A proposal of potent inhibitor for cancer metastasis blocking the pocket of urokinase receptor: \textit{ab initio} molecular simulations

Tomoyo Kasumi\textsuperscript{1}, Kohta Araki\textsuperscript{1}, Tatsuroh Mizushima\textsuperscript{1}, Hiroshi Kobayashi\textsuperscript{2} and Noriyuki Kurita\textsuperscript{1*}

\textsuperscript{1}Department of Computer Science of Engineering, Toyohashi University of Technology, Tempaku-cho, Toyohashi, Aichi, 441-8580, Japan
\textsuperscript{2}Department of Obstetrics and Gynecology, Nara Medical University, 840 Shijio-cho, Kashihara, Nara, 634-8521, Japan

\textsuperscript{*}E-mail: kurita@cs.tut.ac.jp

Abstract. Recent biochemical experiments have elucidated that a variety of proteases play important roles in cancer invasion and metastasis. In particular, binding of urokinase-type plasminogen activator (uPA) to uPA receptor (uPAR) existing on the surface of a cancer cell is considered to be a trigger for cancer invasions. Therefore, the blocking of the binding is expected to inhibit cancer invasion. In previous experiments, several peptides of amino acids were proposed as a potent inhibitor for blocking the binding. In the present study, we obtained stable structures of the solvated complexes with uPAR and the peptides and investigated the specific interactions between uPAR and the peptides by \textit{ab initio} molecular simulations. Based on these results, we clarified which peptide can bind more strongly to uPAR and proposed a novel potent peptide which can inhibit the uPAR-uPA binding efficiently.

1. Introduction
Recently, the number of deaths due to cancer has been increasing continuously. The main cause for the death of cancer patients is cancer metastasis. Cancer can spread rapidly to organs in the body by metastasis and robs the patient of basis functions necessary for supporting the life. Therefore, the inhibition of cancer metastasis is expected to improve the patient’s prognosis dramatically. However, there is no effective treatment for inhibiting the metastasis, and the development of effective inhibitors for the metastasis is a pressing research issue to solve [1]. In general, it is difficult for cancer patients to take strong medicine with side effect and to visit a hospital every day for taking injection drugs. It is thus necessary to develop medicines of tablet that inhibit cancer metastasis with relatively few side effects.

During the processes of invasion and metastasis, cancer cells produce various proteases to dissolve cell tissues such as extracellular matrix (ECM) and basement membrane and facilitate invasion into blood vessels. The cancer cell passes through the blood vessel to reach the endothelial cells of other organs and invades these cells to develop a new metastases focus [2]. By repeating this process, cancer cells grow proliferously. Therefore, it is expected that the inhibition of cancer invasion can suppress cancer metastasis and lead to a good prognosis of the cancer patients.

The recent experiment [3] elucidated several proteases playing essential roles in the cancer invention. In particular, urokinase-type plasminogen activator (uPA) [4, 5] has been found to play a
key role in cancer invasion and metastasis. The specific binding of the amino-terminal fragment (ATF) of uPA to the uPA receptor (uPAR) [6] existing on the surface of cancer cell is considered a trigger for promoting cancer invasions [7]. Consequently the blocking of binding between uPA and uPAR is expected to inhibit cancer invasions and metastasis. Based on this idea, Kobayashi et al. synthesized a chimeric protein ATF-HI-8 [8], which is composed of bikunin existing in amniotic fluid and the ATF of uPA. The effectiveness of this chimeric protein on inhibiting cancer invasion has been confirmed by the biological studies using laboratory animals [1]. In order to propose which parts of ATF-HI-8 are important for the specific interactions between ATF-HI-8 and uPAR, we investigated the stable structure and its electronic properties of ATF-HI-8 by semiempirical molecular simulations [9, 10].

As for the specific interactions between uPAR and uPA, the previous structure analyses [11, 12] for the complex of uPAR with uPA elucidated that the amino acid residues 1-132 of uPA bind specifically to uPAR. In particular, the Ω-loop (residues of 19-31) in the ATF of uPA was found to be responsible for the high-affinity interactions between uPAR and uPA [13]. On the other hand, we investigated the specific interactions between uPAR and ATF at electronic level by ab initio molecular orbital calculations and elucidated that the positively charged Lys46, Lys61, and Lys98 residues as well as the Ω-loop of ATF contribute to the binding between uPAR and ATF [14].

Many types of peptide drugs inhibiting the uPAR-uPA binding were synthesized based on the Ω-loop structure of ATF. In the ATF around the Ω-loop, two disulfide bonds (Cys11-Cys19 and Cys13-Cys31) are formed, indicating that the distance between Cys19 and Cys31 is affected by the positions of Cys11 and Cys13. In fact, the distance between Cys19 and Cys31 is 6.1 Å, which is significantly longer than those (5.2 Å) of the Cys11-Cys19 and Cys13-Cys31 disulfide bonds. Considering this flexibility of Cys19 and Cys31 positions, two types of the 19-31 peptides with and without the disulfide bond between Cys19 and Cys31 were synthesized. Their effects on the inhibition were investigated experimentally to find that both the peptides have an equal inhibition rate for uPAR. Based on the result, 8 kinds of ring-shaped peptides having a disulfide bond were proposed as a peptide drug inhibiting the uPAR-uPA binding [15]. However, their effects on the inhibition and their specific interactions with uPAR are not investigated yet. In order to develop potent peptide inhibitors, it is necessary to clarify the specific binding properties between uPAR and the peptides.

In the present study, to elucidate the specific interactions between uPAR and the 8 types of peptides proposed in the previous experiment [15], we obtained stable structures for the solvated complexes with uPAR and these peptides by classical molecular mechanics (MM) simulations. In addition, the specific interactions between uPAR and the peptides were investigated at an electronic level by ab initio fragment molecular orbital (FMO) calculations. From these results, we clarified which peptide can bind more strongly to uPAR and proposed a new peptide with large binding affinity to uPAR. These results will become useful for the development of novel potent inhibitors.

2. Details of molecular simulations

2.1. Construction and optimization of solvated structure for uPAR-peptide complexes

In our current molecular simulations, we employed the X-ray crystal structure of the ternary complex of uPAR-ATF-SMB, in which SMB is somatomedin B domain of vitronectin. This structure is registered in the Protein Data Bank (PDB ID: 3BT1 [16]), and we used it because it has only missing two amino acids 83 and 84 in uPAR. This structure is of uPAR (residues 1-275) in complex with the ATF (residues 8-132) of uPA, SMB (residues 2-41) and other co-factors. We extracted only the uPAR structure from this PDB structure and predicted the structures of these missing residues in uPAR, by using the homology modeling program SWISS-MODEL [17]. The N- and C-terminals of uPAR were terminated by NH$_3^+$ and COO$^-$, respectively.

Since the X-ray crystal structure of uPAR has no information about hydrogen atoms, hydrogen atoms should be added to the original PDB structure in an appropriate manner. In particular, it is not unique to determine the positions of hydrogen atoms for His amino acid, because His has some different protonated structures depending on the pK$a$ value around the His. We here evaluated the pK$a$
values of His residues contained in uPAR by using the PROPKA Web Interface 3.0 [18-20] and determined the protonation based on the pK\textsubscript{a} values. His amino acid has three types of protonated structures: Hid has a hydrogen atom at the \(\delta\)-site of its imidazole ring, Hie has a hydrogen atom at the \(\varepsilon\)-site of the imidazole ring, and positively charged Hip has hydrogen atoms at both the \(\delta\)- and \(\varepsilon\)-sites. His residues with pK\textsubscript{a} value larger than 6 have the Hip protonation, while those with pK\textsubscript{a} value smaller than 6 have the Hid or Hie protonation. uPAR has 12 His residues and their pK\textsubscript{a} values are listed in Table 1, indicating that 7 His residues have pK\textsubscript{a} values smaller than 6 in uPAR. These His residues have Hid or Hie protonation. Because it is not practical to consider all the combinations of Hid/Hie for the 7 His residues included in uPAR, we here considered that the 7 His residues have Hie protonation, while the other 5 His residues have Hip protonation. As for the other ionizable amino acid residues contained in uPAR, the ionized state was used.

### Table 1. pK\textsubscript{a} values and protonation states of each His residue in uPAR.

| Amino acid | pK\textsubscript{a} | Protonation |
|------------|----------------------|-------------|
| His47      | 4.18                 | Hie         |
| His110     | 5.23                 | Hie         |
| His128     | 4.40                 | Hie         |
| His143     | 6.18                 | Hip         |
| His160     | 3.90                 | Hie         |
| His166     | 1.14                 | Hie         |
| His203     | 6.38                 | Hip         |
| His229     | 5.05                 | Hie         |
| His249     | 6.64                 | Hip         |
| His251     | 5.26                 | Hip         |
| His260     | 6.40                 | Hip         |
| His273     | 6.32                 | Hip         |

We added solvating water molecules in an 8 Å shell around uPAR and considered them explicitly in our classical MM simulations. This solvated uPAR structure was optimized using the classical MM program AMBER9 [21], in which the Parm99SB [22] and TIP3P [23] force fields were used for uPAR and the water molecules, respectively. The threshold value of the energy-gradient for the convergence of the AMBER9 optimization was set as 0.001 kcal/mol/Å.

The structures of the 8 kinds of the peptides, each of which have amino acid sequence listed in Table 2, were optimized in solvating water molecules by the AMBER9 program [21]. The N- and C-terminals were terminated by an acetyl group (-CH\textsubscript{3}CO) and an amine group (-NHCH\textsubscript{3}), respectively. We considered that the His residues of the peptides have Hie protonation, in order to unify the total charge of all peptides.

### Table 2. Amino acid sequences of the peptides investigated in the present study.

The common sequence is underlined.

| Peptide | Amino acid sequence (from N-terminal to C-terminal) | Number of residues |
|---------|----------------------------------------------------|--------------------|
| P1      | Cys, Asn, Ser, Phe, Tyr, Lys, Cys                  | 7                  |
| P2      | Trp, Cys, Asn, Ser, Phe, Tyr, Lys, Cys             | 8                  |
| P3      | Trp, His, Cys, Asn, Ser, Phe, Tyr, Lys, Cys        | 9                  |
| P4      | Trp, Cys, Asn, Ser, Phe, Tyr, Lys, Asn, Cys        | 9                  |
| P5      | Trp, Cys, Ile, Asn, Ser, Phe, Tyr, Lys, Cys        | 9                  |
| P6      | Trp, Cys, His, Ile, Asn, Ser, Phe, Tyr, Lys, Asn, Cys | 11                |
| P7      | Trp, Cys, His, Ile, Asn, Ser, Phe, Tyr, Lys, Asn, Ser, Cys | 12                |
| P8      | Trp, Cys, Ile, Asn, Ser, Phe, Tyr, Lys, Asn, Ser, Cys | 11                |
To obtained candidate structures for the complex of uPAR-peptide, we docked the peptide to a variety of sites around uPAR using the automated protein-ligand docking program Autodock 4.2 [24]. In the docking procedure, the grid box was set as the 40.0 × 126.0 × 40.0 Å³ centered on the gravity center of uPAR, and the spacing between the nearest neighboring grid points was set to 0.375 Å, which is the default value of Autodock 4.2. The structure of uPAR was fixed, and all dihedral angles of the peptide were freely rotated in the docking procedure to search for a variety of stable configurations for the peptide docked to uPAR. We here created 200 candidate structures of the complex by using the genetic algorithm of Autodock 4.2. They were classified into some clusters according to the RMSD value (5.0 Å) between each of the structures created.

The representative structures in each cluster were fully optimized by the AMBER9-MM method. To consider the solvation effect on the complex properly, we added water molecules with a 8 Å layer around the complex and optimized the solvated structure by using AMBER9 [21], in which the Parm99SB [22] and TIP3P [23] force fields were assigned for the complex and water molecules, respectively. The threshold value of the energy-gradient for the convergence in the AMBER9 optimization was set as 0.001 kcal/mol/Å.

2.2. Ab initio FMO calculations for the solvated structures of uPAR-peptide complexes

The electronic properties for the most stable structure of the solvated uPAR-peptide complexes were investigated by the ab initio FMO calculation [25], to elucidate which amino acid residues in uPAR and peptide are important for the specific binding between uPAR and peptide. In the FMO method, the target molecule is divided into units called “fragment”, and the electronic properties of the target molecule are estimated from the electronic properties of the monomers and dimers of the fragments. The specific interactions between the fragments were elucidated from the interaction energies obtained by the FMO calculation.

In the present study, we used the FMO calculation program ABINIT-MP Ver.4.3 [26]. In the FMO calculation, ab initio MP2[27,28]/6-31G method was employed to investigate accurately the π-π stacking, NH-π and CH-π interactions between the amino acid residues of uPAR and peptides. Each amino acid residue of uPAR and peptide, each water molecule were assigned as a fragment, because this fragmentation enables us to evaluate the interaction energies between the amino acid residue of uPAR and peptide. From the comparison of the interaction energies, it is possible to propose which amino acid residues of uPAR and peptide have large contribution to the specific interactions between uPAR and peptide. In addition, we considered the water molecules existing within a 5 Å from the peptide explicitly, in order to elucidate the influence of solvating water molecules on the specific interactions. The number of solvating water molecules was set as 100 for all the complexes to investigate the relative stability among the solvated structures.

In order to investigate the biding energy between uPAR and peptide, the solvated uPAR-peptide complex was divided into the following four structural domains: uPAR-peptide containing solvating water molecules (uPAR-peptide+water), uPAR containing solvating water molecules (uPAR+water), the peptide containing solvating water molecules (peptide+water) and solvating water molecules (water). From total energies (T.E.) obtained by FMO calculations, the binding energy (B.E.) between uPAR and the peptide mediated by solvating water molecules was estimated as B.E. = − T.E. (uPAR-peptide+water) + T.E. (uPAR+water) + T.E. (Peptide+water) − T.E. (water).

3. Results and discussion

3.1. Stable structures of the solvated complexes of uPAR-peptide

To check the validity of the solvated structure of uPAR optimized by AMBER9, we compared it with the uPAR structure in the X-ray structure of the uPAR-uPA-SMB complex registered in PDB (PDB ID: 3BT1 [16]). The RMSD for all Cα atoms of uPAR is 1.3 Å, indicating that the uPAR structure is changed significantly by the optimization in water. We furthermore investigated the displacement of Cα atom for each amino acid residue of uPAR. As shown in Figure 1, among the 275 residues of
uPAR, 10 residues have displacement larger than 2.5 Å. Among them, 6 residues exist on the surface of uPAR to be affected by the solvation. In contrast, the other 4 residues exist near the 83rd and 84th residues, whose structures are missing in the X-ray structure. Therefore, the effect of solvation is likely to cause the change in structure of these residues, resulting in the RMSD of 1.3 Å. In the previous biochemical experiment [13], the residues 24-70, 122-171 of uPAR were found to contribute to the specific interactions between uPAR and uPA. Figure 1 indicates that the Cα positions of these residues are not affected significantly by the optimization in water. It is thus elucidated that the structure of uPA binding site of uPAR optimized in water by the AMBER9-MM method is reasonable.

![Figure 1. Displacement of each amino acid residue of uPAR between the optimized structure in water and the X-ray structure [17].](image)

The peptides were docked to uPAR by the docking program AutoDock4.2 [24] to create the 200 candidate structures for each uPAR-peptide complex. For the peptides P1, P6 and P7, all of the 200 structures were grouped into a single cluster, while for the other peptides, the created structures were grouped into some clusters. The representative structures in each cluster were fully optimized by the AMBER9-MM method, and their total energies calculated by the \textit{ab initio} MP2/6-31G method in FMO [26] are listed in Table 3.
Table 3. Total energies (T.E.) (kcal/mol) for the optimized structures of uPAR-peptide complexes evaluated by ab initio FMO method, and their ranking in stability. P1 and C1 mean the peptide 1 and the cluster 1, respectively.

| Structure     | T.E.     | ∆ T.E. | Ranking |
|---------------|----------|--------|---------|
| uPAR-P1       | -79706094.5 | 0.0    | 1       |
| uPAR-P2-C1    | -80087576.3 | 0.0    | 1       |
| uPAR-P2-C2    | -80086969.4 | 606.9  | 2       |
| uPAR-P3-C1    | -80382434.2 | 0.0    | 1       |
| uPAR-P3-C2    | -80382276.3 | 157.9  | 2       |
| uPAR-P4-C1    | -80347053.4 | 547.0  | 3       |
| uPAR-P4-C2    | -80347600.4 | 0.0    | 1       |
| uPAR-P4-C3    | -80347451.0 | 149.4  | 2       |
| uPAR-P5-C1    | -80315784.2 | 0.0    | 1       |
| uPAR-P5-C2    | -80315454.9 | 329.3  | 2       |
| uPAR-P5-C3    | -80315405.7 | 378.5  | 3       |
| uPAR-P6       | -80870742.6 | 0.0    | 1       |
| uPAR-P7       | -81071961.0 | 0.0    | 1       |
| uPAR-P8-C1    | -80777063.6 | 0.0    | 1       |
| uPAR-P8-C2    | -80776590.0 | 473.6  | 2       |

To clarify the reason for the difference in total energy listed in Table 3, we analyzed the optimized structures of the uPAR-peptide complexes in detail as shown in Figure 2, where the boundary between the inside and the outside of the ligand-binding pocket of uPAR is defined by drawing a line between the Cα atoms of Glu34 and Glu132, because both the residues exist near the entrance of the pocket. Based on Figure 2, we attempted to elucidate how the peptide enters into the pocket of uPAR.

There are three candidate structures for the uPAR-P4 complex as listed in Table 3. The most stable structure is at least 150 kcal/mol more stable than the other ones. In this structure, P4 goes deeply into the pocket of uPAR and binds strongly to the residues (Thr127 and Asp140) of uPAR, which exist at the bottom of the pocket. As a result, almost all parts of P4 enter into the pocket as shown in Figure 2d. In the other two clusters of the uPAR-P4 complex, P4 exists around the entrance of the pocket and is not stabilized enough. For the uPAR-P5 complex, the situation is the same as the uPAR-P4 complex, and almost all parts of P5 enter into the pocket as shown in Figure 2e. In contrast, in the complexes with uPAR and the other peptides, the peptides are stabilized at the intermediate position of the pocket as shown in Figure 2. In the uPAR-uPA complex, the Ω-loop of uPA composed of the residues 19-31 enters completely into the pocket and binds strongly to uPAR [16]. Consequently, the present molecular simulations elucidate that only P4 and P5 can bind to uPAR in the similar way as the uPAR-uPA complex.

To elucidate the reason why only P4 and P5 can enter deeply into the pocket of uPAR, we compared the structures of all peptides. They are stabilized by the disulfide bond between Cys residues and some hydrogen bonds between residues. In the P4 and P5, the amino group of the Lys side chain and the oxygen atom of the backbone between Lys and Tyr are hydrogen bonded each other. As a result, P4 and P5 are stabilized into a compact form to enter easily into the pocket. In contrast, in the other peptides, the Lys side chain extends outward. And it seems that this side chain inhibits the peptide to enter into the pocket. Therefore, it is clarified that the conformation of the Lys side chain in the peptide has large influence on the binding position of the peptides in the uPAR pocket.
Figure 2. Close-up views around the uPA-binding pocket of uPAR and peptide for the optimized structure of the solvated uPAR-peptide complexes; (a) uPAR-P1, (b) uPAR-P2, (c) uPAR-P3, (d) uPAR-P4, (e) uPAR-P5, (f) uPAR-P6, (g) uPAR-P7 and (h) uPAR-P8. The horizontal line, which connects the Cα atoms of the Glu34 and Glu132 residues of uPAR, indicates a boundary between the inside and the outside of the uPA-binding pocket of uPAR.

3.2. Electronic properties of the solvated complexes of uPAR-peptide

We first evaluated the binding energy between uPAR and the peptide by \textit{ab initio} FMO calculations and predicted the binding affinity between them. As listed in Table 4, P4 has the largest binding energy among the 8 peptides investigated here. To elucidate which amino acid residues of P4 are important for this binding energy, the interaction energy between each amino acid residue of P4 and uPAR was investigated by FMO. As shown in Figure 3a, Lys has the strongest attractive interaction with uPAR. We furthermore investigated the interaction energy between Lys and each residue of uPAR. Table 5 lists the residues of uPAR having strong attractive interaction with Lys in the decreasing order of the magnitude of interaction. The positively charged Lys residue of P4 has attractive interactions with negatively charged Asp and Glu residues of uPAR by the electrostatic interactions. In fact, as shown in Figure 4a, the amino group of the Lys side chain of P4 and the COO group of Asp140 of uPAR are hydrogen bonded, in addition, a hydrogen bond is formed between the amino group and the OH group of the Thr127 side chain. As a result, non-charged Thr127 has strong attractive interaction with the Lys of P4. It is thus elucidated that Asp140 and Thr127 as well as the negatively charged Asp and Glu residues of uPAR contribute to the strong binding between uPAR and P4.

In our previous molecular simulations for the uPAR-uPA complex [14], Lys23 included in the Ω-loop of uPA interacts strongly with the Asp140 and Thr127 residues of uPAR. These residues are located on the bottom of the ligand-binding pocket of uPAR. In addition, the interacting structure between Lys of P4 and Asp140/Thr127 of uPAR shown in Figure 4a is similar to that in the uPAR-uPA complex. Accordingly, P4 is expected to be a potent inhibitor for the uPAR-uPA binding.
Table 4. Total energies (T.E.) for each component of the optimized structures of uPAR-peptide complexes, and estimated binding energies (B.E.) between uPAR and peptide. The B.E. was evaluated as 
\[ B.E. = - T.E.(uPAR-peptide+water) + T.E.(uPAR+water) + T.E.(peptide+water) - T.E.(water). \]

| Structure | T.E.(kcal/mol)  | B.E. (kcal/mol) |
|-----------|-----------------|-----------------|
| uPAR-P1 (7)* | -79706094.5 | -7109012.4 | -4777168.3 | 246.3 |
| uPAR-P2 (8) | -80087576.3 | -7490053.3 | -4777250.4 | 670.3 |
| uPAR-P3 (9) | -80382434.2 | -7785173.4 | -4777269.0 | 464.1 |
| uPAR-P4 (9) | -80347600.4 | -7749953.9 | -4777260.3 | 760.4 |
| uPAR-P5 (9) | -80315784.2 | -7718130.2 | -4777206.1 | 632.8 |
| uPAR-P6 (11) | -80777063.6 | -8179758.8 | -4777292.6 | 579.7 |

*The number of residues contained in each peptide is shown in parenthesis.

Table 5. Interaction energies (I.E.) (kcal/mol) between Lys of the peptide 4 and the amino acid residues of uPAR in the uPAR-P4 complex. The residues having strong attractive interaction are listed in the decreasing order of the magnitude of interaction.

| Residues | I.E. |
|----------|------|
| Asp140   | -90.4 |
| Asp254   | -45.2 |
| Glu68    | -38.8 |
| Asp141   | -29.9 |
| Thr127   | -29.2 |
| Glu42    | -27.3 |
| Glu183   | -27.1 |
| Asp124   | -26.1 |
| Glu134   | -25.8 |
| Glu106   | -24.1 |

As listed in Table 4, P6 has the second largest binding energy to uPAR. The difference in specific interactions between uPAR and peptide was investigated for P4 and P6. Figure 3b shows the interaction energies between each residues of P6 and uPAR. The attractive interaction between Lys and uPAR is about 100 kcal/mol weaker compared with the uPAR-P4 complex shown in Figure 3a. To elucidate the reason for this difference, the position of Lys in both the uPAR-Peptide complexes were compared in Figure 5. In the uPAR-P4 complex, Lys exists in the pocket of uPAR, while the Lys of P6 exists out of the pocket, resulting in the weaker interaction between uPAR and Lys. We considered that this difference in the Lys position is caused from the effect of the charged residues around Lys139 existing at the ligand-binding pocket of uPAR. As shown in Figure 6, negatively charged residues Glu34, Glu36, Asp140, Asp141, Glu132, Glu134 and Glu135 exist around Lys139 to surround the pocket of uPAR. Positively charged parts are marked in blue, while negatively charged parts are marked in red, indicating that there is a large polarization around Lys139. Because the peptides have only one positively charged Lys residue, this residue feels sensitively the electrostatic potential of the
charged residues existing around the pocket of uPAR, when the peptides approach to the pocket. Accordingly, the position of Lys and the conformation of its side chain in the peptide is seems to have large effect on the binding properties of peptide to uPAR. In fact, Lys of P6 cannot enter the pocket of uPAR as shown in Figure 5b. In this complex, Asn existing at the opposite side of Lys interacts with Asp140 and Lys139 of uPAR, as shown in Figure 4b, playing an important role in the binding between uPAR and P6. Therefore, Asn as well as Lys in the peptides are expected to be important for the binding to uPAR.

Figure 3. Interaction energies (I.E.) between each amino acid residue of peptide and uPAR for (a) uPAR-P4 and (b) uPAR-P6 complexes.
Figure 4. Hydrogen bonding structures and distances (Å) between amino acid residues of uPAR and (a) Lys of Peptide 4 for the uPAR-P4 complex, (b) Asn of Peptide 6 for the uPAR-P6 complex.

Figure 5. Position of Lys residue of peptide in the uPAR-peptide complex: (a) uPAR-P4 and (b) uPAR-P6 complexes.
Figure 6. Surface electric charge distribution around the ligand binding pocket of uPAR. Positively charged parts are marked in blue, while negatively charged parts are marked in red.

3.3. Proposal of potent inhibitor for uPAR-uPA binding

The 8 kinds of peptides proposed in the previous experiment [15] have a common sequence of Asn-Ser-Phe-Tyr-Lys, as listed in Table 2. This sequence comes from the central 5 residues of the 7 residues (N-terminal-Ile-Asn-Ser-Phe-Tyr-Lys-Asn-C-terminal) constructing the Ω-loop of uPA. The Ω-loop is considered to be flexible and contribute mainly to the uPAR-uPA binding.

Our present calculations elucidate that P4 and P6 bind strongly to uPAR as listed in Table 4. These peptides have the common sequence of Lys-Asn-Cys on the C-terminal. This Asn residue corresponds to the terminal Asn of the 7 residues of the flexible Ω-loop. To investigate the effect of this residue on the binding between uPAR and peptide, we compared the results for P2 and P4, because P2 has the same sequence as P4 except for the missing of Asn, as shown in Table 2. The evaluated binding energy for P2 is 90 kcal/mol smaller than that for P4, because P2 has the same sequence as P4 except for the missing of Asn, as shown in Table 2. The evaluated binding energy for P2 is 90 kcal/mol smaller than that for P4, because P2 has the same sequence as P4 except for the missing of Asn, as shown in Table 2. The evaluated binding energy for P2 is 90 kcal/mol smaller than that for P4, because P2 has the same sequence as P4 except for the missing of Asn, as shown in Table 2. Therefore, it can be concluded that the Asn residue on the C-terminal is important for strong binding of peptide to uPAR.

On the other hand, from the comparison between the results for P2 and P5 listed in Table 4, it was elucidated that the Ile residue on the N-terminal of the 7 residues weakens the binding between uPAR and peptide. The binding energy of P5 is 37.5 kcal/mol smaller than that of P2, indicating that the addition of Ile to P2 causes the weakening of the uPAR-peptide binding. From the above comparison of the binding energies listed in Table 4, we concluded that the peptides having the sequence such as (N-terminal)-Xn-Lys-Asn-Cys-(C-terminal) may bind strongly to uPAR and inhibit the uPAR-uPA binding.

Based on these results obtained by the present calculations, we attempted to propose some potent inhibitors for blocking the binding between uPAR and uPA and investigated their binding properties to
uPAR. Because P4 has the largest binding energy, we mutated some residues of P4 by the other type of residue to improve the binding affinity to uPAR.

We first constructed a peptide by mutating Asn of the N-terminal of P4 by the negatively charged Asp. As shown in Figure 3b and Figure 4b, this Asn residue plays important roles in the binding between P6 and uPAR, although Asn is an uncharged residue. In order to enhance the interaction between this residue and uPAR, we mutated it by the negatively charged Asp residue, which has similar side chain as Asn. This peptide is defined as P4-ND. Figure 7a shows the optimized structure of the uPAR-P4-ND complex, indicating that P4-ND enters completely into the pocket of uPAR and binds to Thr127 and Asp140 in the similar way as P4. In addition, the mutated Asp residue of P4-ND is hydrogen bonded with Lys139 of uPAR. As a result, the interaction energies between each residue of P4-ND and uPAR are slightly enhanced by the mutation, as shown in Figure 8a. Therefore, it is expected that our proposed P4-ND peptide can bind strongly into the ligand-binding pocket of uPAR.

We furthermore constructed a new peptide by mutating Ser of P4 by the other residue, since the interaction energies between each residue of P4 and uPAR shown in Figure 3a indicates that Ser interacts weakly with uPAR. This Ser residue was mutated by Thr, which has a similar side chain as Ser, and the specific interactions between uPAR and the mutated peptide were investigated. This mutated peptide is defined as P4-ST. As shown in Figure 7b, P4-ST can enter into the pocket of uPAR completely. This result comes from the compact structure of P4-ST, in which the amino group of Lys side chain is hydrogen bonded to the two oxygen atoms of the backbones between Lys and Tyr and between Trp and Cys. In addition, the side chain of Asn is hydrogen bonded to the nitrogen atoms of the backbones between Trp and Thr. These hydrogen bonds make P4-ST compact, and P4-ST enters easily into the pocket of uPAR. However, P4-ST dose not have enough side chains for interacting with uPAR, so that the interaction energies between each residue of P4-ST and uPAR are significantly smaller than those for the uPAR-P4 complex, as shown in Figure 8b. Therefore, P4-ST is not expected to be a potent inhibitor for the uPAR-uPA binding.

Figure 7. Stable structures of the solvated complexes with uPAR and our proposed peptides optimized by AMBER9; (a) uPAR-P4-ND and (b) uPAR-P4-ST.
Figure 8. Interaction energies (I.E.) between each amino acid residue of peptide and uPAR for (a) uPAR-P4-ND and (b) uPAR-P4-ST complexes.
4. Conclusions

In the present study, we have investigated at atomic and electronic levels the binding properties between uPAR and the 8 kinds of peptides proposed in the previous experiment [15] and elucidated the following points.

1) The peptide with the Trp-Cys-Asn-Ser-Phe-Tyr-Lys-Asn-Cys sequence has the largest binding energy with uPAR and can be a potent inhibitor for the uPAR-uPA binding.

2) Lys and Asn residues included in the above peptide are important for its binding to uPAR.

3) The peptides having the sequence of N-terminal-Xn-Lys-Asn-Cys-C-terminal can bind strongly to uPAR.

4) In the uPAR, Thr127, Lys139 and Asp140 contribute mainly to the uPAR-peptide binding.

Based on these results, we furthermore proposed new peptides and investigated their binding properties to uPAR. The results indicate that the peptide with the Trp-Cys-Asp-Ser-Phe-Tyr-Lys-Asn-Cys sequence binds strongly to uPAR and can be a potent inhibitor for the uPAR-uPA binding.

Acknowledgments

This work was supported in part by the grants from Iketani Science and Technology Foundation, Tatematsu Foundation, CASIO Science Promotion Foundation and Toukai Foundation for Technology.

References

[1] Kobayashi H, Gotoh J, Hirashima Y, Fujie M, Sugino D and Terao T 1995 *The Journal of Biological Chemistry* 270 8361-8366

[2] Ruoslahti E 1996 *Scientific American* 9 48-55

[3] Baker E A and Leaper D J 2003 *European Journal of Cancer* 39 981-988

[4] Skelly M M, Troy A, Duffy M J, Mulcahy H E, Duggan C, Connell T G, O’Donoghue D P and Sheahan K 1997 *Clinical Cancer Research* 3 1837-1840

[5] Herszényi L, Plebani M, Carraro P, Paoli M De, Roveroni G, Cardin R, Tulassay Z, Naccarato R and Farinati F 1999 *Cancer* 86 1135-1142

[6] Stephens R W, Nielsen H J, Christensen I J, Thorlacius-Ussing O, Sorensen S,Dano K and Brunner N 1999 *Journal of the National Cancer Institute* 91 869-874

[7] Nagase K, Kobayashi H, Nagata S, Yamamoto H, Natsume T, Dedachi K, Tsukamoto T and Kurita N 2007 *Clinical Cancer Research* 13 387-404

[8] Kobayashi H, Terao T, Sugino D and Okushima M 2007 *Clinical Cancer Research* 13 387-404

[9] Sugiuura H, Tsumura N, Kobayashi H and Kurita N 2004 *Journal of Computer Aided Chemistry* 5 62-69

[10] Tsumura N, Sugiuura H, Kobayashi H and Kurita N 2004 *Journal of Computer Aided Chemistry* 5 70-76

[11] Appella E, Robinson E A, Ullrich S J, Stoppelli M P, Corti A, Cassani G and Blasi F 1987 *The Journal of Biological Chemistry* 262 4437-4440

[12] Bdeir K, Kuo A, Sachais B S, Rux A H, Bdeir Y, Mazar A, Higazi A A and Cines D B 2003 *Blood* 102 3600-3608

[13] Barinka C, Parry G, Callahan J, Shaw D E, Kuo A, Bdeir K, Cines D B, Mazar A and Lubkowski J 2006 *Journal of Molecular Biology* 363 482-495

[14] Tsuji S, Kasumi T, Nagase K, Yoshikawa E, Kobayashi H and Kurita N 2011 *Journal of Molecular Graphical and Modelling* 29 975-984

[15] GmbH R D, Burgle M, Graeff H, Kessler H, Magdelen V, Konig B, Koppitz M, Riemer C, Schmitt M, and Weidle U 1998 WO98/46731
[16] Huai Q, Zhou A, Lin L, Mazar A P, Callahan J, Shaw D E, Furie B C and Huang M 2008 Nature Structure & Molecular Biology 15 422-423
[17] Schwede T, Kopp J, Gues N and Peitsch M C 2003 Nucleic Acids Res 31 3381-3385
[18] Li H, Robertson A D and Jensen J H 2005 Proteins 61 704-721
[19] Delphine C B, Rogers D M and J. H. Jensen J H 2008 Proteins 73 765-783
[20] Olsson M H M, Søndergard C R, Rostkowski M and Jensen J H 2011 Journal of Chemical Theory and Computation 7 525-537
[21] Case D A, Cheatham III T E, Darden T, Gohlke H, Luo R, Merz Jr. K M, Onufriev A, Simmerling C, Wang B and Woods R J 2005 Journal of Computational Chemistry 26 1668-1688
[22] Wang J, Cieplak P and Kollman P A 2000 Journal of Computational Chemistry 21 1049-1074
[23] Jorgensen W L, Chandrasekhar J, Madura J and Klein M L 1983 Journal of Chemical Physics 79 926-935
[24] Morris G M, Huey R, Lindstrom W, Sanner M F, Belew R K, Goodsell D S and Olson A I 2009 Journal of Computer Chemistry 16 2785-2791
[25] Kitaura K, Sawai T, Asada T, Nakano T and Uebayashi M 1999 Chemical Physics Letter 312 319-324
[26] Ito M, Fukuzawa K, Ishikawa T, Mochizuki Y, Nakano T, and Tanaka S 2008 The Journal of Physical Chemistry B 112 12081-12094
[27] Mochizuki Y, Nakano T, Koikegami S, Tanimori S, Abe Y, Nagashima U and Kitaura K 2004 Theoretical Chemistry Accounts 112 442-452
[28] Mochizuki Y, Koikegami S, Nakano T, Amari S and Kitaura K 2004 Chemical Physics Letter 396 473-479