Supporting Information

for

Structure of the Brain N-Acetylaspartate Biosynthetic Enzyme NAT8L Revealed by Computer Modeling

Igor V. Polyakov, [a,b] Artem E. Kniga, [a] Bella L. Grigorenko, [a,b] Alexander V. Nemukhin *,[a,b]

[a] Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991, Russian Federation

[b] Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 119334, Russian Federation

* Corresponding author: Dr. Alexander V. Nemukhin, e-mail: anem@lcc.chem.msu.ru, anemukhin@yahoo.com

Section S1. The salt-bridge pattern Glu101-Arg220 in NAT8L.

Fig. 3 in the main text shows the residues including Glu101 and Arg220, importance of which was disclosed in experimental studies. Our simulations demonstrate the significance of Arg220 as a residue responsible for positioning acetyl-CoA, whereas Glu101 forms a salt-bridge pattern with Arg220. Fig. S1 illustrates that this salt-bridge is stable along the molecular dynamics trajectory.

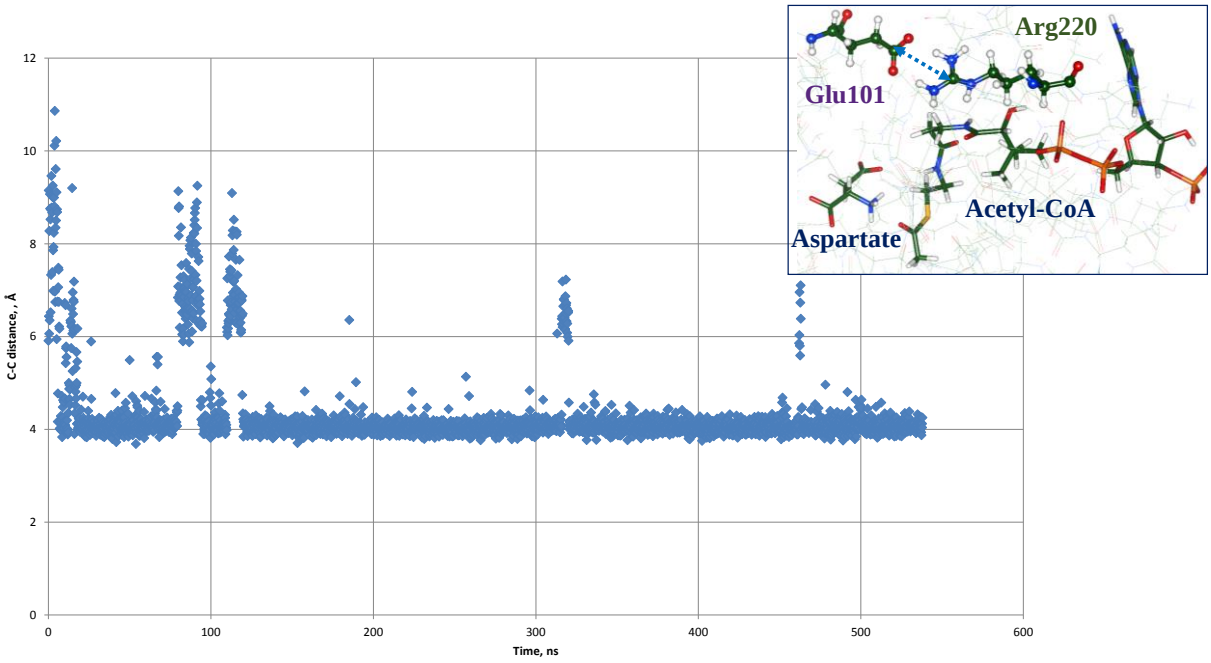

Figure S1. Distance between the Cδ atom of Glu101 and CZ atom of Arg220 along the molecular dynamics trajectory showing the stability of the stable salt-bridge interaction.

Section S2. Details of structure predictions (see section Methods in the main text).

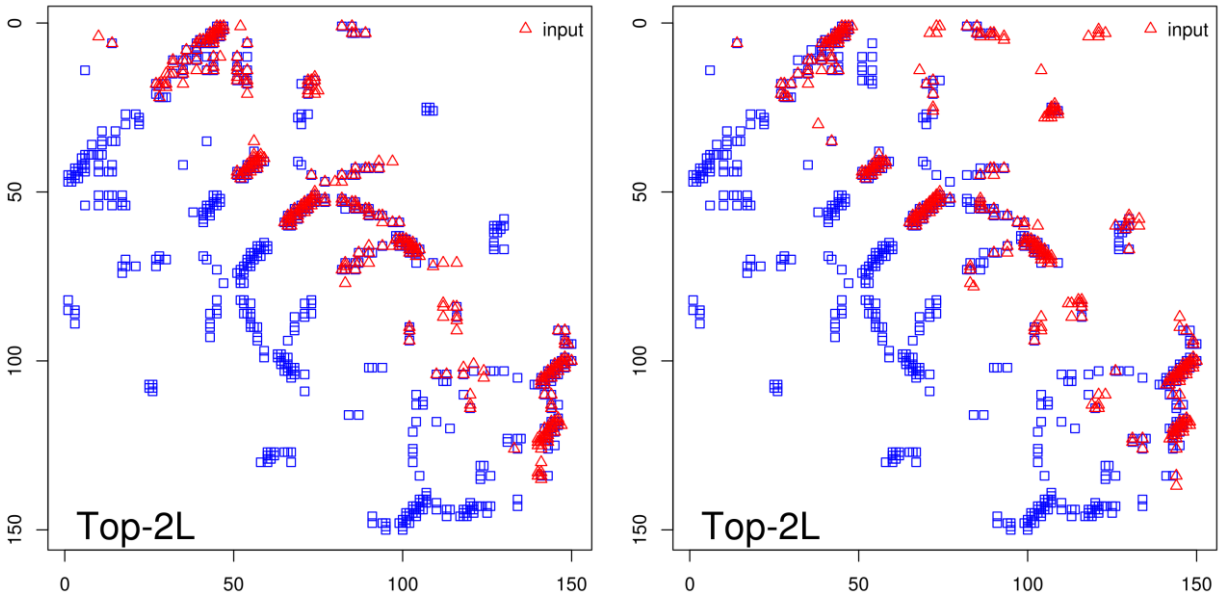

Figure S2. Predicted (red) top-2L filtered contacts in NAGS compared to the contacts in the PDB ID 3B8G structure (blue). Left panel: predictions with TripletRes. Right panel: results of ResPRE with our MSA.

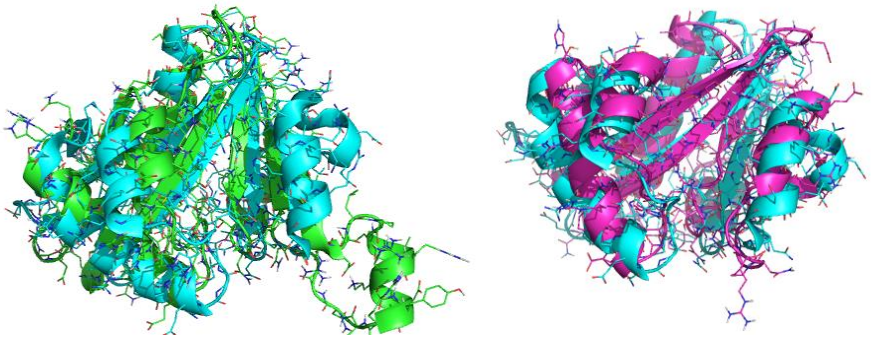

Figure S3. Alignment of the CONFOLD models for NAGS with the reference structure PDB ID 3B8G shown in cyan. Left: the model built on the TripletRes server contacts (green); right: the model built on the ResPRE contacts (purple).

Table S1. Templates used to produce the NAT8L model with the RosettaCM protocol.

PDB ID	PSI-BLAST	Needleman-Wunsch	Smith-Waterman	MapAlign identity, %
1YVK	c:61 pi:25.42	# Identity: 29/172 (16.9%) # Similarity: 48/172 (27.9%) # Gaps: 60/172 (34.9%)	# Identity: 20/66 (30.3%) # Similarity: 31/66 (47.0%) # Gaps: 1/66 (1.5%)	13.22
1Z4R	c:80 pi:19.61	# Identity: 26/186 (14.0%) # Similarity: 44/186 (23.7%) # Gaps: 83/186 (44.6%)	# Identity: 25/106 (23.6%) # Similarity: 41/106 (38.7%) # Gaps: 14/106 (13.2%)	14.05
2WPX	c:73 pi:14.29	# Identity: 30/348 (8.6%) # Similarity: 50/348 (14.4%) # Gaps: 236/348 (67.8%)	# Identity: 19/68 (27.9%) # Similarity: 28/68 (41.2%) # Gaps: 2/68 (2.9%)	14.05
3B8G (ngNAGS)	NA	# Similarity: 46/478 (9.6%) # Gaps: 379/478 (79.3%)	# Identity: 11/47 (23.4%) # Similarity: 23/47 (48.9%) # Gaps: 0/47 (0.0%)	13.22
3R9F	NA	# Identity: 24/212 (11.3%) # Similarity: 48/212 (22.6%) # Gaps: 115/212 (54.2%)	# Identity: 10/44 (22.7%) # Similarity: 22/44 (50.0%) # Gaps: 1/44 (2.3%)	11.57
3TE4	c:74 pi:21.11	# Identity: 24/265 (9.1%) # Similarity: 44/265 (16.6%) # Gaps: 194/265 (73.2%)	# Identity: 23/91 (25.3%) # Similarity: 33/91 (36.3%) # Gaps: 43/91 (47.3%)	14.88
4U9V	NA	# Identity: 21/219 (9.6%) # Similarity: 48/219 (21.9%) # Gaps: 120/219 (54.8%)	# Identity: 12/51 (23.5%) # Similarity: 21/51 (41.2%) # Gaps: 12/51 (23.5%)	13.22
4UA3	c:85 pi:14.02	# Identity: 22/215 (10.2%) # Similarity: 42/215 (19.5%) # Gaps: 115/215 (53.5%)	# Identity: 18/92 (19.6%) # Similarity: 32/92 (34.8%) # Gaps: 21/92 (22.8%)	11.57
4X5K	NA	# Identity: 22/193 (11.4%) # Similarity: 35/193 (18.1%) # Gaps: 96/193 (49.7%)	# Identity: 19/91 (20.9%) # Similarity: 31/91 (34.1%) # Gaps: 3/91 (3.3%)	12.40
5ICW	c:88 pi:15.32	# Identity: 32/205 (15.6%) # Similarity: 50/205 (24.4%) # Gaps: 107/205 (52.2%)	# Identity: 28/115 (24.3%) # Similarity: 42/115 (36.5%) # Gaps: 34/115 (29.6%)	16.53

Section S3. RMSD plots of protein backbones in NAGS and NAT8L along molecular dynamics trajectories.

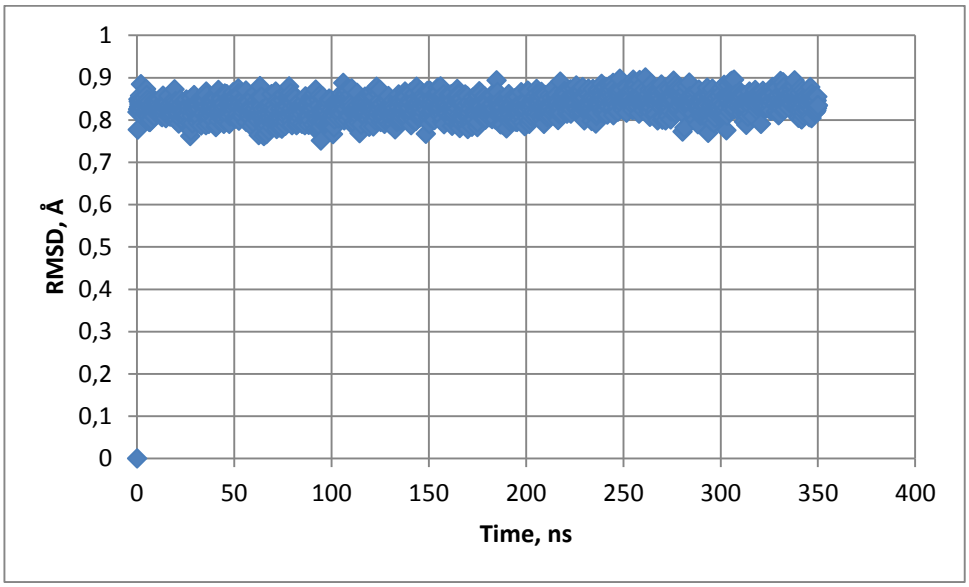

Figure 4S. The trajectory RMSD plot of protein backbone in NAGS.

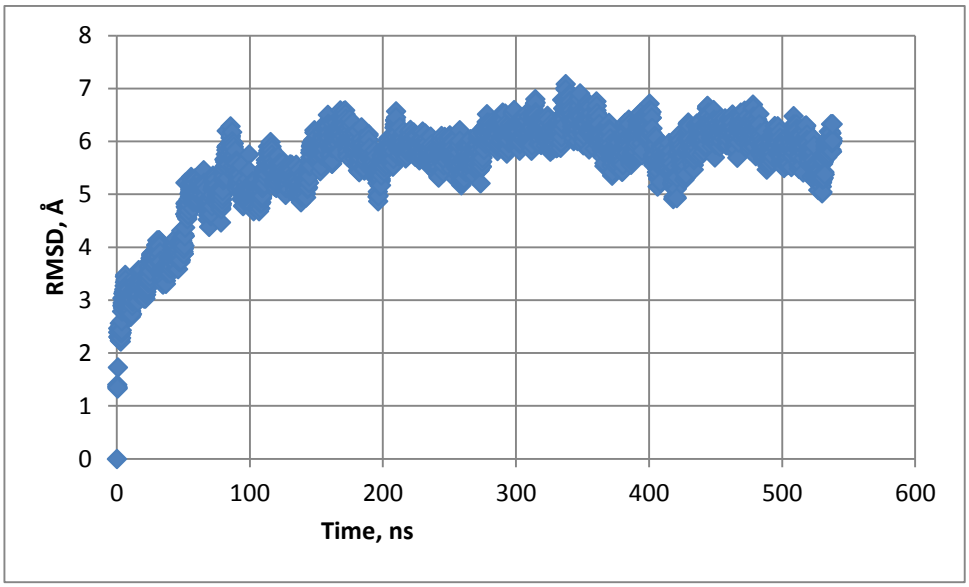

Figure 5S. The trajectory RMSD plot of protein backbone in NAT8L.