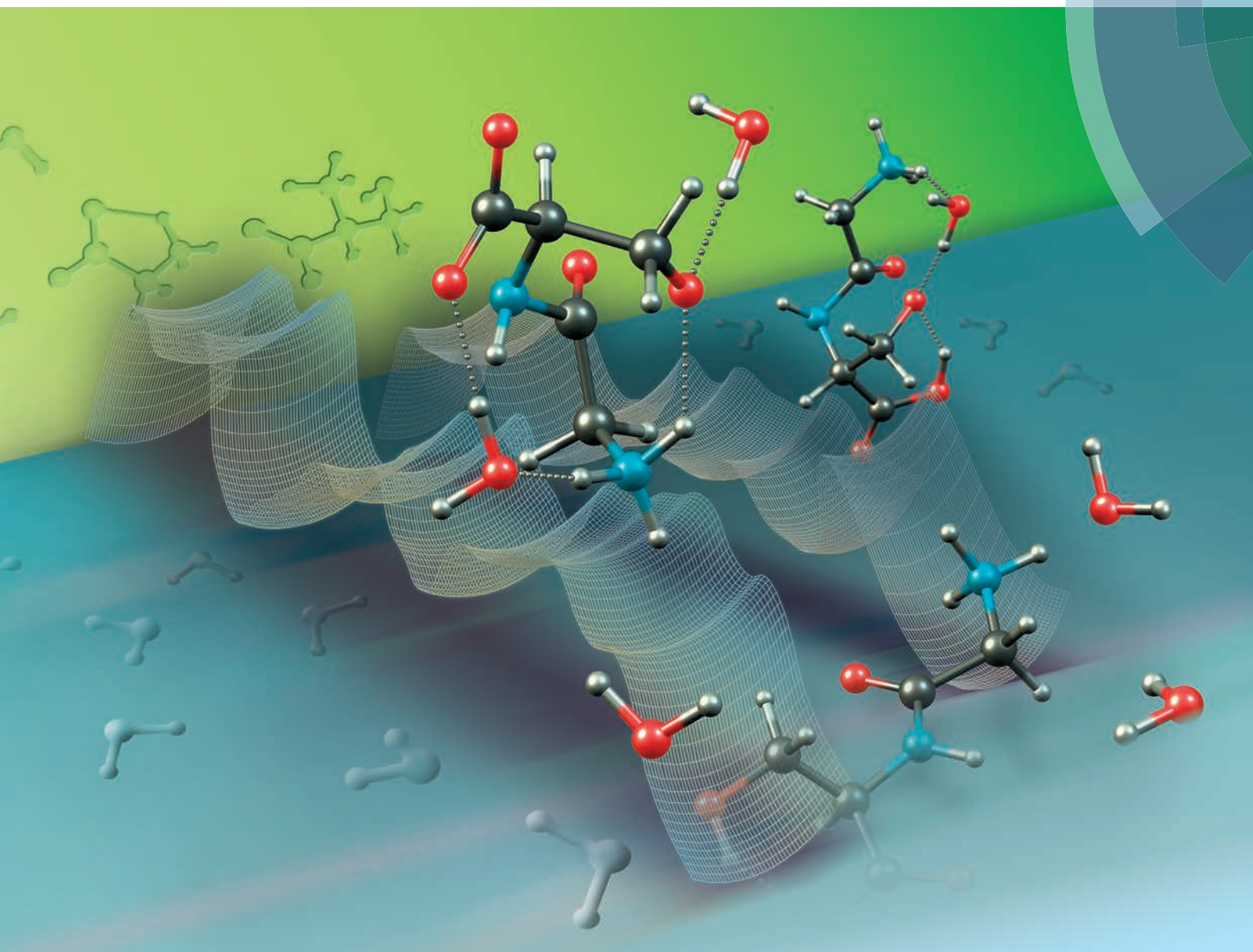


Organic & Biomolecular Chemistry

www.rsc.org/obc



ISSN 1477-0520



PAPER

Tzvetan T. Mihaylov *et al.*

A computational study of the glycolserine hydrolysis at physiological pH: a zwitterionic versus anionic mechanism

A computational study of the glycyserine hydrolysis at physiological pH: a zwitterionic versus anionic mechanism†

Cite this: *Org. Biomol. Chem.*, 2014, **12**, 1395

Tzvetan T. Mihaylov,* Tatjana N. Parac-Vogt and Kristine Pierloot

The hydrolysis of GlySer at physiological pH was investigated by modeling the most feasible reaction mechanisms in aqueous phase at the MP2/6-311+(2df,2p)//SMD-M06/6-311+(2df,2p) level of the theory. To refine the energies of the most relevant transition states along the reaction paths the cluster-continuum concept was adopted. The hydrolytic process could proceed through two competitive mechanisms involving either the zwitterionic or the anionic form of GlySer. The calculations suggest that at physiological pH the actual mechanism is most probably mixed, anionic–zwitterionic. In this reaction scheme the first stage of N→O acyl transfer involves the anionic form whereas the second stage, during which the resultant ester is hydrolyzed, most likely involves the zwitterionic ester form of GlySer. The energy requirement for the first reaction stage is estimated to be slightly lower than for the second one. The calculated activation parameters (e.g. $\Delta G^\ddagger = 27.8 \text{ kcal mol}^{-1}$) for the nucleophilic addition of a water molecule to the ester carbonyl group of the zwitterionic ester are in good agreement with the experimentally determined values at pD 7.4 ($\Delta G^\ddagger = 28.7 \text{ kcal mol}^{-1}$).

Received 26th November 2013,
Accepted 18th December 2013

DOI: 10.1039/c3ob42372g

www.rsc.org/obc

Introduction

Several classes of intrinsically reactive proteins are known to cleave an internal peptide bond as part of their cellular function, without the need for an auxiliary enzyme or cofactor.¹ Typical examples of such proteins are the intein domain in protein splicing,² the pyruvoyl-dependent enzymes,³ the *hedgehog* domain,⁴ the N-terminal nucleophile (Ntn) hydrolases,⁵ the nucleoporins,⁶ and the SEA domain.⁷ Despite the diverse biological functions of these proteins, the self-cleaving processes are in all of them initiated by an N→O or N→S acyl shift (transfer) of a peptide bond involving the amino group of a serine (Ser), threonine (Thr) or cysteine (Cys) residue.¹ The mechanism of this rearrangement involves the intramolecular nucleophilic attack by an –OH or –SH group, present in the side chain of such a residue, on the adjacent amide carbon of the protein backbone, resulting in a transient oxazolidine or thiazolidine ring intermediate. The latter breaks down into a (thio)ester which is further resolved by a subsequent reaction that differs between different autoproteolytic systems.⁸ Under physiological conditions the equilibrium of an N→X (X = O, S) acyl shift lies in favor of the amide bond, thus limiting the

peptide bond cleavages.⁹ However, the protein environment around the scissile bond may significantly influence this process by promoting the formation of a (thio)ester through deprotonation of the nucleophilic –XH group,^{2,10} stabilization of the transient ring intermediate *via* an oxyanion hole,^{10a,11} protonation of the leaving amino group,^{10b,c} or ground-state destabilization of the scissile peptide bond through conformational strain.^{7,12}

Several theoretical studies have already revealed important insights into the self-catalyzed N→X acyl shift mechanism in proteins. They are mainly devoted to protein splicing and SEA domain autoproteolysis, where the process is accelerated either through deprotonation of the nucleophilic –XH group/protonation of the leaving amino group^{10b} or through conformational strain of the peptide bond.^{7b,13} In contrast, the molecular mechanism of the N→X acyl shift initiated peptide bond cleavage of systems where the aforementioned catalytic factors are absent has not yet been investigated. The main advantage of studying the self-cleaving process in such catalytic systems is that it will provide the opportunity to extract essential mechanistic insights about the hydrolytic process, not being influenced by the protein environment surrounding the scissile peptide bond in a protein.

Glycyserine (GlySer) by itself is the simplest serine containing dipeptide that undergoes an N→O acyl shift initiated self-cleavage of the peptide bond, analogous to the one in proteins, and therefore a very suitable model system for such a study. The N→O acyl shift mechanism in GlySer is strongly supported

KU Leuven, Department of Chemistry, Celestijnenlaan 200F, B-3001 Leuven, Belgium.
E-mail: tzmihay@svr.igic.bas.bg; Fax: +32-16-327992; Tel: +32-16-327347

† Electronic supplementary information (ESI) available: Additional Fig. S1 and S2. Optimized molecular structures of the most relevant species in Cartesian coordinates. See DOI: 10.1039/c3ob42372g

by the fact that peptide bond hydrolysis is much faster in the GlySer sequence than in the SerGly one.¹⁴ A logical explanation for this phenomenon could be found in the aforementioned mechanism: the intramolecular attack of the Ser –OH group on the amide carbon is likely in GlySer but unlikely in SerGly, because in the first case it gives rise to a five-membered ring transition state while in the latter case it leads to an unfavorable four-membered ring transition state.

We have recently reported a mechanistic study on the hydrolysis mechanisms of the zwitterionic form of GlySer, where special attention was paid to the possible N→O acyl shift pathways which could be relevant for proteins.¹⁵ The theoretically predicted activation energy for the hydrolysis of the GlySer zwitterion (**ZW**) was found to be in excellent agreement with the experimental results obtained at pD 5.4. Interestingly, a comparison between the experimental activation parameters estimated at pD 5.4 and at pD 7.4 revealed that despite the close similarity between the activation Gibbs energies, the activation entropies and enthalpies exhibit significant differences. These results strongly suggest that the mechanism of the rate-limiting step at pH 5.4 and at neutral pH might be thoroughly different. At pH 5.4 only the **ZW** is supposed to be present in the reaction solution since the isoelectric point of GlySer (5.68) is very close to this pH value. This is not the case, however, at physiological pH. The pK_a value of the GlySer amino group is 8.28¹⁶ which means that at pH 7.4, according to the Henderson–Hasselbalch equation, this dipeptide should exist as a mixture of about 89% **ZW** and 11% anion (**A**). Hence, the rate of GlySer hydrolysis at physiological pH (or in slightly basic solution) might be influenced by **A** through competition with the **ZW** mechanism. This possibility, however, has been essentially unexplored in previous work. Therefore, in the present paper we set out to obtain detailed insight into the hydrolysis process at physiological pH by modeling the most feasible reaction mechanisms involving both **A** and **ZW**.

Computational details

The hydrolysis mechanism of both the anionic and zwitterionic forms of GlySer was modeled in aqueous solution by explicit consideration of a limited number of water molecules, directly interacting with the solute, and implicit treatment of the rest of the solvent with continuum solvation. Such a cluster-continuum approach allows one to reveal the role of the solvent in the reaction mechanisms pointed out as important in previous experimental studies on similar systems.¹⁷ Its efficiency has been shown in a number of studies for both organic and organometallic reactions in solution.¹⁸ According to the cluster-continuum model proposed by Pliego *et al.*,¹⁹ for chemical reactions, the number of explicit solvent molecules (*n*) in the transition state has to be varied in order to attain the lowest possible activation energy. Following this strategy, we performed extensive molecular modeling of transition states (TS) with up to three water molecules. For practical reasons, however, this procedure was applied only for the most relevant

TSs whereas the rest of the reaction mechanisms were modeled with one explicit solvent molecule. In all cases, the arrangement of the water molecules in each TS-*n*H₂O cluster presented in the paper is determined by a systematic search and optimization of many different models so as to ensure maximum system stabilization.

All molecular geometries were optimized with the M06 functional²⁰ and 6-311+G(2df,2p) basis sets, conjugated with an SMD²¹ continuum solvation model. To confirm the character of the stationary points on the potential energy surface (PES), vibrational frequency calculations at the same level of theory were carried out. Intrinsic reaction coordinate (IRC) calculations were also performed to verify the expected connections between TSs and local minima on the PES. The M06 electronic energies were further adjusted by single-point calculations with the MP2 method. This strategy was previously found by us to be most accurate in predicting the activation free energy of the GlySer zwitterion hydrolysis.¹⁵

The free energy profile of the reactions studied is evaluated as described below. The aqueous phase free energy of a stationary point X, along the reaction, may be expressed as the sum of its gas-phase free energy $G^0(\text{X})$ and the free energy of solvation $\Delta G_{\text{solv}}^*(\text{X})$.

$$G_{\text{aq}}(\text{X}) = G^0(\text{X}) + \Delta G_{\text{solv}}^*(\text{X}) \quad (1)$$

To calculate the first term in eqn (1) the zero-point vibrational energies, thermal corrections and entropy terms were derived from the M06 frequency calculations at the aqueous phase optimized geometries, using the standard rigid rotor-harmonic oscillator approximation. Recent investigations have shown that using partition functions computed for molecules optimized in solution is a correct and useful approach for averaging over solute degrees of freedom when computing free energies of solutes in solution.²² The energy shift due to the hydration (the second term in eqn (1)) is calculated by means of the SMD solvation model. As the $G^0(\text{X})$ values are computed with respect to a standard state of 1 atm (denoted by the superscript “0”), whereas the continuum solvation models like SMD usually deliver hydration free energies which refer to a standard state of 1 M²¹ (denoted by the superscript “*”) a correction term, $\Delta G^{0 \rightarrow *}$, should be included in eqn (1). A discussion about the standard state corrections could be found in ref. 23–25.

$$\begin{aligned} \Delta G^{0 \rightarrow *} &= -T\Delta S^{0 \rightarrow *} = -RT \ln \left(\frac{V^*}{V^0} \right)_{P,T} = -RT \ln \left(\frac{1 \text{ L}}{24.46 \text{ L}} \right) \\ &= RT \ln(24.46) = 1.89 \text{ kcal/mol} \quad (T = 298.15 \text{ K}) \end{aligned} \quad (2)$$

Similar energy contributions must also be included if solvated molecules reside in a different standard state than 1 M, *e.g.*, water molecules solvated by water, which has a standard state concentration of 55.34 M.

$$\begin{aligned}\Delta G^{0 \rightarrow \text{conc.}} &= -T\Delta S^{0 \rightarrow \text{conc.}} = -RT \ln\left(\frac{V^{\text{conc.}}}{V^0}\right)_{P,T} \\ &= -RT \ln\left(\frac{(1/55.34) \text{ L}}{1 \text{ L}}\right) \\ &= RT \ln(55.34) = 2.38 \text{ kcal/mol } (T = 298.15 \text{ K})\end{aligned}\quad (3)$$

Thus, at a standard state of 1 M, taking the energy of the infinitely separated reactants (GlySer and $n\text{H}_2\text{O}$) as a reference, the relative aqueous-phase Gibbs energy of X, microsolvated by n water molecules, could be expressed as

$$\Delta G_{\text{aq}}^*(X(\text{H}_2\text{O})_n) = G_{\text{aq}}(X(\text{H}_2\text{O})_n) - G_{\text{aq}}(\text{GlySer}) - n(G_{\text{aq}}(\text{H}_2\text{O}) + \Delta G^{0 \rightarrow \text{conc.}} + \Delta G^{0 \rightarrow *}) \quad (4)$$

Eqn (4) was used throughout this report to evaluate the free energy profile of the reactions studied. All calculations were performed with the Gaussian 09²⁶ suite of programs.

Results and discussion

The reaction of GlySer hydrolysis consists of two stages, which will be discussed separately below. In the first reaction stage the scissile peptide bond interconverts into an ester *via* an N→O acyl shift mechanism. The resultant ester is then hydrolyzed, in the second stage, to give the reaction products Gly and Ser.

N→O acyl shift mechanism

The discussion below will be based mainly on the N→O acyl shift mechanism proposed for the GlySer anion **A**. As mentioned in the introductory part, the same mechanism for the zwitterion **ZW** was thoroughly discussed in our previous work.¹⁵ Therefore, it will be briefly presented here only for the sake of comparison. However, as could be expected both mechanisms are quite similar. As shown for **ZW** the highest energy barrier that limits the N→O acyl transfer corresponds to the Ser –OH nucleophilic attack on the amide carbon. It has been found that the energy requirement for this attack mostly depends on the (aqueous phase) basicity of the assisting atomic site that abstracts the –OH proton during the addition.¹⁵ As the –NH₂ nitrogen is the strongest internal proton acceptor of **A**, it may be safely assumed that this group should play an active role here. Therefore, we have modeled possible TSs where the Ser –OH attack on the amide carbon is accompanied by a proton transfer (PT) from –OH to the –NH₂ group, by gradually changing the number of explicit solvent molecules, n , in the TS. The results are presented in Fig. 1. The proposed N→O acyl shift mechanisms in **A** and in **ZW** as well as the corresponding reaction energy profiles are given in Fig. 2 and 3.

In the **ZW** form the first TS of addition, zw-TS1-1 (Fig. 1), is stabilized by one water molecule hydrogen bonded to both the alcoholate oxygen and a –NH₃⁺ hydrogen. Here, the nucleophilic attack of the Serine –OH is accompanied by a PT of the OH proton to the carboxyl group (this process is denoted as channel A1 in our previous study¹⁵). Although the latter

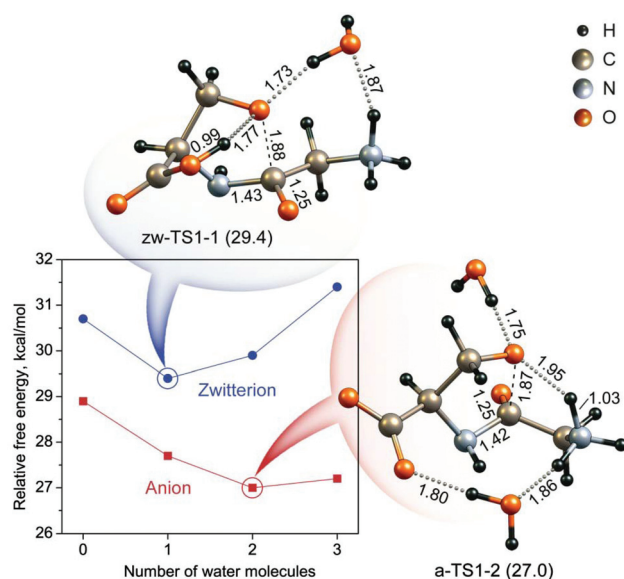


Fig. 1 Molecular structures (obtained from M06) of the lowest energy transition states for intramolecular nucleophilic attack by the Ser –OH group on the amide carbon. Energy (MP2, in brackets) is in kcal mol^{−1}, interatomic distances are in Å. Alternative, higher-energy structures with a different number of H₂O are given in the ESI, Fig. S1.†

process is already accomplished in the TS (O–H distance = 0.99 Å), IRC calculations pointed out that the PT and the nucleophilic attack are concerted. In the **A** form, however, the PT occurs from the –OH to the –NH₂ group, and IRC calculations indicate that the PT precedes the nucleophilic attack. As shown in Fig. 2, starting from the reactant complex (a-RC-2) the reaction system passes through a TS, corresponding to the –OH proton abstraction by the amino nitrogen (a-TS1-1-2), to reach an intermediate structure with a bare O-nucleophile (a-I1-1-2). In the next TS, a-TS1-2, the –O[−] attacks the amide carbon at a distance of 1.87 Å (1.88 in **ZW**) to form a five-member oxyoxazolidine ring intermediate (a-I1-2). Two solvent molecules stabilize the a-TS1-2 structure: one is H-bonded to the alcoholate oxygen and the second is H-bonded to both the amino and the carboxylate groups. Although the a-I1-1-2 structure was located as a minimum on the PES it is not a minimum on the free energy surface (FES). Hence, the PT and the nucleophilic O-attack are in fact concerted after all. The calculated free energy barrier for this process is 27.0 kcal mol^{−1}, which is 2.4 kcal mol^{−1} lower than for the **ZW** (29.4 kcal mol^{−1}).

These are, in fact, the highest barriers along the N→O acyl shift reaction. The following TSs in the reaction stage have relatively lower energy requirements. Therefore, their free energies were not refined through variation of n .

In order to collapse the oxyoxazolidine intermediate to a linear ester, with the release of a new –NH₂ terminus, the ring nitrogen needs to be protonated whereas the lateral (out of the ring) oxygen cannot bare a proton. By modeling different possible rearrangements it has been found that the lowest energy path on FES for a ring opening passes through two PTs,

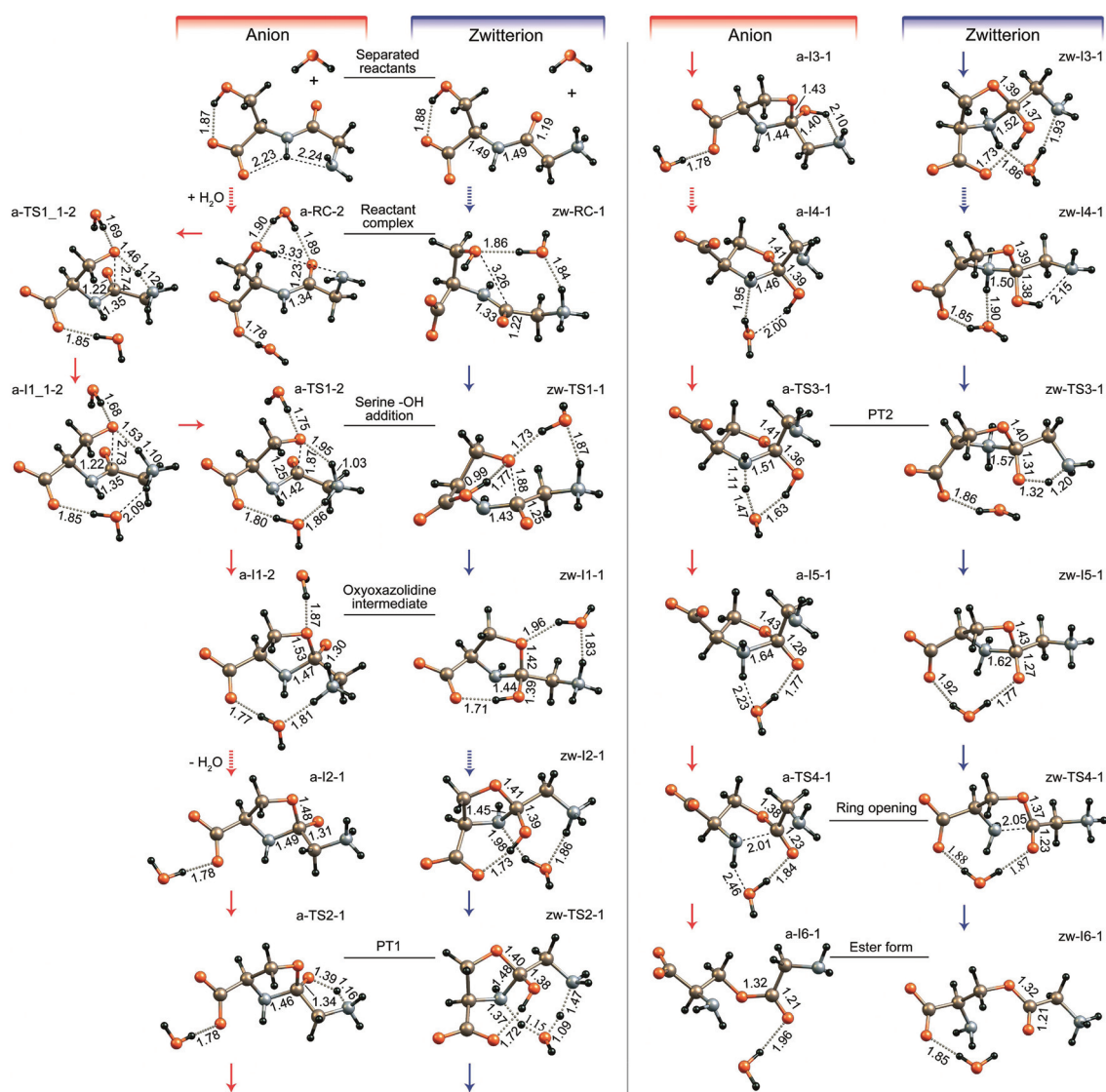


Fig. 2 Proposed mechanisms for the N→O acyl transfer in **A** and **ZW**. Selected interatomic distances are given in Å. The structures obtained by following intrinsic reaction coordinates are connected with a solid arrow. Conformational interconversions²⁷ are indicated with a dashed arrow.

designated as PT1 and PT2 in Fig. 2 and 3. PT1 (a-TS2-1) is an H⁺ transfer from the terminal amino group, protonated in the first reaction step, to the lateral oxygen atom. According to the free energy calculations PT1 proceeds spontaneously ($\Delta G_{\text{aq}}^{\ddagger}$ of a-I2-1 is slightly higher than a-TS2-1, see Fig. 3). The proton is then transferred in the subsequent PT2 (a-TS3-1) to the ring nitrogen through the medium of a solvent water molecule. The transition state of PT2 has a relative energy of 26.2 kcal mol⁻¹ as compared to the initial reactants. The resulting intermediate (a-I5-1, at 19.7 kcal mol⁻¹, in which a net PT from the -NH₃⁺ group to the ring nitrogen has occurred) does not decompose spontaneously. Instead, it needs to overcome a barrier of ~3 kcal mol⁻¹ to reach the ester form. In the **ZW** the lateral oxygen spontaneously abstracts a proton from the -COOH group, which becomes protonated in the first addition step (zw-I1-1). As such, the lowest energy path here proceeds through a different mechanism. PT1 (zw-TS2-1) now involves a

PT from the terminal -NH₃⁺ group to the ring nitrogen *via* a bridging water molecule, while the NH₃⁺ form is recovered in the subsequent PT2 by means of a direct PT from the lateral -OH to the amino nitrogen. The energy requirements for both PT1 and PT2 in **ZW** were calculated to be 22.7 kcal mol⁻¹. Also in this case the resulting intermediate (zw-I5-1) needs to overcome an extra barrier (~4 kcal mol⁻¹) to reach the ester form.

Ester hydrolysis mechanism

The GlySer ester hydrolysis of both the **A** and **ZW** forms proceeds *via* a two-step process, involving the nucleophilic addition of a water molecule to the ester carbonyl group and the decomposition of the resulting tetrahedral intermediate to the products. In our previous work,¹⁵ the corresponding transition states were modeled with one explicit water molecule only. Here, the TS mechanisms/energies have been further refined by a variation of *n*, as shown in Fig. 4. The proposed

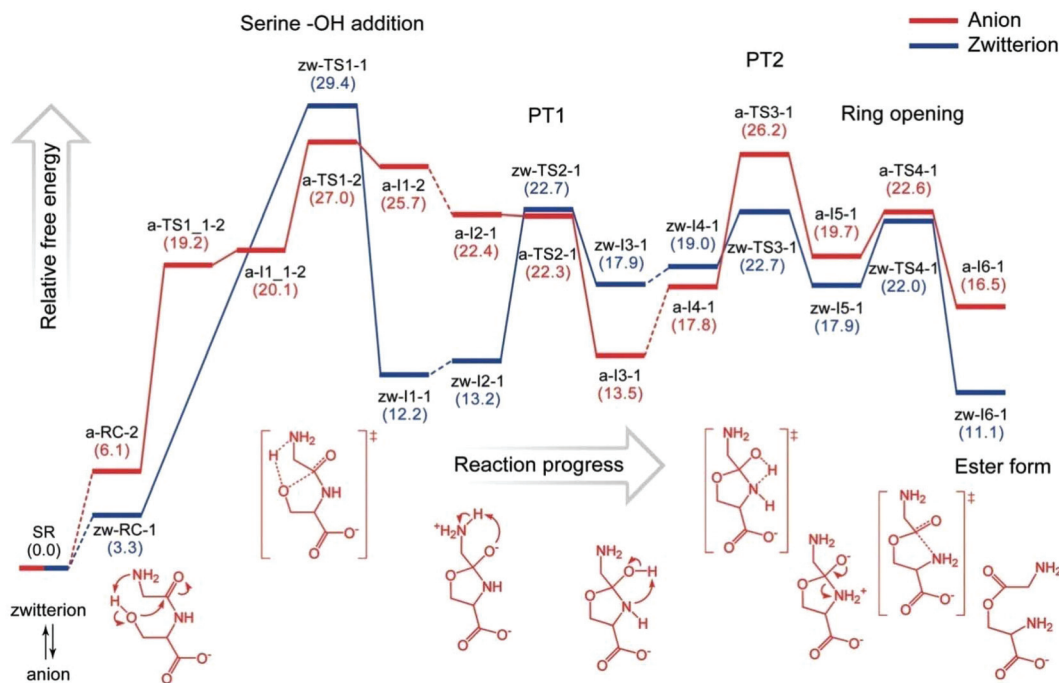


Fig. 3 Free energy profiles (kcal mol⁻¹) of the N→O acyl transfer in **A** and **ZW**. The energy levels corresponding to structures successively obtained by following the intrinsic reaction coordinate are connected with a solid line. The energy levels of structures linked through conformational inter-conversion²⁷ are connected with a dashed line. The water molecules in the structural formulas representing the reaction process of **A** are omitted for simplicity. For more detailed structures we refer to Fig. 2.

molecular mechanisms of **A** and **ZW** ester hydrolysis are depicted in Fig. 5 and the corresponding energy profiles are plotted in Fig. 6.

As shown in Fig. 4, for both **A** and **ZW** the energy barrier corresponding to the first step of addition minimizes at $n = 1$ (only the attacking water). Many different arrangements of the solvent molecules around the solute were considered, but we could not find any specific extra solute–solvent interactions resulting in extra stabilization of the corresponding TSs. We find that during the water addition process the amino group that was released in the oxazolidine ring opening acts as a general base to abstract a proton from the water molecule, and the incipient hydroxide ion attacks the C-atom at a distance of ~ 1.9 Å to form a new O–C bond in the resulting intermediate (a-I8-1, zw-I8-1). A similar intramolecular base catalyzed hydrolytic mechanism has been previously proposed for ester compounds of 2-aminobenzoic acid.²⁸ The calculated TS free energies for the water addition process of **A** and **ZW** are 31.8 kcal mol⁻¹ (a-TS-wa-1) and 27.8 kcal mol⁻¹ (zw-TS-wa-1) respectively (with respect to the initial reactants). In the second step, the tetrahedral intermediate breaks down to release the Ser moiety. For this step, the most likely TS structures of **ZW** (zw-TS-d-1) and **A** (a-TS-d-2) involve one and two solvent molecules respectively. In both cases the explicit water molecules act to stabilize the activated complexes through H-bonding with the alcoholate oxygen of the leaving Ser. Moving downhill from this TS to the product complex, the alcoholate oxygen of **ZW** abstracts a proton from the $-\text{NH}_3^+$ group of Gly

to produce a zwitterionic Ser and a neutral Gly (zw-I10-1). The Gly zwitterionic form could subsequently be recovered through an intramolecular PT from the $-\text{COOH}$ group to the amino nitrogen. In the **A** form, however, such a proton abstraction is not possible in the proposed reaction scheme and the TS of decomposition collapses to a high energy product complex consisting of an anionic Ser with a bare alcoholate oxygen and neutral Gly (a-I10-2). The $-\text{O}^-$ oxygen of Ser could be further protonated either by the Gly $-\text{COOH}$ group (spontaneous PT) or by the solvent. The estimated activation energies for the second (decomposition) step of **A** and **ZW** forms are lower by only 1.4 and 0.3 kcal mol⁻¹ with respect to the TSs of the first step. However, this is not surprising, as the Ser alkoxide ion (as a strong base, $\text{p}K_{\text{a}}(\text{Ser-OH}) = 14.4$ ²⁹) is known to be a poor leaving group.

Proposed reaction path of GlySer hydrolysis at physiological pH

At physiological pH, GlySer hydrolysis could, in principle, proceed through two competitive mechanisms, involving either a zwitterionic or an anionic species. According to our calculations, activation energy of 29.4 kcal mol⁻¹ is required to initiate the N→O acyl transfer in the **ZW**, while the subsequent **ZW** ester hydrolysis is slightly less energy intensive, with activation energy of 27.8 kcal mol⁻¹. Hence, the N→O acyl shift, or more precisely the Ser $-\text{OH}$ attack on the amide C-atom, should be considered as the rate-determining step of the overall **ZW** hydrolysis. This scenario is relevant when only the

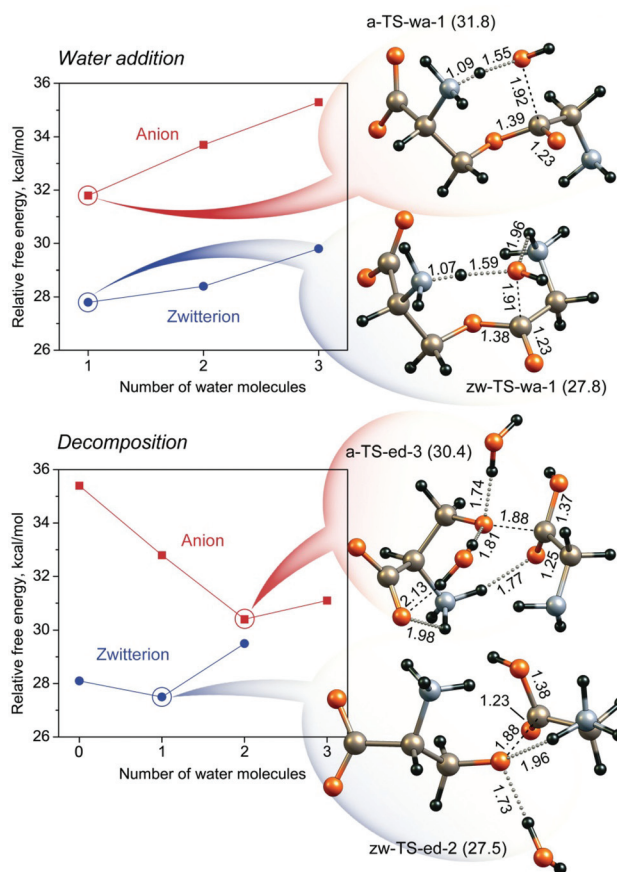


Fig. 4 Molecular structures of the lowest energy transition states of the $-NH_2$ catalyzed water addition and the Td-intermediate decomposition. Energy (in brackets) is in kcal mol $^{-1}$, interatomic distances are in Å. Higher energy structures with different numbers of water molecules are given in the ESI, Fig. S2.†

ZW form of GlySer is present in the reaction solution, as is the case at pD 5.4. At this pH, excellent agreement has been found¹⁵ between the predicted and experimental activation Gibbs function of GlySer hydrolysis by taking into account only the reaction of the **ZW** form.

In the **A** form of GlySer, the $-NH_2$ promoted intramolecular Ser $-OH$ addition on the amide carbon lowers the energy barrier by 2.4 kcal mol $^{-1}$ (to 27.0 kcal mol $^{-1}$) as compared to the $-COO^-$ promoted addition in the **ZW** form (29.4 kcal mol $^{-1}$). The activation energy associated with hydrolysis of the resultant anionic ester (31.8 kcal mol $^{-1}$) is found to be higher by 4.8 kcal mol $^{-1}$ than the barrier of the $N \rightarrow O$ acyl shift. It is also 4.0 kcal mol $^{-1}$ higher than the calculated barrier for **ZW** ester hydrolysis. This means that if **A** would become the predominant form in the reaction medium the rate-limiting step of the overall hydrolysis would rather be the nucleophilic addition of a water molecule on the ester carbonyl group. In order for **A** to become the major form, a pH higher than 10 is required ($pK_a(-NH_3^+) = 8.28$ ¹⁶) where, however, the reaction is largely dominated by the more efficient alkaline hydrolysis mechanism (direct OH^- attack on the ester C-atom). As is well known, the rate constants of intramolecular $-NH_2$ catalyzed

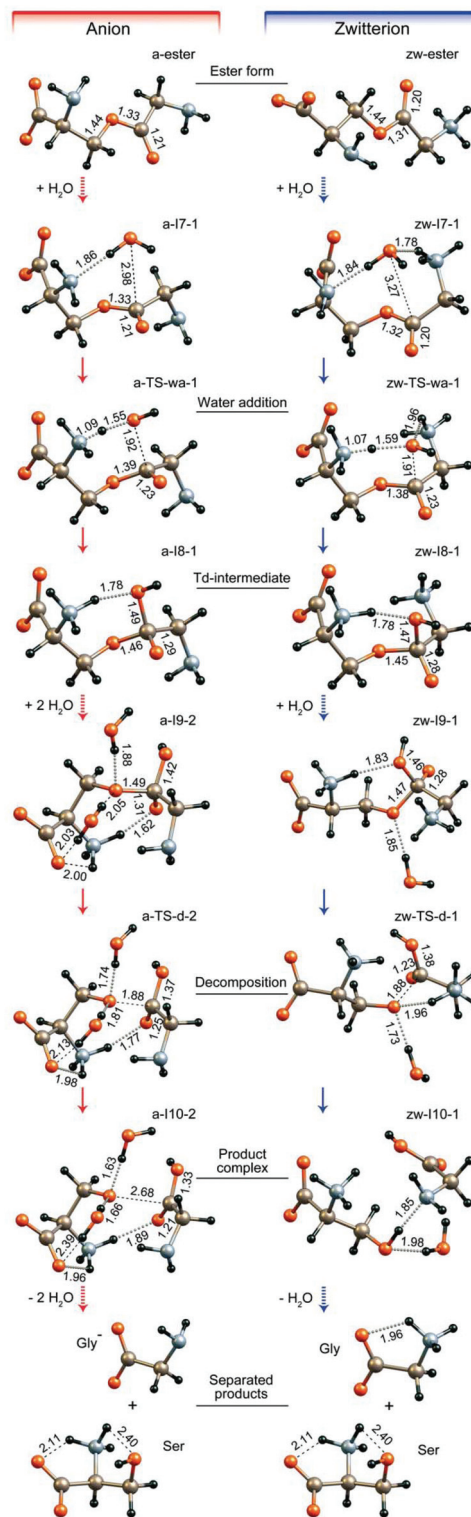


Fig. 5 Proposed ester hydrolysis mechanisms of **A** and **ZW**. Selected interatomic distances are given in Å. The structures obtained by following intrinsic reaction coordinates are connected with a solid arrow. Conformational interconversions²⁷ are indicated with a dashed arrow.

ester hydrolysis are pH-independent in the region from pH 4 to 8.5.²⁸ At values higher than 8.5 the hydrolysis rate constant starts to increase due to the OH^- nucleophilic attack which

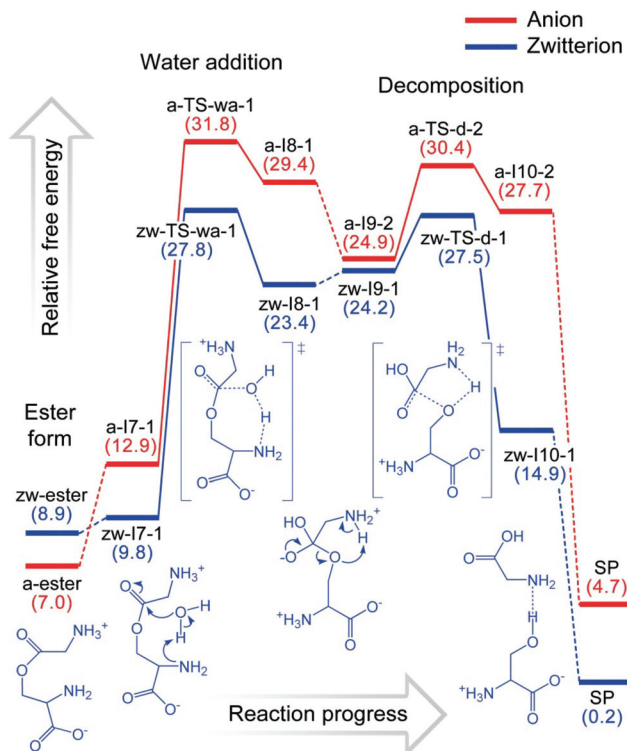


Fig. 6 Free energy profiles (kcal mol^{-1}) of the proposed ester hydrolysis mechanisms of **A** and **ZW**. Selected interatomic distances are given in Å. The energy levels corresponding to structures successively obtained by following the intrinsic reaction coordinate are connected with a solid line. The energy levels of structures linked through conformational interconversion²⁷ are connected with a dashed line. The water molecules in the structural formulas representing the reaction process of **ZW** are omitted for simplicity. For more detailed structures we refer to Fig. 5.

requires a much lower activation energy. Moreover, the predicted rate-determining TS energy of $31.8 \text{ kcal mol}^{-1}$ is significantly higher than the experimental activation Gibbs function

determined at pD 7.4 ($28.7 \text{ kcal mol}^{-1}$).¹⁵ Therefore, it seems evident to suggest that at physiological pH the GlySer ester hydrolysis most probably does not involve the **A** form.

Separate tracing of the reaction paths involving either the **A** or **ZW** form of GlySer only makes sense if one may assume that two forms do not interconvert into each other during the reaction course. At physiological pH, however, the initial **A** and **ZW** forms coexist in equilibrium. In some of the reaction intermediates the terminal Gly $-\text{NH}_2/-\text{NH}_3$ group is free and thus able, in principle, to abstract/release a proton from/to the water solvent under reaction conditions. Among these intermediates, the ester form of GlySer is the most stable. This ester has two amino groups belonging to the Gly and Ser moieties respectively. Unfortunately, their pK_a values have not yet been reported, but one may expect the pK_a value of the Gly $-\text{NH}_3^+$ to be very close to that of the original dipeptide. Hence, it is reasonable to assume that at a slightly basic pH the **A** and **ZW** ester forms could also coexist in equilibrium. This then also means that the reaction system could find the most favorable path for both the $\text{N} \rightarrow \text{O}$ acyl transfer and the ester hydrolysis by switching between the **A** and the **ZW** mechanisms.

Given the above, our suggestion is that GlySer hydrolysis at a physiological pH most likely proceeds *via* a mixed **A-ZW** mechanism, with the first stage of the $\text{N} \rightarrow \text{O}$ acyl transfer involving the **A** form with an activation energy of $27.0 \text{ kcal mol}^{-1}$ (a-TS1-2), whereas the second stage of ester hydrolysis should involve the **ZW** form with an activation energy of $27.8 \text{ kcal mol}^{-1}$ (zw-TS-wa-1). As the computed activation Gibbs energies for the two stages are almost equal, it is hard to judge which of the two is rate-determining. Therefore, to support our hypothesis and to identify the character of the rate-limiting step we have compared the predicted activation enthalpies and entropies with the experimentally obtained ones. Selected, calculated and experimental data are given in Table 1. As continuum solvation models only provide free energies of solvation,

Table 1 Calculated activation parameters (energies in kcal mol^{-1} , entropy in $\text{cal mol}^{-1} \text{K}^{-1}$) of the most relevant transition states along the reaction of GlySer hydrolysis

<i>n</i>	A/ZW	Transition state	ΔH^\ddagger	$\Delta\Delta G_{\text{sol}}^\ddagger$	$\Delta H_{\text{aq}}^\ddagger$ ^a	ΔS^\ddagger	$\Delta S^{*\ddagger}$ ^b	$\Delta G_{\text{aq}}^{*\ddagger}$
<i>Ser -OH addition</i>								
1	ZW	zw-TS1-1	-15.1	36.8	21.6	-40.2	-25.9	29.4
0	ZW	zw-TS1 ^c	2.7	25.6	28.3	-8.0	-8.0	30.7
2	A	a-TS1-2	28.1	-14.0	14.1	-72.2	-43.5	27.0
<i>Water addition</i>								
1	ZW	zw-TS-wa-1	-5.9	25.0	19.2	-43.3	-29.0	27.8
1	A	a-TS-wa-1	31.7	-8.0	23.7	-41.5	-27.2	31.8
<i>Decomposition</i>								
2 ^e	ZW	zw-TS-d-1	5.6	11.4	17.0	-63.7	-35.0	27.5
3 ^e	A	a-TS-d-2	5.0	6.0	11.0	-108.1	-65.1	30.4
<i>Experiment^d</i>								
		pD 5.4			25.6		-12.7	29.4
		pD 7.4			20.7		-27.0	28.7

^a This is not strictly speaking an aqueous phase activation enthalpy – it is the gas phase activation enthalpy, corrected by $\Delta\Delta G_{\text{sol}}^\ddagger$. ^b $\Delta S^{*\ddagger} = \Delta S^\ddagger - n(\Delta S^{0 \rightarrow * \ddagger} + S^{0 \rightarrow \text{conc}})$. ^c TS structure without explicit waters shown in Fig. S1 in the ESI. ^d Experimental values are taken from ref. 15. ^e Unlike in Fig. 4 (*Decomposition*), where *n* corresponds to the number of explicit water molecules microsolvating the solute, here *n* corresponds to the number of water molecules used in the calculation of the relative to the initial reactants, TS energy contributions are according to eqn (4). Note that in the decomposition step one water molecule is already part of the solute.

not separate enthalpies and entropies, it is impossible to reveal the contributions of the latter two. However, if one may assume that the dominant effect of solvation is enthalpic, relative solvation free energies $\Delta\Delta G_{\text{sol}}^{\ddagger}$ may be used to obtain a good estimate of the activation enthalpy in solution $\Delta H_{\text{aq}}^{\ddagger}$ from the corresponding number in gas phase ΔH^{\ddagger} .³⁰

As seen from Table 1, the calculated activation enthalpy (19.2 kcal mol⁻¹) and entropy (-29.0 cal mol⁻¹ K⁻¹) of the -NH₂ catalyzed water addition to the ester carbonyl group of the **ZW** (zw-TS-d-1) are in very good agreement with the corresponding experimental data at pD 7.4 (20.7 kcal mol⁻¹ and -27.0 cal mol⁻¹ K⁻¹ respectively). In contrast, the computed values for the -NH₂ promoted Ser -OH addition of **A** (a-TS1-2) are significantly different from the experiment (14.1 kcal mol⁻¹ and -43.5 cal mol⁻¹ K⁻¹ respectively). Based on this, one may conclude that the overall reaction rate of GlySer hydrolysis at a physiological pH is most probably determined by the hydrolysis of the **ZW** ester form.

In order to gather more information about the thermodynamics and kinetics of the ester hydrolysis process itself, so as to be able to compare our results with analogous processes of other ester compounds, the calculated energy profiles have to be reevaluated with respect to the ester form rather than to the initial reactants. The lowest energy ester conformations of **A** and **ZW** were found to be the a-ester and zw-ester structures depicted in Fig. 5. Their free energies were found to be higher than those of the initial peptide forms by 7.0 kcal mol⁻¹ for **A** and 8.9 kcal mol⁻¹ for **ZW**. When using these structures as the reference the calculated energy barriers for -NH₂ assisted H₂O addition at the ester C=O group in **A** and in **ZW** become 24.8 and 18.9 kcal mol⁻¹ respectively. Rate constants corresponding to activation Gibbs energies of 23–24 kcal mol⁻¹ have previously been reported for intramolecular -NH₂ catalyzed hydrolysis of 2-aminobenzoic acid esters.²⁸ That the TS energy of **ZW** is significantly lower (by 5.9 kcal mol⁻¹) than for **A** might be attributed to the assistance of the Gly -NH₃⁺ group in stabilizing the activated complex through the formation of a H...O hydrogen bond with the nascent OH⁻ nucleophile. The amino group is otherwise inactive in the corresponding TS of **A** or even absent in the quoted 2-aminobenzoic acid esters.

It is noteworthy that despite the excellent agreement between the experimental activation Gibbs energy determined at pD 5.4 and the calculated $\Delta G_{\text{aq}}^{*\ddagger}$ value for the Ser -OH addition of **ZW** (zw-TS1-1) the predicted $\Delta H_{\text{aq}}^{\ddagger}$ and $\Delta S^{*\ddagger}$ terms differ considerably from the experiment (Table 1). The entropic cost of the actual transition state (-12.7 cal mol⁻¹ K⁻¹) is smaller than the calculated value for the zw-TS1-1 model (-25.9 cal mol⁻¹ K⁻¹) involving one explicit water molecule and closer to that found for the TS model without explicit water, zw-TS1. The latter, however, has a slightly higher $\Delta G_{\text{aq}}^{*\ddagger}$ value of 30.7 kcal mol⁻¹. Its molecular structure is depicted in Fig. S1 in the ESI.† Thus, it might be suggested that the -COO⁻ catalyzed Ser -OH addition of **ZW**, which is essentially a unimolecular reaction, might not necessarily require the specific assistance of a solvent water molecule.

Conclusions

Based on our computations, we suggest that under a physiological pH and in a slightly basic water environment GlySer hydrolysis proceeds through a mixed mechanism involving both the zwitterionic and anionic species. At this pH both forms are assumed to readily interconvert into each other, as they are in the peptide form and in the ester form. More specifically, we suggest that the first reaction stage, the N→O acyl shift, preferentially involves anionic GlySer whereas the second stage, ester hydrolysis, most likely involves zwitterionic GlySer. In the anionic form the Ser -OH nucleophilic addition on the amide carbon is predicted to proceed more easily as compared to the zwitterionic form. In **A**, it is catalyzed by the Gly -NH₂ group ($\Delta G_{\text{aq}}^{*\ddagger} = 27.0$ kcal mol⁻¹) while in **ZW** it is promoted by the Ser -COO⁻ group ($\Delta G_{\text{aq}}^{*\ddagger} = 29.4$ kcal mol⁻¹). In contrast, the second hydrolytic stage is predicted to proceed more easily through the zwitterionic form ($\Delta G_{\text{aq}}^{*\ddagger} = 27.8$ kcal mol⁻¹) rather than *via* the anionic form ($\Delta G_{\text{aq}}^{*\ddagger} = 31.8$ kcal mol⁻¹). Therefore, we suspect that the anionic ester intermediate abstracts a proton from the solvent and that the actual ester hydrolysis involves the zwitterionic species. In such a mixed **A-ZW** reaction scheme the N→O acyl transfer has to overcome a barrier of 27.0 kcal mol⁻¹ (**A**) to give an ester intermediate while the ester hydrolysis has to acquire 27.8 kcal mol⁻¹ (**ZW**) to give the products. Although the calculated Gibbs energies are quite similar, a comparison between the estimated and the experimental activation enthalpies and entropies pointed out that the second stage, ester hydrolysis, is a rate-determining step.

Acknowledgements

We acknowledge the Flemish Science Foundation (FWO) for the financial support under project G.0260.12. All calculations were performed on the HPC cluster VIC of the KU Leuven.

References

- 1 F. B. Perler, M.-Q. Xu and H. Paulus, Protein Splicing and Autoproteolysis Mechanisms, *Curr. Opin. Chem. Biol.*, 1997, **1**, 292–299.
- 2 C. J. Noren, J. Wang and F. B. Perler, Dissecting the Chemistry of Protein Splicing and its Applications, *Angew. Chem., Int. Ed.*, 2000, **39**(3), 450–466.
- 3 P. D. van Poelje and E. E. Snell, Pyruvoyl-Dependent Enzymes, *Annu. Rev. Biochem.*, 1990, **59**, 29–59.
- 4 (a) J. A. Porter, D. P. v. Kessler, S. C. Ekker, K. E. Young, J. J. Lee, K. Moses and P. A. Beachy, The Product of Hedgehog Autoproteolytic Cleavage Active in Local and Long-Range Signalling, *Nature*, 1996, **374**, 363–366; (b) J. A. Porter, K. E. Young and P. A. Beachy, Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development, *Science*, 1996, **274**, 255–259.

- 5 J. A. Brannigan, G. Dodson, H. J. Duggleby, P. C. E. Moody, J. L. Smith, D. R. Tomchick and A. G. Murzin, A Protein Catalytic Framework With an N-Terminal Nucleophile is Capable of Self-Activation, *Nature*, 1995, **378**, 416–419.
- 6 A. E. Hodel, M. R. Hodel, E. R. Griffis, K. A. Hennig, G. A. Ratner, S. Xu and M. A. Powers, The Three-Dimensional Structure of the Autoproteolytic, Nuclear Pore-Targeting Domain of the Human Nucleoporin Nup98, *Mol. Cell*, 2002, **10**(2), 347–358.
- 7 (a) A. Sandberg, D. G. A. Johansson and B. Macao, SEA Domain Autoproteolysis Accelerated by Conformational Strain: Energetic Aspects, *J. Mol. Biol.*, 2008, **377**, 1117–1129; (b) D. G. A. Johansson, B. Macao, A. Sandberg and T. Härd, SEA Domain Autoproteolysis Accelerated by Conformational Strain: Mechanistic Aspects, *J. Mol. Biol.*, 2008, **377**, 1130–1143.
- 8 H. Paulus, The Chemical Basis of Protein Splicing, *Chem. Soc. Rev.*, 1998, **27**, 375–386.
- 9 K. Iwai and T. Ando, N → O Acyl Rearrangement, *Methods Enzymol.*, 1967, **11**, 263–282.
- 10 (a) K. Michalska, A. Hernandez-Santoyo and M. Jaskolski, The Mechanism of Autocatalytic Activation of Plant-type L-Asparaginases, *J. Biol. Chem.*, 2008, **283**(19), 13388–13397; (b) Z. Du, P. T. Shemella, Y. Liu, S. A. McCallum, B. Pereira, S. K. Nayak, G. Belfort, M. Belfort and C. Wang, Highly Conserved Histidine Plays a Dual Role in Protein Splicing: A pK_a Shift Mechanism, *J. Am. Chem. Soc.*, 2009, **131**, 11581–11589; (c) Z. Du, Y. Zheng, M. Patterson, Y. Liu and C. Wang, pK_a Coupling at the Intein Active Site: Implications for the Coordination Mechanism of Protein Splicing with a Conserved Aspartate, *J. Am. Chem. Soc.*, 2011, **133**, 10275–10282.
- 11 Q. Xu, D. Buckley, C. Guan and H.-C. Guo, Structural Insights into the Mechanism of Intramolecular Proteolysis, *Cell*, 1999, **98**(5), 651–661.
- 12 (a) T. Klabunde, S. Sharma, A. Telenti, W. R. Jacobs Jr. and J. C. Sacchettini, Crystal Structure of GyrA Intein From *Mycobacterium Xenopi* Reveals Structural Basis of Protein Splicing, *Nat. Struct. Biol.*, 1998, **5**(1), 31–36; (b) B. W. Poland, M.-Q. Xu and F. A. Quiocho, Structural Insights into the Protein Splicing Mechanism of PI-SceI, *J. Biol. Chem.*, 2000, **275**, 16408–16413.
- 13 (a) D. G. A. Johansson, G. Wallin, A. Sandberg, B. Macao, J. Åqvist and T. Härd, Protein Autoproteolysis: Conformational Strain Linked to the Rate of Peptide Cleavage by the pH Dependence of the N→O Acyl Shift Reaction, *J. Am. Chem. Soc.*, 2009, **131**, 9475–9477; (b) G. Wallin, T. Härd and J. Åqvist, Folding-Reaction Coupling in a Self-Cleaving Proteins, *J. Chem. Theory Comput.*, 2012, **8**, 3871–3879.
- 14 (a) P. H. Ho, K. Stroobants and T. N. Parac-Vogt, Hydrolysis of Serine-Containing Peptides at Neutral pH Promoted by [MoO₄]²⁻ Oxyanion, *Inorg. Chem.*, 2011, **50**, 12025–12033; (b) P. H. Ho, T. Mihaylov, K. Pierloot and T. N. Parac-Vogt, Hydrolytic Activity of Vanadate toward Serine-Containing Peptides Studied by Kinetic Experiments and DFT Theory, *Inorg. Chem.*, 2012, **51**, 8848–8859.
- 15 T. T. Mihaylov, T. N. Parac-Vogt and K. Pierloot, A Mechanistic Study of the Spontaneous Hydrolysis of Glycylserine as the Simplest Model for Protein Self-Cleavage, *Chem.–Eur. J.*, 2014, **20**, 456–466.
- 16 V. B. Naidoo and M. Sankar, Thermodynamic Constants of Glycyl Peptides: Solvation and Buffers for the Physiological pH Range, *J. Solution Chem.*, 1998, **27**(6), 553–567.
- 17 L. Bethencourt and O. Nunez, Serine Protease Mechanism-Based Mimics. Direct Evidence for a Transition State Bridge Proton in Stable Potentials, *J. Org. Chem.*, 2008, **73**, 2105–2113.
- 18 R. B. Sunoj and M. Anand, Microsolvated Transition State Models for Improved Insight into Chemical Properties and Reaction Mechanisms, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12715–12736.
- 19 (a) J. R. Pliego Jr. and J. M. Riveros, A Theoretical Analysis of the Free-Energy Profile of the Different Pathways in the Alkaline Hydrolysis of Methyl Formate in Aqueous Solution, *Chem.–Eur. J.*, 2002, **8**(8), 1945–1953; (b) G. I. Almerindo and J. R. Pliego, Ab Initio Investigation of the Kinetics and Mechanism of the Neutral Hydrolysis of Formamide in Aqueous Solution, *J. Braz. Chem. Soc.*, 2007, **18**(4), 696–702.
- 20 Y. Zhao and D. G. Truhlar, The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
- 21 A. V. Marenich, C. J. Cramer and D. G. Truhlar, Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.
- 22 R. F. Ribeiro, A. V. Marenich, C. J. Cramer and D. G. Truhlar, Use of Solution-Phase Vibrational Frequencies in Continuum Models for the Free Energy of Solvation, *J. Phys. Chem. B*, 2011, **115**(49), 14556–14562.
- 23 J. R. Pliego Jr. and J. M. Riveros, The Cluster-Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species, *J. Phys. Chem. A*, 2001, **105**, 7241–7247.
- 24 V. S. Bryantsev, M. S. Diallo and W. A. Goddard III, Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models, *J. Phys. Chem. B*, 2008, **112**, 9709–9719.
- 25 J. A. Keith and E. A. Carter, Quantum Chemical Benchmarking, Validation and Prediction of Acidity Constants for Substituted Pyridinium Ions and Pyridinyl Radicals, *J. Chem. Theory Comput.*, 2012, **8**, 3187–3206.
- 26 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta,

- F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 (Revision A.02)*, Gaussian, Inc., Wallingford CT, 2009.
- 27 In order to find the most probable reaction mechanism our efforts were mainly focused at finding the lowest energy TSs, involving bond breaking and/or bond formation, which might occur during the hydrolysis reaction. Every TS found bridges between two local minima (intermediate structures) obtained by following the intrinsic reaction coordinate in the forward and in the reverse direction respectively. Thus, the two local minima between two consecutive TSs (each of which is derived either from one or from the other TS) are usually in slightly different conformations and to be linked in the reaction path the conformational interconversion should in principle also be tracked. However, as they involve relatively unhindered single bond rotations we may safely assume that the energy barriers of conformational interconversion are smaller than the barriers associated with bond breaking/formation and will not affect the general picture of the reaction energy profile.
- 28 T. H. Fife, R. Singh and R. Bembi, Intramolecular General Base Catalyzed Ester Hydrolysis. The Hydrolysis of 2-Aminobenzoate Esters, *J. Org. Chem.*, 2002, **67**, 3179–3183.
- 29 E. Bottari, D. Cellulosi and M. R. Festa, On the Acid Behaviour of the Alcoholic Group of Serine, *Talanta*, 1999, **50**, 993–1002.
- 30 J. N. Harvey, Ab initio Transition State Theory for Polar Reactions in Solution, *Faraday Discuss.*, 2010, **145**, 487–505.