# Physical Chemistry

6

Wouter Hinrichs and Renske van Gestel

## Contents




**Abstract**

Many pharmaceutical dosage forms are liquids or semi-liquids such as solutions, colloidal systems, suspensions and emulsions. This chapter deals with the physico-chemical backgrounds that are important for the preparation of these types of dosage forms. Successively solubility, rheology, phase behavior, interfaces, surface active agents, disperse systems and osmosis are addressed. Physical chemistry plays an important role in the design of liquid or semi-liquid pharmaceutical dosage forms. By changing physico-chemical parameters intentionally or unintentionally, the biopharmaceutical properties and thus the therapeutic activity of an active substance can drastically change. Many physico-chemical properties are related to each other. For example, changing the composition of a solvent mixture does not only affect its solubility for an active substance but also its surface tension, osmotic value, etc. The chapter is primarily intended to explain physico-chemical aspects described in other chapters. Many examples of the design of pharmaceutical preparations are described to clarify and illustrate the concepts, considering excipients as well as active substances, and small molecules as well as proteins. Also attention is given to the stabilisation of proteins by freeze drying them together with sugars.


### What Is New?

This chapter was based on the chapter Physical Chemistry, by Wouter Hinrichs, Suzy Dreijer-van der Glas, in the 2015 edition of Practical Pharmaceutics. New in this chapter is that all pharmaceutical preparations have been updated. Furthermore, a section on phase behaviour has been added and last, a section on the stabilisation of proteins by freeze drying them in the presence of sugars has been added.

### Learning Objectives

- To obtain a better understanding of the physico-chemical backgrounds of pharmaceutical preparations
- To be able to design pharmaceutical preparations based on physico-chemical principles

Based upon the chapter Physical Chemistry, by Wouter Hinrichs, Suzy Dreijer-van der Glas, in the 2015 edition of Practical Pharmaceutics.


W. Hinrichs (✉)
Department of Pharmaceutical Technology and Biopharmacy, University of Groningen, Groningen, the Netherlands
e-mail: w.l.j.hinrichs@rug.nl

R. van Gestel
Department of Pharmaceutics, Utrecht University, Utrecht, the Netherlands
e-mail: r.a.vangestel@uu.nl


© Springer Nature Switzerland AG 2023

P. Le Brun et al. (eds.), *Practical Pharmaceutics*, https://doi.org/10.1007/978-3-031-20298-8_6

## 6.1 Solubility

Aqueous solubility is an important property of active substances. For example, incompatibility reactions leading to undesired precipitation in parenteral formulations can be avoided with a good knowledge of solubility properties [1–3]. Furthermore, there can be several reasons to influence the aqueous solubility of an active substance. It may be required to increase the solubility to be able to prepare dosage forms in which the active substance is dissolved or it may be more useful to disperse an active substance as particles in a suspension, thus in the undissolved form.

Bringing active substances into solution can have important advantages:

- The preparation method is much simpler than that of disperse systems.
- The dosage accuracy is much better than disperse systems.
- The rate of absorption and the bioavailability will be higher than those of disperse systems.
- Solutions in certain dosage forms are much better tolerated than suspensions. Suspensions (except for nanosuspensions) are not suitable for parenteral infusion for instance.

In a number of cases however, suspensions are preferable:

- If an active substance is unpalatable, the undissolved form can mask the taste.
- If the active substance is unstable in solution.
- If a reduction of the dissolution rate of an active substance is desired in order to slow down the absorption.

The pH of Tetracycline Mouthwash 5% FNA (Table 6.1) is adjusted to 5.0–5.5 with sodium citrate to minimise the solubility of tetracycline hydrochloride. Tetracycline hydrochloride solutions are unstable.

**Table 6.1** Tetracycline mouthwash 5% [4]

| | |
|---|---|
| Tetracycline hydrochloride | 5.0 g |
| Methyl parahydroxybenzoate | 0.1 g |
| Sodium citrate | 6.5 g |
| Sorbitol, liquid (crystallising) | 65.5 g |
| Tragacanth | 0.5 g |
| Water, purified | 40.0 g |
| **Total** | 118.2 g (= 100 mL) |

In the European Pharmacopoeia (Ph. Eur.) various definitions for solubilities in water are described, i.e. from "very soluble" to "practically insoluble" (Table 6.2) [1]. The solubility is also indicated by the amount of water, in millilitres, necessary to dissolve 1 g of the substance at room temperature. The solubility of a substance can also be specified in grams per litre.

Although the terms in Table 6.2 to describe the solubility of an active substance are still used in many handbooks, from a clinical view they are no longer used. The principle of solubility in perspective is now used. That is, the solubility of an active substance is connected to the dose that should be administered to the patient. Thus, solubility in perspective is not a pure physico-chemical characteristic of the active substance. For example, the solubility will not cause any problems when an active substance with a low solubility has to be given to a patient as a solution at a very low concentration. In Chap. 4 Product design and in Chap. 5 Biopharmaceutics and Pharmacokinetics this concept will be explained in more detail.

**Table 6.2** Explanation of the definitions for the solubility in water as specified in the Ph. Eur

| | Solubility[a] |
|---|---|
| Very soluble | Less than 1 |
| Freely soluble | From 1 to 10 |
| Soluble | From 10 to 30 |
| Sparingly soluble | From 30 to 100 |
| Slightly soluble | From 100 to 1000 |
| Very slightly soluble | From 1000 to 10,000 |
| Practically insoluble | More than 10,000 |

[a]Amount of water in millilitres required to dissolve 1 g of an active substance

The rate at which an active substance dissolves does not only depend on the solubility of the active substance. The dissolution of an active substance is described by the Noyes-Whitney equation:

$$\frac{dm}{dt} = \frac{D \cdot A}{h}\left(C_s - C\right) \tag{6.1}$$

where dm/dt is the dissolution rate, D the diffusion coefficient of the active substance in solution, A the contact surface area of the active substance with the solvent, h the thickness of the diffusion layer, $C_s$ the solubility of the active substance and C the concentration of the active substance in the bulk of the solution.

Amongst other strategies, the solubility of a substance can be influenced by variation of the pH, salt formation, variation of the solvent, complex formation, or derivatisation. These concepts, with pharmaceutical preparations as examples, are explained below.

### 6.1.1 Solubility and pH

Most substances are more easily dissolved in water when they are ionised. Many active substances are weak acids or weak bases. The degree of ionisation for many substances depends on the pH. The degree of ionisation, and thus the solubility, can be influenced by adjusting the pH of the medium. The solubility of a non-ionised weak acid can be increased by raising the pH, while the solubility of a non-ionised weak base can be increased by lowering the pH.

The $pK_a$ value of an acid indicates how weak the acid is: the higher the $pK_a$ value, the weaker the acid. Thus when a weak acid is dissolved in pure water, the decrease of the pH and the extent of dissociation of the acid will be less when the $pK_a$ value increases. The relationship of the pH, $pK_a$, and the concentrations of a non-ionised acid [HA] and its salt [A-] is given by the Henderson-Hasselbalch equation:

$$pH = pK_a + \log \frac{[A^-]}{[HA]} \tag{6.2}$$

When the $pK_a$ of a substance is known, the fraction ionised active substance as a function of the pH can be calculated using the following derivatives of this equation:

$$\text{fraction ionised} \; 1 - \left(10^{pH - pK_a} + 1\right)^{-1} \quad (\text{for a weak acid}) \tag{6.3}$$

$$\text{fraction ionised} \; 1 - \left(10^{pK_a - pH} + 1\right)^{-1} \quad (\text{for a weak base}) \tag{6.4}$$

The $pK_a$ of phenobarbital is 7.4. When the pH is adjusted to 5.4, 7.4 or 9.4, the fraction ionised active substance will be: $1 - (10^{5.4-7.4} + 1)^{-1} = 0.01$; $1 - (10^{7.4-7.4} + 1)^{-1} = 0.5$ or $1 - (10^{9.4-7.4} + 1)^{-1} = 0.99$, respectively. This example emphasises that by varying the pH around the $pK_a$, the fraction of ionised active substance can be strongly affected: up to 2 pH units below the $pK_a$, only 1% is ionised, whereas at 2 pH units above the $pK_a$ 99% is ionised.

For a further understanding of the dissolution of a substance by salt formation, some basic concepts of analytical chemistry are summarised (Fig. 6.1).

When an aqueous solution of a weak monovalent acid in water is titrated with NaOH, the pH strongly increases with the first amounts of added NaOH, but when the pH almost reaches the $pK_a$ value of the acid, the pH changes to a lesser extent. When the pH is one to two units above the $pK_a$, the pH again rises sharply upon adding more NaOH. For a titration, the amount of added reagent (in this case NaOH) divided by the quantity of the substance to be determined (both in moles) has been defined as λ. The pH of the initially acidic solution (λ = 0) is dependent on the concentration and the degree of dissociation of the acid. When the pH is around the $pK_a$, the solution acts as a buffer because the addition of a small amount of acid or base will hardly affect the pH. When λ = 0.5, the pH equal to the $pK_a$, there is an equal amount of acid and salt present. In this situation, the buffer capacity is maximal. The equivalence point of the titration is found when λ = 1.

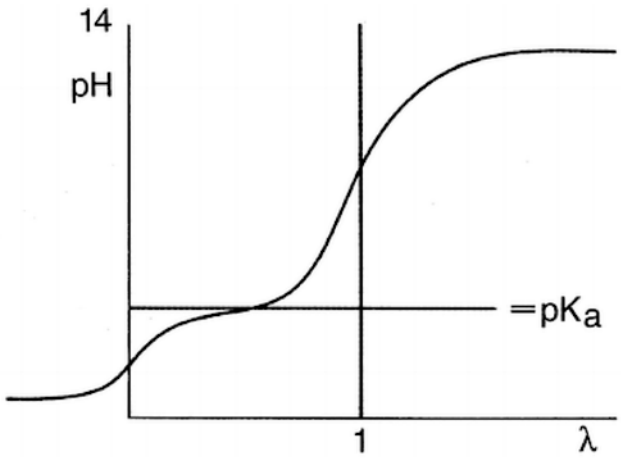


**Fig. 6.1** Titration of a weak monovalent acid with a strong base

A salt composed of a weak base and a strong acid lowers the pH in aqueous solution. Conversely, a salt composed of a weak acid and a strong base raises the pH in aqueous solution. Examples are lidocaine hydrochloride (acid reaction) and sodium phenobarbital (alkaline reaction), respectively. When the pH of aqueous solutions of these substances is adjusted to pH = 7, lidocaine and phenobarbital, respectively, are partially formed again and precipitate.

The so-called pHp Eqs. (6.5 and 6.6) can be used to calculate the solubility of a weak acid or a weak base as a function of the pH [2, 5]. pHp is an abbreviation for pH precipitation, the pH at which just no precipitation occurs. At a given concentration, the weak acid will precipitate when the pH is lower than the pHp and the weak base will precipitate when the pH is higher than the pHp.

$$pHp = pK_a + \log\left(\frac{S - S_0}{S_0}\right) \quad (\text{for a weak acid}) \tag{6.5}$$

$$pHp = pK_a + \log\left(\frac{S_0}{S - S_0}\right) \quad (\text{for a weak base}) \tag{6.6}$$

where $S_0$ is the solubility of the undissociated acid or base and S the concentration of the acid or base which has been added initially (both in mol/L).

The pHp equation can be illustrated in more detail with an example of a calculation. Suppose a 1% w/v solution of sodium phenobarbital should be prepared. The following information can be obtained from the Merck Index [6]:

MW phenobarbital = 232.32 g/mol;
MW sodium phenobarbital = 254.22 g/mol;
Solubility phenobarbital in water = 1 g/L;
$pK_a$ (phenobarbital) = 7.4.

It follows:

$$S_0 = 1/232.32 = 0.0043 \text{ mol/L};$$

1% w/v corresponds to 10 g/L; thus:

$$S = 10/254.22 = 0.0393 \text{ mol/L};$$

Thus the pHp becomes:

$$\text{pHp} = 7.4 + \log\left(\frac{0.0393 - 0.0043}{0.0043}\right) = 8.3.$$

It follows that in order to obtain a 1% w/v sodium phenobarbital solution, the pH must be at least 8.3. At a pH lower than 8.3, phenobarbital will remain partially undissolved.

However, the choice of pH cannot be unlimited. In particular at a high or a low pH, the active substance can for example be chemically unstable. In addition, an extreme pH can be incompatible with other components in the solution, for example, preservatives. In addition, the selected pH should also be compatible with the route of administration.

Thus with the correct setting of the pH, either a solution or a suspension can be obtained. The preparation of solutions and suspensions by manipulation of the pH will be illustrated with a number of examples.

There are several ways to adjust the pH to a certain value, which are essentially not different. Either NaOH or HCl can be added to obtain a suitable pH at which the salt is soluble, or the pHp equation can be used to calculate the amount of an acid or a base with its corresponding salt that will be soluble.

Examples of pharmaceutical preparations in which the active substance is dissolved by influencing the pH are infusion or inhalation solutions with acetylcysteine and a furosemide oral solution (Table 6.3). In both cases, a weak acid is brought into solution by raising the pH. Usually sodium hydroxide is used for acetylcysteine and in the formulation of Table 6.3 the primary amine trometamol is used for furosemide.

Tetracycline hydrochloride easily dissolves in water, to give an acidic solution. In Tetracycline hydrochloride cream 3% FNA, the pH is adjusted to about 5.5 by adding sodium citrate to a tetracycline hydrochloride solution (Table 6.4). At this pH, tetracycline is insoluble and therefore a suspension is formed. In an undissolved state tetracycline is chemically more stable than in the dissolved state. Moreover, a pH of 5.5 is only slightly lower than the third pKa value of citrate ($pK_a$ = 6.4 at 25 °C), so a citrate buffer is formed. This saves tetracycline from dissolving again due to slight pH decrease.

Tetracycline is an amphoteric substance which forms salts in both an acidic and an alkaline environment. The substance dissolves at a pH higher than about 8. This was made use of in the preparation of a solution of tetracycline for eye drops (Table 6.4). By the addition of borax, a pH of about 8.2 is reached. As a consequence, tetracycline readily dissolves (Table 6.5).

Is a bumetanide infusion solution 5 mg in 50 mL (syringe for a pump) possible as a solution? The pH of the licensed product (injection solution 0.5 mg/mL) is 7.0. According to information from the manufacturer, the concentration should be maximally 0.1 mg/mL when mixed with infusion fluids. The reason is that when the concentration is higher than 0.1 mg/mL bumetanide will precipitate unless the pH of the preparation is maintained. The solubility of (non-ionised) bumetanide is 0.1 mg/mL or even lower. The licensed product contains alkaline additives in order to reach a concentration of 0.5 mg/mL. If these additives are diluted, the pH becomes too low, which may cause precipitation of non-ionised bumetanide.

Theophylline can be dissolved by salt formation with ethylenediamine or other amines. Previously, ethylenediamine theofyllinate (aminophylline) was used in oral preparations. Soluble double salts of theophylline can be prepared using sodium acetate and sodium glycinate. However, excellent

**Table 6.3** Furosemide oral solution 2 mg/mL [7]

| | |
|---|---|
| Furosemide | 0.2 g |
| Methyl parahydroxybenzoate | 0.15 g |
| Propylene glycol | 0.91 g |
| Saccharin sodium | 0.1 g |
| Trometamol | 0.1 g |
| Water, purified | Ad 100 mL |

**Table 6.4** Tetracycline hydrochloride cream 3% [8]

| | |
|---|---|
| Tetracycline hydrochloride | 3 g |
| Sodium citrate | 4 g |
| Water, purified | 11 g |
| Cetomacrogol cream FNA[a] | 82 g |
| **Total** | 100 g |

[a]Cetomacrogol emulsifying wax (BP) 15 g, Sorbic acid 200 mg, Decyl oleate 20 g, Sorbitol, liquid (crystallising) 4 g, Water, purified 60,8 g. Total 100 g

**Table 6.5** Tetracycline hydrochloride eye drops solution 0.5% [9]

| | |
|---|---|
| Tetracycline hydrochloride | 0.5 g |
| Borax | 0.5 g |
| Sodium chloride | 0.7 g |
| Water, purified | Ad 100 mL |

absorption is achieved after oral administration of theophylline as such, for example in capsules, so administration as a solution does not seem to be necessary from a biopharmaceutical viewpoint. After absorption, at a physiological pH of approximately 7.4 there is obviously no difference in the degree of ionisation of the substance whether it was administered as such or as a salt. Therefore, no differences in physiological activity between an active substance and its salt are expected. However differences in solubility and dissolution rate and differences in absorption and absorption rate may generally influence the earlier stages of administration and lead to a difference in bioavailability. In addition, the counter ion may induce (undesirable) side-effects. The use of ethylenediamine was abandoned because of the sensitising properties, even after oral use [10]. In addition, it might be toxic, being a secondary amine.

### 6.1.2 Solubility and Salt Formation

Not all salts exhibit a good aqueous solubility. A number of inorganic salts having a relatively low molecular weight and a low water solubility are listed in Table 6.6. Also the solubility product often is given (Table 6.7). This allows us to determine the effect of other ions in the solution on the solubility of a given substance, for example, when sodium carbonate is added to a solution of magnesium chloride. Also the reduction of the solubility of the poorly soluble calcium carbonate can be calculated when a certain amount of the readily soluble sodium carbonate or calcium chloride is added. This reduction in the solubility of, in this case, calcium carbonate is also known as the common ion effect.

**Table 6.6** Solubility of some mineral salts in water in grams per litre

| | 22 °C | 38 °C |
|---|---|---|
| $CaSO_4.2H_20$ | 2.41 | 2.22 |
| $CaCO_3$ | 0.014 | 0.018 |
| $Ca_3(PO_4)_2$ | 0.02 | Not available |
| $CaHPO_4.2H_2O$ | 0.316 | Not available |
| $Ca(H_2PO_4)2.H_2O$ | 18 | Not available |
| $MgSO_4.7H_2O$ | 710 | 910 |
| $MgCO_3$ | 0.106 | Not available |
| $MgHPO_4.7H_2O$ | 3 | 2[a] |

[a]Magnesium and phosphate ions can co-exist in mineral infusions. As can be seen in the table, the solubility of magnesium hydrogen phosphate decreases when the temperature is increased. As a result, during sterilisation a precipitate can be formed in a solution that was originally clear. After cooling, however, this precipitate will dissolve again

**Table 6.7** Solubility product of some mineral salts in water

| | Solubility product (S)[a] | pS (= − logS) |
|---|---|---|
| $CaSO_4$ | $7.1 \times 10^{-5}$ | 4.15 |
| $CaCO_3$ | $5.0 \times 10^{-9}$ | 8.30 |
| $Ca_3(PO_4)_2$ | $2.1 \times 10^{-33}$ | 32.7 |
| $MgCO_3$ | $6.8 \times 10^{-6}$ | 5.17 |
| $Mg_3(PO_4)_2$ | $6.5 \times 10^{-5}$ | 4.18 |
| $Mg_3(PO_4)_2$ | $3.5 \times 10^{-5}$ [b] | 4.46 |

[a]Room temperature
[b]38 °C

An example of this can be found in solutions for parenteral nutrition that contain calcium and phosphate ions (see Chap. 21). At the pH of parenteral nutrition mixtures dihydrogen phosphate and monohydrogen phosphate will both be present. The solubility of calcium dihydrogen phosphate is 18 g/L and that of calcium monohydrogen phosphate 0.3 g/L.

As long as the pH of the mixture remains below 6.4, precipitation is not likely to occur. But for instance at a pH of 7.4, 60% of the phosphate will be present in the form of monohydrogen phosphate, with an increasing risk of precipitation. Therefore, it is always best to check the solubility of the product.

Solubilities of many active substances are listed in Merck Index [6] and in Martindale [11]. When a pharmaceutical preparation contains various salts there may be (multiple) combinations of ions that are incompatible and may result in the formation of a precipitate. For example, after addition of a chloride salt to a chlorhexidine digluconate solution, the insoluble salt chlorhexidine chloride is formed resulting in precipitation. Exact data on this are hard to find. In general, the risks of precipitation are high for combinations of large positive and negative ions. Examples are carmellose anion and lauryl sulfate, which is a component of lanette wax.

Insoluble salts can be made to mask an unpleasant taste or to prevent local irritation of the gastrointestinal tract. Well-known examples are ferrous fumarate suspension and potassium hydrogen tartrate suspension. In Chap. 13 more examples of insoluble salts to mask unpleasant tastes are discussed.

### 6.1.3 Solubility in Non-aqueous Solvents

When the aqueous solubility of an active substance is too low, it can be dissolved in a different solvent or mixture of solvents, which is compatible with the route of administration. The solubility of a lipophilic substance (a substance which dissolves well in the oil or fat but poorly in water) can be increased by making the dissolution medium (water) less polar by the addition of less polar but water-miscible solvents. Often mixtures of water, ethanol and propylene glycol are used. Also glycerol and macrogol (polyethylene glycol; PEG) can be used. The polarity of a solvent can be expressed by its dielectric constant, ε (for examples see Table 6.8, the higher the dielectric constant, the higher the polarity).

The dielectric constant of a mixture of solvents can be calculated as the sum of the dielectric constants of its com-

**Table 6.8** Dielectric constants of some solvents

| Solvent: | Dielectrical constant, ε (25 °C) |
|---|---|
| Water | 79 |
| Glycerol | 43 |
| Propylene glycol | 32 |
| Ethanol | 24 |
| Macrogol 400 | 12 |

ponents, each multiplied by the volume fraction of that solvent. The dielectric constant of a mixture of solvents A, B, … is thus:

$$\varepsilon_{\text{mixture}} = f_A \times \varepsilon_A + f_B \times \varepsilon_B + \ldots \quad (6.7)$$

where $f_A$, $f_B$ … are the volume fractions of the solvents A, B, respectively, and $\varepsilon_A$, $\varepsilon_B$, … the dielectric constants of the solvents A, B, …, respectively.

This equation can be used when the composition of a liquid mixture is to be changed, while keeping the dielectric constant the same. For example, assume that an active substance dissolves in a mixture of 20% v/v water and 80% v/v ethanol. As such a large volume percentage of ethanol is not desirable for many pharmaceutical applications; it has to be reduced to 20% v/v. In order to maintain the dielectric constant, macrogol 400 can be used to replace a large part of the ethanol. The volume percentages of water and macrogol 400 can be calculated as follows:

Original mixture:

$$\varepsilon_{\text{mixture}} = f_{water} \times \varepsilon_{\text{water}} + f_{ethanol} \times \varepsilon_{\text{ethanol}} = 0.2 \times 79 + 0.8 \times 24 = 35$$

Adjusted mixture:

$$\varepsilon_{\text{mixture}} = f_{water} \times \varepsilon_{\text{water}} + f_{ethanol} \times \varepsilon_{\text{ethanol}} + f_{PEG} \times \varepsilon_{\text{PEG}} = y \times 79 + 0.2 \times 24 + (0.8 - y) \times 12 = 35$$

It follows that the volume fraction of water, y, is 0.31 and the volume fraction of the macrogol 400, (0.8 – y), 0.49. The composition of the adjusted mixture by volume percentage is thus: water/ethanol/ macrogol 400 = 31/20/49.

Pharmaceutical examples of solutions where a relatively poorly water soluble substance is dissolved in a mixture of solvents are injections with diazepam or digoxin and oral solutions with lorazepam and paracetamol (Tables 6.9 and 6.10). In these preparations a mixture of ethanol, propylene glycol and water or a mixture of Macrogol 400, propylene glycol and glycerol is used as the solvent.

Paracetamol in an oral solution could at first be dissolved in 85% glycerol (Table 6.10).

**Table 6.9** Lorazepam oral solution 1 mg/mL [12]

| | |
|---|---|
| Lorazepam | 0.1 g |
| Macrogol 400 | 10 g |
| Propylene glycol | 3 g |
| Orange flavouring | 0.1 g |
| Glycerol 85% | 108.1 g |

**Table 6.10** Qualitative composition of paracetamol Oral Solution 500 mg/5 mL [13]

| | |
|---|---|
| Paracetamol (500–90) | |
| Citric acid monohydrate<br>Erythrosine | |
| Glycerol (85 per cent)<br>Macrogol 400<br>Propylene glycol<br>Methyl parahydroxybenzoate<br>Propyl parahydroxybenzoate | |
| Raspberry flavouring<br>Saccharin sodium<br>Sodium citrate | |
| Purified water | |

**Table 6.11** Qualitative composition of phenobarbital Sodium Solution for Injection 60 mg/mL [14]

| | |
|---|---|
| Phenobarbital sodium | |
| Disodium edetate | |
| Propylene glycol | |
| Water for injections | |

The relationship between the solvent medium, the pH and the $pK_a$ can be clearly illustrated using barbiturates as examples. Phenobarbital can be dissolved in water as its sodium salt. This requires a pH of 10 or higher. At this pH (it appears that) barbiturates decompose by ring-opening under the formation of malonylurea derivatives. The shelf life of such aqueous barbiturate solutions for oral administration is therefore limited to approximately 1 week. In addition, heat sterilisation of an aqueous solution for injection is not possible at this pH. In the formulation of injections, this problem has been solved by improving the solubility at slightly alkaline pH through the addition of a less polar solvent mixture e.g. a mixture of propylene glycol and ethanol or the addition of propylene glycol.

An example of a formula for a phenobarbital injection of 60 mg/mL is given in Table 6.11. The pH is usually adjusted to a value between 7 and 8. Phenobarbital in these injections is thus partially in the ionised form. The non-ionised form is dissolved by the organic solvent.

An alternative way to prepare solutions of lipophilic substances is to use solvents that are not miscible with water such as oils or esters of wax alcohols.

### 6.1.4 Solubility and Complex Formation

Sometimes two substances strongly interact (non-covalent) with each other in solution. In these cases complexes are formed whose solubility is different from that of the individual components. For example, iodine is only soluble in water in the form of a complex with iodide ions or with povidone. Iodinated povidone has the additional advantage that its solutions do not irritate the skin, as is the case for solutions where iodine is dissolved by complexation with potassium iodide.

Cyclodextrins are another category of substances that are used for complex formation. Cyclodextrins are ring-shaped oligosaccharides consisting of six, seven or eight glucose units referred to as alpha-, beta-, and gammacyclodextrin, respectively (see Fig. 6.2 for the chemical structure of betacyclodextrin) [15].

The Ph. Eur. has a monograph for betacyclodextrin: Betadex. Cyclodextrin forms a hollow truncated cone structure which is hydrophilic at the outside and contains a non-polar cavity into which lipophilic molecules fit (Fig. 6.3).

The diameter of the non-polar cavity increases when cyclodextrin contains more glucose units. As a consequence, small lipophilic substances form better complexes with cyclodextrins having a small non-polar cavity and large lipophilic substances better with cyclodextrins having a large non-polar cavity. On the outside, cyclodextrins are polar by which they are fairly soluble in water. The aqueous solubility of cyclodextrins can be affected by derivatisation. Hydroxypropylbetacyclodextrin (Hydroxypropylbetadex Ph. Eur.) for example dissolves much better in water than the non-derivatized betadex (betacyclodextrin). Thus, the aqueous solubility of lipophilic molecules increases after complexation with cyclodextrins. By selecting the right combination

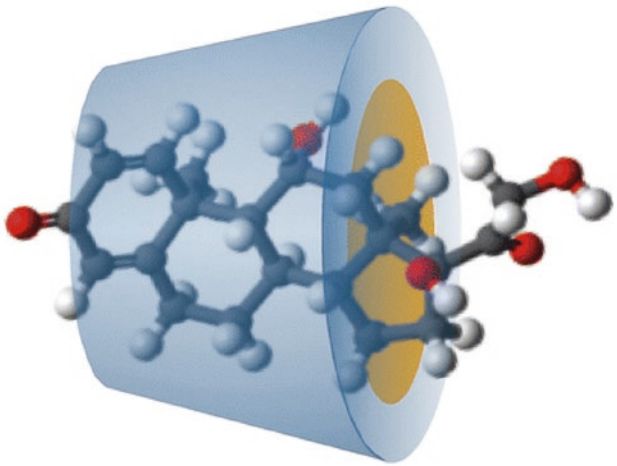

**Fig. 6.3** Schematic representation of a betadex–prednisolone complex

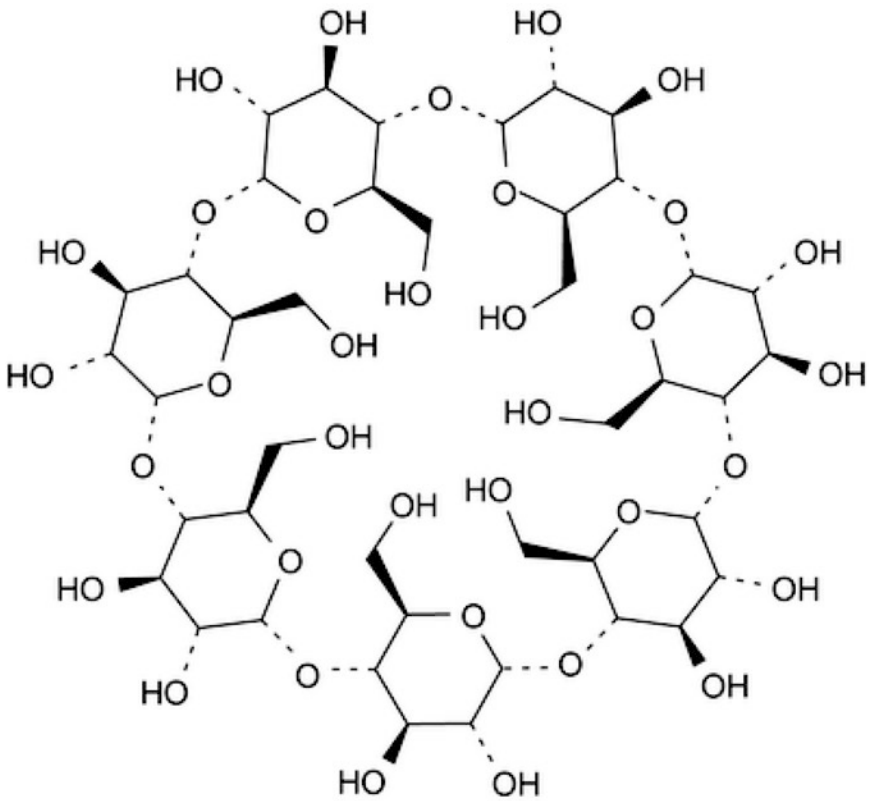


**Fig. 6.2** Chemical structure formula of betadex (betacyclodextrin)

of active substance and cyclodextrin, the solubility of poorly water-soluble substances can be increased.

Amongst other applications, cyclodextrins are used for the oral administration of lipophilic active substances. By complexing piroxicam with cyclodextrin a product is formed which shows an enhanced dissolution rate and therefore a more rapid absorption. An itraconazole oral mixture containing cyclodextrin has been developed to guarantee sufficient absorption of the active substance. Cyclodextrins are also used in parenteral preparations. Fluasteron is an antineoplastic agent which is preferably administered by injection to the patient at a concentration of 1000 micrograms per millilitre. Its aqueous solubility, however, is only 0.045 microgram per millilitre. With a 20% w/v hydroxypropylbetacyclodextrin solution, the desired concentration can be achieved [16].

### 6.1.5 Solubility of Derivatives

Derivatisation is a method of changing the properties of a molecule by means of a chemical reaction where an additional group is covalently coupled. By derivatisation, the solubility of substances can be either decreased or increased. In pharmacy, esters are important derivatives. This will be explained in greater detail on the basis of corticosteroids of which the C21-alcohol group can be esterified. These esters are not effective but should first be hydrolysed in the body. The rate of hydrolysis of esters in solution is strongly pH dependent. In a strongly acid or alkaline environment hydrolysis rapidly occurs. *In vivo*, the hydrolysis is catalysed by esterases.

Esters of corticosteroids with a polyvalent acid group, for example succinic acid or phosphoric acid, are soluble in their salt form. This is because most of the non-esterified acid groups are deprotonated when sufficiently high pH is chosen. Examples of corticosteroids which are esterified with polybasic acids are hydrocortisone sodium succinate, prednisolone sodium phosphate and dexamethasone sodium phosphate. In an aqueous medium with a pH of about 7 or higher, hydrocortisone sodium succinate is largely ionised and thus fairly soluble. At this pH, however, the hydrolysis of the ester is quite fast (this ester is most stable at a pH of about 4.5). As a result, dissolved in water, this substance is unstable and therefore its shelf life is limited.

In aqueous solution, prednisolone sodium phosphate and dexamethasone sodium phosphate are more stable than hydrocortisone sodium succinate. At a pH of about 8, the acid groups are sufficiently ionised to make the substance very soluble, while the phosphate ester is reasonably stable at this pH. Water soluble corticosteroids are mainly used in parenteral preparations, eye drops, oral mixtures and enemas. Examples are prednisolone oral mixtures and dexamethasone injections.

In order to increase their lipid solubility, corticosteroids have been esterified with monocarboxylic (fatty) acids. Examples include hydrocortisone acetate, beclomethasone dipropionate and betamethasone isovalerate. For triamcinolone, the following method has been developed to make the substance lipophilic: the hydroxyl groups at the C16 and C17-position of triamcinolone are used to form a cyclic acetal (Fig. 6.4). Although the formed substance, triamcinolone acetonide, is not a fatty acid ester, with regard to lipid solubility it behaves as such.

Fatty acid esters of corticosteroids may be used for the preparation of depot injections, either as oily solutions or aqueous suspensions. The nature of the fatty acid, in particular the length of the carbon chain, determines, to a large extent, its solubility and therefore its release and absorption rate into the bloodstream. In addition, in the case of corticosteroids, solutions or microcrystalline suspensions in either oil or water are used for intra- or periarticular injection.

Dermatology is another important field where fatty acid esters of corticosteroids are applied. Examples are creams with hydrocortisone acetate and triamcinolone acetonide, respectively. These fat-soluble derivatives are used because they show better penetration into the lipophilic stratum corneum of the skin.

Lipophilic esters of corticosteroids either suspended in water or dissolved in ethanol, propylene glycol, or macrogol 300 are applied in eye and ear drops, respectively. Table 6.12 gives an example of an aqueous suspension containing 20 mg/mL triamcinolone hexacetonide.

Flumetasone pivalate and fludrocortisone acetate are commercially available as solutions in macrogol and a mixture of glycerol and propylene glycol, respectively.

The aqueous solubility of free, not esterified, corticosteroids and their lipophilic esters do not differ substantially. However, their solubility in ethanol is better. Examples are: hydrocortisone, prednisolone and dexamethasone. These free corticosteroids are mainly used in solid oral dosage forms. Hydrocortisone is applied in Acid Ear Drops with Hydrocortisone 1% FNA (Table 6.13), because the underivatised steroid dissolves in propylene glycol, in contrast to its acetylated derivative. Although the lipophilic versions/variants are preferred in dermatological preparations, the free corticosteroids may also be applied.

### 6.1.6 Solubility and Supersaturation

If the solubility of a substance increases with temperature, its dissolution rate can be enhanced by heating. When a saturated solution is obtained at an elevated temperature, it will be supersaturated after cooling. Also solutions that are prepared at room temperature can become supersaturated when they have to be stored in the refrigerator because of their

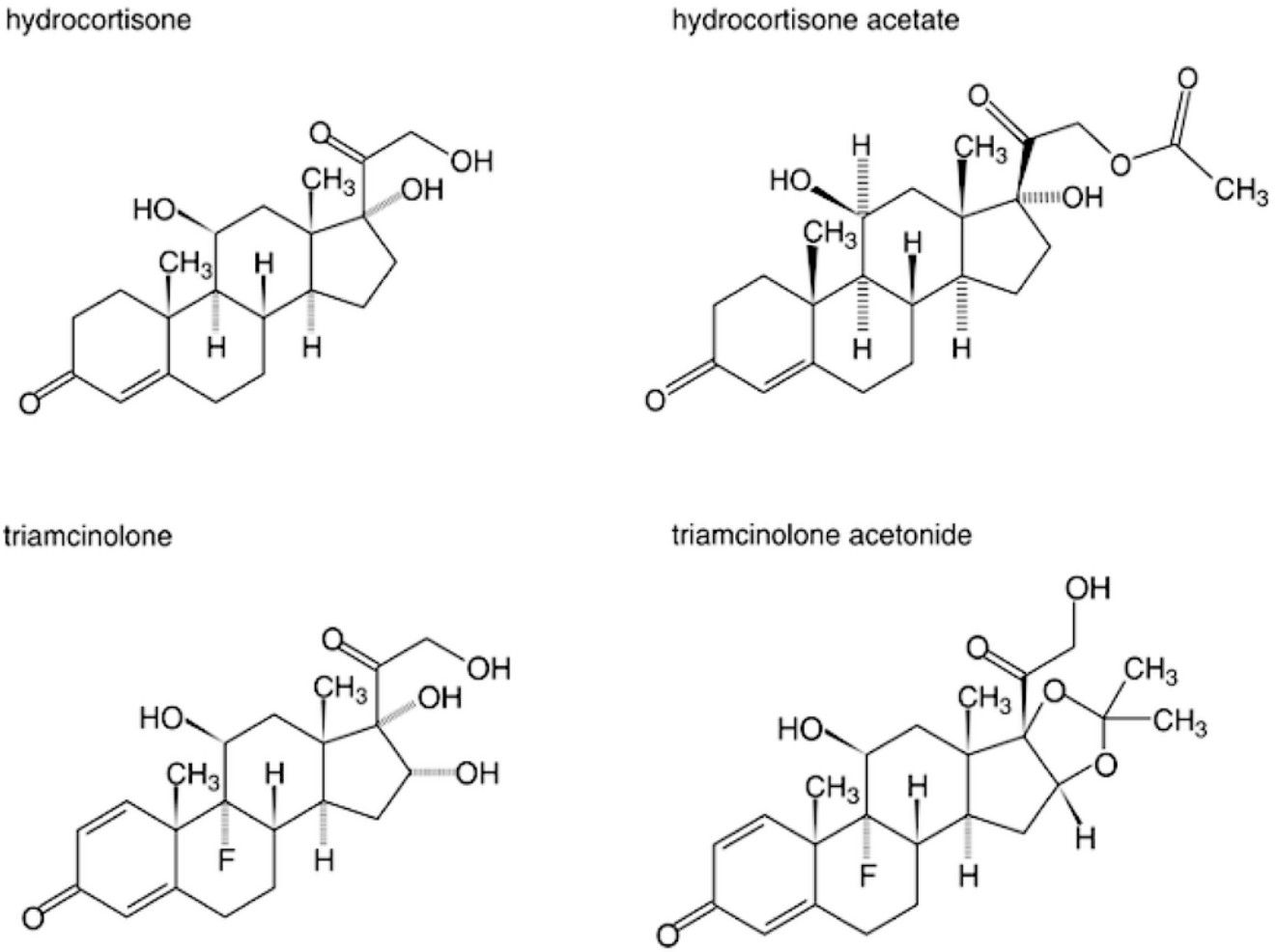


**Fig. 6.4** Chemical structure formulas of hydrocortisone, hydrocortisone acetate, triamcinolone and triamcinolone acetonide

**Table 6.12** Qualitative composition of triamcinolone Hexacetonide 20 mg/ml suspension for injection [17]

| |
|---|
| Triamcinolone hexacetonide |
| Liquid sorbitol |
| Polysorbate 80 |
| Benzyl alcohol |
| Water for injection |

**Table 6.13** Acid ear drops with hydrocortisone 1% [18]

| | |
|---|---|
| Hydrocortisone (micronised) | 1 g |
| Acetic acid (30%) DAC | 2.4 g |
| Propylene glycol | 96.6 g |
| **Total** | 100 g |

chemical instability. It may happen that supersaturation does not immediately result in crystallisation. Such a solution is called metastable and sooner or later crystallisation will occur. This process can proceed faster if solid particles are present in the solution. These solid particles may act as crystallisation nuclei and so initiate the crystallisation process.

The rate of crystallisation increases when the difference between the temperature at which the solution is prepared and the storage temperature increases. This is because the driving force for crystallisation will increase when the temperature difference increases. Supersaturation can occur in starting materials or intermediates, but supersaturated compositions are also used in therapy. Some relevant examples will be briefly discussed below.

Liquid Sorbitol 70% (crystallising) is a supersaturated starting material. Before using/processing this product in preparations the operator should check whether crystallisation has occurred. Crystals can be dissolved by heating and cooling again. Using a crystallised solution in several portions at different times will lead to a too low sorbitol concentration in the first portions, while in subsequent portions the concentration will be too high if the crystals have been dissolved again. Thus, when using a crystallised material in

preparations, there is a risk that the content in the final product will not be correct.

Calcium gluconate 100 mg/mL is an example of a supersaturated injection solution.

Mannitol 20% w/v is an example of a supersaturated solution for infusion. The solution should be inspected for the presence of crystals. As said, the crystals can be dissolved by heating and cooling again.

Magnesium Citrate Oral Solution USP (Table 6.14) is a supersaturated solution of magnesium citrate. The patient should be warned that crystallisation can take place after about 2 weeks, or earlier when the preparation is stored in the refrigerator.

Given the risks of crystallised products, the use of supersaturated solutions in pharmaceutical preparations should be avoided as much as possible.

## 6.2 Rheology

The rheology describes the flow behaviour of fluids. When a force is exerted on a liquid, it will start to flow. The resistance to flow is called dynamic viscosity but in practice it is usually simply referred to as viscosity. Consider a cube of liquid composed of slices like a stack of cards (Fig. 6.5).

When the bottom of the cube is immobilised and a shear force is applied at the top, the upper slice will get the maximum speed, the slice underneath a little less speed, and so on, until the bottom slice does not move. Consequently, there is a velocity gradient, $\frac{dv}{dy} = D$ (shear rate). According to Newton's law (6.8), viscosity (η) is defined as the force exerted per unit area (τ) (shear stress) divided by the shear rate (D). The unit of viscosity is Pascal.second (Pa.s), but in practice the derived unit millipascal.second (mPa.s) is generally used.

$$\eta = \frac{\tau}{D} \tag{6.8}$$

Rheology and viscosity are important for a proper understanding of the stability of disperse systems and for the accuracy of the dosage of liquid preparations.

**Table 6.14** Magnesium Citrate Oral Solution [19]

| | |
|---|---|
| Magnesium carbonate | 15 g |
| Anhydrous citric acid | 27.4 g |
| Lemon oil | 0.1 mL |
| Syrup | 60 mL |
| Talc | 5 g |
| Potassium bicarbonate | 2.5 g |
| Water, purified | Ad 350 mL |

### 6.2.1 Rheograms

Two types of flow behaviour can be distinguished, namely Newtonian and non-Newtonian flow. Non-Newtonian flow is further divided into plastic, pseudo-plastic and dilatant flow.

When the shear rate of a Newtonian fluid is plotted as a function of the shear stress, a straight line is obtained that passes through the origin (Fig. 6.6a). This means that at any shear stress, the viscosity of the fluid is the same, since shear stress divided by shear rate is constant. In this case, one may speak of the viscosity of a Newtonian fluid.

Conversely, the viscosity of non-Newtonian fluids is dependent on the applied shear stress and is referred to as apparent viscosity. When a fluid exhibits plastic flow, a certain minimum shear stress must be applied, called yield stress, before the fluid starts to flow (Fig. 6.6b). At a shear stress of less than the yield stress, the viscosity is thus infinitely large and the liquid behaves like a solid. Above the yield stress, the viscosity decreases with increasing shear stress. Also in the straight part of the curve in Fig. 6.6c, the viscosity decreases with increasing shear stress. This is because viscosity is shear stress divided by shear rate ($\tau/D$), and not the change of the shear stress divided by the change of the shear rate ($d\tau/dD$). Pseudo-plastic flow strongly resembles plastic flow. The difference is that pseudo-plastic fluids do not exhibit a yield stress. The consequence of this is

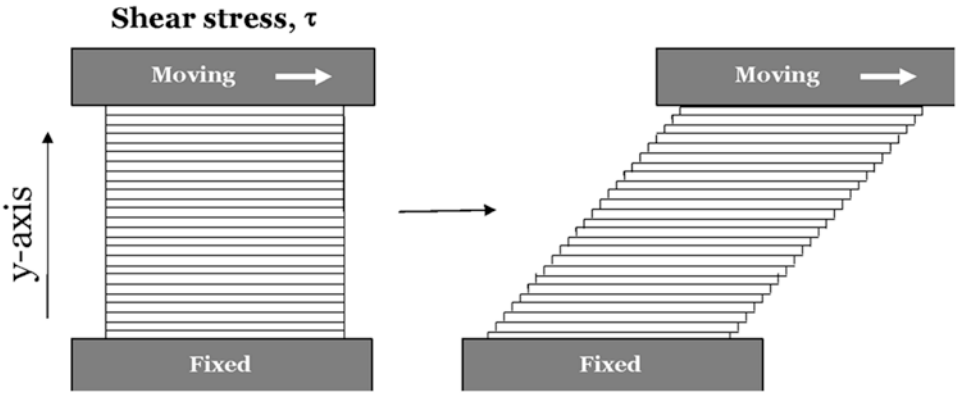


**Fig. 6.5** Effect of shear stress on a cube of liquid

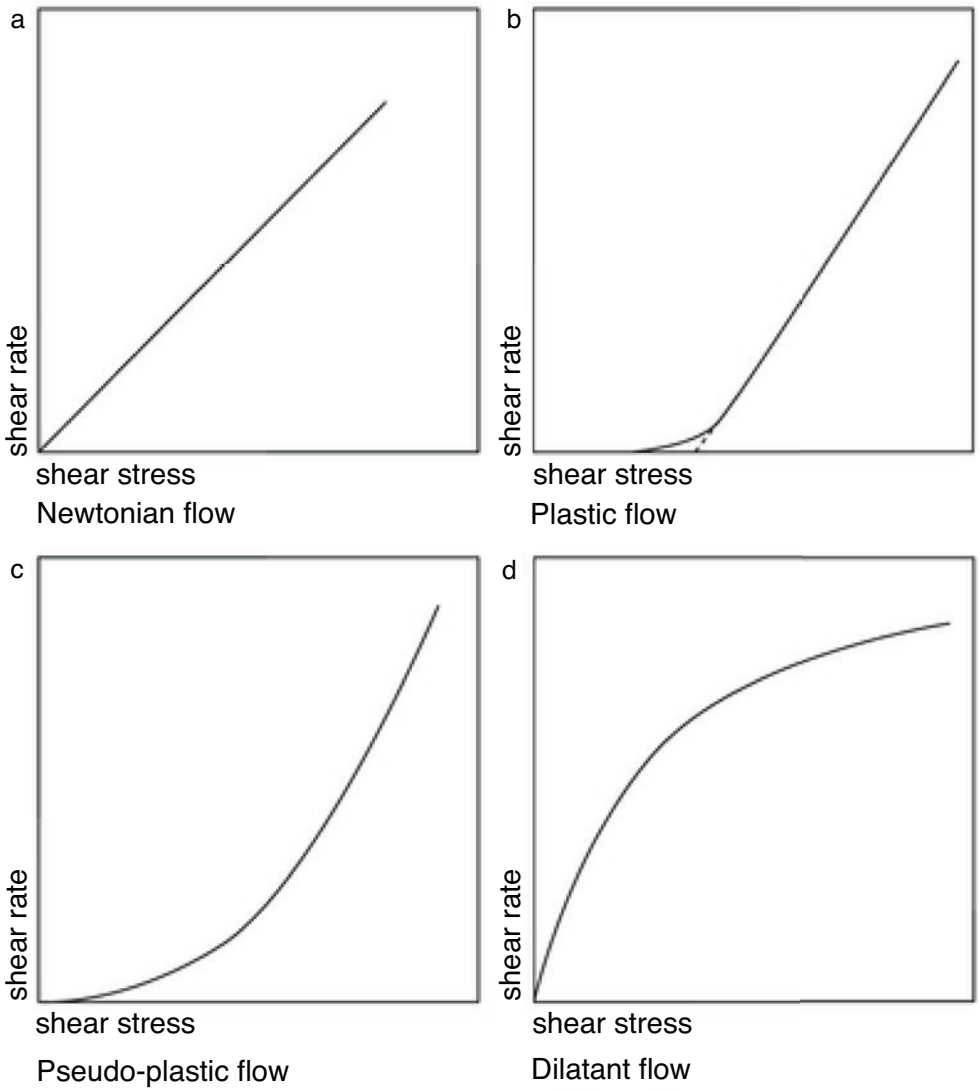


**Fig. 6.6** Rheograms of liquids with Newtonian, plastic, pseudo-plastic and dilatant flow behaviour. (Source: Recepteerkunde 2009, ©KNMP)

that the fluid never exhibits solid state behaviour because even at a very low shear stress flow takes place.

Gels and emulsions show plastic and pseudo-plastic behaviour. At rest, these systems are more viscous than when they flow. Creams and ointments should be easily spreadable, but they should not drip from the skin. Emulsions and suspensions should be as stable as possible and pouring should be easy.

This aim is even better achieved if, after a temporary high shear stress, the recovery of the higher viscosity is delayed. This phenomenon is called thixotropy. This phenomenon can often be seen in daily practice. For example, highly viscous emulsions like tomato ketchup and other cooking sauces temporarily have a lower viscosity after vigorous shaking. Also cutaneous emulsions often exhibit thixotropic behaviour.

Finally, if the viscosity increases with increasing shear stress, the flow behaviour is called dilatant (Fig. 6.6d). It is characteristic for pastes with a high content of solids. The pharmaceutical applicability of fluids with this flow behaviour is limited because they are often poorly spreadable. Only in a few cases this may be an advantage, namely when the paste must possess abrasive properties, as in cases where chalk is the main component. During rubbing, the skin is vigorously massaged due to the increasing viscosity of the paste.

### 6.2.2 Measurement Methods

The techniques to measure viscosity can be subdivided into two groups.

In the first group of methods, the relationship between the shear stress, $\tau$, and the velocity gradient, D, is measured. The main device within this group is the rotational viscometer. The method is described in the European

Pharmacopoeia [1]. In a vessel, a measuring body (spindle) is rotated in the test sample. The resistance to the rotation speed is measured as torque in the shaft. Because several combinations of vessel and spindle can be chosen for this equipment, the torque of both fluids with a very high and very low viscosity can be measured accurately. For each combination, the manufacturer supplies a table or often computer software, which can be used to derive the viscosity from the torque and the rotational speed. Thus, with this method the viscosity can be calculated directly. The method is applicable to both Newtonian and non-Newtonian fluids. By varying the rotational speed, complete rheograms can be obtained. This provides information about both non-Newtonian behaviour and thixotropy. The method is especially useful in the investigation of the stability of viscous systems.

In the second group of methods, gravity is used as a force to bring the fluid in motion. The suspended level viscometer as described in [20] is based on this principle. The Ford cup, a sort of funnel through which the fluid flows, is another example. The flow rate of a fluid per unit area of the outlet opening of, for example a capillary tube or cup is proportional to the viscosity, η, but inversely proportional to the relative density, ρ, of the fluid. The rate at which the fluid level decreases thus depends on η/ρ. Viscosity divided by density, η/ρ, we call kinematic viscosity, with $m^2/s$ as unit. By calibrating the device with a fluid having a known kinematic viscosity, the kinematic viscosity of the fluid can be calculated. The equipment used for these methods are usually much cheaper than those in the first group. Moreover, they are useful for designing preparations. The disadvantage of these viscometers is that they are only suitable for measuring the viscosity of Newtonian fluids. However, with these viscometers information about the pouring behaviour can be achieved which is, for example, relevant to emptying a bottle.

With an extensometer and a penetrometer, the viscosity is not directly measured but information about the flowability of materials can be achieved. In an extensometer, an ointment or cream is allowed to spread between two glass plates which yields the spreading capacity. With a penetrometer, the resistance during the penetration of a pin connected to a cone into a product is measured. This method can be used to characterise highly viscous preparations, e.g. ointments.

## 6.3 Phase Behaviour

### 6.3.1 Gibbs' Phase Rule

Many pharmaceutical preparations are heterogeneous systems, meaning that they consist of multiple phases, e.g. emulsions and suspensions. To physically describe such systems in equilibrium, Gibbs' phase rule can be applied:

$$F = 2 + C - P \tag{6.9}$$

Where F is the number of independent intensive parameters or the number of degrees of freedom, C is the number of components, and P is the number of phases.

The three parameters in this equation require some further clarification. An intensive parameter is independent of the size of the system under evaluation while an extensive parameter is dependent on the size of a system. For example, when one liter of water of 300 K is added to another liter of water of 300 K, two litres of water of 300 K is obtained, obviously not one litre of 600 K. It means that temperature is an intensive parameter while volume is an extensive parameter. Other examples of intensive parameters are pressure, density, composition, refractive index, and (partial) molar quantities like molar volume and chemical potential of a compound. Other examples of extensive parameters are mass and energy. The number of components is the number of chemical compounds minus the number of relationships between these compounds. E.g. a mixture of water and ethanol, contains two components. However, when considering e.g. a mixture of gaseous $NO_2$ and $N_2O_4$, there is only one component as these substances are in equilibrium: $2\ NO_2 \rightleftharpoons N_2O_4$ and evidently there is a relationship between the two chemical compounds. A phase is defined as an area in which all intensive parameters are the same while at the surface or interface at least one intensive parameter changes abruptly. Liquid water and liquid ethanol mix on a molecular level and hence forms one phase. Liquid water and liquid chloroform, however, do not mix and hence two phases are present. Density and composition are amongst other intensive parameters that change abruptly at the interface between water and chloroform. A phase can be solid, liquid or gaseous. It should be noted, however, that in a system in equilibrium, there can only be one gaseous phase, e.g. while liquid water and liquid chloroform do not mix, they do so in the gaseous phase. Below, the application of the Gibbs' phase rule will be illustrated for one and two component systems.

### 6.3.2 Application of the Gibbs' Phase Rule to One Component Systems

The application of the Gibbs' phase rule will be illustrated using water as an example. The phase diagram of water is shown in Fig. 6.6a. Passing one of the curves indicates that a phase transition occurs, which means that on the curve two phases are in equilibrium. The point where the three curves meet (also referred to as the triple point) indicates the equilibrium between the solid, liquid and gaseous phase. When water is only present as one phase, then C = 1 and P = 1, thus F = 2 + 1 − 1 = 2. Assuming water in the gaseous phase

behaves as a perfect gas, the equation: $pV_m = RT$ can be applied where p is the pressure, $V_m$ the molar volume, T is the temperature (in K) and R is the gas constant. The equation contains three intensive parameters: pressure, molar volume and temperature. However, as the number of independent intensive parameters is two, not all three parameters can be freely adjusted at the same time, but only two. The reason for this is that these three parameters are related in the equation for a perfect gas:

$$p = \frac{RT}{V_m} = f\left(T, V_m\right) \tag{6.10}$$

Thus, the temperature and the molar volume can be adjusted but then the pressure is fixed.

$$V_m = \frac{RT}{p} = f\left(T, p\right) \tag{6.11}$$

Thus, the temperature and the pressure can be adjusted but then the molar volume is fixed.

$$T = \frac{pV_m}{R} = f\left(p, V_m\right) \tag{6.12}$$

Thus, the pressure and molar volume can be adjusted but then the temperature is fixed.

Normally, the temperature and the pressure are selected as independent intensive parameters (as also shown in Fig. 6.7a). Although the equation for a perfect gas can obviously not be applied for the solid or the liquid phase, also when either one of these phases is present, the pressure and the temperature can be adjusted freely with the molar volume being fixed.

When two phases are in equilibrium, then C = 1 and P = 2, thus, F = 2 + 1 – 2 = 1. In the phase diagram, the curves indicate that two phases are in equilibrium. Therefore, it can be easily deduced that either the pressure can be adjusted but then the temperature is fixed or the temperature can be adjusted but then the pressure is fixed (see also Fig. 6.7b).

When three phases are in equilibrium, then C = 1 and P = 3, thus, F = 2 + 1 – 3 = 0. In the phase diagram, where the three curves meet, indicates that the three phases are in equilibrium (see also Fig. 6.7b). Obviously, now neither the temperature nor the pressure can be adjusted, as they are both fixed.

### 6.3.3 Application of the Gibbs' Phase Rule to Two Component Systems

Generally, when considering two component systems, the gaseous phase is ignored and the pressure is kept constant. This is also referred to as a condensed system. In particular for pharmaceutical preparations this is justified because the gaseous phase is often not relevant and formulations are usually exposed to ambient pressure. Keeping the pressure constant implies that the Gibbs' phase rule now becomes F = 1 + C – P. Generally, the temperature is plotted as a function of the composition. The composition on the x-axis can be determined following the lever rule as shown in Fig. 6.8. Application of the Gibbs' phase rule will be illustrated using a water/phenol exhibiting liquid-liquid phase separation and a dapsone (DAP)/3-hydroxybenzoic acid (3HBA) combination exhibiting both solid-liquid and solid-solid phase separation.

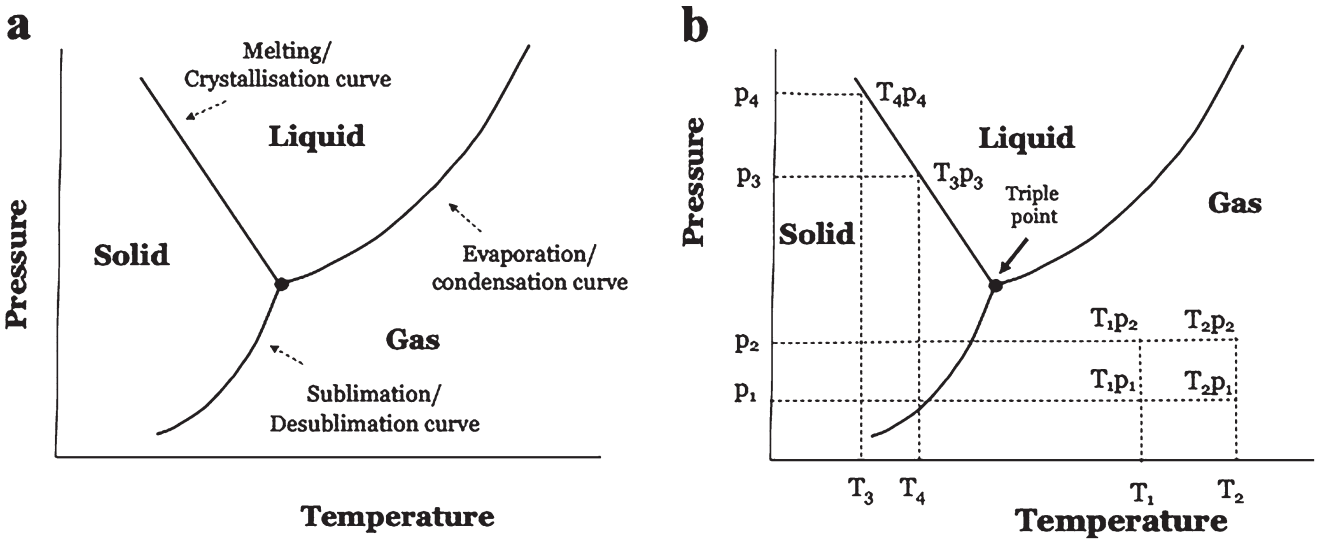


**Fig. 6.7** Plain phase diagram of water (**a**), and phase diagram of water showing that when only one phase is present (e.g. only gas) both temperature and pressure can be chosen independently, when two phases are present (e.g. solid and liquid) either temperature or pressure can be chosen independently; when all three phases are present neither temperature nor pressure can be chosen (**b**)

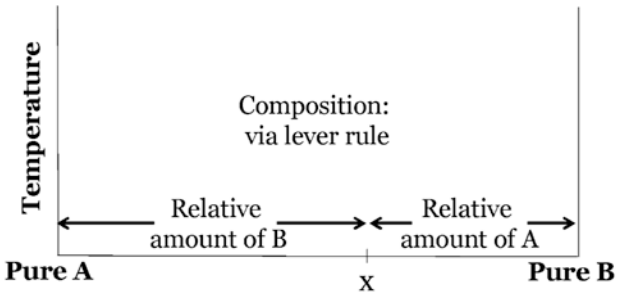


**Fig. 6.8** Set-up of a phase diagram for a condensed system

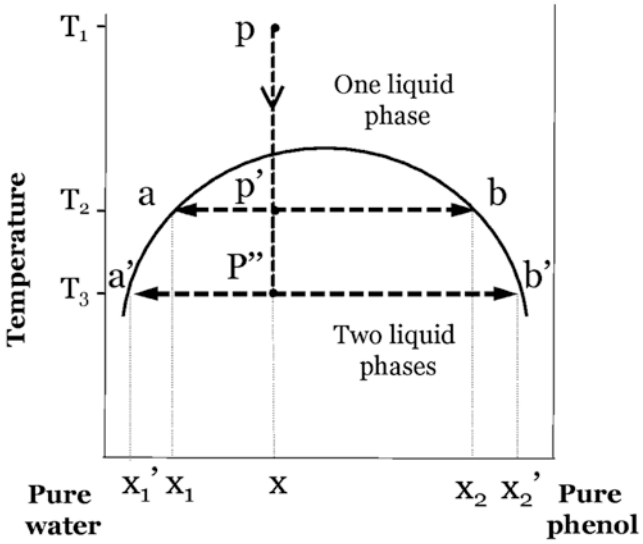


**Fig. 6.9** Phase diagram for a condensed system composed of a mixture of water and phenol

The phase diagram of the water/phenol system is shown in Fig. 6.9. As can be seen, at high temperature, water and phenol form one liquid phase. According to the Gibbs' phase rule the number of independent variables becomes: $F = 1 + C - P = 1 + 2 - 1 = 2$. Therefore, temperature and composition, i.e. water/phenol ratio, can be adjusted freely. When cooling this homogeneous mixture liquid-liquid phase separation (also referred to as liquid-liquid demixing) occurs when passing the solid curve, which is also referred to as the binodal. E.g. when cooling a homogeneous mixture with composition X in Fig. 6.9 from temperature $T_1$ to $T_2$, (thus from point p to point p'), the system is not in equilibrium anymore and two new liquid phases are formed having composition $X_1$ (point a) and $X_2$ (point b), meaning that one phase is rich in water and the other phase is rich in phenol, respectively. When cooling further to $T_3$, liquid-liquid phase separation becomes more pronounced: the newly formed liquid phases now have a composition $X_1$' (point a"; richer in water) and $X_2$' (point b'; richer in phenol). According to the Gibbs' phase rule, when two phases are formed the number of independent intensive parameters is: $F = 1 + 2 - 2 = 1$. In this situation, the temperature can be adjusted, but then the composition of the two phases is fixed. Alternatively, the composition of one of the two phases can be adjusted, but then the composition of the other phase and the temperature are fixed.

The phase diagram of the DAP/3HBA combination is a bit more complicated as shown in Fig. 6.10a. It is a typical example of a so-called eutectic mixture. Similar to the water/phenol combination, both components form a homogeneous liquid mixture at high temperatures and consequently, temperature and composition can be adjusted freely as obviously also in this situation the number of independent intensive parameters is two. The curve at the left side of Fig. 6.10a represents the crystallization curve of DAP. This curve decreases with an increasing 3HBA/DAP ratio. When passing the crystallization curve of DAP by cooling a homogeneous solution, solid-liquid phase separation occurs (see Fig. 6.10b). Thus, when a homogeneous solution with composition $W(T_0)$ is cooled from $T_0$ to $T_1$, pure DAP crystals and a solution of composition $W(T_1)$ are formed. Obviously, the 3HBA/DAP ratio of the remaining solution is higher than before phase separation, as DAP molecules are removed from the solution in the form of crystals. When further cooling to $T_2$, more DAP crystals are formed and therefore the 3HBA/DAP ratio in the remaining solution is further increased. The curve at the right side of Fig. 6.10c represents the crystallization curve of 3HBA and basically a similar phase separation occurs when cooling a homogeneous solution, but now 3HBA crystallizes and the 3HBA/DAP ratio decreases upon further cooling. The number of independent intensive parameters is one; either the composition of one of the two liquid phases can be adjusted but then the temperature is fixed or the temperature can be adjusted but then the composition of one of the two phases is fixed. The crystallization curves of DAP and 3HBA meet each other in the eutectic point. Cooling to below the eutectic temperature (Te), both DAP and 3HBA crystallise and thus two solid phases are formed (Fig. 6.10d). Again, the number of independent intensive parameters is one, but in this situation, only the temperature can be adjusted as the composition of the two phases are fixed, being pure DAP and pure 3HBA crystals.

## 6.4 Interfaces and Surface Active Agents

Many pharmaceutical preparations consist of several phases. For example, a suspension consists of solid particles dispersed in a liquid and an emulsion is composed of drops of a liquid dispersed in a second liquid which is immiscible with the first. Interfaces exist between the different phases. In this section the properties of interfaces are discussed.

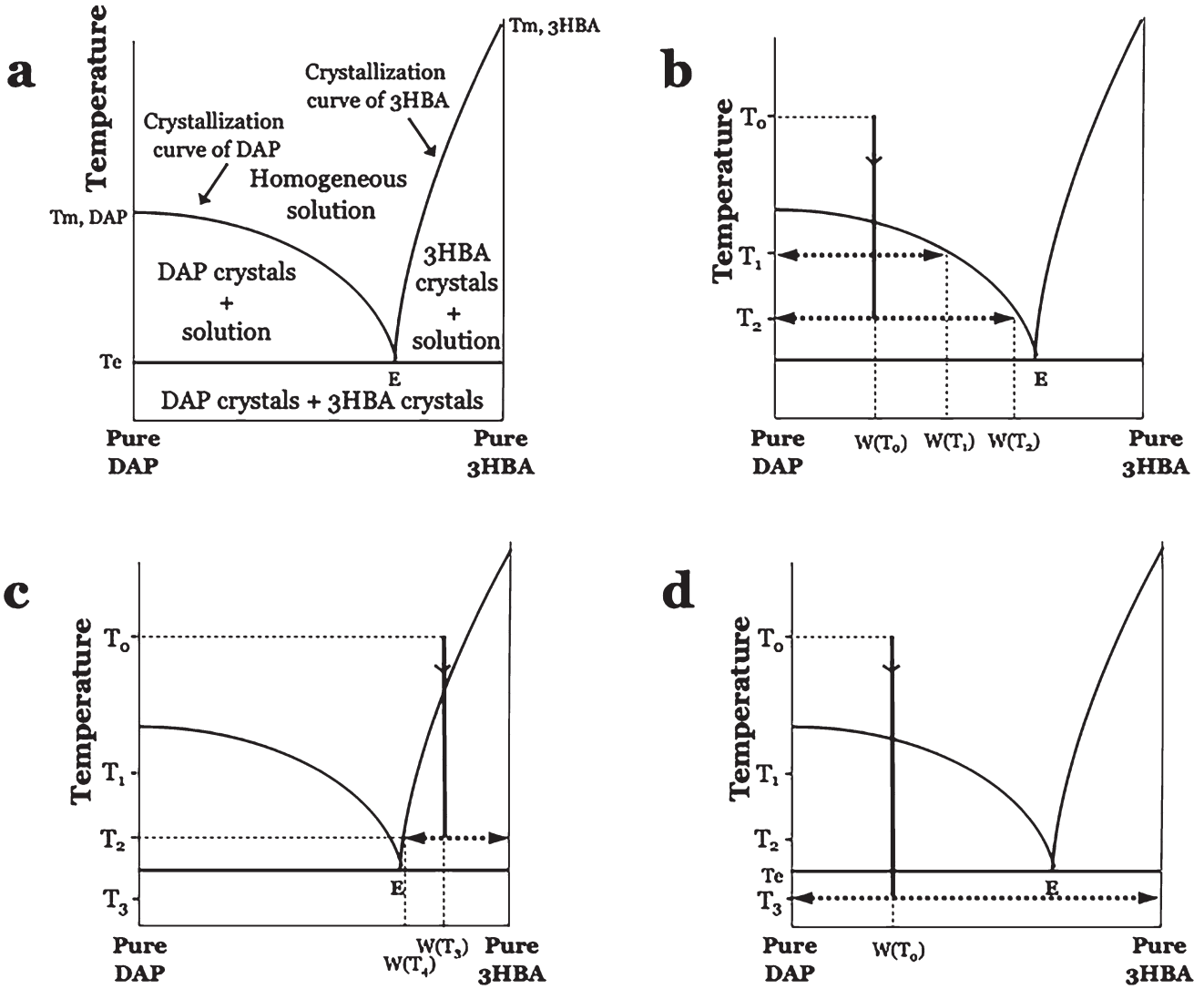


**Fig. 6.10** Phase diagram for a condensed system composed of a mixture of DAP and 3HBA

### 6.4.1 Surface and Interfacial Tension

Molecules exert attractive (Van der Waals) forces to each other. In the bulk of a liquid, this force is equal in all directions. Because fewer molecules per unit of volume are present in the air than in a fluid, the Van der Waals forces at the interface of liquid and air are smaller in the direction of air than in the direction of liquid. As a result, there is a net force (or pressure if force per unit area is considered) perpendicular to the liquid surface in the direction of the bulk of the liquid. Due to this pressure, which is called surface tension, the molecules located at the surface are in a higher Gibbs free energy state than in the bulk of the liquid. This implies that if a new surface is created, molecules are in a higher Gibbs free energy state. In other words, it takes Gibbs free energy to create a new surface. A similar situation occurs at the interface of two non-miscible liquids and the interface between a liquid and a solid. In these cases one speaks of interfacial tension. The terms surface tension and interfacial tension are often used interchangeably. The unit of surface and interfacial tension is $J/m^2$ [5, 21].

Molecules consisting of a highly water soluble or hydrophilic moiety and a highly oil soluble or lipophilic moiety are called amphiphilic. These molecules accumulate at interfaces in such a way that the water soluble or hydrophilic part is oriented to the hydrophilic (aqueous) phase and the oil soluble part to the lipophilic (oil) phase. This causes lowering of the surface or interfacial tension. Therefore, the Gibbs free energy required to create a new surface or interface can be reduced by using amphiphilic substances. In the context of this application, these substances are referred to as surface active agents or surfactants. An overview of surfactants used in pharmaceutical preparations can be found in Chap. 7.

In order to make a proper choice of a surfactant for a specific preparation, a system has been developed in which the relative contribution of the hydrophilic and the lipophilic part

of the molecule is expressed as a number. This system is called hydrophilic-lipophilic balance, abbreviated as HLB. Surfactants that are equally soluble in oil and in water have an HLB value of 7. Surfactants which are more soluble in oil have a HLB-value < 7 and surfactants which are more soluble in water have an HLB-value > 7. Surfactants may have an HLB value of 1–40. In designing of a formula, the HLB-value is generally used as a number to characterise surfactants.

### 6.4.2 Wetting

Wetting means bringing a solid in close contact with a liquid. This process is of importance for example for the preparation of suspensions and the disintegration of a tablet in the gastrointestinal tract.

When a drop of a liquid is placed on a solid (smooth) surface, the drop will spread depending on the properties of the liquid and of the solid. For example, a drop of water completely spreads out on a (clean, grease-free) glass but hardly on a Teflon® surface. The angle formed by a liquid at the three phase boundary where a liquid, air and solid intersect as shown in Fig. 6.11, is called the contact angle θ. Young's law describes the relationship between the surface and interfacial tension of the three phases: air (A), solid (S) and liquid (L) and contact angle (θ). In equilibrium, the forces will compensate each other at the point P, so:

$$\gamma_{SA} = \gamma_{SL} + \gamma_{LA} cos\theta. \tag{6.13}$$

or

$$cos\theta = \left(\gamma_{SA} - \gamma_{SL}\right) / \gamma_{LA} \tag{6.14}$$

When the contact angle θ is larger than 90° it is called poor wetting and when it is smaller than 90° it is called good wetting. Wetting can be improved by the addition of surfactants to the aqueous phase. As a result, the contact angle θ becomes smaller (thus cos θ becomes larger) and finally approaches 0. In complete wetting (θ = 0°) the drop spreads out completely on the surface.

When an aqueous suspension is prepared, the air at the surface of the particles should be replaced by the aqueous phase. When a poorly wettable substance is used (θ > 90°) this will not happen. The particles can float when the air is not replaced. This phenomenon is called flotation. Poorly wettable and floating particles often adhere to the wall of the bottle neck. As a result, dosing is difficult with such systems. Pharmaceutical examples of poorly wettable substances are: phenytoin, sulfur, zinc oxide and barium sulfate. By adding a surfactant, the interfacial pressure can be reduced by which wetting is improved.

To reduce the risk of foaming, surfactants that lower the interfacial pressure only moderately are mostly used.

To improve wetting, surfactants with an HLB larger than 15 are used, such as sodium dioctyl sulfosuccinate (Aerosol OT ®), ethoxylated castor oil, poloxamer, sodium salts of higher alcohol sulfates (sodium lauryl sulfate) and polysorbates, but also polymers with interfacial activity as, for example hydrogel formers as methylcellulose, hypromellose, etcetera. A further discussion of these substances can be found in Chap. 7. In practice, propylene glycol or a thickening agent will lower the interfacial tension to a sufficient extent.

### 6.4.3 Micelle Formation and Solubilisation

Surfactants are not only applied to reduce the surface or interfacial tension. They can also be used to increase the solubility of substances. When dissolved in water above a certain minimum concentration, aggregates of surfactants are formed, also called micelles. Lipophilic substances can be brought into solution (solubilised) using these micellar solutions. Micelles can be prepared as follows.

Consider an aqueous solution of which the surfactant concentration is gradually increased. At low concentration of the surfactant, the molecules are predominantly located at the surface. If the concentration is increased, more of the surfactant molecules will be present at the surface and the surface tension will decrease. Above a certain concentration, however, the surface will be full and the extra molecules will migrate into the bulk of the solution. When this occurs, the surface tension will not further decrease.

The surfactants molecules that migrate into the bulk of the solution form micelles. The structure of the micelles is such that the lipophilic part of the surfactant molecules is at the inside of the micelle, so that the thermodynamically unfavourable contact of the lipophilic part of the surfactant mol-

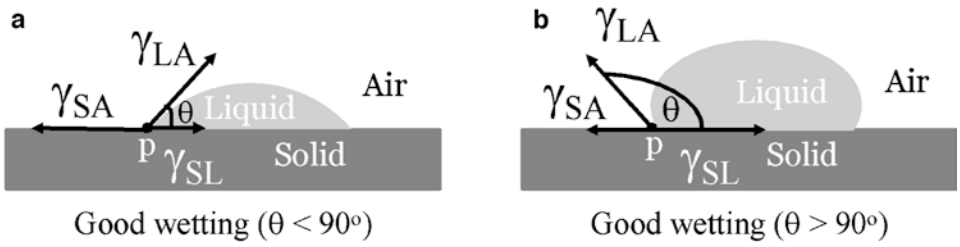


**Fig. 6.11** Equilibrium of forces at the point P due to surface and interfacial tensions. The force vectors are shown as arrows

ecules with the water molecules is minimal. The minimum concentration at which micelles are formed is called the critical micelle concentration (CMC). The exact shape of the micelles and the CMC differ from substance to substance but are usually spherical. Micelles have the size of colloids and also behave physico-chemically as such. This behaviour is discussed in more detail in Sect. 6.5.1.

In an oil phase, micelles with the inverted structure can be formed, i.e. the hydrophilic part of the surfactant molecules is oriented towards the inside. These micelles are also called reverse micelles.

Many poorly water soluble substances can, molecularly disperse or absorb into the core of the micelles. These micellar solutions are optically clear. This is called solubilization (Fig. 6.12). Examples from the Dutch formulary FNA are oral aqueous preparations with vitamin A (retinol palmitate) or D (cholecalciferol (Table 6.15)). Another example is a licensed oral mixture with ciclosporin.

The order of mixing the components in this preparation is important in order to achieve solubilisation: first polysorbate to cover the flask with a layer, then carefully adding and mixing the cholecalciferol concentrate and star anise oil. The ratio between the amount of polysorbate and the cholecalciferol concentrate is also important.

Solubilisation may also be undesirable. Solutions containing polysorbate cannot be preserved with for example methyl or propyl parahydroxybenzoate. The preservative effect of these substances is inhibited by solubilisation or even nullified. With sorbic acid/sorbate, preservation is possible, provided that a higher than conventional concentration is used.

## 6.5 Disperse Systems

Many liquid and semi-liquid pharmaceutical preparations are disperse systems. Disperse systems are defined as systems in which a substance is distributed as particles (discontinuous) into a dispersion medium (continuous). Three types of disperse systems will be discussed which are pharmaceu-

**Table 6.15** Cholecalciferol oral solution[a] 50.000 IU/mL [22]

| | |
|---|---|
| Cholecalciferol concentrate (oily form) 1000,000 IU/g | 5.0 g |
| Citric acid monohydrate | 0.48 g |
| Polysorbate 80 | 25 g |
| Potassium sorbate | 0.6 g |
| Star anise oil | 0.22 g |
| Syrup BP | 12.5 g |
| Water, purified | 60.2 g |
| **Total** | 104 g (= 100 mL) |

[a]This solution is actually a solubilisate

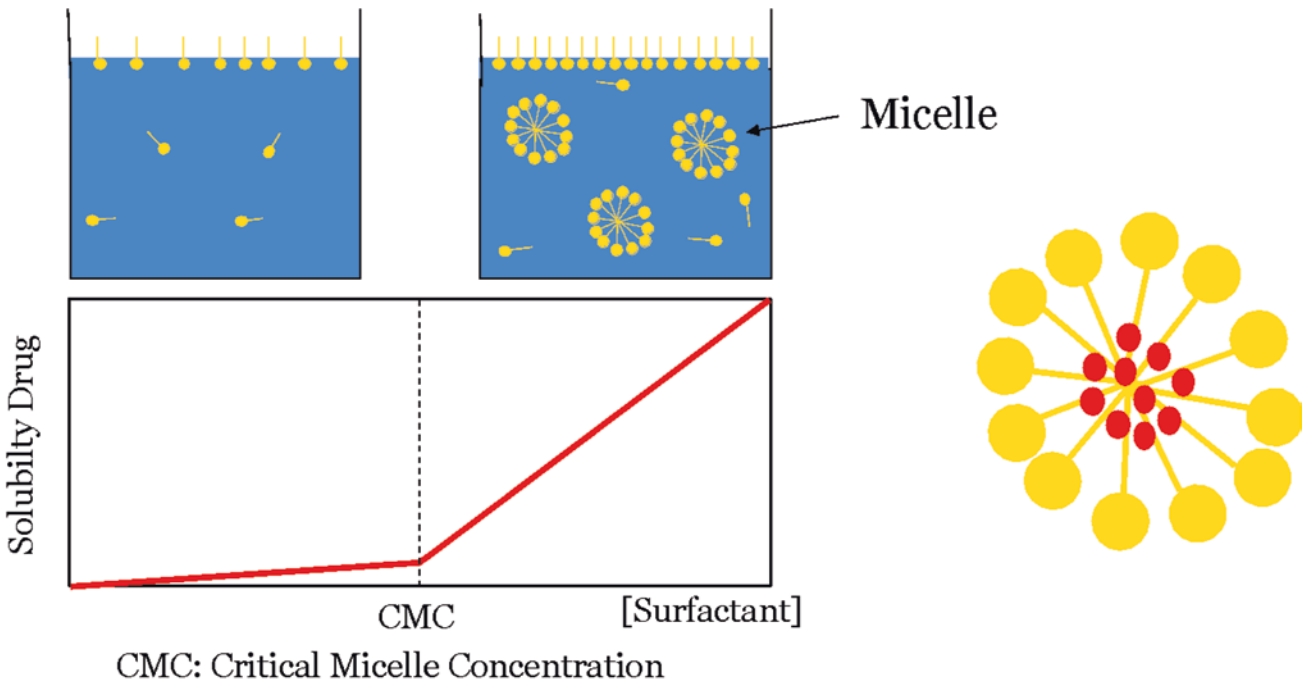


**Fig 6.12** Micelle formation and its effect on the solubility of poorly water soluble substances

tically relevant: colloidal systems, suspensions and emulsions. In both colloidal systems and suspensions, solid particles are dispersed in a liquid. The difference is that in colloidal systems the particles do not settle, while they do in suspensions. This difference is caused by the size of the particles. In colloidal systems, the particles are so small (1 nm – 1 μm) that the Brownian motion (diffusion caused by thermal energy) is stronger than the force of gravity so that they remain suspended in the liquid and do not settle. In suspensions, the particles are larger (>1 μm) and as a consequence the force of gravity is stronger than the Brownian motion which makes them settle (if the density of the particles is larger than that of the dispersion medium). Emulsions consist of non-miscible liquids. Two types of emulsions will be discussed: oil drops disperse in water (oil-in-water emulsion or o/w emulsion) and water drops disperse in oil (water-in-oil emulsion or w/o emulsion). There are also more complex structures such as w/o/w emulsions and bi-continuous systems. However, these systems will not be discussed.

## 6.5.1 Colloidal Systems

### 6.5.1.1 Lyophilic and Lyophobic Systems

Colloidal systems can be divided into lyophilic and lyophobic systems. Lyophilic colloids have a strong affinity with the dispersion medium by which a solvation shell around the particle is formed. This process is called solvation and if the dispersion medium is water it is called hydration. A polysaccharide dissolved in water is an example of a lyophilic colloidal system. The solvation shell is formed by hydrogen bonds between the hydroxyl groups of the polymer molecules and the water molecules. Pharmaceutical examples are solutions of dextran, used as plasma expanders. Micelles are also lyophilic colloids. Example of such a system is the aqueous cholecalciferol oral mixture (Table 6.15). In these preparations, a lipophilic fluid is dissolved in an aqueous medium by incorporating it in micelles. Because this type of colloids falls apart on dilution to concentrations below the CMC, they are also known as association colloids. Lyophobic colloids have no affinity with the dispersion medium. Thus, in this type of colloids no solvation shell is formed around the particles. An example of lyophobic particles are colloidal gold particles (with a diameter of 1 nm – 1 μm) disperse in water. There are no hydrogen bonds or other interactions between the gold particles and the water molecules, so the solvation shell is missing. If the dispersion medium is water, lyophilic colloids and lyophobic are also referred to as hydrophilic and hydrophobic colloids, respectively.

### 6.5.1.2 Stabilisation of Colloidal Systems

Basically, colloidal systems are not thermodynamically stable. The particles have the tendency to attract each other by Van der Waals forces and aggregation can take place. Yet there are many colloidal systems that can be stored for extended periods of time without this happening. This is because two stabilising mechanisms may play a role.

Around lyophilic colloids a solvation shell is formed, which acts as a protective layer around the particles. This protective layer prevents the two particles from approaching each other too closely. In addition, the particles repel each other when they are electrostatically charged. This repulsion is not fully determined by the charge at the surface of the particle (Nernst potential) but by the charge at a small distance from the particle which is called zeta or ζ-potential. During the diffusion of the particle through the dispersion medium a layer of the dispersion medium around the particle is dragged along with it. Therefore, it is not the charge of the particle, but the charge of the particle together with this layer of dispersion medium that is relevant for the stability of the system. If the particle is charged and when ions are present in the dispersion medium, there will be more ions of opposite charge (counter ions) in the near vicinity of the particle than ions of the same charge due to electrostatic attraction. The charge of the particles is therefore neutralised to a certain extent. This neutralisation increases with increasing ionic strength. This implies that the zeta-potential is smaller than the Nernst potential and decreases with increasing ionic strength. If the dispersing medium contains polyvalent counter-ions, the zeta-potential and the Nernst potential can even have opposite charges.

The zeta-potential can also be influenced by the absorption of specific ions from the dispersion medium onto the surface of the colloidal particle. For example, if a positively charged surfactant adsorbs onto a positively charged colloidal lyophobic particle, the zeta-potential becomes larger than the Nernst potential.

Moreover, the adsorption of surfactants onto lyophobic particles has a second effect. Because only the lyophobic part of the surfactant adsorbs onto the lyophobic particle, its lyophilic part is oriented towards the dispersion medium. This lyophilic part forms a protective layer by which the particles can approach each other less easily. This effect is called steric stabilisation. Surfactants usually do not adsorb onto lyophilic particles. However, by covalently linking hydrophilic polymers to their surface, we can also achieve steric stabilisation of lyophilic colloidal particles. A detailed description of the forces of attraction and repulsion can be found in literature [5, 21].

So a lyophilic colloidal system can be stabilised by two mechanisms, namely by a solvation shell and by electrostatic repulsion. This implies that a lyophilic colloidal system with a zero zeta-potential does not necessarily have to be unstable. This is because the stabilising effects of the solvation shell may be sufficient. A lyophobic colloidal system, however, lacks a solvation shell. A lyophobic colloidal system can therefore only be stable if the zeta-potential is sufficiently high (positive or negative).

### 6.5.1.3 Destabilisation of Colloidal Systems

A sol is a colloidal system in which the repulsion forces between the colloidal particles dominate in such a way that they can move freely with respect to each other. Lyophilic colloidal particles can be destabilised either by making the particles more lyophobic or by reducing the zeta-potential or both. Lyophobic colloidal particles, however, can only be destabilised by reducing the zeta-potential. Lyophilic colloidal particles can be rendered more lyophobic by adding a fluid which is miscible with the dispersion medium but in which the colloidal particles are lyophobic (for example, ethanol when the dispersion medium is water). The zeta-potential of colloidal particles (lyophilic or lyophobic) can be reduced by adding an electrolyte to the system. When a sol is partially destabilised, the forces of attraction are stronger than the forces of repulsion. As a result, the particles will no longer be able to move freely with respect to each other, but they will form a continuous three-dimensional network extending throughout the dispersion medium. Such a structure is called a gel and it is called a hydrogel if the dispersion medium is water. This sol-gel transition can be observed by the flow behaviour. Because the colloidal particles in sols move more or less freely with respect to each other, Newtonian or pseudo-plastic flow behaviour can be observed. Gels exhibit a yield stress because first the continuous three-dimensional structure must be broken down before flow can occur. In this case plastic flow behaviour can be observed. Because the colloidal particles in a gel cannot move freely with respect to each other, a gel can be considered as a partially or controlled destabilised sol. However, when sols are extensively destabilised a compact aggregate will be formed, which will start to float or sediment depending on the density difference with the dispersion medium. They are no longer considered as colloidal systems. This process is called salting-out when it is induced by the addition of an electrolyte.

Carbomer is a special hydrogel former. The chemical name for carbomer is polyacrylic acid. As such the polymer is poorly soluble in water. However, when monovalent bases (e.g. NaOH) are added, the carboxylic acid groups are deprotonated. By deprotonation the polymer becomes negatively charged and thereby hydrophilised and forms a hydrogel.

Carbomer in combination with monovalent bases, however, is not extremely hydrophilic as it also forms a gel in ethanol.

Addition of salts may have two effects:

1. Divalent ions such as calcium and magnesium form cross-links between deprotonated carboxylic acid groups by which the negative charge is neutralised and the hydrophilisation is counteracted. To prevent this, disodium edetate is added in (aqueous) carbomer gel pH 6.5 NRF (Table 6.16).
2. At high ionic strength the (absolute) zeta-potential of carbomer will decrease, which may result in precipitation and is an example of salting-out of a colloid.

**Table 6.16** Carbomer Gel pH 6.5 [23]

| | |
|---|---|
| Carbomer 35000 | 1 g |
| Disodium edetate | 0.1 g |
| Propylene glycol | 10 g |
| Trometamol | 1 g |
| Water, purified | 87.9 g |
| **Total** | 100 g |

Carbomer, carmellose sodium, hydroxypropylcellulose and other cellulose derivatives are examples of well-known polymers that form hydrogels. An overview of different gel formers can be found in Chap. 7.

#### DLVO-Theory

The destabilisation of colloidal systems can also be described with the DLVO (Deryagin-Landau-Verwey-Overbeek) theory. This theory has been proposed for lyophobic colloidal systems but can also be applied qualitatively to lyophilic colloidal systems. If the potential energy is plotted as a function of the distance of two particles, a curve is obtained as shown in Fig. 6.13.

This curve is established by adding the potential energies resulting from the attraction and repulsion forces to each other (with a negative potential energy meaning attraction and a positive potential energy repulsion). Suppose that two particles approaching each other have too little thermal energy to pass the maximum. In this situation they approach each other to a distance where the secondary minimum is located and thus attract each other. Now two situations can occur:

1. The particles have sufficient thermal energy and they spontaneously diffuse away from each other. This kind of system is a sol. Such a situation is also called deflocculation.
2. The particles have insufficient thermal energy and remain at the same distance. This kind of system is

(continued)

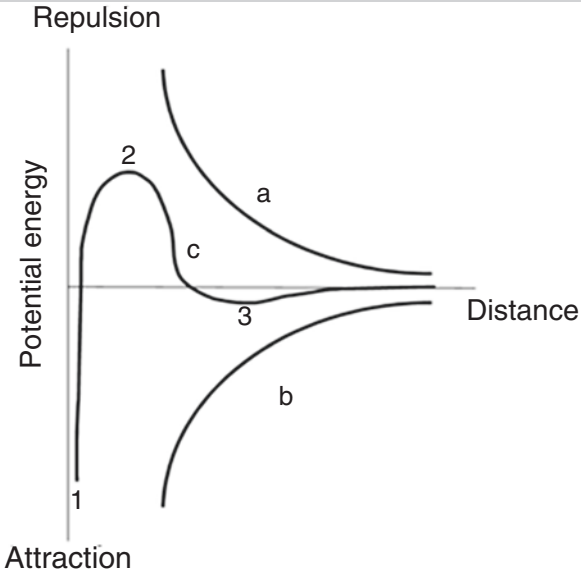


**Fig. 6.13** Potential energy between two particles as a function of their distance. (**a**): repulsion due to zeta-potential; (**b**): attraction as a result of Van der Waals forces; (**c**): net potential energy curve; primary minimum, maximum and secondary minimum are indicated by 1, 2 and 3, respectively. (Source: Recepteerkunde 2009, reprinted by permission of the copyrights holder)

a gel, because external energy (in the form of yield stress) is needed to bring the particles at a greater distance from each other. This is called reversible aggregation or flocculation.

Suppose a lyophobic colloidal system that behaves like a sol and it is destabilised by decreasing the zeta-potential. At a certain point, a sol-gel transition will take place. The reduction of the zeta-potential by the addition of electrolyte is expressed in the energy curve by a lowering of the potential energy in the secondary minimum. The addition of a certain amount of electrolyte will result in such a lowering of the potential energy in the secondary minimum that the particles have insufficient thermal energy to spontaneously diffuse away from each other. However, addition of electrolyte also results in a lowering of the maximum of the potential energy curve. Therefore, when a huge amount of electrolyte is added, it is possible that the particles have sufficient thermal energy to pass the maximum. The particles approach each other now to a distance where the primary minimum is located. Compact aggregates are formed and there is no longer a colloidal system. Because the attractive forces are so high, the original colloidal system cannot be restored by adding external energy. This process is called irreversible aggregation or coagulation.

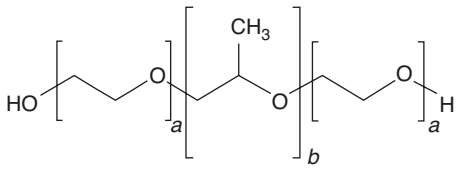


**Fig. 6.14** Chemical structure of poloxamer

A great deal of research has been performed into the use of poloxamers for the controlled release of active substances [24]. Poloxamers, also referred to as Pluronics or Lutrols, consist of triblock copolymers having a central hydrophobic polypropylene oxide block and on both sides a hydrophilic polyethylene oxide block (Fig. 6.14).

By varying the length of the blocks, polymers with different physical properties can be achieved. At low temperatures, these polymers are generally soluble in water and form sols. When the temperature is increased, however, a sol-gel transition may take place. This is caused because, with increasing temperature, the thermal energy and thus also the Brownian motion of the molecules increases. As a result, the hydrogen bonds between the water molecules and the polymer become weaker. The stabilising effects of the solvation shell will therefore decrease and the attraction forces will become dominant, resulting in gelation. The temperature at which the sol-gel transition takes place depends on the composition of the polymer (length of the three blocks) and the concentration. In addition, the sol-gel transition temperature can be influenced by the addition of other substances. For example, the sol-gel transition temperature will be greatly reduced by the addition of a small amount of carmellose sodium. This makes it possible to prepare a solution of poloxamer (and an active substance) that behaves as a sol at a low temperature, for example room temperature or lower, but transforms into a gel at body temperature. Therefore, the solution can be administered at low temperature as a free flowing liquid to a patient after which a gel is formed in situ by the increase of the temperature. Because the gel slowly erodes *in vivo*, the active substance is slowly released. In principle, different routes of administration for this delivery system are possible. Preparations based on poloxamers have been studied for e.g. parenteral (subcutaneous and intramuscular injection), rectal, vaginal, nasal, ocular, and dermal administration. E.g. a poloxamer based morphine-containing hydrogel for the treatment of large-scale skin wounds has been developed [25].

#### 6.5.1.4 Protein Solutions as an Example of Colloidal Systems

Until the beginning of the twenty-first century, in pharmacy, the particles in a colloidal system usually did not consist of active substances but of excipients, such as viscosity enhancers. The number of publications in the pharmaceutical literature on colloidal systems in which the

dispersed particles solely consist of an active substance or consist of carrier systems in which an active substance has been incorporated, however, has increased dramatically in recent years.

Important medicines within this so-called nanotechnology are biologicals, in particular therapeutic proteins. Protein solutions are becoming ever more popular within pharmacy. This is due to the fact that the unravelling of the human genome and the developments in the field of biotechnology offer access to a growing number of proteins that can be used for the treatment of various diseases and disorders. Due to their specific physico-chemical properties, these medicines must be handled in a different way from the classical medicines, which are usually relatively small organic molecules. The specific three-dimensional structure of proteins is essential for their therapeutic action.

The structure of proteins can be described at four levels:

1. Proteins are polymers built up from amino acids. The sequence of the amino acids is called the primary structure.
2. Parts of the protein form specific three-dimensional structures, for example alpha-helices (spiral like structures) and beta-sheets (plate like structures), which are called the secondary structure. The secondary structure is mainly stabilised by hydrogen bonds.
3. The relative orientation of the structural elements with respect to each other is called the tertiary structure. This structure is stabilised by hydrogen bonds, disulfide bonds, electrostatic interactions and hydrophobic interactions.
4. In some cases, several chains of amino acids, also referred to as polypeptide chains, form complexes. For example, insulin forms hexamers in the presence of zinc ions. The relative orientation of the polypeptide chains with respect to each other in such a complex is called the quaternary structure.

Some proteins also contain, besides amino acids, oligo- or polysaccharides. These substances are called glycoproteins.

One of the major problems with proteins is that they are usually not stable. Physical or chemical changes may lead to changes in the three-dimensional structure. This may not only cause a loss of efficacy but it can also have dramatic effects such as the induction of antibodies and severe immune responses.

Some practical advice to improve the stability of proteins is given below.

Protein solutions should not be stored at a pH which is equal to their iso-electric point. At the iso-electric point, the amount of deprotonated carboxylic acid groups and protonated amino groups are equal and thus the net charge of the particle is zero. In that situation the zeta-potential is zero and irreversible aggregation can easily occur. For the same reason, the electrolyte concentration in the solution should not be too high. A very low or high pH is not recommended because the hydrolysis of proteins is both acid and base catalysed. The oxidation of many proteins is catalysed by the divalent metal ions, in particular $Fe^{2+}$ and $Cu^{2+}$. Divalent metal ion catalysed oxidation of proteins can be prevented by the addition of disodium edetate to complex these ions. However, certain divalent metal ions can also act as stabilisers for specific proteins.

Proteins can in particular irreversibly adsorb onto hydrophobic surfaces. Storage of a protein solution in polypropylene vials is therefore not recommended. In addition, protein solutions are sensitive to shear forces. The use of peristaltic pumps during the preparation of formulations should therefore be avoided. By refrigerated storage and transportation, the degradation processes of protein solutions can be slowed down, and thus their shelf life increases. However, freezing must be prevented since ice formation can damage proteins. Alternatively, proteins can be stabilised by freeze-drying them together with sugars. The stabilizing action of sugars as well as the freeze drying procedure is elucidated in Sect. 6.5.1.5.

Summarising, physical degradation of proteins can be caused by:

- Aggregation
- Denaturation
- Adsorption onto surfaces
- Precipitation

Chemical degradation can be caused by:

- Deamidation
- Oxidation
- Hydrolysis
- Racemisation
- Reduction disulfate bridges/disulfide exchange

### 6.5.1.5 Stabilisation of Proteins by Freeze Drying Them Together with Sugars

As mentioned in Sect. 6.5.1.4 proteins in solution are unstable. As most degradation pathways require molecular mobility, it is obvious to bring the protein in the dry state, e.g. by freeze drying. The freeze drying process, however, can be detrimental for the protein and therefore stabilising excipients are required. It is well known that sugars can act as such excipients, with the disaccharides sucrose and trehalose being the two most often applied sugars. Two theories have been described to explain the stabilising effects of sugars: the water replacement theory and the vitrification theory (although refinements to these theories have been proposed [26]). According to the water replacement theory, hydrogen bonds between the protein and water molecules are gradually replaced by hydrogen bonds between the protein and the hydroxyl groups of the sugar molecules by which the original three-dimensional structure of the protein is maintained. According to the vitrification theory, the protein is incorporated in a sugar matrix in which the translational molecular mobility is strongly reduced, thereby also reducing most

degradation processes. For both stabilisation mechanisms, the sugar molecules should form a tight coating around the protein molecules. Therefore, the sugar molecules should be oriented towards the irregular surface of the protein, and thus not towards each other in a specific way like in a crystal lattice. Such random orientation of molecules is also referred to as amorphous. An amorphous material can be either in the rubbery or in the glassy state, i.e. when a glass is heated to above a certain temperature (the glass transition temperature; Tg) it changes into a rubber. While in the rubbery state the molecules have a high translational mobility, in the glassy state it is very low. The rubbery or glassy state, however, are no equilibrium states. Consequently, the material is prone to crystallisation. Due to the high translational mobility of the molecules in a rubber, crystallisation is indeed likely to occur sooner or later. Due to its low translational molecular mobility, this is not the case for a glass during pharmaceutical relevant time scales. Therefore, a material in the glassy state is called kinetically stable. For two reasons, the sugar should be in the glassy state to optimally stabilise the protein. First, according to the vitrification theory, the sugar molecules should have a low translational mobility and therefore in the glassy state. Second, with respect to the water replacement theory, the sugar should also be in the glassy state, as crystallization of the sugar will break up hydrogen bonds between the sugar and the protein. As a consequence, the sugar molecules do not form a tight coating anymore and protein stabilisation is lost. In addition, mechanical forces during the crystallization process may deteriorate the protein structure. As a glass or a rubber do not represent equilibrium states, they are not included in a phase diagram, instead a state diagram applies. Detailed information on such state diagrams can be found in [27].

Basically, freeze drying concerns a process where a solution is cooled until it is fully solidified after which water is removed by reducing the pressure. During cooling, water will crystallize to form ice. However, not all water will crystallize as part of it will form a glass together with the sugar, the protein and possibly other components of the solution, e.g. buffers. The Tg of solutions containing sucrose or trehalose is around −30 °C. Therefore, such solutions should be cooled to below this temperature to avoid crystallization of the sugar. For the same reason, cooling should also be fast. However, extremely fast cooling is not preferred as due to fast nucleation, small ice crystals are formed which may be detrimental for the protein due to a large specific interfacial area where protein molecules can accumulate and unfold. As a rule of thumb, a cooling rate of 1–2 °C/min is advised [28].

Freeze drying of protein/sugar solutions thus implies the removal of these two types of water; ice crystals and water within the glass. This can be achieved by applying a reduced pressure. Water from ice crystals is removed by sublimation while water in the glass is removed by evaporation. Sublimation of water molecules from ice crystals is a relatively fast process as these molecules are directly exposed to the environment. On the other hand, evaporation of water molecules from the glass is a relatively slow process as the diffusion of these molecules through the glass to the environment is hindered by the presence of the sugar, protein and potentially other molecules. Therefore, the removal of water molecules from the solidified solution can be divided into two subsequent stages: primary drying which is the sublimation of water from the ice crystals, and secondary drying which is the removal of water from the glass, although this distinction may in practice not be that strict. Nevertheless, the freeze-drying process should be initiated at a temperature below the Tg of the solution to avoid crystallisation of the sugar. However, primary drying should be performed at a temperature as high as possible to achieve a drying process of acceptable rate. The pressure should not be too low as at very low pressures the capacity of the atmosphere in the freeze dryer for water vapour is very low. As a result, the driving force for sublimation is low and therefore also the sublimation rate. However, the pressure should be below the sublimation curve (Fig. 6.15). This implies that when the temperature is below −30 °C, a pressure lower than around

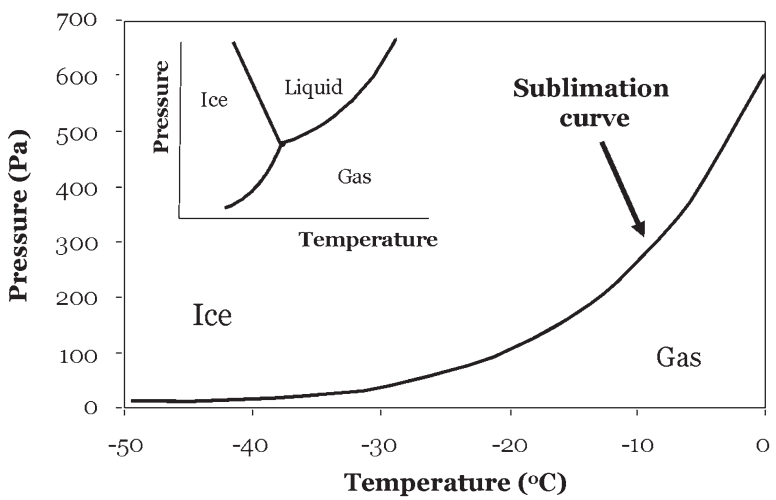


**Fig. 6.15** Phase diagram of water with emphasis on the sublimation curve

40 Pa should be applied. After primary drying is completed, the slower secondary drying process continues. Because during drying the sugar/water ratio of the glass increases, the Tg of the glass gradually increases (the Tg's of pure sucrose and trehalose are around 85 and 120 °C, respectively). This means that the temperature of the sample can be gradually increased as long as the glass transition curve is not passed. Furthermore, the pressure should be decreased as much as possible, as the capacity of atmosphere above the samples for water vapour is not rate limiting anymore but the diffusion of water molecules through the glass. After the freezing-drying process is completed, a highly porous cake with a white appearance and a volume equal to the original solution will be obtained. The high porosity is the result of the fact that during freezing ice crystals are formed everywhere in the solution, and ultimately, the glass will be situated in the interstices between these ice crystals. As a consequence, after the freeze-drying process, a mirror image of the ice crystals will be obtained as the molecules in the glass show a low translational mobility. Such a highly porous cake is very advantageous, as due to its very large specific surface area, reconstitution, e.g. for the preparation of an injection, proceeds fast with no or minimal agitation. In this respect, it should be emphasised that vigorously shaking a protein solution should be avoided as it creates a large liquid-air interface where protein molecules tend to accumulate and unfold. Furthermore, the cake appearance can also indicate whether or not the freeze-drying process was successful. E.g. when the Tg is passed during secondary drying, the glass will be turned into a rubber. Due to the high translational mobility of the molecules in the rubbery state, the porous cake will collapse and the product will appear as a translucent or white crystalline layer at the bottom.

Below, some practical advice for freeze-drying protein/sugar solutions using lab-scale freeze dryers is given. Usually, glass vials are used to charge them with solutions to be freeze-dried, with potentially the possibility to stopper them within the freeze dryer after the drying process depending on the type of freeze dryer used. Freezing of solutions in the glass vials can be achieved by immersing them in liquid nitrogen or by placing them on a pre-cooled shelf of the freeze dryer. When the samples are frozen on the shelf, the shelf temperature is usually set at −50 to −60 °C to obtain a sufficiently fast cooling rate. The drying process can be regarded as an erosion process, meaning that water is gradually removed from the top to the bottom of the vial. Therefore, the height of the solution in the vial should not be too large to avoid a long drying time. Typically, a maximal height of approximately 1 cm is recommended. In that case, using a robust freeze-drying program, the total drying time would be around 2 days, i.e. 1 day of primary drying and 1 day of secondary drying. However, the drying process time can be shortened by adjusting the program settings depending on the formulation characteristics, i.e. type of sugar, sugar/protein ratio, total solid concentration, etcetera and the height of the solution. After the freeze-drying process is completed, the samples can be collected in two ways. First, when the vials can be stoppered in the freeze dryer, the pressure within the freeze dryer is increased to about 0.8 atm with either air or nitrogen via an inlet after which the vials are stoppered. After further increasing the pressure to ambient conditions, the freeze dryer is opened and the vials are removed. Second, the pressure within the freeze dryer is increased to ambient conditions via an inlet. Thereafter, the freeze dryer is opened and the vials are removed and subsequently stoppered. It should be emphasised that, depending on the relative humidity of the environment, stoppering of the vials should be performed immediately after collection. During the freeze-drying process, not all water will be removed, and the final product usually contains a few percent of water. Therefore, the Tg of a freeze dried product will be substantially lower than that of a fully anhydrous product. Furthermore, sugar glasses are highly hygroscopic and can therefore absorb substantial amounts of water from the environment. This can result in a drop in the Tg to below ambient temperature followed by crystallisation. E.g. sucrose and trehalose already crystallise at 25 °C within a few hours when exposed to 30 and 45% relative humidity, respectively [29].

### 6.5.2 Suspensions

Suspensions are regularly used as a dosage form. Examples can be found in oral suspensions (co-trimoxazol suspension), dermatological preparations (zinc oxide or calamine lotions like Zinc oxide lotion (Table 6.17)), parenteral preparations (corticosteroid injections, medroxyprogesterone injection)

**Table 6.17** Zinc oxide cutaneous suspension [30]

| | |
|---|---|
| Zinc oxide | 13.8 g |
| Propylene glycol | 13.8 g |
| Ethanol (90%) BP | 23 g |
| Talc | 13.8 g |
| Water, purified | 50.6 g |
| **Total** | 115 g (=100 mL) |

and a suspension in the form of a solid dispersed in a melted fat base as in the case of suppositories.

As with colloidal systems, suspensions consist of particles dispersed in a liquid. As a result, the physico-chemical properties of suspensions are, in principle, similar to those of colloidal systems. As described in the introduction to Sect. 6.5, the main difference is that the particles in a suspension are larger (>1 μm) than in a colloidal system (1 nm – 1 μm). As a result, the particles settle in a suspension (if the density of the particles is greater than the density of the dispersion medium, which is generally the case) while this is not the case in a colloidal system. Fast sedimentation of particles in a suspension has major drawbacks. Pouring out a partially settled suspension in several portions or at different times leads to too low particle concentrations in the first portions and too high in later portions. In the past this has had fatal consequences, when a 4 months old boy got a threefold dose of spironolactone [31].

#### 6.5.2.1 Sedimentation Behaviour

The sedimentation rate of the particles in a suspension can be calculated using Stokes' law:

$$v = \frac{2r^2\left(\rho_1 - \rho_2\right)\mathrm{g}}{9\eta} \tag{6.15}$$

where $v$ is the sedimentation rate, $r$ the radius of the particle, $\rho_1$, the density of the particle, $\rho_2$, the density of the medium, g the acceleration due to gravity and, $\eta$ the viscosity of the medium.

From this equation, it can be deduced that the rate of sedimentation decreases as the particle size decreases, the difference in the density of the particles and the medium decreases and the viscosity increases.

When applying this equation, the zeta-potential of the particles has also to be taken into account. Again the curve can be used in which the potential energy is plotted as a function of the distance of two particles, as discussed for colloidal systems (Fig. 6.13). The difference with colloidal systems is that the maximum in the curve for suspensions is generally higher. As a result, coagulation or irreversible aggregation almost never occurs in practice. Usual situations are either reversible flocculation or aggregation when the zeta-potential is small or deflocculation when the zeta-potential is large. The sedimentation behaviour is largely affected by whether or not flocculation or aggregation occurs, as described below.

In a deflocculated system the particles will not aggregate and therefore settle separately from each other. Because, according to Stokes' law, the sedimentation rate increases with particle size; large particles will arrive earlier on the bottom of the container than smaller particles. Because the particles do not attract each other, the voids between the large particles are gradually being filled up with the small particles. Therefore, a very compact sediment (cake) is slowly built up from the bottom of the container (Fig. 6.16a).

If the suspension also contains particles smaller than 1 μm (the colloidal fraction of the suspension), these particles will not settle but will continue to be suspended in the liquid so that the liquid above the sediment remains cloudy. The sediment exhibits dilatant flow behaviour.

In a flocculated system, the particles do attract each other and aggregates will be formed anywhere in the fluid.

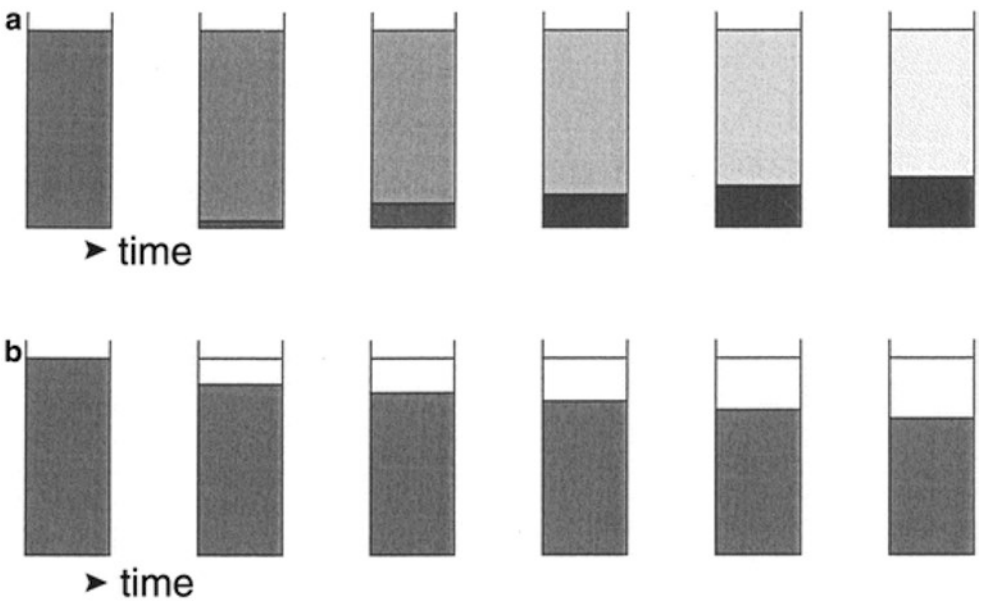


**Fig. 6.16** Schematic representation of deflocculated (**a**) and flocculated (**b**) sedimentation. (Source: Recepteerkunde 2009, ©KNMP)

These aggregates have very open structures. This is because when a particle approaches an aggregate, it will be immobilised by the attractive forces at the outside of the aggregate. This makes diffusion of the particle to possible void spaces in the inside of the aggregate impossible. As everywhere in the fluid aggregates are formed, a large and loose sediment is formed (Fig. 6.16b). This typical way of settling of flocculated suspensions may be called sedimentation, but more precisely subsidencing as is suggested by [5]. Because the aggregates rapidly increase in size the subsidencing also progresses rapidly. Any further settling and compaction of the sediment is unlikely to occur. Once the sedimentation has been completed, a very open sediment is obtained due to the cavities in the aggregates. As a result, at the same volume fraction of particles, the volume of the sediment of a flocculated suspension will be much larger than in a deflocculated suspension. The liquid above the sediment is clear because particles smaller than 1 μm are included in the aggregates. The sediment exhibits plastic flow behaviour.

Both deflocculated and flocculated systems have advantages and disadvantages. In a deflocculated system sedimentation proceeds slowly but once it is completed, it is very difficult to disperse the very compact sediment. The disadvantage of a flocculated system is the high sedimentation rate. The sediment, however, is easy to disperse because it has a very open structure.

In practice, therefore, the objective is to achieve an intermediate form by the addition of a controlled amount of electrolyte or surfactant. When the particles strongly repel each other, an electrolyte can be added. By decreasing the zeta-potential, the repulsive forces will decrease. When the particles attract each other too strongly a surfactant can be added. As the lyophobic part of the surfactant molecule adsorbs onto the surface of lyophobic colloids its lyophilic part will be oriented into the dispersion medium. By steric stabilisation, the attraction forces are decreased. The properties of flocculated and deflocculated suspensions are summarised in Table 6.18.

**Table 6.18** Properties of a deflocculated and a flocculated suspension

| Deflocculated suspension | Flocculated suspension |
|---|---|
| Sediment built up from the bottom | Subsidencing sediment |
| Slow sedimentation | Fast sedimentation |
| Compact sediment | Open sediment |
| Small volume sediment | Large volume sediment |
| Possibly cloudy fluid above sediment | Clear fluid above sediment |
| Flow behaviour sediment: Dilatant | Flow behaviour sediment: Plastic |
| Sediment difficult to disperse | Sediment easy to disperse |

**Table 6.19** Chloramphenicol Oral Suspension 33 mg/mL [32]

| | |
|---|---|
| Chloramphenicol palmitate | 5.75 g |
| Carmellose sodium M | 1 g |
| Polysorbate 80<br>Propylene glycol | 0.5 g<br>2 g |
| Cacao syrup (local standard) | 15 g |
| Syrup B.P. | 15 g |
| Ethanol (90%) BP | 1 mL |
| Vanillin | 0.02 g |
| Water, purified | Ad 100 mL |

A preparation that illustrates the versatile function of a surfactant in a suspension is the Chloramphenicol Oral Suspension 33 mg/mL ex-FNA (Table 6.19).

In this preparation, first a gel of carmellose is prepared in a portion of the cold water using a rotor-stator mixer. Then chloramphenicol palmitate is dissolved in a mixture of hot polysorbate 80 and propylene glycol. Polysorbate 80 thus functions here as a part of the solvent mixture. The hot clear solution is added to the aqueous gel under intensive stirring with the rotor-stator mixer. During this process the chloramphenicol palmitate crystallises to form microcrystals. The size of these microcrystals is not only determined by the intensity of stirring but also by the presence of polysorbate 80, which acts as the surfactant. During storage polysorbate 80 also inhibits any crystal growth. In addition, in this suspension design, polysorbate 80 acts as a deflocculating agent and prevents, by wetting chloramphenicol palmitate, flotation and sticking of the active substance to the bottleneck.

### 6.5.2.2 Influencing Sedimentation Behaviour

Ideally the sedimentation rate of pharmaceutical preparations is as low as possible. On the basis of Stokes' law, it is clear which variables can be varied to accomplish this. The gravitational acceleration cannot be reduced, nor the density of the particles. However, the particle size, the density and the viscosity of the dispersion medium may be adjusted.

A first method to reduce the sedimentation rate is particle size reduction. This can be accomplished by milling. On a small scale, this is possible by using a mortar and pestle, if electrostatic charging and agglomerating can be managed. On a larger scale, high-tech equipment has been developed for this (see Chap. 28). The precipitation method is another

method to reduce the particle size. It is a useful method when no milling equipment is available or when electrostatic charging will be a problem. In the precipitation method, at first the active substance is dissolved and then its solution is brought into the supersaturated state. As a result, the dissolved substance will precipitate. The size of the precipitated particles will decrease when the degree of supersaturation increases or when the rate at which supersaturation has been achieved increases, or both. Supersaturation can be created in different ways. The solubility of many active substances is pH dependent. As described in Sect. 6.1.1, active substances with for example one or more acid groups are generally poorly soluble at a low pH but their solubility will be better at a high pH. This property can be used by preparing a solution of the active substance at high pH and then adding an acid to the solution by which it becomes supersaturated. Obviously, the opposite strategy can be used for active substances containing one or more amine groups: first the active substance is dissolved at low pH and then a base is added. Thus in both cases a solution of the active substance in ionised form is prepared. Then, the active substance in the form of its free acid or base (non-ionised form) is formed by changing the pH and supersaturation is achieved.

In another precipitation method, two different liquids are used that are miscible with each other, but in which the solubility of the active substance substantially varies. Firstly, the active substance is dissolved in the liquid in which the active substance dissolves well. Subsequently, supersaturation is achieved by adding the second liquid to the solution in which the active substance is poorly soluble.

In a third precipitation method, the fact that most substances are more soluble at a high than at a low temperature is to be made use of. A saturated solution is made at a high temperature after which it is cooled until supersaturation is achieved. This last method is the least suitable because in practice it is often difficult to cool rapidly. The chloramphenicol palmitate suspension as described in Table 6.19 is an example of a suspension prepared by precipitation. Chloramphenicol palmitate precipitates when the solution in a hot mixture of polysorbate 80 and propylene glycol is mixed with the cold aqueous gel. By vigorous stirring during the final step, small particles are obtained.

A second method to reduce the sedimentation rate is to increase the density of the dispersion medium, for example, by the addition of syrups or a solution of sorbitol.

A third method to reduce the sedimentation rate is to increase the viscosity of the dispersion medium, which is almost always achieved by the application of polymers. But it must be kept in mind that accurate dosing by the patient of a specific amount of the suspension is more difficult when the viscosity increases. To avoid inadequate dosing the preparation can be delivered together with a dosing/measuring syringe instead of a measuring cup or spoon.

> Most corticosteroids nasal sprays (licensed preparations) are suspensions in which croscarmellose sodium is used as a viscosity enhancer. The inhalation liquids for nebulisation with the same type of active substances however only contain polysorbate and sorbitanlaureate to stabilise the suspension. For atomisation in jet nebulisers, the liquid should not be too viscous, in order to prevent clogging of the nebuliser.

#### 6.5.2.3 Particle Size Stability

The size of the particles in suspensions is not stable. During time, the particle size will increase by temperature fluctuations and by the so-called Ostwald ripening.

Because most substances are more soluble at a high than at a low temperature, small particles present in a suspension will dissolve when the temperature increases. When subsequently the temperature decreases again the solution will become supersaturated. The undissolved larger particles will act as nuclei for precipitation and grow.

Ostwald ripening is happening because the solubility of a substance in the near vicinity of small particles is greater than in the near vicinity of large particles. The relationship between the solubility and the size of the particles is given by the Ostwald-Freundlich equation:

$$C_{s,curved} = C_{s,flat} \cdot \exp\left(\frac{2 \cdot \gamma \cdot \mathrm{M}}{\mathrm{R} \cdot \mathrm{T} \cdot \rho \cdot r}\right) \qquad (6.16)$$

where $C_{s,curved}$ and $C_{s,flat}$ is the solubility near a small particle and an infinitely large particle, respectively, γ the interfacial tension between the particles and the dissolution medium, M the molecular weight of the substance, R the gas constant, T is the temperature (in K), ρ the density of the substance, and *r* the radius of curvature of the particles.

Due to the differences in solubility, differences in concentration in the dispersion medium arise and the dissolved molecules diffuse from the small particles to the large particles. As a consequence, the dispersion medium around the large particles becomes supersaturated and the dissolved molecules will precipitate onto these particles. Because the dissolved molecules around the small particles are diffusing away, the dispersion medium is no longer saturated and molecules from the small particles will dissolve. Thus, small particles become smaller and eventually disappear and large particles are getting bigger. A second effect is that irregularly shaped particles in suspensions become spherical.

From the above it can be concluded that by both temperature fluctuations and Ostwald ripening the particle size will increase faster when the particle size distribution is larger. It is therefore desirable that the particle size distribution in a suspension is as small as possible.

#### 6.5.2.4 Polymorphism, Pseudo-Polymorphism, Glassy State

Solids may exhibit polymorphism, which means that they can exist in different crystal modifications which differ in their physical properties [3, 20, 33, 34]. No less than eleven different crystal modifications of phenobarbital are known. It depends on the physical conditions such as temperature which crystal modification is stable. The other modifications are metastable. For substances that exhibit polymorphism, the solubility of the metastable form is higher than that of the stable form. This means that if the metastable modification is in equilibrium with the solution (i.e. saturated for the metastable form), the solution is supersaturated for the stable modification. In other words, the stable form can grow over time at the expense of the metastable form. Some fats also exhibit polymorphism [35]. The preparation method and the storage temperature will influence for example the melting behaviour of suppositories, and thereby probably also the release rate.

Besides polymorphism, pseudo-polymorphism also exists. In pseudo-polymorphism, a substance is found in crystal modifications whose hydration or solvation state differs. There may therefore be crystal lattices in which more or less water or other solvent molecules are included. Similar to polymorphism, the substance in the form of one pseudo-polymorph is more soluble than in the other. Erythromycin is an example of a substance which exhibits pseudo-polymorphism. It exists in an anhydrous form and as a hydrate. In Erythromycin Eye Ointment FNA (Table 6.20) the anhydrate is used, because this form dissolves faster in the fatty ointment base.

Besides in the crystalline form, a substance may also exist in the glassy state. In the glassy state, the molecules are not oriented in a specific manner towards each other as they are in a crystal lattice, but randomly (amorphous). The aqueous solubility and thereby also the dissolution rate of a substance in the glassy state is better and higher than in crystalline form [37, 38]. Therefore, the bioavailability of lipophilic active substances after oral administration, which is limited due to their slow dissolution, can be improved by converting them into the glassy state.

**Table 6.20** Erythromycin Eye Ointment 0.5% [36]

| | |
|---|---|
| Erythromycin, anhydrous | 0.5 g |
| Cetostearyl alcohol | 2.49 g |
| Paraffin, liquid | 39.8 g |
| Paraffin, white soft | 51.2 g |
| Wool fat | 5.97 g |
| **Total** | 100 g |

### 6.5.3 Emulsions

Emulsions can often be found as dermatological preparations, and sometimes as injections and oral preparations. They make combinations of immiscible liquids possible, typically of fatty/lipophilic components and water. The fat/lipophilic ingredients can act as an active substance, for example, in creams to keep the skin hydrated. They can also serve as a solvent for other substances such as diazepam in Diazemuls® injection or for fat soluble vitamins in parenteral nutrition (see Chap. 21).

The physico-chemical properties of emulsions are basically similar to those of suspensions, with the essential difference that in emulsions the dispersed phase consists of a liquid instead of a solid. This difference has important consequences. Similar to a suspension, an emulsion exhibits a large interface between the dispersed phase and the dispersing medium. It requires energy input to create an interface. This implies that energy is liberated when the interfacial area is reduced. Since this is thermodynamically advantageous, the system will attempt to minimize the interfacial area. Because liquids are 'deformable' the dispersed drops in emulsions will therefore have a strong tendency to coalesce. This coalescence may eventually result in 'breaking' of the emulsion. When breaking occurs, the two phases will be present as two liquid layers. The stability of an emulsion can be improved by adding surfactants. The driving force of the disperse drops to coalesce becomes smaller because less energy will be liberated.

Specifically applied in emulsions, surfactants are referred to as emulsifying agents. According to the rule of Bancroft, the HLB of the emulsifying agent determines which type of emulsion is obtained, water in oil (w/o) or oil in water (o/w). According to this rule, the phase in which the emulsifying agent dissolves better will be the dispersion medium. An emulsifying agent with an HLB < 7 thus gives a w/o emulsion, and an emulsifying agent with an HLB > 7 an o/w emulsion. This can be explained as follows. When an emulsifying agent is more soluble in oil than in water (HLB < 7), the lipophilic part of the molecule occupies a larger volume than the hydrophilic part. Since at the outside of a drop there is more space than at the inside, it is sterically more favourable when the (larger) lipophilic part is directed towards the outside and (smaller) hydrophilic part to the inside of the drop. In this case a w/o emulsion will be obtained. With an emulsifying agent having an HLB > 7, the hydrophilic part occupies a larger volume and an o/w emulsion is obtained.

The stability of an emulsion increases when the emulsifying agent molecules form a more compact layer at the inter-

face. A very compact layer can be achieved by making use of two different emulsifying agents, one of the o/w-type and one of the w/o-type. One emulsifying agent occupies the cavities in the interface that the other emulsifying agent cannot fill. Such combinations of emulsifying agents are referred to as mixed layer emulsifying agents or emulsifying agent complexes. The HLB of two emulsifying agents can be calculated by multiplying the weight fraction of each emulsifying agent by its HLB value and adding them together. Thus the HLB of a mixture of two emulsifying agents A and B can be calculated as follows:

$$\mathrm{HLB}_{\mathrm{mixture}} = f_A \times \mathrm{HLB}_A + f_B \times \mathrm{HLB}_B \qquad (6.17)$$

where $f_A$ and $f_B$ are the weight fractions of the emulsifying agents A and B, respectively; $\mathrm{HLB_A}$ and $\mathrm{HLB_B}$ are the HLB values of the emulsifying agents A and B, respectively.

When two emulsifying agents are combined, one with an HLB of 4.7 and the other with an HLB of 10.3, in a weight ratio of 40/60, the HLB of the whole will become 8.1 ($\mathrm{HLB_{mixture}} = 0.40 \times 4.7 + 0.60 \times 10.3 = 8.1$). According to the rule of Bancroft, with this combination an o/w emulsion will be obtained.

The Bancroft rule should be interpreted as a general rule. In practice, however, there are many exceptions. In addition, the type of emulsion that is formed will also depend on the volume ratio of the two phases, the method of preparation, the electrolyte concentration, etc.

## 6.6 Osmosis

Osmosis is the transport of water through a semi-permeable membrane as a result of a difference in the concentration of solutes on either side of the membrane. A semi-permeable membrane is only permeable to water; dissolved dissociated or undissociated substances cannot pass through it. Living cells are provided with a membrane through which water transport can take place. This must be taken into consideration when the envisaged dosage form for an active substance is a solution. If a solution is administered to a patient, water transport across the cell membrane of the cells in the near vicinity of the site of administration should be avoided as much as possible. This is because extensive water transport across cell membranes may lead to irritation and cell damage. The risks for this are, in particular, present with parenteral preparations and irrigations, but also in the case of preparations for eye, middle ear, and nose. In this section the water transport across membranes is dealt with in more detail. Methods are given to prevent net water transport across cell membranes after the administration of solutions.

### 6.6.1 Osmotic Pressure

When an aqueous solution and pure water in two different compartments are separated from each other by a semipermeable membrane, a spontaneous transport of water molecules across the membrane into the solution will take place. This spontaneous transport is caused by the attractive forces between the solute molecules and the water molecules. As a result, water molecules are forced through the pores of the semi-permeable membrane. By this water transport, the liquid level in the compartment of the solution will rise, while it will fall in the compartment of pure water. The rise of the liquid level in the solution compartment will, however, not continue indefinitely because the difference of the fluid levels creates a hydrostatic pressure difference that leads to a driving force for water transport in the opposite direction, i.e. across the membrane to the compartment containing pure water. At a given moment the driving forces for transport of water from the solution to pure water and from pure water to the solution are equal, and the liquid levels in the two compartments do not change anymore.

The process of transport of water through a semipermeable membrane due to a concentration difference is called osmosis and the final pressure difference between both sides of the membrane is called osmotic pressure. Basically, the osmotic pressure should be expressed in Pascal, but in practice the words osmolality or osmolarity are used to indicate osmotic pressure, both with osmole as unit.

Osmotic pressure is a colligative property. A colligative property solely depends on the concentration of the dissolved molecules or ions and is independent of the nature of the solute. Freezing point depression is also a colligative property and can be indirectly used to determine osmotic pressure. In practice, it is much more difficult to determine the osmotic pressure of a solution than to measure its freezing point depression. The freezing point of a solution can be measured and the osmotic pressure can be calculated from it.

The molar freezing point depression of water is 1.86 °C. Blood freezes on average at −0.54 °C. Plasma and tear fluid have the same freezing point. The osmolarity of blood, plasma and tear fluid is thus equal to:

$$\frac{0.54}{1.86} = 0.290 \text{ osmole} \left(290 \text{ mosmole}\right) \qquad (6.18)$$

In literature, also a freezing point depression of blood of 0.52 °C or 0.56 °C has been reported. Based on these values, it can be calculated that the osmolarity of blood, plasma and tears will be 280 or 300 mosmol, respectively.

The osmotic pressure has the unit of Pascal ($N/m^2$) but in clinical practice this unit is not used as such. Measurements and calculations are performed with concentrations expressed as osmols or milliosmols, which are abbreviated as "osmol" and "mosmol", respectively.

In the clinical setting 1 osmole means a concentration of 1 mol of a non-dissociable substance per kg of solvent or per litre of solution. If this concentration is expressed as mol per kg of solvent (molality) it is called osmolality. If the concentration is expressed as mol per litre of solution (molarity) it is called osmolarity.

Since the concentration of, in particular, parenteral preparations is more commonly expressed as mole per litre solution than mole per kilo solvent, usually the osmolarity of injections and infusions is given. The difference between osmolarity and osmolality in dilute aqueous solutions is usually not significant, because the density of water is 1 kg per litre, and the volume fraction of the solute is usually negligible. In highly concentrated solutions, however, there is a clear difference between osmolality and osmolarity. But highly concentrated solutions are clinically hardly relevant.

### 6.6.2 Iso-osmotic and Isotonic

If two aqueous solutions with different concentrations of dissolved substances are separated from each other by a semipermeable membrane, there is a net transport of water from the solution with lower concentration molecules or ions to the solution with the higher concentration of molecules or ions. The solution with the lower concentration is called hypo-osmotic, while the solution with the higher concentration is hyper-osmotic in relation to the other. When the concentrations of dissolved substances in the solutions on either side of the semi-permeable membrane are equal to each other, there will be no net transport of water across the membrane. In such a case we call the two solutions iso-osmotic.

When cells are brought into contact with a solution and a cell membrane would behave as a semi-permeable membrane, the solution within the cell will try to become iso-osmotic to that outside of the cell. This water transport will cause damage to the cell. It must therefore be ensured that there is little or no net water transport from the solution into the cell or vice versa. If an active substance is administered in a low concentration solution, the osmotic value will be less than that of the blood. To achieve an iso-osmotic concentration, NaCl or glucose can be added. When the active substance has to be administered in a high, hyper-osmotic concentration, consideration should be given to reducing the concentration by dilution. If this is not possible, then, under certain conditions a solution with a hyperosmotic concentration can be administered, preferably into a vein with a good flow so dilution will take place swiftly (see Chap. 21).

The cytoplasmic membrane of an erythrocyte or a corneal epithelial cell and other physiological membranes, however, do not always behave as a semi-permeable membrane. Cell membranes are to some extent also permeable to some molecules other than water. Some molecules or ions, such as urea, ethanol, and ammonium salts, are able to pass through a cell membrane at a relatively high rate. If an erythrocyte for example is placed in an iso-osmotic solution of ammonium chloride, the transport of ammonium chloride across the cell membrane occurs quickly until its concentration in- and outside the cell is equal. For this reason, solutions containing that type of molecules may be iso-osmotic with the cell content, but nevertheless show water transport when brought into contact with cells. On the other hand, a net water transport does not necessarily occur when the cell content is not iso-osmotic with the environment. If certain ions or molecules do not stay at one side of the (cell) membrane, their contribution to the pressure difference will not be the same as in the case of a 'perfect semi-permeable' membrane. If net water transport occurs from the environment to the cell the solution outside the cell is called hypotonic. And vice versa, when a net water transport occurs from the cell to the environment the solution outside the cell is hypertonic. If no net water transport takes place, both solutions are called isotonic.

Iso-osmotic is a physico-chemical concept and only depends on the concentration of dissolved molecules and ions. Isotonicity is the concept that takes into account, as well, the properties of the biological membrane in relation to the type of dissolved substances. Thus isotonicity should be interpreted as a physiological concept. Therefore, in this context it is better to speak of selectively permeable instead of semi-permeable. For most applications or routes of administration (bio-membranes), the number of substances for which there is a difference between iso-osmotic and isotonic is limited. For this reason, terms such as hypertonic and hypotonic are commonly used while actually hyper- or hypo-osmotic, respectively, are meant.

### 6.6.3 Non-ideal Solutions

On the basis of paragraph Sect. 6.6.1, it would be expected to calculate as follows: if the composition of the fluid is given in millimoles, determine whether or not the dissolved sub-

stances dissociate and if so how many ions are formed. For example, a solution of 1 mmol of glucose in 1 L of water yields 1 mosmol, but a solution of 1 mmol of NaCl or 1 mmol $CaCl_2$ in 1 L of water yields 2 or 3 mosmol, respectively. After summing the contributions of all components it can be verified whether or not the total strength is approximately 300 mosmol.

This way of calculating only applies to so-called ideal solutions. In an ideal solution of substance A in a solvent B, the interactions between A and B, A and A, and B and B are equal. In practice, however, non-ideal solutions are more common. As a result, at an equal concentration of molecules or ions, the osmolarity of a non-ideal solution can be different from that of an ideal solution. This difference in osmolarity is expressed by a correction factor f. This correction factor is specific to each substance and in dilute solutions independent of the concentration. At very high, clinically irrelevant, concentrations, the correction factor may change due to association of the dissolved components. The osmolarity of a solution can be calculated as follows:

$$\frac{G}{\mathrm{M}} \times \mathrm{f} \tag{6.19}$$

where $G$ is the concentration of the solute in grams per litre and M the molecular weight of the dissolved substance. The f-values are mentioned in Table 6.21 as group averages.

**Table 6.21** Dissociation types and f-values of various substances

| Dissociation type | Molar freeze point depression (in °C/mol) | f-value | Examples |
|---|---|---|---|
| Non-dissociating | 1.9 | 1.0 | Glycerol |
| | | | Glucose |
| | | | Sorbitol |
| | | | Urea |
| Weak electrolytes | 2.0 | 1.05 | Alkaloids |
| | | | Bases |
| | | | Boric acid |
| Di-divalent electrolytes | 2.0 | 1.05 | Magnesium sulfate |
| | | | Zinc sulfate |
| Mono-monovalent electrolytes | 3.4 | 1.8 | Sodium chloride |
| | | | Silver nitrate |
| | | | Phenobarbital sodium |
| Mono-divalent electrolytes | 4.3 | 2.3 | Sodium sulfate |
| Di-monovalent electrolytes | 4.8 | 2.5 | Zinc chloride |
| Mono-trivalente electrolytes | 5.2 | 2.7 | Sodium citrate |
| Tri-monovalent electrolytes | 6.0 | 3.2 | Aluminium chloride |
| Tetraborate | 7.6 | 4.0 | Borax |

In a preparation, the active substance (if present in dissolved form) and the excipients (e.g. buffering agents, preservatives, antioxidants, and disodium edetate) all contribute to the osmotic value of a preparation.

If 290 mosmol is taken as the iso-osmotic value, a solution of the substances A, B, … is iso-osmotic with blood under the following condition:

$$\frac{G_A}{\mathrm{M_A}} \times \mathrm{f}_A + \frac{G_B}{\mathrm{M_B}} \times \mathrm{f}_B + \ldots + \frac{G_H}{\mathrm{M_H}} \times \mathrm{f}_H = 0.290 \tag{6.20}$$

where $G_A$, $G_B$, .... are the concentrations of the solutes in grams per litre, $M_A$, $M_B$, .... the molecular weights of the dissolved substances, and $f_A$, $f_B$, ..... the correction factors associated with dissociation of the solutes. To make a solution iso-osmotic, an excipient can be added. In (6.20) this excipient is indicated with an H.

### 6.6.4 Calculation of Osmotic Value

If an active substance or excipient substantially contributes to the osmotic value, it may be necessary to calculate this contribution. This can be done in practice by three different calculation methods. The choice of the method depends on which physical characteristics of the substances are available.

**Method 1:**
If the molecular weight and the dissociation type (f-value, Table 6.21) are known, the osmotic value can be calculated using a part of Eq. 6.16, namely:

$$\frac{G_A}{\mathrm{M_A}} \times \mathrm{f}_A$$

**Method 1 Is Exemplified by the Calculation of the Osmotic Value of Betaxolol 1% Eye Drops**
Suppose that no iso-osmotic concentration of the substance (betaxolol) is known. The osmotic value can then be calculated using the molecular weight and the type of dissociation. Betaxolol hydrochloride has a molecular weight of 343.9. The degree of dissociation and the $pK_a$ cannot be easily found in the literature. Based on its chemical structure, however, it can be concluded that this is a salt of a secondary amine. The so-called f-value of this type of molecule is 1.8. The contribution of betaxolol hydrochloride is then (G/M) × f = (10/343.9) × 1.8 = 52 mosmol per litre. Betaxolol hydrochloride therefore contributes for (52/290) × 100% = 18% of the iso-osmosis. As a consequence, excipients should contribute for the remaining 82% to make an isotonic solution.

**Method 2:**
If the iso-osmotic concentration of a substance is known, the osmotic value can be easily calculated. In Martindale, these values are often specified in the description of substances [11]. In the Merck Index and in the Handbook of Injectable Drugs, the iso-osmotic concentrations for a large number of substances are listed in a table [6, 39].

Method 2 is exemplified by the calculation of the osmotic value of Pilocarpine Eye Drops 2% FNA (Table 6.22).

**Table 6.22** Pilocarpine eye drops solution 2% [40]

| | |
|---|---|
| Pilocarpine hydrochloride | 2 g |
| Benzalkonium chloride | 0.01 g |
| Borax | 0.375 g |
| Boric acid | 0.7 g |
| Disodium edetate | 0.1 g |
| Water, purified | Ad 100 mL |

A 4.1% w/v solution of pilocarpine hydrochloride is iso-osmotic (293 mosmol as determined by measuring its freezing point depression). A 2% w/v pilocarpine hydrochloride thus contributes to about 50% of the iso-osmotic value (the osmolarity as determined at pH 6 is 147 mosmol). The other 50% should be provided by the excipients. Iso-osmotic stock solutions for the preparation of eye drops could be very useful for the purpose of easy calculation. The stock solution Boric acid-benzalkonium solution FNA is nearly iso-osmotic (boric acid is iso-osmotic at a concentration of 19 mg/mL). The contribution to the osmotic value of the benzalkonium chloride 100 mg/L is too small and can be neglected in the calculations. Without adjusting the pH, 50% v/v of Boric acid-benzalkonium solution would be needed. But because the stability of pilocarpine is optimal at pH 6.5, the pH is adjusted with 3.75 mg/mL borax to 6.5. As 3.75 mg/mL borax contributes 15% to the osmotic value, 35% is left for the Boric acid-benzalkonium solution.

**Method 3:**
If no iso-osmotic concentration is known, it can be calculated with the aid of the sodium chloride equivalent, also known as tonicic equivalent or E-value. The sodium chloride equivalent is defined as:

$$E = \frac{\text{freeze point depression per gram of compound A}}{\text{freeze point depression per gram NaCl}} \qquad (6.21)$$

For each substance, the E value can be calculated if the dissociation type of that substance is known [41]. In practice, tables are used in which the E-values for a large number of substances are listed. These tables can be found in references [6, 39]. The calculation using E-values proceeds as follows. The concentration of each of the solutes, expressed as a percentage, is multiplied with the corresponding E value. The product thus obtained, gives for each substance in the amount used the amount of NaCl which corresponds to the osmotic value. The sum of the products of all individual components represents the strength in NaCl equivalents of the total solution. Because the overall strength of the solution should be 0.9 NaCl equivalents, the relative contribution of the active substance to the osmotic value is known. The amount of iso-osmotic stock solution to be added can be calculated as described above.

Method 3 is exemplified by the calculation of E value and NaCl equivalents for a thiamine-injection 25 mg/mL (Table 6.23):

**Table 6.23** Calculation of E value an NaCl equivalents for a thiamine injection solution

| | Per 100 mL | E | NaCl-eq |
|---|---|---|---|
| Thiamine-hydrochloride | 2.5 g | 0.21 | 0.525 |
| Disodium edetate | 0.01 g | 0.20 | 0.002 |

This gives a total of at E = 0.527 sodium chloride equivalents. 0.9 g of sodium chloride per 100 mL is iso-osmotic. This means that almost 0.373 g per 100 mL should be added.

### 6.6.5 Importance of Osmotic Value in Dosage Forms

With the aid of the three methods described in Sect. 6.6.4 it can be calculated whether or not a pharmaceutical preparation is iso-osmotic. Hypo-osmolarity can usually be avoided as it can be compensated by the addition of excipients in calculated quantities. Hyper-osmolarity may be inevitable due to dosage reasons, for example when a high dose of an active substance has to be administered in a small volume. The extent to which hyper-osmolarity is tolerated will depend on the route of administration and administration site. The tolerance for parenteral administration, for example, increases in the order: subcutaneous < intramuscular < intravenous. This has to do with the fact that of these three routes, the intravenously administered dose spreads most rapidly, and thus dilutes most rapidly in the body and

the subcutaneously administered dose most slowly. For the same reason, the tolerance is greater when the solution is injected into a large blood vessel than in a small blood vessel. The tolerance is also determined by the volume infused. In Chap. 21, more information is given about the relative importance of iso-osmosis for different types of parenterals. The osmolarity of an eye wash should be much more accurate than that of an eye drop. This is because the eye drops, which are administered in small volumes, are rapidly diluted by tear fluid. But with an eye wash, all protective tear fluid is washed out of the eye. The osmotic value of nasal drops and aqueous ear drops should not deviate too far from the physiological value. Limits of osmolarity for various dosage forms are discussed in the chapters on dosage forms of this book. These limits, however, are never absolute. Adverse effects are more readily accepted for serious indications.

**Questions**

1. What is the difference between Newtonian and non-Newtonian flow behavior?
2. What would be the number of independent variables of a condensed system of two components and two phases?
3. Why should a lyophobic colloidal particle be electrostatically charged to be stable and why is that not always necessary for a lyophilic colloidal particle?
4. Why can proteins be stabilised by drying them in the presence of sugars?
5. Describe the differences between a flocculated and a deflocculated suspension.
6. Why is it more important that an injectable is iso-osmotic when it is injected intramuscularly than when it is injected intravenously?

## References


1. Anonymous (2014) European Pharmacopoeia, 8th edn. Council of Europe, Strasbourg
2. Florence AT, David AD (2011) Physicochemical principles of pharmacy, 5th edn. Pharmaceutical Press, London
3. Newton DW (2009) Drug incompatibility chemistry. Am J Health-Syst Pharm 66:348–357
4. Tetracyclinehydrochloridemondspoeling 5% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
5. Sinko PJ (2006) Martin's physical pharmacy and pharmaceutical sciences, 5th edn. Lippincott, Williams & Wilkins, Philadelphia etc
6. O'Neil MJ red (2013) The Merck Index, 15th edn. Royal Society of Chemistry, New York
7. Furosemide 2 mg/ml FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
8. Tetracyclinehydrochloridecrème 3% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
9. Tetracycline-oogdruppels 0,5% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
10. Zuidema J (1985) Ethylenediamine, profile of a sensitizing excipient. Pharm Weekbl Sci 7:134–140
11. Brayfield A (ed) (2014) Martindale, the complete drug reference, 38th edn. London, Pharmaceutical Press
12. Lorazepamdrank 1mg/mL FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
13. Paracetamol 500mg/5mL Oral Solution in Electronic Medicines Compendium. https://www.medicines.org.uk/emc/product/3233#EXCIPIENTS
14. Phenobarbital Sodium 60mg/mL Injection in Electronic Medicines Compendium. https://www.medicines.org.uk/emc/product/3606#EXCIPIENTS
15. Loftsson T, Jarho P, Masson M, Jarvinen T (2005) Cyclodextrins in drug delivery. Expert Opin Drug Deliv 2:335–351
16. Zhao L, Li P, Yalkowsky SH (1999) Solubilization of fluasteron. J Pharm Sci 88:967–969
17. Triamcinolone hexacetonide 20 mg/ml suspension for injection Injection in Electronic Medicines Compendium. https://www.medicines.org.uk/emc/product/5408
18. Zure oordruppels met hydrocortison 1% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
19. Magnesium Citrate Oral Solution, United State Pharmacopeia 44 – NF 39 (2021)
20. Ubbelohde L (1937) The principle of the suspended level: applications to the measurement of viscosity and other properties of liquids. Ind Eng Chem Anal Ed 9(2):85–90
21. Aulton ME (2013) Aulton's pharmaceutics – the design and manufacture of medicines, 4th edn. Churchill, Livingstone, Elsevier, Edinburgh etc
22. Colecalciferoldrank 50.000 IE/ml, waterig FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
23. Carbomeerwatergel 1% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
24. Dumortier G, Grossiord JL, Agnely F, Chaumeil JC (2006) A review of Poloxamer 407 pharmaceutical and pharmacological characteristics. Pharm Res 23:2709–2728
25. Heilmann S, Küchler S, Wischke C, Lendlein A, Steind C, Schäfer-Kortinga M (2013) A thermosensitive morphine-containing hydrogel for the treatment of large-scale skin wounds. Int J Pharm 444:96–102
26. Mensink MA, Frijlink HW, Voort Maarschalk van der K, Hinrichs WLJ (2007) How sugars protect proteins in the solid state and during drying (review): mechanisms of stabilization in relation to stress conditions, Eur J Pharm Biopharm 114:288–295
27. Amorij JP, Huckriede A, Wilschut J, Frijlink HW, Hinrichs WLJ (2008) Development of stable influenza vaccine powder formulations: challenges and possibilities. Pharm Res 24:1256–1273
28. Tang X, Pikal MJ (2004) Design of Freeze-Drying Processes for pharmaceuticals: practical advice. Pharm Res 2:191–200
29. Drooge van DJ, Hinrichs WLJ, Frijlink HW (2004) Incorporation of lipophilic drugs in sugar glasses by lyophilization using a mixture of water and tertiary butyl alcohol as solvent. J Pharm Sci 93:713–725

30. ZInkoxideschudsel, alcoholisch FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
31. Croonen H, Stolker E (2006) Suspensies: "Weet waar je aan begint". Pharm Weekbl 139(15):512–514
32. Chlooramfenicolsuspensie 33 mg/ml ex-FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
33. Thoma VK, Serno P (1984) Physikalische Instabilität von Arzneimitteln als Folge von polymorphen Veränderungen der Kristalstruktur. Dtsch Apoth Ztg 124:2162–2170
34. Borka L (1991) Review on crystal polymorphism of substances in the European pharmacopoeia. Pharm Acta Helv 66:16–22
35. Himawan C, Starov VM, Stapley AGF (2006) Thermodynamic and kinetic aspects of fat crystallization. Adv Colloid Interf Sci 122:3–33
36. Erytromycine-oogzalf 0,5% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
37. Hancock BC, Parks M (2000) What is the true solubility advantage for amorphous pharmaceuticals? Pharm Res 17(4):397–404
38. Chawla G, Bansal AK (2007) A comparative assessment of solubility advantage from glassy and crystalline forms of a water-insoluble drug. Eur J Pharm Sci 32(1):45–57
39. Trissel LA (2011) Handbook on injectable drugs, 16th edn. American Society of Health-System Pharmacists®, Bethesda
40. Pilocarpine-oogdruppels 1%;2%;3%;4% FNA. Jaar 2022. Formularium der Nederlandse Apothekers. Den Haag: Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie (KNMP)
41. Sinko PJ (2006) Martin's physical pharmacy and pharmaceutical sciences, 5th edn. Williams & Wilkins, Lippincott, pp 221–225