J|A|C|S ARTICLES

Published on Web 09/02/2009

Peptide Synthesis in Aqueous Environments: The Role of Extreme Conditions on Peptide Bond Formation and Peptide Hydrolysis

Eduard Schreiner,*,† Nisanth N. Nair,‡ and Dominik Marx

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

Received April 23, 2009; E-mail: eschrein@ks.uiuc.edu

Abstract: The mechanisms and free energetics underlying the formation of peptides from α-amino acids and α-amino acid *N*-carboxyanhydrides (NCAs) in bulk water at both ambient and extreme temperature and pressure conditions were investigated using accelerated *ab initio* molecular dynamics. In particular, peptide bond formation using an activated amino acid in form of its NCA, subsequent decarboxylation, as well as hydrolysis of the formed peptide were studied using glycine. It is shown to what extent thermodynamic conditions affect the reaction mechanisms qualitatively and the energetics quantitatively in solution. In particular, the zwitterionic intermediate in the peptidization step found in ambient water degenerates into a transient species in hot-pressurized water, whereas the hydrolysis reaction is found to follow qualitatively different pathways at ambient and extreme conditions. The work also quantifies the impact of extreme solvent conditions on both peptide bond formation and peptide hydrolysis in aqueous media. Beyond the specific case, the results provide important insights into how elevated temperatures and increased pressures affect organic reactions in aqueous solutions.

1. Introduction

The polymerization of amino acids into peptides and proteins is of outstanding importance not only in biology but also in technological applications. For decades, α-amino acid *N*-carboxyanhydrides (NCA), also called "Leuchs anhydrides",[1] are used in the stepwise synthesis of high-molecular weight polypeptides.[2,3] NCAs can be considered as the activated form of amino acids with the merits of a reactive CO group and a simultaneously protected amino site.[3] New developments in the controlled polypeptide synthesis using NCAs[4] have increased the scope of these compounds in the industrial production of polypeptide hybrid copolymers, which will also be useful in various biomedical applications.[5] Comprehensive reviews on NCA chemistry and its broad applications in various fields are available in refs 3−7 and in Kricheldorf's book.[2]

Early kinetic studies[8] on the polymerization of NCAs at aqueous conditions have demonstrated the efficacy of amino acid-initiated polymerization.[9] However, synthesis of polypeptides from NCAs in aqueous solution is complicated by several side reactions, a major one being hydrolysis.[8] Later, experimental works[10,11] have shown that under controlled experimental conditions NCAs can be used to productively synthesize peptides in aqueous solutions of amino acids. The ring-opening of NCAs during the reaction seems to be catalyzed by bases,[8] and high pH conditions are necessary for the formation of noncharged amino acids or peptide end groups for subsequent reactions using NCA. But basic conditions can trigger other side reactions like hydrolysis and the formation of hydantoic acids calling for a delicate control of experimental conditions.[11] In this respect, our recent work[12] has shown that it is possible to increase the concentration of the noncharged form of amino acids at *neutral pH conditions* by increasing temperature and pressure. Thus, working at appropriately chosen extreme thermodynamic conditions might open up novel opportunities to improve the productivity of peptide synthesis using NCAs instead of working at basic aqueous conditions.

Currently, water at extreme thermodynamic conditions is gaining an increasing attention in the field of organic chemistry since many compounds tend to react very differently compared to water at ambient conditions. In particular, chemical reactions can be accelerated efficiently under such conditions without

† Present Address: Theoretical and Computational Biophysics Group, Beckman Institute, Urbana, IL 61801.

‡ Present Address: Department of Chemistry, Indian Institute of Technology, Kanpur 208016, India.

(1) Leuchs, H. *Ber. Dtsch. Chem. Ges.* **1906**, *39*, 857–861.
(2) Kricheldorf, H. R. *α-Aminoacid-N-Carboxy-Anhydrides and Related Hetrocycles*; Springer-Verlag: Berlin Heidelberg, 1987; pp 3−157.
(3) Kricheldorf, H. R. *Angew. Chem., Int. Ed.* **2006**, *45*, 5752–5784.
(4) Deming, T. J. *Adv. Polym. Sci.* **2006**, *202*, 1–18.
(5) Deming, T. J. *Prog. Polym. Sci.* **2007**, *32*, 858–875.
(6) Szwarc, M. *Adv. Polym. Sci.* **1965**, *4*, 1–65.
(7) Pascal, R.; Boiteau, L.; Commeyras, A. *Top. Curr. Chem.* **2005**, *259*, 69–122.
(8) Bartlett, P. D.; Richard, H. J. *J. Am. Chem. Soc.* **1957**, *79*, 2153–2159.
(9) Bartlett, P. D.; Dittmer, D. C. *J. Am. Chem. Soc.* **1957**, *79*, 2159–2160.
(10) Denkewalter, R. G.; Schwam, H.; Strachan, R. G.; Beesley, T. E.; Veber, D. F.; Schoenewaldt, E. F.; Barkemeyer, H.; Paleveda, W. J., Jr; Jacob, T. A.; Hirschmann, R. *J. Am. Chem. Soc.* **1966**, *88*, 3163–3164.
(11) Hirschmann, R.; Strachan, R. G.; Schwam, H.; Schoenewaldt, E. F.; Joshua, H.; Barkemeyer, B.; Veber, D. F.; Paleveda, W. J., Jr; Jacob, T. A.; Beesley, T. E.; Denkewalter, R. G. *J. Org. Chem.* **1967**, *32*, 3415–3425.
(12) Nair, N. N.; Schreiner, E.; Marx, D. *J. Am. Chem. Soc.* **2008**, *130*, 14148–14160.

10.1021/ja9032742 CCC: $40.75

using any catalyst at all.[13−17] Moreover, it is interesting to note that the dielectric constant of water at its critical point (T_c = 647.1 K and p_c = 22.1 MPa) is only $\varepsilon \approx 6$ instead of about 80 at room temperature and pressure.[15] Thus, water at such conditions behaves more like a protic organic solvent with a high ionic product and a low viscosity.[15] Detailed reviews on organic reactions in subcritical water at extreme thermodynamic conditions as well as supercritical water can be found in refs 14, 16, 18, and 19.

Furthermore, using NCA-activated amino acids does not lead to racemization at the chiral centers during polymerization, which is an interesting feature of NCA-based peptide synthesis.[2] It is noted in passing that such stereoselective polymerization reactions are hypothesized by some authors to have played a role in the origin of homochirality in biology.[20,21] On the basis of this stereoselectivity, together with the fact that NCAs can be synthesized by reactions between amino acids and simple "inorganic molecules" like NCO-, NO/O_2,[22,23] or COS,[24,25] it has been suggested and demonstrated that the rich NCA-chemistry can be relevant for prebiotic chemistry as well.[7,24−26] At this point, it is important to recall that NCAs have been proposed not only to be key components in peptide formation[24,25] as such, but also in nucleic acid synthesis[7,26] depending on suitable prebiotic conditions including extreme temperatures and pressures.

Recently, we have studied a sequence of reactions leading to the formation of NCA in aqueous solution at ambient and extreme thermodynamic conditions.[12] The main observation was that in hot pressurized water charged molecular species are destabilized with respect to their neutral counterparts. Accordingly, reaction mechanisms were observed to change at extreme thermodynamic conditions in a way avoiding charged intermediates. Additionally, all of the studied reactions were found to occur faster at extreme conditions, mainly due to the high temperature. However, also additional reaction channels were observed, for example, for the interconversion of neutral and zwitterionic glycine. These results have important implications for various synthetic and practical applications.[17,27]

After having clarified the influence of extreme conditions on NCA synthesis itself, the present work aims at extending this line of research to the next step, which is to investigate peptide formation using the synthesized NCA and available amino acids at aqueous reaction conditions. Concerning this process several mechanisms have been proposed for the reaction of an amino acid or peptide with an NCA, among them a "carbamate mechanism", "activated monomer mechanism", and a "zwitterionic mechanism" as reviewed in ref 2. Detailed analyses of the mechanism of this "elongation" reaction conducted by experimental techniques, however, strongly suggest the "amine mechanism" as the preferred one for primary-amine initiated polymerizations, which originally was put forward by Wessley and co-workers.[2,28] It involves a ring-opening chain growth after the nucleophilic attack of the amino group of amino acids or polypeptides on NCA. Based on the reaction rates in various solvents, the transition state is assumed to be a noncharged species,[2] although there also are predictions of a zwitterionic form of the intermediate structure.[29] The rate determining step of this reaction seems to be the nucleophilic attack of the amino group and not the subsequent ring-opening.[2] Both the free energy barrier for the initial nucleophilic attack as well as the nature of the transition state are expected to depend on the solvent environment. These important insights, however, are completely lacking up to this day.

Due to recent developments of novel simulation methods[30] in combination with advances in computer technology it is now possible to systematically and directly study such complex chemical reactions in condensed phases fully *in silico*. In particular numerical simulations have been shown to allow to dissect reaction mechanisms[31−33] and to estimate the reaction rates at various conditions. Especially, high temperature and high pressure conditions can be easily modeled in computer simulations,[34,35] while it is still challenging in experiments. Molecular dynamics simulations with concurrent electronic structure calculations to compute the interactions, that is, *ab initio* molecular dynamics,[30] offer several simulation techniques in order to study chemical reactions in solution, especially when solvent molecules need to be treated explicitly as reactive species which is often the case for water.[31] Different methodologies can be coupled with such *ab initio* molecular dynamics simulations in order to study "rare events" and thereby extracting free energies and reaction mechanism. Like in our previous investigations[12,36] we use the metadynamics sampling technique[37] which can be combined with Car−Parrinello molecular dynamics[38] in an efficient way; see refs 30, 39, and 40 for reviews. The technique has been shown to be reasonably

(13) Nagai, Y.; Morooka, S.; Matubayasi, N.; Nakahara, M. *J. Phys. Chem. A* **2004**, *108*, 11635–11643.
(14) Akiya, N.; Savage, P. E. *Chem. Rev.* **2002**, *102*, 2725–2750.
(15) Weingärtner, H.; Frank, E. U. *Angew. Chem., Int. Ed.* **2005**, *44*, 2672–2692.
(16) Li, C.-J. *Chem. Rev.* **2005**, *105*, 3095–3165.
(17) Sheldon, R. A. *Green Chem.* **2005**, *7*, 267–278.
(18) Simoneit, B. R. T. *Origins Life Evol. Biospheres* **1995**, *25*, 119–140.
(19) Katritzky, A. R.; Nichols, D. A.; Siskin, M.; Murugan, R.; Balasubramanian, M. *Chem. Rev.* **2001**, *101*, 837–892.
(20) Brack, A. *Origins Life* **1987**, *17*, 367–379.
(21) Hitz, T.; Luisi, P. L. *Helv. Chim. Acta* **2003**, *86*, 1423–1434.
(22) Collet, H.; Bied, C.; Mion, L.; Taillades, J.; Commeyras, A. *Tetrahedron Lett.* **1996**, *37*, 9043–9046.
(23) Commeyras, A.; Collet, H.; Boiteau, L.; Taillades, J.; Vandenabeele-Trambouze, O.; Cottet, H.; Biron, J.-P.; Plasson, R.; Mion, L.; Lagrille, O.; Martin, H.; Selsis, F.; Dobrijevic, M. *Polym. Int.* **2002**, *51*, 661–665.
(24) Leman, L.; Orgel, L.; Ghadiri, M. R. *Science* **2004**, *306*, 283–286.
(25) Huber, C.; Wächtershäuser, G. *Science* **1998**, *281*, 670–672.
(26) Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2006**, *128*, 20–21.
(27) Hailes, H. C. *Org. Process Res. Dev.* **2007**, *11*, 114–120.
(28) Wessely, F.; Riedl, K.; Tuppy, H. *Monatsh. Chem.* **1950**, *81*, 861–872.
(29) Ballard, D.; Bamford, C. *Proc. R. Soc. Lond, Ser A* **1954**, *223*, 495–520.
(30) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, 2009.
(31) Stubbs, J. M; Marx, D. *Chem.—Eur. J.* **2005**, *11*, 2651–2659.
(32) Blumberger, J.; Ensing, B.; Klein, M. L. *Angew. Chem., Int. Ed.* **2006**, *45*, 2893–2897.
(33) Michel, C.; Laio, A.; Mohamed, F.; Krack, M; Parrinello, M.; Milet, A. *Organometallics* **2007**, *26*, 1241–1249.
(34) Cavazzoni, C.; Chiarotti, G. L.; Scandolo, S.; Tosatti, E.; Bernasconi, M.; Parrinello, M. *Science* **1999**, *283*, 44–47.
(35) Schreiner, E.; Nair, N. N.; Marx, D. *J. Am. Chem. Soc.* **2008**, *130*, 2768–2770.
(36) Nair, N. N.; Schreiner, E.; Marx, D. *J. Am. Chem. Soc.* **2006**, *128*, 13815–13826.
(37) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566.
(38) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, . 238302−1-4.
(39) Ensing, B.; Vivo, M. D.; Liu, Z.; Moore, P.; Klein, M. L. *Acc. Chem. Res.* **2006**, *39*, 73–81.
(40) Laio, A.; Parrinello, M. Computing Free Energies and Accelerating Rare Events with Metadynamics. In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*; Ferrario, M., Ciccotti, G., Binder, K., Eds.; Springer Verlag: Berlin Heidelberg, 2006; Vol. 1, pp 315−347.

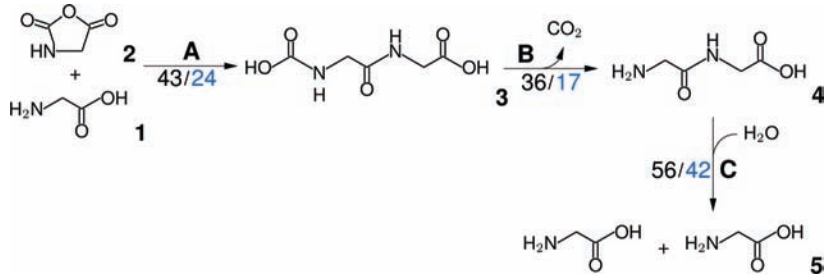

Figure 1. Investigated mechanism of peptide bond formation by the reaction of α-amino acid *N*-carboxyanhydride (NCA) **2** and glycine **1**, and the hydrolysis of the formed diglycine **4**. The numbers at the arrows correspond to effective free energy barriers in units of k_BT for the corresponding reaction at ABW (black) and HPW (blue, after /) thermodynamic conditions.

accurate[12] for the present purpose (error around 1 k_BT) and since the same methodology is used the results from the two studies can be directly compared.

Employing these techniques, the present study focuses on the general question of *formation of peptides from amino acids and NCAs in aqueous environment* using glycine as the simplest example as sketched in Figure 1. In particular, the emphasis is on the influence of thermodynamic conditions on the studied reactions, that is, the effect on going from ambient water to hot-pressurized water. As a major back reaction in the synthesis of peptides we are also considering hydrolysis. Overall, a sequence of three chemical reactions (see **A**, **B**, and **C** in Figure 1) has been investigated in condensed phase environments, each at two thermodynamic conditions.

2. Methods and Models

2.1. Simulations and Free Energy Sampling. The general simulation protocol is identical to that used in our previous work on NCA synthesis.[12] All calculations were performed using *ab initio* molecular dynamics[30] using the efficient Car−Parrinello propagation.[41] The electronic structure was described by spin-restricted Kohn−Sham density functional theory in its plane-wave/pseudopotential formulation as implemented in the CPMD program package.[42,43] Using ultrasoft pseudopotentials,[44] a plane-wave cutoff of 25 Ry is sufficient to achieve reasonable accuracy with the chosen PBE[45,46] exchange-correlation functional. Molecular dynamics simulations were performed using a time step of 0.145 fs and the fictitious mass for the orbitals was set to 700 au as before; all hydrogen masses in the system were substituted by deuterium masses, allowing a larger time step.[30] Nosé−Hoover chain[47] thermostats were used for both nuclei and electronic orbitals.

In this work, the free energy landscapes and the mechanisms of all reaction steps leading from the reaction of glycine-NCA with glycine up to the formation of diglycine and its hydrolysis (see Figure 1) are studied in aqueous bulk environments at two thermodynamic conditions: ambient bulk water (ABW) and hot-pressurized bulk water (HPW). This corresponds to a total of three independent reaction steps to be investigated or a total of six distinct simulations including two thermodynamic conditions. The two different thermodynamic conditions were established by choosing the particle number density based on the experimental equation of state[48] and performing canonical ensemble calculations with the periodicity applied on the simulation supercell. For the ABW and HPW conditions, the supercells used were orthorhombic, having the fixed dimensions of 10.8 × 10.8 × 9.5 Å^3 and 10.8 × 10.8 × 10.8 Å^3, respectively, each containing 36 water molecules. The density of water is ∼1.00 g/cm^3 and $p \approx 0.1$ MPa for the ABW case at the thermostatted temperature of 300 K. In the case of HPW conditions, the temperature was thermostatted to 500 K and the density of water is ∼0.85 g/cm^3, resulting into $p \approx 20$ MPa according to the experimental equation of state.[48]

The starting configurations for each simulations were taken either from the final structure of the previous simulations, or by the following equilibration protocol. From a pre-equilibrated water system at any thermodynamic condition, one water molecule was replaced by the solute and equilibrated at the target temperature for 1 ps keeping the atomic positions of the solute fixed. Eventually, all constraints were removed and all (Cartesian) degrees of freedom were thermostatted individually for about 3 ps. For the subsequent metadynamics simulations all the nuclei were coupled to one thermostat only. In total, about one nanoseconds of *ab initio* metadynamics trajectories were produced for all the calculations presented in this paper.

The metadynamics technique[37] is used for accelerating the reactions, being rare events in view of the large barriers involved, and to extract the reaction mechanism from the computed free energy surfaces. This technique enables one to efficiently sample even multidimensional free energy landscapes simultaneously without providing *a priori* the type of reaction product(s). Essentially a reduction of the dimensionality of the underlying reaction coordinate space is done by an appropriate selection of a suitable set of so-called collective coordinates, $\{S_\alpha(\mathbf{R})\}$, which allow to discriminate reactant and product states and describe the slowest modes in the process of interest.[30,39,40] In the extended Lagrangian formulation of metadynamics[38] used here, a set of auxiliary (fictitious) particles $\{s_\alpha\}$ equal in number to the selected collective coordinates extends the physical degrees of freedom. Moreover, their fictitious motion is coupled to that of the collective coordinates, $\{S_\alpha\}$, by a harmonic potential with coupling constant of strength k_α. A repulsive time-dependent potential $V(t,\mathbf{s})$, which is slowly constructed during the dynamics fills the free energy landscape spanned by the collective coordinates S_α, forcing the system to escape from the minima. Rare-events can thus be accelerated and the information of the added potential can be used to eventually map the underlying free energy surface; see refs.[30,39,40] for reviews and ref.[12] for technical details that are closely related to this work such as the choice and the dynamics of $V(t,\mathbf{s})$.

(41) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
(42) Marx, D.; Hutter, J. Ab initio molecular dynamics: Theory and Implementation. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; John von Neumann Institute for Computing (NIC): Forschungszentrum Jülich, Germany, 2000; Vol. 3, pp 301−449.
(43) CPMD, Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart 1997−2001; see also http://www.cpmd.org.
(44) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892–7895.
(45) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(46) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396.
(47) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
(48) Lown, D. A.; Thirsk, H. R.; Lord, W.-J. *Trans. Faraday Soc.* **1970**, *66*, 51–73. Note: the typographical error in the specific volume of 1.1178 cm^3/g for 200 bar pressure and 225 °C temperature in Table 5 (page 62) is corrected to 1.1778 cm^3/g.

We used three different types of collective coordinates, $S_\alpha(\mathbf{R})$, in this study. One of them is the distance between two atoms A and B, defined as

$$d[\mathrm{A}-\mathrm{B}] = |\mathbf{R}_\mathrm{A} - \mathbf{R}_\mathrm{B}| \tag{1}$$

and the second type is the coordination number[49] c[A − B] between an atom A with respect to a set of other atoms B belonging to the same species

$$c[\mathrm{A}-\mathrm{B}] = \sum_{I\in \mathrm{B}} \frac{1-(R_{\mathrm{A}I}/R^0_{\mathrm{AB}})^p}{1-(R_{\mathrm{A}I}/R^0_{\mathrm{AB}})^{p+q}} \tag{2}$$

with the particular choice $p = 6$ and $q = 6$. Here, $R_{\mathrm{A}I}$ is the distance between atom A and atom I of the set of B atoms and R^0_{AB} is a fixed cutoff parameter based on bond distance between A and B. For each pair of atoms belonging to A and B, the function c[A − B] is nearly unity when the actual bond distance $R_{\mathrm{A}I} < R^0_{\mathrm{AB}}$ and approaches zero rapidly when $R_{\mathrm{A}I} > R^0_{\mathrm{AB}}$. The used cutoff distances R^0_{AB} depend on the atomic species A and B and we chose $R^0_{\mathrm{XH}} =$ 1.4 Å for X = N, O, and S, while $R^0_{\mathrm{CO}} = 1.8$ Å. The third type of collective coordinate is a more general coordination number, c^{tot}[A − B], which measures the total coordination between two sets of atoms A and B and is defined as

$$c^{\mathrm{tot}}[\mathrm{A}-\mathrm{B}] = \sum_{J\in \mathrm{A}}\sum_{I\in \mathrm{B}} \frac{1-(R_{IJ}/R^0_{\mathrm{AB}})^p}{1-(R_{IJ}/R^0_{\mathrm{AB}})^{p+q}} \tag{3}$$

again choosing $p = q = 6$. The specific sets of A and B atoms depend on the reaction studied and will be defined when presenting the results. In order to focus the sampling to the most relevant regions of the free energy landscape spanned by these collective coordinates, repulsive potentials were applied to the auxiliary variables;[50] details about these potentials will be stated when presenting the results. Convergence of the computed free energies were finally studied using different Gaussian heights while keeping all other parameters fixed following the same refinement protocol as introduced in ref.[12]

2.2. Assessing the Electronic Structure Method. Clearly, the quality of the free energy landscapes hinges on the quality of describing the interactions based on the underlying electronic structure method. Total energies and nuclear gradients must be computed on the order of 100 000 to 1 000 000 times using periodic electronic structure calculations in order to sample sufficiently reactant, solvent, and temperature fluctuations of the liquid system. Thus, only most efficient electronic structure methods can be used in practice for such simulations. Here, we continue to use an exchange-correlation functional, PBE,[45,46] which is most efficiently evaluated within a plane-wave/pseudopotential framework being a semilocal functional of the Generalized Gradient Approximation (GGA) family of functionals.[30]

In order to access the error of the used density functional for the particular question and system studied here we have performed a set of single-point MP2 calculations using Dunning's augmented correlation-consistent triple-ζ basis set (aug-cc-pVTZ).[51] In particular, two important reactions have been scrutinized, which is the peptidization step (see **A** in Figure 1) as well as the hydrolysis of the dipeptide (see **C** in Figure 1). In order to be most realistic, seven instantaneous configurations consisting of reactants or products and the most important solvating water molecules have been sampled along the minimum free energy pathways for reactions **A** and **C**. The choice of the number of water molecules, five and three, respectively, was motivated by the specific chemistry

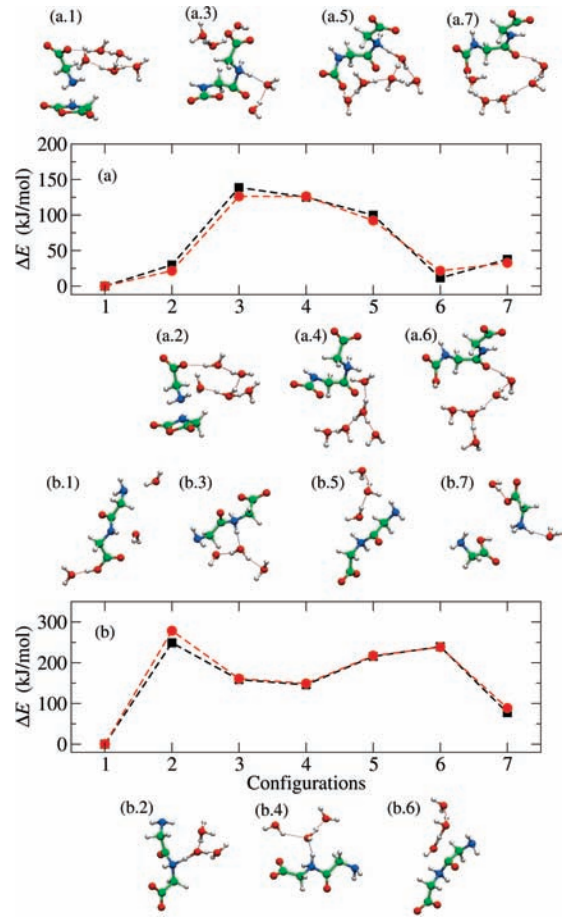

Figure 2. Total electronic energies ΔE (kJ/mol) computed for seven instantaneous configurations sampled along the minimum free energy pathways for reaction step **A** in (a) and step **C** in (b) according to Figure 1; the particular configurations are finite complexes that include five and three water molecules, respectively, as depicted. The energies obtained using PBE (black squares) and MP2 (red dots) are reported relative to the energy of the first configuration; see text for details.

of the system, including sufficient solvation of a hydronium ion, and the total charge of all complexes is zero. For this set of frozen configurations the total electronic energies have been computed using PBE. Here, the technical details are identical to those used for the full metadynamics simulations except for employing cluster boundary conditions[30,53] (with a cubic box of 15 Å) as implemented in CPMD[42,43] in order to yield energies that are directly comparable to the MP2 single-point calculations using the same configurations; the latter have been carried out using the Turbomole package[52] (see www.turbomole-gmbh.com).

The comparison of the total energies relative to first configuration is shown in panels a and b of Figure 2 along pathway **A** and **C**, respectively. It is clear from this comparison that PBE and MP2 energies are in very good agreement along both reaction pathways. The largest deviation is observed for configuration (b.2) where the PBE energy deviates by about 10% from the MP2 reference value. Other than this particular point the maximum deviation was in the order of 10 kJ/mol (i.e., < 4%), which is consistent with a different error estimation performed for a similar system in ref 32. Importantly, the mean-average deviation between PBE and MP2 energies is about 7 kJ/mol along these pathways which has to be judged with respect to relative changes on the scale of 200−300

(49) Sprik, M. *Chem. Phys.* **2000**, *258*, 139–150.
(50) Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. *J. Phys. Chem. B* **2005**, *109*, 6676–6687.
(51) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
(53) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810–2821.
(52) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

kJ/mol. We note in passing that this comparison is necessarily carried out using rather small microsolvated complexes in vacuum. Thus, the variation of the total electronic energy is rather high along both pathways where charge separation is involved; bulk solvation stabilizes such situations in the simulations which yields smaller free energy variations in the condensed phase.

In summary, the data suggest that PBE is able to describe the underlying energetics of both crucial reactions, peptidization and hydrolysis, sufficiently accurate for the present purpose.

3. Results and Discussion

3.1. Formation of the Peptide Bond. The first step to consider is the reaction of NCA **2** with glycine **1** which is route **A** in Figure 1. The chosen subspace to describe this reaction is spanned by three collective coordinates. Since the mechanism is expected to involve a nucleophilic attack of the amino nitrogen of glycine **1**at the C5 carbon atom of the NCA **2**, the distance between these two atoms, $d[N_{Gly} - C5_{NCA}]$, defines the first coordinate. In the resulting peptide group the nitrogen of glycine carries only one hydrogen atom. To capture the proton exchange with the solvent in a unbiased way the coordination number of N_{Gly} to *all* nonaliphatic hydrogen atoms in the system $c[N_{Gly} - H_{na}]$ is introduced. For the attack at C5 the amino group must not be protonated. Thus, in order to prevent an additional protonation of the amino group a repulsive potential was set along $c[N_{Gly} - H_{na}]$ at values larger than 2. Effectively, the repulsive potential decouples the peptide bond formation from the interconversion between the different protonation states of glycine that has already been studied in detail previously.[12] The necessary ring-opening of **2** during the reaction was captured by the third coordinate which is defined to be the distance between the oxygen and carbon atoms within the NCA denoted as $d[O1_{NCA} - C5_{NCA}]$.

The reaction mechanisms produced by *ab initio* metadynamics at ambient (ABW) and extreme (HPW) conditions are shown in Figure 3 together with the corresponding schematic free energy profiles whereas the reconstructed three-dimensional free energy surfaces underlying this mechanism are depicted in Supporting Information (SI) Figure 1. Starting from the reactants **2** and **1** in the first step the nitrogen of the amino group of glycine attacks C5 of NCA **2** resulting in the intermediate **2.1** at ABW conditions. The free energy barrier for this process is about 88 kJ/mol (which corresponds to a thermal energy of 35 k_BT_{300} at $T = 300$ K), while the reverse reaction is associated with an activation energy of about 11 kJ/mol (i.e., 4 k_BT_{300}).

At variance with the scenario observed at ambient conditions, **2.1** is found to be only a transient species, **2.1′**, in hot-pressurized water which thus undergoes a very fast ring-opening reaction resulting into **2.2**. The activation free energy barrier for this sequence is about 74 kJ/mol (18 k_BT_{500}). Furthermore, while the ring-opening appears to be a barrierless process in hot-pressurized water it is associated in ABW with a pronounced barrier of 25 kJ/mol (10 k_BT_{300}). The subsequent deprotonation of **2.2** at the NH_2 moiety yields the *N*-carboxyl dipeptide **3**. The activation free energy barrier for this deprotonation step is 13 kJ/mol (5 k_BT_{300}) at ABW but 32 kJ/mol (8 k_BT_{500}) at HPW conditions. Water is not an inert solvent but plays an active role in this deprotonation step by accepting the detached proton thus forming a hydronium, H_3O^+, which is observed to undergo Grotthuss structural diffusion.[54]

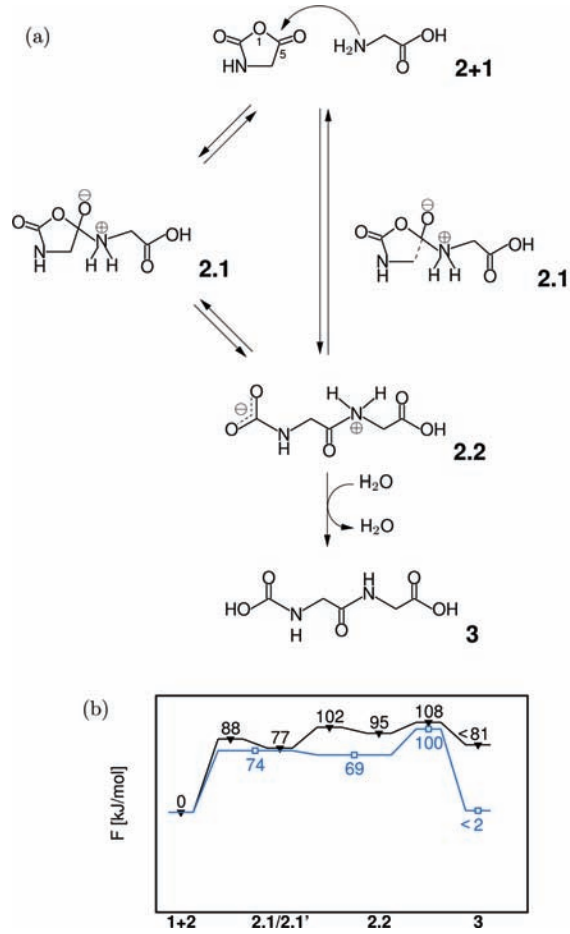

Figure 3. (a) Mechanisms for the reaction of a glycine molecule **1** with NCA **2** (with atom labeling). At ambient conditions (ABW, left) the reaction occurs through intermediate **2.1** whereas in hot-pressurized water (HPW, right) the reaction proceeds via intermediate **2.1′** (note the dashed line in the heterocycle). (b) Corresponding schematic free energy profiles at ambient (black line and filled triangles) and at extreme conditions (blue line and open squares).

Here, the changing properties of water when going from ABW to HPW conditions are found to play a role: the intermediate **2.2** is found to be less stabilized in hot-pressurized water which can be understood in view of the much lower dielectric constant of water at HPW conditions. This decreases the reverse barrier in the formation of the intermediate, that is, going back from **2.2** to reactants at HPW conditions, when compared to the same scenario in ambient bulk water. As the reverse barrier for the intermediate to reactants is only about one k_BT_{500} at $T = 500$ K transitions between reactants and intermediate become very likely and are actually more probable than deprotonation at the nitrogen due to a solvent water molecule.

The overall free energy barrier of activation for this reaction is 43 k_BT_{300} and 24 k_BT_{500} at ambient and extreme conditions, respectively, which implies that the reaction is greatly accelerated in hot-pressurized water. Although the barrier going back from **3** to **2** + **1** was not determined, the reverse reaction involving a cyclization would be very similar to the direct route of NCA formation itself. The latter reaction has been studied previously[12] using the same methods. These barriers were determined to be more than twice as high at both conditions (i.e., 100 k_BT_{300} and 54 k_BT_{500}) compared to those observed for the peptidization step.

(54) Marx, D. *ChemPhysChem* **2006**, *7*, 1848–1870Addendum: Marx, D. *ChemPhysChem* **2007**, *8*, 209–210.

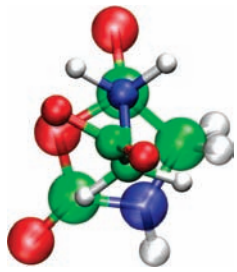

Figure 4. Transition state for the reaction of a glycine **1** molecule with NCA **2** leading to a *cis* peptide. The carbon atoms are green, oxygen red, nitrogen blue, and hydrogen atoms are shown in white.

It is found that the *cis* peptide is formed preferentially as a result of this mechanism and Figure 4 shows the respective transition state leading to the *cis* isomer of diglycine. Since glycine lacks bulky side chains at its C_α carbon it can orient directly above the ring of NCA without strong steric hindrance which provides a favorable *staggered* configuration of the N−H bonds within the amino group of the glycine molecule with respect to the C5−O1 and C5−C_α bonds of NCA. Note that the elimination of any of the amino group protons at the glycine molecule will result in the *cis* form. This structure suggests that, at least for glycine, it is possible to adapt a transition state that leads exclusively to a *cis* peptide. For amino acids having side groups other than hydrogen unfavorable steric and/or electrostatic interactions between the residues at C_α and the NCA ring can arise if the attacking amino acid is directly above the ring of NCA. A transition state leading only to the *trans* isomer should have an *eclipsed* configuration. It is noted in passing that the alternative *trans*-route has also been observed in one of the simulations, but only in a preliminary run using a fast sampling of the free energy surface for initial exploration, i.e. when using "large" Gaussians to build up the repulsive time-dependent potential $V(t,\mathbf{s})$. It is thus reasonable to assume that this run did not allow the system to follow the lowest free energy pathway to the dipeptide.

3.2. Decarboxylation of Gly_2COOH. Since the peptide **3** is the product of the reaction of a free amino acid and NCA, its amino terminus is carboxylated. For the complete formation of the diglycine, the *N*-carboxyl group has to be eliminated according to route **B** in Figure 1. This decarboxylation reaction can be captured within three collective coordinates: the distance between the nitrogen and carbon atoms of the carbamate group, $d[N - C]$, and the coordination numbers of the carbamate nitrogen and oxygen atoms to all nonaliphatic hydrogen atoms, $c[N - H_{na}]$ and $c^{tot}[O - H_{na}]$, respectively. A repulsive potential for $d[N - C] \geq 4$ was applied to restrict the products from diffusing apart too much.

For this step, the observed mechanisms are qualitatively the same at both conditions and are thus shown together with the different schematic free energy profiles in Figure 5; the reconstructed free energy surfaces are shown in SI Figure 2. The reaction starts with a proton transfer from the oxygen to the nitrogen atom within the carbamate group (structure **3.1**). Again, water is seen to play an active role by mediating this proton transfer step.[54] This process is associated with a barrier of 57 kJ/mol (23 k_BT_{300}) at ABW conditions whereas it is 81 kJ/mol (19 k_BT_{500}) at HPW conditions. The reason for the significantly larger free energy to reprotonation at HPW conditions is due to the lower dielectric constant of HPW compared to ABW and its respective influence on the destabilization of the necessarily charged species involved. Nevertheless, although protonation at the nitrogen requires to surmount a higher free energy barrier in the case of HPW compared to ABW (i.e., 81 kJ/mol versus 57 kJ/mol) this process occurs faster at HPW in view of the higher temperature which reduces the *relative* barrier slightly when measured in the appropriate thermal energy units (i.e., 23 k_BT_{300} versus 19 k_BT_{500}).

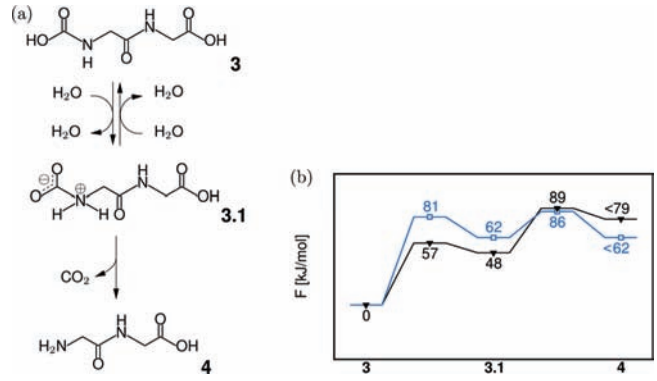

Figure 5. (a) Reaction mechanism for the decarboxylation of *N*-carboxyldiglycine **3** at both ambient conditions (ABW) and in hot-pressurized water (HPW). (b) Corresponding schematic free energy profiles at ambient (black line and filled triangles) and at extreme conditions (blue line and open squares).

In the next and final step of the peptidization sequence CO_2 is eliminated which results in a glycine dipeptide **4**. Since HPW disfavors charged species, the activation free energies for this decarboxylation step is considerably lower at HPW conditions, 24 kJ/mol (6 k_BT_{500}) than at ambient conditions, 41 kJ/mol (16 k_BT_{300}). Thus, the obtained energetics clearly indicate that in HPW this reaction step is much faster than in ABW.

3.3. Hydrolysis of the Dipeptide. To draw conclusions about the capability of the polymerization of amino acids and NCAs at ABW and HPW conditions it is necessary to investigate the back reaction, which is hydrolysis of formed peptides as indicated by route **C** in Figure 1. The chosen reaction coordinate space to describe this process is spanned by the distance of the nitrogen and carbon atoms comprising the peptide bond, $d[N - C]$, and the coordination numbers between the peptide group nitrogen atom and *all* nonaliphatic hydrogen atoms and between peptide group carbon atom and *all* oxygen atoms except those of the carboxyl group, $c[N_{pept} - H_{na}]$ and $c[C_{pept} - O_{free}]$, respectively. A repulsive potential at values exceeding 4Å along $d[N - C]$ restricted the products from diffusing out of the reactive region. During this reaction the collective coordinate $c[N_{pept} - H_{na}]$ is expected to raise from 1 to 2 whereas $c[C_{pept} - O_{free}]$ should increase from 0 to 1. The distance $d[N - C]$ will also increase if the free amino acid molecules diffuse apart.

The different reaction mechanisms for hydrolysis at ambient conditions and in hot-pressurized water along with the corresponding schematic free energy profiles are shown in Figure 6 (for the underlying free energy surfaces see Figure 6d and SI Figure 3). In ABW the first step is the protonation of the peptide **4** at the peptide nitrogen atom, which is associated with an activation barrier of approximately 90 kJ/mol (36 k_BT_{300}) to reach **4.1** that is observed in a flat plateau on the free energy surface (see SI Figure 3). In the next step a water molecule attacks the peptide carbon atom of **4.1** and the diglycine breaks into two glycine molecules **5**. The free energy barrier for this step is 53 kJ/mol (21 k_BT_{300}).

In contrast to the investigations at ambient conditions, in hot-pressurized water the *cis* form of the dipeptide was used as the

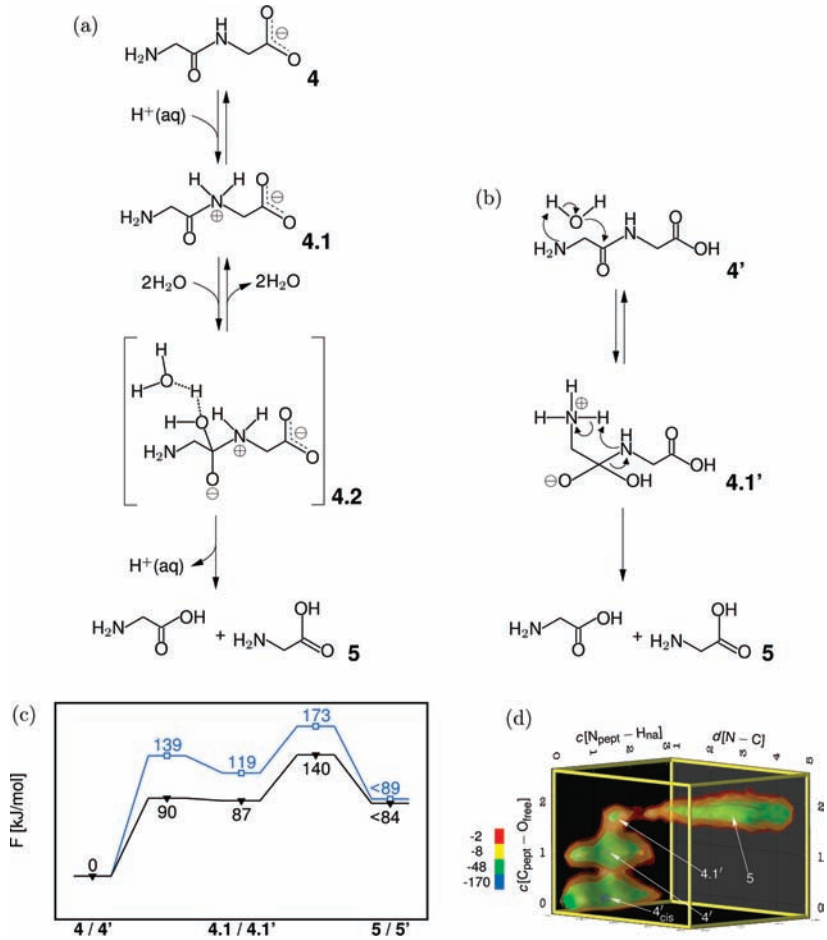

Figure 6. Reaction mechanisms for the hydrolysis of diglycine **4** (a) at ambient conditions (ABW) and (b) in hot-pressurized water (HPW). (c) Corresponding schematic free energy profiles at ambient (black line and filled triangles) and at extreme conditions (blue line and open squares). (d) Reconstructed free energy surface for this reaction at HPW is presented as a volumetric data for selected free energy (contour) values (in kJ/mol); $\mathbf{4'}_{cis}$ is having a *cis*, and **4** and **4′** is having a *trans* peptide bond.

starting structure in order to probe possible influences of the peptide bond's conformation on hydrolysis. However, during the simulation, the *cis*-diglycine transforms into the *trans* isomer before hydrolysis sets in (this part of the mechanism is not depicted in Figure 6). In the first step of the initial *cis* to *trans* isomerization a water molecule attacks the amide carbon atom and the activation barrier for this process is about 140 kJ/mol (33 k_BT_{500}); see Figure 6d. After several protonation/deprotonation steps of the oxygen atoms of this intermediate a hydroxyl group is eliminated regenerating the peptide but in terms of the *trans* isomer **4′**. Since the eliminated oxygen atom stems from the peptide this process corresponds to an O-exchange at the peptide group. The protonation/deprotonation steps at the nitrogen atom are associated with a free energy barrier of about 40 kJ/mol (~10 k_BT_{500}).

The mechanism of peptide hydrolysis itself is very different at HPW. Hydrolysis starts at HPW conditions with an attack of a water molecule on the peptide group carbon atom of **4′**, like the oxygen exchange mechanism, and not with a protonation of the amide nitrogen atom as observed in ABW. Note that unlike in **4**, the initial structure in the ABW scenario, the C-terminus of **4′** is protonated in high-pressurized water which is in accordance with its much lower dielectric constant. Moreover, in contrast to the corresponding process at ambient conditions the resulting excess proton of the attacking water molecule is not accepted by the solvent, but directly transferred instead to the N-terminus of the dipeptide resulting in the intermediate **4.1′**. Again, the low dielectric constant of HPW tends to suppress the involvement of charged species in this reaction step. The free energy of activation is about 140 kJ/mol (33 k_BT_{500}) which is thus the same barrier as found for the first step of the O-exchange. Making use of an intramolecular proton transfer step from the N-terminus the peptide group nitrogen atom is protonated and the peptide bond breaks apart thus yielding **5′**. This last process is associated with an activation barrier of 54 kJ/mol (13 k_BT_{500}). It is found that here intramolecular proton transfer is preferred over water-mediated proton transfer as generally observed in ambient water with its high dielectric constant.

The barrier at ambient conditions, 140 kJ/mol, is in accordance with a previous theoretical study[55] investigating *N*-methylacetamide hydrolysis in aqueous solution by constrained dynamics where the barrier was determined to be 147 kJ/mol in water. In the present investigations a stepwise mechanism starting with the protonation of the amide nitrogen atom is observed, whereas in ref 55 a concerted water attack on the amide carbon atom and proton transfer to nitrogen is suggested. However, unlike in the case of formamide investigated in ref 55, the system studied here has additional amino and carboxyl groups and a one-dimensional reaction coordinate

(55) Zanh, D. *Eur. J. Org. Chem.* **2004**, *2004*, 4020–4023.

has been used previously to enforce the reaction whereas here a much more flexible three-dimensional reaction coordinate space is spanned in terms of rather general collective coordinates.

How does this *in silico* scenario compare to experimental knowledge about hydrolysis of peptides in water? Overall, the free energy barrier for the hydrolysis of a peptide bond in diglycine is higher in HPW than in ABW, that is, 173 vs 140 kJ/mol. Nevertheless, due to the 200 deg temperature increase from 300 to 500 K and thus the increasing thermal energy k_BT, $\Delta F/k_BT$ is lowered from 56 k_BT_{300} in ABW to 42 k_BT_{500} in HPW which implies that the hydrolysis reaction is expected to be *faster* in hot-pressurized water. First of all the computed free energy of activation at ambient conditions is in accordance with experiment.[56] At $T = 298$ K and pH 6.8 a reaction rate of $k = 6.3 \times 10^{-11}s^{-1}$ was measured for the uncatalyzed hydrolysis of diglycine in sealed quartz tubes,[56] which corresponds to an experimental free energy of activation of 131 kJ/mol compared to 140 kJ/mol computed at 300 K. Raising the temperature by 125 deg to $T = 423$ K the reported[56] reaction rate increases to $k = 9.8 \times 10^{-6}s^{-1}$ despite the fact that the free energy of activation, about 146 kJ/mol, is *larger* than at room temperature; note that a further increase of the activation barrier is expected when increasing the temperature from 423 to 500 K as used in the simulation. Thus, also the experimentally found trend upon raising the temperature agrees with the simulation results in that the increasing thermal fluctuations outweigh the increasing barrier height and thus lead to a faster hydrolysis at extreme conditions.

4. Conclusions and Outlook

The presented investigations assess both the detailed mechanistic aspects and the free energetics involved in the formation of peptides via NCA-activated amino acids as well as peptide hydrolysis in ambient bulk water (ABW) and also at hot-pressurized bulk water (HPW) extreme thermodynamic conditions. Both peptide bond formation by the reaction of glycine-NCA with glycine and the subsequent decarboxylation reaction are found to be accelerated at HPW conditions. This effect can mainly be attributed to the decrease in the effective forward barriers due to increasing thermal fluctuations at HPW conditions. At both thermodynamic conditions peptidization proceeds *via* a zwitterionic species, which is an intermediate in ABW whereas it is a transient structure at extreme conditions. The observed mechanistic differences at ABW and HPW conditions are the result of a decreasing stabilization of charged species in hot-pressurized water due to the substantial lowering of the dielectric constant compared to ambient water.

In addition to peptidization also the back reaction, that is, hydrolysis of diglycine, was investigated again comparing ABW to HPW conditions. This hydrolysis reaction is found to be also accelerated by extreme thermodynamic conditions though the free energy barrier in kJ/mol is higher for HPW compared to ABW in agreement with experimental data. Interestingly, the observed mechanisms of hydrolysis for both conditions are shown to be qualitatively different. At ABW conditions, a proton transfer to the nitrogen of the peptide occurs as the first step while this is the second step in HPW. Water addition to the peptide carbon is the second step in ABW but is the first step at HPW conditions. Furthermore, proton transfer is seen to be water-mediated in ABW whereas intramolecular proton transfer mediated by the NH_2 terminus of diglycine itself is observed in hot-pressurized water, thus avoiding solvation of long-lived charged species. The effective barrier of hydrolysis of diglycine in neutral water is higher than any forward barrier to peptide formation at both ABW and HPW conditions, even taking into account the formation of NCAs by the activation of amino acids with COS.[12] These results suggest that the presented reaction sequence allows for peptide production at both ABW and HPW conditions.

Being an important first step, a more comprehensive understanding of peptide synthesis along the lines of the NCA-route certainly calls for analyses of alternative pathways, competing reactions and side reactions, as well as homogeneous or heterogeneous catalytic effects in the presence of transition metal salts or mineral surfaces, respectively. Some of these investigations are on the way in our laboratory and will certainly contribute more exciting insights into peptide formation in aqueous environments.

Acknowledgment. We are grateful to G. Wächtershäuser and H. R. Kricheldorf for stimulating discussions as well as to J. Hutter for technical help. Partial financial support was provided by DFG (Normalverfahren MA 1547/7) and by FCI. The simulations were carried out on the IBM Blue Gene system at John von Neumann Institute for Computing (NIC) at Forschungszentrum Jülich (Germany).

Supporting Information Available: Reconstructed free energy surfaces for steps **A** to **B** to **C** according to the scheme in Figure 1. This material is available free of charge via the Internet at http://pubs.acs.org.

JA9032742

(56) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105–6109.