**Supporting Information**

# Probing the local conformational flexibility in receptor recognition: Mechanistic insight from an atomic-scale investigation

Fei Ding[ab] and Wei Peng*[cd]

[a] *School of Environmental Science and Engineering, Chang'an University, Xi'an 710064, China*

[b] *Key Laboratory of Subsurface Hydrology and Ecological Effect in Arid Region of Ministry of Education, Chang'an University, No. 126 Yanta Road, Yanta District, Xi'an 710064, China*

[c] *College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China*

[d] *Department of Chemistry, China Agricultural University, Beijing 100193, China*

*Corresponding Author
Phone/fax: +86-29-87092367
E-mail: crystalw.peng@outlook.com

**Supplementary Introduction:**

**A brief introduction to the ligand bixin**

Of the crudely occurring carotenoids, bixin (structure shown in Fig. S1) extracts from annatto seeds of the tree *Bixa orellana* L. (family: *Bixaceae*), which ranks first in economic status in Latin America.[1] Today the general idea is that over 80% of the carotenoid in the annatto seed epidermis consists of bixin, which has mainly imparted yellow to red color to a diversity of dairy products, e.g. butter mixes, cheese, dressings, margarine and sausages.[2] The largest consumer markets of bixin are found to be located in the developed world, i.e. Europe, Japan and United States; for example, bixin has been observed to be the most routinely consumed natural color carotenoid in the Netherlands and the United Kingdom, where the primary use is in the cheese industry.[3]

**Supplementary Protocol:**

**Molecular dynamics simulation**

Molecular dynamics (MD) simulation of AGP-bixin complex was conducted using Gromacs program, version 2019, with the Gromos96 ffG43a1 force field.[4-7] Simulation processes were operated under physiological conditions (pH＝7.4), and the amino acid residues possessed acidity and basicity were adjusted to the protonation states at neutrality condition. Initial conformations of AGP and bixin were, respectively, taken from the original X-ray diffraction crystal structure that was solved at 2.1 Å resolution (entry codes 3APU) and the optimal structure originated from molecular docking.[8] The topology of protein were produced by Gromacs package directly, whereas bixin by PRODRG2.5 Server.[9,10] The simulation system was solvated with a periodic cubic box (the volume is 4.699×4.538×4.707 $nm^3$) filled with TIP3P water molecules and an approximate number (5) of sodium counterion to neutralize the charge.[11] Totally, there are 12,842 crystallographic solvent molecules, and the shortest distance between the adduct and the edge of the box is set to 10 Å. Simulations were performed utilizing the isothermal-isobaric (NPT) ensemble with an isotropic pressure of 1 bar,[12] and the temperature of the ligand, protein and solvent (water and counterion) was separately coupled to an external bath held at 300 K, using the Berendsen thermostat with 0.2 ps relaxation time.[13] The LINCS algorithm was employed to constrain bond lengths,[14] and the long-range electrostatic interactions beyond 10 Å were modeled hiring the Particle Mesh Ewald

(PME) method with a grid point density of 0.1 nm and an interpolation order of 4.[15,16] A cutoff of 14 Å was rented for van der Waals' interactions. The MD integration time step was 2.0 fs and covalent bonds were not constrained, and the system configurations were saved every 2.0 ps. To decrease the atomic collisions with each other, both gradient descent and conjugate gradient algorithm were applied to optimize the whole system.[17,18] First the solvated starting structure was preceded by a 1,000-step gradient descent and then by conjugate gradient energy minimization. Subsequently, 100 ps equilibration with position restraints executes to remove possible unfavorable interactions between solute and solvent, and after thorough equilibration, MD simulations were carried out for 100 ns. Moreover, the top two ligand structures in the reaction system were selected to execute the parallel MD simulations, in order to confirm the rationality of the optimal energy conformation for MD simulation. The pure protein was also selected to run a time period (50 ns) MD simulation so as to compare with the first-rank molecular docking conjugate. The results of MD simulations were eventually manifested by Visual Molecular Dynamics 1.9.4,[19,20] and the program Discovery Studio Visualizer 4.0 (BIOVIA, San Diego, CA) was exploited to reveal the profiles of the MD simulations.

**Supplementary Results and Discussion:**

**Elucidation of the overall energy of the receptors-bixin adducts**

The Lennard-Jones potential is a mathematically fundamental model that tackles the interaction between a pair of neutral atoms or molecules,[21] and such method is utilized widely in molecular modeling owing to its computational convenience and comparatively nice approximation. Whereas the Coulomb potential is a valid pair potential that delineates the interaction between two point charges, and it operates along the line relating the two charges.[22] In computational chemistry, the two parameters have ordinarily been hired to scrutinize the electrostatic interaction and the critical interaction energy in the biomacromolecule-ligand reaction, namely the electrostatic interaction is estimated using the Coulomb's law and the dispersion and repulsion forces utilizing the Lennard-Jones potential. Thus the Lennard-Jones potential and Coulomb potential obtained by molecular modeling are collected in Table S1.

Clearly, the significant energy contributions of the proteins-bixin reactions derives from the Lennard-Jones term, and the contributions of the Coulomb term are relatively small, confirming that electrostatic interaction energy did not play a major role in the total energy of the biopolymers-bixin conjugations. This event correlates closely with the structural property of bixin, or, more exactly, bixin is constructed mainly by hydrophobic unit (polyene chain), and in consequence both conjugated effects and hydrophobic effects are observed to be far more important to the

biointeractions of the proteins with bixin. Meanwhile, it should be noted evidently from Table S1 that the bioreaction intensity of the AGP-bixin complex is slightly lower than the HSA-bixin bioconjugate, but the difference is not really quite remarkable. Undoubtedly, this coheres well with the experimental results of molecular docking and MD simulations.

Table S1 here about

**Supplementary Table:**

Table S1
Interaction energies (kJ $mol^{-1}$) of the proteins-bixin bioreaction processes

| Biosystems | Lennard-Jones energy | Coulomb energy | Total energy |
|---|---|---|---|
| HSA-bixin | −27.36 | −18.68 | −46.04 |
| AGP-bixin | −25.92 | −16.19 | −42.11 |

**Supplementary Figures:**

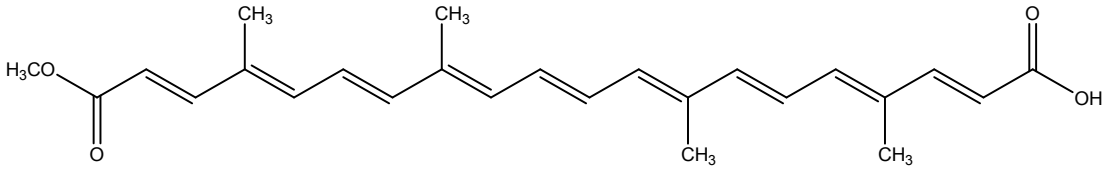


Fig. S1. Molecular structure of bixin

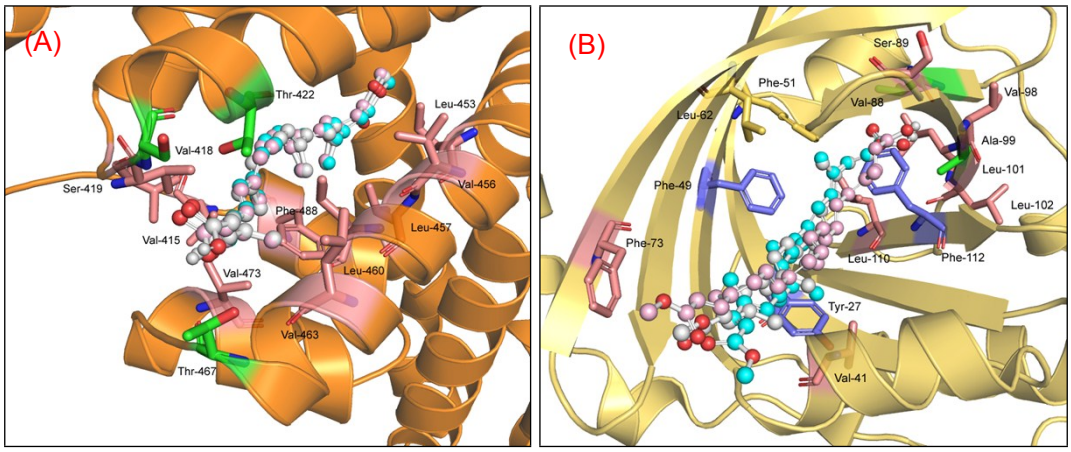


Fig. S2. Superimposition of the top three bixin conformations among the docking conformations resulting from the HSA-bixin complexes (panel (A)) and the AGP-bixin conjugates (panel (B)). The first conformation of bixin displayed in cyan ball-and-stick model, the second conformation of bixin expressed in white ball-and-stick model, and the third conformation of bixin indicated in pink ball-and-stick model, respectively.

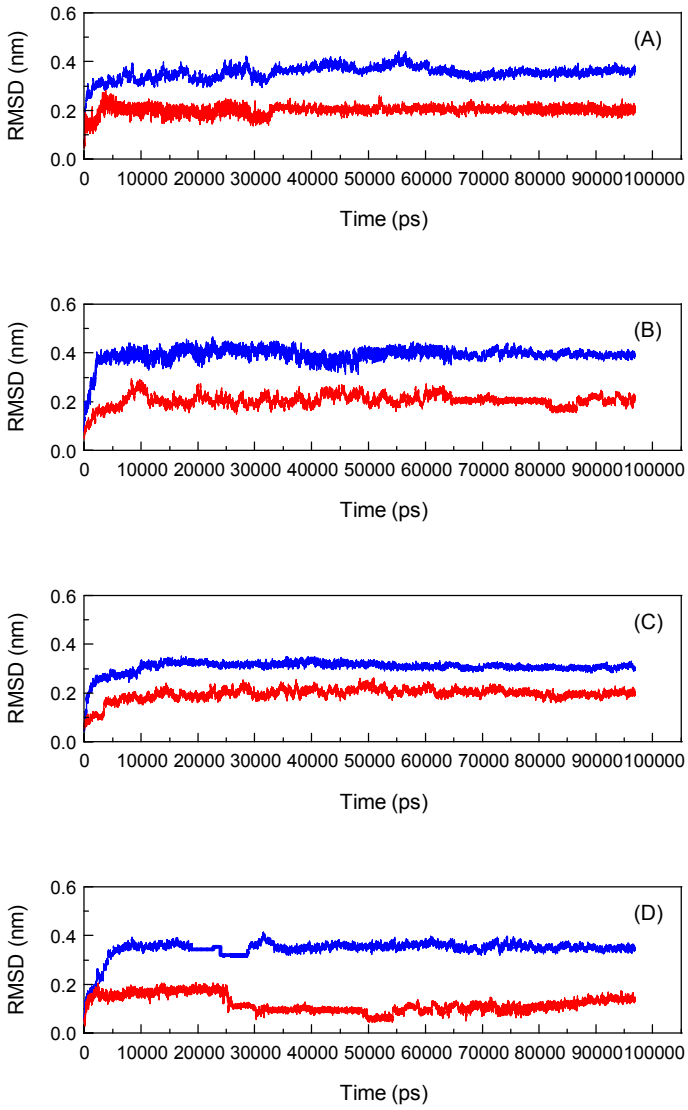


Fig. S3. Calculated Root-Mean-Square Deviation (RMSD) for the bixin and the backbone $C_\alpha$ atoms of HSA (panel (A) and panel (B)) and AGP (panel (C) and panel (D)) from MD simulation at temperature of 300 K with respect to their docking results as a function of the simulation time. The blue and red trajectories exhibit RMSD data for the backbone $C_\alpha$ atoms of proteins and the bixin, respectively.

**Supplementary References:**


1. C. R. Cardarelli, M. de T. Benassi and A. Z. Mercadante, *LWT - Food Sci. Technol.*, 2008, **41**, 1689-1693.
2. A. Z. Mercadante, in *Chemistry and Physiology of Selected Food Colorants*, ed. J. M. Ames and T. Hofmann, American Chemical Society, Washington, D.C., 2001, vol. 775, pp 92-101.
3. G. Giuliano, C. Rosati and P. M. Bramley, *Trends Biotechnol.*, 2003, **21**, 513-516.
4. D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701-1718.
5. D. van der Spoel, P. J. van Maaren and C. Caleman, *Bioinformatics*, 2012, **28**, 752-753.
6. M. M. Ghahremanpour, S. S. Arab, S. B. Aghazadeh, J. Zhang and D. van der Spoel, *Bioinformatics*, 2014, **30**, 439-441.
7. W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger and W. F. van Gunsteren, *J. Phys. Chem. A*, 1999, **103**, 3596-3607.
8. K. Nishi, T. Ono, T. Nakamura, N. Fukunaga, M. Izumi, H. Watanabe, A. Suenaga, T. Maruyama, Y. Yamagata, S. Curry and M. Otagiri, *J. Biol. Chem.*, 2011, **286**, 14427-14434.
9. D. M. F. van Aalten, R. Bywater, J. B. C. Findlay, M. Hendlich, R. W. W. Hooft and G. Vriend, *J. Comput.-Aided Mol. Des.*, 1996, **10**, 255-262.

10. A. W. Schüttelkopf and D. M. F. van Aalten, *Acta Crystallogr. D: Biol. Crystallogr.*, 2004, **D60**, 1355-1363.

11. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.

12. H. A. Stern, *J. Comput. Chem.*, 2004, **25**, 749-761.

13. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.

14. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463-1472.

15. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.

16. T. Darden, L. Perera, L. P. Li and L. Pedersen, *Structure*, 1999, **7**, R55-R60.

17. M. R. Hestenes and E. Stiefel, *J. Res. Natl. Bur. Stand.*, 1952, **49**, 409-436.

18. J. A. Snyman, in *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, Springer Science＋Business Media, New York, NY, 2005, vol. 97, pp 1-231.

19. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33-38.

20. J. Hsin, A. Arkhipov, Y. Yin, J. E. Stone and K. Schulten, *Curr. Protoc. Bioinformatics*, 2008, **24**, 5.7.1-5.7.48.

21. J. E. Jones, *Proc. R. Soc. A: Math., Phys. Eng. Sci.*, 1924, **106**, 441-462.

22. T. M. C. Faro, G. P. Thim and M. S. Skaf, *J. Chem. Phys.*, 2010, **132**, 114509.