Supporting Text to *In silico* study of amyloid β-protein folding and oligomerization

B. Urbanc* [1], L. Cruz*, S. Yun*, S. V. Buldyrev*, G. Bitan‡, D. B. Teplow‡, and H. E. Stanley*

*Center for Polymer Studies, Department of Physics, Boston University, Boston, MA 02215;

‡Center for Neurologic Diseases, Brigham and Women's Hospital, and Department for Neurology, Harvard Medical School, Boston, MA 02115.

I. METHODS

A. Hydrogen bond interaction parameters

The hydrogen bond between the N_i atom and the C'_j atom is formed if the two atoms are at a distance of 4.0 Å $< r_{ij} <$ 4.2 Å and the four pairs of atoms N_i—$C_{\alpha\ j}$, N_i—N_{j+1}, $C_{\alpha\ i+1}$—C'_j, and C'_i—C'_j are simultaneously at energetically favorable distances (Fig. S-1a). These constraints are introduced to correctly model strong angular dependence of the hydrogen bond. The energetics of the constraints is governed by a double-step potential U(r) (Fig. S-1b). The parameters of this potential are inter-atom distances of r_{min}, r_0, r_1, and r_{max}. The change in the potential energy at these distances is proportional to the hydrogen bond potential energy E_{HB}, as shown in Fig. S-1b. The distances r_{min}, r_0, r_1, and r_{max} are determined empirically by making statistical averages over the corresponding inter-atom distances within the hydrogen bond using a database of about 7700 proteins with known all-atom structures from PDB. The values of these distances used in this study are given in Table S-I. A more detailed description of the hydrogen bond implementation is given in Ref. [30] of the manuscript.

1. Matrix of amino acid-specific interactions

The matrix of hydrophobic/hydrophilic interactions is given in Table S-II, where the relative strengths of effective hydrophobic attraction (negative values) and effective hydrophilic

[1] Corresponding author. Email: brigita@bu.edu

repulsion (positive values) between the pairs of side chain atoms are shown. The amino acid-specific attraction or repulsion takes place only if the center-to-center distance between the two side chain atoms is smaller than 0.75 nm. The relative hydropathies of individual amino acids are based on the Kyte-Doolittle hydropathy scale (Ref. [35] of the manuscript). The interaction strength zero is assigned to pairs that interact only through hard-core and hydrogen bond interaction. The real strength of the amino acid-specific interactions is obtained by multiplying the relative strengths given in Table S-II by the strength of the hydropathic interaction, which in our study is equal to 0.3 E_{HB}.

II. RESULTS

A. Time dependence of monomer and oligomer numbers

In Fig. S-2 we show the time-dependence of the number of monomers, dimers, trimers, tetramers, pentamers and higher order oligomers (hexamers and larger oligomers). We find significant differences in occurrence probabilities for dimers and pentamers. In addition, Aβ40 has a tendency to form more tetramers relative to Aβ42, which is most likely a consequence of increased level of dimers in Aβ40 (two dimers on further association form a tetramer). There is also a consistent tendency of Aβ42 to form more higher-order oligomers (hexamers, heptamers, etc) than Aβ40.

B. Secondary structure analysis

In Fig. S-3 we present the average propensities for turn formation (left panels) and β-strand formation (right panels) in monomers, dimers, trimers, tetramers, and pentamers of Aβ40 (black curve) and Aβ42 (red curves). The average turn propensity of a folded monomer shows that the Aβ42 monomer is characterized by a turn centered at Gly37-Gly38 (the TRB region), while in Aβ40 monomer, this turn is not significant. Comparing the turn propensities within the TRB region (Val36-Val39) in monomers ($N = 1$), dimers ($N = 2$), trimers ($N = 3$), tetramers ($N = 4$), and pentamers ($N = 5$) shows that in Aβ42 the turn is equally present in all assembly states, while in Aβ40, the turn becomes significant in dimers and then gets more pronounced as the oligomer size increases.

Examination of the average β-strand propensity of folded monomers reveals that the region Ala2-Phe4 in Aβ40 has a pronounced β-strand structure which is absent in Aβ42. This difference persists in dimers, trimers, tetramers and pentamers. In addition, Aβ42 monomers have more β-strand structure within the MHR (Gly33-Met35) than Aβ40. This difference disappears as the peptides start to assemble. There is a significant β-strand structure in the CTR (Val39-Ala42) in all Aβ42 oligomers (dimers through pentamers) which is almost absent in Aβ40.

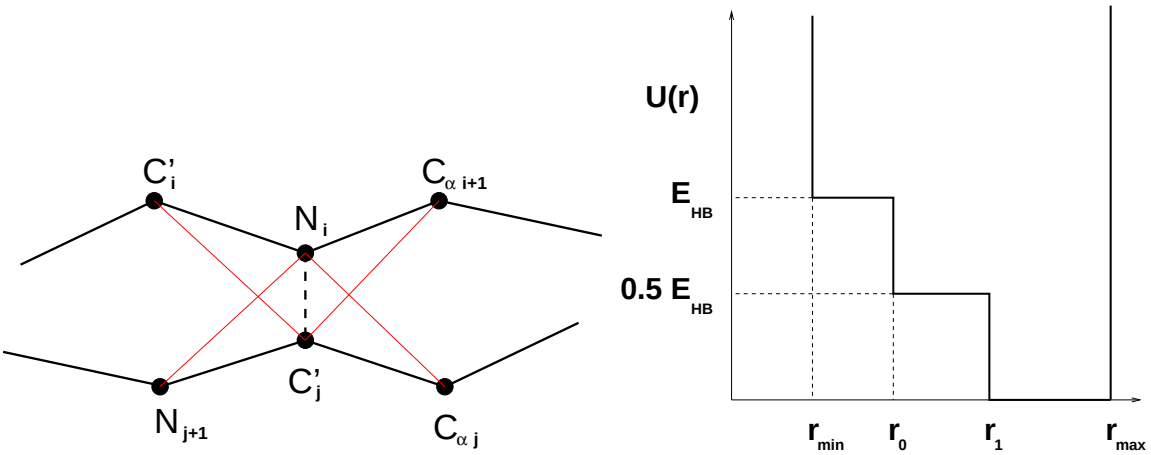
C. Analysis of intra- and intermolecular contacts in monomers and oligomers

Fig. S-4 (columns 1 and 2) shows intramolecular contact maps of monomer, dimer, trimer, tetramer, and pentamer conformations, averaged over three fixed simulation steps (9, 9.5, and 10 million simulation steps). The main feature of the intramolecular contact maps is a turn-like element centered at Gly25-Ser26 (black squares). There are strong contacts centered around Val36-Val39 (red squares), which are more prominent in Aβ42. Comparison of Aβ40 and Aβ42 (Fig. S-4, column 1 and 2) shows that there is a difference in intramolecular contacts at the N-terminus—in Aβ40 the N-terminus is in contact with the CHC (Leu17-Ala21) as well as with the C-terminal hydrophobic region (Ala30-Val40), while in Aβ42 the N-terminus has no contacts with the CHC and has significantly less contacts with the region Ala30-Val40. This difference between Aβ40 and Aβ42 is more pronounced in tetramers and pentamers (Fig. S-4, rows $N = 4$ and $N = 5$).

Fig. S-4 (columns 3 and 4) shows intermolecular contact maps of dimer, trimer, tetramer, and pentamer conformations, averaged over three fixed simulation steps (9, 9.5, and 10 million simulation steps). The intermolecular contact maps of Aβ40 and Aβ42 show distinct behavior. Similar to the differences in intramolecular contact maps described above, there are less contacts of the NTR with the CHC, MHR, and CTR in Aβ42 compared to Aβ40 (Fig. S-4, columns 3 and 4, black rectangles). In addition, the strongest contacts in Aβ40 are connecting pairs of CHC regions (Fig. S-4, column 3, red squares), while the contacts of the CTR with the CHC and the MHR dominate in Aβ42 (Fig. S-4, column 4, red rectangles). The relative importance of the hydrophobic CTR for intermolecular contact formation in Aβ42 increases with the oligomer size and is strongest in a pentamer (Fig. S-4, column 4, compare lower parts of the red rectangles).

TABLE I: The inter-atom distances r_{min}, r_0, r_1, and r_{max} as defined in Fig. S-1, associated with the four constraints of the hydrogen bond formation between the N_i atom and the C'_j atom.

Pairs	r_{min} [Å]	r_0 [Å]	r_1 [Å]	r_{max} [Å]
N_i—$C_{\alpha\ j}$	4.43	4.87	4.96	5.34
N_i—N_{j+1}	4.19	4.43	4.65	5.03
$C_{\alpha\ i+1}$—C'_j	4.33	4.67	4.77	5.17
C'_i—C'_j	4.38	4.60	4.66	4.98



(a) Hydrogen bond with constraints (b) Potential for constraints

FIG. 1: (a) The strong angular dependence of the hydrogen bond between the N_i atom and the C'_j atom (dashed black line) is implemented through four constraints (red solid lines). (b) The potential U(r) associated with the four constraints is determined by four inter-atom distances r_{min}, r_0, r_1, and r_{max}.

TABLE II: The matrix of amino acid-specific interaction strengths between pairs of side chains. The attractive hydrophobic interaction has a negative sign, whereas the repulsive hydrophilic interaction has a positive sign. Zero is assigned to those pairs of side chains that do not interact via hydropathic interaction. Because the matrix is by definition symmetric, we present only the strengths above the diagonal.

A.A.	Ile	Val	Leu	Phe	Cys	Met	Ala	Gly	Thr	Ser	Trp	Tyr	Pro	His	Gln	Asn	Glu	Asp	Lys	Arg
Ile	-1.00	-0.97	-0.92	-0.81	-0.78	-0.71	-0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Val	N/A	-0.93	-0.89	-0.78	-0.75	-0.68	-0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Leu	N/A	N/A	-0.84	-0.73	-0.70	-0.63	-0.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Phe	N/A	N/A	N/A	-0.62	-0.59	-0.52	-0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cys	N/A	N/A	N/A	N/A	-0.56	-0.49	-0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Met	N/A	N/A	N/A	N/A	N/A	-0.42	-0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ala	N/A	N/A	N/A	N/A	N/A	N/A	-0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gly	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Thr	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ser	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trp	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tyr	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pro	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
His	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	+0.71	+0.75	+0.75	+0.75	+0.75	+0.79	+0.86
Gln	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	+0.78	+0.78	+0.78	+0.78	+0.83	+0.89
Asn	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	+0.78	+0.78	+0.78	+0.83	+0.89
Glu	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00	0.00
Asp	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00	0.00
Lys	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	0.00
Arg	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00

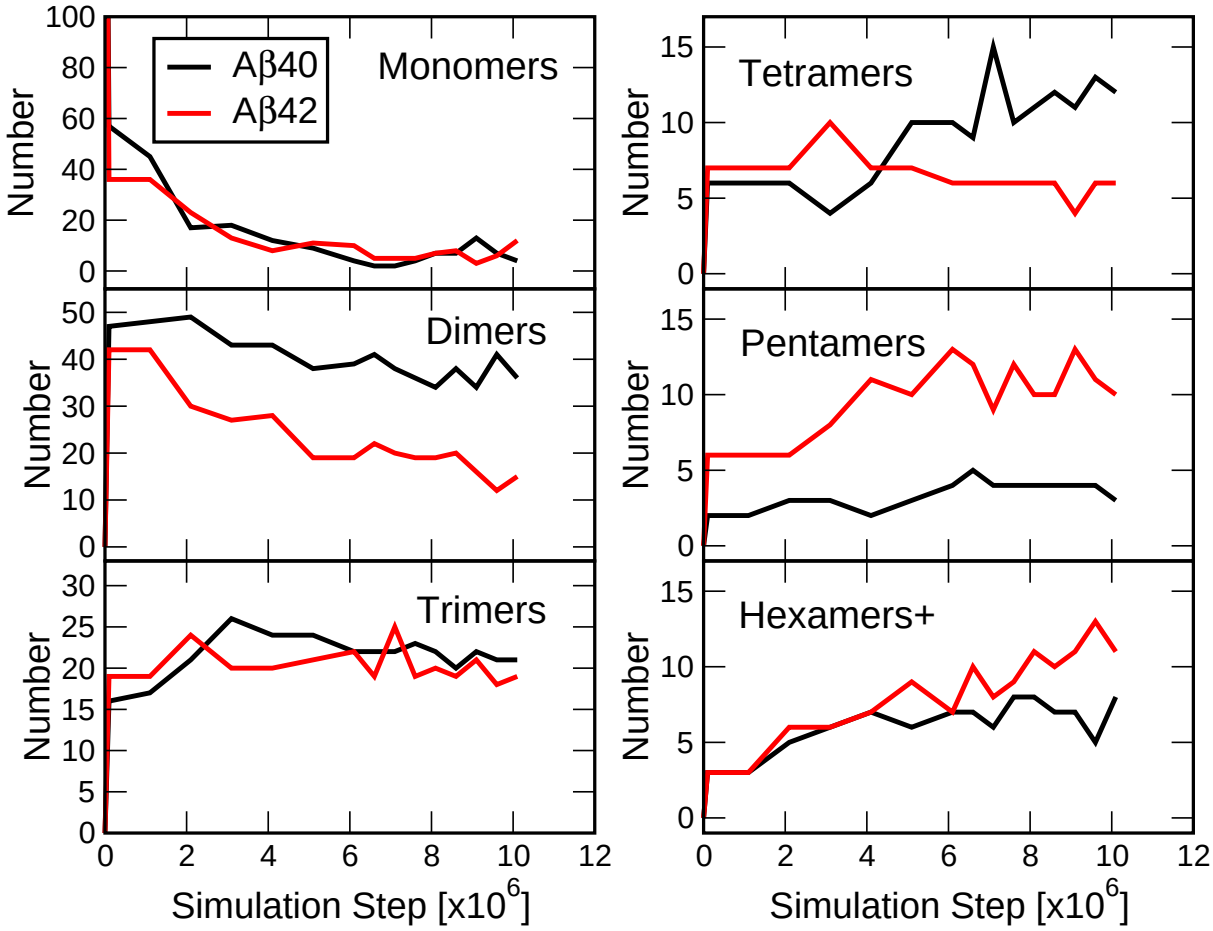


FIG. 2: Time-dependence of monomer and oligomer numbers.

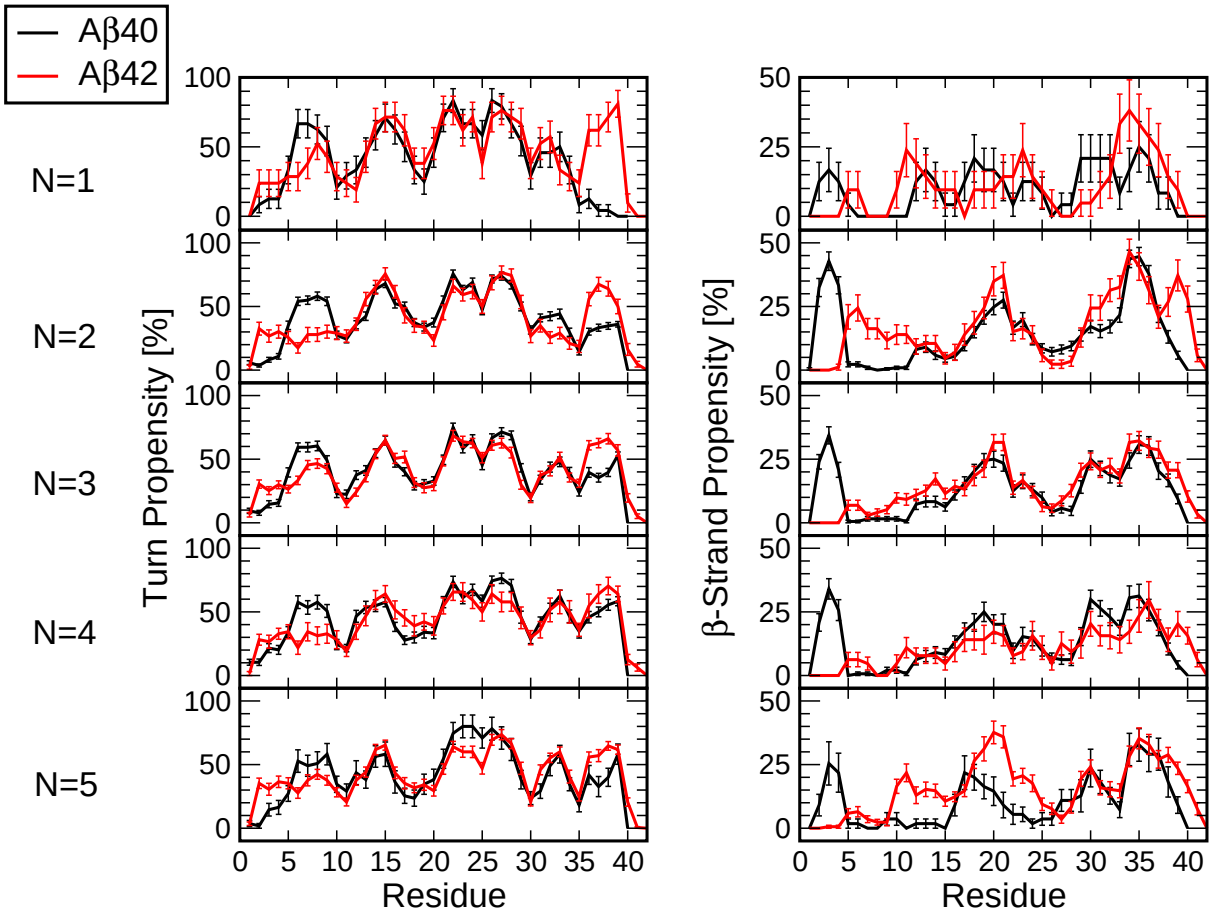


FIG. 3: Secondary structure analysis for different assembly states averaged over 9, 9.5, and 10 million simulation steps. Propensities for two secondary structure elements, a turn (left column) and a β-strand (right column) per residue for Aβ40 (black curve) and Aβ42 (red curve). The secondary structure analysis is done for each assembly state separately: monomers ($N = 1$), dimers ($N = 2$), trimers ($N = 3$), tetramers ($N = 4$), and pentamers ($N = 5$). The secondary structure is calculated using STRIDE (Refs. [39,40] of the manuscript) within the VMD visualization software (Ref. [41] of the manuscript). The error bars represent the standard errors of the mean values.

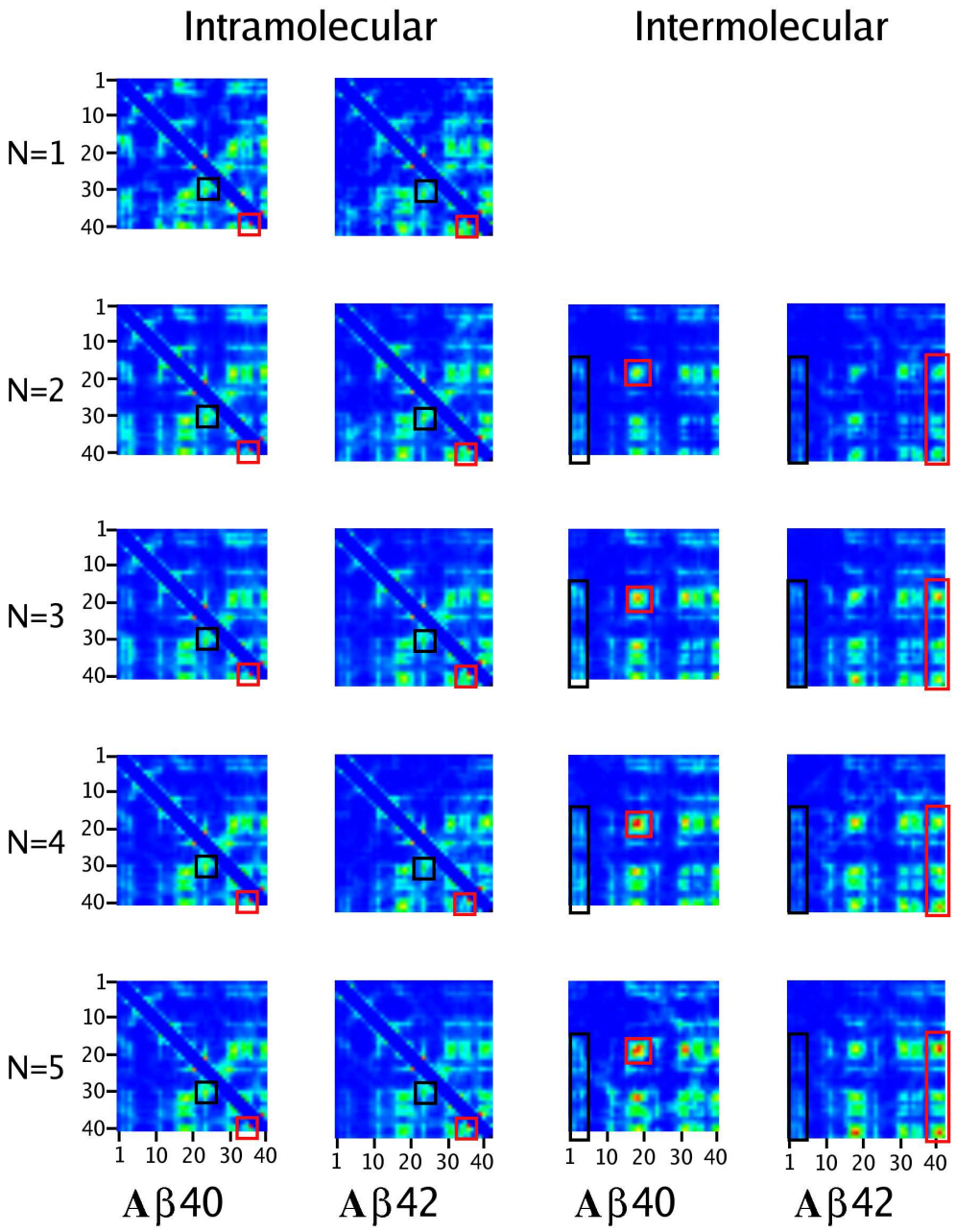


FIG. 4: Intra- and intermolecular contact maps averaged over 9, 9.5, and 10 million simulation steps for each assembly state: monomers ($N = 1$), dimers ($N = 2$), trimers ($N = 3$), tetramers ($N = 4$), and pentamers ($N = 5$). The contact maps are averages over 24/21 monomers, 111/43 dimers, 64/58 trimers, 36/17 tetramers, and 11/34 pentamers of Aβ40/Aβ42. Columns 1 and 3 correspond to Aβ40 and columns 2 and 4 to Aβ42. (Asp1,Asp1) is at the top left corner of the contact maps, while (Val40,Val40) for Aβ40 or (Ala42,Ala42) for Aβ42 is at the bottom right corner. The strength of contacts is color coded as in Fig. 2 of the manuscript. In columns 1 and 2, the black and red squares mark the centers of the TRA and TRB regions, respectively. In columns 3 and 4, the black rectangles mark the contacts of the NTR with the CHC, MHR, and CTR. In column 3, the red squares mark the contacts between pairs of central hydrophobic clusters. In column 4, the red rectangles mark the contacts of the CTR with the CHC and the MHR.