Ab Initio Quantum Mechanical Model Calculations on the Catalytic Mechanism of Aspartylglucosaminidase (AGA): A Serine Protease-Like Mechanism with an *N*-terminal Threonine and Substrate-Assisted Catalysis

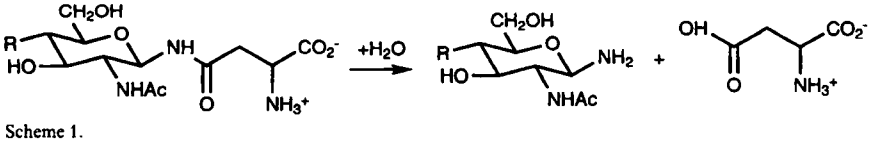
Mikael Peräkylä* and Juha Rouvinen

Abstract: Ab initio quantum mechanical model calculations were used in studying the acylation and deacylation steps of the catalytic mechanism of aspartylglucosaminidase (AGA). AGA catalyses the hydrolysis of an amide linkage between oligosaccharide and asparagine by utilising an *N*-terminal threonine as a catalytic amino acid. Results are reported for the model enzyme reaction at the MP2/6-31+G*//HF/6-31+G* + ΔZPE level. Contribution of aqueous solvation to the relative energies was estimated by using the continuum solvation model of Tomasi. The serine protease-like catalytic mechanism was found to be feasible for AGA. The protonated α-amino group of the substrate of AGA was suggested to enhance the catalysis by stabilising the anionic oxygen of the substrate, which is formed in the reaction, and by lowering the pK_a of the nucleophilic oxygen of the *N*-terminal threonine. Finally, the similarities in the catalytic mechanisms of AGA and other amidohydrolases were discussed.

Keywords
ab initio calculations · enzymatic catalysis · enzyme models · mechanistic studies

1. Introduction

Aspartylglucosaminidase (AGA) is a lysosomal hydrolase that catalyses the hydrolysis of the amide linkage between oligosaccharide and asparagine (Scheme 1).[1] The mutations of the AGA amino acid sequence are responsible for the human genetic disease aspartylglucosaminuria, which causes severe mental retardation.[1, 2] The X-ray crystal structure of human AGA complexed with the natural reaction product aspartate has been recently determined.[3] AGA belongs to the newly recognized structural superfamily of *N*-terminal nucleophilic (Ntn) amidohydrolases, which use the side chain of the *N*-terminal amino acid as the nucleophile in the catalytic attack on the carbonyl carbon of the substrate.[4] The crystal structures of three other Ntn hydrolases—penicillin acylase (PA),[5] proteasome (PRO)[6] and glutamine PRPP amidotransferase (GAT)[7]—have been recently determined. Ntn hydrolases share a similar three-dimensional fold in which the nucleophile and other catalytic groups occupy equivalent sites. The common catalytic framework of these enzymes suggest that the four enzymes share a common catalytic mechanism.[4] Based on the three-dimensional structures and biochemical data, the well-known serine protease-like catalytic mechanisms[8–11] have been proposed for the enzymes.[5, 6] The *N*-terminal threonine of AGA and PRO, the serine of PA, and the cysteine of GAT act as the nucleophile and the α-amino group of the *N*-terminal amino acid as the base in the catalytic reaction, corresponding to serine and histidine, respectively, of the serine proteases. Furthermore, the oxyanion binding site, the oxyanion cavity, has been identified from the structures of Ntn hydrolases. The pH optima for the rate of the catalytic reactions are between 7–9 for Ntn hydrolases.[5, 12, 13] This is consistent with the suggestion that the α-amino group, where the pK_a in a protein is 6.8–7.9,[14] acts as a base in the catalytic reaction.



Scheme 1.

In this paper we present the first detailed theoretical study on the catalytic mechanism for the hydrolysis of a peptide bond by an enzyme utilising an *N*-terminal amino acid as a catalytic group. Specifically, we report the catalytic mechanism for the acylation and deacylation steps of AGA-catalysed reactions, based on ab initio quantum mechanical model calculations and the three-dimensional structure of AGA. Active-site models of two different sizes were used. First, the mechanisms of the acylation and deacylation steps of the catalytic reaction were calculated with a model that comprised the model substrate formamide, the *N*-terminal threonine as modelled by 2-aminoethanol, and the oxyanion cavity, as modelled by 3-hydroxypropionamide (Figs. 1a and 1b). Second, to make a more represen-

[*] Dr. M. Peräkylä, Dr. J. Rouvinen
Department of Chemistry, University of Joensuu
POBox 111, FIN-80101, Joensuu (Finland)
e-mail: Mikael.Perakyla@csc.fi

0947-6539/96/0212-1548 $ 15.00+.25/0 Chem. Eur. J. 1996, 2, No. 12

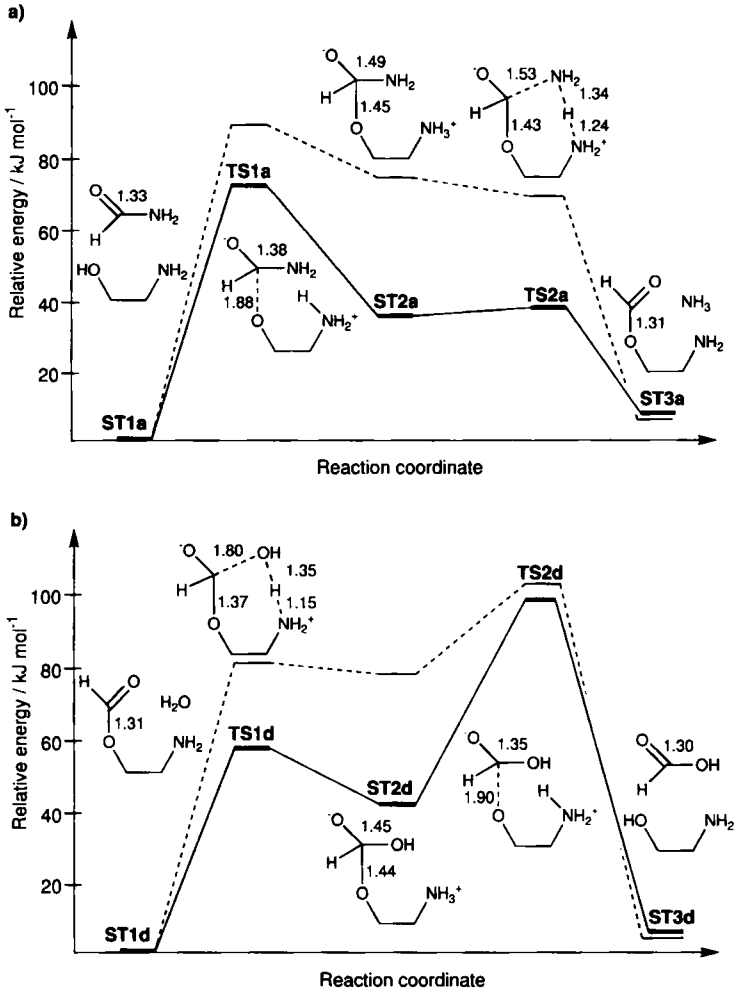


Fig. 1. Structures and the relative energies (MP2/6-31+G*//HF/6-31+G* + ΔZPE) of stationary points in a) the acylation step and b) the deacylation step of the model catalytic reaction of AGA in the gas phase (- - -) and in solution (—, $\varepsilon = 78.3$, water). The oxyanion cavity in the model is omitted from the Figure.

tative model of the active site and to include the most important interactions in the calculations, we used N^{δ}-methylasparagine as the substrate.

It must be stressed here that, owing to the model nature of the calculations, the results reported do not aim to reproduce the catalytic reaction within the enzyme. However, the qualitative features of the reaction mechanism presented in this paper are thought to be representative for the AGA-catalysed reaction. It will be shown that a serine protease-like catalytic mechanism is feasible both energetically and structurally for the Ntn amidohydrolases and that the protonated α-amino group of the substrate may play an important role in the AGA-catalysed reaction.

2. Results and Discussion

2.1. Catalytic mechanism: The energies of the stationary points and the selected key geometric parameters of the model enzyme reaction are shown in Figure 1a for the acylation step and in Figure 1b for the deacylation step. Both the gas-phase energies and the energies including solvation contributions are shown in the Figures. In the acylation step of the enzyme reaction the nucleophilic oxygen ($O\gamma$ of Thr183 of AGA) first attacks the carbonyl carbon of the substrate (TS1a). The α-amino group acts here as a base and enhances the nucleophilic character of the hydroxyl group. The transfer of the proton from the oxygen to the amino group is completed at the transition state. TS1a was calculated to be the highest point in the acylation step and the corresponding structure TS2d the highest point in the deacylation step. The energetics in Figure 1 are values for isolated active-site models, and it is therefore probable that inclusion of protein environment in the calculations would lower the reaction barriers. In the case of penicillin acylase it has been proposed that the transfer of the proton from the nucleophilic oxygen to the α-amino group is mediated by a bridging water molecule.[5] In the case of the crystal structure of AGA no such water was found in the active-site. Furthermore, modelling of the enzyme reaction suggested that there is no room for a bridging water molecule in the active site. The tetrahedral intermediate ST2a loses NH_3 from the model substrate and collapses in the acylated enzyme intermediate through the transition state TS2a. At TS2a the proton is between the nitrogens, and the C–N distance is slightly elongated. TS2a was identified at the HF/6-31+G* level, but it disappeared when electron correlation and ΔZPE energies were added. In contrast, the corresponding transformation (ST2d → TS1d) in the deacylation part of the overall reaction has a small barrier, and the C–O bond between the oxygen of the attacking water and the carbonyl carbon is significantly longer (1.80 Å) at TS1d than the C–N bond (1.53 Å) at TS2a. The deacylation step of the reaction is analogous to the acylation step, and the energies of the corresponding points are rather similar. Inclusion of solvation energies stabilises the transition states and the intermediates ST2a(d) by 10–40 kJ mol^{-1}. The stabilised species have ionised catalytic groups and are presumably stabilised by the protein environment as well.

2.1. Substrate-assisted catalysis: To further elucidate the reaction catalysed by AGA we optimised the geometry (HF/6-31G*) of the tetrahedral intermediate corresponding to ST2a by using a larger substrate model, N^{δ}-methylasparagine. The optimized structure is shown in Figure 2, top, and the corresponding amino acid residues and the reaction product aspartate as taken from the AGA–aspartate X-ray structure[3] in Figure 2, bottom.

As can been seen from Figure 2, the atoms of the *N*-terminus and the oxyanion cavity of the calculated model and the X-ray structure are in similar positions. Furthermore, the NH_3^+ and CO_2^- groups of the model and the aspartate of the X-ray structure are located in a similar fashion. It can be seen that all the direct interactions between the amide bond, which is cleaved in the reaction, and the enzyme are present in the tetrahedral intermediate model of Figure 2, top. This structure nicely reveals the key catalytic interactions and the perfectly arranged catalytic machinery of AGA. In addition to the anionic oxygen of the tetrahedral intermediate being hydrogen-bonded to the NH and OH groups of the oxyanion cavity, it makes a close intramolecular contact with the protonated α-amino group of the substrate. The distance between the anionic oxygen and the NH_3^+ nitrogen is 2.71 Å. Such an intramolecular interaction between charged groups probably provides a significant amount of transition-state stabilisation. A similar intramolecular transition-state stabilisation may be important for other asparaginases having a free α-amino group, as well. In addition, the NH_3^+ group of the substrate is close to the nucleophilic oxygen and

  0947-6539/96/0212-1549 $ 15.00+.25/0

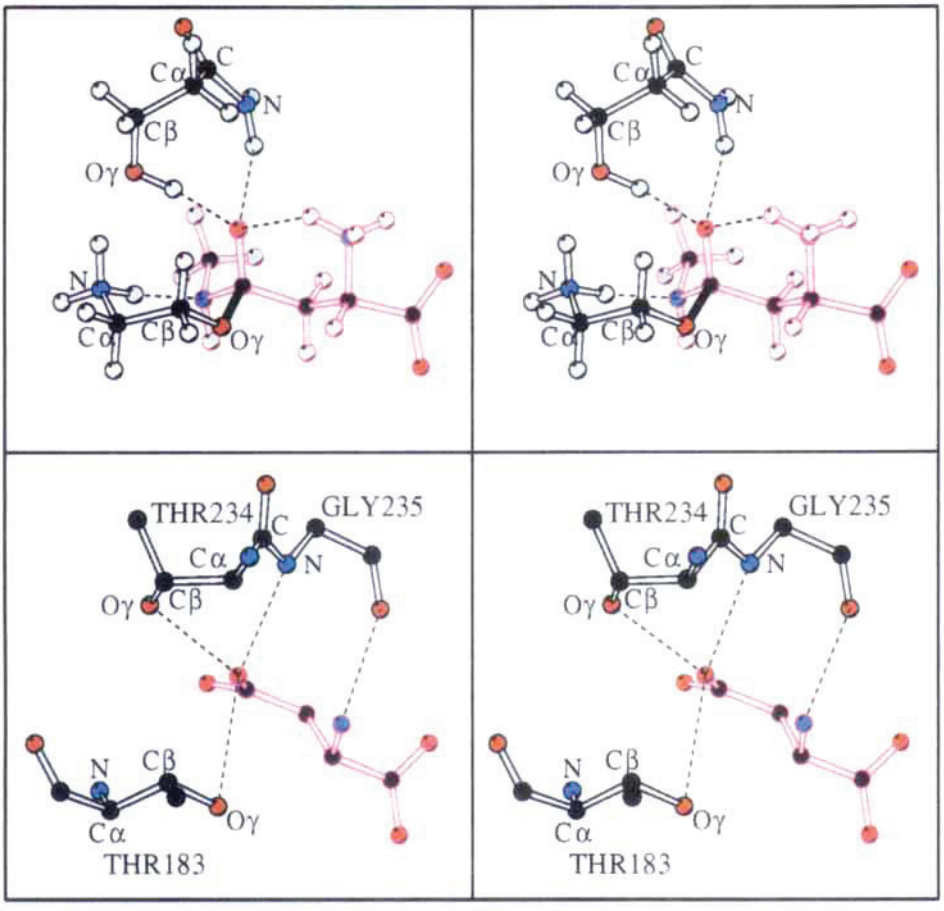


Fig. 2. Stereo representation of the optimised geometry (HF/6-31G*) of the tetrahedral intermediate of the acylation step using N^{δ}-methylasparagine as a substrate (corresponding to ST2a of Fig. 1a) (top) and the selected active-site residues and the reaction product as taken from the X-ray structure of the AGA-aspartate complex (bottom).

those of the other Ntn hydrolases, too, resembles closely the well-know reaction mechanisms of serine proteases. However, in the case of Ntn hydrolases there is no catalytic triad present in the active site, but an *N*-terminal amino acid, which functions as a nucleophile and a catalytic base. In the case of AGA, the protonated α-amino group of the substrate plays an active part in the catalytic mechanism by stabilising the oxyanion that is formed in the reaction and, presumably, by lowering the pK_a of the nucleophilic oxygen of the *N*-terminal threonine by electrostatic effects. This special feature of the catalytic mechanism of AGA, which can be described as substrate-assisted catalysis,[19] may be responsible for the high catalytic activity of AGA, even in acidic pH. At pH 5, 70% of the AGA activity remains.[20] This may be important for the function of AGA in lysosomes, because although AGA exhibits a broad pH maximum between 7 and 9, the lysosomal enzymes usually have an acidic pH optimum.

therefore may enhance the catalysis by lowering the pK_a of the oxygen by electrostatic effects. Such a role has been proposed for the NH_3^+ group of lysine of other amidohydrolyses. In the active site of proteasome a Lys 33 is in close proximity to the Oγ of the catalytic threonine. Lys 33 is conserved in the proteasome family, and the mutation of Lys33 to alanine or arginine produces a catalytically inactive enzyme.[6, 15] In the case of L-asparaginase the nucleophile threonine[16, 17] and in the case of *N*-carbamoylsarcosine amidohydrolase the nucleophile cysteine[18] are probably hydrogen-bonded to NH_3^+ group of lysine. Figure 2, top, also shows that in the tetrahedral intermediate the N-Cα-Cβ-Oγ atoms of the *N*-terminus, and the Cγ and Nδ (designation of the atoms is taken from the corresponding amino acids) of the substrate form a six-membered ring in a chair conformation. The two nitrogens of the ring are in close proximity with an N–N distance of 2.76 Å. This arrangement facilitates the proton transfer between the nitrogens and the cleavage of the Cγ–Nδ bond (ST2a → TS2a → ST3a). In analogy, the addition of H_2O to the carbonyl carbon of the acylated enzyme intermediate (TS1d) is catalysed by the perfectly positioned α-amino group of the *N*-terminus, which acts as a base, and the protonated α-amino group of the substrate which stabilises the oxyanion that is formed in the reaction (ST1d → TS1d → ST2d).

3. Conclusion

The ab initio quantum mechanical calculations presented here suggest that the catalytic mechanism of AGA, and propably

Computational Methods

The geometries of the minima and the transition states of the model enzyme reaction were minimized at the HF/6-31G* and HF/631+G* level with the Gaussian 94 program [21]. The starting structures were constructed by using the X-ray structure of AGA [3] as a model. No significant geometrical changes occurred in the geometries during the optimisations. Vibrational frequencies were calculated at the HF/6-31G* level to conform with the nature of all the stationary points and to achieve zero-point vibrational energies (ZPE). Frequencies were scaled by 0.89. Energies were further calculated at the MP2/6-31+G*//HF/6-31+G* level. To account for the effect of solvation on the energies of the models, the electrostatic solvation energies were calculated with the solvation model of Tomasi [22,23] as implemented in the Gaussian 94 (IPCM-option) [24]. Gas-phase geometries were used in these calculations. In the solvation calculations a dielectric constant (ε) of 78.3 was given for the solvent (water) and a value of 0.0004 $e B^{-3}$ was used for the charge density in the determination of the cavity boundary. Coordinates and the energies of the optimised structures are available from the authors upon request.

Received: May 29, 1996 [F376]

[1] I. Mononen, K. Fischer, V. Kaartinen, N. N. Aronson, Jr., *FASEB J.* **1993**, *7*, 1247–1256.
[2] P. J. Artymiuk, *Nat. Struct. Biol.* **1995**, *2*, 1035–1037.
[3] C. Oinonen, R. Tikkanen, J. Rouvinen, L. Peltonen, *Nat. Struct. Biol.* **1995**, *2*, 1102–1108.
[4] J. A. Brannigan, G. Dodson, H. J. Duggleby, P. C. E. Moody, J. L. Smith, D. R. Tomchick, A. G. Murzin, *Nature* **1995**, *378*, 416–419.
[5] H. J. Duggleby, S. P. Tolley, C. P. Hill, E. J. Dodson, P. C. E. Moody, *Nature* **1995**, *373*, 264–268.
[6] J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, *Science* **1995**, *268*, 533–539.
[7] J. L. Smith, E. J. Zaluzec, J.-P. Wery, L. Niu, R. L. Switzer, H. Zalkin, Y. Satow, *Science* **1994**, *264*, 1427–1433.
[8] V. Daggett, S. Schröder, P. Kollman, *J. Am. Chem. Soc.* **1991**, *113*, 8926–8935.
[9] A. Warshel, S. Russell, *J. Am. Chem. Soc.* **1986**, *108*, 6569–6579.
[10] S. Nakagawa, H.-A. Yu, M. Karplus, H. Umeyama, *Proteins*, **1993**, *16*, 172–194.
[11] J. Kraut, *Ann. Rev. Biochem.* **1977**, *46*, 331–358.

[12] V. Kaartinen, J. C. Williams, J. Tomich, J. R. Yates III, L. E. Hood, I. Mononen, *J. Biol. Chem.* **1991**, *266*, 5860–5865.
[13] E. Seemüller, A. Lupas, F. Zuhl, P. Zwickl, W. Baumeister. *FEBS Lett.* **1995**, *359*, 173–178.
[14] C. Tanford, *Adv. Protein Chem.* **1962**, *17*, 69–165.
[15] E. Seemüller, A. Lupas, D. Stock, J. Löwe, R. Huber, W. Baumeister, *Science* **1995**, *268*, 579–582.
[16] A. L. Swain, M. Jaskólski, D. Housset, J. K. M. Rao, A. Wlodawer, *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 1474–1478.
[17] M. Miller, J. K. M. Rao, Wlodawer, M. R. Gribskov, *FEBS Lett.* **1993**, *328*, 275–279.
[18] M. J. Ramão, D. Turk, F.-X. Gomis-Ruth, R. Huber, G. Schumacher, H. Mollering, L. Russmann, *J. Mol. Biol.* **1992**, *266*, 1111–1130.
[19] P. Carter, J. A. Wells, *Science* **1987**, *237*, 394–399.
[20] R. Tikkanen, A. Riikonen, C. Oinonen, J. Rouvinen, L. Peltonen, *EMBO J.* **1996**, *15*, 2954–2960.
[21] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petterson, J. A. Montgomery, K. Rachavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. J. P. Stewart, M. Head-Gordon, C. Gonzales, J. A. Pople, *Gaussian 94*, Revision B. 3. Gaussian Inc., **1995**.
[22] J. Tomasi, M. Persico, *Chem. Rev.* **1994**, *94*, 2027–2094.
[23] S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys.* **1981**, *55*, 117–129.
[24] K. B. Wiberg, P. R. Rablen, D. J. Rush, T. A. Keith, *J. Am. Chem. Soc.* **1995**, *117*, 4261–4270.