

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 13 日現在

機関番号：12601

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24570179

研究課題名(和文) マルチスケールシミュレーションによるタンパク質へのリガンド結合過程の研究

研究課題名(英文) Multi-scale simulation of protein-ligand binding process

研究代表者

寺田 透 (Terada, Tohru)

東京大学・農学生命科学研究科・准教授

研究者番号：40359641

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：粗視化分子動力学シミュレーションを用いて、リガンドがタンパク質のリガンド結合ポケットに結合する過程を明らかにする方法を確立した。この方法を、リガンド結合ポケットの形状や物理化学的性質、リガンドの物理化学的性質が異なる多数のタンパク質・リガンドペアに適用した。この結果の比較から、リガンドは特定のパスウェイを経由してリガンド結合ポケットに入ることが示唆された。また、全原子モデルを用いてパスウェイを精密化し、これに沿った自由エネルギープロファイルを計算する方法を確立した。

研究成果の概要(英文)：I established a method to reproduce the protein-ligand binding process in a coarse-grained molecular dynamics (MD) simulations. This method was applied to many protein-ligand pairs that differ in terms of the shape of the ligand-binding pocket and the physicochemical properties of the pocket and the ligand. The comparison of the results suggested that the ligands enter the ligand-binding pockets through specific pathways. Furthermore, I also established a method to refine the ligand-binding pathway with all-atom MD simulations and to calculate free-energy profile along the pathway.

研究分野：計算生物物理学

キーワード：分子動力学シミュレーション、タンパク質、リガンド、パスウェイ、マルチコピー・マルチスケール法、粗視化モデル、自由エネルギー計算

1. 研究開始当初の背景

タンパク質への低分子化合物(リガンド)の結合は、代謝系や細胞内・外の情報伝達系に数多く見られる、生命活動を特徴付ける重要な化学過程の一つである。リガンドは、一般に基質ポケットと呼ばれる、タンパク質表面にある特定の窪みに結合する。反応を起こさないが、基質ポケットに結合するリガンドを用いれば、酵素の機能を阻害することができる。ノイラミニダーゼの基質ポケットに結合してこの酵素の働きを阻害する抗インフルエンザ薬タミフルのように、多くの薬剤がこの原理に基づいて開発されてきた。

リガンドのタンパク質への結合能を精度良く予測するためには、リガンド結合過程に対する理解が欠かせない。実験による結合アッセイに倣い、タンパク質の周囲に水分子とともに候補化合物を一定の割合で配置してシミュレーションを行い、リガンド結合過程を追跡することができれば理想的である。しかし、リガンドのタンパク質への結合過程の時間スケールは、一般にミリ秒のオーダーであり、通常分子動力学法を用いてこの過程を再現することは困難である。これは、系を構成する数万以上の原子間に働く力の計算に時間がかかるためと、運動方程式を数値的に解く際に用いる時間刻みがフェムト秒程度であるため、膨大な(1ミリ秒で 10^{12})ステップの逐次計算が必要となるためである。粗視化シミュレーションはこれら問題を解決する上で有効な方法である。複数の原子を1つの粒子にマップすることで、相互作用点の数を減らし1ステップの計算に必要な計算量を減らすことができる。また、エネルギー曲面が滑らかになることで、長い時間刻みを使用することが可能になり、粒子間の摩擦も減るため、同じ時間でもより広い立体構造空間を探索できるようになる。

MARTINI は、このような目的で開発された粗視化モデルの1つである。脂質と水分子のモデルをランダムに配置し、MARTINI 力場パラメータを用いてシミュレーションを行うと、自発的に脂質2重膜を形成することが示されている[1]。これは従来の全原子モデルを用いたシミュレーションでは不可能であり、粗視化モデルの潜在能力の高さを示している。このモデルはタンパク質に拡張され、膜タンパク質のシミュレーションに適用されている[2]。

研究代表者らは、MARTINI の可能性を探るため、2つの異なるタンパク質・リガンド系を対象にMARTINIを用いた予備的な粗視化シミュレーションを行った。ここでは、タンパク質の周囲に複数のリガンドをランダムに配置した構造を初期構造とした。シミュレーションの結果、リガンドは基質ポケット以外のタンパク質表面上の窪みにも結合するが、基質ポケットに最も安定に結合することが示された。ここから、MARTINI はリガンドの基質ポケットへの結合過程を再現しう

る能力を持っていると期待される。

<引用文献>

- [1] Marrink S. J. et al. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
- [2] Monticelli, L. et al. *J. Chem. Theory Comput.* **4**, 819–834 (2008).

2. 研究の目的

本研究では、粗視化分子動力学シミュレーションに基づき、リガンドがタンパク質表面上のリガンド結合ポケットに結合する過程を再現する方法を確立する。この方法を、リガンド結合ポケットの形状や表面の物理化学的性質、リガンドの物理化学的性質が異なる多数のタンパク質・リガンドペアに系統的に適用し、結果を比較することで、リガンドによるリガンド結合ポケットの認識機構や、リガンドのリガンド結合ポケットへの結合過程を支配する要因を明らかにする。また、一般に粗視化モデルは全原子モデルに比べて精度が劣るといわれているが、粗視化分子動力学シミュレーションのトラジェクトリデータから、結合・解離速度定数および解離定数を計算し、実験データと比較することで、本手法の精度を評価する。また、粗視化分子動力学シミュレーションによって得られたリガンド結合パスウェイを、全原子モデルを組み合わせたマルチスケールシミュレーションを用いて精密化し、リガンド結合過程の自由エネルギー地形を高い精度で求める。

3. 研究の方法

(1) タンパク質・リガンドペアの選定

立体構造データベースに登録されているタンパク質・リガンド複合体に対する網羅的な調査を行い、性質が異なる多様なタンパク質・リガンドペアをシミュレーション対象として選定した。具体的な手順は以下の通り。生物学的に意味のあるリガンドを結合したタンパク質立体構造データベース Binding MOAD [1]から、分子量550程度以下のリガンドを1種類のみ結合し、かつタンパク質単体の立体構造も決定されているタンパク質・リガンドペアを選定した。

リガンド結合ポケットの形状(広い/狭い)、基質ポケット表面の物理化学的性質(親水性・静電相互作用あり/親水性・静電相互作用なし/疎水性)、リガンドの物理化学的性質(親水性・静電相互作用あり/親水性・静電相互作用なし/疎水性)に従ってタンパク質・リガンドペアを分類した。基質ポケットの形状の評価には、CAVER 3.0 [2]を用いた。類似した配列を持つタンパク質・リガンドペアを除いた後、数の多い分類から計算対象とするペアを2つずつ選定した。

(2) 立体構造変化への対応

MARTINI を用いたタンパク質の粗視化シミュレーションでは、距離が9 Å以内の粗視化

粒子間をバネで仮想的に結合する elastic network model を用いて立体構造を保持する[4]。しかし、このままではリガンド結合に伴う立体構造変化を再現することができない。そこで、本研究では、multiple-basin model [5]を組み合わせたことで、複数の安定な立体構造をとることができるように、分子動力学シミュレーションプログラム Gromacs [6]の改良を行った。

(3) 粗視化シミュレーションの実施と解析
上で選定したタンパク質・リガンドペアに含まれるリガンドのうち、既存のパラメータ（タンパク質、脂質、糖など）から類推できないものについて粗視化パラメータを決定した。ここでは、MARTINI の力場パラメータ決定のプロトコール[3]に従った。まず、リガンドの概ね 4 つの重原子に対して 1 つの粗視化粒子をマップし、粗視化モデルを構築した。リガンドの全原子モデルのシミュレーションを行い、粗視化モデルの共有結合パラメータを、この結果を再現するように決めた。次いで、原子種、Lennard-Jones パラメータ、電荷を決定した。ここでは、水中およびオクタノール中におけるリガンドの粗視化シミュレーションに基づいて、油相と水相の間の分配係数を計算した。これを分配係数の実験値または予測値と比較し、原子種の割り当てやパラメータの調整を行った。

タンパク質のリガンド非結合構造の周りに水分子とともにリガンド分子を 4 ~ 50 個ランダムに配置し、1 ~ 5 μ s の粗視化シミュレーションを、初速とリガンド配置を変えながら 50 ~ 100 回程度繰り返した（表 3 参照）。次いで、粗視化シミュレーションのトラジェクトリデータを並進・回転させてタンパク質の立体構造を重ね合わせた後、系を 6 Å の格子に切り、各格子に含まれるリガンド重心の密度からリガンドの分布を計算した。結晶構造におけるリガンド結合部位に結合したリガンドの構造ポーズをクラスタに分類し、結晶構造に近いポーズを含むクラスタについて、結合・解離速度定数および解離定数を計算した。これを、実験データと比較することで、粗視化分子動力学シミュレーションの精度を評価した。また、系を 2.5 Å の格子に切り、各格子に含まれるリガンド重心の速度からリガンドの流束密度を計算した。これをタンパク質・リガンドペアごとに可視化し、タンパク質表面の形状や物理化学的性質、リガンドの物理化学的性質との関係に着目して比較を行った。

(4) マルチスケールシミュレーション
タンパク質・リガンドペアを 1 つ選定し、マルチコピー・マルチスケール分子シミュレーション法の 1 つである string 法[7]を用いて、リガンド結合パスウェイの精密化を行った。まず、粗視化シミュレーションの結果得られた結合パスウェイを用いて初期パスを構築した。このパスを 32 個のイメージ（タンパク質、リガンド、イオン、水分子からなる全

原子モデルのコピー）を用いて離散化し、全原子分子動力学シミュレーションによって計算された自由エネルギー勾配を用いて、最小自由エネルギーパスに向けて最適化を行った。パスが収束した後、これを反応座標としたアンブレラサンプリングを行い、multistate Bennett acceptance ratio (MBAR)法 [8]を用いて、この座標に沿った自由エネルギー地形を求めた。これらの計算には、研究代表者らが開発したマルチコピー・マルチスケール分子シミュレーションソフトウェア μ^2 lib (<http://www.mu2lib.org/>)を用いた。

<引用文献>

- [1] Hu, L. *et al. Proteins* **60**, 1097–10134 (2005); Benson, M. L. *et al. Nucleic Acids Res.* **36**, D674–D678 (2008).
- [2] Petrek, M. *et al. BMC Bioinformatics* **7**, 316 (2006); Chivancova, E. *et al. PLoS Comput. Biol.* **8**, e1002708 (2012).
- [3] Marrink, S. J. *et al. J. Phys. Chem. B* **111**, 7812–7824 (2007).
- [4] Periole, X. *et al. J. Chem. Theory Comput.* **5**, 2531–2543 (2009).
- [5] Maragakis, P. and Karplus, M. *J. Mol. Biol.* **352**, 807–822 (2005); Okazaki, K. *et al. Proc. Natl. Acad. Sci. USA* **103**, 11844–11849 (2006).
- [6] van der Spoel, D. *et al. J. Comput. Chem.* **26**, 1701–1718 (2005).
- [7] Maragliano, L. and Vanden-Eijnden, E. *Chem. Phys. Lett.* **446**, 182–190 (2007).
- [8] Shirts, M. R. and Chodera, J. D. *J. Chem. Phys.* **129**, 124105 (2008).

4. 研究成果

(1) タンパク質・リガンドペアの選定
生物学的に意味のある分子量 550 程度以下のリガンドを 1 種類のみ結合し、かつタンパク質単体の立体構造も決定されているタンパク質・リガンドペアを、リガンド結合ポケットの形状、基質ポケット表面の物理化学的性質、およびリガンドの物理化学的性質に従って分類した結果を表 1 に示す。

表 1: タンパク質・リガンドペアの分類結果

リガンド結合ポケット	リガンド	リガンド					
		親水性・静電相互作用あり	親水性・静電相互作用なし	疎水性	疎水性		
親水性	狭い	179	I	46	III	11	
	広い	260	II	132	IV	37	
疎水性	狭い	26		12		142	V
	広い	39		21		105	VI

このうち、ペア数の多いI~VIの分類から2つずつ、タンパク質・リガンドペアを選定した(表2)。

表 2: 選定されたタンパク質・リガンドペア

分類	タンパク質 ^a	PDB ID	
		リガンド非結合	リガンド結合
I	DAHPS	1OFB	1HFB
I	AST	1AAW	1ASM
II	6PGL	2J0E	3EB9
II	HAP	3IT2	3IT1
III	PNP	1ODL	1ODJ
III	NdRT	1F8X	1F8Y
IV	Lev	1OYG	1PT2
IV	PAL	1S1A	2ARE
V	LinB	1IZ7	1G5F
V	PGA	1PNK	1AI7
VI	KSI	3VSY	2PZV
VI	PDE4D	3SL3	1Y2B

^aDAHPS: 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; AST: aspartate aminotransferase; 6PGL: 6-phosphogluconolactonase; HAP: histidine acid phosphatase; PNP: purine nucleoside phosphorylase; NdRT: nucleoside 2-deoxyribosyltransferase; Lev: levansucrase; PAL: Pterocarpus angolensis lectin; LinB: haloalkane dehalogenase; LinB; PGA: penicillin G acylase; KSI: ketosteroid isomerase; PDE4D: phosphodiesterase 4D.

(2) 立体構造変化への対応

分子動力学シミュレーションソフトウェア Gromacs 4.0.7 に dual-basin model のポテンシャルエネルギー計算ルーチンを実装した。これを、リガンド結合に伴って立体構造が開いた構造から閉じた構造に変化する ribose binding protein (リガンド非結合構造: 2FN9; リガンド結合構造: 2FN8) と adenylate kinase (リガンド非結合構造: 2RH5; リガンド結合構造: 2RGX) に適用した。Dual-basin model には2つのエネルギー極小状態の間のエネルギー差を制御するパラメータと、エネルギー障壁の高さを制御するパラメータが存在する。リガンド非結合状態とリガンド結合状態の2つの結晶構造を用いて elastic network を構築した後、この2つのパラメータを変えながらシミュレーションを行った。この結果、1 μ s のシミュレーション中に複数回立体構造遷移を起こす条件を見出すことに成功した。

(3) 粗視化シミュレーションの実施と解析

計算対象としたタンパク質(表2)は、いずれもリガンド結合に伴う立体構造変化が小さいため、これらに対するリガンド結合粗視化分子動力学シミュレーションでは、通常の elastic network model を組み合わせることとした。以下に、シミュレーション条件を示す(表3)。

表 3: シミュレーション条件

タンパク質	リガンド ^a	分子数	長さ × 回数
DAHPS	PEP	10	5 μ s × 50
AST	MA	50	2 μ s × 50
6PGL	CA	10	5 μ s × 50
HAP	TA	30	2 μ s × 50
PNP	Gua	10	5 μ s × 50
NdRT	dCyt	15	5 μ s × 50
Lev	sucrose	5	4 μ s × 50
PAL	glucose	5	1 μ s × 100
LinB	1,2-DCE	10	1 μ s × 100
PGA	phenol	30	5 μ s × 50
KSI	phenol	4	2 μ s × 50
PDE4D	EDPC	5	5 μ s × 50

^aPEP: phosphoenolpyruvate; MA: maleic acid; CA: citric acid; TA: tartaric acid; Gua: guanosine; dCyt: 2'-deoxycytidine; 1,2-DCE: 1,2-dichloroethane; EDPC: ethyl 3,5-dimethyl-1H-pyrazole-4-carboxylate.

計算対象のリガンドのうち、glucose と sucrose には MARTINI の糖のパラメータを用いた。Phenol については tyrosine の側鎖のパラメータを用いた。1,2-DCE には、実験値に近い油水分配係数を持つ C4 粗視化粒子を用いた。その他のリガンドについては、上述のプロトコールに従ってパラメータを決定した。

粗視化分子動力学シミュレーションの結果いずれに系においても、実験的に明らかになっているリガンド結合部位への多数回のリガンドの結合が観測された(表4)。

表 4: リガンド結合・解離イベントの回数

タンパク質	結合回数 ^a	解離回数 ^a
DAHPS	127	52
AST	313	262
6PGL	45	0
HAP	769	693
PNP	128	44
NdRT	94	20
Lev	49	18
PAL	1054	1045
LinB	80	60
PGA	22	3
KSI	245	145
PDE4D	111	62

^a リガンド重心とリガンド結合部位との間の距離に基づく結合・解離判定

計算対象のタンパク質・リガンドペアのうち、結合・解離速度定数 k_{on} 、 k_{off} の実験値が得られている Lev について、計算値との比較を行ったところ、実験値 $k_{on} = 6.6 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ 、 $k_{off} = 8.3 \times 10^3 \text{ s}^{-1}$ に対して計算値は、 $k_{on} = 7.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ 、 $k_{off} = 6.5 \times 10^4 \text{ s}^{-1}$ となった。計算値は実験値より 8~11 倍大きい。粗視化モデルでは物質が速く拡散することが知られているため、これを考慮すれば粗視化分子動力学シミュレーションは、リガンド結合の速度論をよく再現しているといえる。

次いで、解離定数の実験値が得られているタンパク質・リガンドペアについて計算値との比較を行った（表5）。

表 5: 解離定数の実験値と計算値の比較

タンパク質	実験値 [mM]	計算値 [mM]
AST	19	21.7
HAP	0.2	16.3
PNP	0.8	2.97
Lev	13	8.80
PAL	0.53	98.1
LinB	2.31	196
PGA	0.82	17.9
KSI	0.135	0.187
PDE4D	0.082	0.282

AST、PNP、Lev、KSI、PDE4D については、実験値と計算値の比が4倍以内となり、粗視化分子動力学シミュレーションの結果は実験値をよく再現しているといえる。一方、下線を付した、HAP、PAL、LinB、PGAについては、計算値は実験値より10倍以上大きな値となった。これらの系について結果を再検討したところ、HAPとPALについては、粗視化モデルでリガンド結合状態の結晶構造を再現すると、リガンドの粗視化粒子がタンパク質の粗視化粒子と衝突してしまうことが明らかとなった。また、LinBについては、リガンドとタンパク質の間で、粗視化モデルでは考慮されていないハロゲン結合を形成していることが明らかとなった。このためこれらについて粗視化モデルの力場パラメータを調整し、再度リガンド結合シミュレーションを実施した。その結果、いずれの系においても、計算値は実験値の1.3~6倍程度に改善した（表6）。

表 6: パラメータ調整後の解離定数の計算値

タンパク質	実験値 [mM]	計算値 [mM]
HAP	0.2	0.273
PAL	0.53	3.19
LinB	2.31	3.02

PGAについては、表4にあるように解離イベントのサンプル数が少ないため十分な精度で解離定数を求めることができなかったと思われる。この例のように、比較的親和性が高いリガンドでは、より長時間のシミュレーションを行うか、リガンドが解離しやすくするポテンシャルをかけ、後で補正する方法を用いる必要があると考えられる。

次いで、リガンドの流束の計算を行い、リガンド結合パスウェイの解析を行った。図1にLevにおけるリガンドの流束を示す。いずれのタンパク質・リガンドペアにおいても、リガンドはタンパク質表面上の特定のパスウェイに沿って移動し、リガンド結合ポケットに入ることが明らかとなった。IおよびIIの分類（リガンドとタンパク質の間に静電相互作用がある場合）では、リガンドが、タンパク質表面上のリガンドと反対の電荷を持つアミノ酸側鎖付近を通る傾向があり、その他の分類では、リガンドはタンパク質表面上の

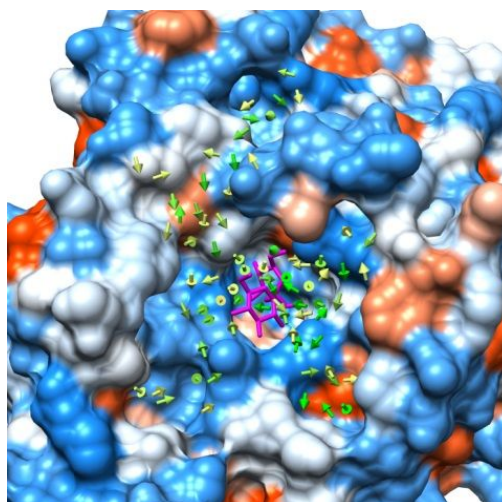


図 1: Lev におけるリガンドの流束

(4) マルチスケールシミュレーション
Levの系について、図1のリガンドの流束に基づいてリガンド結合パスウェイ構築し、string法を用いて精密化を行った。リガンド結合パスウェイは、20 nsのシミュレーションでほぼ収束した。図2に最適化前と最適化後のリガンド結合パスウェイを示す。

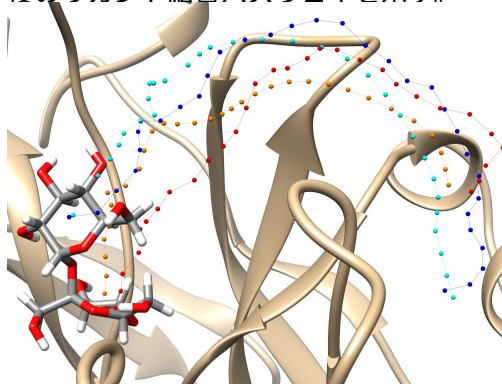


図 2: 最適化前（橙・水色）と最適化後（赤・青色）のリガンド結合パスウェイ

最適化後もパスウェイはタンパク質表面上の同じ溝の中にとどまっていることから、粗視化分子動力学シミュレーションによって予測されたリガンド結合パスウェイは十分な精度を持っていたといえる。次いで、最適後のパスウェイに沿って自由エネルギー地形を求めた（図3）。

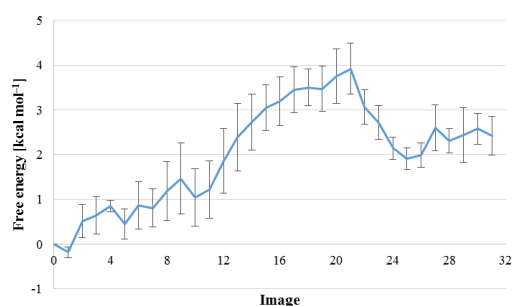


図 3: 最適化後のリガンド結合パスウェイに沿った自由エネルギー地形

ここから、イメージ21からイメージ0（リガ

ンド結合部位)に至るまで、パス上には大きな自由エネルギー障壁はなく、リガンドをリガンド結合ポケットに導く自由エネルギー勾配が存在することが明らかになった。粗視化分子動力学シミュレーションの結果、リガンド結合パスウェイはタンパク質表面上の溝に沿う傾向があることが示されているが、タンパク質表面上の溝は、リガンドをリガンド結合ポケットに向けてスムーズに流す役割を担っていると考えられる。

5. 主な発表論文等

〔雑誌論文〕(計1件)

Tatsuki Negami, Kentaro Shimizu, and Tohru Terada, “Coarse-Grained Molecular Dynamics Simulations of Protein–Ligand Binding” *Journal of Computational Chemistry*, 査読有, vol. 35, pp. 1835–1845 (2014), DOI: 10.1002/jcc.23693.

〔学会発表〕(計5件)

Tohru Terada, Tatsuki Negami, Kentaro Shimizu, “Protein–ligand binding processes studied by coarse-grained molecular dynamics simulations” 日本生物物理学会第52回年会, 2014年9月25日~2014年9月27日, 札幌コンベンションセンター, 北海道.

Tatsuki Negami, Tohru Terada, Kentaro Shimizu, “The factors determining protein–ligand binding processes revealed by comparative coarse-grained simulations” 日本生物物理学会第52回年会, 2014年9月25日~2014年9月27日, 札幌コンベンションセンター, 北海道.

Tatsuki Negami, Kentaro Shimizu, Tohru Terada, “Protein–ligand binding simulation with the MARTINI coarse-grained force field” Biophysical Society 58th Annual Meeting, 2014年2月15日~2014年2月19日, Moscone Center, San Francisco, California, USA.

Tatsuki Negami, Tohru Terada, Kentaro Shimizu, “Comparative simulations of protein–ligand binding processes using MARTINI coarse-grained force field” 日本生物物理学会第51回年会, 2013年10月28日~2013年10月30日, 国立京都国際会館, 京都府.

Tatsuki Negami, Tohru Terada, Kentaro Shimizu, “Coarse-grained MD simulations of ligand binding to proteins: Effect of the conformational changes of the proteins” 日本生物物理

学会第50回年会, 2012年9月22日~2012年9月24日, 名古屋大学, 愛知県.

〔その他〕
ホームページ等
<http://www.iu.a.u-tokyo.ac.jp/~tterada/>

6. 研究組織

(1) 研究代表者

寺田 透 (TERADA, Tohru)
東京大学・大学院農学生命科学研究科・准教授
研究者番号: 40359641

(2) 研究協力者

根上 樹 (NEGAMI, Tatsuki)