A single aqueous reference equilibrium constant for amide synthesis–hydrolysis

Rein V. Ulijn,*[a] Barry D. Moore,[b] Anja E. M. Janssen[c] and Peter J. Halling[b]

[a] *Department of Chemistry, The Edinburgh Centre for Protein Technology, The University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, Scotland, UK EH9 3JJ. E-mail: Rein.Ulijn@ed.ac.uk; Fax: (+)0131 650 6452; Tel: 0131 650 4792*
[b] *Department of Pure and Applied Chemistry, Thomas Graham Building, 295 Cathedral Street, Glasgow, Scotland, UK G1 1XL*
[c] *Department of Food Technology and Nutritional Sciences, Food and Bioprocess Engineering Group, Wageningen University, PO Box 8129, 6700 EV Wageningen, The Netherlands*

Received (in Cambridge, UK) 5th September 2001, Accepted 22nd February 2002
First published as an Advance Article on the web 22nd March 2002

Experimentally measured equilibrium constants at a given pH in part reflect the contributions of ionisation of acidic and basic groups present. These contributions can be isolated from the equilibrium constant by expressing all reactant concentrations in terms of the uncharged forms only. This article presents methods to calculate uncharged reference equilibrium constants for amide synthesis/hydrolysis reactions. For zwitterions in particular these methods are not always straightforward. It is explained how (microscopic) pK_a values can be estimated where experimental values are not available. A large number of equilibrium data are analysed for hydrolysis or synthesis of protected and unprotected di- and tri-peptides, beta-lactam antibiotics, and acyl acids and amides. This reveals just how similar the reference equilibrium constants are when ionisation is properly accounted for ($K^0_{ref} = 10^{3.6}$ M^{-1}) regardless of the molecular form of the reactants involved.

Introduction

In synthesis or hydrolysis reactions of amides, the reactants involved usually exist in a number of different ionised forms. Experimentally measured equilibrium constants at a given pH in part reflect the contributions of ionisation of acid and basic groups present. These contributions can be isolated from the equilibrium constant by expressing all reactant concentrations in terms of the uncharged forms only. This results in the following reference reaction:

$$[R^1\text{–}CO_2H]^0 + [H_2N\text{–}R^2]^0 \rightleftharpoons [R^1\text{–}CONH\text{–}R^2]^0 + H_2O \quad (1)$$

In 1960, Carpenter found that the Gibbs free energy changes in terms of this uncharged reference reaction varied within a narrow range for a series of peptide and amide hydrolysis reactions.[1] This indicates that for the above reaction a single reference equilibrium constant (K^0_{ref}) can be calculated. Two decades later, Svedas *et al.* found that for a range of beta-lactam antibiotics (these are also amides) the (K^0_{ref}) values were similar, but significantly different from values observed for peptides.[2]

Today, a much larger number of experimentally obtained equilibrium constants are available for synthesis and hydrolysis reactions. In order to establish whether a single reference value for K^0_{ref} for amide bond synthesis or hydrolysis exists, these need to be recalculated in terms of the uncharged reference reaction. For reactions involving zwitterions in particular, the calculations to account for ionisation are not straightforward. Some of the *microscopic* ionisation constants required can only be measured with extreme difficulty. Experimental microscopic pK_a values are therefore only available for a limited number of compounds and usually have to be estimated.[3]

In this article, the existing methods to estimate unknown pK_a values are compared. In addition, the mathematical tools to calculate neutral concentrations for reactions involving any number of ionisable groups are presented. For a large number of amide synthesis and hydrolysis reactions reported in the literature (these include protected and unprotected di- and tri-peptides, beta-lactam antibiotics, and other primary amine acylations) K^0_{ref} values were calculated.

Theory

In order to make valid comparisons between equilibrium constants of different reactions it is important that the concentrations of all reacting species are expressed in terms of one type of ionic species only. Use of total concentrations gives apparent equilibrium constants that vary with pH and are difficult to interpret directly. Here we choose the uncharged reference reaction (1) of synthesis, where all concentrations are in terms of the uncharged species (this includes groups that are not involved in the reaction). The thermodynamic reference equilibrium constant for this reaction is given by:

$$K^0_{ref.th} = \frac{a^0_{eq.product} a_{water}}{a^0_{eq.acid} a^0_{eq.base}} \quad (2)$$

If the reactants are present at dilute concentrations in water then the equilibrium can be described reasonably accurately by a concentration-based constant as follows:

$$K^0_{ref} = \frac{C^0_{eq.product}}{C^0_{eq.acid} C^0_{eq.base}} \quad (3)$$

where C^0_{eq} are equilibrium concentrations of uncharged species. In the following it is explained how concentrations of neutral species can be calculated from measured total concentrations.

How to calculate the concentration of uncharged species

The following treatment is based on well-known theories of acid, base and zwitterion ionisation. For more details the reader

 DOI: 10.1039/b108041e

This journal is © The Royal Society of Chemistry 2002

should refer to text books such as those of Greenstein and Winitz[4] or Mahler and Cordes.[5]

Acids and bases. Reactants such as C- or N-protected amino acids with neutral side chains are simple acids or bases. When the solid neutral forms (not salts) are dissolved in pure water they will largely remain in their uncharged forms. If however solubilities are measured in buffer of a known pH, the solution will consist of fractions of charged and uncharged species. The concentrations of the uncharged basic and acidic forms can be calculated from measured concentrations (C_{total}), the pH of the solution and pK values using the standard equations, eqns. (4) and (5).

$$C^0_{\text{acid}} = \frac{C^{\text{total}}_{\text{acid}}}{(1+10^{\text{pH}-\text{p}K_{\text{acid}}})} \quad (4)$$

$$C^0_{\text{base}} = \frac{C^{\text{total}}_{\text{base}}}{(1+10^{\text{p}K_{\text{base}}-\text{pH}})} \quad (5)$$

Throughout this treatment, H^+ dissociation constants will be used (K_a in the original nomenclature). Hence pK_{base} refers to the pK_a of the conjugate acid of the base, not pK_b.

Titration experiments with di-acids reveal two separate (macroscopic) pK values. These are the cumulative results of the three microscopic pK values. When only macroscopic pK values are known, the complete ionisation behaviour cannot be described as there are two different singly ionised species (except in symmetrical molecules) that cannot be distinguished. However, macroscopic pKs do allow for the calculation of the concentration of uncharged species using eqn. (6), or (7) for bases.

$$C^0_{\text{acid}} = \frac{C^{\text{total}}_{\text{acid}}}{(1+10^{\text{pH}-\text{p}K_{\text{acid}_1}})(1+10^{\text{pH}-\text{p}K_{\text{acid}_2}})} \quad (6)$$

$$C^0_{\text{base}} = \frac{C^{\text{total}}_{\text{base}}}{(1+10^{\text{p}K_{\text{base}_1}-\text{pH}})(1+10^{\text{p}K_{\text{base}_2}-\text{pH}})} \quad (7)$$

Zwitterions. Measured concentrations of zwitterionic compounds in pure water are predominantly of the zwitterionic (charged) form. Fig. 1 shows the complete distribution of differently charged species in the ionisation of glycine. It is clear that at low pH the protonated form predominates, at intermediate pH values the zwitterionic form predominates, while at higher pH values the main form is negatively charged. The neutral form is by far the smallest fraction, and is usually ignored for this reason. It is however this fraction that is needed for calculation of K^0_{ref}.

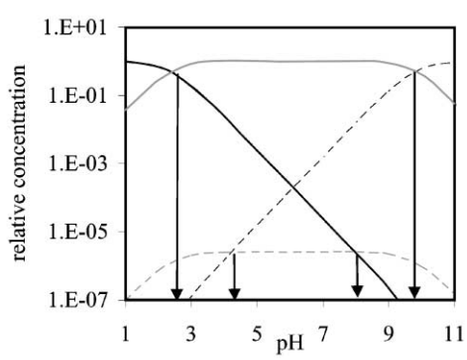

Fig. 1 Distribution of differently charged species of glycine as a function of pH. Grey is the zwitterionic form, black is the positively charged form, the dotted line is the negatively charged form and the dotted grey line represents the uncharged form. Arrows indicate the microscopic pK values, from left to right these are p$K^{\pm}_{\text{acid}}$, pK^0_{acid}, pK^0_{base}, and p$K^{\pm}_{\text{base}}$.

In order to interconvert total concentrations and those of the neutral form alone, the full ionisation pattern of zwitterions has to be considered. The two ionisation constants that are measured in titration experiments are the cumulative result of four *microscopic* ionisation constants. The complete ionisation pattern of zwitterionic molecules is illustrated in Fig. 2.

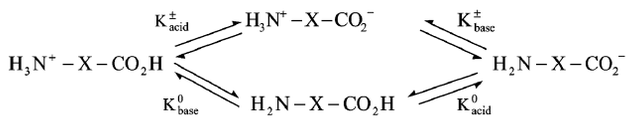

Fig. 2 Ionisation scheme of zwitterions with four different microscopic ionisation constants

Ionisation of each group is more favoured when a zwitterion rather than an uncharged species is formed. This is due to a combination of simple electrostatic effects (opposite charges stabilise each other) and effects through chemical bonds (electron donation and withdrawal). pKs of the neutral form are therefore somewhat higher for the acid group (pK^0_{acid} > p$K^{\pm}_{\text{acid}}$) and lower for the base (pK^0_{base} < p$K^{\pm}_{\text{base}}$) (Fig. 1). The concentration of the zwitterionic form can be calculated from:

$$C^{\pm} = \frac{C^{\text{total}}}{(10^{\text{p}K^{+}_{\text{acid}}-\text{pH}} + 10^{\text{p}K^{0}_{\text{acid}}-\text{p}K^{\pm}_{\text{base}}} + 10^{\text{pH}-\text{p}K^{-}_{\text{base}}} + 1)} \quad (8)$$

while the concentration of the neutral form follows from:

$$C^0 = \frac{C^{\text{total}}}{(10^{\text{p}K^{0}_{\text{base}}-\text{pH}} + 10^{\text{p}K^{0}_{\text{base}}-\text{p}K^{+}_{\text{acid}}} + 10^{\text{pH}-\text{p}K^{0}_{\text{acid}}} + 1)} \quad (9)$$

Fortunately, it is not necessary to try to estimate directly all four of these microscopic pK values. Firstly, the identity of the two paths in Fig. 2 leads to a relationship between the microscopic pK values, so that only 3 are independent:

$$\text{p}K^{\pm}_{\text{acid}} + \text{p}K^{\pm}_{\text{base}} = \text{p}K^{0}_{\text{base}} + \text{p}K^{0}_{\text{acid}} \quad (10)$$

Secondly, the readily available macroscopic pKs can be used to eliminate two more of the microscopic values. The simplest relationships are in terms of K rather that pK values:

$$K_{\text{acid}} = K^{\pm}_{\text{acid}} + K^{0}_{\text{base}} \quad (11)$$

$$\frac{1}{K_{\text{base}}} = \frac{1}{K^{\pm}_{\text{base}}} + \frac{1}{K^{0}_{\text{acid}}} \quad (12)$$

From eqns. (8), (9) and (10) the relationship between the concentrations of the zwitterionic and neutral forms can be derived:

$$\frac{C^{\pm}}{C^0} = 10^{\text{p}K^{0}_{\text{base}}-\text{p}K^{+}_{\text{acid}}} = 10^{\text{p}K^{+}_{\text{base}}-\text{p}K^{0}_{\text{acid}}} \quad (13)$$

From eqn. (13) it follows that when pK values of the acidic and basic groups are far apart, the concentration of the zwitterionic form predominates in solution (as seen in Fig. 1).

How to obtain pK values

General. pK values for single compounds commonly vary by several tenths of a pH unit when values from different literature

Table 1 Measured and estimated microscopic pK values of glycine derivatives[a]

	pK_{acid}^{0}			pK_{base}^{0}		pK_{ester}	$pK_{acid}^{\pm}$		$pK_{base}^{\pm}$	
Zwitterion	Exp.	SPARC	Est.	Exp.	SPARC	Ester	Exp.	SPARC	Exp.	SPARC
Gly[b]	4.4[3]	4.1[c]	4.4	8[3]	7.6[c]	7.7	2.4[3]	2.1	9.8[3]	9.6
di-Gly		3.6	3.4		7.4	7.8	3.1	3.2	8.1	7.8
tri-Gly		3.4	3.4		7.0	7.9	3.2	3.2	8.1	7.3
tetra-Gly		3.3	3.3		6.9	7.8	3.1	3.2	8.0	7.1

[a] 'Exp.' indicates experimental values while SPARC values are calculated. 'Est.' indicates values that are based on experimental pKs of alkyl esters and eqn. (10). 'Ester' indicates pK values of the ethyl ester of the acid. Values were taken from the *CRC Handbook of Chemistry and Physics*,[6] unless otherwise indicated. [b] Note that the experimental values for Gly do not satisfy eqn. (10) perfectly. [c] The values we obtained with the web SPARC calculator differed slightly from the published values.[3]

sources are compared. Throughout the paper deviations of <0.5 pH units are taken as acceptable, while deviations >0.5 are branded unreliable.

Acids and bases. Macroscopic pK values are commonly available from research papers, databases such as Beilstein Crossfire or printed reference literature such as the *CRC handbook for Chemistry and Physics*.[6] When they are not available, pKs of structurally similar compounds can be used. In addition, a mathematical model based on structure–activity relationships and perturbed molecular orbital theory has been developed to calculate pKs (SPARC: SPARC Performs Automated Reasoning in Chemistry).[7] A SPARC pK calculator is freely available on the internet (http://ibmlc2.chem.uga.edu/sparc). It was found to give values slightly different from published SPARC predictions.[3] SPARC values given in Table 1 are all taken from the web calculator.

Zwitterions. The microscopic pK values that are needed to calculate concentrations of uncharged species are usually not available. In fact, they have been determined for fewer than 100 compounds because measuring microscopic pK values (using spectral, calorimetric, or NMR based methods) is very complicated.[3,8] However, there are reasonable procedures to estimate them, as will now be explained.

Estimation of microscopic pK values. The differences between microscopic pK values for a given group depend strongly on the spatial distance between the acidic and basic groups within the zwitterionic molecule. This is because of a combination of two effects. First, there are 'through bond' electronic interactions between the acidic and basic groups. These effects wear off over 3 bond lengths. Second, there are 'through space' electrostatic effects between formal charges. They also tend to fall off with increasing separation between groups, but the extent to which this occurs depends on the conformation adopted in solution, which is difficult to predict.

Both of these effects are at least partly accounted for in a traditional approach for estimating microscopic pK values. These are approximated by the experimental pKs of molecularly similar, but non-zwitterionic, analogues of the molecule of interest.[4] In particular, the magnitude of pK_{base}^{0} is similar to the dissociation constant of an alkyl ester of the molecule of interest. Table 1 shows that the value for glycine ethyl ester is very similar to the experimental pK_{base}^{0}.

Unfortunately, experimental alkyl ester pKs are by no means always available. With some care, amide derivatives of an amino group can be used to estimate pK_{acid}^{0}, although the electron-withdrawing effect of the amide carbonyl will affect the acidity if there are fewer than three bonds between them. The pK values for Ac-, Bz- and Z-Gly (3.7, 3.8, and 3.7)[9] are significantly lower than pK_{acid}^{0} of Gly (4.4).[3] But the pK of 3.5 for Ac–Gly–Gly[9] is much closer to the pK_{acid}^{0} of Gly–Gly (3.7 based on the alkyl ester pK and eqn. (10)).

When no experimental pKs of non-zwitterionic analogues are available, SPARC can be used with some caution. pK_{acid}^{0} and $pK_{acid}^{\pm}$ were estimated reasonably by SPARC (see Table 1). Rather poor results were however obtained for $pK_{base}^{\pm}$ and pK_{base}^{0} of zwitterions with distant chargeable groups, such as the polyglycines (Table 1) and beta-lactam antibiotics (see Results and discussion section). SPARC predicts a continuing decline in $pK_{base}^{\pm}$ values as the chain length increases, while the experimental values remain constant between Gly–Gly and Gly–Gly–Gly–Gly. The same increasing difference was observed between SPARC predicted pK_{base}^{0} and the experimental pK_{ester} values. These observations suggest that SPARC does not properly account for some effects related to distance between the ionisable groups.

Results and discussion

Correct treatment of ionisation reveals a single non-ionised reference equilibrium constant for amide synthesis

With the tools described in the theoretical section it is now possible to calculate K_{ref}^{0} (eqn. (1)) from experimental equilibrium constants K_{eq} in water. This involves identification of the ion-type of each of the substrates and the products. Then the appropriate equations, eqns. (4)–(7) and (9), are inserted into general equation, eqn. (3).

$$K_{ref}^{0} = \frac{C_{eq,product}^{0}}{C_{eq,sub1}^{0} C_{eq,sub2}^{0}} = K_{eq} \frac{Y_{sub1} Y_{sub2}}{Y_{product}} \qquad (14)$$

Here, the Y terms are the denominators of eqns. (4)–(7) and (9). As an example, consider the formation of acetylglycine from acetate and glycine. For this reaction, the substrates are an acid and a zwitterion while the reaction product is an acid (entry 13, Table 2). Hence, take $Y_{sub1} = (1 + 10^{pH - pK_{acid}})$, $Y_{sub2} = (10^{pK_{base}^{0} - pH} + 10^{pK_{base}^{0} - pK_{acid}^{\pm}} + 10^{pH - pK_{acid}^{0}} + 1)$ and $Y_{product} = (1 + 10^{pH - pK_{acid}})$. The next step is then to identify the correct pK values for all reactants. For the current reaction the values for both substrates (4.8 for acetate, 2.4, 9.8 as the macroscopic ionisation constants for Gly and 8.0 as the pK_{base}^{0}) and the product (3.7) were available from the literature and databases. Whenever three zwitterion pK values are known the fourth follows directly from eqn. (10). When pK values are not available they can be estimated as explained in the theoretical section. The values are then simply inserted into eqn. (14) together with the pH of the reaction and the measured equilibrium constant.

A number of equilibrium data were obtained from the biocatalysis literature and the NIST database,[10] which lists a large number of equilibrium data for enzyme-catalysed reactions. All values for amide hydrolysis and synthesis found in the database and recent literature are shown in Table 2, except for a few special cases noted shortly. pK values were obtained as indicated in the footnotes. Because of the possible complications with zwitterions, Table 2 has separate sections depending on whether zwitterionic molecules were absent, present only as reactants, or present as both reactants and products. The data set covers a wide range of molecules and includes formation of charged (entries 1–3, 24) or neutral di- and tri-peptides (4–10),

Table 2 Equilibrium constants for amide formation and hydrolysis obtained from the literature and recalculated in terms of the uncharged form of the reactants only (log K^0_{ref}). pK values listed are macroscopic values in the case of non-zwitterions, while for zwitterionic compounds two macroscopic values and pK_{base} are given. Sources of pKs are indicated

Product[a]	Acyl donor pKs	Nucleophile pKs	Product pKs	pH	K^{obs}_{eq}	log K^0_{ref}
Non-zwitterion + non-zwitterion ⟶ non-zwitterion						
1 Boc-Asp—Phe-OMe	3.2[b], 4.8[b]	7.0[9]	4.1[c,13]	7.2	1.7[14]	3.7
2 Z-Asp—Met-OMe	3.2,[13] 4.8[13]	7.1[7]	4.1[c,2]	7.5	0.32[15]	3.3
3 Z-Asp—Phe-OMe	3.2,[13] 4.8[13]	7.0[9]	4.1[13]	6.0, 6.5	1.4,[16] 0.88[16]	3.2
4 Ac-Phe—di-Br-Tyr-OEt	3.6[9]	7.3[9]		4.1	0.4[10]	3.4
5 Ac-Phe—Gly-NH_2	3.6[9]	7.5[7]		5.5, 7.3, 8.2	0.5, 0.4, 0.18[10]	3.8
6 Ac-Phe—Phe-Gly-OMe	3.6[9]	7.3[9]		4.6	0.8[10]	3.7
7 Z-Gly—Phe-NH_2	3.7[7]	7.3[9]		7.2, 7.1, 7.0	0.6, 0.7, 1.0[10]	3.7
8 Z-Trp—Gly-NH_2	3.6[7]	7.5[9]		5, 6.7	0.45,[17] 0.45[18]	3.6
9 Bz-Tyr—Gly-NH-Ph	3.4[7]	7.3[7]		6.5	0.1[10]	3.0
10 Bz-Tyr—Gly-NH_2	3.4[7]	7.5[9]		8	0.2[10]	4.0
Zwitterion + non-zwitterion ⟶ non-zwitterion						
11 Ac—Ala	4.8[9]	2.4,[9] 9.9,[9] 7.8[c,9]	3.7[19]	Range	Range[10]	3.5
12 Ac—4-aminobutyrate	4.8[9]	4.2,[9] 10.4,[9] 9.7[c,9]	4.4[7]	7.5	0.18[10]	4.4
13 Ac—Gly	4.8[9]	2.4,[9] 9.8,[9] 8.0[3]	3.7[9]	7.5	0.22[10]	3.8
14 Ac—Met	4.8[9]	2.3,[9] 9.2,[9] 7.1[d,9]	3.5[7]	7.5	0.27[10]	2.9
15 Ac—NorLeu	4.8[9]	2.3,[19] 9.8,[19] 8.1[d,7]	3.7[7]	7.5	0.08[10]	3.6
16 Ac—NorVal	4.8[9]	2.3,[19] 9.8,[19] 8.1[d,7]	3.7[7]	7.5	0.10[10]	3.7
17 Ac—Val	4.8[9]	2.3,[19] 9.8,[19] 7.5[d,9]	3.7[7]	Range	Range[10]	3.3
18 Cephaloridine[e]	4.2[2]	2.2,[7] 5.8,[7] 5.3[7]	2.7[7]	5	67[10]	3.6
19 Cephalotin[f]	4.2[2]	2.5,[h,2] 4.8[h,2] 4.5[h]	3.7[7]	5	143[10]	4.1
20 Cephamandole[g]	3.9[9]	2.5,[h,2] 4.9[h,2] 4.6[h]	3.6[7]	7	0.26[20]	3.3
21 Penicillin G[i]	4.2[9]	2.5,[11] 4.9,[11] 4.6[2,j]	2.8[9]	Range	Range[10]	3.4
22 PhAc—7-ADCA	4.2[9]	2.5,[2] 4.8,[2] 4.5[2]	2.8[k]	5, 6	(1.4 10^3)10, 9.5[12]	3.5
23 PhAc—Gly	4.2[9]	2.4,[9] 9.8,[9] 8.0[3]	3.6[9]	5	0.8[10]	5.0
24 Ac-Phe—Phe-Gly	3.6[9]	3.3[l], 8.3[l], 7.3[c,1]	3.39[9]	4.6	0.2[10]	3.1
Zwitterion + zwitterion ⟶ zwitterion						
25 Amoxicillin	2.2,[11] 9.2,[11] 8.1[d,21]	2.5,[11] 4.9,[11] 4.6[g]	2.9,[11] 7.4,[11] 7.5[g]	Range	Range[11]	3.6
26 Cephalexin	2,[2] 9,[9] 7.1[d,2]	2.5,[2] 4.8,[2] 4.5[g]	2.9,[11] 7.4,[11] 7.5[g]	Range	Range[12]	3.6
Average						3.6

[a] '—'indicates amide bond formed, for beta-lactam antibiotics the bond formed is that between the side chain and the antibiotic nucleus. [b] Taken as Z-Asp. [c] Taken as Z-Asp-Phe-OMe. [d] Taken as alkyl ester. [e] The antibiotic nucleus and antibiotic have a permanent positive charge.[22] [f] This is 2-thienylacetic acid linked to an antibiotic nucleus.[22] [g] Cephamandole consists of *d*-mandelic acid linked to TET-ACA 7-amino-3-(1-methyltetrazole-5-yl)thiomethylceph-3-em-4-carboxylic acid.[20] [h] Taken as 7-ADCA (7-aminodeacetoxycephalosporinic acid). [i] Penicillin G consists of phenylacetic acid linked to 6-APA (6-aminopenicillanic acid). [j] Assuming that pK^0_{acid} is the same as pK_{acid} of penicillin G, pK^0_{base} obtained from eqn. (10). [k] Taken as penicillin G. [l] Values for Leu-Gly.[6]

zwitterionic (25,26) and singly charged antibiotics (18–22) as well as acylations of amino acids (11–17; 23).

Some of the equilibrium constants in the literature have been excluded from this table. The NIST database lists a value for benzylpenicilloic acid. The nucleus of this antibiotic has 4 ionisable groups, which makes calculation of its uncharged concentration very complex. Since experimental pKs were not available and SPARC tends to perform poorly with these compounds we decided to leave this entry out. For zwitterionic beta-lactam antibiotics, such as ampicillin, cephalexin and amoxicillin, it has been reported[11,12] that some of the older equilibrium values in the literature were overestimated. These were left out of the current analysis. Where experimental values were available for beta-lactam antibiotics, SPARC predictions were very poor (often out by more than one unit). Hence in those cases where pK values were not known the experimental values of similar compounds were used (as indicated).

From Table 2 it is clear that log K^0_{ref} is largely constant, for the three categories of reactions, regardless of the presence of zwitterions in the reaction. It is doubtful whether the experimental accuracy of the data used is sufficient to consider as significant any of the deviations from the average value of $10^{3.6}$ M^{-1} since reported experimental pKs often vary by several tenths of a unit for the same compound. The most reliable equilibrium constants will naturally be obtained when equilibrium is approached from both directions (hydrolytic and synthetic) at a range of different pH values (entries 11, 17, 21, 25, 26). The only K^0_{ref} that was out by more that 1 order of magnitude was for phenylacetyl—glycine (entry 23). This equilibrium constant was obtained from hydrolysis only,[2] so it may well be overestimated (equilibrium had not yet been reached).

The reference equilibrium constants calculated here directly reflect the free energy change of condensation of an amino group and a carboxylic acid group to form an amide bond. The small deviations observed show that this value is largely independent of the nature of remote parts of the reacting molecules. From the calculated average value of log K^0_{ref} the free energy change involved with amide bond formation could be directly calculated, ΔG^0_{ref} = 20.5 kJ mol^{-1}. The value ranges between 16 and 25 kJ mol^{-1} (excluding only the clear outlier PhAc—Gly). This agrees closely with the values obtained for a number of amides calculated by Carpenter (9 to 22 kJ mol^{-1}).[1] It also overlaps with the values calculated for a series of beta-lactam antibiotics by Svedas *et al.* (17 to 28 kJ mol^{-1}).[2] The current analysis is both more rigorous and based on a much larger number of reactions.

General

The existence of a single reference equilibrium constant can be of importance in various areas. Applied biocatalysis is one field in which the structure-independent reference equilibrium constant will be of value. In the last ten years researchers have rediscovered water as a solvent for enzyme-catalysed reverse hydrolysis reactions. This is particularly true for systems where precipitation of the synthesis products led to very good yields. The feasibility of product precipitation in these systems can be assessed straightforwardly by comparison of the ratio of

product over substrate solubilities to the equilibrium constant of the reaction.[23] Clearly, such an assessment is greatly simplified when a constant reference equilibrium value can be used. In addition, the final yields of these reactions are naturally determined by the equilibrium constant and the tools developed here can be used to predict at what pH the highest yield can be expected. This does not require any experimental information on the actual reaction of interest.

Another research area where the existence of a single free energy change of peptide synthesis and hydrolysis can be of importance is in understanding cellular metabolism. Many biological molecules contain amide bonds, and Gibbs energy changes for their synthesis have usually not been measured. The constant reference value found here, together with the method presented to allow for ionisation, will allow accurate understanding of the energetics of these important biochemical reactions.

The observation of a constant (uncharged) reference equilibrium constant is probably quite general, and it should be possible to extend it to other types of biochemical reactions such as (reverse) hydrolysis of esters or glycosides.

Conclusions

In this paper we show that the equilibrium constant for amide bond formation (expressed in terms of the concentrations of the uncharged form of reactants) is largely constant, regardless of the molecular form and ionisation behaviour of the reacting molecules. This provides a reference value ($K^0_{ref} = 10^{3.6}$ M^{-1}, $\Delta G^0_{ref} = 20.5$ kJ mol^{-1} at 298 K) that in combination with pK values (either experimental or predicted) can be used to estimate the equilibrium position of amide synthesis reactions from any starting materials.

Acknowledgements

The authors would like to thank the BBSRC, Wageningen University, and DSM Research for funding.

References

1 F. H. Carpenter, *J. Am. Chem. Soc.*, 1960, **82**, 1111.
2 V. K. Svedas, A. L. Margolin and I. V. Berezin, *Enzyme Microb. Technol.*, 1980, **2**, 138.
3 S. H. Hilal, S. W. Karickhoff and L. A. Carreira, *Talanta*, 1999, **50**, 827.
4 J. P. Greenstein and M. Winitz, *Chemistry of the Amino Acids, Volume 1*, Wiley, New York, 1961, p. 435.
5 H. R. Mahler and E. H. Cordes, *Biological Chemistry, Int. Ed.*, Harper & Row, New York, 1966, p. 193.
6 *CRC Handbook of Chemistry and Physics*, ed. D. R. Lide, CRC Press, Boca Raton, 74th Edn., 1994, pp. 8–43.
7 SPARC web calculator to be found at http://ibmlc2.chem.uga.edu/sparc.
8 J. C. D'Angelo and T. W. Collette, *Anal. Chem.*, 1997, **69**, 1642.
9 Beilstein Crossfire Database, Version 4.0, MDL Information Systems GmbH, Frankfurt.
10 NIST to be found at http://wwwbmcd.nist.gov:8080/enzyme/enzyme.html.
11 M. B. Diender, A. J. J. Straathof, L. A. M. Van der Wielen, C. Ras and J. J. Heijnen, *J. Mol. Catal. B: Enzym.*, 1998, **5**, 249.
12 C. G. P. H. Schroën, V. A. Nierstrasz, P. J. Kroon, R. Bosma, A. E. M. Janssen, H. H. Beeftink and J. Tramper, *Enzyme Microb. Technol.*, 1999, **24**, 498.
13 M. Erbeldinger, X. Ni and P. J. Halling, *Biotechnol. Bioeng.*, 2001, **72**, 69.
14 H. Kitano, Y. Maeda, M. Yamamoto and R. Izumida, *Macromol. Chem. Phys.*, 1996, **197**, 4173.
15 S. Nakamura and S. Nakai, *Nahrung*, 1998, **42**, 131.
16 K. Inouye, *J. Biochem.*, 1992, **112**, 335.
17 G. A. Homandberg, J. A. Mattis and M. Laskowski, *Biochemistry*, 1978, **14**, 181.
18 H. D. Jakubke, P. Kuhl and A. Konnecke, *Angew. Chem., Int. Ed. Engl.*, 1985, **24**, 85.
19 SPARC Database, 2001, to be found at http://ibmlc2.chem.uga.edu/sparc.
20 V. A. Nierstrasz, C. G. P. H. Schroën, R. Bosma, P. J. Kroon, H. H. Beeftink, A. E. M. Janssen and J. Tramper, *Biocatal. Biotransform.*, 1999, **17**, 209–224.
21 M. B. Diender, A. J. J. Straathof, T. Van der Does, M. Zomerdijk and J. J. Heijnen, *Enzyme Microb. Technol.*, 2000, **27**, 576.
22 A. A. Petrauskas and V. K. Svedas, *J. Chromatogr.*, 1991, **585**, 3.
23 R. V. Ulijn, A. E. M. Janssen, B. D. Moore and P. J. Halling, *Chem. Eur. J.*, 2001, **7**, 2089.