption coefficient of BSA (a very commonly employed

protein standard) at 280 nm (ε_{280}) is 0.66 (mg mL^{-1})$^{-1}$ cm^{-1}, whereas that of hen egg white lysozyme is 2.65 (mg mL^{-1})$^{-1}$ cm^{-1}, reflecting the greater proportion of aromatic amino acids in the latter.

The relationship between the absorption coefficient of a protein at 280 nm and the content of aromatic amino acids is discussed in Chapter 6, section 6.1.5.

For proteins, the absorption coefficients are often referred to concentrations in terms of mg mL^{-1} (i.e. w/v) rather than in terms of molarity. If molarity were to be used as the basis for expressing concentrations, not only would the quoted absorption coefficients be inconveniently large (see worked example below), but in addition there also could be ambiguities about the molarity in the case of proteins with multiple polypeptide chains. Does the molarity refer to the individual polypeptide chains or to the intact protein? This should be clearly stated.

WORKED EXAMPLE

The molecular mass of lysozyme is 14.3 kDa. What is the molar absorption coefficient of the protein at 280 nm, given that the A_{280} of a 1 mg mL^{-1} solution in a cuvette of 1 cm path length is 2.65?

STRATEGY

This calculation relies on evaluating the properties of a 1 M solution of the protein, and scaling the absorbance accordingly.

SOLUTION

A 1 M solution of lysozyme would be 14 300 mg mL^{-1}. Using the Beer-Lambert law, the A_{280} of this solution in a 1-cm pathlength cuvette is $2.65 \times 14\,300 = 37\,900$. Thus, the absorption coefficient of lysozyme is 37 900 M^{-1} cm^{-1}.

Many biochemical compounds, including proteins, nucleic acids and small molecules such as ATP and NADH, have convenient absorption bands. If at all possible, it is a good idea to use spectrophotometry to check the concentrations of solutions.

WORKED EXAMPLE

The successful ligation of a DNA fragment (insert) into a cloning vector (DNA) requires a vector/insert ratio of 1/10 in the ligation reaction. Following the careful preparation of a 3.5 kilobase pair (kbp) vector and a 1.0 kbp insert, A_{260} measurements were made to determine the concentration of DNA in each preparation, using the relationship that a value of 1.0 for the A_{260} in a cuvette of 1 cm path length corresponds to 50 μg mL^{-1} DNA. If the A_{260} values of the vector and insert preparations are 0.80 and 0.60, respectively, calculate the volumes of the two preparations required for a ligation reaction which contains 20 fmol vector and 200 fmol insert.

STRATEGY

The DNA concentrations of vector and insert are worked out in terms of μg mL^{-1}, and hence mg mL^{-1}. The molecular masses of the two samples of DNA can be estimated from the number of bases (see Chapter 2, section 2.1), and this can be used to express the concentrations in molar terms. From these values, the volumes of solution which contain the required number of moles can be calculated.

SOLUTION

The DNA concentration of the vector is $(0.80/1.0) \times 50\ \mu g\ mL^{-1} = 40\ \mu g\ mL^{-1}$ (i.e. $0.04\ mg\ mL^{-1}$). Similarly, the DNA concentration of the insert = $30\ \mu g\ mL^{-1}$ (i.e. $0.03\ mg\ mL^{-1}$).

The molecular mass of the vector can be estimated from the relationship (Chapter 2, section 2.1) that each nucleotide in a strand of DNA contributes 330 Da to the molecular mass. Thus, the molecular mass of the vector = $3500 \times 2 \times 330$ Da = 2.31×10^6 Da (note that we have multiplied by 2 because there are 3500 base pairs in this sample of double-stranded DNA). Similarly the molecular mass of the insert = $1000 \times 2 \times 330$ Da = 0.66×10^6 Da.

Hence the concentration of the vector = concentration (in $mg\ mL^{-1}$)/molecular mass (in Da) = $0.04/(2.31 \times 10^6)$ M = 1.73×10^{-8} M = 17.3 nM. The concentration of insert = $0.03/(0.66 \times 10^6)$ M = 45.5 nM.

The solution of vector contains 17.3 nmol L^{-1}; this is equivalent to 17.3 fmol μL^{-1}. Thus, the volume of solution that contains 20 fmol = 20/17.3 μL = 1.16 μL.

The solution of insert contains 45.5 nmol L^{-1}; this is equivalent to 45.5 fmol μL^{-1}. Thus, the volume of solution that contains 200 fmol = 200/45.5 μL = 4.40 μL.

This worked example provides good practice in being able to interconvert the different prefixes for units in the SI system (see section 3.3.1). It also highlights the very small volumes of reagents used in many molecular biology techniques; great care should be taken in handling such small volumes.

SELF TEST

Check that you have mastered the key concepts at the start of the section by attempting the following questions.

ST 3.9 A solution of *N*-acetyltyrosine has an A_{280} of 0.265 in a cuvette of path length 0.5 cm. The absorption coefficient of the compound at this wavelength is $1280\ M^{-1}\ cm^{-1}$. What is the concentration of the solution?

ST 3.10 A solution of bovine insulin, when diluted 20-fold into buffer, had an A_{280} of 0.732 in a cuvette of path length 1 cm. If the absorption coefficient of insulin at 280 nm is $0.97\ (mg\ mL^{-1})^{-1}\ cm^{-1}$ and the molecular mass of the protein is 5734 Da, what is the concentration of the original solution expressed in terms of $mg\ mL^{-1}$ and in terms of molarity?

ST 3.11 4-Nitroaniline (4NA) is a product of the action of proteases on 4-nitroanilide substrates; the absorption coefficient of 4NA at 405 nm is $9500\ M^{-1}\ cm^{-1}$. In a cuvette of 1 cm path length, the A_{405} of a solution of 4NA whose volume is 2.5 mL is 0.27. What is the concentration of the solution and how many moles of 4NA are present?

In ST 3.11, it should be noted that a nitroanilide is a good substrate for a proteolytic enzyme (protease) as it contains an amide bond, which links the amino acid units in a protein (see Chapter 9, section 9.9.3).

Answers

ST 3.9 The concentration = 0.414 mM or 414 μM.

ST 3.10 The concentration = $15.1\ mg\ mL^{-1}$ or 2.63 mM.

ST 3.11 The concentration = 2.84×10^{-5} M or 28.4 μM; there are 7.11×10^{-8} mol or 71.1 nmol present.

Attempt Problems 3.16 and 3.17 at the end of the chapter.

3.7 Acids, bases, and buffers

KEY CONCEPTS

- Using the Henderson-Hasselbalch equation, calculate the balance of an ionization equilibrium of known pK_a at a given pH
- Understanding how buffers operate and how they can be prepared

As mentioned in Chapter 2, section 2.3.1, the pH scale provides a convenient measure of the acidity of a solution; pH = –log [H^+]. At 25°C (298K), pure water has a pH of 7.

The definition of pH means that as [H^+] increases, pH decreases. Distilled water usually contains some dissolved CO_2 (*i.e.* carbonic acid) from the atmosphere, so that the observed pH is lower than 7, typically in the range 5.5–6.0.

An acid can be defined as a substance which has a tendency to donate H^+ ions and a base as a substance which has a tendency to accept H^+ ions. Strong acids such as HCl are completely dissociated (to give H^+ and Cl^- ions), whereas weak acids (such as ethanoic acid) are only partially dissociated. In a similar manner, strong and weak bases will differ in terms of their readiness to accept H^+ ions. Thus, NaOH is a strong base since it is completely dissociated to yield Na^+ and OH^- ions; the latter will readily accept H^+ ions to form water.

The strength of an acid HA (which dissociates to give H^+ and A^- ions) can be expressed quantitatively in terms of its acid dissociation constant K_a, which is defined in eqn. 3.5 as:

$$K_a = ([H^+]\,[A^-])/[HA] \qquad 3.5$$

By analogy with the definition of pH, we define $pK_a = -\log K_a$.

By analogy with the definition of pH, the lower the value of pK_a of an acid, the stronger acid it is.

Taking logarithms of both sides of eqn. 3.5, we obtain:

$$\log K_a = \log [H^+] + \log ([A^-]/[HA])$$

Rearranging the terms:

$$-\log [H^+] = -\log K_a + \log ([A^-]/[HA])$$

which can be expressed as eqn. 3.6 (remembering that pH = –log [H^+] and that $pK_a = -\log K_a$):

$$pH = pK_a + \log ([A^-]/[HA]) \qquad 3.6$$

Eqn. 3.6 is known as the *Henderson–Hasselbalch equation*; it provides a simple way of visualizing the predominant species in the equilibrium as the pH varies with respect to the pK_a; this is depicted in Fig. 3.2.

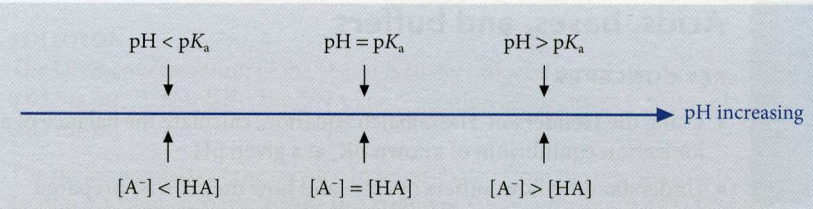

Fig. 3.2 The effect of pH on the position of the HA $\rightleftharpoons$ H$^+$ + A$^-$ equilibrium. When pH = pK_a, exactly 50% of the HA has dissociated to give H$^+$ and A$^-$. The exact ratio of [A$^-$]:[HA] at any pH can be calculated using the Henderson–Hasselbalch equation (eqn. 3.6).

WORKED EXAMPLE

What percentage of the acid is in the form of A$^-$ (a) when the pH is 1 unit below the pK_a, (b) when the pH = pK_a, and (c) when the pH is 1 unit above the pK_a?

STRATEGY

This calculation applies eqn. 3.6; it is important to keep track of the signs and to use the 10^x key on the calculator reliably. To calculate the percentage of each form we must remember that in each case, the sum of [A$^-$] and [HA] is 100%.

SOLUTION

From the Henderson–Hasselbalch equation (eqn. 3.6), it is found that ([A$^-$]/[HA]) is 0.1, 1.0 and 10 for (a), (b), and (c), respectively. This would give the percentages of A$^-$ as 9.1, 50, and 90.9%, respectively.

The pK_a values for some acids of biochemical interest are 4.76 for ethanoic acid (acetic acid), 7.2 for the second dissociation of phosphoric acid ($H_2PO_4^- \rightleftharpoons H^+ + HPO_4^{2-}$), and 8.1 for Tris base (TrisH$^+$ $\rightleftharpoons$ Tris + H$^+$).

The structure of Tris is given in Chapter 1, section 1.8.

Since the structures and hence biological activities of many proteins are very sensitive to pH, we usually try to maintain the pH in the solution by the use of a buffer, which is a system designed to minimize the change in pH on addition of H$^+$ or OH$^-$ ions. A buffer typically consists of a mixture of a weak acid (HA) and its salt (A$^-$); the latter is usually supplied as its sodium salt, for example. (Note that a buffer can also be prepared from a mixture of a weak base and its acid.)

We can explain the action of buffers by considering the equilibrium HA $\rightleftharpoons$ H$^+$ + A$^-$. If H$^+$ ions are added to the system, they will combine with the A$^-$ ions to form HA and thus effectively be removed. When OH$^-$ ions are added, these will react with HA to generate H_2O and A$^-$, again resisting the change in pH. Buffers will be most effective when there are reasonable concentrations of HA and A$^-$ present, and thus should be used within a narrow range (usually taken to be 1 pH unit on either side) of the relevant pK_a. Thus, the phosphate buffer system based on $H_2PO_4^- \rightleftharpoons H^+ + HPO_4^{2-}$ (i.e. using the salts NaH_2PO_4 and Na_2HPO_4) would be suitable for buffering in the range of pH between 6.2 and 8.2. For effective buffering the concentration of the buffer should be kept high.

WORKED EXAMPLE

The Tris buffer system ($TrisH^+ \rightleftharpoons Tris + H^+$, based on the base Tris and its hydrochloride TrisHCl) has a pK_a of 8.1. What are the concentrations of the $TrisH^+$ and Tris forms at pH 8.3, if the total concentration of Tris species is 50 mM? During a reaction, 5 mM H^+ ions are formed. What is the new pH?

STRATEGY

This calculation depends on applying eqn. 3.6 to evaluate the concentrations of Tris and $TrisH^+$ and then consider the effect of adding 5 mM H^+ to the system.

SOLUTION

From the Henderson–Hasselbalch equation (eqn. 3.6), the ratio $[Tris]/[TrisH^+] = 10^{0.2} = 1.58$. Thus, $[Tris] = 61.2$ mM and $[TrisH^+] = 38.8$ mM. On addition of 5 mM H^+, [Tris] will be reduced by 5 mM to 56.2 mM and $[TrisH^+]$ will be raised by 5 mM to 43.8 mM. Application of the Henderson–Hasselbalch equation shows that the new pH would be 8.21. (Note if no buffer had been present, 5 mM H^+ would give a pH of 2.3.)

WORKED EXAMPLE

In the previous worked example, if the total concentration of Tris species is 20 mM, what would be the new pH on adding 5 mM H^+ ions?

STRATEGY

This follows the same steps as the previous example, with the addition of 5 mM H^+ having a proportionately greater effect.

SOLUTION

At pH 8.3, the concentrations of Tris and $TrisH^+$ are 12.2 and 7.8 mM, respectively. On addition of 5 mM H^+, the concentrations become 7.2 and 12.8 mM, respectively; the new pH is 7.85. Comparison with the previous example clearly demonstrates the beneficial effect of using a high concentration of buffer.

WORKED EXAMPLE

A phosphate buffer system can be made up by mixing appropriate quantities of the two salts NaH_2PO_4 (supplied as the dihydrate $NaH_2PO_4.2H_2O$; molecular mass 156 Da) and Na_2HPO_4 (supplied as the anhydrous salt; molecular mass 142 Da). You wish to make up 1 L of a buffer of pH 7.5 with a total phosphate concentration of 0.1 M. Explain how you would make up 0.1 M stock solutions of the two salts in water and then in what proportion you would mix them to obtain the required pH. The appropriate pK_a for the $H_2PO_4^- \rightleftharpoons H^+ + HPO_4^{2-}$ ionization is 7.2.

STRATEGY

This is a straightforward application of eqn. 3.6 to calculate the ratio of the two salts required for the buffer.

SOLUTION

0.1 M stock solutions of $NaH_2PO_4.2H_2O$ and Na_2HPO_4 can be made up by weighing out 15.6 (i.e. 156×0.1) g and 14.2 (i.e. 142×0.1) g of the salts, respectively, dissolving each in water and making up to 1 L in each case. From the Henderson–Hasselbalch equation (eqn. 3.6) the ratio of $[HPO_4^{2-}]/[H_2PO_4^-]$ at pH 7.5 is 1.995 (i.e. $10^{0.3}$). Thus, to obtain 1 L of the required buffer, we need to mix 666 mL ($(1.995/(1 + 1.995) \times 1000)$ mL) of the 0.1 M Na_2HPO_4 solution and 334 mL ($(1/(1 + 1.995) \times 1000)$ mL) of the 0.1 M $NaH_2PO_4.2H_2O$ solution.

In practice, we could also make up the buffer in the last worked example (0.1 M sodium phosphate, pH 7.5) in a rather simpler way. The appropriate amount of the acidic form of the buffer ($NaH_2PO_4.2H_2O$) is weighed out to make 1 L of the buffer, i.e. 0.1×156 g = 15.6 g. This is then dissolved in about 900 mL water, and titrated to the required pH using concentrated (5–10 M) NaOH. Finally, the volume is made up to the required mark (1 L) with water. One advantage of this second method is that if other components are added, e.g. NaCl, they could affect the pK_a of the buffer system, which would alter the ratio of the two solutions to be added in the first method.

We would commonly use a titration method to make up Tris buffers. The required amount of the solid (Tris(hydroxymethyl)methylamine) is weighed out and dissolved in water before being titrated down to the required pH with the appropriate concentrated acid (HCl, ethanoic acid, etc.) before being made up to the required volume with water.

? SELF TEST

Check that you have mastered the key concepts at the start of the section by attempting the following questions.

ST 3.12 The pK_a of the amino group in the side chain of lysine is 10.5. What percentage of lysine is in the unprotonated ($-NH_2$) form at pH 8.2?

ST 3.13 The pK_a of a particular lysine side chain in a protein is 7.8. What percentage of this lysine side chain is in the unprotonated form at pH 8.2?

ST 3.14 The pK_a of ethanoic acid is 4.76. What concentrations of ethanoic acid and sodium ethanoate would be required to make up a buffer at pH 5.2, if the total ethanoic acid and ethanoate concentration is 0.2 M?

In ST 3.14, the Henderson–Hasselbalch equation (eqn. 3.6) is used to calculate the ratio [ethanoate]/[ethanoic acid] at the stated pH. The total concentration of the two species is 0.2 M.

Answers

ST 3.12 0.5% of the lysine is unprotonated.

ST 3.13 71.6% of the lysine side chain is unprotonated.

ST 3.14 The concentrations are: 0.053 M ethanoic acid and 0.147 M sodium ethanoate.

Attempt Problems 3.18 and 3.19 at the end of the chapter.

3.8 Specific activities of proteins and enzymes

KEY CONCEPTS

- Understanding how to calculate the specific activity of a protein
- Expressing the result in the appropriate units

A very important quantity when studying a protein is a measure of its biological activity, usually called an assay of its activity. In the case of a binding protein, we will want to know how much ligand it can bind and with what affinity. If the

protein was able to elicit some more complex response such as cell growth, we would need to measure the ability of a given quantity of the protein to cause cell proliferation in a suitable system. In the case of an enzyme, the activity will be measured in terms of how much substrate is converted to product in a given time.

A very commonly used unit for enzyme activity is the amount that converts 1 μmol substrate to product in 1 min; in fact this is often abbreviated simply to 1 unit (U) or sometimes 1 International Unit (IU).

In the SI system, the unit of enzyme activity (the katal) is defined as the amount which converts 1 mol substrate to product in 1 s. This is, in fact, a very large amount of activity and in practice we would often be dealing with nkatal amounts.

An important property is the specific activity of the protein, which is defined as the amount of biological activity per given weight of the protein. For an enzyme, this might be measured in terms of μmol min^{-1} mg^{-1} (U mg^{-1} or IU mg^{-1}) or using the SI definition in terms of katal kg^{-1}.

WORKED EXAMPLE

Show that for an enzyme a specific activity of 1 katal kg^{-1} corresponds to 60 μmol min^{-1} mg^{-1}.

STRATEGY

The important point here is to keep a very careful track of the units.

SOLUTION

The specific activity of 1 katal kg^{-1} corresponds to 1 mol substrate s^{-1} kg^{-1}, i.e. to 60 mol substrate min^{-1} kg^{-1}. This is 60×10^6 μmol substrate min^{-1} kg^{-1} and hence to 60 μmol substrate min^{-1} mg^{-1}. Thus, despite the fact that both the katal and the kg are large and impractical amounts of activity and protein, the SI definition of specific activity gives numbers that are of similar orders of magnitude to the more widely used definition (U mg^{-1}).

The specific activity of a protein is essentially a measure of the quality and purity of the sample and will be discussed further in section 3.9.

WORKED EXAMPLE

The activity of the enzyme lactate dehydrogenase can be measured by monitoring the reaction in the direction pyruvate + NADH → lactate + NAD^+. NADH absorbs radiation at 340 nm (absorption coefficient = 6220 M^{-1} cm^{-1}); NAD^+ does not absorb at this wavelength. In an assay, 25 μL of a sample of enzyme (containing 5 μg protein mL^{-1}) was added to a mixture of pyruvate and NADH to give a total volume of 3 mL in a cuvette of 1-cm pathlength. The decrease in A_{340} was 0.14 min^{-1}. What is the specific activity of the enzyme expressed in terms of μmol min^{-1} mg^{-1} and in SI units (katal kg^{-1})?

STRATEGY

This calculation depends on evaluating the amount of activity in terms of μmol min^{-1} and the amount of enzyme added. It is advisable to break down the calculation into small steps and set the working out clearly.

Strictly speaking the equation for the reaction catalysed by lactate dehydrogenase should include H^+ on the left-hand side. It is assumed that the reaction is carried out in a suitable buffer, so that H^+ is often omitted.

SOLUTION

The amount of enzyme added is 0.025×5 μg (note that 25 μL = 0.025 mL), i.e. 0.125 μg.

The rate of the reaction (change in A_{340} min^{-1}) can be converted to a concentration change min^{-1}, i.e. 0.14/6220 M min^{-1} = 22.51 μM min^{-1}.

In a 3 mL assay mixture this corresponds to change of $22.51 \times 3/1000$ μmol min^{-1} = 0.0675 μmol min^{-1}.

The specific activity is thus 0.0675 μmol min^{-1} per 0.125 μg, or $0.0675 \times 1000/0.125$ μmol min^{-1} mg^{-1} = 540 μmol min^{-1} mg^{-1}.

In SI units this would be 540/60 katal kg^{-1}, i.e. 9.0 katal kg^{-1}.

We shall see in Chapter 4, section 4.5 how the specific activity can be used to calculate the turnover number (k_{cat}) of an enzyme, provided the molecular mass of the enzyme is known.

? SELF TEST

Check that you have mastered the key concepts at the start of the section by attempting the following questions.

ST 3.15 The rate of an enzyme-catalysed reaction as measured by the formation of product is 7.3 μM min^{-1} in a solution of volume 2 mL. What does this rate correspond to in terms of μmol min^{-1}? If 0.15 μg enzyme had been added to the reaction, what is the the specific activity of the enzyme expressed as μmol min^{-1} mg^{-1} and as katal kg^{-1}?

ST 3.16 The specific activity of a certain enzyme is 745 μmol min^{-1} mg^{-1}. How much enzyme should be added to a reaction mixture of volume 1 mL to give a rate of reaction of 12 μM min^{-1}?

In ST 3.16 the rate should be calculated in terms of mol/min from the concentration change/min and the volume of solution. Knowing the specific activity of the enzyme allows you to work out how much enzyme is needed to give the required rate.

Answers

ST 3.15 The rate corresponds to 0.0146 μmol min^{-1}; the specific activity is 97.3 μmol min^{-1} mg^{-1} (or 1.62 katal kg^{-1}).

ST 3.16 The amount of enzyme to be added is 0.0161 μg.

Attempt Problem 3.20 at the end of the chapter.

3.9 Purification tables

KEY CONCEPTS

- Defining the terms specific activity, yield and purification factor
- Calculating these quantities to draw up a purification table

A very important part of the study of proteins involves their purification. Only when purified protein is available can we determine its structure and investigate the basis of its biological activity. Traditionally, purification would have involved

initially identifying a suitable source in which the protein of interest was abundant, for example yeast would be an excellent source of enzymes of the glycolytic pathway.

After breakage of the cells to provide a crude extract, various purification procedures would be applied in turn until pure, active protein was obtained. Recently, the trend has been to overexpress a recombinant version of the protein (often with some tag incorporated to facilitate purification) in a convenient host organism, such as *Escherichia coli*, yeast or insect cells, which can often be grown on a large scale using relatively simple growth media (see Chapter 5, section 5.4.2).

Before undertaking a purification process, it is essential to devise a reliable assay system for the protein in question, this will usually involve a measurement of the biological activity (catalysis, binding, etc.); see Chapter 5, section 5.3.

Whatever method of purification is used, it is important to record the progress of the procedure in a suitable way; this involves drawing up a purification table, which will record the following information:

- Step involved
- Total protein
- Total activity
- Specific activity (i.e. the activity per unit weight of protein)
- Yield (i.e. the recovery of activity expressed as a per cent of that in the initial extract)
- Purification factor (the factor by which the specific activity has increased).

It is essential for you to be able to calculate the specific activity, yield, and purification factor given data on the total protein and total activity.

To illustrate the concepts involved, consider the following data from a purification of an enzyme.

Step 1 refers to the crude extract. The total protein is 940 mg, and the total activity is 56 780 units (1 unit equals 1 μmol substrate consumed per min).

Step 2 refers to the extract which has been subjected to ion-exchange chromatography on DEAE-cellulose. The total protein in the pooled peak fractions is 53 mg and the total activity is 47 640 units.

We can then draw up our initial purification table (Table 3.2).

Of course, in a real example it is likely that there will be several more steps before the protein is completely purified.

Table 3.2 Purification table for a protein

Step	Total protein (mg)	Total activity (units)	Specific activity (units mg^{-1})	Yield (%)	Purification factor
Crude extract	940	56 780	60.4	100	1.0
Ion-exchange	53	47 640	899	83.9	14.9

The *specific activity* is obtained by dividing the total activity by the total protein at that step (e.g. 56 780/940 = 60.4 units mg^{-1}).

The *yield* refers to the recovery of activity, *not* to the recovery of protein. After all, the purpose of the purification is to get rid of contaminating proteins and retain the protein of interest. In this case, the yield at step 2 is (47 640/56 780) × 100% = 83.9%.

The *purification factor* refers to the factor by which the specific activity has increased, in this case 899/60.4 after step 2 = 14.9-fold. (Note in a multi-step purification, we can express this either in terms of the factor for each individual step or as a running total factor relative to the initial crude extract as 1.0).

Some examples of protein purification experiments are described in Chapter 7.

SELF TEST

Check that you have mastered the key concepts at the start of the section by attempting the following question.

ST 3.17 In a purification procedure for a certain enzyme, the total protein and total activity at each step was as follows:

Step	Total protein (mg)	Total activity (units)
Crude extract	53.5	780
Ion-exchange chromatography	8.7	525
Gel filtration	1.6	490

Calculate the specific activity, yield and purification factor at each step in the procedure.

Answer

ST 3.17 The values are: specific activity, 14.6, 60.3, 306.3 units mg^{-1}; yield, 100, 67.3, 62.8%; purification factors (per step), 1.0, 4.1, 5.1).

Attempt Problem 3.21 at the end of the chapter.

3.10 Problems

Full solutions to odd-numbered problems are available to all in the student section of the Online Resource Centre at www.oxfordtextbooks.co.uk/orc/price/. Full solutions to even-numbered problems are available to lecturers only in the lecturer section of the Online Resource Centre.

3.1 Hair is composed of the protein α-keratin, which is α-helical in structure. How many turns of α-helix are produced per second in growing hair? Assume that the rate of hair growth is 18 cm $year^{-1}$ and that each turn of α-helix corresponds to a length of 0.54 nm.

3.2 How many moles of GTP are present in 2 µl of a 15 µM solution?

3.3 GTP (sodium salt) has a molecular mass of 545 Da. What mass of GTP is present in 3 mL of a 15 µM solution?

3.4 A certain analytical technique is said to be capable of determining 100 pmol of the amino acid tyrosine (molecular mass 181 Da). What mass of tyrosine does this correspond to?

3.5 The concentration of haemoglobin in red blood cells is estimated to be 250 mg mL^{-1}. Haemoglobin consists of four polypeptide chains each of molecular mass 16.5 kDa (16 500 Da); what is the molar concentration of haemoglobin chains in the red blood cell?

3.6 The molecular mass of chymotrypsin is 24 500 Da (24.5 kDa). What is the concentration of a 0.35 mg mL^{-1} solution expressed in molar terms?

3.7 A solution of aspartate aminotransferase is stated to be 13.6 µM in terms of the dimeric (two polypeptide chain) form of the enzyme (the molecular mass of each polypeptide chain is 44 kDa). What is the concentration of the enzyme in terms of mg mL^{-1}?

3.8 The enzyme phosphoglycerate mutase from *Saccharomyces cerevisiae* (polypeptide molecular mass 27 kDa) is typically assayed at a concentration of 0.05 µg mL^{-1}. What is the molar concentration of enzyme polypeptide chains in the assay?

3.9 Explain how you would make up 15 mL of a 10 M solution of urea in water. The molecular mass of urea is 60 Da.

3.10 For the preparation of a certain gel for electrophoresis, 20 mL of an 8% (w/v) acrylamide solution is required. What volume of a 30% (w/v) stock solution of acrylamide is required for this?

3.11 A stock solution of a protein in buffer is 4.3 mg mL^{-1}. How would you prepare 1 mL of a solution containing 0.2 mg mL^{-1} protein?

3.12 The concentration of glacial acetic (ethanoic) acid is stated to be 17.3 M. How would you prepare 200 mL of a 1 M solution of the acid in water?

3.13 IPTG (isopropyl-β-D-thiogalactopyranoside) is widely used as an inducer for the expression of recombinant proteins in the bacterium *E. coli*. Its molecular mass is 238.3. How much IPTG is required to make up 5 mL of a 0.5 M stock solution in water? What volume of this solution should be added to 50 mL of bacterial culture to give a final concentration of 1.2 mM?

3.14 The four deoxyribonucleotide triphosphates (NTPs), i.e. dATP, dCTP, dGTP, and dTTP, are essential components of the polymerase chain reaction (PCR) used to amplify a selected section of DNA. The NTPs are often supplied as separate stock solutions, each at a concentration of 100 mM in sterile H_2O. How would you make up a working solution (in sterile H_2O) of final volume 1 mL which contains the four NTPs each at a concentration of 2 mM?

3.15 The concentration of a stock solution of NADH is 3 mg mL^{-1}. What does this correspond to in molar terms? (The molecular mass of NADH (disodium salt) is 709 Da). What volume of this stock solution would you add to an assay mixture (of final volume 1.0 mL) to obtain a final concentration of 0.1 mM?

3.16 The absorbance of a solution of ATP was checked by performing two successive dilutions, the first 50-fold and the second 40-fold. The A_{260} of this final solution was 0.47 in a cuvette of 1 cm path length. What is the concentration of the original solution of ATP? The absorption coefficient of ATP at 260 nm is 15 000 M^{-1} cm^{-1}.

3.17 A solution of BSA is made up by weighing out 20 mg of the lyophilized (freeze dried) powder of the protein supplied and dissolving in 2 mL water. The concentration of the protein was checked by adding 0.04 mL of the solution to 0.96 mL water; the A_{280} of the resulting solution relative to a water blank was 0.237. The absorption coefficient for BSA is 0.66 (mg $mL^{-1})^{-1}$ cm^{-1}. Calculate the actual concentration of the original solution of BSA and comment on your answer.

3.18 What would be the change in pH on adding OH^- ions to a final concentration of 20 mM to a 0.1 M sodium phosphate buffer, pH 7.0? The pK_a for the $H_2PO_4^- \rightleftharpoons H^+ + HPO_4^{2-}$ ionization is 7.2.

3.19 Explain how you would make up 1 L of a 50 mM Tris/HCl buffer at pH 7.5, containing 20 mM NaCl. The molecular masses of Tris(hydroxymethyl)methylamine and NaCl are 121.1 and 58.45 Da, respectively.

3.20 A sample of malate dehydrogenase (3.2 mg mL^{-1}) was diluted 500-fold for assay. 20 µL of this solution was added to an assay system of total volume 1 mL in a cuvette of 1 cm path length. The increase in A_{340} (corresponding to the conversion malate + NAD^+ → oxaloacetate + NADH + H^+) was 0.23 min^{-1}. What is the specific activity of the enzyme expressed in terms of µmol min^{-1} mg^{-1} and in SI units? (The absorption coefficient of NADH is 6220 M^{-1} cm^{-1}; NAD^+ does not absorb at this wavelength.)

3.21 The enzyme adenylate kinase catalyses the interconversion of adenine nucleotides in the reaction AMP + ATP ⇌ 2ADP. The enzyme was purified from pig muscle by procedures including precipitation at low pH, elution from phosphocellulose by a pulse of AMP and gel filtration on Sephadex G-75. The following results were obtained starting from 1 kg muscle:

Step	Total protein (mg)	Total activity (units)*
Crude extract	72 500	413 000
Low pH	18 700	365 000
Phosphocellulose	290	223 000
Gel filtration	77	200 000

*1 unit of enzyme activity corresponds to 1 µmol substrate consumed per minute.

Complete the purification table for this procedure.

Appendix

Appendix 3.1 The SI units and physical constants

SI units are based on the metre-kilogram-second system of measurement, and are very widely accepted among scientists. SI units for various physical quantities are listed below.

Quantity	SI unit	Notes
Amount of substance	mole (mol)	This quantity contains 1 Avogadro number of the basic units (atoms, molecules etc.)
Electric charge	coulomb (C)	
Length	metre (m)	Å (Ångstrom) = 1×10^{-10} m = 0.1 nm
Mass	kilogram (kg)	Multiple units are based on g (e.g. mg) although kg is the basic unit
Molecular mass	dalton (Da)	1 Da = 1 atomic mass unit ≈ mass of 1 hydrogen atom (1.66×10^{-24} g); more exactly it is 1/12 of the mass of the ^{12}C isotope of carbon
Temperature	degree Kelvin (K)	0°C = 273.15K
Time	second (s)	
Volume	cubic metre (m^3)	1 litre (L) = 1 dm^3; 1 mL = 1 cm^3
Electric potential	volt (V) (J C^{-1})	
Energy	joule (J) (1 m^2 kg s^{-2})	1 cal = 4.18 J
Force	newton (N) (1 m kg s^{-2})	
Frequency	hertz (Hz) (1 s^{-1})	
Power	watt (W) (J s^{-1})	
Pressure	pascal (Pa) (1 N m^{-2})	1atm = 101.325 kPa
Radioactive decay	becquerel (Bq) 1 s^{-1}	1 disintegration per second; 1 curie (Ci) = 3.7×10^{10} s^{-1}
Concentration	mol kg^{-1}	Moles per kg of solvent, also called molality Concentration is more usually quoted in terms of molarity (M), moles per L of solution For dilute aqueous solutions, molarity is very nearly equal to molality.

The following prefixes are used in conjunction with the SI units to bring the numerical values of quantities into a convenient range (see Chapter 3, section 3.3.1).

Prefix	G, giga	M, mega	k, kilo	Base unit	m, milli	μ, micro	n, nano	p, pico	f, femto
Numerical value	10^9	10^6	10^3	1	10^{-3}	10^{-6}	10^{-9}	10^{-12}	10^{-15}

Physical constants used in this book are:

Avogadro number (L)	6.022×10^{23} mol^{-1}
Faraday constant (F)	9.648×10^4 C mol^{-1}
Gas constant (R)	8.314 J K^{-1} mol^{-1}
Planck constant (h)	6.626×10^{-34} J s
Speed of light in a vacuum (c)	2.997×10^8 m s^{-1}