Two Rules on the Protein-Ligand Interaction

Xiaodong Pang[1, 2†], Linxiang Zhou[1†], Lily Zhang[3], Lina Xu[3] & Xinyi Zhang[1, 2*]

So far, we still lack a clear molecular mechanism to explain the protein-ligand interaction on the basis of electronic structure of a protein. By combining the calculation of the full electronic structure of a protein along with its hydrophobic pocket and the perturbation theory, we found out two rules on the protein-ligand interaction. One rule is the interaction only occurs between the lowest unoccupied molecular orbitals (LUMOs) of a protein and the highest occupied molecular orbital (HOMO) of its ligand, not between the HOMOs of a protein and the LUMO of its ligand. The other rule is only those residues or atoms located both on the LUMOs of a protein and in a surface pocket of a protein are activity residues or activity atoms of the protein and the corresponding pocket is the ligand binding site. These two rules are derived from the characteristics of energy levels of a protein and might be an important criterion of drug design.

[1]Surface Physics Laboratory (National Key Laboratory), Department of Physics, Fudan University, Shanghai 200433, China. [2]Synchrotron Radiation Research Center, Fudan University, Shanghai 200433, China. [3]Department of Electrical and Computer Engineering，Rice University，Houston, TX77005, USA.

†These authors contributed equally to this work.

*To whom all correspondences should be addressed. E-mail: xy-zhang@fudan.edu.cn

Understanding ruling principles of interactions between a target protein and a ligand is of paramount importance in drug discovery efforts. So far, in finding a real ligand for a given target protein, we are limited to experimental screening from a large number of small molecules, or through free energy calculation of assessing a ligand. However, we still lack a clear molecular mechanism to explain the protein-ligand interaction on the basis of electronic structure of a protein. Here we report two rules on the protein-ligand interaction using the full electronic structure calculation of a protein along with the perturbation theory of the interaction between two wave functions and the pocket calculation as a new means of drug design and exploring enzyme reaction.

The full electronic structure calculation of a protein has been a challenge as the system of a protein is huge and complicated. We use the overlapping-dimer approximation (ODA)[1-3] and the extended negative factor counting (ENFC)[4] methods to calculate the electronic structure of a protein. These two methods were first developed by J. Ladik and Y. J. Ye *et al*[1,4]. We further clarified and revised some parts to facilitate calculations of proteins with over 600 residues.

The ODA uses the divide and conquer (D&C) algorithm, which works by recursively breaking down a complex problem into many subproblems of the same (or related) type of the same scale. These subproblems are independent of each other, and their characters are the same as the original problem. We can use the solutions of these small subproblems to get the solution of the original problem. First of all, the whole chain of a protein is divided into many dimers, each containing two amino acid

residues, and then we can perform any quantum calculation methods for these dimers; finally, construct the full Hamiltonian matrix using the solution of each dimer and use the ENFC math method to get the energy eigenvalues and the orbital coefficients of the wave function of a protein. Additionally, the surrounding environment of each dimer and the water environment of a protein are applied in the process.

We take the protein-ligand complex Cyclophilin A (CypA) and its ligand Cyclosporin A (CsA) and FKBP12 of the protein family FKBPs and its immunosuppressant FK506 as examples, and the results obtained are in good agreement with experiments.

The system of a protein molecule is so large that the energy levels are extremely compact. Therefore, for a protein system, we need to extend the definition of the frontier orbitals to be a series of HOMOs and LUMOs, not just limited to be a single HOMO or LUMO. Ordinarily, the first 10 to 20 occupied molecular orbitals and the unoccupied ones are considered to be HOMOs and LUMOs. For example, the protein CypA has a total of 4807 occupied molecular orbitals. We considered the first 10 highest occupied molecular orbitals to be its HOMOs and the first 10 lowest unoccupied molecular orbitals as its LUMOs. The 10 HOMOs comprise only 0.2% among total 4807 occupied molecular orbitals.

RESULTS

First rule of the protein-ligand interaction

Figure 1 and **Tables 1** and **2** express the partial results of the frontier orbitals and

orbital coefficients for CypA/CsA and FKBP12/FK506 separately (see **Supplementary Tables 1** and **2** for more details about the frontier orbitals).

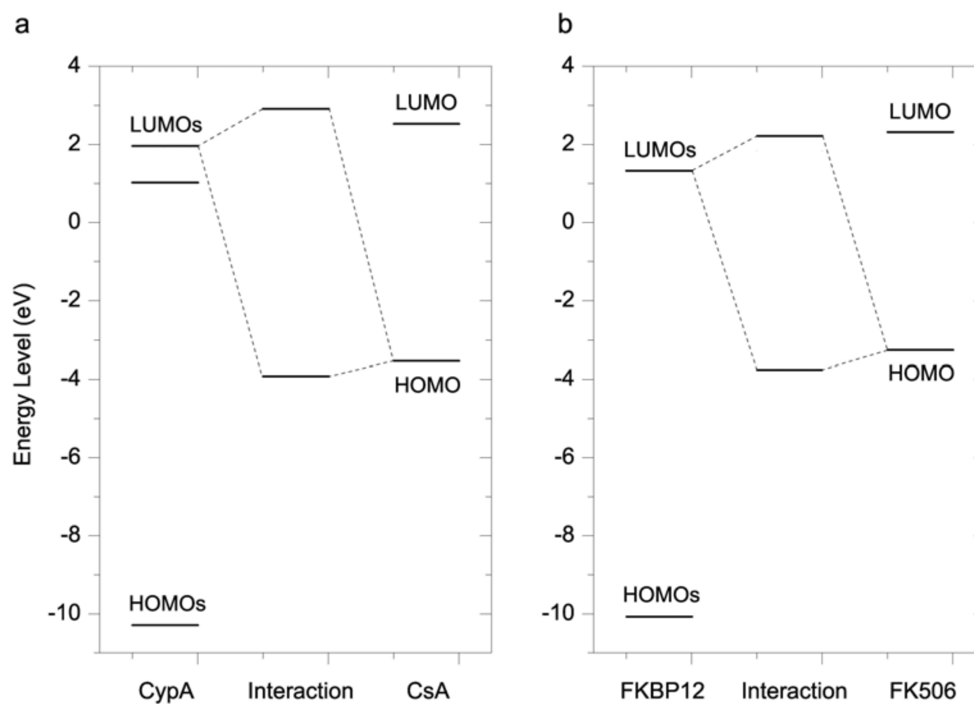

Figure 1 Schematic diagrams of the energy level. (**a**) Protein CypA interacts with ligand CsA at ground state. (**b**) Protein FKBP12 interacts with ligand FK506 at 300 K. The characteristics of energy level structure of a protein are that the LUMOs of a protein are always about 1.0 eV above the zero-energy-level, but its HOMOs are always far away below the zero-energy-level (about -10.0 eV). It is the very characteristics that result in the first rule.

Table 1 The first 10 LUMOs of CypA and FKBP12. The definition of occupancy at nth residue is $a(n) = \sum_{j=1}^{m_n} C_j^2(n) / \sum_{n=1}^{n=N} \sum_{j=1}^{m_n} C_j^2(n)$, where C_j are the orbital coefficients. Those residues that have $a(n) < 0.05$ are neglected. The wave function localizes on a single residue, since the occupancy of each orbital is larger than 99%. For CypA, residue PHE60 and PHE113 are activity residues. For FKBP12, residue TRY26, PHE36, PHE46, PHE59, PHE82 and PHE99 are activity residues.

	Molecular orbital	Energy band (eV)	Position of wave function (residues)	Occupancy
CypA	4817	1.959	PHE36	0.996
	4816	1.956	***PHE60***	0.996
	4815	1.954	***PHE113***	0.995
	4814	1.930	PHE145	0.998
	4813	1.880	PHE53	0.998
	4812	1.866	PHE67	0.996
	4811	1.850	PHE53	0.997
	4810	1.849	TYR79	0.998
	4809	1.690	TYR48	1.000
	4808	1.020	TRP121	0.996
FKBP12	3169	2.261	PHE48	0.993
	3168	2.141	**TYR26**	0.998
	3167	2.109	**PHE36**	0.995
	3166	2.101	**PHE82**	0.998
	3165	1.988	PHE15	0.999
	3164	1.932	**PHE46**	0.999
	3163	1.811	PHE48	0.998
	3162	1.754	**PHE99**	0.997
	3161	1.620	**PHE36**	0.992
	3160	1.328	**TRP59**	0.999

Table 2 The HOMO and LUMO of the ligand CsA and FK506.

	Molecular orbital	Energy band (eV)	Position of wave function (residues)
CsA	247	2.520	LEU6,ALA7,DLA8,LEU9
	246	-3.519	MVA11
FK506	842	2.314	C8,C9,C10,C14,C19,C20,C22, C27,C28,N7,O3,O4,O5,O9
	841	-3.255	C2,C8,C9,C10,C11,N7,O3,O4,O6

We define the energy difference between a protein and its ligand as

$$\Delta E_1 \equiv E_{proteinLUMO} - E_{ligandHOMO} \ , \ \Delta E_2 \equiv E_{ligandLUMO} - E_{proteinHOMO} \ ,$$

where, $E_{proteinLUMO}$ is the average value of the activity residues and $E_{proteinHOMO}$ is the highest value of HOMOs, $E_{ligandLUMO}$ and $E_{ligandHOMO}$ are values of the LUMO and the HOMO of ligand, respectively.

The calculation result demonstrates as follows:

for CypA/CsA, $\Delta E_1 = 1.96 - (-3.52) = 5.48eV < \Delta E_2 = 2.52 - (-10.29) = 12.80eV$,

for FKBP12/FK506, $\Delta E_1 = 1.33 - (-3.26) = 4.59eV < \Delta E_2 = 2.31 - (-10.07) = 12.38eV$.

Especially, we have to point out that the results of CypA/CsA and FKBP12/FK506 are examples only. We calculated other protein/ligand systems and obtained the same resulting relationship $\Delta E_1 < \Delta E_2$ as the above two systems without exception[5,6]. Therefore, the relationship $\Delta E_1 < \Delta E_2$ might be a generally correct in protein-ligand systems.

According to the frontier orbital theory, the perturbation energy of interaction between two molecules mainly comes from the interaction between HOMO and LUMO. If the energy of a protein is different from the energy of its ligand, the effect of the interaction produces energy split. As a result, the smaller the energy difference

between two orbitals, the stronger the interaction. (Whether the interaction between two molecules can form a bond, it is related to the symmetry of the molecular orbital.) We can, therefore, deduce the first rule on the protein-ligand interaction:

The protein-ligand interaction only occurs between the lowest unoccupied molecular orbitals (LUMOs) of a protein and the highest occupied molecular orbital (HOMO) of its ligand, not between the HOMOs of a protein and the LUMO of its ligand.

Actually, it is the very characteristics of energy level structure of a protein that result in the first rule. The characteristics are that the LUMOs of a protein are always about 1.0 eV above the zero-energy-level, but its HOMOs are always far away below the zero-energy-level (about -10.0 eV). It is due to this kind of structure that the LUMOs of a protein are always near the HOMO of its ligand and the HOMOs of a protein are always far away from the LUMO of its ligand. That is why always there is the relationship $\Delta E_1 < \Delta E_2$ in the protein-ligand interaction. The first rule may be regarded as one of criteria of identifying a ligand. An interesting aside is that the wave function of a protein usually localizes on one or few residues (see the occupancy item on **Table 1**).

Second rule of the protein-ligand interaction

According to the first rule, as a necessary condition, the activity residues or activity atoms of a protein must be located on the LUMOs of the protein, but conversely, not every residue on the LUMOs is certainly the activity residue. When a ligand interacts with its target protein, it must enter into a pocket on the surface of the protein. For the

same reason, only the pocket which has its forming atoms located on the LUMOs is the ligand binding site. Therefore, on the basis of the full electronic structure of a protein along with the pocket calculation, we can derive the second rule that **only those residues or atoms located both on the LUMOs of a protein and in a surface pocket of a protein are activity residues or activity atoms of the protein and the corresponding pocket is the ligand binding site.** This is the necessary and sufficient condition for being the activity residues or activity atoms, as well as the ligand binding site.

Now we can use this rule to check the activity atoms of a protein. We selected CASTp method[7] to calculate the pockets. The CASTp is a geometry-based method of pocket detection, based on the alpha shape and discrete flow theory, and a related suite of programs. It can provide full descriptions of pockets on protein surface, including the atoms, residues, and volume of each pocket. For example, the CypA has a total of 19 pockets and the largest pocket is number 19. The FKBP12 has a total of 13 pockets. **Table 3** and **Supplementary Table 3** show the pocket 19 of CypA and pocket 13 of FKBP12, respectively. **Tables 4** and **5** and **Supplementary Table 4** describe activity atoms and their orbital coefficients of protein CypA, ligand CsA and FK506, and protein FKBP12, respectively.

Table 3 The forming atoms in the pocket 19 of CypA. Only the residues PHE60 and PHE113 are located on the LUMOs of CypA, and all other residues on the LUMOs of CypA are not located in other pockets. So, the pocket 19 is the ligand binding site and the PHE113 and PHE60 are the activity residues (From **Supplementary Table 3**, we know that the pocket 13 is the ligand binding site of FKBP12 and residue TRY26, PHE36, PHE46, PHE59, PHE82 and PHE99 are activity residues).

Atom	Res	ID	Atom	Res	ID
NH1	ARG	55	CA	ASN	102
HH11	ARG	55	C	ASN	102
CE2	**PHE**	**60**	O	ASN	102
CZ	**PHE**	**60**	N	ALA	103
CG	MET	61	CB	ALA	103
SD	MET	61	CB	THR	107
CG	GLN	63	O	THR	107
CD	GLN	63	N	GLY	109
OE1	GLN	63	CA	GLY	109
NE2	GLN	63	N	SER	110
HE21	GLN	63	N	GLN	111
O	GLY	72	CB	GLN	111
CA	THR	73	CG	GLN	111
C	THR	73	CD	GLN	111
O	THR	73	OE1	GLN	111
CA	GLY	74	NE2	GLN	111
C	GLY	74	HE21	GLN	111
O	GLY	74	HE22	GLN	111
CA	GLY	75	**CB**	**PHE**	**113**
C	GLY	75	**CG**	**PHE**	**113**
O	GLY	75	**CD1**	**PHE**	**113**
CA	LYS	76	**CD2**	**PHE**	**113**
OE1	GLU	81	**CE1**	**PHE**	**113**
O	GLU	81	**CE2**	**PHE**	**113**
CG	LYS	82	**CZ**	**PHE**	**113**
CE	LYS	82	CD1	LEU	122
CA	ALA	101	HD13	LEU	122
CB	ALA	101	CD2	LEU	122
C	ALA	101	HD21	LEU	122
O	ALA	101	HD22	LEU	122
N	ASN	102	CE1	HSD	126

Table 4 Orbital coefficient of the activity residue PHE113 and PHE60 of CypA on the LUMOs. Those atoms that have the absolute value of orbital coefficient less than 0.1 are neglected. Therefore, atom CE2 of PHE60 and atom CG, CD1, CD2, CE1, CE2 and CZ of PHE113 are the activity atoms. (AR: Activity Residue; AO: Atomic Orbital; OC: Orbital Coefficient)

AR	Atom	AO	OC
PHE60	CD1	2X	-0.440
		2Y	0.114
		2Z	-0.403
	CD2	2X	0.439
		2Y	-0.112
		2Z	0.398
	CE1	2X	0.434
		2Y	-0.112
		2Z	0.396
	CE2	2X	-0.457
		2Y	0.116
		2Z	-0.419
PHE113	CA	2S	-0.145
		2X	0.100
		2Z	0.102
	HB	1S	0.147
	CG	2X	0.211
		2Y	-0.494
		2Z	0.420
	CD1	2X	-0.153
		2Y	0.386
		2Z	-0.307
	CD2	2Y	0.121
	CE1	2Y	0.112
	CE2	2X	-0.149
		2Y	0.387
		2Z	-0.293
	CZ	2X	0.192
		2Y	-0.496
		2Z	0.384

Table 5 Orbital coefficient of activity atoms of CsA and FK506 on the HOMO. Those atoms that have the absolute value of orbital coefficient less than 0.1 are neglected. The residue MVA11 is the activity residue of CsA with four activity atoms: N, CA, C and O. FK506 has nine activity atoms: C2, C8, C9, C10, C11, N7, O3, O4 and O5.

CsA			FK506					
MVA11: N	2Pz	-0.144	C2	2Py	-0.123	N7	2S	0.100
MVA11: CA	2S	0.211	C8	2S	-0.151		2Px	-0.120
	2Px	0.160		2Py	-0.235		2Pz	0.123
	2Py	-0.313	C9	2Py	0.151	O3	2Px	-0.111
	2Pz	-0.154		2Pz	0.132		2Py	0.378
MVA11: C	2S	-0.380	C10	2S	0.147	O4	2Py	-0.451
	2Py	0.337		2Py	-0.117		2Pz	-0.441
	2Pz	-0.443		2Pz	-0.331	O6	2Pz	0.112
MVA11: O	2Py	-0.341	C11	2Pz	0.113			
	2Pz	0.362						

We can see that none but the residues PHE60 and PHE113 are located not only on the LUMOs of CypA, but also in the pocket 19 (compare **Table 1** and **Table 3**). All other residues on the LUMOs of CypA are not located in any other pocket. Therefore, we can say that the pocket 19 is the ligand binding site and the PHE113 and PHE60 are the activity residues or that there are seven activity atoms: PHE113: CG, CD1, CD2, CE1, CE2, and CZ (these six atoms form a quincunx in **Fig. 2a,b**), and PHE60:CE2 as shown in **Table 4**. Though the atoms PHE113: CB, PHE113: CZ and PHE60: CZ are located on the LUMOs and in the pocket 19, their orbital coefficients are zero (**Table 4**), thus they are not activity atoms. Similarly, though the atoms PHE113: CA and PHE113: HB located on the LUMOs of CypA, they are not activity atoms because they are not located in the pocket 19.

It has been reported[8] that CsA interacted with **ARG55**, **PHE60**, **MET61**, **GLN63**, **GLY72**, **ALA101**, **ASN102**, ALA103, **GLN111**, **PHE113**, TRP121, **LEU122** and **HIS126** of CypA. The ligand binding pocket obtained by our method is number 19, which comprises **ARG55**, **PHE60**, **MET61**, **GLN63**, **GLY72**, THR73, GLY74, GLY75, LYS76, GLU81, LYS82, **ALA101**, **ASN102**, **GLN111**, **PHE113**, **LEU122** and **HIS126**. They are in good agreement.

We can obtain the activity residues of FKBP12 by the same way: TYR26, PHE36, PHE46, TRP59, PHE82 and PHE99 (compare **Table 1** and **Supplementary Table 3**), as well as the activity atoms as shown in **Supplementary Table 4**. Therefore, the pocket 13 is the ligand binding site of FKBP12.

X-ray experiment[9] reported that the binding pocket of FKBP12 comprised the residues **TYR26**, **PHE46**, **PHE99**, VAL55, ILE56 and **TRP59.** The activity residues we worked out are **TYR26**, PHE36, **PHE46**, **TRP59**, TRY82 and **PHE99**. They are also in good agreement.

The residue MVA11 of the ligand CsA is on the HOMO of CsA. Therefore, MVA11 is the activity residue of CsA. Four atoms of residue MVA11 on the HOMO with the absolute value of its atomic coefficient larger than 0.1 are considered to be the activity atoms of the CsA: MVA11:N, MVA11:CA, MVA11:C and MVA11:O. And the activity atoms of FK506 are: C2, C8, C9, C10, C11, N7, O3, O4 and O6 (**Table 5**).

Figure 2 shows the spatial configuration of the ligand binding pocket and the activity atoms of CypA, CsA, FKBP12 and FK506. The activity atoms of ligand are

fitting so well with the activity atoms of protein that indicates strong interactions between them (**Fig. 2b, d**). Especially, as NMR experiment pointed out[10], the CsA changed its configuration so that its activity residue MVA11 (the projecting part of the loop as shown in **Fig. 2a**) clearly trended toward the activity part of CypA.

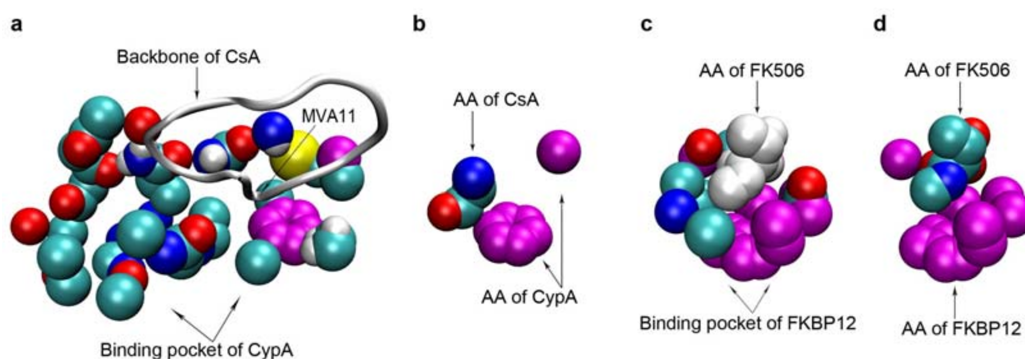

Figure 2 Spatial configurations of binding pockets and activity atoms (AA). (**a**) The binding pocket of protein CypA and backbone of ligand CsA. The activity residue MVA11 (the projecting part of the loop) of CsA clearly trends toward the activity part of CypA. (**b**) There are seven activity atoms of CypA, PHE113: CG, CD1, CD2, CE1, CE2, CZ (these six atoms form a quincunx in the bottom of the pocket.) and PHE60: CE2; and four activity atoms of CsA, MVA11: N, CA, C, O. (**c**) AA of FK506 (colored in gray to distinguish them from the binding pocket) fit well into the binding pocket of protein FKBP12. (**d**) There are nine activity atoms of FK506 and sixteen activity atoms of FKBP12 (see **Table 5** and **Supplementary Table 4**). The activity atoms of FK506 are enclosed by the activity atoms of FKBP12

DISCUSSION

Assuming we now have a new target protein, we can work out its LUMOs and pockets only from its coordinate file called pdb file, and we can immediately obtain its ligand binding site, activity residues and activity atoms. Then we can help identify which small molecules can be its ligand by checking its HOMO energy. If its HOMO energy is far away from the LUMOs energy of the target protein, we are sure this small molecule can not be its ligand.

These two rules enable us to have a clearer understanding of the mechanism of protein-ligand interaction on the basis of full electronic structure of a protein. The real reason which created these rules is due to the characteristics of energy level structure of a protein. They might be an important criterion for identifying or designing a ligand for a target protein. The first rule tells us the necessary condition for being the activity residues or activity atoms that they must locate on the LUMOs of a protein, and permits us to identify a ligand by checking its HOMO energy level. We can use the energy eigenvalues and orbital coefficients of the wave function to calculate many other characteristics of a protein. The second rule points out the necessary and sufficient condition for being the activity residues/atoms and the ligand binding site. Therefore, it allows us to identify not only the ligand binding site, but also the activity residues and activity atoms of a protein. When detecting the activity pocket, its volume and mouth-area can also help us identify a ligand.

We anticipate our method to be a starting point, which contributes to the study of the protein-ligand interaction and can assist the drug design on the basis of electronic

structure, and provides us with a more detailed understanding of the mechanism of protein-ligand interaction.

METHODS

Molecular Dynamic (MD) Simulation. Complex CypA/CsA and FKBP12/FK506 were performed MD simulation to obtain the ground state and the state of 300 K, respectively. Their starting coordinates were taken from the protein database bank[11] (PDB) with entry 1CWK and 1FKJ, respectively. The parallel MD program we used is NAMD[12] and all the molecules were solvated before the MD simulation.

However, for a ligand or other small molecules, there is often no topology and parameter information in the standard CHARMM's library. In that case, the web site HIC-UP[13] can help us to find the topology and parameter files for these small molecules. These files are X-PLOR version[14], not CHARMM version. The ligand FK506 was treated in this way.

For ligand CsA, a cyclic undecapeptide with the following primary structure: MeBmt1-Abu2-Sar3-MeLeu4-Val5-MeLeu6-Ala7-D-Ala8-MeLeu9-MeLeu10-MeVal 11. Residues 1, 3, 4, 6, 9, 10 and 11 are N-methylated on the amide nitrogen. Here, we created its topology and parameter information based on existing topology information from other molecules.

The basic idea to create the topology file of CsA is as follows:

(1). For residues MeBmt1 and Abu2, of which both not belong to the 20 kinds of

standard resides of protein, we need to write out their topology files. Details of creating those two topology files can be found at the book[6] from page 125 to 139.

(2). We used program PSFGEN[12] to build the protein data base (PDB) file and protein structure file (PSF) of CsA by using a patch LINK to link MeBmt1 and MeVal11 since it's a cyclic molecule and a patch METH to methylate the amide nitrogen of resides 1, 3, 4, 6, 9, 10 and 11.

After we get the topology file of CsA, we can use the program PSFGEN to build its new PDB and PSF files. PSFGEN, a CHARMM version, can generate a revised PDB and PSF files with hydrogen atoms added, as well as the water environment

For protein CypA, consisting of 165 residues and 2053 atoms, we also used the PSFGEN to get its new PDB and PSF files.

Finally, we set up the configuration file to run the MD simulation. The parameters for CHARMM force field was selected as follows: exclude = scaled 4, 1-4 scaling = 1.0, switching = on, switchdist = 8 Å, cutoff = 12 Å, pairlestdist = 13.5. We minimized the CypA and CsA for 80ps, separately.

For ligand FK506, with chemical formula C44H69NO12, we get its topology and parameter files from HIC-UP server and run program XPLOR64[13] to get its new PDB and PSF file. The XPLOR64, a XPLOR version, can generate a revised PDB and PSF files using topology file obtained from CNS[13].

For protein FKBP12, we also used XPLOR to generate its new PDB and PSF files.

We select the same MD parameters as CypA/CsA for FKPB12 and FK506

separately. They were firstly relaxed to near 20 K by running 20 ps. And then it was heated continuously from 20K to 290K with the temperature step 30 K, running 30 ps for each step. And we performed a smaller temperature step 5 K from 295 K to 300 K with 30 ps each step. Finally, run 50 ps at 300 K. The total time scale is 520 ps for the whole process using 16 CPUs.

Full Electronic Structure Calculation. We used ODA and a math method ENFC to calculate the electronic structure of protein CypA (165 residues) and FKBP12 (107 residues). After MD simulation, we obtained the coordinates of the proteins and ligands at temperature of interest. We needed to transform the format of these pdb files into ODA format. Then, we followed the four steps of ODA mentioned above to calculate their eigenvalues, orbital coefficients of wave functions and density of state (DOS).

We used program ArgusLab[15] to calculate the HOMO and the LUMO of ligand CsA and FK506.

ACKNOWLEDGEMENTS

This work was supported by National Basic Research Program of China (no.2006CB504509), and the Project of the State Key Program of National Natural Science Foundation of China (Grant No. 10635060). The authors thank Ye Yuanjie, Wang Xun and Ye Ling for their kind help and the Modern Applied Mathematical Key Laboratory in Shanghai (the Department of Mathematics at Fudan University) and Shanghai Supercomputer Center (SSC) for providing the parallel computer.

COMPETING INTEREST STATEMENT

The authors declare that they have no competing financial interests.

References

1. Ladik, J., Seel, M., Otto, P. & Bakhshi, A.K. On the electronic structure and conduction properties of aperiodic DNA and proteins. I. Strategy and methods of investigations. *Chem. Phys.* **108**,203-214 (1986).
2. Bakhshi, A.K., Otto, P., Ladik, J. & Seel, M. On the electronic structure and conduction properties of aperiodic DNA and proteins. II. Electronic structure of aperiodic DNA. *Chem. Phys.* **108**,215-222 (1986).
3. Otto, P., Bakhshi, A.K., Ladik, J., Seel, M. & Chin, S. On the electronic structure and conduction properties of aperiodic DNA and proteins. III. The band structures of periodic polypeptides. *Chem. Phys.* **108**,223-231 (1986).
4. Ye, Y.J. The extended negative factor counting method for tridiagonal block matrices with cross links. *J. Math. Chem.* **14**,121-139 (1993).
5. Xu, L.N., Zhang, L.L., Guo, Z. & Zhou, L.X. Full electronic structure calculation of the biological activity in P173L enzyme. *Int. J. Quantum Chem.* **106**,1544-1550 (2006).
6. Zhou, L.X. & Shuai, J.W. *Protein Computer Simulation* (Xiamen Univ. Press, Xiamen, China, 2008).
7. Dundas, J. *et al.* CASTp: computed atlas of surface topography of protein with structural and topographical mapping of functionally annotated residues. *Nucleic*

Acids Res. **34**, W116-W118 (2006).

8. Ke, H.M. & Huai, Q. Crystal structure of cyclophilin and its partners. *Frontiers in Bioscience.* **9**, 2285-2296 (2004).
9. Van Duyne, D.G., Standaert, R.F., Karplus, P.A., Schreiber, S.L. & Clardy, J. Atomic structure of FKBP-FK506, an immunophilin-immunosuppressant complex. *Science.* **252**,839-842 (1991).
10. Theriault, Y. *et al.* Solution structure of the cyclosporin A/cyclophilin complex by NMR. *Nature.* **361**,88-91 (1993).
11. Berman, H.M. *et al.* The Protein Data Bank**.** *Necleic Acids Res.* **28**,235-242 (2000).
12. Phillips, J.C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**,1781-1802 (2005).
13. Kleywegt, G.J., Henrick, K., Dodson, E.J. & *van* Aalten, D.M.F. Pound-wise but penny-foolish: how well do micromolecules fare in macromolecular refinement? *Sturcture.* **11**, 1051-1059 (2003).
14. Brunger, A.T. *et al.* Crystallography & NMR System (CNS), A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**,905-921 (1998).
15. Thompson, M.A. & Zerner, M.C. A theoretical examination of the electronic structure and spectroscopy of the photosynthetic reaction center from rhodopseudomonas. *J. Am. Chem. Soc.* **113**, 8210-8215 (1991).