CRYSTAL GROWTH & DESIGN
2008
VOL. 8, NO. 12
4268–4276

# Rational Design for Crystallization of β-Lactoglobulin and Vitamin $D_3$ Complex: Revealing a Secondary Binding Site†

Ming Chi Yang,‡ Hong-Hsiang Guan,§,# Jinn-Moon Yang,‡ Cheng-Neng Ko,‡ Ming-Yih Liu,§ Yih-Hung Lin,§ Yen-Chieh Huang,§ Chun-Jung Chen,*,§,||,▽ and Simon J. T. Mao*,‡,⊥

*Department and College of Biological Science and Technology, National Chiao Tung University, 75 Po-Ai Street, Hsinchu 30068, Taiwan, ROC, Life Science Group, Scientific Research Division, National Synchrotron Radiation Research Center, 101 Hsin-Ann Road, Hsinchu 30076, Taiwan, ROC, Department of Physics, Institute of Bioinformatics and Structural Biology, National Tsing Hua University, 101, Section 2 Kuang Fu Road, Hsinchu 30013, Taiwan, ROC, Department of Biotechnology, Asian University, 500, Liufeng Road, Wufeng, Taichung 41354, Taiwan, ROC, and Institute of Biotechnology, National Cheng Kung University, 1 University Road, Tainan 70101, Taiwan, ROC*

*Received June 30, 2008; Revised Manuscript Received September 30, 2008*



**ABSTRACT:** β-Lactoglobulin (LG) is a major milk whey protein containing primarily a calyx for vitamin $D_3$ binding, although the existence of another site beyond the calyx is controversial. Using fluorescence spectral analyses in the previous study, we showed the binding stoichiometry for vitamin $D_3$ to LG to be 2:1 and a stoichiometry of 1:1 when the calyx was "disrupted" by manipulating the pH and temperature, suggesting that a secondary vitamin D binding site existed. To help localize this secondary site using X-ray crystallography in the present study, we used bioinformatic programs (Insight II, Q-SiteFinder, and GEMDOCK) to identify the potential location of this site. We then optimized the occupancy and enhanced the electron density of vitamin $D_3$ in the complex by altering the pH and initial ratios of vitamin $D_3$/LG in the cocrystal preparation. We conclude that GEMDOCK can aid in searching for an extra density map around potential vitamin D binding sites. Both pH (8) and initial ratio of vitamin $D_3$/LG (3:1) are crucial to optimize the occupancy and enhance the electron density of vitamin $D_3$ in the complex for rational-designed crystallization. The strategy in practice may be useful for future identification of a ligand-binding site in a given protein.

## 1. Introduction

Bovine β-lactoglobulin (LG) is a major protein in milk comprising about 10−15%.[1] Because of its thermally unstable and molten-globule nature, LG has been studied extensively for its physical and biochemical properties.[2−7] Some essential functions of LG including hypocholesterolemic effect,[8] retinol transport,[9,10] and antioxidant activity[11−14] have been reported. According to the crystal structure, LG comprises predominantly a β-sheet configuration containing nine antiparallel β-strands from A to I (Figure 1a).[15−17] Topographically, β-strands A−D form one surface of the barrel (calyx), whereas β-strands E−H form the other. The only α-helical structure with three turns is located at the COOH-terminus, which follows the β-strand H beyond the calyx.[18] A remarkable feature of the calyx is its ability to bind hydrophobic molecules such as retinol, fatty acids, and vitamin $D_3$ (Figure 1b).[19−22]

The location of vitamin D binding sites has been controversial, yet most evidence points toward the calyx. It has been postulated that another site exists,[23−25] but it could not be verified by the crystal structure using the vitamin $D_2$−LG complex.[10] In a previous study, we first conducted a ligand binding assay using the fluorescence changes by retinol, palmitic

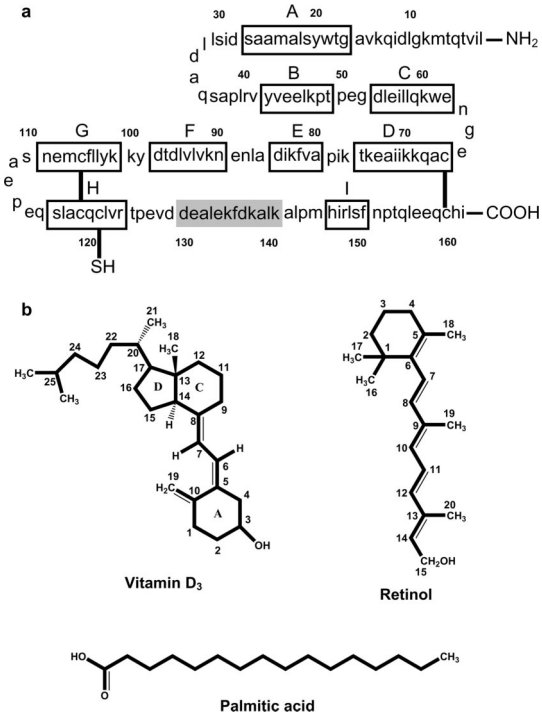


**Figure 1.** Primary structure of LG and chemical structure of its binding ligands. (a) LG comprises 162 amino acids with nine antiparallel β-sheet strands (A−I) and one α-helix (in gray). There are two disulfide bonds located between β-strand D and carboxyl-terminus (Cys-66 and Cys-160) and between β-strands G and H (Cys-106 and Cys-119), while a free buried thiol group is at Cys-121. (b) Chemical structure of ligands which bind to LG, such as vitamin $D_3$, retinol, and palmitic acid.

† Part of the special issue (Vol 8, issue 12) on the 12th International Conference on the Crystallization of Biological Macromolecules, Cancun, Mexico, May 6−9, 2008.

* To whom correspondence should be addressed. (S.J.T.M.) E-mail: mao1010@ms7.hinet.net; (C.-J.C.) E-mail: cjchen@nsrrc.org.tw.
‡ National Chiao Tung University.
§ National Synchrotron Radiation Research Center.
# Institute of Bioinformatics and Structural Biology, National Tsing Hua University.
|| Department of Physics, National Tsing Hua University.
▽ National Cheng Kung University.
⊥ Asian University.

10.1021/cg800697s CCC: $40.75 
Published on Web 10/31/2008

**Table 1. Relative Ligand Binding Ability of LG at the Various pH and Preheated Temperatures Extracted from Our Published Paper[26]**

| relative binding (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| titration experiment | | | | effect of pH | | | | effect of heating (preheated for 16 min) | | | |
| [ligand]/ [native LG] | vitamin $D_3$ | retinol | palmitic acid | pH | vitamin $D_3$ | retinol | palmitic acid | temperature (°C) | vitamin $D_3$ | retinol | palmitic acid |
| 0.125 | 7.1 | 13.2 | 15.1 | 2.0 | **29.1** | 2.2 | 2.0 | 50 | 93.4 | 92.1 | 92.6 |
| 0.25 | 14.1 | 26.0 | 30.0 | 3.0 | **29.0** | 0 | 0 | 60 | 91.1 | 90.2 | 90.4 |
| 0.375 | 21.2 | 38.8 | 44.9 | 4.0 | **29.3** | 0 | 0 | 70 | 83.0 | 67.0 | 72.3 |
| 0.5 | 28.1 | 50.6 | 59.0 | 5.0 | **29.3** | 0 | 0 | 80 | 49.2 | 10.2 | 14.4 |
| 0.625 | 35.3 | 62.0 | 71.5 | 6.0 | 46.7 | 8.5 | 38.0 | 90 | 41.9 | 7.5 | 6.8 |
| 0.75 | 42.3 | 72.7 | 80.4 | 7.0 | 58.1 | 27.0 | 49.4 | 100 | **40.7** | 6.7 | 5.8 |
| 0.875 | 49.3 | 81.3 | 86.9 | 8.0 | 100 | 87.7 | 100 | | | | |
| 1.0 | **56.2** | 88.0 | 91.9 | 9.0 | 78.2 | 100 | 82.6 | | | | |
| 1.25 | 70.6 | 91.8 | 96.0 | 10.0 | 77.7 | 97.6 | 54.3 | | | | |
| 1.5 | 82.9 | 95.0 | 98.5 | | | | | | | | |
| 1.75 | 92.0 | 96.7 | 99.2 | | | | | | | | |
| 2.0 | 96.0 | 98.0 | 99.7 | | | | | | | | |

**Table 2. Possible Sites Predicted by Insight II, Q-SiteFinder, and GEMDOCK**

| | | | GEMDOCK | |
|---|---|---|---|---|
| programs | Insight II | Q-SiteFinder | vitamin $D_3$ docking for whole protein | vitamin $D_3$ docking for other parts of protein besides sites 1 and 2 |
| ranking fitness (ratio) | 1 | 1 | −87.2 (50%) | |
| Site 1 (calyx) | | | | |
| Site 2 | 2 | 2 | −88.4 (20%) | |
| Site 3 | 3 | 8 | −85.4 (6%) | −83.9 (40%) |
| Site 4 | | 3 | −82.4 (6%) | −81.2 (20%) |
| Site 5 | | 5 | −75.4 (3%) | −80.1 (3%) |
| Site 6 | | 4 | −83.4 (3%) | −87.0 (3%) |



acid, and vitamin $D_3$ to address the existence of another site for vitamin D binding. This data demonstrated that the maximal binding stoichiometry of vitamin $D_3$ with LG was 2:1, whereas that of retinol or palmitic acid was 1:1 (Table 1),[26] suggesting that there was another binding site for the vitamin $D_3$ molecule. Second, we manipulated the binding capability by switching off the gate of calyx at low pH (2−6)[15] to further substantiate the "two site hypothesis". As expected, the binding capability of retinol and palmitic acid diminished under this condition, but it retained the vitamin D binding with a vitamin $D_3$ to LG stoichiometry of 1:1 (Table 1).[26] We also used a strategy to denature the conformation of the calyx by heat treatment[7] and conducted the binding assay after the calyx was thermally "disrupted". Under this condition, the binding stoichiometry of vitamin $D_3$ to heated LG (100 °C for 16 min) was found to be 1:1 (Table 1).[26] These previous binding experiments[26] suggest that there is a secondary site for vitamin $D_3$ binding distinct from the calyx.

In the present study, we located the secondary vitamin $D_3$ binding site of LG, which has been a controversy among many researchers, using bioinformatic analysis to narrow the region of potential binding sites for crystallographic verification. We first utilized well-known programs (ActiveSite_Search Insight II and Q-SiteFinder) to predict the ligand-binding site in search of a possible location for vitamin D binding with the geometric or energetic criteria. The Insight II[27] is based on the size of surface cavities of a given protein without a specified ligand; it searches the location and extent of the pocket according to the geometric criteria. The Q-SiteFinder,[28] however, defines a binding pocket only by energy calculations using a methyl probe for van der Waals interactions with a given protein. We next used a more accurate docking program, GEMDOCK, which was developed in our laboratories,[29] to dock a specific ligand with a given protein based on a nonbiased search for their interactions. GEMDOCK is specifically designed for a flexible ligand that may best fit into possible binding sites. Under this condition the docked ligand conformations are generated. All three programs require an established 3D structure of a given protein. We show that Insight II and Q-SiteFinder predicted three and six major potential regions, respectively, available for any nondefined ligand binding on LG. Docking using GEMDOCK predicted six possible sites for vitamin $D_3$ binding. Following the analyses of Insight II, Q-SiteFinder, and GEMDOCK cross-docking, we identified that there were two potential secondary sites for vitamin D binding located near the C-terminal α-helical region. This led us to study the crystal structure of LG−vitamin $D_3$ complex to further identify the secondary vitamin D binding site, if any.

Furthermore, we cocrystallized the LG−vitamin $D_3$ complex which was prepared at pH 7 with a vitamin $D_3$/LG ratio of 2 according to the previous crystallographic study.[20] A search for an extra electron density for vitamin $D_3$ around potential secondary binding sites was based on the predicted data obtained from bioinformatic analysis. We found a weak extra electron density that was located near the C-terminal α-helical region. With respect to cocrystallization, the maximum occupancy of the ligand should provide a better opportunity in growing high-quality crystals of the ligand−protein complex. In general, the affinity, solubility, and concentrations of added hydrophobic ligands would influence the ligand occupancy at equilibrium. For 90% occupancy, the amount of added ligands must be greater than the amount of protein so that the free ligands at equilibrium are not depleted.[30] In this work, we reported a rationally designed approach for preparing the complex of LG and vitamin $D_3$ at various pH and vitamin $D_3$/LG ratios to optimize the occupancy of vitamin $D_3$ and improve the electron density of the secondary binding site. Finally, we identified an exosite for vitamin D binding to be located near the α-helix and $\beta$-strand I of LG using a crystal prepared at pH 8 with a vitamin $D_3$/LG ratio of 3:1. The biological significance of the revealed exosite for vitamin D binding in milk LG was also discussed.

## 2. Experimental Section

**2.1. Materials.** LG was purified from fresh raw milk using 40% saturated ammonium-sulfate, followed by a G-150 column chromatography of the supernatant as described previously.[4] Vitamin $D_3$ (cholecalciferol) was purchased from Sigma (St. Louis, MO).

**2.2. Predicting Possible Binding Site by ActiveSite_Search Insight II and Q-SiteFinder.** To identify the possible binding sites, the ActiveSite_Search Insight II[27] and Q-SiteFinder[28] software were

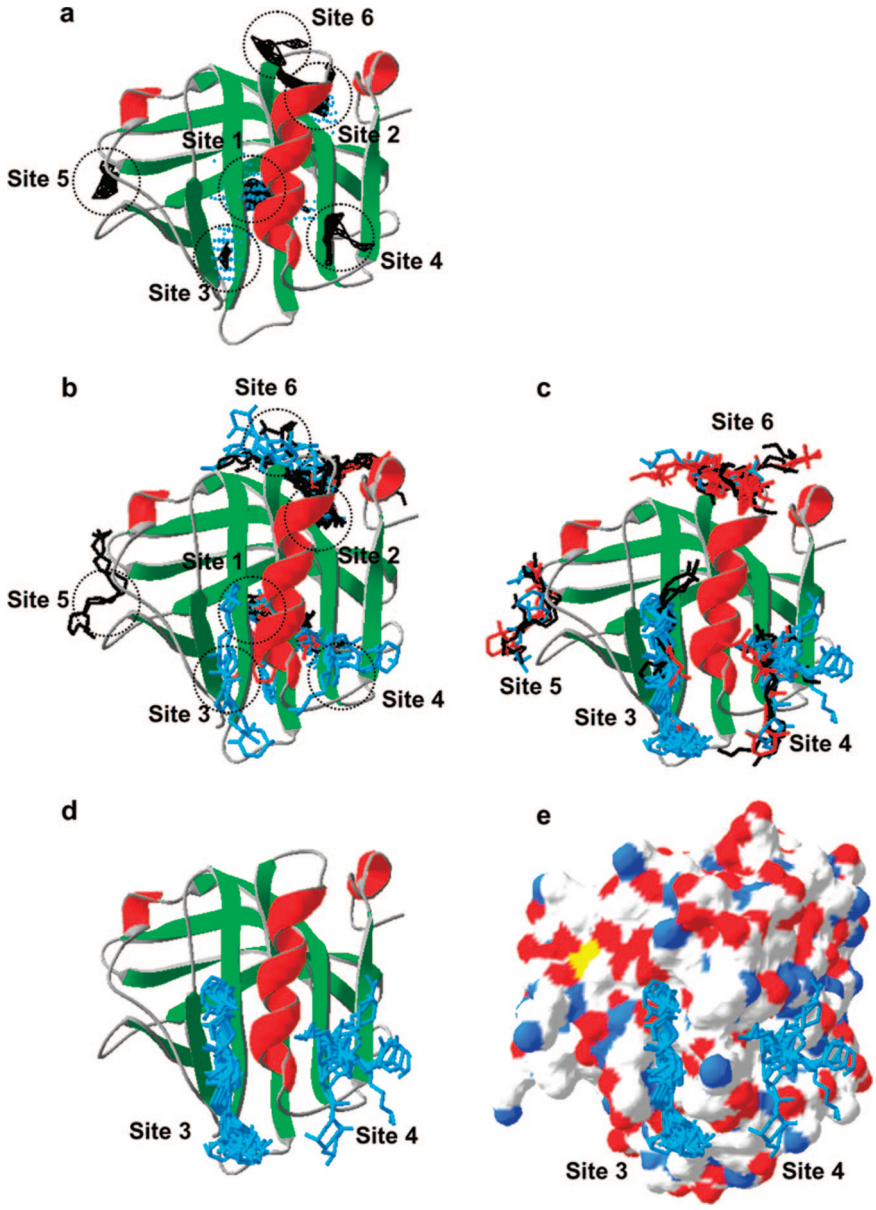




**Figure 2.** Prediction of potential binding pockets of LG using Insight II, Q-SiteFinder, and GEMDOCK. (a) Binding pockets of LG predicted by Insight II (color in blue) and Q-SiteFinder (in black). There are three and six major binding sites predicted by Insight II or Q-SiteFinder, respectively, for any nondefined ligand. (b) There are total six potential binding sites for retinol, palmitic acid, and vitamin $D_3$ (red, black, and blue) in LG following the GEMDOCK docking. LG structure in 3D used was based on LG−retinol complex (PDB code: 1GX8). Retinol and palmitic acid are able to dock exactly into the calyx (site 1), but not to the other sites except for palmitic acid interacting with site 2. The probability of vitamin $D_3$ fitting into the calyx is about 50% (see Table 1). It indicates that the calyx is a superior binding site for vitamin D. (c) Cross-docking of retinol, palmitic acid, and vitamin $D_3$ into the cavity of LG with sites 1 and 2 blocked. The docked orientation of the vitamin $D_3$ in sites 3 and 4 are distinctly different from retinol and palmitic acid. (d) Comparison of the docked orientation of vitamin $D_3$ between sites 3 and 4. The docked conformation of vitamin $D_3$ binding in site 3 seems to be more stable than in site 4. (e) The electrostatic surface model of LG depicting the positively charged (in blue), negatively charged (in red), and hydrophobic (white) side chains, in which the hydrophobic surface of site 4 seems to be large relative to that of site 3.

**Table 3. Percentage Probability of Different Ligands Docking into Possible Sites of LG**

| conditions | docking ligands into 1GX8 | | | | | |
|---|---|---|---|---|---|---|
| | docking search on whole protein | | | docking search on other parts of protein | | |
| fitness (ratio) | retinol | palmitic acid | vitamin $D_3$ | retinol | palmitic acid | vitamin $D_3$ |
| Site 1 (calyx) | −86.3 (93%) | −75.7 (43%) | −87.2 (50%) | | | |
| Site 2 | −86.4 (6%) | −77.6 (46%) | −88.4 (20%) | | | |
| Site 3 | | | −85.4 (6%) | −76.2 (3%) | −69.7 (10%) | −83.9 (40%) |
| Site 4 | | | −82.4 (6%) | −73.3 (10%) | −69.4 (16%) | −81.2 (20%) |
| Site 5 | | −76.1 (6%) | −75.4 (3%) | −72.9 (6%) | −67.7 (10%) | −80.1 (3%) |
| Site 6 | | | −83.4 (3%) | −77.2 (43%) | −72.5 (13%) | −87.0 (3%) |

used, and the prediction was performed according to the standard protocol. For ActiveSite_Search Insight II, we used the size within 50 Å as a cutoff site for the smallest cavity (in grid points). For Q-SiteFinder, the parameters for identified protein residues involved

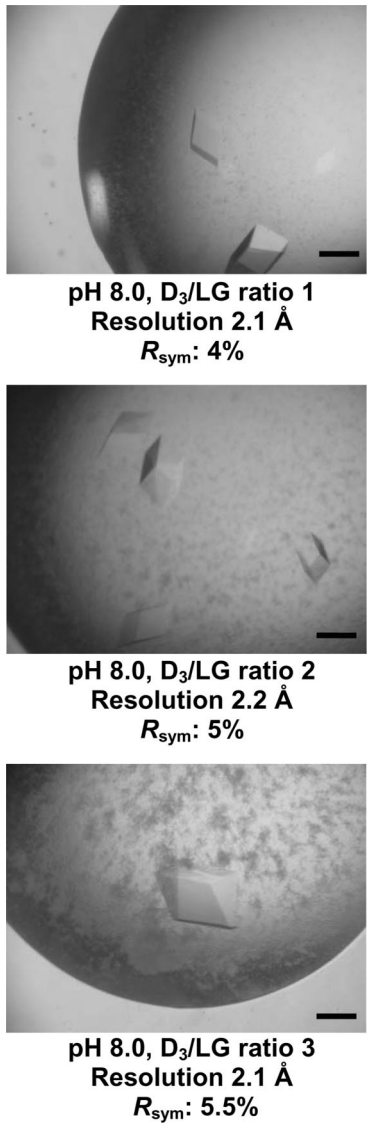


**Figure 3.** Typical example of LG−vitamin $D_3$ crystals prepared with various vitamin $D_3$/LG ratios at pH 8. Crystals of LG−vitamin $D_3$ complexes were grown using a hanging-drop vapor-diffusion method. Complex crystals of rhombohedral shape prepared at vitamin $D_3$/LG ratios of 1, 2, or 3 appeared at day 11, 10, or 6 with 2.1, 2.2, or 2.1 Å resolution diffraction were observed, respectively. The crystals grew slowly to a dimension of 0.2−0.4 mm in 18 days. Each bar represents 0.2 mm.

in the van der Waals interactions with the methyl probe, 5.0 Å was used. The coordinates of possible sites predicted by Insight II and Q-SiteFinder were saved in PDB format and depicted using the Swiss-Pdb Viewer program.[31]

**2.3. Docking Analysis between Ligands and LG by GEMDOCK.** We extrapolated the 3D structure of retinol (PDB ID 1GX8, LG-retinol complex), palmitic acid (1GXA, LG-palmitic acid complex), and vitamin $D_3$ (modified from 25(OH)-vitamin $D_3$; 1MZ9, cartilage oligomeric matrix protein−vitamin $D_3$) from the Protein Data Bank (PDB). Energy minimization of these three compounds was performed using a SYBYL program package. An established 3D structure of LG (1GX8) was used as the target for cross-docking. Since GEMDOCK is able to search the whole protein for exploring the binding site of a given ligand, we presumed those atoms near the charged surface and hydrophobic areas of LG were the potential binding sites for vitamin $D_3$. The predicted area(s) was then compared with the established binding pocket (calyx) for palmitic acid and retinol. Following the removal of all structured water molecules in LG molecule, GEMDOCK search was then conducted according to the procedures previously described.[29] Using an empirical energy function which consists of the electrostatic, steric, and hydrogen bonding potentials for docking, GEMDOCK[29] seemed to be more accurate than some conventional approaches, such as GOLD and FlexX, based on a diverse data set of 100 protein−ligand complexes proposed by Jones et al.[32] The accuracy of GEMDOCK was demonstrated when screening the ligand database for antagonist and agonist ligands of the estrogen receptor (ER).[33]

In this work, we used an empirical scoring function to estimate interaction energies between LG and ligands. The parameters used in the flexible docking included the initial step size ($\sigma = 0.8$ and $\psi = 0.2$), family competition length ($L = 2$), population size ($N = 1000$), and recombination probability ($p_c = 0.3$). For each docked ligand, optimization was stopped when either the convergence was below a certain threshold value or the iterations exceeded the maximal preset value of 80. Therefore, GEMDOCK produced 3600 solutions in one generation and terminated when it exhausted to 324 000 solutions for each docked ligand.

**2.4. Crystallization.** Purified LG was concentrated to 20 mg/mL in 20 mM acetate (pH 4), cacodylate (pH 6), HEPES (pH 7), or Tris buffer (pH 8). Vitamin $D_3$ stock solution prepared as 50 mM in 100% ethanol was added to LG solution to give a molar ratio of 3:1, 2:1, or 1:1 with a final ethanol concentration less than 7% and incubation for three hours at 37 °C. Crystallization of the LG−vitamin $D_3$ complex was achieved using the hanging-drop vapor-diffusion method at 18 °C with 2 $\mu$L hanging drops containing equal amounts of LG−vitamin $D_3$ complex and a reservoir solution (0.1 M HEPES containing 1.4 M trisodium citrate dehydrate, pH 7.5).

**2.5. Crystallographic Data Collection and Processing.** The crystals were mounted on a Cryoloop (0.1−0.2 mm), dipped briefly in 20% glycerol as a cryoprotectant solution, and frozen in liquid nitrogen. X-ray diffraction data at 2.1−2.2 Å resolution were collected at 110 K using synchrotron radiation on the Taiwan contracted beamlines BL12B2 at SPring-8 (Harima, Japan) and BL13B at NSRRC (Hsinchu, Taiwan). The data were indexed and processed using a *HKL2000* program.[34]

## 3. Results and Discussion

Although a secondary vitamin D binding site of LG has been proposed from some physicochemical experiments,[23−25,35,36] its existence and location have remained elusive and controversial. Several studies have clearly demonstrated that the binding stoichiometry between retinol or palmitic acid and LG is 1 where the central calyx of LG is responsible for retinol and palmitic acid binding,[21−23] but whether the binding of vitamin D to LG is 1 or 2 remains uncertain. Wang et al. proposed that LG possesses two potential binding sites for vitamin D: one is in the calyx formed by a $\beta$-barrel and the other is near an external hydrophobic pocket between the α-helix and the $\beta$-barrel.[24,25] In the previous study, we first showed that the relative binding of vitamin D to LG was 56% of the maximal binding at a vitamin $D_3$/LG ratio of 1:1 and the binding stoichiometry of vitamin $D_3$ to LG was 2:1 relative to that 1:1 of retinol or palmitic acid verified using extrinsic fluorescence emission, fluorescence enhancement and quenching methods (Table 1).[26] Our previous data tended to support the view that LG comprises two vitamin D binding sites. Second, we monitored the vitamin $D_3$ binding of LG by utilizing a unique structural change property of LG at various pH to further substantiate the "two site hypothesis". The EF loop of LG is known to act as a gate over the calyx;[15] at pH values lower than 6 the loop is in a "closed" position. Of remarkable interest, LG still retained about 30% of the maximal binding for vitamin $D_3$ at pH below the transition (2−6) (Table 1).[26] It suggested that there might be another vitamin D binding site that is independent of pH. Our previous study has shown that thermally denatured LG (heated 100 °C for 5 min) is unable to bind to retinol and palmitic acid owing to the deterioration of the calyx.[7] We tested the hypothesis whether the "secondary binding site" for vitamin D (if any) still existed after heating. Since LG is a molten globule with a

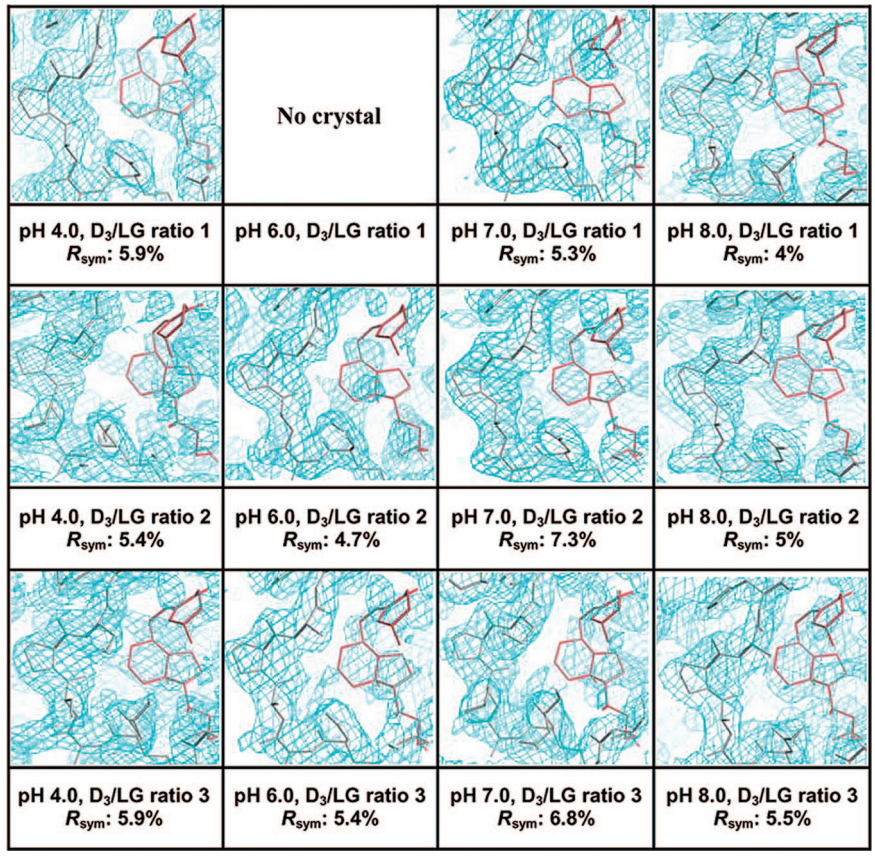


**Figure 4.** Initial electron densities map around the exosite of LG−vitamin $D_3$ complex prepared at indicated conditions. The $|2F_{obs} - F_{calc}|$ electron density maps generated by the initial LG model with omitted vitamin $D_3$ show that the electron density (in cyan) of vitamin $D_3$ molecule (in red) in the exosite is more visible when the complex was prepared at pH 8 with a vitamin $D_3$/LG ratio of 3:1. The images were generated using the program O.[46]

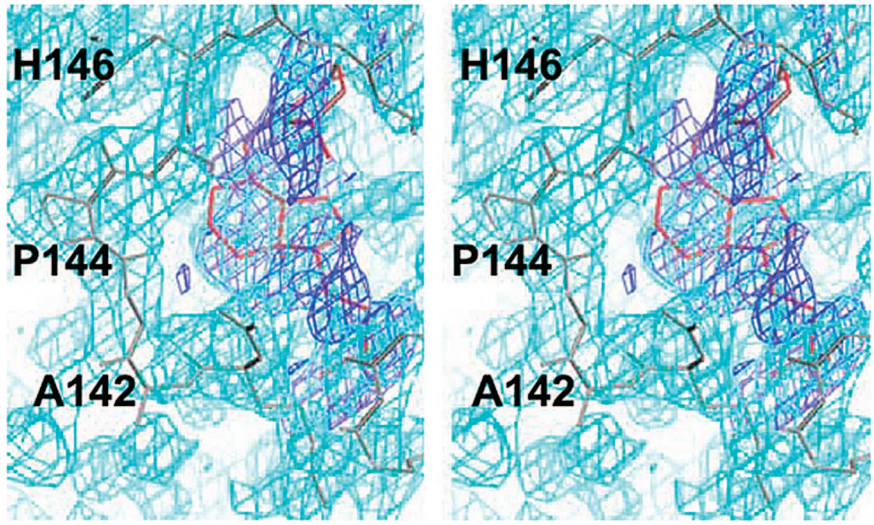


**Figure 5.** The electron density maps around the exosite of LG-vitamin $D_3$ complex prepared at pH 8 with an initial vitamin $D_3$/LG ratio of 3:1 in stereoview. The $|F_{obs} - F_{calc}|$ (in blue) and $|2F_{obs} - F_{calc}|$ (in cyan) electron density maps generated by the initial LG model with vitamin $D_3$ omitted show that the electron density of the vitamin $D_3$ molecule (in red) in the exosite is clearly visible.

heating transition between 70 and 80 °C,[5,7] we monitored the vitamin $D_3$ binding with LG preheated at different temperatures including one that could denature the calyx structure. Interestingly, it showed a dramatic and sharp decrease in retinol, palmitic acid, and vitamin $D_3$ binding near the LG transition temperature. At the temperature above 80 °C, it almost completely abolished the retinol and palmitic acid binding, while still retaining 40% of the vitamin $D_3$ binding even at 100 °C heating for 16 min (Table 1).[26] To confirm it further, we monitored the binding of heated LG (100 °C for 16 min) with various amounts of vitamin $D_3$, and a stoichiometry of 1:1 was observed between heat-denatured LG and vitamin $D_3$ using the same fluorescence quenching analysis.[26] Taking the previous pH and thermal experiments together (Table 1), we concluded that a thermally stable site (defined as an exosite) beyond the calyx exists for vitamin $D_3$ binding. In the present study, we

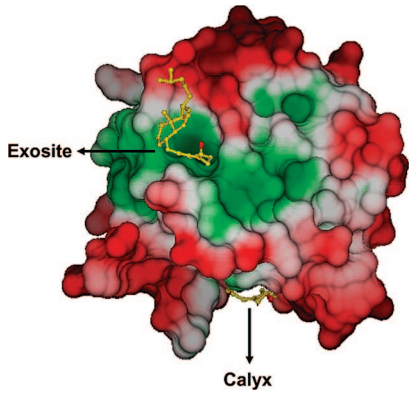


**Figure 6.** Binding characteristics of vitamin $D_3$ in the exosite of LG. The surface model of vitamin $D_3$ interacting LG with the hydrophobic (green) or hydrophilic region (red) of LG is displayed. The data reveal that the exosite of LG provides the hydrophobic force to bind vitamin $D_3$ stably. Notably, the 3-OH group (red) of vitamin $D_3$ sticks out from the pocket.



located the secondary vitamin $D_3$ binding site of LG using bioinformatic analysis to narrow the region of potential binding sites, the search of extra electron density around the potential binding sites of the cocrystal prepared on previous crystallographic study,[20] and a rationally designed crystallographic approach for obtaining cocrystals with sufficient quality and ligand occupancy.

**3.1. Structural Analysis of the Binding Pockets of LG Using Insight II and Q-SiteFinder.** We attempted to identify the regions of LG that might be available for the interaction with any nondefined molecules using well-known programs (Insight II[27] and Q-SiteFinder[28]), which are based on the size of the binding pocket and the binding energy between the methyl group and the pocket of a given protein, respectively, for prediction of a ligand binding site. The former and latter predicted that there were 3 and 10 possible "binding sites" for any ligand, respectively. Table 2 and Figure 2a show the ranking and the location of six major predicted sites. Essentially both programs predicted that site **1** is exactly located at the calyx with the ranking superior to any others. This site is also known for all the vitamin D, retinol, and palmitic acid binding. Location of other predicted sites **2** (among C-terminal loop, $\beta$-strands C and D), **3** (among the pocket C-terminal α-helix, $\beta$-strands F, G, H, and A), **4** (at the side α-helix and $\beta$-strands I), **5** (near loop H), and **6** (between CD and DE loops) is depicted in Figure 2a.

**3.2. Docking Analysis of the Interaction between LG and Retinol or Palmitic acid Using GEMDOCK.** Because both Insight II and Q-SiteFinder are not able to perform the specific interaction between an assigned ligand and a given binding site, we next used a more accurate docking program, GEMDOCK,[29] for molecular docking to search for a possible number of vitamin D binding sites and to further assess the interaction of a specific ligand with each site. GEMDOCK appeared to be more accurate than comparative approaches, such as GOLD and FlexX, on a diverse data set of 100 protein−ligand complexes[32] for docking and two cross-docking experimental sets.[29] The screening accuracies of GEMDOCK were also better than GOLD, FlexX, and DOCK on screening the ligand database for antagonist and agonist ligands of estrogen receptor (ER).[33] GEMDOCK is a better docking tool for assessing the interaction of a specific ligand with each site of LG. We generated 30 docked conformations of each retinol or palmitic acid with whole LG. Based on the docking protocol,[29] we considered that a correct binding-mode was reproducible when the root-mean-square deviation (rmsd) between the best energy-scored conformation and crystal coordinates was less than 2 Å.[32] In order to meet this criterion, we then evaluated the feasibility of GEMDOCK by cross-docking retinol and palmitic acid into the calyx or site 1 of LG (PDB code: 1GX8), known to be an established site for the binding of retinol or palmitic acid. Our data reveal that these two ligands could be docked exactly into the calyx with a rmsd less than 2.0 Å. The average fitness of docked energy of retinol and palmitic acid was −86.3 and −75.7 kcal/mol (Table 3), respectively. Following the analyses of all the docking data, we found that none of the other sites 2−6 were fitted by either retinol or palmitic acid (<10% probability), except for palmitic acid interacting with site 2 (46% of total 30-docked conformations) (Table 3). For example, none of the docked conformations of retinol and palmitic acid could fit into sites 3, 4, 5, and 6. This result is almost consistent with the current knowledge that the calyx (site 1) is the major pocket for interacting with these two ligands. It is thus feasible to use a similar strategy for studying the interaction between vitamin $D_3$ and LG.

**3.3. Docking Analysis of the Interaction between LG and Vitamin $D_3$ Using GEMDOCK.** Each docked conformation of vitamin $D_3$ ($n = 30$) obtained from GEMDOCK was then evaluated for a possible vitamin D binding ability. Table 2 shows that vitamin $D_3$ could fit into the calyx with about 50% of the probability. Interestingly, the docked energy of vitamin $D_3$ into the calyx site 1 was similar to that of retinol. The result suggests that GEMDOCK is suitable in conducting interactions between vitamin $D_3$ and LG. It also pointed out site 2 (Figure 2b) as a possible second binding site for palmitic acid (46%) and vitamin $D_3$ (20%). However, since the binding stoichiometry between palmitic acid and LG is 1 (Table 1) and the calyx is the only available binding site for palmitic acid;[21] we ruled out site 2 as a particular site for palmitic acid binding. This site is located among the C-terminal loop, $\beta$-strand C, and $\beta$-strand D. Since $\beta$-strand D is thermally unstable,[7] it is not consistent to the proposed secondary vitamin D binding site that is supposed to be thermally stable in nature (Table 1). Site 2 is thus unlikely for vitamin D interaction.

In order to search for a potential secondary vitamin $D_3$ binding site location, we blocked sites 1 and 2 with the reason mentioned above and then performed GEMDOCK, while using retinol and palmitic acid as a "negative control". Table 3 shows that, under this condition, sites 3 and 4 possessed high probability for vitamin $D_3$ interaction (40 and 20%) relative to retinol and palmitic acid. The other domains (sites 5 and 6) were considered nonsignificant due to the low probability (<3%). Notably, the docked orientation of the vitamin $D_3$ in sites 3 and 4 was distinctly different from that of other ligands (Figure 2c). Figure 2d tempted to suggest that the docked orientation of vitamin $D_3$ in site 3 is more stable (with a low rmsd) than that in site 4 due to its large pocket size that could stabilize ligand docking. On the other hand, the hydrophobicity of site 4 is high relative to site 3 based on the electrostatic potential surface model between LG and docked vitamin $D_3$ conformation (Figure 2e). The prediction of site 4 being involved in vitamin $D_3$ interaction is consistent with ranking predicted by Q-SiteFinder which is based on van der Waals interaction (Table 2). Taken together, sites 3 and 4 are the potential candidates for the secondary vitamin D binding site and thus led us to localize it via a high-quality LG-vitamin $D_3$ crystal. We searched for an extra density around C-terminal α-helix (sites 3 and 4) of LG−vitamin $D_3$ complex which was prepared at pH 7 with a vitamin $D_3$/LG

ratio of 2 according to the previous crystallographic study.[20] We found a weak extra electron density located near the C-terminal α-helical region. Later, we used a rationally designed crystallography experiment for improving electron density of vitamin $D_3$ of the secondary binding site. In the bioinformatic prediction of the secondary vitamin $D_3$ binding site, we excluded false-positive site 2 on the previous biochemical findings. It was confirmed that there is no extra density around site 2 in the crystal structure of LG−vitamin D complex that was prepared using the previous condition. For this reason, bioinformatic prediction excluded site 2 as a potential candidate for the secondary vitamin D binding site based on the previous biochemical findings that is proper for this study.

**3.4. Crystallization and Diffraction of LG−vitamin $D_3$ Complex Prepared at Various pH and Vitamin $D_3$/LG Ratios.** In any event, starting the cocrystallization experiment with maximum ligand occupancy should provide a better opportunity in growing high-quality ligand−protein crystals. The affinity and concentrations of added ligands, as well as ligand solubility, would influence the occupancy of ligands at equilibrium. For 90% occupancy, the amount of added ligands must be greater than the amount of protein so that free ligands at equilibrium is not depleted to less than about $10 \times K_d$.[30] In practice, ratios of ligands to a given protein up to 10:1 or more are commonly used, but large excesses should be avoided owing to the possibility of ligand binding to nonspecific sites. In the case of a weak binding affinity, the concentrations of ligands (or ligand solubility) may have to be 10 mM or higher in order to observe crystallographic occupancy.[37,38] We suspect that the lower occupancy of vitamin $D_2$ might explain an early study in which no electron density was found beyond the calyx using the crystal structure of LG−vitamin $D_2$ complex.[10] To optimize the binding and occupancy of vitamin $D_3$ to LG, we prepared this complex at various pH values ranging from 4 to 8 with a vitamin $D_3$/LG ratio ranging from 1 to 3. Cocrystallization was conducted using the hanging-drop vapor-diffusion method in the same reservoir solution. Although the complexes were prepared at various pH, all were crystallized in the same bottom reservoir at pH 7.5. It is difficult to assess the true effect of pH on maximizing ligand occupancy, but one can prepare a LG−vitamin $D_3$ complex with the maximal ligand binding for cocrystallization, and avoid precipitation of water-insoluble vitamin $D_3$ during preparation of the LG−vitamin $D_3$ complex. The LG−vitamin $D_3$ complex crystals of rhombohedral shape appeared in 6−11 days and continued to grow slowly to a dimension of 0.2−0.4 mm in 18 days (Figure 3). In general, crystal could form except for the condition when prepared at pH 6 with a vitamin $D_3$/LG ratio of 1. Crystals grew in larger quantity when the complex was prepared at pH 8 with a vitamin $D_3$/LG ratio of 3 (data not shown). Although some precipitation was observed while the ratio of vitamin $D_3$/LG increased, this was probably due to an excess of water-insoluble vitamin $D_3$. Analysis of the diffraction pattern indicated that the crystals exhibited trigonal symmetry and systematic absences suggested that the space group was $P3_221$, with unit-cell parameters $a = b = 53.78$ Å and $c = 111.57$ Å.

**3.5. Initial Electron Density for Secondary Vitamin $D_3$ Binding Site.** The structures of the LG−vitamin $D_3$ complexes prepared in various conditions were determined by molecular replacement[39] as implemented in *CNS v1.1*[40] using the crystal structure of bovine LG (PDB code 2BLG)[15] as a search model. The LG molecule was located in the asymmetric unit after the rotation and translation function searches. The initial electron density maps of complexes prepared at various conditions were generated from the molecular-replacement phase. One elongated extra electron density in the calyx was clearly visible. We then searched the extra electron density around the α-helix, which was predicted by bioinformatic programs, to explore the existence and location of the exosite. We found one extra electron density located near the α-helix and β-strand I as "site 4" which was predicted by bioinformatic programs (Figure 4). Interestingly, the electron densities around the exosite of LG−vitamin $D_3$ crystals prepared at pH 8 were the best defined (Figure 4). Figure 4 reveals the greater the vitamin $D_3$/LG ratio prepared for crystallization at pH 8 the more visible electron densities were around the exosite. The result indicates that the occupancy of vitamin $D_3$ in complexes is an essential requirement for exploring the location of a secondary vitamin D binding site. Stereoviews of the $|F_{obs} - F_{calc}|$ and $|2F_{obs} - F_{calc}|$ electron density maps of LG−vitamin $D_3$ complex prepared at pH 8 with a vitamin $D_3$/LG ratio 3:1 were generated by the initial model with omitted vitamin $D_3$ (Figure 5). The electron densities around vitamin $D_3$ in the exosite were all more visible in the $|F_{obs} - F_{calc}|$ and $|2F_{obs} - F_{calc}|$ electron density maps (Figure 5). This result supported the notion that the secondary vitamin D binding site exists and is located between α-helix and β-strand I. The final refined structure determined using the LG−vitamin $D_3$ complex prepared at pH 8 with a vitamin $D_3$/LG ratio of 3:1 was given in detail and discussed elsewhere.[26] The surface model of the LG−vitamin $D_3$ complex reveals that vitamin $D_3$ almost perpendicularly inserts into the calyx cavity, while the other vitamin $D_3$ is located near the carboxyl terminus of LG (residues 136−149 including part of the α-helix and β-strand I) (Figure 6). It shows the exosite containing most hydrophobic residues fitted parallel with one vitamin $D_3$ molecule (Figure 6).

In the present study, with bioinformatic analysis we were able to suggest two sites (sites 3 and 4) around the C-terminal α-helix as potential binding sites on LG. Subsequently, we searched for an extra density around the C-terminal α-helix (sites 3 and 4) in the electron density map of the cocrystal prepared in a previous crystallographic study.[20] A weak electron density around site 4 (located between α-helix and β-strand I) of LG−vitamin $D_3$ complex was found. The rationally designed crystallographic approach resulted in obtaining cocrystals with sufficient quality and ligand occupancy so as to allow experimental validation of the second vitamin $D_3$ binding site on LG. We believe that our strategy is useful for others in the search for a ligand-binding site on a given protein.

**3.6. Biological Significance.** Recent studies indicate that increased plasma vitamin $D_3$ concentrations are associated with decreased incidence of cancer[41] and osteoporotic fractures.[42] Vitamin D is found in only a few foods, such as fish oil, liver, milk, and eggs. As the level of vitamin D in bovine milk has been reported to be low, dairy products are fortified with vitamin $D_3$ to a level of about 0.35 μM in many sophisticated food industries.[43] Therefore, the vitamin D-fortified milk is a major source for vitamin D in the diet. Because intact LG is acid resistant with a superpermeability to cross the epithelium cells of the gastrointestinal tract via a receptor mediated process,[44] this unique property of LG is worthy of consideration for transport of vitamin D in milk. There are two advantages for the presence of an exosite for vitamin $D_3$ binding. First, the central calyx of LG is primarily occupied by the fatty acids in milk.[45] Second, many dairy products are currently processed under excessive heat treatment for the purpose of sterilization. The presence of a

thermally stable exosite may provide another route for transporting vitamin D.

## 4. Conclusion

We concluded that there is a heat-resistant exosite beyond the calyx responsible for vitamin $D_3$ binding in a previous binding study. GEMDOCK is a helpful tool to localize and search for an extra density map around a possible secondary vitamin D binding site. Both pH and the initial ratio of vitamin $D_3$/LG are crucial to optimize the occupancy and enhance the electron density of vitamin $D_3$ in the complex for rational-designed crystallization. This strategy may be useful for future identification of a ligand-binding site in a given protein.

**Acknowledgment.** We are grateful to Dr. Yuch-Cheng Jean and the supporting staff for technical assistance and the discussion of the synchrotron radiation X-ray data collected at BL13B1 of NSRRC, Taiwan, and Dr. Jeyaraman Jeyakanthan at BL12B2 of SPring-8, Japan. We also thank Mr. James Lee for the critical editing. This work was supported in part by the National Science Council NSC 92-2313-B-009-002, 93-2313-B009-002, 94-2313-B-009-001 & 95-2313-B-009-001 to S. J. T. Mao and NSC 94-2321-B-213-001, 95-2321-B-213-001-M and the National Synchrotron Radiation Research Center 944RSB02 & 954RSB02 to C.-J. Chen, Taiwan, ROC.



## References

(1) Hambling, S. G.; MacAlpine, A. S.; Sawyer, L. In *Advanced Dairy Chemistry*; Fox, P. F. Eds.; Elsevier: Amsterdam, 1992; pp 141−190.
(2) Qi, X. L.; Brownlow, S.; Holt, C.; Sellers, P. Thermal denaturation of beta-lactoglobulin: effect of protein concentration at pH 6.75 and 8.05. *Biochim. Biophys. Acta* **1995**, *1248*, 43–49.
(3) Sawyer, L.; Kontopidis, G. The core lipocalin, bovine beta-lactoglobulin. *Biochim. Biophys. Acta* **2000**, *1482*, 136–148.
(4) Chen, W. L.; Huang, M. T.; Liu, H. C.; Li, C. W.; Mao, S. J. T. Distinction between dry and raw milk using monoclonal antibodies prepared against dry milk proteins. *J. Dairy Sci.* **2004**, *87*, 2720–2729.
(5) Chen, W. L.; Huang, M. T.; Liau, C. Y.; Ho, J. C.; Hong, K. C.; Mao, S. J. T. $\beta$-Lactoglobulin is a thermal marker in processed milk as studied by electrophoresis and circular dichroic spectra. *J. Dairy Sci.* **2005**, *88*, 1618–1630.
(6) Chen, W. L.; Liu, W. T.; Yang, M. C.; Hwang, M. T.; Tsao, J. H.; Mao, S. J. T. A novel conformation-dependent monoclonal antibody specific to the native structure of beta-lactoglobulin and its application. *J. Dairy Sci.* **2006**, *89*, 912–921.
(7) Song, C. Y.; Chen, W. L.; Yang, M. C.; Huang, J. P.; Mao, S. J. T. Epitope mapping of a monoclonal antibody specific to bovine dry milk: involvement of residues 66−76 of strand D in thermal denatured beta-lactoglobulin. *J. Biol. Chem.* **2005**, *280*, 3574–3582.
(8) Nagaoka, S.; Futamura, Y.; Miwa, K.; Awano, T.; Yamauchi, K.; Kanamaru, Y.; Tadashi, K.; Kuwata, T. Identification of novel hypocholesterolemic peptides derived from bovine milk beta-lactoglobulin. *Biochem. Biophys. Res. Commun.* **2001**, *281*, 11–17.
(9) Zsila, F.; Bikadi, Z.; Simonyi, M. Retinoic acid binding properties of the lipocalin member beta-lactoglobulin studied by circular dichroism, electronic absorption spectroscopy and molecular modeling methods. *Biochem. Pharmacol.* **2002**, *64*, 1651–1660.
(10) Kontopidis, G.; Holt, C.; Sawyer, L. Invited review: beta-lactoglobulin: binding properties, structure, and function. *J. Dairy Sci.* **2004**, *87*, 785–796.
(11) Salvi, A.; Carrupt, P.; Tillement, J.; Testa, B. Structural damage to proteins caused by free radicals: asessment, protection by antioxidants, and influence of protein binding. *Biochem. Pharmacol.* **2001**, *61*, 1237–1242.
(12) Chevalier, F.; Chobert, J. M.; Genot, C.; Haertle, T. Scavenging of free radicals, antimicrobial, and cytotoxic activities of the Maillard reaction products of beta-lactoglobulin glycated with several sugars. *J. Agric. Food Chem.* **2001**, *49*, 5031–5038.
(13) Marshall, K. Therapeutic applications of whey protein. *Altern. Med. Rev.* **2004**, *9*, 136–156.
(14) Liu, H. C.; Chen, W. L.; Mao, S. J. T. Antioxidant nature of bovine milk beta-lactoglobulin. *J. Dairy Sci.* **2007**, *90*, 547–555.
(15) Qin, B. Y.; Bewley, M. C.; Creamer, L. K.; Baker, H. M.; Baker, E. N.; Jameson, G. B. Structural basis of the Tanford transition of bovine beta-lactoglobulin. *Biochemistry* **1998**, *37*, 14014–14023.
(16) Qin, B. Y.; Bewley, M. C.; Creamer, L. K.; Baker, E. N.; Jameson, G. B. Functional implications of structural differences between variants A and B of bovine beta-lactoglobulin. *Protein Sci.* **1999**, *8*, 75–83.
(17) Kuwata, K.; Hoshino, M.; Forge, V.; Era, S.; Batt, C. A.; Goto, Y. Solution structure and dynamics of bovine beta-lactoglobulin A. *Protein Sci.* **1999**, *8*, 2541–2545.
(18) Uhrinova, S.; Smith, M. H.; Jameson, G. B.; Uhrin, D.; Sawyer, L.; Barlow, P. N. Structural changes accompanying pH-induced dissociation of the beta-lactoglobulin dimer. *Biochemistry* **2000**, *39*, 3565–3574.
(19) Narayan, M.; Berliner, L. J. Fatty acids and retinoids bind independently and simultaneously to beta-lactoglobulin. *Biochemistry* **1997**, *36*, 1906–1911.
(20) Qin, B. Y.; Creamer, L. K.; Baker, E. N.; Jameson, G. B. 12-Bromododecanoic acid binds inside the calyx of bovine beta-lactoglobulin. *FEBS Lett.* **1998**, *438*, 272–278.
(21) Wu, S. Y.; Perez, M. D.; Puyol, P.; Sawyer, L. beta-lactoglobulin binds palmitate within its central cavity. *J. Biol. Chem.* **1999**, *274*, 170–174.
(22) Kontopidis, G.; Holt, C.; Sawyer, L. The ligand-binding site of bovine beta-lactoglobulin: evidence for a function. *J. Mol. Biol.* **2002**, *318*, 1043–1055.
(23) Wang, Q.; Allen, J. C.; Swaisgood, H. E. Binding of retinoids to beta-lactoglobulin isolated by bioselective adsorption. *J. Dairy Sci.* **1997**, *80*, 1047–1053.
(24) Wang, Q.; Allen, J. C.; Swaisgood, H. E. Binding of vitamin D and cholesterol to beta-lactoglobulin. *J. Dairy Sci.* **1997**, *80*, 1054–1059.
(25) Wang, Q.; Allen, J. C.; Swaisgood, H. E. Binding of lipophilic nutrients to beta-lactoglobulin prepared by bioselective adsorption. *J. Dairy Sci.* **1999**, *82*, 257–264.
(26) Yang, M. C.; Guan, H. H.; Liu, M. Y.; Lin, Y. H.; Yang, J. M.; Chen, W. L.; Chen, C. J.; Mao, S. J. T. Crystal structure of a secondary vitamin $D_3$ binding site of milk beta-lactoglobulin. *Proteins* **2008**, *71*, 1197–1210.
(27) Xiao, L.; Cui, X.; Madison, V.; White, R. E.; Cheng, K. C. Insights from a three-dimensional model into ligand binding to constitutive active receptor. *Drug Metab. Dispos.* **2002**, *30*, 951–956.
(28) Laurie, A. T.; Jackson, R. M. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **2005**, *21*, 1908–1916.
(29) Yang, J. M.; Chen, C. C. GEMDOCK: a generic evolutionary method for molecular docking. *Proteins* **2004**, *55*, 288–304.
(30) Danley, D. E. Crystallization to obtain protein-ligand complexes for structure-aided drug design. *Acta Crystallogr.* **2006**, *D62*, 569–575.
(31) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714–2723.
(32) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(33) Yang, J. M.; Shen, T. W. A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. *Proteins* **2005**, *59*, 205–220.
(34) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326.
(35) Ragona, L.; Fogolari, F.; Zetta, L.; Pérez, D. M.; Puyol, P.; De Kruif, K.; Löhr, F.; Rüterjans, H.; Molinari, H. Bovine beta-lactoglobulin: interaction studies with palmitic acid. *Protein Sci.* **2000**, *9*, 1347–1356.
(36) Muresan, S.; van der Bent, A.; de Wolf, F. A. Interaction of beta-lactoglobulin with small hydrophobic ligands as monitored by fluorometry and equilibrium dialysis: nonlinear quenching effects related to protein-protein association. *J. Agric. Food Chem.* **2001**, *49*, 2609–2618.
(37) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* **2005**, *48*, 403–413.
(38) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **2000**, *18*, 1105–1108.

(39) Rossmann, M. G. The molecular replacement method. *Acta Crystallogr.* **1990**, *A46*, 73–82.

(40) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **1998**, *D54*, 905–921.

(41) Lappe, J. M.; Travers-Gustafson, D.; Davies, K. M.; Recker, R. R; Heaney, R. P. Vitamin D and calcium supplementation reduces cancer risk: results of a randomized trial. *Am. J. Clin. Nutr.* **2007**, *85*, 1586–1591.

(42) Bischoff-Ferrari, H. A.; Willett, W. C.; Wong, J. B.; Giovannucci, E.; Dietrich, T.; Dawson-Hughes, B. Fracture prevention with vitamin D supplementation: a meta-analysis of randomized controlled trials. *JAMA* **2005**, *293*, 2257–2264.

(43) Holick, M. F.; Shao, Q.; Liu, W. W.; Chen, T. C. The vitamin D content of fortified milk and infant formula. *N. Engl. J. Med.* **1992**, *326*, 1178–1181.

(44) Fluckinger, M.; Merschak, P.; Hermann, M.; Haertlé, T.; Redl, B. Lipocalin-interacting-membrane-receptor (LIMR) mediates cellular internalization of beta-lactoglobulin. *Biochim. Biophys. Acta* **2008**, *1778*, 342–347.

(45) Pérez, M. D.; Díaz de Villegas, C.; Sánchez, L.; Aranda, P.; Ena, J. M.; Calvo, M. Interaction of fatty acids with beta-lactoglobulin and albumin from ruminant milk. *J. Biochem.* **1989**, *106*, 1094–1097.

(46) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **1991**, *A47*, 110–119.

CG800697S