http://dx.doi.org/10.5012/bkcs.2012.33.5.1597

The Identification of Binding Mode for *Arabidopsis thaliana* 7-Keto-8-aminopelargonic Acid Synthase (*At*KAPAS) Inhibitors

Jae Eun Cho, Sun Young Kang, Jung Sup Choi,† Young Kwan Ko,† In Taek Hwang,† and Nam Sook Kang*

Graduate School of New Drug Discovery and Development, Chungnam National University, Deajeon 305-764, Korea
**E-mail: nskang@cnu.ac.kr*
†*Korea Research Institute of Chemical Technology, P.O. Box 107, Yuseong-gu, Daejeon 305-600, Korea*
Received January 18, 2012, Accepted February 9, 2012

In this study, we determined the 3D-structure of *Arabidopsis thaliana* KAPAS by homology modeling. We then investigated the binding mode of compounds obtained from in-house library using computational docking methods. From the flexible docking study, we achieved high dock scores for the active compounds denoted in this study as compound **3** and compound **4**. Thus, we highlight the flexibility of specific residues, Lys 312 and Phe 172, when used in active sites.

Key Words : KAPAS, Herbicides, Homology modeling, Docking

Introduction

Research on the topic of herbicides has advanced during the past 50 years to the point that it can now protect crops and elevate the quality and quantity of agricultural products. However, the successful development of herbicides has decreased recently owing to new environmental regulations. To overcome this problem, there is an urgent need for new herbicidal targets and new techniques.[1]

While the traditional approach to discover lead compounds heavily depends on serendipity given the poor understanding of biological modes of actions, the structure-based approach utilizes the structure of appropriate target proteins which have well-known binding sites for possible rational designs. The structure-based approach uses only appropriate target proteins instead of the entire plant for *in vivo* testing.[2] In order to perform a structure-based assay, it is necessary to determine a potent target with a thorough understanding of the mechanism of action of the target.

Several enzymes in plants are known to be essential enzymes, meaning that they are crucial for the plant's survival. Disrupting a single essential enzyme leads to severe disorder in the metabolism of the plant, ultimately causing a lethal phenotype to arise. 7-keto-8-aminopelargonic acid synthase from *Arabidopsis thaliana* plants (*At*KAPAS), introduced in this research, is a new potent herbicide target which is involved in the early steps of the creation of the biotin biosynthesis pathway. *At*KAPAS as pyridoxal 5'-phosphate-dependent enzyme catalyzes the decarboxylative condensation of L-alanine with pimeloyl-CoA in a stereo specific manner to form KAPA, coenzyme A, and carbon dioxide in the first committed step ofbiotin biosynthesis. Inhibiting *At*KAPAS leads to significant changes in the phenotype, such as growth inhibition, severe growth retardation, and the creation of lethal phenotypes.[3]

Although the physiological systems of human and plants are different in various ways, the misuse of agricultural chemicals can be extremely harmful to humans. Therefore, herbicides must follow strict toxicity regulations that are in place to prevent harm to humans. As mentioned above, the novel herbicidal target 7-keto-8-aminopelargonic acid synthase functions in the initial steps of the biosynthetic pathways of biotin (vitamin H) in plants and microorganisms. Because biosynthetic steps of biotin exist only in plants, we expect that the inhibition of the potent target *At*KAPAS will not affect the human metabolic system.[1,3] A few publications have also reported beneficial effects of *At*KAPAS as a potential herbicidal target.Hwang *et al.* described the possibility of *At*KAPAS as a potential herbicide target enzyme and chemical validation of triphenyltin acetate as a lead compound for the *At*KAPAS inhibition *in vitro* and *in vivo*.[1] They also suggested *At*KAPAS can be a useful target for the rational design of inhibitors in the hope of developing new herbicides. Therefore, we aim to obtain potential *At*KAPAS inhibitors using the knowledge-based computational informatics method in this research. We described the 3D-structure of *At*KAPAS *via* theoretical method and the binding mode for *At*KAPAS and its inhibitors obtained from *in vitro* assay for in house compounds.

Experimental Section

Building a Homology Model of *At*KAPAS. To apply the structure-based drug design (SBDD) method using current knowledge of protein and drug interactions, a three-dimensional protein structure is necessary.[4] Because the known protein crystal structural information of 7-keto-8-aminopelargonic acid synthase from an experiment was absent, a homology model of *At*KAPAS was constructed from its amino acid sequence (Table 1). To obtain the 3D-structure of *At*KAPAS, a hierarchical protein structure modeling approach was adopted based on secondary-structure enhanced Profile-Profile threading Alignment (PPA) and iterative implementation by the Threading ASSEmbly

Table 1. Single letter amino acid sequence for *At*KAPAS

10	20	30	40	50	60
madhswdktv	eeavnvlesr	qilrslrpic	msrqneeeiv	ksranggdgy	evfdglcqwd
70	80	90	100	110	120
rtsvevsvsi	ptfqkwlhde	psngeeifsg	dalaecrkgr	fkkllfsgn	dylglsshpt
130	140	150	160	170	180
isnaaanavk	eygmgpkgsa	licgyttyhr	llesslaqlk	kkedclvcpt	gfaanmaamv
190	200	210	220	230	240
aigsvaslla	asgkplknek	vaifsdalnh	asiidgvrla	erqgnvevfv	yrhcdmyhln
250	260	270	280	290	300
sllsnckmkr	kvvvtdslfs	mdgdfapmee	lsqlrkkygf	llviddahgt	fvcgengggv
310	320	330	340	350	360
aeefnceadv	dlcvgtlska	agchggfiac	skkwkqliqs	rgrsfifsta	ipvpmaaaay
370	380	390	400	410	420
aavvvarkei	wrrkaiwerv	kefkelsgvd	isspiislvv	gnqekalkas	ryllksgfhv
430	440	450	460	470	480
mairpptvpp	nscrlrvtls	aahttedvkk	litalsscld	fdntathips	flfpkl

Refinement (TASSER) program.[5] We obtained five candidate models for the three-dimensional structure of *At*KAPAS and then performed molecular dynamics simulations with the use of the CHARMM force field (version 27.0)[6] with default parameters interfaced with Accelrys Discovery Studio 3.1. To identify binding sites, we collected the crystal structures of homologous proteins with *At*KAPAS as templates from the RCSB Protein Data Bank (PDB) based on three different categories: structural similarity, binding site similarity, functional similarity. Ten different protein crystal structures originating from different species but structurally similar to *At*KAPAS were selected. The binding site sequence and conformation of the target protein are particularly important here compared to other sites. Therefore, we found 10 different PDB hits which are similar in terms of their binding site to *At*KAPAS. A comparison between each PDB hit and the *At*KAPAS biding site sequence was done. The results are denoted using root mean square deviation (RMSD) values ranging from 1.83 to 3.51, as shown in Table 2. To identify the binding mode of *At*KAPAS, 10 different enzymes (PDB code; 1FC4,[7] 1DJE,[8] 2WKA,[9] 2BWO,[10] 3KKI,[11] 3DXV,[12] 3DXW,[12] 2OAT,[13] 1GBN,[14] and 1MLY[15]) with different functions originating from different species were superimposed (Figure 1). Consequently, we found that several residues were crucial for protein-ligand interactions. His 210 mainly forms a π-π interaction or a π-cation interaction with the ligands, and all of the reference proteins contain histidine residue at a position homologous to the *At*KAPAS. Adjacent to His 210, Phe 172 forms a π-π interaction with compounds that have an aromatic ring moiety as well. 50% of proteins have phenyl residue at a similar position; however, its conformation was considerably different. As shown in Figure 1, conformation of Phe 172 residue of our homology model, shown in yellow, is uniquely folded toward the active site cavity as opposed to the phenyl residues of the reference PDB. To confirm the effect of Phe 172, we undertook a flexible docking. Lys 312 also plays an important role as a hydrogen bond donor in the active site. All reference

Table 2. Sequence similarity and RMSD between 10 reference PDB proteins

PDB code	Sequence	RMSD
1FC4	36.3	1.83
1DJE	39.5	1.66
2BWO	39.6	1.59
2WKA	32.2	2.42
3KKI	30.3	2.64
3DXV	16.6	3.26
3DXW	16.8	3.25
2OAT	15.7	3.38
1GBN	15.6	3.43
1MLY	15.5	3.51

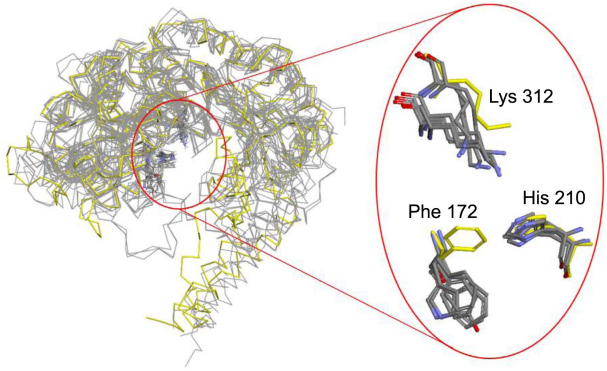


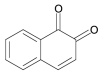
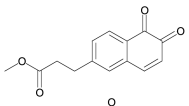
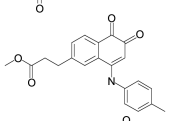
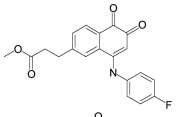
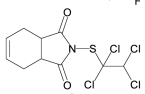
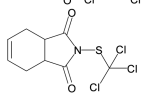
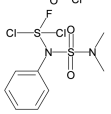
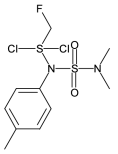
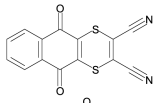
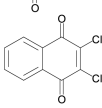
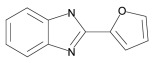
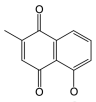
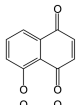
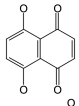
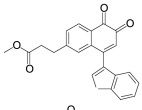
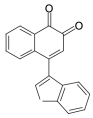
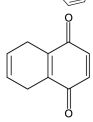
Figure 1. The superimposed 10 known enzymes (PDB code: 1FC4, 1DJE, 2WKA, 2BWO, 3KKI, 3DXV, 3DXW, 2OAT, 1GBN, and 1MLY). The yellow colored structure represents our homology model and the overlaped predominantly important resides in active site are elaborated on the right side of the entire protein alignment.

enzymes showed the lysine residue at an analogous location which led to π-cation interaction and hydrogen bonding interaction with its ligand. In addition, due to its comparatively free long aliphatic chain,lysine residue was very flexible. Lysine residues superimposed onto the analogous site verified its flexibility. By superimposing the binding sites of other similar proteins,we reached the conclusion that the residues His 210, Phe 172 and Lys 312 play major roles at the binding site. This research will therefore highlight the ligand binding site residues and its flexibility to search for potent hits.

Rigid Docking. As a structure-based drug design, we used an automated docking method.[16] To dock the compounds as shown in Table 3 into the protein active site, we used the rigid docking method implemented in Discovery Studio 3.1 (Accelrys, Inc.), which adopts a Monte-Carlo algorithm to generate ligand conformations and docks the generated ligands into the active site using a shape-based filtering method. The rigid docking process consists of two main steps: defining a binding cavity and docking ligands onto the defined cavity.[17] The *At*KAPAS protein as obtained from the homology model needs to be prepared before the docking process is initiated. To prepare the protein, the CHARMM force field was assigned and the docking cavity was defined using the advanced define and edit binding site tool module. Ligands also should be prepared with low-energy conformations so as to be docked onto a 'clean' protein. The protocol 'Generate Conformations' in Discovery Studio 3.1 was used to obtain three-dimensional conformations of each ligand. In-house 17 ligands were generated by the protocol 'Generate Conformation' with the conformation method 'BEST', and the CHARMM force field was applied. For an interaction filter, Lys 312 was selected.

Flexible Docking. Various docking methods are utilized by researchers. Each approach was developed by focusing on different aspects of docking. One of the factors determining the accuracy of docking is protein flexibility. Much emphasis has been placed on the conformational changes of protein binding sites, where different ligands form interactions. The Flexible Docking protocol of Discovery Studio 3.1 allows receptor flexibility during the docking of ligands.[18-20] To confirm the flexibility of the selected residues in the *At*KAPAS active site, three different sets of residues were defined as flexible docking 1; Lys 312, flexible docking 2; Phe 172, and flexible docking 3; Lys 312 and Phe 172. For the flexible docking 1 set, only the residue Lys 312 was assigned to move when flexible docking was underway. 17 pre-processed ligands were docked into the prepared *At*KAPAS homology model. The 'BEST' conformation method was selected to generate three-dimensional ligands with stable energy levels, and other parameters were left with the default values. The residue Phe 172 was selected for the second flexible docking trial with the same parameters used with the flexible docking 1 set. To validate the effect of both residues, Lys 312 and Phe 172 were set as the flexible docking 3 group, and these residues were moved while flexible docking was underway.

Table 3. The structures of the used 17 compounds and its biological activies

Compounds	Structure	pIC_{50}
1		5.48
2		5.36
3		6.23
4		6.20
5		4.25
6		4.27
7		4.39
8		4.33
9		4.95
10		4.95
11		4.38
12		5.68
13		5.02
14		5.36
15		5.97
16		5.90
17		5.36

***In vitro* Assay.** Pimeloyl CoA was synthesized according to the method described previously.[21] *At*KAPAS activity was determined according to the method described previously[22] usinga linked assay by monitoring the increase in absorption of NADH at 340 nm using a Microplate Spectrophotometer (Benchmark Plus, Bio-rad, USA), thermostatically controlled at 30 °C. A typical assay contained 20 mM potassium phosphate (pH 7.5), 1 mM α-ketoglutarate, 0.25 mM thiamine pyrophosphate, 1 mM NAD^+, 3 mM $MgCl_2$, 0.1 unit of α-ketoglutaratedehydrogenase, and 2 to 10 μg of *At*KAPAS in a total volume of 200 μL. L-Alanine and pimeloyl-Co A were added to give the desirable final concentrations. Prior to analysis, enzyme samples were dialyzed for 2 hours at 4 °C against 20 mM potassium phosphate (pH 7.5) containing 100 μM PLP. The KAPAS concentration in allanalyses was 10 μM in 20 mM potassium phosphate (pH 7.5) and the concentrations of each compound were 0.1 to 250 μM. Reference corvettes contained all other compounds except inhibitor.

Results and Discussion

Significant results were obtained from the docking processes undertaken in this study. The rigid docking output scores of the in-house compounds are shown in Table 4. The most active compounds, in this case compound **3** and compound **4**, obtained high dock scores of 104.21 and 105.47, respectively. Moreover, active ligands which have IC_{50} (μM) values 1.07 and 1.26 (compound 15 and compound 16, respectively) formed a stable docking pose (Figure 2) with high dock scores 106.03 and 72.5, respectively. However, other active compounds, specifically compound **1**, compound **9** and compound **10**, showed rather low dock scores, as shown in Table 3.

To obtain a better docking result, we used a flexible docking strategy, as mentioned in the experimental section. In the result for flexible docking 1, which stipulated that the Lys 312 residue of the *At*KAPAS model is set to move, most of the compounds formed a stable conformation and formed interactions with the *At*KAPAS homology model (Table 5).

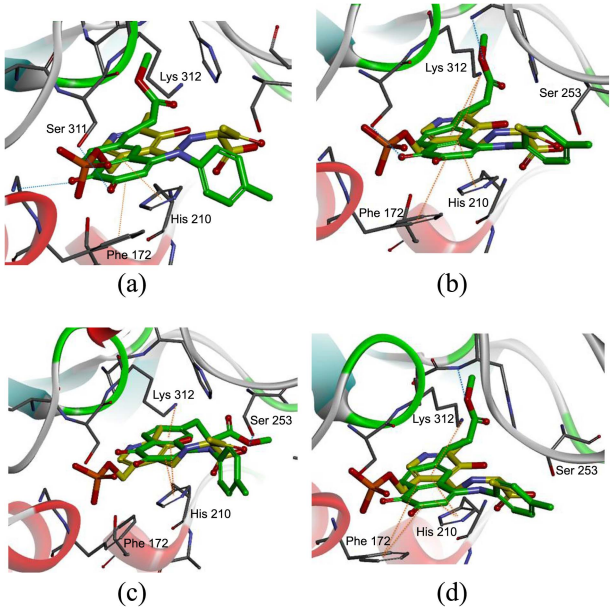


Figure 2. The docking pose from the process of rigid docking and flexibie docking (a~d). The blue dashed line represents hydrogen bonding interaction and the orange dashed line represents the π-cation and π-π interaction. The docked ligand, compound **3**, is colored with green and the crystal ligand of 1FC4 colored with yellow is overlapped as a reference compound. The docking pose fromthe process of rigid docking is shown in (a). From the process of flexible dockset 1, the docking pose of compound **3** was obtained as shown in (b). The docking pose shown in (c) achieved from the process of flexible docking set 2. In addition, the docking pose of compound **3** shown in (d) obtained from the process of flexible docking set 3.

Table 4. The result of rigid docking

Compound	Dock score	pIC_{50}[a]
1	40.40	5.48
2	67.64	5.36
3	104.21	6.23
4	105.47	6.20
5	55.73	4.25
6	42.30	4.27
7	50.67	4.39
8	51.67	4.33
9	58.72	4.95
10	45.79	4.95
11	49.91	4.38
12	63.13	5.68
13	52.84	5.02
14	52.99	5.36
15	106.03	5.97
16	72.51	5.90
17	52.31	5.36

[a]$pIC_{50} = -\log IC_{50}$

Table 5. The result of flexible docking set 1

Compound	Dock score	pIC_{50}
1	74.00	5.48
2	83.14	5.36
3	134.03	6.23
4	125.27	6.20
5	53.40	4.25
6	50.83	4.27
7	59.46	4.39
8	54.40	4.33
9	105.54	4.95
10	76.74	4.95
11	74.36	4.38
12	83.56	5.68
13	76.92	5.02
14	80.04	5.36
15	131.67	5.97
16	96.88	5.90
17	75.41	5.36

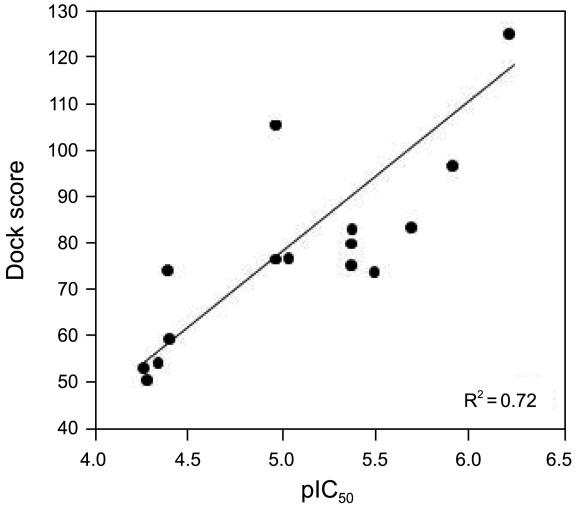


Figure 3. Correlation of IC_{50} and dock score for the flexible docking set 1.

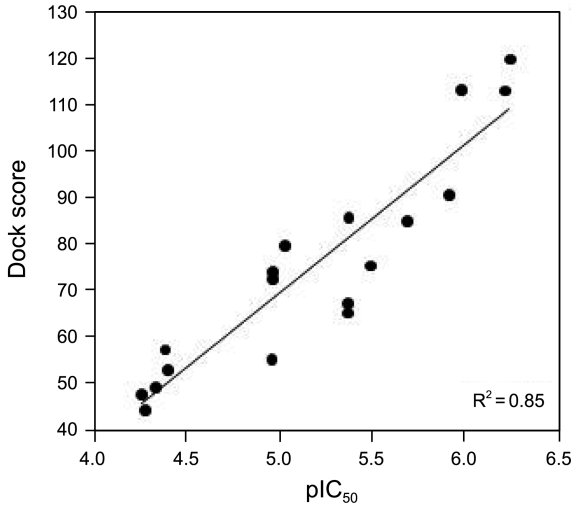


Figure 4. Correlation of IC_{50} and dock score for the flexible docking set 2.

According to the docking result, the diverse conformation of Lys 312 directly affects the pose of the ligands and the related activity. The Lys 312 residue forms a hydrogen bond or undergoes the π-cation interaction mostly with the oxygen moiety of the ligand by flexibly moving through the protein cavity (Figure 2). Interestingly, we obtained relatively high dock scores for compound **1**, compound **9** and compound **10**. This was unobtainable with the rigid docking process. The correlation coefficient between pIC_{50} and the dock score for the flexible docking 1 set was 0.72, as shown in Figure 3.

By flexibly moving the residue Phe 172 while the flexible docking 2 process was underway, a better result than the flexible docking 1 process was obtained (Table 6). The flexibility of the Phe 172 residue has a significant effect on the ligand binding at the *At*KAPAS active site. The phenyl ring is particularly important to form the π-π interaction with compounds. According to the docking result, high dock scores of 119.91 and 113.06 were the result with the most active compounds, compound **3** and compound **4**. In addition, active compounds with IC_{50} values lower than 1.3 μM received dock scores higher than 90 with a stable conformation. Conversely, the inactive compounds of compound **5** through compound **8** as well as compound **11** obtained relatively low dock scores (Table 6).

The two specific *At*KAPAS residues, Lys 132 and Phe 172, can be moved mutually for the flexible docking 3 process. The result shows the clear discrepancy between an active compound and an inactive compound (Table 7). The flexible docking 3 results showed a higher correlation between pIC_{50} and the dock score compared to flexible docking 2, whereas it had a slightly lower R^2 value than the flexible docking 2 result (Figure 5). Because the two residues were set to move at the same time, more diverse results could be obtained. As a result of several flexible

Table 6. The result of flexible docking set 2

Compound	Dock score	pIC_{50}
1	75.60	5.48
2	85.84	5.36
3	119.05	6.23
4	113.06	6.20
5	48.11	4.25
6	44.80	4.27
7	53.49	4.39
8	49.60	4.33
9	72.97	4.95
10	74.43	4.95
11	57.66	4.38
12	85.24	5.68
13	79.99	5.02
14	67.73	5.36
15	113.29	5.97
16	90.84	5.90
17	65.49	5.36

Table 7. The result of flexible docking set 3

Compound	Dock score	pIC_{50}
1	71.47	5.48
2	95.12	5.36
3	120.48	6.23
4	119.46	6.20
5	40.92	4.25
6	40.03	4.27
7	54.90	4.39
8	61.54	4.33
9	80.87	4.95
10	72.13	4.95
11	75.00	4.38
12	78.83	5.68
13	74.82	5.02
14	73.04	5.36
15	124.28	5.97
16	91.92	5.90
17	70.70	5.36

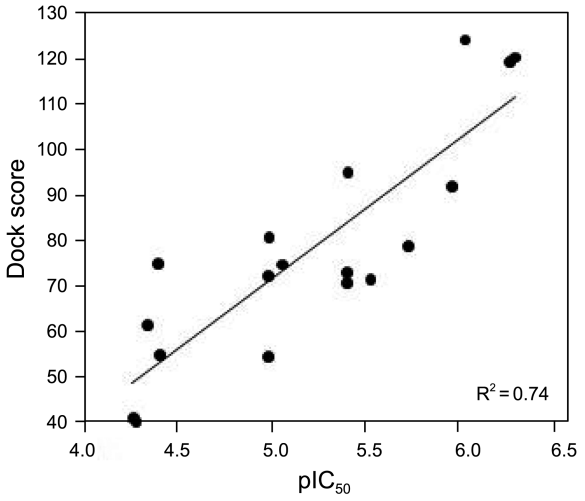


Figure 5. Correlation of IC_{50} and dock Score for the flexible docking set 3.

docking processes, this study emphasizes the flexibility of several residues. Noticeably higher dock scores were obtained from flexible docking as compared to rigid docking. The rigid Lys 312 residue mostly tends to form π-cation interaction with the aromatic moiety of compounds, allowing a certain amount of space for compounds during the rigid docking process. However, the flexible Lys 312 forms either the π-cation interaction or the hydrogen bonding interaction with hydrogen bond acceptors. The flexibility of the residue, including Lys 312 and Phe 172, allows more space for compounds, thus offering a better docking pose and dock scores.

Conclusion

In this study, we determined the 3D-structure of *At*KAPAS by homology modeling. We then investigated the binding mode of our in-house library using computational docking methods. In the rigid docking of the in house compounds as shown in Table 3, the most active compounds, in this case compound **3** and compound **4**, obtained high dock scores of 104.21 and 105.47, respectively. However, some active compounds showed rather low dock scores. To obtain a better docking result, we used a flexible docking strategy. From the flexible docking study, we achieved high dock scores and stable binding conformations for the active compounds denoted in this study. Thus, we highlight the flexibility of specific residues, Lys 312 and Phe 172, when used in active sites. Furthermore, we are going to optimize compound **3** and compound **4** using this homology model for *At*KAPAS.

References

1. Hwang, I. T.; Choi, J. S.; Song, H. Y.; Cho, S. J.; Lim, H. K.; Park, N. J.; Lee, D. H. *Pesticide Biochemistry and Physiology* **2010**, *97*, 24.
2. Hwang, I. T.; Lee, D. H.; Park, N. J. *InTec.* **2011**, *471*, 22.
3. Choi, J. S.; Lim, H. K.; Seo, B. R.; Kim, J. S.; Choi, C. W.; Kim, Y. S.; Ryu, S. Y. *The Korean Society of Weed Science* **2011**, *31*, 240.
4. Krieger, E.; Nabuurs, S. B.; Vriend, G. *Structural Bioinformatics* **2003**, *25*, 507.
5. Zhang, Y. *Proteins* **2007**, *69*, 108.
6. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
7. Schmidt, A.; Sivaraman, J.; Li, Y.; Larocque, R.; Barbosa, J. A.; Smith, C.; Matte, A.; Schrag, J. D.; Cygler, M. *Biochemistry* **2001**, *40*, 5151.
8. Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. *Biochemistry* **2000**, *39*, 516.
9. Jahan, N.; Potter, J. A.; Sheikh, M. A.; Botting, C. H.; Shirran, S. L.; Westwood, N. J.; Taylor, G. L. *J. Mol. Biol.* **2009**, *3*, 763.
10. Astner, I.; Schulze, J. O.; van den Heuvel, J.; Jahn, D.; Schubert, W. D.; Heinz, D. W. *EMBO J.* **2005**, *18*, 3166.
11. Kelly, R. C.; Bolitho, M. E.; Higgins, D. A.; Lu, W.; Ng, W. L.; Jeffrey, P. D.; Rabinowitz, J. D.; Semmelhack, M. F.; Hughson, F. M.; Bassler, B. L. *Nat. Chem. Biol*. **2009**, *12*, 891.
12. Okazaki, S.; Suzuki, A.; Mizushima, T.; Kawano, T.; Komeda, H.; Asano, Y.; Yamane, T. *Biochemistry* **2009**, *5*, 941.
13. Storici, P.; Capitani, G.; Müller, R.; Schirmer, T.; Jansonius, J. N. *J. Mol. Biol.* **1999**, *1*, 297.
14. Shah, S. A.; Shen, B. W.; Brünger, A. T. *Structure* **1997**, *8*, 1067.
15. Sandmark, J.; Mann, S.; Marquet, A.; Schneider, G. *J. Biol. Chem.* **2002**, *227*, 43352.
16. Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151.
17. Montes, M.; Miteva, M. A.; Villoutreix, B. O. *Protein* **2007**, *15*, 712.
18. Welch, W.; Ruppert, J.; Jain, A. N. *Chem. Biol.* **1996**, *3*, 449.
19. Koska, J.; Spassov, V. Z.; Maynard, A. J.; Yan, L.; Austin, N.; Flook, P. K.; Venkatachalam, C. M. *J. Chem. Inf. Model* **2008**, *48*, 1965.
20. Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. *J. Mol. Graph. Model* **2003**, *21*, 289.
21. Ploux, O.; Marquet, A. *Biochem. J.* **1992**, *283*, 327.
22. Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. *Biochemistry* **2000**, *39*, 516.