Available online at www.sciencedirect.com

ScienceDirect

Mendeleev Commun., 2022, **32**, 739–741

Mendeleev Communications

Mechanism of chemical reactions in the active site of aspartate *N*-acetyltransferase NAT8L revealed by molecular modeling

Igor V. Polyakov,[a] Maria G. Khrenova,*[a,b] Bella L. Grigorenko[a] and Alexander V. Nemukhin[a,c]

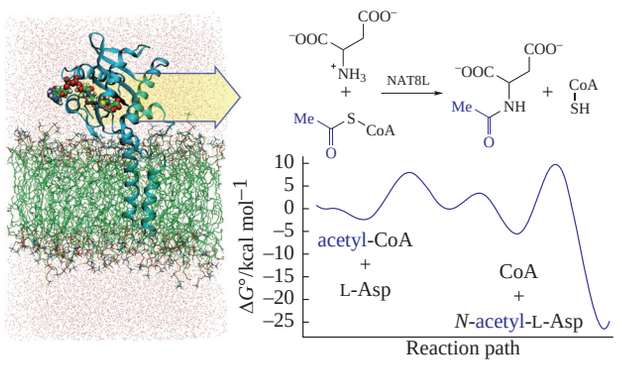
[a] *Department of Chemistry, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation. E-mail: khrenova.maria@gmail.com*

[b] *A. N. Bach Institute of Biochemistry, Federal Research Centre 'Fundamentals of Biotechnology' of the Russian Academy of Sciences, 119071 Moscow, Russian Federation*

[c] *N. M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 119334 Moscow, Russian Federation*

DOI: 10.1016/j.mencom.2022.11.010

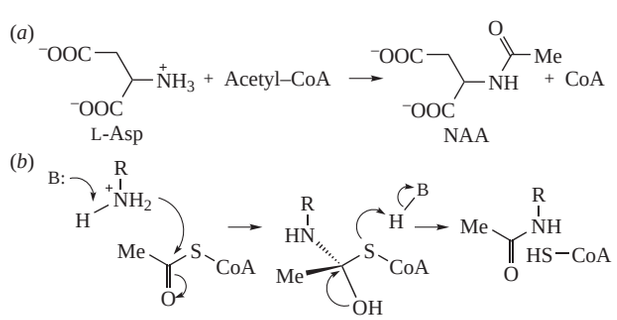
The results of a computational study of the synthesis of a key brain metabolite, *N*-acetyl-L-aspartate, catalyzed by aspartate *N*-acetyltransferase, encoded by the *NAT8L* gene, are reported. The reaction Gibbs energy profiles were computed using molecular dynamics simulations with interaction potentials estimated on-the-fly by the quantum mechanics/molecular mechanics QM(PBE0/6-31G)/MM(CHARMM) approach. The revealed reaction mechanism includes four elementary steps with corresponding activation energies not exceeding 14 kcal mol^{-1}.**



Keywords: *N*-acetyl-L-aspartate, aspartate *N*-acetyltransferase, NAT8L, reaction mechanism, Gibbs energy profiles, molecular dynamics, QM/MM potentials.

The chemistry of the enzyme-catalyzed synthesis of a key brain metabolite, *N*-acetyl-L-aspartate (NAA),[1] is poorly characterized in experiments and simulations. Synthesis is carried out by a membrane-associated *N*-acetyltransferase 8-like protein (NAT8L) encoded by the *NAT8L* gene.[2,3] The reaction is presented in Scheme 1 in accordance with the general knowledge of GCN5-related *N*-acetyltransferases (where GCN5 means 'general control of amino acid synthesis').[4] However, the mechanism of chemical transformations in the active site of NAT8L on the way from reactants (REAC), aspartate and acetyl-coenzyme A (acetyl-CoA), to products (PROD), NAA and CoA, is unknown.

Crystal structures of NAT8L, neither in its apo form nor in complex with ligands, have been reported despite extensive efforts to purify and characterize this important protein.[3] A tentative prediction of the secondary structure of NAT8L was proposed by



Scheme 1 (*a*) Reactants and products of NAA synthesis by the NAT8L enzyme. (*b*) Previously proposed reaction mechanism in the active site of *N*-acetyltransferases.[4]

Tahay *et al.*[5] after careful analysis of scarce kinetic and mutational experiments. In particular, the authors were able to determine which regions of the protein are important for catalytic activity. A three-dimensional all-atom model of NAT8L containing reactants in the active site of the enzyme was constructed using molecular modeling tools, when starting from the primary protein sequence.[6] The methods included multiple sequence alignment, deep neural network modeling, comparative modeling with RosettaCM and classical molecular dynamics simulations. This model is in good agreement with the experimentally validated structure regarding the location of the catalytic and membrane-associated regions.[5] According to these findings, the two catalytic regions spanning amino acid residues 80–118 and 149–302 are separated by the membrane-associated region 119–148. The principal features of the computationally derived structure are shown in Figure S1 (see Online Supplementary Materials).

The structure[6] of the protein with the reactants in the active site provides a basis for considering the reaction mechanism. We assume that the reaction is initiated by the formation of an enzyme–substrate complex, REAC, containing the substrate with a protonated amino group and negatively charged carboxylates (Figure S2). The reaction products (PROD) contain acetylated aspartate in the active site of the enzyme. The goal of the present simulations is to dissect the reaction REAC → PROD into a sequence of elementary steps and evaluate the corresponding Gibbs energy profiles using an advanced modeling strategy based on molecular dynamics (MD) calculations with interaction potentials estimated using quantum mechanics/molecular mechanics (QM/MM). The application of the high-level quantum chemistry DFT(PBE0/6-31G**) method for a large quantum subsystem (131 atoms) for on-the-fly evaluation of

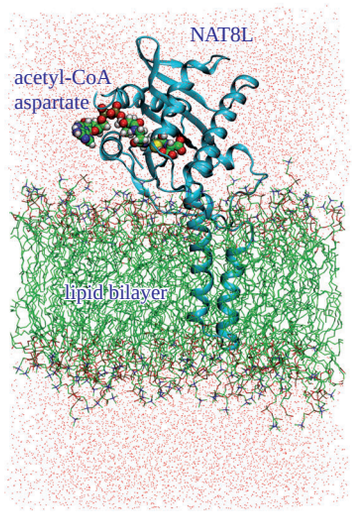


Figure 1 Model system used in calculations. The protein is shown as a cyan cartoon image. The oxygen atoms of the water molecules surrounding the protein are shown as red dots, the phospholipid layers are shown in sticks. The atoms of the reactants, aspartate and acetyl-CoA, are represented as space-filled spheres. Atoms are colored as follows: carbon – green, oxygen – red, nitrogen – blue, phosphorus – orange, sulfur – yellow, hydrogen – white.

energies and energy gradients in MD simulations promises reliable estimates of activation energy barriers, which are strongly required for the kinetics of NAA transformations in the human brain.[7,8] In turn, these data are needed in biomedical research, as over-expression of *N*-acetylaspartate synthase and the subsequent increase in NAA levels are associated with various life-threatening cancers.

Figure 1 illustrates the model system used to calculate the Gibbs energy profiles of the reaction.[†]

In Figure 2, balls and sticks show the molecular groups of the reactants directly involved in chemical transformations. Two-letter symbols and single-letter symbols designate the key atoms of aspartate and acetyl-CoA, respectively. The atoms of these groups are assigned to the QM subsystem together with the selected atoms from the amino acid residues surrounding the immediate participants in the reaction. The names of these residues are shown in boxes in the main part of Figure 2. Several water molecules are also included into the QM subsystem. The complete molecular model of the active site is illustrated in Figure S1. When starting from the reactants, the N(1)–C chemical bond is formed and the C–S bond is cleaved, as indicated by the light blue arrows in the main part of Figure 2, eventually leading to the products shown in the inset.

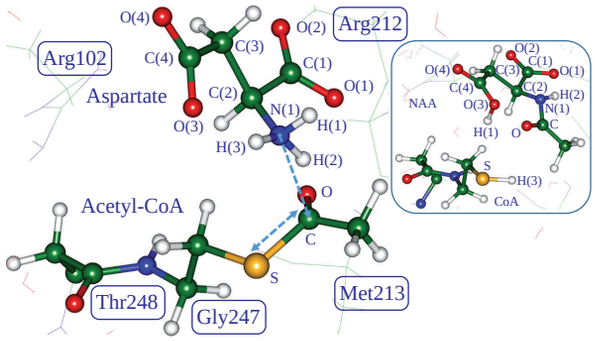


Figure 2 Molecular models of the reactants and (inset) reaction products in the NAT8L active site. Atoms are colored as follows: carbon – green, oxygen – red, nitrogen – blue, phosphorus – orange, sulfur – yellow, hydrogen – white.

As is commonly assumed in *N*-acetyltransferase catalysis,[4] the NH_3^+ group of aspartate should be deprotonated before nucleophilic attack on acetyl-CoA (see Scheme 1). However, there is no suitable candidate for the role of a general base in the active site of this enzyme. Therefore, we propose to explore a variant of the substrate-assisted (or substrate-as-base) mechanism,[9] according to which the functional groups of the substrate contribute to the catalysis of the enzyme. Specifically, we consider the formation of an aspartate tautomer by transferring the H(1) proton to the oxygen atom O(1) of the carboxyl group O(1)–C(1)–O(2). We note that an alternative pathway that considers the second carboxyl group O(3)–C(4)–O(4) of aspartate as a proton acceptor (see Figure 1) was rejected in the preliminary calculations due to the much higher energy barrier along this tentative route.

Thus, the first step of the reaction, REAC → REAC*, leads to an activated form of the reactants (REAC*), which is capable of initiating the nucleophilic attack of aspartate on acetyl-CoA. The conformation of REAC* corresponds to the complex of the aspartate tautomer and acetyl-CoA in the active site of the enzyme. The corresponding collective variable CV1 for this step includes the distances N(1)–H(1) and H(1)–O(1) and the dihedral angle N(1)–C(2)–C(1)–O(1):

$$\mathrm{CV1} = 0.5d_{\mathrm{H(1)–O(1)}} - 0.5d_{\mathrm{N(1)–H(1)}} - 0.015\chi_{\mathrm{N(1)–C(2)–C(1)–O(1)}}. \quad (1)$$

The computed Gibbs energy profile and corresponding structural changes are illustrated in Figure S2. The profile shows a tiny energy barrier of less than 1 kcal mol^{-1}, whereas the energy of REAC* is about 2 kcal mol^{-1} below the level of REAC. The structure of REAC* remains stable in the unconstrained QM/MM MD simulations on trajectories longer than 10 ps.

On the way from REAC* to PROD, we locate two reaction intermediates, designated INT1 and INT2. The elementary step REAC* → INT1 is described by the collective variable CV2:

$$\mathrm{CV2} = d_{\mathrm{N(1)–C}} + d_{\mathrm{H(1)–O}}. \quad (2)$$

The computed Gibbs energy relief and the corresponding structural changes are illustrated in Figure S3. The energy barrier for this elementary step is 10 kcal mol^{-1}, whereas the energy level of INT1 is +2 kcal mol^{-1} relative to REAC*. The energy of INT1 matches the level of the initial REAC structure.

The next step INT1 → INT2 leads to a typical tetrahedral reaction intermediate in the nucleophilic attack of the primary amine on the acyl carbon of the acetyl group.[4] The collective variable at the step INT1 → INT2 includes the H(3) proton transfer distances and the torsion angle (Figure 2):

$$\mathrm{CV3} = d_{\mathrm{H(3)–O(3)}} - d_{\mathrm{N(1)–H(3)}} + 0.03\chi_{\mathrm{N(1)–C(2)–C(3)–C(4)}}. \quad (3)$$

[†] *Computational protocol.* To compute the Gibbs energy profiles for the elementary steps of the reaction REAC → PROD, we use MD simulation with QM/MM potentials. Specifically, we employ the implementation[10] of the QM/MM MD approach *via* a modified interface[11] of the NAMD MM/MD suite[12,13] and the Terachem QC software.[14] This approach makes it possible to efficiently calculate the QM/MM MD trajectories using graphical processing units with atom-centered Gaussian basis sets and hybrid density functionals (PBE0-D3/6-31G** in this work) necessary for an accurate description of nucleophilic reactions.[15] The present simulations use CHARMM36[16,17] all-atom force fields for the protein and lipid layers, TIP3P[18] for water molecules and CGenFF[19,20] topology and parameters for acetyl-CoA. The entire model system consisted of 66167 atoms. The trajectories were simulated for T = 310 K, P = 1 atm and an integration step of 1 fs, and periodic boundary conditions were used along with the particle mesh Ewald method. The system was equilibrated *via* an extensive classic MD refinement.[6] Umbrella sampling biased QM/MM MD trajectory calculations were performed for 5–10 ps trajectories in each window, umbrella integration and WHAM approaches[21,22] were applied to obtain free energy profiles. The parameters of harmonic potentials (force constants and CV values at which the potentials were centered) were chosen to provide overlaps between the distributions of CV values calculated in different windows.

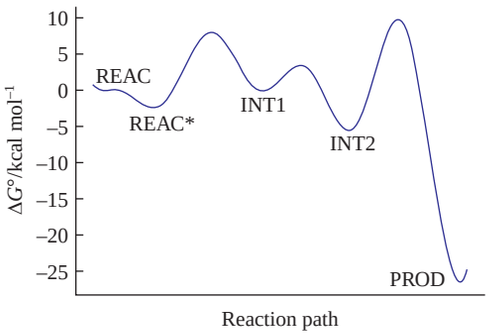


Figure 3 The computed Gibbs energy profile.

The calculated Gibbs energy profile and the corresponding structural changes are shown in Figure S4. It is important to note that both conformations, INT1 and INT2, remain stable for 13 ps along the unconstrained QM/MM MD trajectories (see Figure S4). The activation barrier at this step is 3 kcal mol^{-1}, the energy of INT2 is −5 kcal mol^{-1} relative to the initial level of REAC.

The last step INT2 → PROD uses the collective variable CV4 which takes into account the final distribution of protons to arrive at the products, CoA and *N*-acetylaspartic acid:

$$CV4 = 0.5d_{H(3)–O(3)} – 0.5d_{S–H(3)} + d_{S–C} + 0.5d_{O–H(1)} – 0.5d_{H(1)–O(3)}. \quad (4)$$

The computed Gibbs energy profile and the details of QM/MM MD calculations at this step are presented in Figure S5. The most pronounced energy changes correspond to the formation of products from the tetrahedral intermediate INT2: the activation barrier is 14 kcal mol^{-1}, and the energy of PROD is 25 kcal mol^{-1} lower than the initial level of REAC. The calculated Gibbs energy diagram of the reaction is shown in Figure 3.

In summary, the results of the described QM/MM MD simulations allow us to propose a detailed mechanism for the reaction of NAA synthesis in the NAT8L active site from aspartate and acetyl-CoA. A distinctive feature of this reaction is substrate-assisted catalysis, according to which no external general base is involved in the aspartate deprotonation step. The estimated activation barriers in the Gibbs energy profile of the reaction do not exceed 14 kcal mol^{-1}, which is consistent with the observed catalytic activity of NAT8L.[5] We emphasize that the molecular models were constructed entirely from the primary sequence of the protein, which paves the way for *ab initio* computational enzymology.

This work was supported by the Russian Science Foundation (project no. 18-13-00030). The research was performed on the equipment of the shared research facilities of HPC computing resources at M. V. Lomonosov Moscow State University.

Online Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi: 10.1016/j.mencom.2022.11.010.

References

1 J. R. Moffett, B. Ross, P. Arun, C. N. Madhavarao and A. M. A. Namboodiri, *Prog. Neurobiol.*, 2007, **81**, 89.
2 P. S. Ariyannur, J. R. Moffett, P. Manickam, N. Pattabiraman, P. Arun, A. Nitta, T. Nabeshima, C. N. Madhavarao and A. M. A. Namboodiri, *Brain Res.*, 2010, **1335**, 1.
3 M. S. Hussain, Q. Wang and R. E. Viola, *Arch. Biochem. Biophys.*, 2021, **703**, 108870.
4 F. Dyda, D. C. Klein and A. B. Hickman, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 81.
5 G. Tahay, E. Wiame, D. Tyteca, P. J. Courtoy and E. Van Schaftingen, *Biochem. J.*, 2012, **441**, 105.
6 I. V. Polyakov, A. E. Kniga, B. L. Grigorenko and A. V. Nemukhin, *ACS Chem. Neurosci.*, 2020, **11**, 2296.
7 S. D. Varfolomeev, V. I. Bykov, N. A. Semenova and S. B. Tsybenova, *ACS Chem. Neurosci.*, 2020, **11**, 763.
8 S. D. Varfolomeev, V. I. Bykov and S. B. Tsybenova, *Dokl. Biochem. Biophys.*, 2020, **492**, 147 (*Dokl. Ross. Akad. Nauk. Nauki o Zhizni*, 2020, **492**, 305).
9 W. Dall'Acqua and P. Carter, *Protein Sci.*, 2000, **9**, 1.
10 M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, F. Neese, K. Schulten and Z. Luthey-Schulten, *Nat. Methods*, 2018, **15**, 351.
11 M. G. Khrenova, I. V. Polyakov and A. V. Nemukhin, *Khim. Fiz.*, 2022, **41** (6), 65 (in Russian).
12 J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot and E. Tajkhorshid, *J. Chem. Phys.*, 2020, **153**, 044130.
13 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781.
14 S. Seritan, C. Bannwarth, B. S. Fales, E. G. Hohenstein, C. M. Isborn, S. I. L. Kokkila-Schumacher, X. Li, F. Liu, N. Luehr, J. W. Snyder, Jr., C. Song, A. V. Titov, I. S. Ufimtsev, L.-P. Wang and T. J. Martínez, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2021, **11**, e1494.
15 M. G. Khrenova, V. G. Tsirelson and A. V. Nemukhin, *Phys. Chem. Chem. Phys.*, 2020, **22**, 19069.
16 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell, Jr., *J. Chem. Theory Comput.*, 2012, **8**, 3257.
17 E. J. Denning, U. D. Priyakumar, L. Nilsson and A. D. Mackerell, Jr., *J. Comput. Chem.*, 2011, **32**, 1929.
18 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
19 K. Vanommeslaeghe and A. D. MacKerrel, Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3144.
20 K. Vanommeslaeghe, E. P. Raman and A. D. MacKerrel, Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3155.
21 J. Kästner, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 932.
22 J. Kästner and W. Thiel, *J. Chem. Phys.*, 2005, **123**, 144104.

Received: 23rd April 2022; Com. 22/6878