

**Supporting Information**  
**For**  
**Unusual Binding-Site-Specific Photophysical Properties of the**  
**Benzothiazole-Based Optical Probe in Amyloid Beta Fibrils**

**N. Arul Murugan<sup>†</sup>, Robert Zalesny<sup>‡</sup>, Hans Ågren<sup>†,II</sup>**

<sup>†</sup> Department of Theoretical Chemistry and Biology, School of Engineering Sciences in  
Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, S-106 91,  
Stockholm, Sweden

<sup>‡</sup> Department of Physical and Quantum Chemistry, Faculty of Chemistry, Wrocław  
University of Technology, Wyb. Wyspiańskiego 27, PL-50370 Wrocław, Poland

<sup>II</sup> Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala,  
Sweden

# 1 Computational details

Below we describe in detail various computational methods employed in this study.

## 1.1 Molecular docking

The molecular structure of BTA-3 was built using the Molden software<sup>1</sup> and the geometry was optimized at the B3LYP/6-31+G\* level of theory as implemented in GAUSSIAN 09 software.<sup>2</sup> Molecular docking of BTA-3 with amyloid fibril has been carried out using AUTODOCK4 software.<sup>3</sup> There are many structures available for amyloid protofibril and we have used the recent structure based on cryogenic-electron microscopy experiments reported for full-length amyloid beta peptide fragment with 42 residues.<sup>4</sup> The reported structure contains nine fragments of amyloid beta peptide where pentamer and tetramer units are intertwined as shown in Figure 2 in the manuscript. The reference ID for this protofibril structure is 5OQV and the resolution for the structure was 4 Å. It is known from the experimental binding assay studies and competition experiments that there are multiple binding sites in the amyloid beta fibril<sup>5-8</sup> and, considering that such information regarding location of the binding sites is not available, we have carried out a blind docking study. The number of grids were chosen as 220×200×130 with a (default grid size of 0.375 Å ) so that both surface and core sites can be identified. At least 500 configurations with larger binding affinity (and least binding free energy) for the complexes were stored for further analysis. It was found that BTA-3 binds to certain sites with larger binding affinity and only the one with least free energy of binding in each sites were further used for the subsequent molecular dynamics simulations. There are at least four potential binding sites and three of them bind to core and the fourth one binding to surface site (refer to Figure 2 of the manuscript).

## 1.2 Molecular dynamics

In order to assess the stability of BTA-3 in different binding sites, molecular dynamics (MD) simulations were carried out by keeping BTA-3 in all four binding sites associated with large binding affinity. Refer to Figure 2 of the manuscript for viewing these high affinity binding sites. A single MD for the fibril complexed with BTA-3 in four different sites has been carried out. The charges for the BTA-3 were obtained using B3LYP/6-31+G\* as implemented in the GAUSSIAN 09 software.<sup>2</sup> In order to mimic protein-like environment, the charge calculations were carried out in chloroform solvent described using polarizable continuum model. The charges are computed by best fitting to the molecular electrostatic potential and by employing the CHELPG procedure<sup>9</sup> as implemented in the GAUSSIAN 09 software. We employed General Amber Force-Field (GAFF)<sup>10</sup> for describing the BTA-3.

For the amyloid fibril, the FF99SB force-field has been used. Each peptide fragment has 3<sup>-</sup> charges and so the whole fibril-BTA-3 system was neutralized with 27 Na<sup>+</sup> ions. As many as 18000 water molecules were added to the complex. The simulations were carried out using AMBER 16 software.<sup>11</sup> The calculations followed the usual protocol for doing MD, starting with energy minimization followed by scaling run and then subsequent finite-temperature and finite-pressure run in isothermal-isobaric ensemble. The calculations were carried out at 1 atm pressure and 300 K. The time step for solving the equation of motion was kept at 2 fs and the length of equilibration run was 5 ns while the production runs were of length scale 50 ns.

### 1.3 Car-Parrinello hybrid QM/MM molecular dynamics

The final configuration from the molecular dynamics has been used as the input structure for four different hybrid QM/MM Car-Parrinello molecular dynamics simulations. The BTA-3 probe in different binding sites has been treated using quantum mechanical theory while the rest of the systems including fibril, solvent and ions were treated using the molecular mechanics force-field as in the aforementioned MD. The interactions between the QM and MM subsystems include dispersion and electrostatic interaction. This type of QM/MM implementation is referred as electrostatic embedding and is available in the CPMD and GROMOS softwares with interface codes.<sup>12-14</sup> The wavefunction is described using plane wave basis sets and with a energy cutoff of 80 Ry. We employ Becke, Lee, Yang and Parr (BLYP) gradient corrected functional<sup>15,16</sup> and the Troullier-Martins norm conserving pseudopotentials<sup>17</sup> to describe the QM region. As the wavefunction for the QM system is optimized in the fibril and solvent environment described using charges and as the system evolves as per the QM/MM Hamiltonian the effect of environment on the electronic structure and molecular structure of BTA-3 probe are accounted for. An orthorhombic simulation box was employed with dimensions 100.2, 94.6, 65.6 Å . The time step for the integration of equation of motion is 5 au while the time scale for the production run was around 30 ps. The calculation involves energy minimization followed by scaling run carried out for a few ps. Final set of simulations were carried out in isothermal ensemble by connecting the system to Nose-Hoover thermostat. Similarly, another pair of hybrid QM/MM molecular dynamics have been carried out for BTA-3 in water solvent and in chloroform solvent for which the starting structure was taken from MD done for a system having single BTA-3 and 33000 water molecules and 2366 chloroform solvents respectively. The simulation box was orthorhombic with dimensions 76.5, 68.9 and 63.7 Å for BTA-3 in water solvent while for the BTA-3 in chloroform solvent the dimensions were 76.0, 66.9 and 63.2Å. Many of the simulation parameters in this set of calculation were set as the same as in the case of BTA-3

in fibril.

## 1.4 Absorption spectra calculations using TD-DFT/MM and RI-CC2/MM approaches

The absorption spectra calculations were carried out using time-dependent density functional theory/molecular mechanics approach which explicitly treats the interaction between the probe and fibril using an electrostatic embedding scheme. The charges for the fibril and solvent were based on the molecular mechanics force-field as employed in the previous set of MD simulations. In particular two levels of theory have been employed namely B3LYP/TZVP and CAM-B3LYP/TZVP. The choice of basis sets was based on our previous experience with one photon spectra calculations.<sup>18-22</sup> Both levels of theory provided the spectra with similar features but the spectra from latter theory was blue-shifted and so for the analysis we only used the results from B3LYP level of theory. The reason for using B3LYP level of theory was justified due its closer agreement with the results from CC2/MM level of theory. Further the computed spatial overlap,  $\Lambda$  justifies (refer to Table S1 below) the use of B3LYP level of theory for describing the low frequency electronic excitation for BTA-3. The value for  $\Lambda$  diagnostics using B3LYP level of theory is larger than 0.6 which suggests that the excitation energies computed at this level of theory will be reliable<sup>23</sup>. The electronic excitations corresponding to the first two bands in the absorption spectra of BTA-3 in water solvent were dominated by HOMO $\rightarrow$ LUMO and HOMO-1 $\rightarrow$ LUMO excitations. The frontier molecular orbitals involved in these low energy excitations are shown in Figure S1.

We also employed the RI-CC2 method<sup>24</sup> and the def-SVP basis set<sup>25</sup> for calculating the one photon absorption spectra of BTA-3 in fibril and in water solvent. The TURBOMOLE program was employed to carry out these calculations.<sup>26,27</sup>

Table S1: Comparison of the absorption maximum as obtained from B3LYP and CAM-B3LYP levels of theory compared to CC2 approach. Also the spatial overlap,  $\Lambda$  computed using two different levels of density functional theory is given.

	excitation wavelength [nm]	$\Lambda$
CAM-B3LYP	514	0.61
B3LYP	580	0.65
CC2	616	—

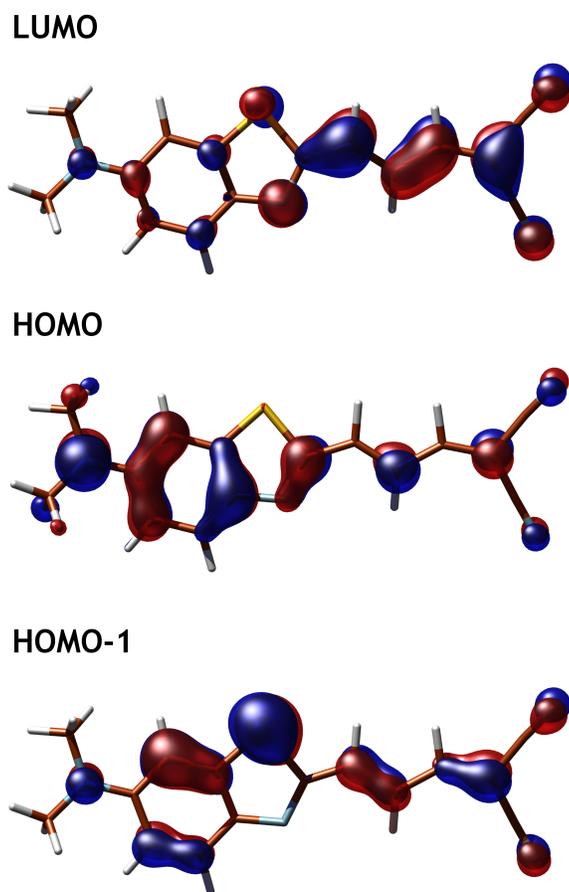


Figure S1: The molecular orbitals involved in the excitations corresponding to first two bands in the absorption spectra of BTM-3 in water. These results are based on single point calculation at TD-DFT/PCM level of theory for the optimized geometry of BTM-3 in water.

## 1.5 Molecular mechanics-Generalized Born Surface Area approach

In this study, free energy calculations were carried out by using molecular mechanics-generalized Born surface area (MM-GBSA)<sup>28</sup> approach to estimate the relative binding affinity of BTM-3 in different binding sites of amyloid fibril. The trajectories for the complex, where BTM-3 bound to different sites in fibril, were extracted independently and the free energy of binding in each of these sites were computed. The free energy of binding is computed as the difference between the free energy of complex and free energy of individual systems namely amyloid fibril and BTM-3 ligand. In this approach, the solvation contribution is computed using an implicit solvent model and so the solvents and ions were usually stripped out and only the coordinates of the complex as extracted from MD are used in the free energy calculations. The free energies for the individual systems as well as for the complex have contributions from van der Waals, electrostatic, and polar and non-polar solvation free energies. The polar solvation free energy is computed by solving the generalized Born equa-

tion. The non-polar part of the solvation free energy is computed using the solvent accessible surface area. The free energies are obtained as an average over 1000 configurations.

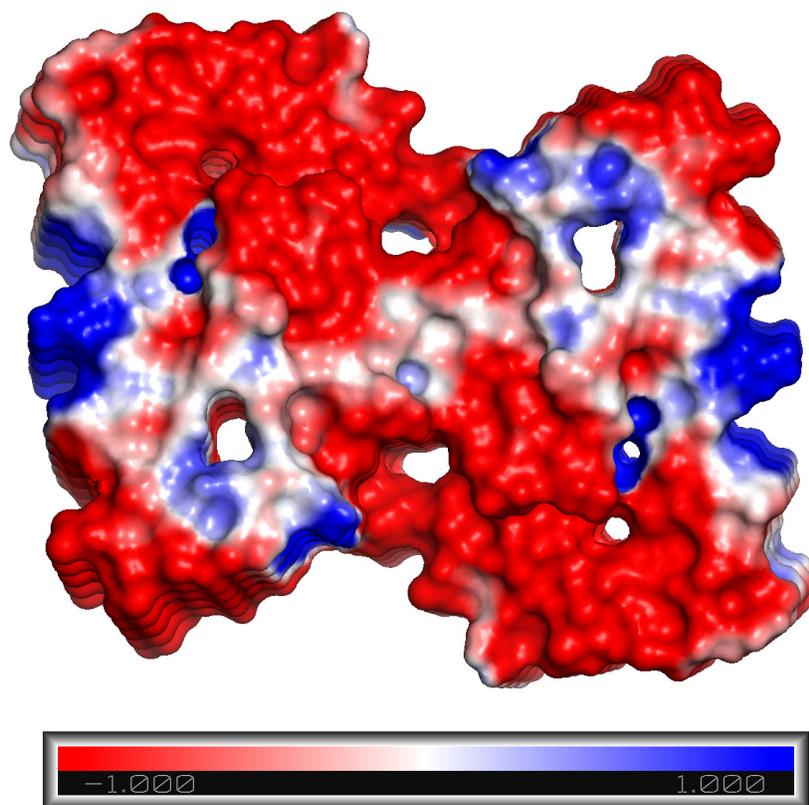


Figure S2: Electrostatic potential calculated for amyloid beta (1-42) fibril

# References

- [1] G. Schaftenaar and J. Noordik, *J. Comput.-Aided Mol. Design*, 2000, **14**, 123–134.
- [2] 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.
- [3] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **16**, 2785–91.
- [4] L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold and G. F. Schröder, *Science*, 2017, **358**, 116–119.
- [5] H. Levine, III, *Amyloid*, 2005, **12**, 5–14.
- [6] A. Lockhart, L. Ye, D. B. Judd, A. T. Merritt, P. N. Lowe, J. L. Morgenstern, G. Hong, A. D. Gee and J. Brown, *J. Biol. Chem.*, 2005, **280**, 7677–7684.
- [7] R. Ni, P.-G. Gillberg, A. Bergfors, A. Marutle and A. Nordberg, *Brain*, 2013, **136**, 2217–2227.
- [8] R. Kawai, M. Araki, M. Yoshimura, N. Kamiya, M. Ono, H. Saji and Y. Okuno, *ACS Chem. Neurosci.*, 2018, 957–966.
- [9] C. Breneman and K. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361–373.
- [10] J. Wang, R. Wolf, J. Caldwell, P. Kollman and D. Case, *J. Comput. Chem.*, 2004, **34**, 1157–1174.
- [11] D. Case, R. Betz, D. Cerutti, T. Cheatham, III, T. Darden, T. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao and P. Kollman, *Amber 16, University of California, San Francisco.*, 2016.
- [12] J. Hutter, M. Parrinello, D. Marx, P. Focher, M. Tuckerman, W. Andreoni, A. Curioni, E. Fois, U. Röthlisberger, P. Gianozzi, T. Deutsch, A. Alavi, D. Sebastiani, A. Laio, J. VandeVondele, A. Seitsonen and S. Billeter, *Computer code CPMD, version 3.11*, Copyright IBM Corp. and MPI-FKF Stuttgart, 1990–2002.
- [13] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, **1996**.
- [14] A. Laio, J. VandeVondele and U. Rothlisberger, *J. Chem. Phys.*, 2006, **116**, 6941.
- [15] A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
- [16] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- [17] N. Trouiller and J. Martins, *Phys. Rev. B*, 1991, **43**, 1993.
- [18] N. A. Murugan, K. Aidas, J. Kongsted, Z. Rinkevicius and H. Ågren, *Chem. Eur. J.*, 2012, **18**, 11677–11684.
- [19] N. Murugan, J. Kongsted, Z. Rinkevicius, K. Aidas and H. Ågren, *J. Phys. Chem. B.*, 2010, **114**, 13349–13357.
- [20] N. A. Murugan, J. Kongsted, Z. Rinkevicius and H. Ågren, *Phys. Chem. Chem. Phys.*, 2012, **14**, 1107–1112.

- [21] N. A. Murugan, J. M. H. Olsen, J. Kongsted, Z. Rinkevicius, K. Aidas and H. Ågren, *J. Phys. Chem. Lett.*, 2013, **4**, 70–77.
- [22] K. Rajasekhar, N. Narayanaswamy, N. A. Murugan, K. Viccaro, H.-G. Lee, K. Shah and T. Govindaraju, *Biosens. Bioelectron.*, 2017, **98**, 54–61.
- [23] M. J. G. Peach, P. Benfield, T. Helgaker and D. J. Tozer, *J. Chem. Phys.*, 2008, **128**, 044118.
- [24] C. Hättig and F. Weigend, *J. Chem. Phys.*, 2000, **113**, 5154–5161.
- [25] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- [26] *TURBOMOLE V6.5, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>, 2013.*
- [27] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, *WIREs Comput. Mol. Sci.*, 2014, **4**, 91–100.
- [28] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.