Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2024

Supporting Information

Prediction for 3D conformation of small peptide vaccine targeting Aβ oligomers

Yvning Guan, Jinfei Mei, Xvzhi Gao, Chuanbo Wang, Mengke Jia, Sajjad Ahmad, Fahad Nouman Muhammad and Hongqi Ai*

School of Chemistry and Chemical Engineering, University of Jinan, Jinan 250022, P.R. China

*To whom correspondence should be addressed, E-mail: chm_aihq@ujn.edu.cn

Contents

MM/PBSA Calculational details

Molecular Mechanics/Poisson-Boltz Area (MM mann Surface /PBSA) is a widely used approach to estimate binding free energies by combining molecular mechanics and continuum solvation models [1-3]. The binding free energy (ΔG_{bind}) between pentamer and vaccine in the present study is calculated by the following equations [4].

$$\Delta G_{bind} = \Delta G_{complex} - \Delta G_{rec} - \Delta G_{lig}$$

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S$$

$$\Delta E_{MM} = \Delta E_{VDW} + \Delta E_{COU}$$

$$\Delta G_{solv} = \Delta G_{P\,B} + \Delta G_{SA} = \gamma\ SASA + \beta$$

$$\Delta G_{SA} = \gamma\ SASA + \beta$$

Where $\Delta G_{complex}$, ΔG_{rec}, ΔG_{lig} represent the free energy values of Aβ pentamer-vaccine complex, receptor Aβ pentamer and ligand vaccine. ΔE_{MM} includes two terms, ΔE_{COU} for electrostatic interaction and ΔE_{VDW} for van der Waals interaction. ΔG_{solv} refers to the total solvation free energy, which includes polar solvation energy ΔG_{PB} and nonpolar solvation energy ΔG_{SA}. SASA refers to the solvent accessible surface area with values of constants γ and β as 2.2 kJ/mol.nm^2 and 3.84 kJ/mol, respectively.

The ΔG_{bind} calculated above often overestimates the screening effect to some extent when studying binding ability [5]. Therefore, Huang et al. proposed a new correction method [6] to modify the electrostatic and polar solvent terms, resulting in a highly applicable fitted equation, as shown below. The entropy contribution in the original method is usually obtained by normal mode analysis. Drastically different from the widely employed but extremely expensive normal mode method for calculating entropy change in protein–ligand binding, the calculation of the interaction entropy simply involves the natural log of an ensemble average, which can be readily extracted along with MD simulation without extra computational cost. That is, the fitting method is based on the interaction energy between ligand and receptor, termed interaction entropy[7].

$$\Delta G_{fittd} = 0.0542\ (\Delta E_{COU} + \Delta G_{PB}) + 0.14852\ \Delta E_{VDW} + 0.05584\ \Delta G_{SA} + 0.11351\ (-T\Delta S) - 4.77148$$

ΔG_{fittd} stands for the binding free energy value after fitting. For each complex, the total binding free energy is obtained by summing up the contributions of each residue in the receptor (pentamer) and the ligand (vaccine) to the free energy[8, 9].

- Length: 14
- Number of predicted TMHs : 0
- Exp number of AAs in TMHs : 0
- Exp number, first 60 AAs : 0
- Total prob of N-in : 0.45149
- TMHMM2.0 inside 1 14

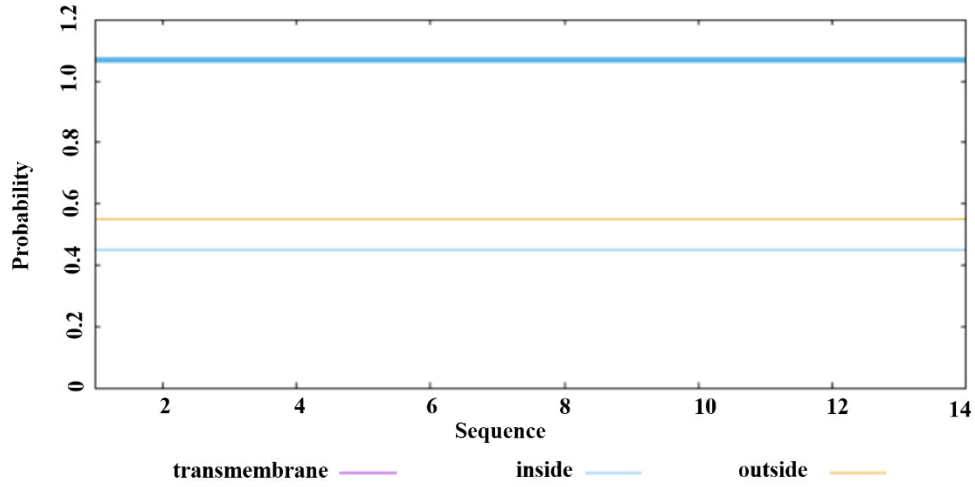

Figure S1 Transmembrane Helix of $A\beta_{1-14}$.

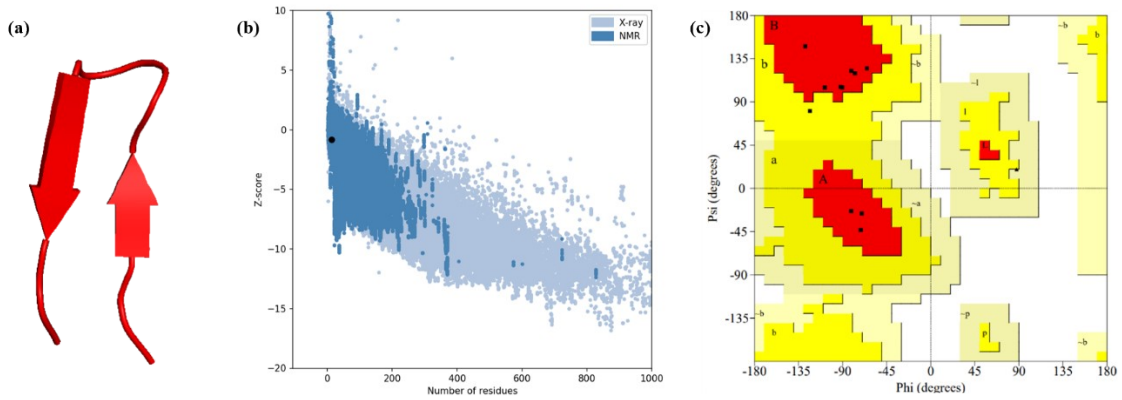

Figure S2 Refined 3D structure from the Galaxy refine web server (a) and structure validation (b, c) of model1$A\beta_{1-14}$. (b) ProSA validation of predicted structure with Z score of -0.84, (c) Ramachandran plot analysis indicates 90.9% of residues are present in the most favored regions, 9.1% of residues are found in the allowed regions.

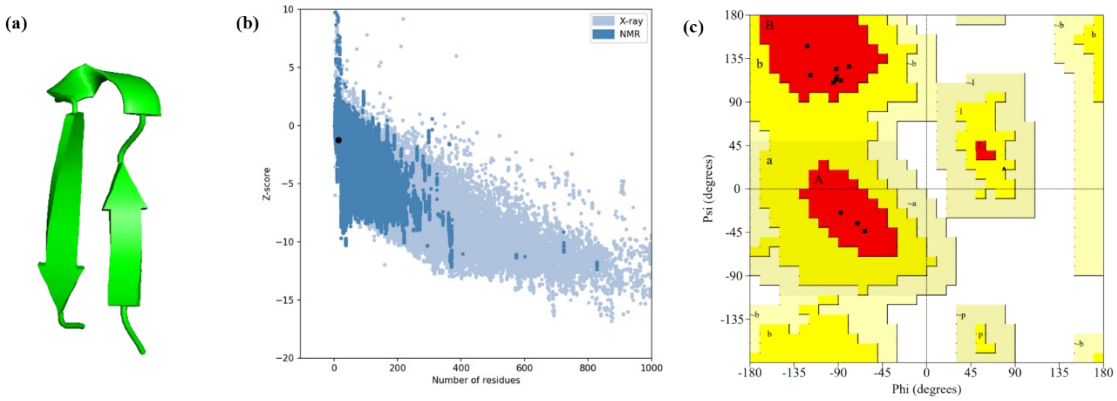

Figure S3 Refined 3D structure from the Galaxy refine web server (a) and structure validation (b, c) of model2A$\beta_{1\text{-}14}$. (b) ProSA validation of predicted structure with Z score of -1.23, (c) Ramachandran plot analysis indicates 100% of residues are present in the most favored regions, 0% of residues are found in the allowed regions.

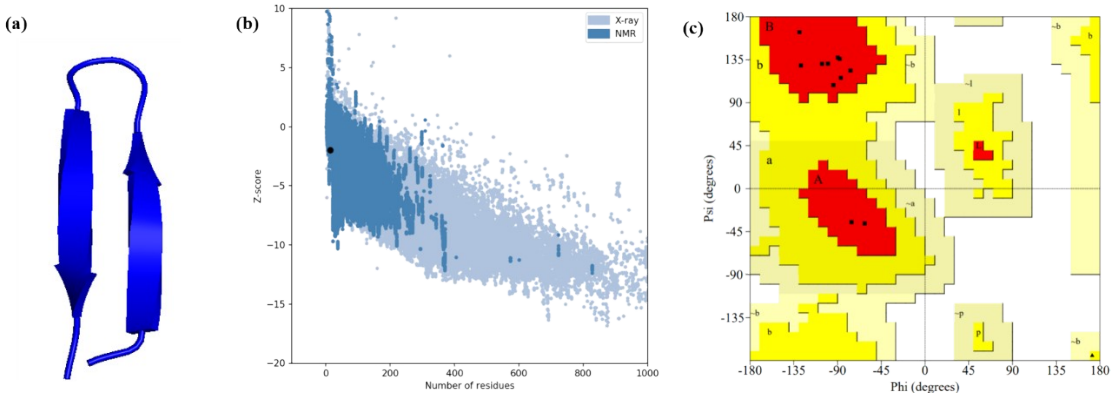

Figure S4 Refined 3D structure from the Galaxy refine web server (a) and structure validation (b, c) of model3A$\beta_{1\text{-}14}$. (b) ProSA validation of predicted structure with Z score of -1.99, (c) Ramachandran plot analysis indicates 100% of residues are present in the most favored regions, 0% of residues are found in the allowed regions.

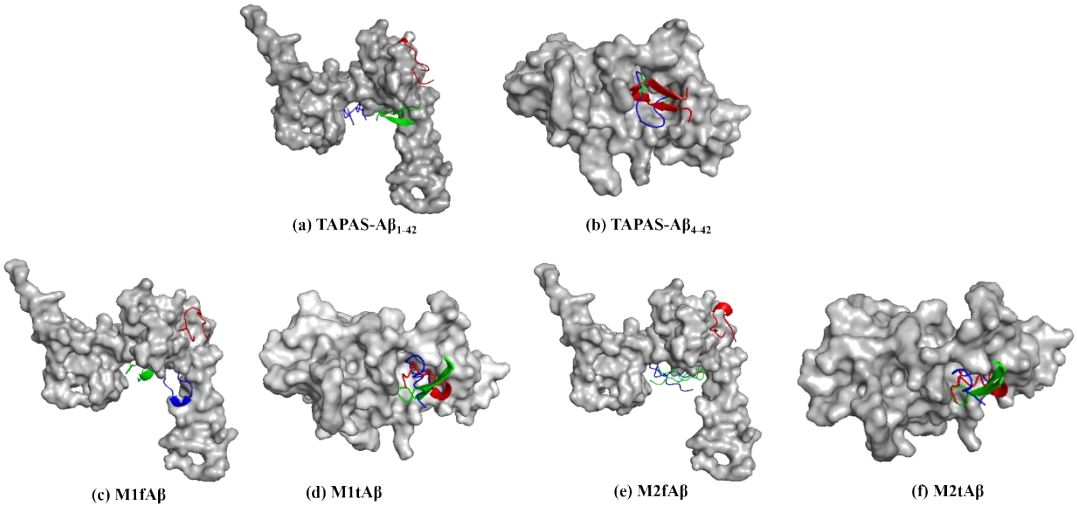

Figure S5 Top 3 positions of two Aβ pentamers (fAβ and tAβ) for vaccines (TAPAS, M1 and M2) docking, in which the three top poses of the vaccine are highlighted in red for Top 1, green for Top 2, and blue for Top 3, respectively.

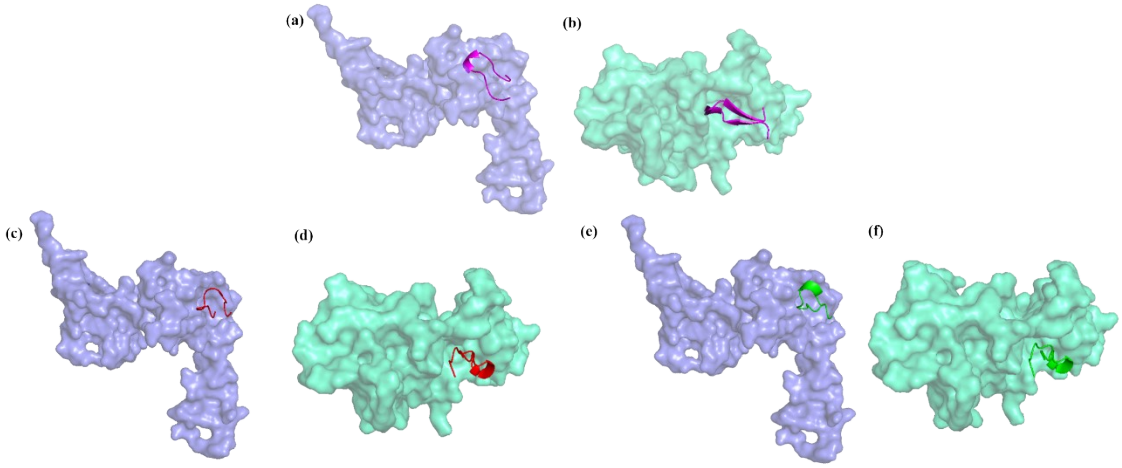

Figure S6 Docked complexes of TAPAS and model*i'*$A\beta_{1-14}$ (*i'*=1 and 2) vaccines to Aβ receptors. (a) and (b) are docked complexes of TAPAS docked with receptors $A\beta_{1-42}$ and $A\beta_{4-42}$, respectively; (c) and (d) stand for the docked complexes of model1$A\beta_{1-14}$ docked with the two receptors, respectively; (e) and (f) refer to the docked complexes of model2$A\beta_{1-14}$ docked with two receptors, respectively. TAPAS, model1$A\beta_{1-14}$, model2$A\beta_{1-14}$, $A\beta_{1-42}$ and $A\beta_{4-42}$ are represented in magenta, red, blue, slate and cyan, respectively.

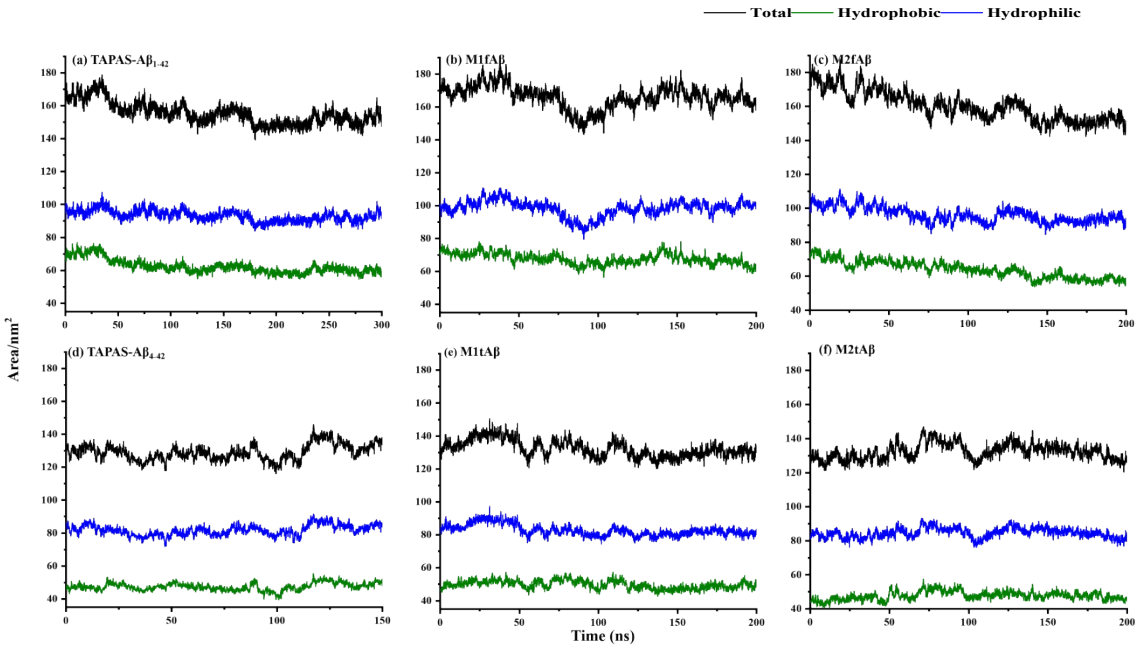

Figure S7 SASA results of the six complexes

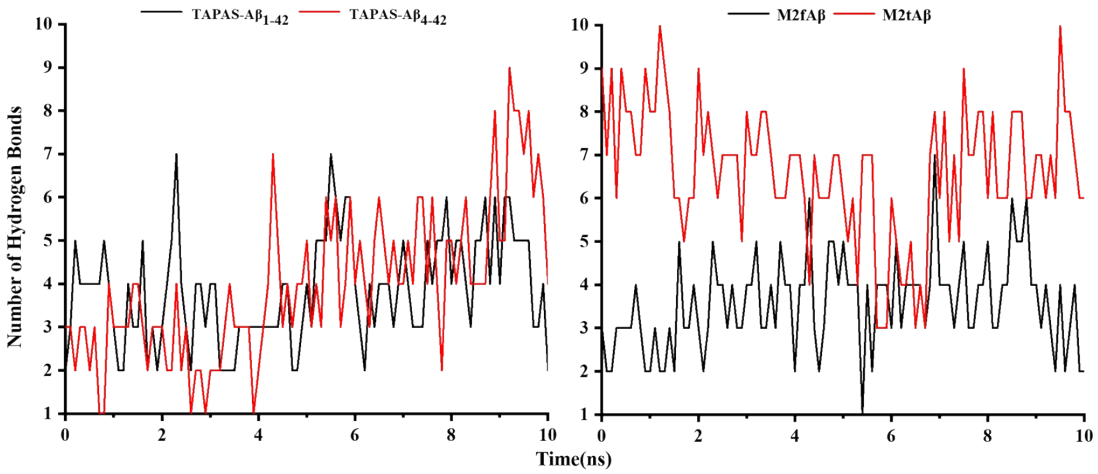

Figure S8 Number of hydrogen bonds formed between the vaccine and the receptor protein obtained from the last 10ns trajectory.

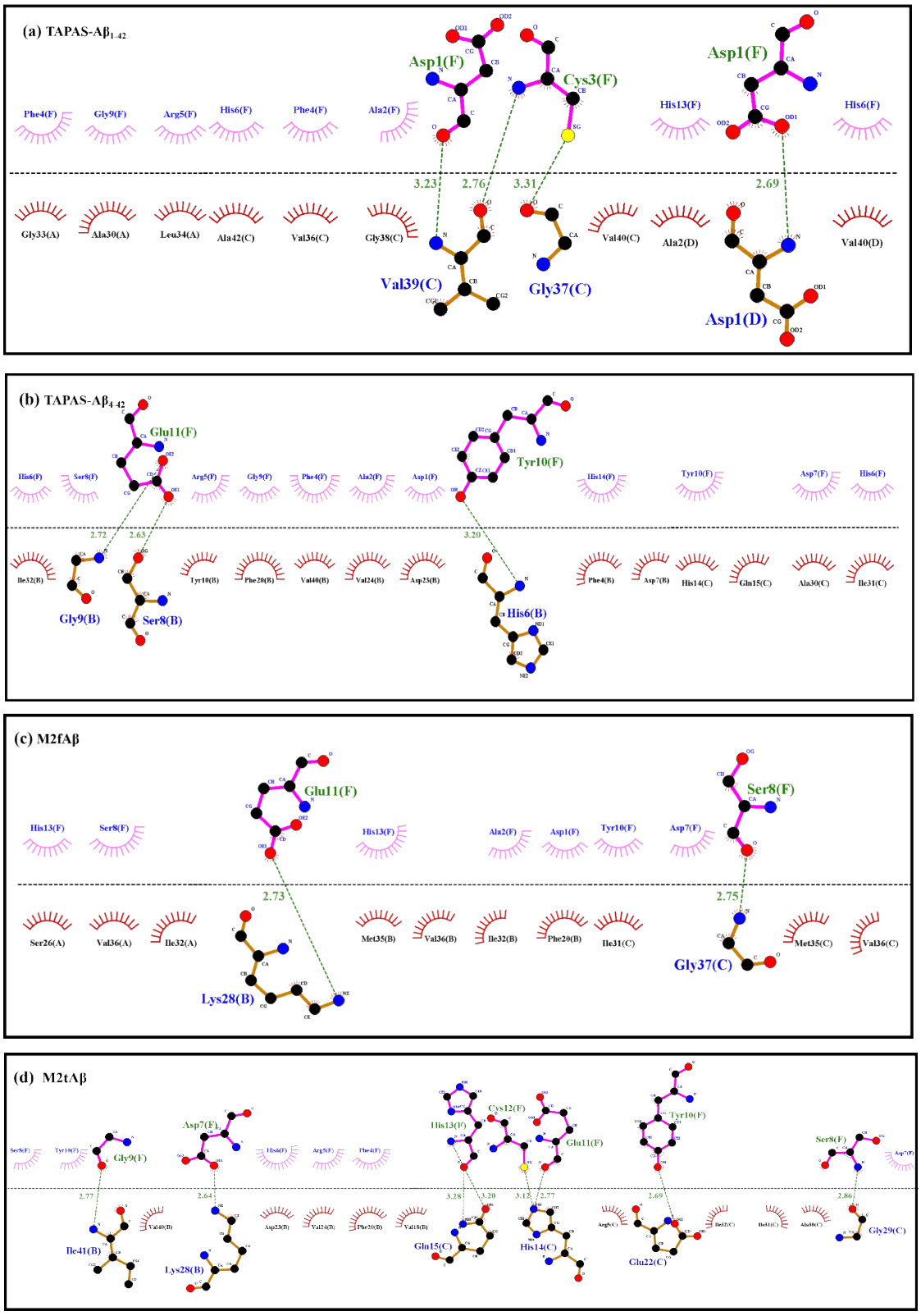

Figure S9 Interaction residues and corresponding distances on the vaccine-receptor contact interface. The symbols of A-E in parenthesis indicate the serial number of five chains in the Aβ pentameric receptors. F in the parenthesis stands for the vaccine. Hydrogen bond is represented by a dotted green line.

- Length: 15
- Number of predicted TMHs : 0
- Exp number of AAs in TMHs : 0
- Exp number, first 60 AAs : 0
- Total prob of N-in : 0.34348
- TMHMM2.0 inside 1 15

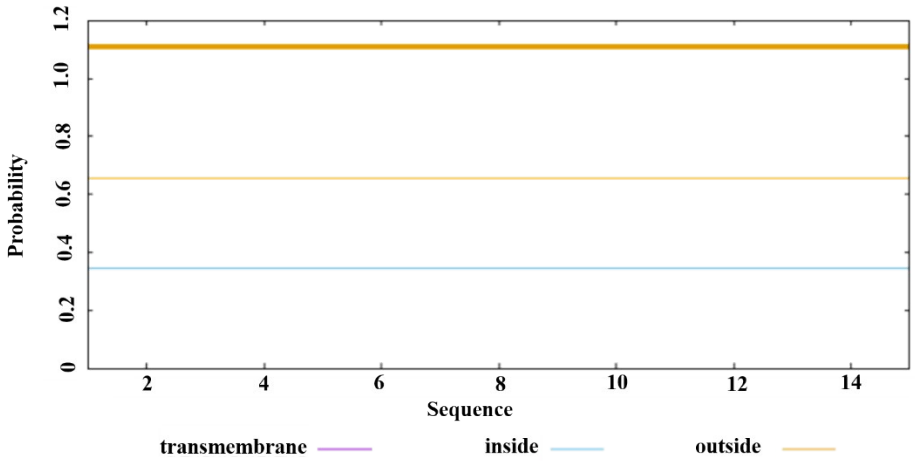

Figure S10 Transmembrane Helix of $A\beta_{1-15}$.

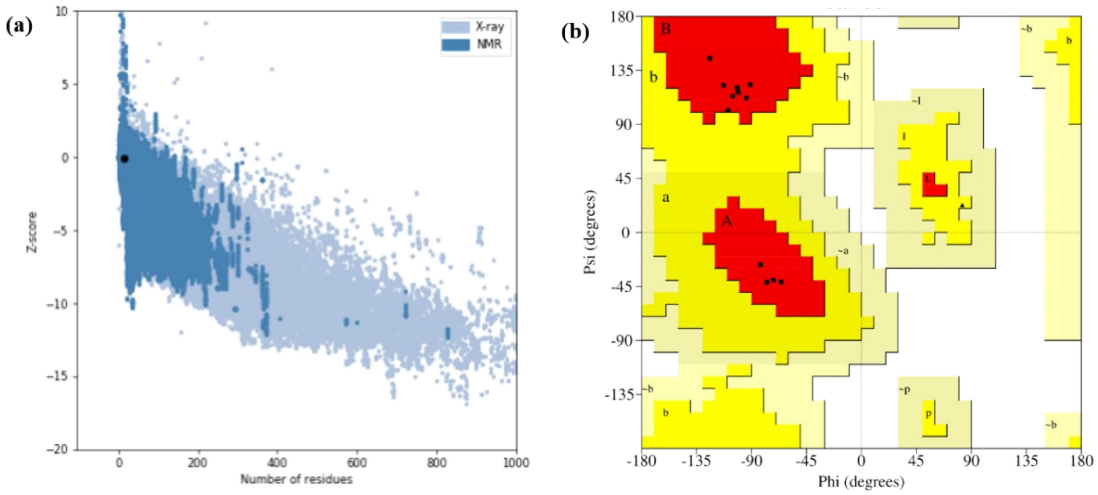

Figure S11 Structure validation of model1$A\beta_{1-15}$. (a) ProSA validation of predicted structure with Z score of -0.1, (b) Ramachandran plot analysis indicates 100% of residues are present in the most favored regions, 0% of residues are found in the allowed regions.

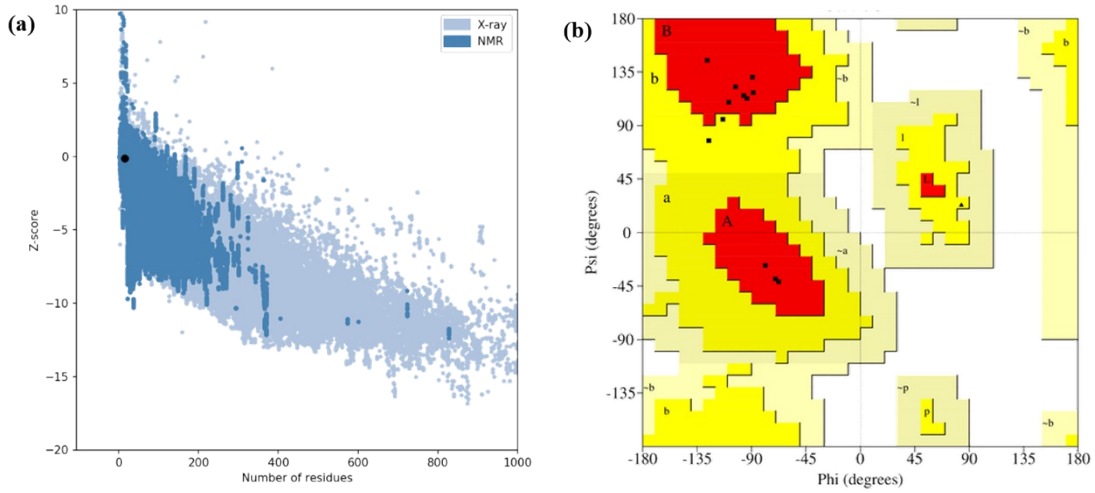

Figure S12 Structure validation of model2Aβ$_{1-15}$. (a) ProSA validation of predicted structure with Z score of -0.14, (b) Ramachandran plot analysis indicates 83.3% of residues are present in the most favored regions, 16.7% of residues are found in the allowed regions.

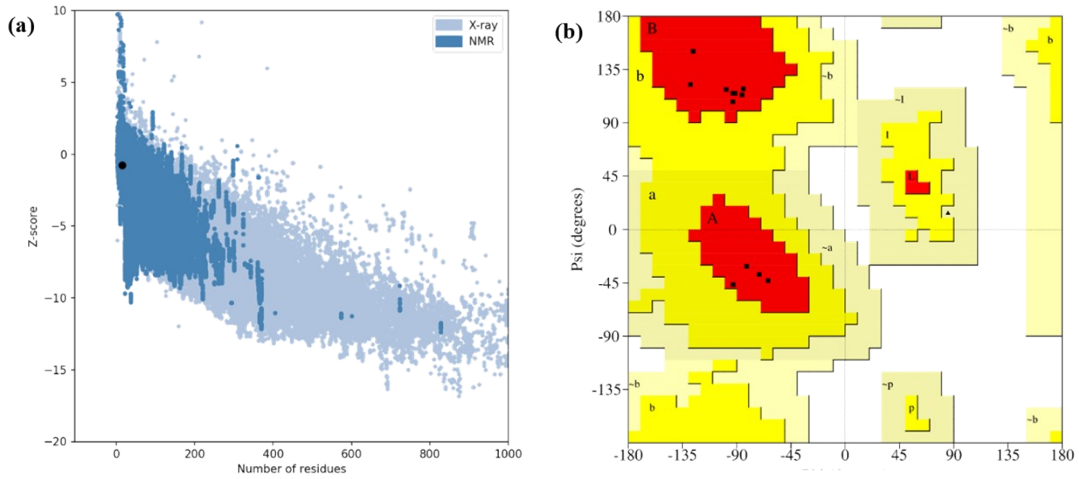

Figure S13 Structure validation of model3Aβ$_{1-15}$. (a) ProSA validation of predicted structure with Z score of -0.77, (b) Ramachandran plot analysis indicates 100% of residues are present in the most favored regions, 0% of residues are found in the allowed regions.

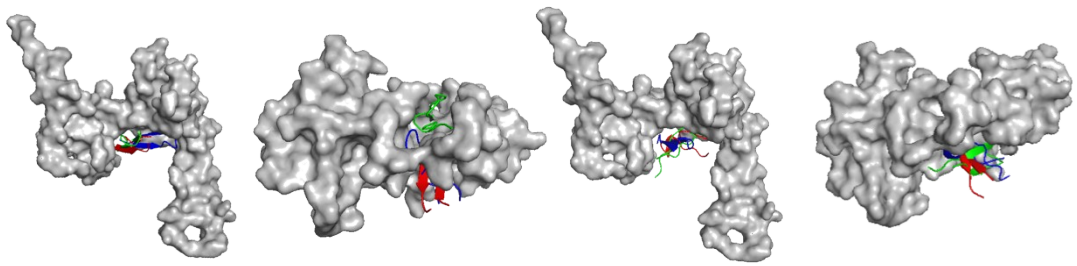

Figure S14 Top 3 positions of two Aβ pentamers (fAβ and tAβ) for vaccines (M′1 and M′3) docking, in which the three top poses of the vaccine are highlighted in red for Top 1, green for Top 2, and blue for Top 3, respectively.

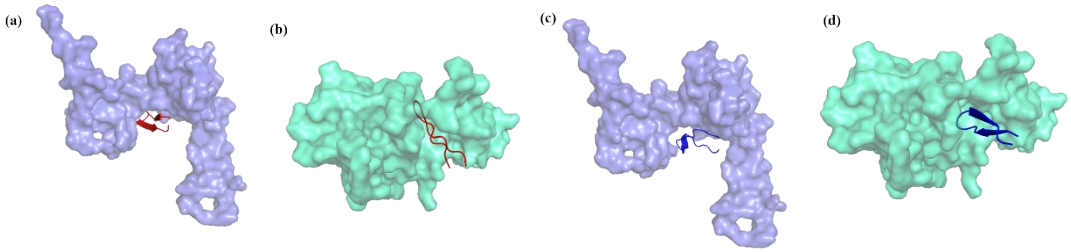

Figure S15 Top 1 docked complexes of modelj'A$\beta_{1\text{-}15}$ (j'=1 and 3) vaccines to Aβ receptors (fAβ and tAβ). (a) and (b) are complexes of model1A$\beta_{1\text{-}15}$ docked with receptors A$\beta_{1\text{-}42}$ and A$\beta_{4\text{-}42}$, respectively; (c) and (d) stand for the complexes of model3A$\beta_{1\text{-}15}$ docked with the two receptors, respectively. Model1A$\beta_{1\text{-}15}$, model3A$\beta_{1\text{-}15}$, A$\beta_{1\text{-}42}$ and A$\beta_{4\text{-}42}$ are represented in red, blue, slate and cyan, respectively.

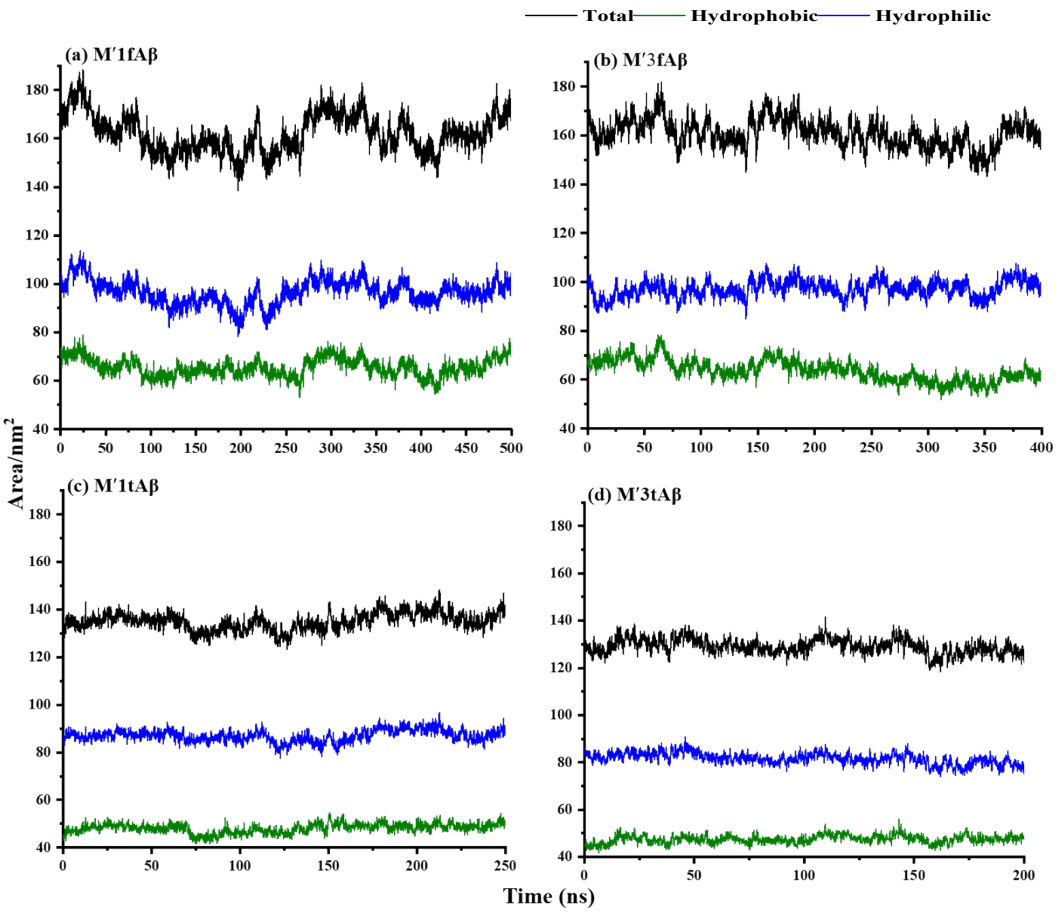

Figure S16 SASA results of the four complexes.

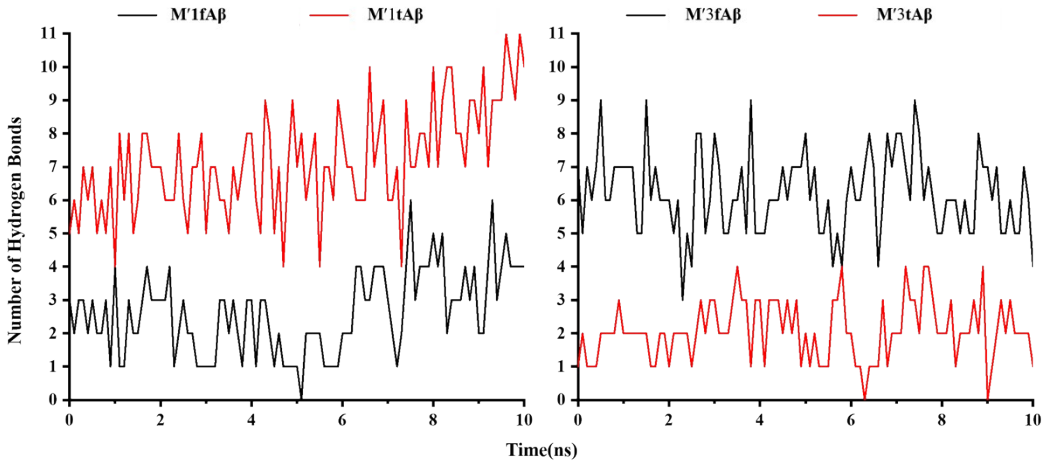

Figure S17 Number of hydrogen bonds formed between the vaccine and the receptor protein obtained from the last 10ns trajectory.

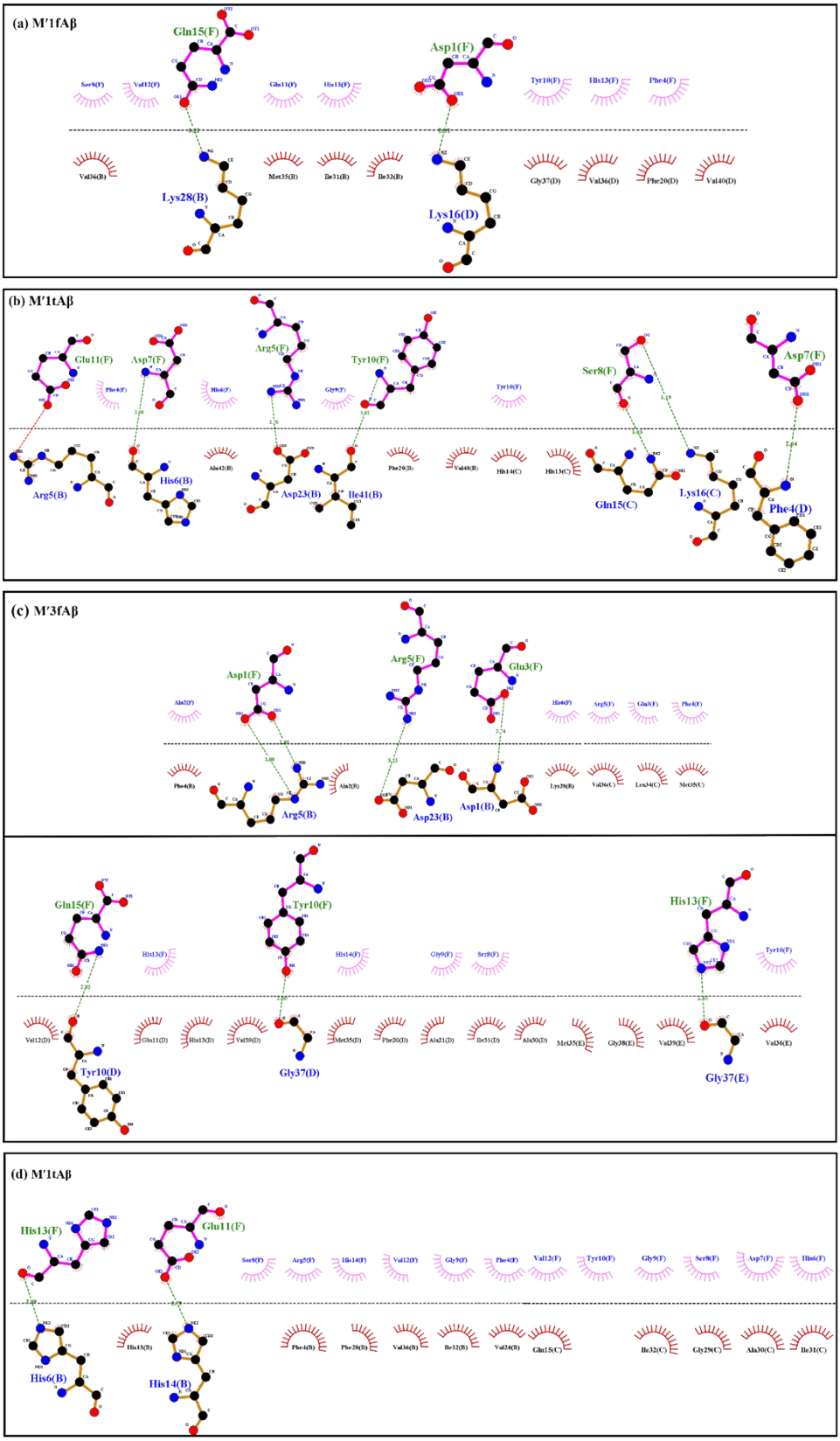

Figure S18 Interaction residues and corresponding distances on the vaccine-receptor contact interface. The symbols of A-E in parenthesis indicate the serial number of five chains in the Aβ pentameric receptors. F in the parenthesis stands for the vaccine. Hydrogen bond and salt bridges are represented by dotted lines in green and red, respectively.

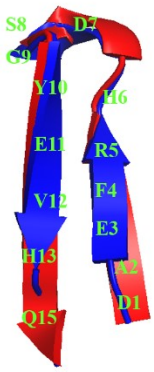

Figure S19 Fitted 3D conformations for model1A$\beta_{1\text{-}15}$ and model3A$\beta_{1\text{-}15}$, where model1A$\beta_{1\text{-}15}$ and model3A$\beta_{1\text{-}15}$ are represented in red and blue, respectively, and the labels of residues 1-15 are represented in green.

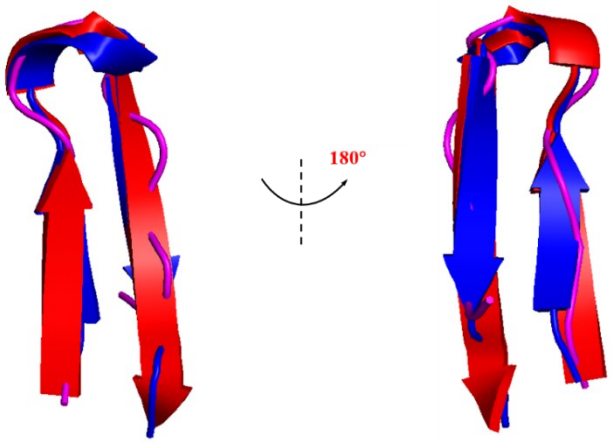

Figure S20 Fitted 3D conformation for modelj'A$\beta_{1\text{-}15}$ (j'=1 and 3) and TAPAS. TAPAS and modelj'A$\beta_{1\text{-}15}$ are represented in magenta, red and blue, respectively.

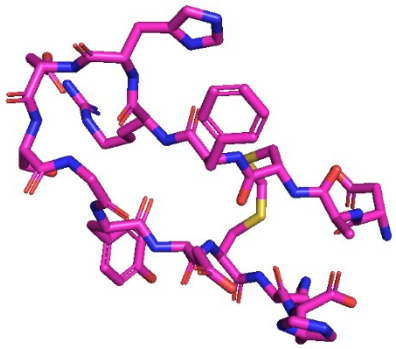

Figure S21 The 3D structure of the free TAPAS

Table S1 Immunogenicity, antigenicity, allergen, toxicity of $A\beta_{1-14}$

immunogenicity	antigenicity	allergen	toxic
0.14	VaxiJen 2.0 0.61 > 0.40	Non-allergenic	non-toxic

Table S2 Physicochemical properties of $A\beta_{1-14}$

Number of amino acids	Molecular weight	Theoretical pI	Estimated half-life	Aliphatic index	GRAVY
14	1676.76	5.73	1.1 hours (mammalian reticulocytes, in vitro). 3 min (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	7.14	-1.25

Table S3 Parameters after fitting model*i*$A\beta_{1-14}$ (*i*=1, 2 and 3) structure to the 3D structure of TAPAS

Model	RMSD(Å)	TM-score	Identity
1	1.36	0.78	92%
2	1.20	0.81	92%
3	2.93	0.53	8%

Table S4 Cluster scores of docked TAPAS-$A\beta_{1-42}$ complex

Cluster	Members	Representative	Weighted Score
1	213	Center	-710.10
		Lowest Energy	-817.90
2	184	Center	-734.90
		Lowest Energy	-795.40
3	180	Center	-793.20
		Lowest Energy	-793.20

Table S5 Cluster scores of docked TAPAS-$A\beta_{4-42}$ complex

Cluster	Members	Representative	Weighted Score
1	273	Center	-808.90
		Lowest Energy	-919.40
2	235	Center	-776.60
		Lowest Energy	-998.20
3	158	Center	-785.20
		Lowest Energy	-866.30

Table S6 Cluster scores of docked M1fAβ complex

Cluster	Members	Representative	Weighted Score
1	413	Center	-663.90
		Lowest Energy	-774.90
2	125	Center	-637.20
		Lowest Energy	-692.90
3	96	Center	-684.10
		Lowest Energy	-684.10

Table S7 Cluster scores of docked M1tAβ complex

Cluster	Members	Representative	Weighted Score
1	312	Center	-610.80
		Lowest Energy	-717.10
2	275	Center	-609.50
		Lowest Energy	-712.70
3	229	Center	-621.00
		Lowest Energy	-702.20

Table S8 Cluster scores of docked M2fAβ complex

Cluster	Members	Representative	Weighted Score
1	215	Center	-635.90
		Lowest Energy	-678.70
2	145	Center	-583.40
		Lowest Energy	-652.40
3	137	Center	-562.80
		Lowest Energy	-635.80

Table S9 Cluster scores of docked M2tAβ complex

Cluster	Members	Representative	Weighted Score
1	312	Center	-610.80
		Lowest Energy	-717.10
2	275	Center	-609.50
		Lowest Energy	-712.70
3	229	Center	-621.00
		Lowest Energy	-702.20

Table S10 Contribution of each residue in the $A\beta_{1-42}$ when binding to TAPAS

Residue name	Num.	probability
VAL	30	0.37
ALA	16	0.20
ILE	11	0.14
GLY	9	0.11
LEU	7	0.09
ASP	5	0.06
GLU	3	0.04

Table S11 Contribution of each residue in the TAPAS when binding to the $A\beta_{1-42}$

Residue name	Num.	probability
HIS	22	0.33
ASP	11	0.17
PHE	10	0.15
CYS	8	0.12
GLY	5	0.08
ALA	5	0.08
ARG	3	0.04
SER	2	0.03

Table S12 Contribution of each residue in the $A\beta_{4-42}$ when binding to TAPAS

Residue name	Num.	probability
PHE	23	0.20
ILE	19	0.16
VAL	16	0.14
HIS	14	0.12
ASP	9	0.08
GLN	8	0.07
LYS	7	0.06
ALA	6	0.05
TYR	4	0.03
GLY	4	0.03
SER	3	0.03
ARG	2	0.02
GLU	1	0.01

Table S13 Contribution of each residue in the TAPAS when binding to the $A\beta_{4-42}$

Residue name	Num.	probability
HIS	24	0.26
PHE	15	0.16
TYR	14	0.15
ASP	11	0.12
SER	9	0.10
GLU	6	0.06
GLY	5	0.05
ALA	4	0.04
CYS	3	0.03
ARG	2	0.02

Table S14 Contribution of each residue in the $A\beta_{1-42}$ when binding to model2$A\beta_{1-14}$

Residue name	Num.	probability
ILE	25	0.26
VAL	23	0.23
MET	11	0.11
LYS	10	0.10
PHE	9	0.09
GLY	8	0.08
SER	6	0.06
ASN	5	0.05
ALA	1	0.01

Table S15 Contribution of each residue in the model2Aβ_{1-14} when binding to the Aβ_{1-42}

Residue name	Num.	probability
ASP	16	0.21
HSE	15	0.19
TYR	15	0.19
SER	9	0.12
GLU	5	0.06
ALA	5	0.06
ARG	5	0.06
GLY	4	0.05
CYS	3	0.04
PHE	1	0.01

Table S16 Contribution of each residue in the Aβ_{4-42} when binding to model2Aβ_{1-14}

Residue name	Num.	probability
VAL	20	0.19
ILE	18	0.17
PHE	16	0.15
HIS	16	0.15
LYS	7	0.07
GLN	6	0.06
GLY	5	0.05
GLU	4	0.04
ALA	4	0.04
ASP	3	0.03
ARG	3	0.03
LEU	1	0.01
SER	1	0.01

Table S17 Contribution of each residue in the model2Aβ_{1-14} when binding to the Aβ_{4-42}

Residue name	Num.	probability
HIS	17	0.22
TYR	15	0.19
PHE	12	0.15
ASP	8	0.10
SER	8	0.10
GLY	5	0.06
GLU	5	0.06
CYS	4	0.05
ARG	4	0.05

Table S18 Immunogenicity, antigenicity, allergen, toxicity of Aβ_{1-15}

immunogenicity	**antigenicity**	**allergen**	**toxic**
0.304	VaxiJen 2.0 0.56 > 0.4	non-allergen	Non-Toxin

Table S19 Physicochemical properties of Aβ_{1-15}

Number of amino acids	**Molecular weight**	**Theoretical pI**	**Estimated half-life**	**Aliphatic index**	**GRAVY**
15	1826.86	5.21	1.1 hours (mammalian reticulocytes, in vitro). 3 min (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	26.00	-1.69

Table S20 Cluster scores of docked vaccine M′1fAβ complex

Cluster	Members	Representative	Weighted Score
1	588	Center	-636.60
		Lowest Energy	-749.00
2	156	Center	-647.60
		Lowest Energy	-719.70
3	50	Center	-660.90
		Lowest Energy	-636.60

Table S21 Cluster scores of docked vaccine M′1tAβ complex

Cluster	Members	Representative	Weighted Score
1	267	Center	-711.00
		Lowest Energy	-780.70
2	189	Center	-701.00
		Lowest Energy	-813.90
3	185	Center	-703.30
		Lowest Energy	-853.90

Table S22 Cluster scores of docked vaccine M′3fAβ complex

Cluster	Members	Representative	Weighted Score
1	379	Center	-591.80
		Lowest Energy	-690.30
2	318	Center	-590.50
		Lowest Energy	-748.60
3	109	Center	-586.40
		Lowest Energy	-661.40

Table S23 Cluster scores of docked vaccine M′3tAβ complex

Cluster	Members	Representative	Weighted Score
1	455	Center	-641.00
		Lowest Energy	-793.30
2	265	Center	-687.90
		Lowest Energy	-779.20
3	175	Center	-669.00
		Lowest Energy	-704.80

Table S24 Contribution of each residue in the $A\beta_{1-42}$ when binding to model1$A\beta_{1-15}$

Residue name	Num.	probability
VAL	25	0.29
LYS	17	0.20
ILE	10	0.12
MET	9	0.11
PHE	9	0.11
GLY	8	0.09
SER	3	0.03
HIS	3	0.03
ALA	1	0.01

Table S25 Contribution of each residue in the model1Aβ_{1-15} when binding to the Aβ_{1-42}

Residue name	Num.	probability
HIS	15	0.20
TYR	13	0.18
VAL	8	0.11
ASP	7	0.10
GLN	7	0.10
SER	6	0.08
PHE	6	0.08
GLY	4	0.05
ALA	3	0.04
GLU	2	0.03
ARG	2	0.03

Table S26 Contribution of each residue in the Aβ_{4-42} when binding to model1Aβ_{1-15}

Residue name	Num.	probability
HIS	21	0.19
PHE	20	0.18
ARG	16	0.15
ILE	12	0.11
VAL	11	0.10
ALA	6	0.06
GLN	6	0.06
ASP	5	0.05
SER	5	0.05
LYS	5	0.05
GLY	1	0.01

Table S27 Contribution of each residue in the model1Aβ_{1-15} when binding to the Aβ_{4-42}

Residue name	Num.	probability
HIS	15	0.20
TYR	13	0.18
VAL	8	0.11
ASP	7	0.10
GLN	7	0.10
SER	6	0.08
PHE	6	0.08
GLY	4	0.05
ALA	3	0.04
GLU	2	0.03
ARG	2	0.03

Table S28 Parameters of fitted 3D structures for model1Aβ_{1-15} and model3Aβ_{1-15}

RMSD	**TM-score**	**Identity**
0.71	0.87	100%

Table S29 Parameters of fitted 3D structures for modelj'Aβ_{1-15} (j'=1 and 3) and TAPAS

Model	**RMSD**	**TM-score**	**Identity**
model1Aβ_{1-15}	1.11	0.65	85%
model3Aβ_{1-15}	1.30	0.59	85%

References

1. M. U. Rahman, H. Liu, A. Wadood and H. F. Chen, *Mol Biosyst*, 2016, **12**, 3280-3293.
2. H. Yu, Y. Fang, X. Lu, Y. Liu and H. Zhang, *Chem. Biol. Drug. Des.*, 2014, **83**, 89-105.
3. T. D. Martin, E. H. Hill, D. G. Whitten, E. Y. Chi and D. G. Evans, *Langmuir*, 2016, **32**, 12542-12551.
4. E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang and T. Hou, *Chem. Rev*, 2019, **119**, 9478-9508.
5. Y.-j. Sheng, Y.-w. Yin, Y.-q. Ma and H.-m. Ding, *Journal of Chemical Information and Modeling*, 2021, **61**, 2454-2462.
6. K. Huang, S. Luo, Y. Cong, S. Zhong, J. Z. H. Zhang and L. Duan, *Nanoscale*, 2020, **12**, 10737-10750.
7. L. Duan, X. Liu and J. Z. H. Zhang, *Journal of the American Chemical Society*, 2016, **138**, 5722-5728.
8. M. R. Lee, Y. Duan and P. A. Kollman, *Proteins*, 2000, **39**, 309-316.
9. R. Kumari, R. Kumar and A. Lynn, *J. Chem. Inf. Model*, 2014, **54**, 1951-1962.