

## ウリジン二リン酸グルクロン酸転移酵素 1A1 の立体構造予測

小田 彰史\*, 小林 佳奈, 高橋 央宜

### Prediction of Three Dimensional Structure of Uridine Diphosphate Glucuronosyltransferase 1A1

Akifumi ODA, Kana KOBAYASHI, and Ohgi TAKAHASHI

(Received November 20, 2008)

In this study, we construct the three dimensional structure of Uridine Diphosphate Glucuronosyltransferase (UGT) 1A1 by using protein homology/analogy recognition engine (phyre) server. Ten PDB entries including UGT2B7 were used for template of the model, and both UDP-glucuronic acid binding domain and substrate binding domain were obtained. By using the modeled structure, structural refinement was carried out by molecular dynamics (MD) simulations. In order to investigate the role of cysteine residues, MD simulations of both UGT1A1 models with and without the Cys127-Cys156 disulfide bond were performed. As the results of the simulations, structural features of UGT1A1 can be investigated, and refined structure of it was obtained.

**Key words** — Uridine Diphosphate Glucuronosyltransferase; Structural Prediction; Drug Metabolism

ウリジン二リン酸グルクロン酸転移酵素 (UDP-グルクロン酸転移酵素, UGT) は内在性物質・外来性物質の両方の代謝に関与する代謝酵素である。<sup>1)</sup> UGT は主に肝細胞などの小胞体膜に局在しており, 補酵素として UDP-グルクロン酸 (UDPGA) を含んでいる。UGT は様々な化合物にグルクロン酸を付加するグルクロン酸抱合反応を触媒する。グルクロン酸抱合を受ける化合物は通常ヒドロキシル基, カルボキシル基, アミノ基, チオール基などを持っており, これら官能基にグルクロン酸が付加する。たとえばヒドロキシル基を持つ化合物へのグルクロン酸抱合は以下のように起こる。



グルクロン酸は水溶性が高いため, グルクロン酸抱合を受けた化合物は水溶性が上昇し, 代謝されやすくなる。

UGT には複数のアイソザイムがあるが, ヒトにおいては UGT1 ファミリーと UGT2 ファミリーの2つのファミリーに分類される。<sup>2)</sup> いずれのファミリーに属するアイソザイムも 500 残基程度のアミノ酸残基から構成されており, その C 末端側残基はアイソザイム間で比較的高い相同性を保っている一方で, N 末端領域には多様性がある。C 末端側に補酵素である UDPGA が結合し, 相同性の

低い N 末端領域が様々な基質を認識する部位であると考えられている。これら 1 次構造に起因して各アイソザイムは異なる基質特異性を示すが, 一般に UGT1 ファミリーのほうが UGT2 ファミリーと比較して幅広い基質特異性を持つ。薬物およびその関連化合物では, 抗がん剤 CTP-11 (イリノテカン) の活性代謝物 SN-38 の代謝においては主に UGT1 ファミリーに属する UGT1A1 および UGT1A7 が関与し, アセトアミノフェンは主に UGT1A6 による抱合を受ける。モルフィンや 3'-アジド-3'-デオキシチミジンのグルクロン酸抱合は主に UGT2 ファミリーに属するアイソザイムである UGT2B7 によって触媒される。

UGT には様々な遺伝多型が存在することが知られており, それら遺伝多型が疾患や医薬品の薬物動態における個人差の原因となっていることが知られている。UGT1 ファミリーにおけるアミノ酸置換や欠損によって起こる Crigler-Najjar I 型, II 型, Gilbert's 症候群といったビリルビンの代謝異常がその代表例である。この代謝異常は特に UGT1A1 のみのアミノ酸変異や欠損でも発生することが知られており, UGT1A1 が特に重要な役割を果たしている。<sup>3)</sup> また, UGT1A1 の遺伝多型は日本人の新生児黄疸にも検出されており,<sup>4)</sup> 乳がんの発症率への影響なども指摘されている。<sup>5)</sup> さらに UGT1A1 は前述の通り抗がん剤 CTP-11 の活性代謝物 SN-38 など

の代謝に対しても触媒として働くため, UGT1A1 の遺伝多型は薬物代謝に対しても影響を及ぼす。

このように UGT は内在性化合物・外因性化合物両方の代謝において重要な役割を果たしており, 遺伝多型の影響を含めてその構造的特徴を解明することが非常に重要であるが, 膜結合タンパク質で結晶化が困難であるため, その立体構造はほとんど実験的に解明されていない。ヒトの UGT については唯一 UGT2B7 のみリガンドを含まない構造が実験によって解析されているが, これについても C 末端側の UDPGA 結合部位のみについて解かれており, N 末端側の基質認識部位については不明のままである。<sup>6)</sup> また, 遺伝多型が重要となる UGT1A1 はもちろんのこと, UGT1 ファミリーのアイソザイムについては全く立体構造が解明されていない。そのため, 計算機的に UGT の 3 次元構造的特徴を解明する研究がなされており, タンパク質立体構造予測<sup>7)</sup> や 3 次元定量的構造活性相関 (3D-QSAR)<sup>8-10)</sup> によってモデル化が行われてきた。これら計算機的研究によって UGT の立体構造的特徴はある程度解明されてはいるものの, 3D-QSAR ではタンパク質の構造まで予測されているわけではない。またタンパク質立体構造予測についても, 文献 7 ではテンプレートとして植物の UGT が使用されており, 近年解明された UGT2B7 の構造は使用されていない。

そこで本研究では, UGT1A1 の立体構造の予測を行う。その際には UGT2B7 の構造も使用し, より高精度なモデルの構築を行う。代謝酵素については, UGT に限らず複数の分子種を持つものが多く, かつそれらに遺伝多型が存在しており, 多数のアイソザイムを考慮する必要がある場合が多い。これらすべてに対して実験的に立体構造を解析するのは困難であるため, 計算機的に構造予測や構造精密化を行うことは重要である。これまで立体構造未知のタンパク質の立体構造を予測する手法は様々に開発されているが, 近年では複数の手法を組み合わせ, それらの長所をあわせることのできるメタサーバが優れた結果を出している。<sup>11)</sup> そこで本研究でもメタサーバを使用し, UGT1A1 の立体構造を予測した。UGT1 ファミリーの UGT1A6 については Cys121 と Cys125 の間のジスルフィド結合が活性発現に重要であることが報告されているが,<sup>12)</sup> UGT1A1 において UGT1A6 の Cys125 に対応する Cys127 がジスルフィド結合に関与するか

否かは活性発現において重要ではないことが報告されている。<sup>13)</sup> そこで UGT1A1 の Cys127 が予測構造中で他のシステイン残基と近接しているかどうか, また近接しているのであればジスルフィド結合の有無が構造に対して影響を与えるかどうか, 分子モデリングおよび分子シミュレーションを行って検討した。

## 方 法

本研究では UGT1A1 の 1 次構造から立体構造を予測したが, その際にはフォールド認識法を用いたメタサーバである protein homology/analogy recognition engine (phyre)<sup>14,15)</sup> を使用した。phyre では 2 次構造予測から立体構造構築まですべて自動で行い, テンプレートとなるタンパク質構造についてもプログラム自身が同定するが, UGT1A1 に対しては 2IYA, 2IYF, 2PQ6, 2C1Z, 2ACV, 2VCU, 1IIR, 1RRV, 2O6L, 2P6P がテンプレートとして使用された。これらはいずれも RCSB Protein DataBank (PDB)<sup>16)</sup> に収載されている構造である。また, PDB ID 2O6L が UGT2B7 の C 端側構造にあたり, 他の構造もすべて転移酵素の立体構造である。このようにして得られた構造に対して, Cys127 に関与したジスルフィド結合を繋いだモデルと繋がずにチオール体のままとしたモデルの 2 種類を作成した。また, 本研究でテンプレートとして使用した PDB 収載の構造は X 線回折によって解明された構造であるため, 得られた構造には水素原子が付加されていないが, ジスルフィド結合に関するモデリングの際に水素原子についても付加した。これらの操作は AMBER9<sup>17)</sup> の Leap モジュールを使用して行った。

このようにして作成したモデルに対して, 分子力学計算による構造最適化および分子動力学 (MD) シミュレーションによる構造精密化を行った。これらの計算はすべて水分子の箱中で行っている。水分子を付加するため, タンパク質の表面から最低でも 8Å の幅を持たせることができる大きさの立方体の箱を用意した。水分子のモデルとしては TIP3P を使用した。また, 系全体で電氣的に中性となるように, ナトリウムイオンを付加した。分子力学計算, MD シミュレーションともに古典的分子力場を使用しているが, 本研究では AMBER の ff99SB 力場を使用した。構造最適化については, まず付加した

水の最適化のみを行った後、系全体の最適化を行った。前者は最大 5000 サイクルでエネルギーの傾きが  $0.1 \text{ kcal/mol \AA}$  以下になれば終了するように設定し、後者については最大 10000 サイクルでエネルギーの傾きについてはデフォルトの条件  $1.0 \times 10^{-4} \text{ kcal/mol \AA}$  を使用した。また、最適化の段階では定積下で計算を行った。最適化された構造に対して、まず系の温度を上昇させる MD シミュレーションを行った。30 ps かけて系の温度を 0 K から 300 K まで上昇させた。この際、水分子のみを運動させ、タイムステップは 0.5 fs を使用した。温度が上昇した後、300 K で平衡化のための MD シミュレーションを行った。タイムステップは 1 fs で、2 ns のシミュレーションを実行した。昇温 MD

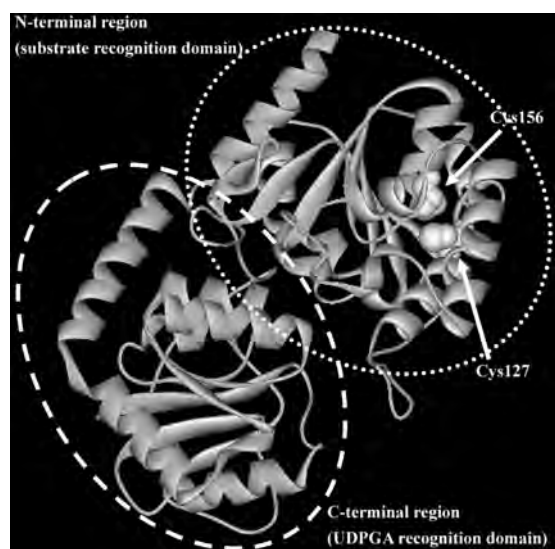


Fig. 1. Predicted structure of UGT1A1 by using phyre.

は定積条件下で、平衡化 MD は定圧条件下で計算を行った。また MD シミュレーションにおいては、水素原子を含む結合に対して SHAKE 法を適用し、相互作用の計算を節約した。構造最適化、MD シミュレーションともに周期境界条件を使用し、静電項の計算には particle mesh Ewald 法を用いた。van der Waals 項の計算には  $10 \text{ \AA}$  のカットオフを使用した。構造最適化および MD シミュレーションは AMBER9 プログラム<sup>17)</sup> を用いた。ハードウェアとしては Xeon X5482 プロセッサ 2 基を CPU とし、メモリを 16 GB 搭載した Linux マシンを使用して計算を行った。

得られた構造に対してリガンド結合部位探索を行い、構築したモデルの構造的特徴を検討した。リガンド結合部位探索には Q-SiteFinder<sup>18)</sup> の web バージョン<sup>19)</sup> を使用した。Q-SiteFinder はそれ自身に水素を付加するルーチンが組み込まれているため、水素を外して計算を行った。リガンド結合部位探索については、phyre によって得られた予測構造そのもの、ジスルフィド結合を形成して MD を行った後の構造、ジスルフィド結合を形成せずに MD を行った後の構造の 3 つすべてに対して実行した。

## 結果および考察

Fig. 1 に、phyre によって予測された UGT1A1 の構造を示す。アミノ酸配列が比較的保存されている C 末端領域の UDPGA 認識領域だけではなく、N 末端領域についても予測がなされている。N 末端

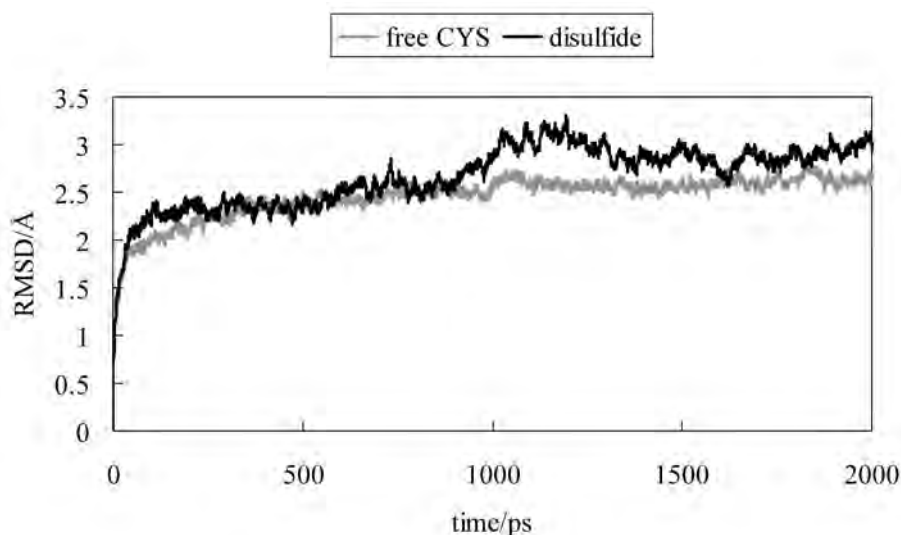


Fig. 2. RMSD during MD simulations.



側の 28 残基についてはフォールド認識法によっても構造を決定することができなかったが、活性発現において重要とされる残基<sup>13)</sup>の周辺については構造が得られている。また、N 末端側ドメインと C 末端側ドメインが接しており、基質と補酵素が反応に関与してグルクロン酸抱合が起こるという UGT の機能と矛盾のない構造となっている。

文献 13 でジスルフィド結合について論じられている Cys127 について、予測構造中での位置を Fig. 1 に示している。図に示したように、Cys127 は Cys156 と近接した位置にあり、2 つの残基の硫黄原子間距離は 5.84 Å となっている。この距離はジスルフィド結合を行うには遠いものの、もし Cys127 がジスルフィド結合に関与するとすれば、結合の相手となる残基としては Cys156 の可能性が高いのではないかと考えられる。そこで、Cys127 と Cys156 の間にジスルフィド結合のあるモデルと

ないモデルを作成し、それら両方に対して MD シミュレーションを行った。

Fig. 2 に、MD シミュレーションにおける偏差二乗平均平方根 (RMSD) の変化を図示した。灰色で示した線がジスルフィド結合のないモデル、黒で示した線がジスルフィド結合のあるモデルでの RMSD を表している。また、RMSD の比較対象としては MD シミュレーションの初期構造を用いた。ここに示したように、いずれのモデルにおいても 2 ns のシミュレーションによって構造が収束していると考えられる。また、両者ともにそれほど大きく構造が崩壊していないものの、ジスルフィド結合のあるモデルのほうがわずかながら RMSD が大きくなる傾向があった。

Table 1 に、Q-SiteFinder で発見されたりガンド結合ポケット候補を示した。ここではポケットを構成する残基を列挙しており、サイトの体積につ

Table 1. Predicted ligand binding sites.

	rank of site	Residues included in sites	Vol/Å <sup>3</sup>
phyre structure	Site 1 (N-terminal)	Thr150, Asp151, Pro152, Phe153, Leu154, Pro155, Pro158, Phe170, Phe171, Leu172, Leu175, Pro176, Cys177, Ser178, Glu182, Tyr230, Asp246, Leu247, Leu248, Ser249, Ser250, Ala251, Trp254, Phe256, Ile268, Met269, Pro270, Asn271, Met272	595
	Site 2 (N-terminal)	Ile33, Pro34, Val35, Asp36, Pro62, Tyr79, Pro80, Val81, Pro82, Phe83, Gln84, Met122, Leu123, Leu124, Ser125, Gly126, Cys127, Ser128, His129, Leu130, Leu131, Leu136, Phe153, Leu154, Pro155, Cys156	357
	Site 3 (inter-domain)	Ser38, Leu41, Asn279, His282, Ser306, Leu307, Gly308, Arg336, Tyr337, Trp354, Leu355, Pro356, Gln357, Asn358, Asp359, Phe369, Thr371, His372, Gly374, His376, Gly377, Glu380	326
without disulfide	Site 1 (N-terminal)	Leu123, Leu124, Pro152, Phe153, Pro155, Phe170, Phe171, Leu172, His173, Ala174, Leu175, Pro176, Cys177, Ser178, Glu182, Phe190, Leu247, Leu248, Ser249, Ser250, Ala251, Trp254, Phe256, Ile268, Met269, Met272	513
	Site 2 (inter-domain)	Phe83, Gln84, Thr112, Lys114, Lys115, Ile116, Asp119, Ser120, Pro187, Asn188, Pro189, Met310, Leu393, Phe394, Gly395, Asp396	334
	Site 3 (N-terminal)	Pro158, Ile159, Ala161, Gln162, Ser165, Leu166, Pro167, Thr168, Val169, Phe170, Arg240, Asp246, Leu247, Ser250, Ala251, Ser252, Val253	261
with disulfide	Site 1 (inter-domain)	Pro34, Val35, Asp36, Gly37, Ser38, Trp40, Leu41, Pro62, Ala64, Ser65, Leu66, Tyr67, Ile68, Glu86, His282, Gly308, Ser309, Arg336, Lys353, Trp354, Leu355, Gln357, His376	422
	Site 2 (inter-domain)	Leu175, Pro176, Cys177, Glu182, Trp254, Ser258, Asp259, Phe260, Asp263, Pro265, Arg266, Ile268, Phe274, Tyr379, Asp396, Asp399, Asn400, Arg403, Met404, Lys407	386
	Site 3 (inter-domain)	Asp36, His39, Phe83, Gln84, Arg85, Glu86, Asp87, Gln107, Arg108, Ile110, Lys111, Thr112, Ser309, Met310, Val311, Ile314, Arg336, Tyr337, Lys353	452

いても示している。また、リガンド結合部位となる可能性の高い上位3位までの予測ポケットを記載した。表中の Site 1 が最もリガンド結合サイトである可能性の高い予測部位、Site 2 が2番目に可能性の高い部位、Site 3 が3番目に可能性の高い部位である。サイト名の後のカッコ内に、N末端ドメイン、C末端ドメイン、ドメイン間のいずれの領域にそのサイトがあるかを示している。また、“phyre structure”に phyre によって得られた構造をそのまま Q-SiteFinder にかけた結果を示し、“without disulfide”にジスルフィド結合を作らずに MD を行った後の構造を用いた結果を示し、“with disulfide”にジスルフィド結合を作った MD を行った後の構造を用いた結果を示している。ここに示したように、phyre による予測構造とジスルフィド結合なしのモデルでは N 末端側と、N 末端ドメインと C 末端ドメインの境界領域の2つの領域に結合サイト候補が発見されているのに対して、ジスルフィド結合ありのモデルではドメイン間の境界領域にのみ結合サイト候補が見いだされている。またサイトの体積を見ても、最もリガンド結合サイトである可能性の高いサイト1の体積が phyre による予測構造やジスルフィド結合なしのモデルと比較してジスルフィド結合ありのモデルでは小さくなっている。これらの結果は、予測構造にジスルフィド結合を加えた結果、構造が大きく変化したことを意味している。特にサイトの体積が狭くなっていることから、ジスルフィド結合がないモデルのほうがリガンドの結合において有利であることが示唆される。上述の RMSD の結果と併せて、少なくとも本研究で phyre を用いて作成したモデルに関しては、ジスルフィド結合を形成しない構造が UGT1A1 のモデルとして妥当ではないかと推測される。この結果は Cys127 のジスルフィド結合が UGT1A1 の活性発現に重要ではないという実験結果<sup>13)</sup>と対応している。

これは Cys127 のみに限った検討であるが、文献13においては Cys127 のみならずすべてのシステインに対して、UGT1A1 の活性発現において分子内のジスルフィド結合が重要ではないという実験結果が得られている。今回作成したモデルにおいては、Cys127 と Cys156 以外に 10 Å 以内に2つのシステイン残基の硫黄原子が近接したペアは存在せず、Cys127 と Cys156 についても上述の通り 5.86 Å の距離であった。すべてのシステインのペ

アがジスルフィド結合可能な距離にないという計算結果は、上述の実験結果を説明しうるものとなっており、本研究で作成したモデルの妥当性を間接的に示唆しているのではないかと考えられる。今後の展開としては、今回作成したモデルに対して UGT の補酵素である UDPGA や SN-38 などの基質を計算機的にドッキングし、構造を精密化することで、UGT1A1 のリガンド認識機構について検討する予定である。

## 謝辞

本研究は科研費（19850019）の助成を受けたものである。

## REFERENCES

- 1) Radomska-Pandya A., Czernik P. J., Little J. M., Battaglia E., Mackenzie P. I., *Drug Metab. Rev.*, **31**, 817-899 (1999).
- 2) Mackenzie P. I., Owens I. S., Burchell B., Bock K. W., Bairoch A., Bélanger A., Fournel-Gigleux S., Green M., Hum D. W., Iyanagi T., Lancet D., Louisot P., Magdalou J., Chowdhury J. R., Ritter J. K., Schachter H., Tephly T. R., Tipton K. F., Nebert D. W., *Pharmacogenetics*, **7**, 255-269 (1997).
- 3) Monaghan G., Ryan M., Seddon R., Hume R., Burchell B., *Lancet*, **347**, 578-581 (1996).
- 4) 赤羽和博, 早坂 清, 生体の化学, **50**, 449-450 (1999).
- 5) Guillemette C., Millikan R. C., Newman B., Housman D. E., *Cancer Res.*, **60**, 950-956 (2000).
- 6) Miley M. J., Zielinska A. K., Keenan J. E., Bratton S. M., Radomska-Pandya A., Redinbo M. R., *J. Mol. Biol.*, **369**, 498-511 (2007).
- 7) Locuson C. W., Tracy T. S., *Xenobiotica*, **37**, 155-168 (2007).
- 8) Sorich M. J., Smith P. A., McKinnon R. A., Miners J. O., *Pharmacogenetics*, **12**, 635-645 (2002).
- 9) Ethell B. T., Ekins S., Wang J., Burchell B., *Drug Metab. Dispos.*, **30**, 734-738 (2002).
- 10) Smith P. A., Sorich M. J., Low L. S. C., McKinnon R. A., Miners J. O., *J. Mol. Graph. Model.*, **22**, 507-517 (2004).
- 11) 朴 聖俊, 千見寺浄慈, 広川貴次, 富井健太郎, 高田彰二, 人工知能学会論文誌, **20**, 0512 (2005).
- 12) Ikushiro S., Emi Yoshikazu, Iyanagi T., *Biochemistry*,

- 41**, 12813-12820 (2002).
- 13) Ghosh S. S., Lu Y., Lee S. W., Wang X., Guha C., Roy-Chowdhury J., Roy-Chowdhury N., *Biochem. J.*, **392**, 685-692 (2005).
- 14) Bennett-Lovsey R. M., Herbert A. D., Sternberg M. J. E., Kelley L. A., *Proteins*, **70**, 611-625 (2008).
- 15) <<http://www.sbg.bio.ic.ac.uk/~phyre/>>, Imperial College Phyre Server.
- 16) Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E., *Nucleic Acids Res.*, **28**, 235-242 (2000).
- 17) Case D. A., Darden T. A., Cheatham T. E., III, Simmerling C. L., Wang J., Duke R. E., Luo R., Merz K. M., Pearlman D. A., Crowley M., Walker R. C., Zhang W., Wang B., Hayik S., Roitberg A., Seabra G., Wong K. F., Paesani F., Wu X., Brozell S., Tsui V., Gohlke H., Yang L., Tan C., Mongan J., Hornak V., Cui G., Beroza P., Mathews D. H., Schafmeister C., Ross W. S., Kollman P. A., AMBER9, University of California, San Francisco, 2006.
- 18) Laurie A. T. R., Jackson R. M., *Bioinformatics*, **21**, 1908-1916 (2005).
- 19) Laurie A. T. R., Jackson R. M., <<http://bmbpcu36.leeds.ac.uk/qsitefinder/>>, Q-SiteFinder: Ligand Binding Site Prediction Server.