

# Docking and Virtual Screening Studies for New Leads of Boar Salivary Lipocalin

Nack-Do Sung, Chang-Sik Park, Hyung-Yeon Park,<sup>†</sup> and Chan Kyung Kim<sup>†\*</sup>

Research Center for Transgenic Cloned Pigs, Division of Applied Biology & Chemistry, Chungnam National University, Daejeon 305-764, Korea

<sup>†</sup>Department of Chemistry, Inha University, Incheon 402-751, Korea. \*E-mail: kckjung@inha.ac.kr

Received March 21, 2008

We have performed FlexX docking experiments to predict the best docking poses of 5-androst-16-en-3-ol or 5-androst-16-en-3-one to Boar salivary lipocalin (SAL). Since no steroids were found inside of the binding pocket of the X-ray structure of 1GM6, we tried to find docking structures after opening the pocket using the random tweak option implemented in SYBYL. This operation allowed the ligand to enter the pocket. The best poses generated from FlexX were different from the structures reported earlier, which calculated docking poses by manual docking followed by minimization. Analysis of docking poses allowed us to identify pharmacophores. From this information, virtual screening experiments using UNITY were performed. Among six candidates, 3-(3,7-dimethyloct-6-enylamino)propane-1,2-diol (Leadquest code name: 5755) was chosen for further development. Future work will involve synthesis of some derivatives of 5755 and biological experiments if any derivatives can control the biostimulation and improve reproductive efficiency in pigs.

**Key Words :** Boar salivary lipocalin, Steroid pheromones, FlexX docking, UNITY virtual screening

## Introduction

Boar salivary lipocalin (SAL) is a porcine odorant-binding protein (POBP) specifically expressed in the submaxillary glands of the boar. While produced in large amount (hundreds of milligrams) in the gland, it is unable to be detected in the sow. When extracted from the natural sources, this protein is associated with boar's pheromones. Chemical communication in the pig has been known for a long time to be mediated by saliva.<sup>1</sup> In particular, steroid pheromones, such as 5-androst-16-en-3-ol (androst-16-en-3-ol, **I**) and 5-androst-16-en-3-one (androst-16-en-3-one, **II**), present in the saliva of mature males, have been shown to act as pheromones on the female.<sup>2</sup>

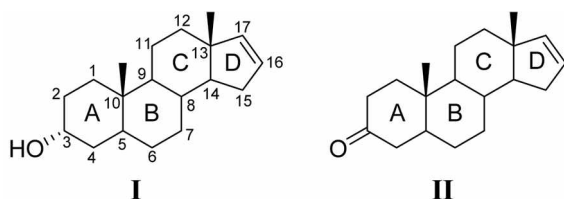
In an earlier study by S. Spinelli and coworkers,<sup>3</sup> an X-ray structure of SAL was determined (PDB code = 1GM6). Surprisingly, in the X-ray structure a glycerol molecule was found in the binding cavity since the cavity might be too small to accommodate the steroids. Therefore, to obtain the docking poses of **I** or **II** to SAL, each ligand was manually introduced into the cavity in two different orientations - C-OH (or C=O) group pointing toward the entrance or the bottom of the cavity. Then each starting structure was minimized to find an optimum position. As a result of optimization, the volume of the cavity increases due to conformational changes of Val45 and Ala73. If this is true,

there still remains a question: how the ligand can enter the binding pocket? To answer this question, we performed the docking studies using different approach, namely FlexX<sup>4</sup> docking, as our continuing works on computer-aided molecular design.<sup>5</sup> From the new docking poses, we can find a better way of docking the steroids and of identifying pharmacophores. Using this information, virtual screening study is also carried out to find out new lead compounds. To search of a new porcine pheromonal odorants for biostimulation control system technologies to offer a potentially useful and practical way to improve reproductive efficiency in pigs, this work will be a first step in finding a new active pig pheromonal odorants.

## Computation

Since the binding site was not large enough to accommodate one of the steroids, we tried to open it by using the tweak option in SYBYL. Random tweak was developed to predict the conformations of loops in proteins and have been applied successfully in several proteins.<sup>6,7</sup> By doing this operation, we could open the entrance of the cavity to allow the ligand to enter the binding pocket without problem. In this sense, our approach might be better than earlier one.

FlexX is a fast, flexible docking method that uses an incremental construction algorithm to place flexible ligands into a rigid active site and is known to perform well to reproduce X-ray structures.<sup>8</sup> Standard parameters of the FlexX program as implemented in SYBYL 6.9<sup>9</sup> were used during docking. Bond lengths and angles were kept constant as given in the optimized structures. As a preliminary step for docking, structures of **I** and **II** were minimized using TRIPOS force field until the root-mean-square (rms) gradient converged to 0.005 kcal/mol<sup>10</sup> and these ligands were



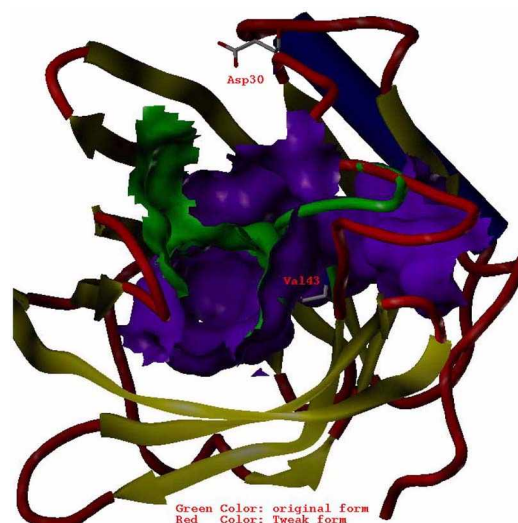
Scheme 1

docked to the tweaked crystal structure of SAL. FlexX dockings of **I** and **II** were performed using the Run-Multiple Ligand option of FlexX. From the best docking poses, important interactions between the ligand and receptor were identified. Based on the analysis of interaction within the binding pocket, we performed virtual screening study using UNITY<sup>11</sup> module implemented in SYBYL packages. UNITY is a search and analysis system for exploring chemical database. Conformationally flexible 3D search rapidly finds molecules that can satisfy queries regardless of the conformation stores in a database. Here, we composed queries based on pharmacophores identified during FlexX docking study. The selected candidates were tested for their toxicity using TOPKAT from Acclerys Inc.<sup>12</sup>

## Results and Discussion

**Binding site.** In order to perform docking experiment, it was necessary to examine the binding region of the X-ray structure of 1GM6.<sup>3</sup> As mentioned above, the binding pocket formed a closed cavity with a glycerin molecule inside and thus its entrance should be open to accommodate a larger ligand. There were two polar residues at the bottom of the active site (Glu121 and Tyr123), which could form strong hydrogen bond and/or polar interaction with the ligand. The active site was enclosed by several hydrophobic residues such as Val45, Phe58, Ala73, Val85, Phe93, Leu106, and Leu108. The active site had volume of 274 Å<sup>3</sup> and Log P of 15.5.

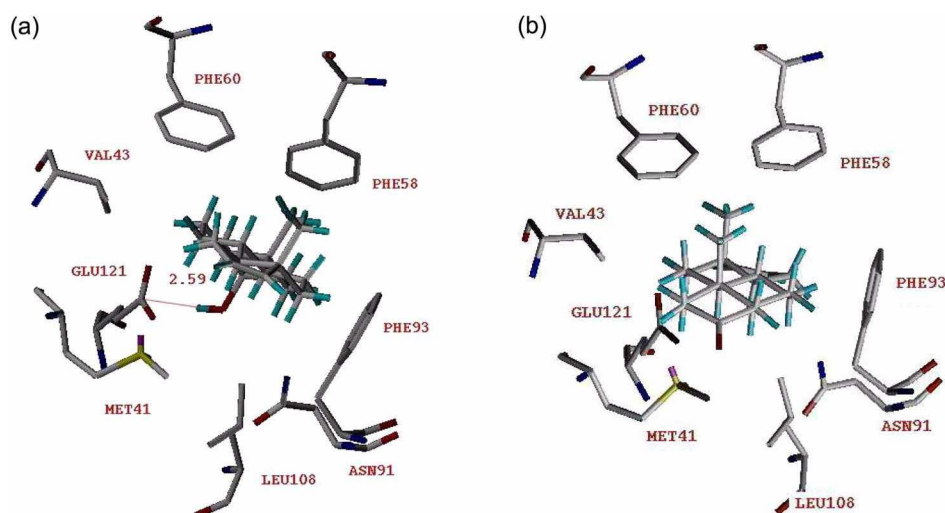
To open the binding pocket, we selected the most flexible loop located near the entrance of the protein for flapping. Two loops connecting two  $\beta$ -sheets were found: loop 1 was consisted of three residues (Tyr87-Gly89) and loop 2 consisted of 14 residues (Asp30-Val43). Since loop 1 was too small to have sufficient mobility, loop 2 was chosen for this purpose. Mobility of the loop was tested using the tweak option<sup>6,7</sup> in Sybyl. The conformational change before and after this operation is shown in Figure 1. FlexX docking was



**Figure 1.** Conformational change of loop2 by random tweak operation.

performed while this loop was open wide.

**Docking of androstenol and androstenone.** FlexX dockings of **I** and **II** were performed after opening the loop 2 at the entrance of the binding region. The best poses are drawn in Figure 2. The two ligands in Figure 2 showed similar orientations in the pocket. The  $-OH$  group of **I** shown in Figure 2(a) was pointing toward the bottom of the pocket to make a strong hydrogen bond with  $-CO_2^-$  group of Glu121. On the other hand, such an orientation could cause electrostatic repulsion between  $-CO_2^-$  group and  $C=O$  group of **II**,  $C=O$  group of **II** moved away from the Glu121 as shown in Figure 2(b). Due to lack of strong hydrogen bond in **II**, FlexX docking scores of **II** ( $-4.50$  kcal/mol) was worse than that of **I** ( $-6.78$  kcal/mol). Regardless of the orientations of the ligands, hydrophobic interaction with the surrounding lipophilic residues was a key factor in positioning the ligands in the binding pocket. However, note that the best poses found in this work were different from those



**Figure 2.** Best docking poses of **I** (a) and **II** (b) at the binding pocket of SAL.

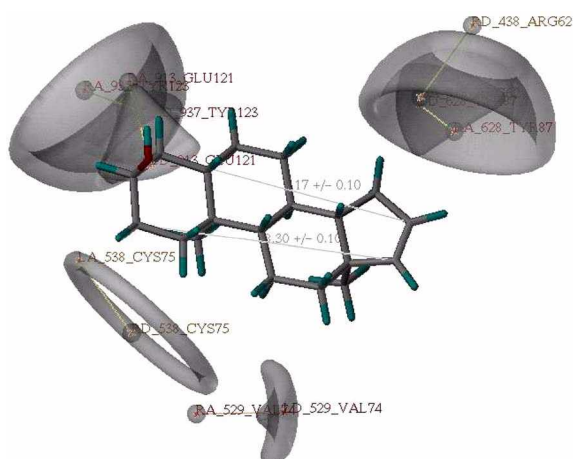


Figure 3. Query option for virtual screening using UNITY.

reported earlier,<sup>3</sup> which manually introduced the ligands in the cavity. In an earlier study, the polar –OH group of androstenol was positioned toward the bottom of the cavity but without interaction with the polar residues (Glu121 or Tyr123) and the polar C=O group of androstenone was pointing toward the entrance of the binding pocket. Therefore, we might think that our approach was more rational in that we did not assume any manual operation during docking experiments.

**Search for new ligands.** From the docking studies, we found that the most important interactions were the hydrogen bonding with Glu121 as well as hydrophobic interaction with residues surrounding the binding pocket. To locate compounds in databases that matched with these pharmaco-

phores, virtual screening using UNITY from the docking pose of **I** was performed. As shown in Figure 3, the following queries were set up: (1) distance constraints – C<sub>2</sub>-C<sub>17</sub>: 8.30 ± 0.10 Å and C<sub>4</sub>-C<sub>16</sub>: 8.17 ± 0.10 Å (atom numbering is shown in Scheme 1). (2) receptor donor and corresponding ligand acceptor sites: Arg62, Val74, Cys75, Tyr123 and Glu121 (3) an additional receptor donor/ligand acceptor site: Tyr87 as hydrogen bond acceptor/donor. Three-dimensional molecular candidates were generated by using UNITY from 30,000 entries of the LeadQuest<sup>13</sup> database from Tripos. 3-(3,7-dimethyloct-6-enylamino)propane-1,2-diol (Leadquest code name: **5755**), N<sup>1</sup>-allyl-N<sup>2</sup>-((tetrahydrofuran-2-yl)methyl)oxalamide (LeadQuest code name: **6060**), 1-allyl-3-(4-(1,1,1,3,3,3-hexafluoro-2-hydroxypropan-2-yl)phenyl)thiourea (LeadQuest code name: **6431**), N-[(4-methoxyphenyl)methyl]-3-[[3,5,6,8-tetrahydro-6,6-dimethyl-4-oxo-3-(2-propenyl)-4H-pyrano[4',3':4,5]thieno[2,3-d]pyrimidin-2-yl]thio]-propanamide (LeadQuest code name: **6663**), N<sup>1</sup>-(4-(allyloxy)phenyl)-N<sup>2</sup>-(2-hydroxypropyl)oxalamide (LeadQuest code name: **7111**), and 2-(4-methyl-2-oxo-2H-chromen-7-yloxy)-N-((tetrahydrofuran-2-yl)methyl)acetamide (LeadQuest code name: **7113**) were found as possible candidates and their docking score were evaluated. The docking scores and two typical docking poses are shown in Table 1 and Figure 4, respectively. In Table 1, results of toxicity prediction using TOPKAT are also summarized. It is interesting to examine the docking poses shown in Figure 4 in detail. The two ligands, **5755** (Figure 4(a)) and **6060** (Figure 4(b)), fitted well into the binding pocket. However, in the cases of other four compounds, some part of the side chain was located outside of the binding pocket as shown for **6431** in

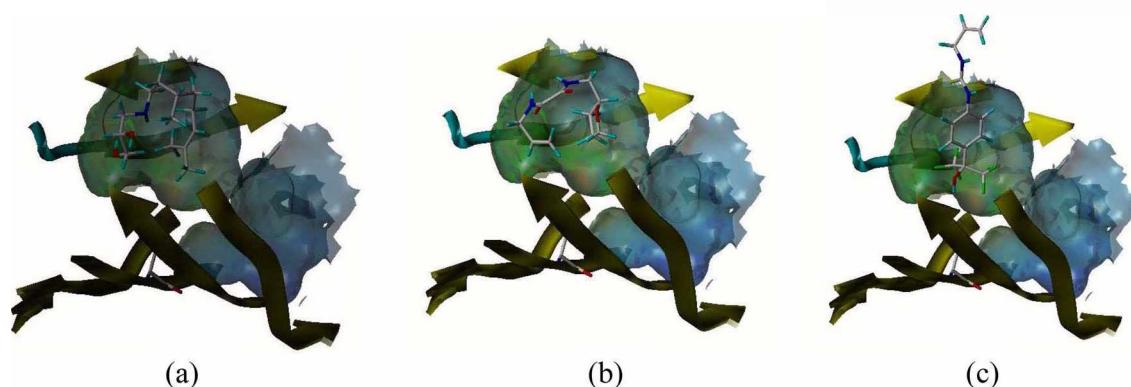


Figure 4. Three typical docking poses of selected candidates. (a) **5755** (b) **6060**, and (c) **6431**.

Table 1. Results of docking score and toxicity prediction for six candidates

Compound No.	Volume <sup>b</sup>	Docking score <sup>a</sup>	FDA <sup>c</sup>	AM <sup>d</sup>	SI <sup>e</sup>	DTP <sup>f</sup>	LD <sub>50</sub> <sup>g</sup>
<b>5755</b>	227.4	-39.4	0	0.00	0.06	0.00	> 10
<b>6060</b>	182.0	-38.1	1	0.98	0.00	0.00	1.8
<b>6431</b>	235.4	-39.9	0	0.00	0.00	0.00	201.7
<b>6663</b>	400.5	-52.7	0	0.00	0.99	0.99	497.6
<b>7111</b>	233.0	-39.0	1	0.64	0.94	0.00	2.8
<b>7113</b>	253.6	-38.5	1	0.00	0.00	0.00	2.8

<sup>a</sup>in kcal/mol. <sup>b</sup>in Å<sup>3</sup>. <sup>c</sup>FDA Rodent carcinogenicity. Non-carcinogen: FDA < 0.3, carcinogen: FDA > 0.7. <sup>d</sup>Ames mutagenicity. Non-mutagen: AM < 0.3, mutagen: AM > 0.3. <sup>e</sup>Rabbit skin irritancy. Non-skin irritancy: SI < 0.3, skin irritancy: SI > 0.7. <sup>f</sup>Developmental toxicity potential. No potential for developmental toxicity: DTP < 0.3, potential developmental toxicity: DTP > 0.7. <sup>g</sup>Rat oral LD<sub>50</sub> (g/kg)

Figure 4(c). As summarized in Table 1, the computed values of toxicity for **5755** were determined to be within non-toxic in all dimensions of a model Optimum Prediction Space. Therefore, we selected **5755** as a possible candidate for further development and experimental studies are in progress in our lab to validate docking/virtual screening results. In future work, we will synthesize some derivatives of **5755** and perform biological experiments if these compounds can control the biostimulation and improve reproductive efficiency in pigs.

### Conclusion

FlexX docking studies were performed to identify key residues in salivary lipocalin. Before docking, loop 2 was open to make the ligands to enter the binding pocket using tweak option in Sybyl. FlexX docking experiments were performed with Run-Multiple Ligand option. The best docking poses found in the work were different from those found earlier. Using the pharmacophores from docking study, virtual screening using UNITY was performed. Six candidates were identified among 30,000 compounds in the database. One ligand, **5755**, is selected as a lead for further synthesis and examination for biostimulation in pigs.

**Acknowledgements.** This work was supported by a grant (No. R11-2002-100-05003) from the excellent research Center of KOSEF.

### References

- Gower, D. B.; Booth, W. D. *Ontog. Olfaction* **1986**, 255.
- (a) Melrose, D. R.; Reed, H. C.; Patterson, R. L. *Br. Veteri. J.* **1971**, 127, 497. (b) Reed, H. C. B.; Melrose, D. R.; Patterson, R. L. S. *Br. Veteri. J.* **1974**, 130, 61.
- Spinelli, S.; Vincent, R.; Pelosi, P.; Tegoni, M.; Cambillau, C. *Eur. J. Biochem.* **2002**, 269, 2449.
- (a) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, 261, 470. (b) Rarey, M.; Kramer, B.; Lengauer, T. *J. Comput.-Aided Mol. Des.* **1997**, 11, 369.
- (a) Lee, D. Y.; Hyun, K. H.; Park, H. Y.; Lee, B.-S.; Kim, C. K. *Bull. Korean Chem. Soc.* **2006**, 27, 273. (b) Kim, C. K.; Lee, K. A.; Zhang, H.; Cho, H.; Lee, B.-S. *Bull. Korean Chem. Soc.* **2007**, 28, 1141. (c) Kim, C.; Kim, J.; Park, H.-Y.; McLean, R. J. C.; Kim, C. K.; Jeon, J.; Yi, S.-S.; Kim, Y. G.; Lee, Y.-S.; Yoon, J. *J. Microbiol. Biotechnol.* **2007**, 17, 1598.
- Fine, R. M.; Wang, H.; Shenkin, P. S.; Yarmush, D. L.; Levinthal, C. *Proteins* **1986**, 1, 342.
- Shenkin, P. S.; Yarmush, D. L.; Fine, R. M.; Wang, H.; Levinthal, C. *Biopolymers* **1987**, 26, 2053.
- Hoffmann, D.; Kramer, B.; Washio, T.; Steinmetzer, T.; Rarey, M.; Lengauer, T. *J. Med. Chem.* **1999**, 42, 4422.
- SYBYL 6.9; TRIPOS Inc.: St. Louis, USA.
- (a) Powell, M. J. D. *Mathematical Programming* **1977**, 12, 241. (b) Clark, M.; Crammer, R. D.; Opdenbosch, N. V. *J. Comput. Chem.* **1989**, 10, 982. (c) Dixon, D. A.; Zeroka, D. J.; Wendoloski, J. J.; Wasserman, Z. R. *J. Phys. Chem.* **1985**, 89, 5334.
- (a) Clark, R. D. *J. Chem. inf. Compt. Sci.* **1997**, 37, 1181. (b) Dixon, S. L.; Koehler, R. T. *J. Med. Chem.* **1999**, 42, 2887.
- TOPKAT, version 6.2; Accelrys Inc.: San Diego, CA, 2001.
- LeadQuest Chemical Compound Libraries; Tripos, Inc.: St. Louis, MO, 2000; Vol. 1-3.