

K301

Chemocavity: Specific Concavity in Protein Reserved for the Binding of Biologically Functional Small Molecules

Shinji SOGA^{*1}, Hiroki SHIRAI^{*1}, Masato KOBORI^{*1}, Noriaki HIRAYAMA^{*2}

1) *Molecular Medicine Research Laboratories, Drug Discovery Research, Astellas
Pharma Inc., 21 Miyukigaoka, Tsukuba, Ibaraki 305-8585, Japan*

2) *Basic Medical Science and Molecular Medicine, Tokai University School of
Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1143, Japan*

shinji.soga@jp.astellas.com

1. Introduction

The idea that there should be a specific site on a protein for a particular functional small molecule is widespread. It is, however, usually not so easy to understand what characteristics of the site determine the binding ability of the functional small molecule. Recently, we have developed an index named propensity for ligand binding (PLB index) to identify the specific concavity in a protein for binding of druglike small molecules¹. The PLB index is based on the characteristic appearances of the 20 standard amino acids at the concavity. In this study, we have expanded the concept of the PLB index in order to make it sensitive enough to apply in identification of various kinds of concavities. We have focused on the concurrence rate of the 20 standard amino acids at such binding sites².

2. Method

We selected the quality X-ray structures of complexes between proteins and small organic molecules from PDB³. This quality data set was selected by use of the following criteria: the R_{free} value of less than or equal to 0.24, the resolution value of less than or equal to 2.5 Å, the occupancy factors of 1.0 for all non-hydrogen atoms of the small molecule, and the atomic displacement parameters of less than 50 Å² for all non-hydrogen atoms of the small molecule. The redundant proteins were removed from the data set by consulting the “Non-redundant PDB chain set (NRPDB) [<http://www.ncbi.nlm.nih.gov/Structure/VAST/>]” resource. We classified the small organic molecules according to their properties, which are molecular weight, SlogP and MACCS structural keys, to obtain a set of canonical molecular groups. The proteins were then classified according to the canonical molecular groups that are bound to the proteins. The specific concavity for a particular canonical molecular group is designated as chemocavity. And then, we analysed the amino acids concurrence rates in the concavities that share the same canonical molecular group. The amino acids concurrence rates can be expressed by concurrence matrixes. By use of linear combination of concurrence rates with frequencies of the two amino acid combinations in the concavity, we expressed the characteristics of the concavities as the chemocavity indices.

3. Results and Discussion

For selection of molecular groups that well represent canonical molecular groups, only the groups consisting of more than nine molecules were included in the consideration. The total 48 independent groups were picked out and used in this study as canonical molecular groups. For example, we show the concurrence matrix of a

chemocavity for a canonical molecular group which mainly included porphyrin molecules in Figure 1. The concurrence rates greater than 1.0 are colored in red. It means that the corresponding concurrent pairs in the chemocavity appear more than expected. The concurrence rates of Leu-Leu, Leu-Phe, Phe-Phe, Cys-Cys, and Cys-Met pairs are greater than 3.0 in the chemocavity for porphyrin group, which is the most

outstanding aspect of the chemocavity. We illustrated porphyrin binding site in Figure 2. Since the chemocavity for porphyrin group is rich in concurrence of Leu, Cys, Met, and Phe, these amino acids are depicted by ball & stick in the two different proteins.

Although the three-dimensional arrangements of these amino acids around the porphyrin group are markedly different, the chemocavity indexes calculated from the concurrence rates of these amino acids clustered around the porphyrin group unequivocally suggest that these proteins share a common chemocavity for the porphyrin group. It is particularly worth noting that the concurrence matrix can characterize the chemocavity, although the concurrence matrix does not directly reflect the three-dimensional structures of the concavities where small molecules are bound. We confirmed that any other chemocavities showed the same tendency.

References

1. Soga S, Shirai H, Kobori M, Hirayama N. *J. Chem. Inf. Model.* **2007**, *47*, 400–406.
2. Soga S, Shirai H, Kobori M, Hirayama N. *J. Chem. Inf. Model.* **2008**, *48*, 1679–1685.
3. Berman H M, Westbrook J, Fenz Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N, Bourne P E. *Nucleic Acids. Res.* **2000**, *28*, 235–242.

	GLY	ALA	VAL	LEU	ILE	HIS	PHE	TYR	TRP	SER	THR	ASP	GLU	ASN	GLN	ARG	LYS	CYS	MET	PRO
GLY	0.4																			
ALA	0.7	1																		
VAL	0.9	1.3	1.8																	
LEU	1.8	2.1	2.8	4.2																
ILE	0.7	1.3	1.5	2.4	1.6															
HIS	0.9	1.4	2.1	3	1.6	1.8														
PHE	1.1	2	2.3	3.7	2.3	2.4	3.1													
TYR	0.4	0.8	1.3	1.7	1	1.3	1.3	0.6												
TRP	0.3	0.7	0.7	1	0.9	0.7	0.9	0.5	0.3											
SER	0.4	0.6	0.7	1.2	0.6	1	1	0.5	0.4	0.3										
THR	0.6	0.9	1.1	2	1.1	1.3	1.8	0.7	0.5	0.5	0.9									
ASP	0.1	0.2	0.3	0.4	0.2	0.4	0.3	0.2	0.1	0.2	0.3	0.1								
GLU	0.1	0.3	0.3	0.6	0.4	0.4	0.5	0.3	0.2	0.2	0.3	0.1	0.1							
ASN	0.3	0.5	0.8	1.1	0.7	0.8	1	0.5	0.2	0.3	0.5	0.1	0.2	0.3						
GLN	0.7	1.4	1.9	3.2	1.9	1.8	0.9	0.7	0.6	1.1	0.2	0.3	0.6	0.8						
ARG	0.5	0.9	1	1.7	1.1	1.1	1.6	0.6	0.5	0.6	0.8	0.2	0.3	0.5	0.9	0.7				
LYS	0.5	0.9	1.3	1.6	1	1.4	1.5	0.9	0.4	0.4	0.9	0.2	0.2	0.4	0.8	0.5	0.8			
CYS	1	1.7	2.3	2.6	2	2.7	2.2	1.4	0.9	1	2	0.6	0.5	1.1	2.1	1.2	1.8	3.5		
MET	1	1.5	1.9	2.6	1.4	2.3	2.2	1.3	0.8	1	1.3	0.3	0.5	0.7	1.8	1.1	1.4	3.1	2.5	
PRO	1	1.5	1.6	2.4	1.5	1.7	1.8	1.1	0.8	0.8	1.1	0.3	0.3	0.7	1.2	1.1	0.9	2.6	2	1.2

Fig.1 concurrence matrix of a chemocavity for a canonical molecular group which mainly included porphyrin molecules. The concurrence rates greater than 1.0 are colored in red.

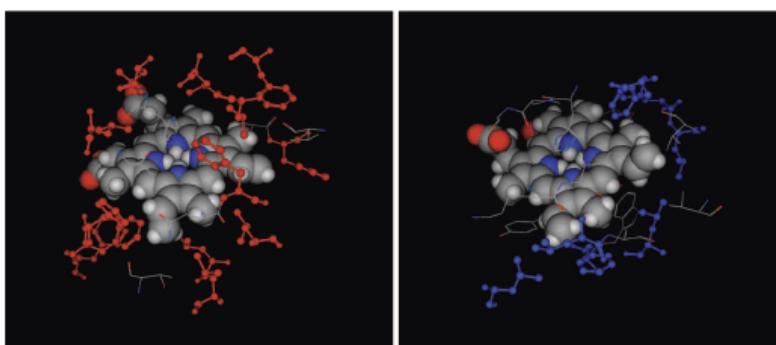


Fig.2 concurrently existing amino acids at chemocavities for porphyrin group. The left and right structures are 1J3Y(2.57) and 1SOX(2.00), respectively. The values in the parentheses indicate the chemocavity indexes. Leu, Cys, Met, and Phe are depicted by ball & stick.