Structural Basis of Functional Conversion of a Floral Repressor to an Activator: A Molecular Dynamics Simulation Study

Sukki Kang,[a] Juyong Lee,[a] Myeong Sup Lee,[†] and Chaok Seok[*]

Department of Chemistry, College of Natural Sciences, Seoul National University, Seoul 151-747, Korea
[*]E-mail: chaok@snu.ac.kr
[†]Department of Biochemistry and Genome Regulation Center, Yonsei University, Seoul 120-749, Korea
Received December 15, 2007

FLOWERING LOCUS T (FT) and *TERMINAL FLOWER 1* (TFL1) in *Arabidopsis* are homologous proteins that perform opposite functions: FT is an activator of flowering, and TFL1 is a repressor. It was shown before that change of a single amino acid (His88) of TFL1 to the corresponding amino acid (Tyr) of FT is enough to convert the floral repressor to an activator. However, structural basis of the functional conversion has not been understood. In our molecular dynamics simulations on modified TFL1 proteins, a hydrogen bond present in native TFL1 between the His88 residue and a residue (Asp144) in a neighboring external loop became broken by change of His88 to Tyr. This breakage induced conformational change of the external loop whose structure was previously reported to be another key functional determinant. These findings reveal that the two important factors determining the functional specificities of the floral regulators, the key amino acid (His88) and the external loop, are correlated, and the key amino acid determines the functional specificity indirectly by affecting the conformation of the external loop.

Key Words : *FLOWERING LOCUS T, TERMINAL FLOWER 1*, Floral regulators, Molecular dynamics simulation, Hydrogen bond

Introduction

In the life cycle of plants, switching from vegetative phase to flowering is a crucial developmental change. *FLOWERING LOCUS T* (FT) and *TERMINAL FLOWER 1* (TFL1) are two key regulators of flowering in *Arabidopsis*, a model organism in plant biology.[1-3] They are homologous to phosphatidylethanolamine-binding proteins (PEBPs),[4] a wide family of proteins performing diverse roles in various organisms: inhibition of Raf kinase and thrombin activities in mammals and action as scaffolds for signaling complexes in plants, for example.[5] Although homologous, FT and TFL1 perform opposite functions: FT promotes flowering strongly while TFL1 represses flowering.[6]

FT and TFL1 are small proteins with 175 and 177 amino acids, respectively. Recently determined crystal structures of the two proteins are very similar,[5] as expected from the relatively high sequence identity of 59%. Although the overall structures are similar, there is a significant structural difference between the external loops composed of 14 residues for FT and of 15 residues for TFL1 (See Figure 1). Interestingly, a chimeric protein of FT whose external loop is replaced with that of TFL1 represses flowering as much as native TFL1 does upon overexpression, although the chimeric protein differs from FT in only 12 amino acids.[5] Therefore, structure of the external loop seems to be crucial in determining the function of TFL1 as a floral repressor.

As well as the external loop, an amino acid (Tyr85 of FT and His88 of TFL1) near the entrance of the potential ligand-binding site was also reported to be a critical determinant of the opposite activities of FT and TFL1 (The key residues are indicated with ovals in Figure 1).[6] Change of the key amino acid (His88) of TFL1 to the corresponding amino acid (Tyr) of FT was shown to be enough to convert the floral repressor to an activator, although changes of several other amino acids in the potential ligand-binding site did not lead to functional change.[6] It was also noted before that this key amino acid (His88) of TFL1 interacts with an amino acid (Asp144) of the external loop mentioned above through a hydrogen bond,[5] but whether this interaction is related to the function of the floral regulators was unknown. To understand structural basis of the functional conversion of the floral regulators due to the single amino acid change, we performed molecular dynamics simulation studies on modified FT and TFL1 proteins. Our simulation results suggest that the change of His88 of TFL1 to Tyr induces conformational change of the neighboring external loop through breakage of the hydrogen bond. Therefore, the key amino acid (His88) may determine the function of TFL1 as a floral regulator through intimate interaction with the external loop instead of by acting independently of the external loop, while the previous experimental study on the chimeric FT with the TFL1 loop[5] implies that the external loop can determine the function of TFL1 independently of the key amino acid.

Materials and Methods

Preparation of proteins for molecular dynamics simulations. The initial atomic coordinates of the native FT and

(a)

```
FT      ---MSINIRDPLIVSRVVGDVLDPFNRSITLKVTYGQREVTNGLDLRPSQVQNKPRVEIG 57
TFL1    MENMGTRVIEPLIMGRVVGDVLDFFTPTTKMNVSYNKKQVSNGHELFPSSVSSKPRVEIH 60
        *.  .: :***:.******** *. : .::*:*.:::*:** :* **.*..******

FT      GEDLRNFYTLVMVDPDVPSPSNPHLREYLHWLVTDIPATTGTTFGNEIVCYENPSPTAGI 117
TFL1    GGDLRSFFTLVMIDPDVPGPSDPFLKEHLHWIVTNIPGTTDATFGKEVVSYELPRPSIGI 120
        * ***.*:****:*****.**:.*:*:*:***:**:**.**.:***:*:*.** * *: **

FT      HRVVFILFRQL-GRQTVYAPGWRQNFNTREFAEIYNLGLPVAAVFYNCQRESGCGGRRL  175
TFL1    HRFVFVLFRQKQRRVIFPNIPSRDHFNTRKFAVEYDLGLPVAAVFFNAQRETAARKR--  177
        **.**:****  *  .    *::****:** *:*********:*.***:.. *
```
<u>External loop</u>

(b)

Figure 1. Alignments of the amino-acid sequences and the crystal structures of FT and TFL1 proteins. (a) Alignment of the amino-acid sequences of FT and TFL1 proteins. Asterisks denote identical residues, colons conserved residues, and dots semi-conserved residues. The region of the external loop is indicated with a brace. The key amino acids (Tyr85 of FT and His88 of TFL1) near the potential ligand-binding site are marked with an oval, and the residues (Gln140 of FT and Asp144 of TFL1) in the external loop that are closely located to the key amino acids with a rectangle. (b) Superimposed crystal structures of FT (yellow) and TFL1 (magenta). The locations of the external loop and the potential ligand-binding site are shown, and the two amino acid positions are marked with the same symbols as in (a). Amino (N) and carboxy (C) termini are also labeled.

TFL1 protein structures for molecular dynamics simulations were obtained from their crystal structures deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB).[7] The PDB IDs for FT and TFL1 are 1WKP and 1WKO, respectively. Due to the missing coordinates of the terminal residues, only the residues 6-167 of FT and 7-171 of TFL1 were considered in the simulations. Initial coordinates of the modified amino acids presented in Results and Discussion were obtained using the XLEAP module of the AMBER9 program package.[8] Assignment of protonation states of His residues is critical because it can affect hydrogen bond formation. In the key residue His88 of TFL1 (and the His residue that substitutes Tyr85 of FT), both side chain nitrogen atoms, N_δ and N_ϵ, were protonated because the local environment is basic due to the closeness of the two amino acids, Asp144 and Glu112. Two favorable hydrogen bonds were formed to stabilize the protein structure in this protonation state. Moreover, Molecular Dynamics simulations with different protonation states of His88 did not lead to the conformational conversion of the external loop presented below, which implies that the use of the current protonation state is proper. All other His residues were protonated only at N_ϵ, following the convention of the AMBER9 package.

Details of molecular dynamics simulations. All the molecular dynamics simulations on the modified proteins and the native proteins were performed with the SANDER module of the AMBER9 package.[8] The PARM99 force field[9] and the generalized Born solvation model[10-12] were employed to properly account for energetics of hydrated proteins. Non-bonded interaction was cutoff at 12 Å. The SHAKE algorithm[13] that constrains the lengths of the bonds involving hydrogen atoms was used for faster simulation with 2.0 femtoseconds (fs) of integration time step. After initial relaxation with 1000 steps of energy minimization, the simulated protein system was gradually heated from 0 K to 300 K during 1.2 nanoseconds (ns) of molecular dynamics simulation. Final molecular dynamics simulation was performed for 5.0 ns at 300 K using the Brendsen thermostat.[14]

Structural changes during the simulations were analyzed with the PTRAJ module of AMBER9.[8] Average structures over the last 1 ns of the simulations were taken as the representative structures for structure comparison.

Calculation of Root-Mean-Square Deviations between the modified and native protein structures. To calculate the best-fit Root-Mean-Square Deviation (RMSD), a measure of structural difference between two proteins, the two protein structures must first be superimposed optimally. The rotation matrix for optimal superimposition was calculated using the quaternion method developed by us previously.[15,16] To calculate the rotation matrix, the alpha carbons of the residues 8-127 for FT and 11-130 for TFL1 were superimposed excluding the external loop and its C-terminal region because the crystal structures of FT and TFL1 are similar and the major structural difference occurs at the external loop. Using the calculated rotation matrix, the whole protein structures were superimposed, and then RMSD was calculated between the alpha carbon coordinates at equivalent positions. This RMSD is referred to as 'overall RMSD'. The equivalent positions were assigned as in Figure 1(a), in which the 132nd residue of TFL1 was deleted to minimize RMSD between the crystal structures of FT and TFL1. In addition to the 'overall RMSD', deviation in the external loop structures ('loop RMSD') was also calculated between the alpha carbon coordinates of the loop residues (128-141 for FT and 131-145 for TFL1).

Results and Discussion

Six modified FT and TFL1 proteins were subject to molecular dynamics simulations. To investigate the roles of the key amino acid and of the interaction of the key amino acid with the neighboring external loop in functional determination of the floral regulators, molecular dynamics simulations were performed on three different modified forms for each of the FT and TFL1 proteins constructed as follows: First, the two key amino acids, Tyr85 of FT and His88 of TFL1 (marked with ovals in Figure 1), were swapped.

Second, the amino acids, Gln140 of FT and Asp144 of TFL1 (marked with rectangles in Figure 1), located in the external loop that were shown to be positioned closely to the above key amino acids in the crystal structure,[5] were swapped. Finally, the two neighboring amino acids were swapped simultaneously. The resulting modified proteins are referred to as FT (Y85H), TFL1 (H88Y), FT (Q140D), TFL1 (D144Q), FT (Y85H, Q140D), and TFL1 (H88Y, D144Q).

Molecular dynamics simulations were carried out on the six modified and the two native proteins. as described in Materials and Methods. All the protein structures became stable after 2-3 ns of simulations (data not shown). The average deviations of the simulated native protein structures from their known crystal structures were 1.7 Å for FT and 2.7 Å for TFL1, respectively. which are within the range of deviations observed in typical molecular dynamics simulations. Therefore, the simulations can be considered reasonable. and all the simulated structures were thus analyzed systematically.

Functions of the modified floral regulators correlate with their external loop conformations. Structure of the external loop was previously shown to be an important factor determining the opposite functions of the floral regulators. FT and TFL1.[5] Therefore, the external loop and the overall structures of each of the six modified proteins were compared with those of each of the two native proteins. FT and TFL1, by computing loop RMSD and overall RMSD, as described in Materials and Methods. The results are listed in Table 1. Both overall RMSDs and loop RMSDs of the three modified FT proteins, FT (Y85H), FT (Q140D), and FT (Y85H, Q140D), from native FT are smaller than those from native TFL1, indicating that the modified FT proteins remain closer in structure to native FT than to native TFL1. The overall and the loop structures of one of the modified TFL1 proteins, TFL1 (D144Q), are also closer to those of native TFL1 than FT (Table 1). However, in the cases of the two modified TFL1 proteins. TFL1 (H88Y) and TFL1 (H88Y, D144Q), the loop structures of the modified TFL1 proteins

are closer to that of FT than TFL1 [Loop RMSDs from FT and TFL1 are 2.1 Å and 2.7 Å. respectively, for TFL1 (H88Y). and 2.3 Å and 3.0 Å. respectively, for TFL1 (H88Y. D144Q)] although the overall structures of the two modified TFL1 proteins deviate from those of FT and TFL1 to about the same degree [Overall RMSDs from FT and TFL1 are 2.5 Å and 2.5 Å, respectively. for TFL1 (H88Y), and 2.8 Å and 2.7 Å, respectively, for TFL1 (H88Y, D144Q)].

The simulation results on TFL1 (H88Y) and FT (Y85H) correlate well with the available experimental findings that (1) TFL1 (H88Y) acts as an activator like FT although native TFL1 is a repressor, and (2) FT (Y85H) acts as a weak activator in controlling flowering time[6]: (1) The external loop structure of the simulated TFL1 (H88Y) protein resembles that of native FT more than TFL1, and (2) the loop structure of the simulated FT (Y85H) protein is closer to that of native FT. It is also notable that the loop structures of the simulated proteins are closely related to the hydrogen bond interaction between the key residue (at the residue position 88 of TFL1 or 85 of FT) and a residue (at the position 144 of TFL1 or 140 of FT) in the external loop, as discussed in more detail in the next section: (1) The hydrogen bond present in TFL1 is absent in TFL1 (H88Y). just as in native FT. and (2) FT (Y85H) does not form a hydrogen bond. just like FT [See Table 1 and Figure 3]. Experimental data for functions of the other modified proteins simulated here are not available. but the current simulation results for them may be useful for predicting their functions as floral regulators because there is a good correlation between the simulation results and the experimental data for the two modified proteins, TFL1 (H88Y) and FT (Y85H).

It is intriguing that TFL1 (H88Y) has the loop structure closer to FT than TFL1 (See Table 1) although it has the same loop sequence as TFL1 and has a different loop length and sequence from FT. In particular, the structure of the segment composed of the first 5 residues (residues 131-136 for TFL1, excluding the 132nd residue) in the N-terminal region of the TFL1 (H88Y) loop is much closer to that of FT

Table 1. Overall and loop RMSDs of the modified protein structures from the native FT and TFL1 protein structures

Simulated Protein	Compared region	RMSD			Function[*]	Hydrogen bond[**]
		From native FT	From native TFL1	Difference of the RMSDs		
FT (Y85H)	Overall	1.46	2.25	0.79	Weak activator	No
	Loop	2.06	3.53	1.47		
FT (Q140D)	Overall	1.60	2.22	0.62	Not known	No
	Loop	2.36	2.63	0.27		
FT (Y85H, Q140D)	Overall	1.83	2.33	0.50	Not known	Yes
	Loop	2.84	2.92	0.08		
TFL1 (H88Y)	Overall	2.49	2.49	0.00	Activator	No
	Loop	2.12	2.74	0.62		
TFL1 (D144Q)	Overall	2.46	2.22	−0.24	Not known	Yes
	Loop	2.92	2.35	−0.57		
TFL1 (H88Y, D144Q)	Overall	2.77	2.73	−0.04	Not known	No
	Loop	2.30	3.00	0.70		

* Experimentally determined function in controlling flowering time [6].
** Existence of the hydrogen bond between the key residue (at residue position 88 of TFL1 or 85 of FT) and the residue (at position 144 of TFL1 or 140 of FT) in the external loop determined from our simulations.
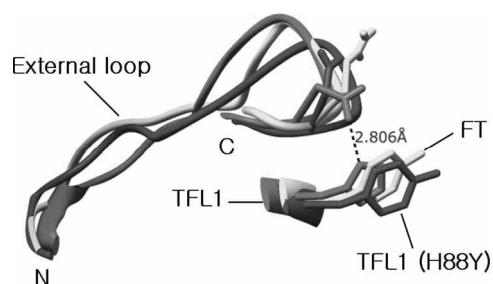
Figure 2. Structures of the external loop and the hydrogen bond in the simulated proteins. Structure of TFL1 (H88Y) (blue) is compared to those of native FT (yellow) and native TFL1 (magenta). The N- and C- termini of the external loop are labeled. A hydrogen bond is formed between the key amino acid (His88) and an amino acid (Asp144) in the external loop of native TFL1, but not between the corresponding residues of native FT or TFL1 (H88Y).

than that of TFL1: RMSD of the 5-residue segment of TFL1 (H88Y) is 1.93 Å from that of FT and 2.70 Å from that of TFL1 [See Figure 2 for comparison of the loop structures]. It is notable that the difference between the crystal structures of native FT and TFL1 in the same segment is also large (RMSD is 2.73 Å). These observations together imply that the structure of the 5-residue segment of the loop may be the key in determining the opposite functions of TFL1 (H88Y) and TFL1, as well as of native FT and TFL1.

The external loop conformations of the modified floral regulators are influenced by formation of the hydrogen bond with the key residue. To study the relationship between the hydrogen bond interaction and the conformational change of the external loop more deeply, change in the distance between the hydrogen bonding residues in the simulated proteins was monitored over simulation time (See Figure 3). The distance between the hydrogen bonding residues is defined as the distance between the center of the sidechain ring of the key residue (His or Tyr) at position 85 of FT or 88 of TFL1 and carboxyl carbon of the residue (C_γ of Asp or C_δ of Gln) at position 140 of FT or 144 of TFL1. In the cases of native TFL1 and TFL1 (D144Q), the distance between the residues stays around 5 Å throughout the simulation [Figure 3(a)], indicating existence of a hydrogen bond. However, in TFL1 (H88Y) and TFL1 (H88Y, D144Q), the distance fluctuates around 7 Å [Figure 3(a)], indicating that the hydrogen bond is broken due to the mutation of His88 of TFL1 to Tyr. Furthermore, the geometry of the two side chains of TFL1 (H88Y) is similar to that of native FT compared to that of native TFL1 [See Figure 2]. The distances between the amino acids for TFL1 (H88Y) and TFL1 (H88Y, D144Q) are shorter during the first 2 ns than during the last 3 ns. This increase in distance at later time of the simulation is correlated with the change in the loop conformation at 1.5 ns shown in Figure 4 (See discussions below), implying close connection of the hydrogen bond and the external loop structure in the interaction network of the protein structures. In native FT, FT (Y85H), and FT (Q140D), the distances between the side chains fluctuates around 7 Å, as shown in Figure 3(b), indicating no hydrogen bond between the residues. However, in FT (Y85H, Q140D) the
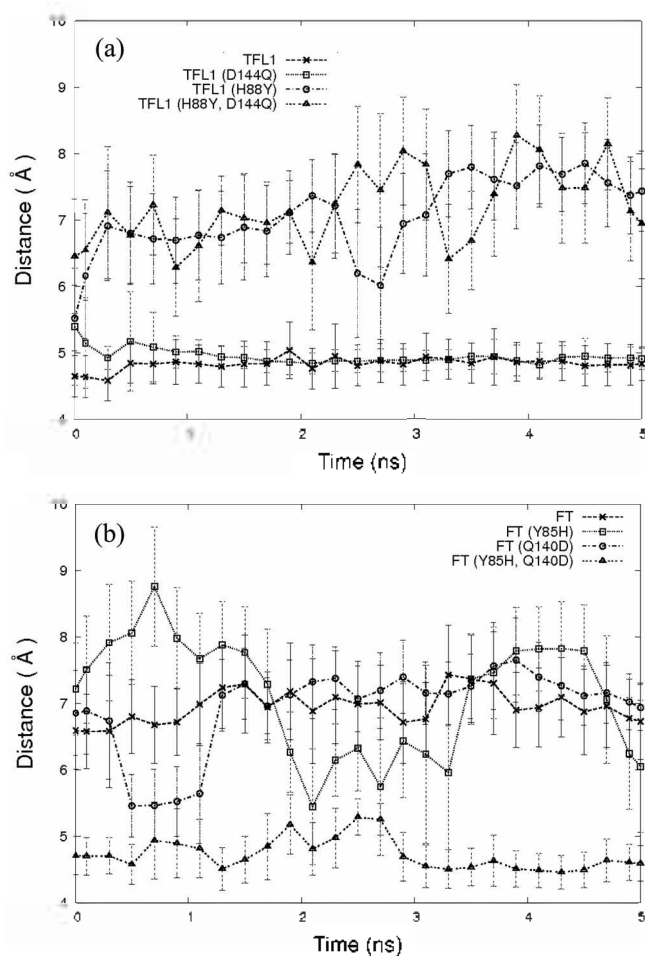

Figure 3. Change in the distance between the key amino acid and the closely located amino acid in the external loop over the simulation time. The time 0 is the point right after energy minimization and 1.2 ns of heating simulation, and thus the initial structures of the modified proteins have changed during this preparation stages after the amino acid changes. (a) Change in the distance (in Angstroms) between the amino acid at the key position 88 (His or Tyr) and that at position 144 (Asp or Gln) in the external loop over the simulation time (in nanoseconds) for the native TFL1 and the three modified TFL1 proteins, TFL1 (H88Y), TFL1 (D144Q), and TFL1 (H88Y, D144Q). (b) Change in the distance between the amino acids at position 85 (His or Tyr) and at position 140 (Asp or Gln) over the simulation time for the native FT and the three modified FT proteins, FT (Y85H), FT (Q140D), and FT (Y85H, Q140D).

distance stays around 5 Å, indicating that a hydrogen bond exists just like in TFL1. The curves for FT (Y85H) and FT (Q140D) show interesting fluctuations, indicating that the amino acids temporarily come closer to form a hydrogen bond involving the sidechain oxygen atom of Gln instead of Asp [FT (Y85H)] or the polar hydrogen atom of the sidechain of Tyr instead of His [FT (Q140D)]. In FT (Y85H, Q140D), the hydrogen bond is stabilized by the simultaneous change of both amino acids. The relationships between the hydrogen bond breakage or formation and the functions of these modified proteins are further discussed below.

The observations that the hydrogen bond present in TFL1, a floral repressor, is not present in the two modified TFL1 proteins, TFL1 (H88Y) and TFL1 (H88Y, D144Q), and that
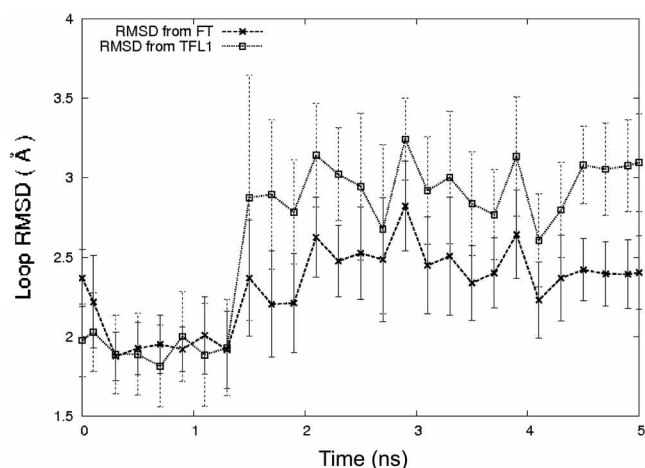
Figure 4. Change in RMSDs (in Angstroms) of the loop structure of TFL1 (H88Y) from those of the native FT and TFL1 proteins over the simulation time (in nanoseconds). The time 0 is defined in the same way as in Figure 3.

the external loop conformations of the two modified TFL1 proteins are closer to that of FT, a floral activator, than TFL1 implies that existence of the hydrogen bond is highly correlated with the external loop conformation and the function of the floral regulators. To understand this correlation better, temporal change of the loop conformation of TFL1 (H88Y) for which experimental data for function is available was examined. In Figure 4, the loop RMSDs of TFL1 (H88Y) from FT and TFL1 are plotted along the simulation time. Initially, the loop structure of TFL1 (H88Y) is slightly closer to that of TFL1 than to FT, but significant conformational change occurs after 1.5 ns of simulation time, during which the loop structure becomes closer to that of FT than to TFL1. Because the breakage of the hydrogen bond in TFL1 (H88Y) due to mutation of His88 to Tyr happens at 0.3 ns, as indicated by a sharp increase in the residue distance from 5.5 Å to 7 Å in Figure 3(a), this loop structure conversion occurs at 1.2 ns after the hydrogen bond breakage. This fast induction of conformational change of the external loop after hydrogen bond breakage implies that the hydrogen bond interaction directly influences the conformation of the neighboring external loop in the intramolecular interaction network of the protein.

The modified FT protein, FT (Y85H), for which experimental data for function is available, did not show TFL1-like characteristics in terms of loop conformation or hydrogen bond formation according to our simulations [Table 1 and Figure 3(b)]. However, the modified FT protein with two amino acid changes, FT (Y85H, Q140D), forms a hydrogen bond between the modified amino acids, His and Asp, just like TFL1. Moreover, deviation of the external loop conformation of FT (Y85H, Q140D) from native TFL1 is smaller than that of FT (Y85H) from native TFL1 (See Table 1). These observations lead to a prediction that the simultaneous change of the two amino acids in FT (Tyr85 to His and Gln140 to Asp) may convert FT to a repressor more completely than a single amino acid change (Tyr85 to His) does.

although the single modification resulted in relatively weak change in function from an activator to a weak activator in terms of controlling flowering time.[5] This prediction may be tested experimentally in the future.

From the current molecular dynamics simulation study on FT, TFL1, and their modified proteins, it can be concluded that the functional conversion of the floral repressor TFL1 to an activator by a single amino acid change (His88 to Tyr) is due to conformational change of the neighboring external loop caused by hydrogen bond breakage between the key amino acid His88 and the amino acid Asp144 located in the external loop. In other words, the structure at the site of the key amino acid or interaction of the key amino acid with the immediate neighbors may not be the determining factor of the function. Features of binding partners that distinguish between FT and TFL1 and thus contributing to their opposite functions have not been discovered yet. The present simulation results suggest that the protein-protein interface involving the external loop, especially the N-terminal region of the loop, may be critical in determining the functional specificity of the floral regulators.

References

1. Kobayashi, Y.; Kaya, H.; Goto, K.; Iwabuchi, M.; Araki, T. *Science* **1999**, *286*, 1960.
2. Kardailsky, I.; Shukla, V. K.; Ahn, J. H.; Dagenais, N.; Christensen, S. K.; Nguyen, J. T.; Chory, J.; Harrison, M. J.; Weigel, D. *Science* **1999**, *286*, 1962.
3. Bradley, D.; Ratcliffe, O.; Vincent, C.; Carpenter, R.; Coen, E. *Science* **1997**, *275*, 80.
4. Schoentgen, F.; Saccoccio, F.; Jolles, J.; Bernier, I.; Jolles, P. *Eur. J. Biochem.* **1987**, *166*, 333.
5. Ahn, J. H.; Miller, D.; Winter, V. J.; Banfield, M. J.; Lee, J. H.; Yoo, S. Y.; Henz, S. R.; Brady, R. L.; Weigel, D. *EMBO J.* **2006**, *25*, 605.
6. Hanzawa, Y.; Money, T.; Bradley, D. *Proc. Nat. Acad. Sci. U.S.A.* **2005**, *102*, 7748.
7. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
8. Case, D. A.; Cheatham Iii, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
9. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
10. Tsui, V.; Case, D. A. *Biopolymers (Nucleic Acid Sciences)* **2001**, *56*, 257.
11. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824.
12. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122.
13. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
14. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
15. Coutsias, E. A.; Seok, C.; Dill, K. A. *J. Comput. Chem.* **2004**, *25*, 1849.
16. Coutsias, E. A.; Seok, C.; Dill, K. *J. Comput. Chem.* **2005**, *26*, 1663.