

MolDesk Screening ver.1.1

インシリコ創薬のためのドラッグデザイン パッケージ ソフトウェア www.moldesk.com

普通の8コアCPU搭載PCを使って **0.5 ~ 2 日**で **200万 化合物**をバーチャルスクリーニング (MolDesk Basicのすべての機能が使えます)

バーチャル スクリーニング (並列計算)

構造に基づくVirtual Screening(VS)のリランキング法として、*myPresto*^{※1}が提供するMultiple Target Screening(MTS)法および機械学習MTS法、リガンドに基づくVSのリランキング法として、同じく機械学習Docking Score Index(DSI)法が利用できます。

LigandBox^{※2}より創薬向けに抽出した200万化合物ライブラリに対して、スレッド並列で高速にスクリーニングを行います。

ユーザが用意したインハウス化合物ライブラリ (2次元SDF形式、数百~数百万分子) をスクリーニング対象とすることができます。

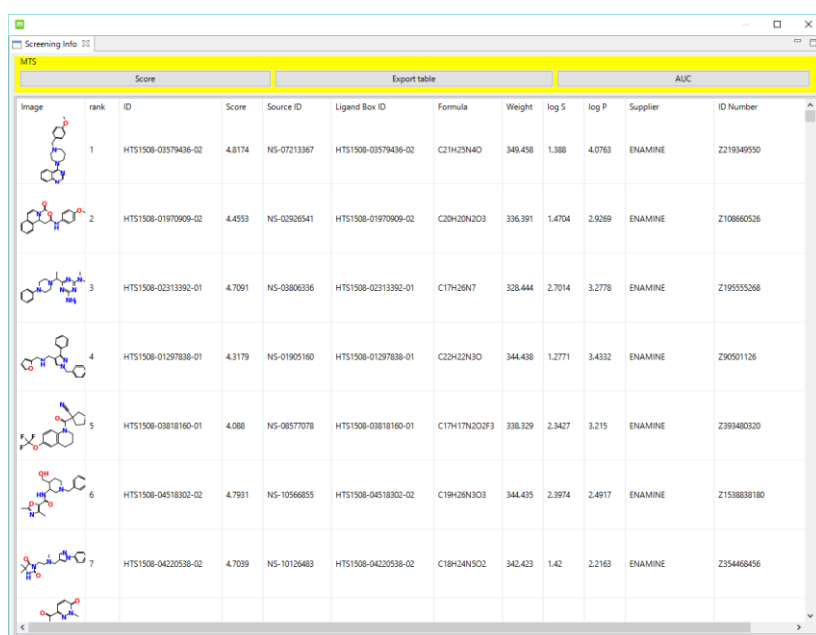









Image	rank	ID	Score	Source ID	Ligand Box ID	Formula	Weight	log S	log P	Supplier	ID Number
	1	HTS1508-03579436-02	4.8174	NS-07213367	HTS1508-03579436-02	C21H25N4O	349.458	1.388	4.0763	ENAMINE	Z219349550
	2	HTS1508-01970909-02	4.4553	NS-03926541	HTS1508-01970909-02	C20H20N2O3	336.391	1.4704	2.9269	ENAMINE	Z108660526
	3	HTS1508-02313392-01	4.7091	NS-03806036	HTS1508-02313392-01	C17H26N7	328.444	2.7014	3.2778	ENAMINE	Z195555268
	4	HTS1508-01297838-01	4.3179	NS-01905160	HTS1508-01297838-01	C22H22N3O	344.438	1.2771	3.4332	ENAMINE	Z90001126
	5	HTS1508-03818160-01	4.088	NS-08570778	HTS1508-03818160-01	C17H17N2O2F3	338.329	2.3427	3.215	ENAMINE	Z393480320
	6	HTS1508-04518202-02	4.7931	NS-10566855	HTS1508-04518202-02	C19H26N3O3	344.425	2.3874	2.4917	ENAMINE	Z1538838180
	7	HTS1508-04220538-02	4.7039	NS-10126483	HTS1508-04220538-02	C18H24N5O2	342.423	1.42	2.2163	ENAMINE	Z354468456

図1. MTSスクリーニング結果リスト

リストには化合物の2次元構造式、ランク、ID、スコア、化学式、分子量、LogS、LogP、電荷、ドナー数、アクセプター数、HOMO、LUMO、Chiral原子数、Source、Source IDが表示されます。各項目についてソート可能です。

リストをCSVまたはHTML形式でファイル出力し、これらをMicrosoft Excelに取り込んで化合物の発注に利用できます。

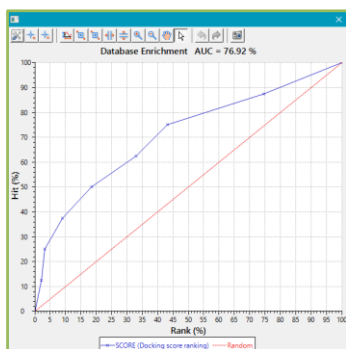


図3. Database enrichment 曲線

スクリーニング結果は指定リランキング順にリストされ、選択した化合物のドッキングポーズも確認できます。

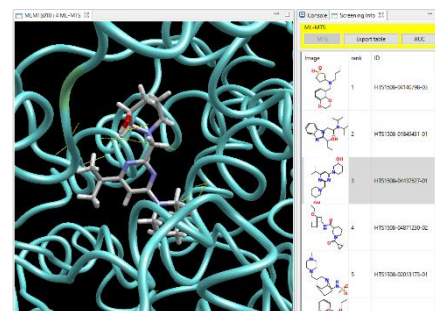


図2. ドッキングポーズ表示

既知活性リガンドをスクリーニング対象に加えて、計算精度の検証ができます。計算精度は、Database enrichment曲線を表示してAUC (Area Under the Curve)によって確認します。

インハウスの化合物ライブラリ (2次元SDF形式) を対象としてスクリーニングを行う場合、その前に各分子の立体構造を構築します。各分子の立体構造構築では、キラリティーを考慮した分子配座が生成されると共に、指定範囲外の分子量を持つ分子および不適当な構造 (例えば反応性の高い構造) を持つ分子のフィルタリングが行われます。

※1 *myPresto*は、AMED・経済産業省・NEDOの支援によって産業技術総合研究所・JBIC・大阪大学蛋白質研究所等によって開発された一般社団法人バイオ産業情報化コンソーシアム (JBIC) のソフトウェアです。

※2 *LigandBox*は、AMED・経済産業省・NEDOの支援によって産業技術総合研究所・JBIC・大阪大学蛋白質研究所等によって*myPresto*の一部として開発された一般社団法人バイオ産業情報化コンソーシアム (JBIC) のデータベースです。

MVO Screening (並列計算)

Image	ID	S_MVO-D	S_MVO-Q	S_MVO-T
	Lig_es_1.mol2	-1.0	-1.0	-1.0
	Lig_es_1.mol2	-0.9972	-0.9917	-0.989
	rofecoxib.M	-0.7668	-0.8197	-0.6561
	Lig_es_1.mol2	-0.6311	-0.7549	-0.5238
	Lig_es_1.mol2	-0.7371	-0.6313	-0.5153
	2qwk.mol2	-0.5445	-0.7083	-0.4448
	lig.mol2	-0.6519	-0.5707	-0.4374

図4. MVO-screening result 結果

MVO(Maximum Volume Overlap)-screening法は、myPrestoのリガンドに基づくスクリーニング法の一つです。

MVO-screening法は、クエリ分子とDB分子の重ね合わせを行い、体積の重なり割合(%)を計算すると共に、併せて分子配座を計算します。

計算されるMVO-screening法のスコアには、クエリ分子とDB分子の重ね合わせた体積が、クエリ分子の体積に占める比率 S_{MVO-Q} ・DB分子の体積に占める比率 S_{MVO-D} ・谷本インデックスに相当する比率 S_{MVO-T} の三種類があります。各スコアの定義は、下表の通りです。

Score Type	Definition	Notes
S_{MVO-D}	$Q \cap D / D \cap D$	For smaller-size molecules selection.
S_{MVO-Q}	$Q \cap D / Q \cap Q$	For larger-size molecules selection.
S_{MVO-T}	$-1.0 * Q \cap D / Q \cup D$	For near-query-molecule-size molecules selection.

表1. MVO-screening スコア

Q = Query molecule's volume
D = Database molecule's volume

Docking-score QSAR (並列計算)

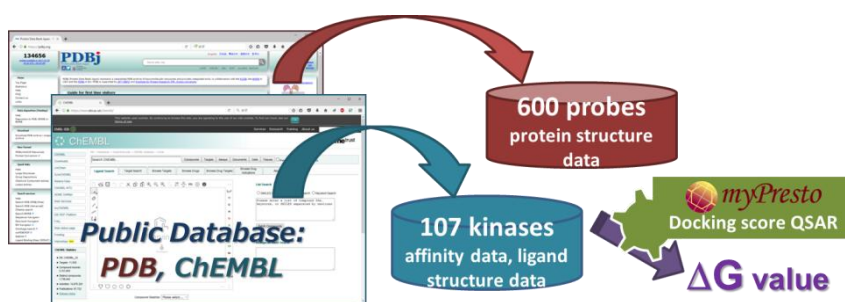


図5. Docking-score QSAR法

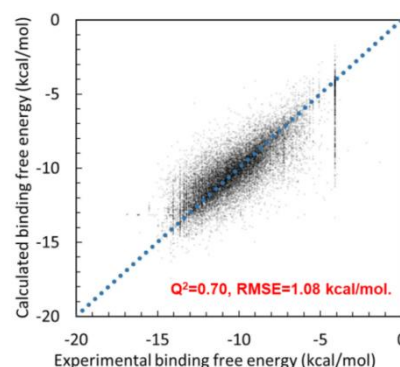


図6. 107キナーゼ予測結果

myPrestoのdocking-score QSAR法を利用して、化合物の結合自由エネルギー ΔG を予測します。Docking-score QSAR法は、600プローブ蛋白質に対するドッキングスコアを記述子として回帰分析を行い、化合物の ΔG を推算する手法です。

回帰に必要な ΔG およびプローブ蛋白質構造は、公的データベース (ChEMBLおよびPDB) より取得します。なお、 ΔG はChEMBLから得た親和性データ (IC_{50} 値, %阻害値, 活性値) から換算したものです。回帰モデルは、正規化項を伴った記述子ベースの重み付きPCRで構築しており、同時にロバスト推定 (M推定) とL2正規化を使用して、外れ値の影響を抑制しています。

例えば、107キナーゼに対する適用においては、 $Q^2 = 0.70$ ($RMSE = 1.08 \text{ kcal/mol}$) の予測精度が得られています (図6.)。

MolDesk Basic / MolDesk Screening 機能一覧

機能		MolDesk Basic	MolDesk Screening
動作OS	Windows 11 / 10 (64bit), Linux (64bit) macOS ver.10.11 以上 (MacBook Air (11-inch, Early 2014)以上)	○	○
入力ファイル	mmCIF pdb mol2 sdf mol SMILES (mol2 / SMILES は multi も可) mmCIF (pdb) と 化合物ファイル はインターネット入力可	○	○
出力ファイル	mmCIF pdb mol2 tplファイル MD 計算設定ファイル ドッキング計算設定 ファイル MD計算、ドッキング計算結果ファイルなど (以上、myPresto 仕 様) 3D画面の png 画像ファイル (ピクセルサイズ変更可能)	○	○
外部分子入力	すでにある系に、任意の mmCIF pdb mol2 sdf mol SMILES (mol2 / SMILES は multi も可) をマウスクリック点、またはファイル座標で読み込み 可	○	○
化合物 (リガンド) 編集	2Dエディター (JChemPaint)、2D化学構造と合成容易性のリスト表示、 H 原子削除・付加、原子 (複数) 削除・置換・付加、構造置換、 結合削除・次数置換・回転、構造最適化 (Clean Geometry) 電荷計算 (Gasteiger または MOPAC7 AM1 / PM3 / AM1-BCC / RESP)	○	○
蛋白質の編集	MD計算等の準備として、H原子付加・不要部分の削除・末端処理 エピゲノム修飾に使える 40種類のアミノ酸変換	○	○
溶媒・イオン付加	TIP3P群の直方体、球(Cap) を分子からのマージンで生成、 または、マウスクリック点中心に任意サイズで付加 イオン (Na+,K+,Cl-) を中和した生理食塩水濃度、濃度/個数指定などで付加	○	○
脂質2重膜系の自動生成	6種類の脂質分子の任意構成比による脂質2重膜系にタンパク質などを自動配置	○	○
化合物の 2D -> 3D 一 括変換	AMBER GAFF2力場でconformerや光学異性体を考慮して3次元化したmol2 ファイルを作成	○	○
構造最適化 (Clean Geometry)	Amber ff99SB力場、GAFF2力場、マウス選択原子の位置拘束、 タンパク質主鎖の位置拘束、リガンドと周辺だけの高速な構造最適計算	○	○
	構造最適化のMPI/GPU (NVIDIA CUDA) による並列計算	○ (cosgene は不可)	○
MD計算	Amber ff99SB力場、GAFF2力場、マウス選択原子の位置拘束可、 タンパク質主鎖の位置拘束、NVT, NPT, NVE アンサンブル、 クーロン力: FMM (溶媒水が球) / PME (溶媒水が直方体) リスタート機能、再計算時の境界条件自動読み込み、GROMACS使用可能	○	○
	MD計算のMPI/GPU (NVIDIA CUDA) による並列計算	○ (cosgene は不可)	○
MD 計算結果解析	トラジェクトリの動画 (GROMACS, AMBERなども可)、各種エネルギー、温度、 任意2点間原子距離、任意2面角の動画と時間連動した時間変化グラフ GROMACSはほぼすべてのトラジェクトリ解析結果の動画と連動したグラフ表 示	○	○
手動ドッキング計算	タンパク質のポケット (と推定される所) の近くに、リガンドを配置しただけ で、ドッキングポーズの構造最適化 結合自由エネルギー ΔG をリスト表示	○	○
ドッキング計算	ドッキング計算で得られたドッキングポーズと各種特性値をリスト表示 リスト表示で、ドッキングポーズの3D図の $\uparrow\downarrow$ キーによる切り替え ポケット側は、ユーザ指定のタンパク質・核酸・イオン・金属・水分子の複合 体が可 ポケット中心原子のマウス選択、ポケットのプロープ点の削除、 候補構造数の指定、複数の候補構造の構造クラスタリング、 多数のリガンドをドッキングする際の高速計算 (グリッドポテンシャル再利 用)	○	○
溶液 NMR 実験シグナル を使ったドッキング計算	溶液 NMR 実験シグナルによるドッキング構造予測	○	○

MolDesk Basic / MolDesk Screening 機能一覧

機能		MolDesk Basic	MolDesk Screening
ポケット探索	簡易版（幾何学的な手法による探索）	○	○
	高精度版（MolSite法による探索、並列計算）	○	○
バーチャルスクリーニング	SBDS（ターゲット蛋白質が必要）：MTS法、機械学習MTS法 LBDS（ターゲット蛋白質が不要）：機械学習DSI法（全て並列計算）		○
	リランキング順にリスト表示（2D構造式および特性値）、ドッキングポーズ表示、CSV/HTMLファイル(2D構造式画像を含む) 出力		○
	Database enrichment curve による精度解析		○
	インハウス化合物ライブラリ（2D構造式、MOL/SDF形式）取り込み機能（立体構造構築、並列計算）		○
化合物の各種特性値計算	合成容易性	○	○
	Docking-score QSAR 法による活性予測（ChEMBL利用、並列計算）		○
	回帰分析による LogP/LogS/LogD/Papp(膜透過係数)/pKa予測		○
類似・部分構造探索	MVO-screening法による類似構造探索（並列計算）		○
	Topology Graph Similarityによる類似構造探索、部分構造検索		○
プロジェクト管理	新規作成、既存のプロジェクトを開く、コピー、エクスポート、インポート	○	○
UNDO REDO	回数制限なし、全工程の保存・復元可 各工程でのユーザが作成した3D表示は、原子単位で完全に保存可能	○	○
分子表示	ワイヤーフレーム、スティック、ボールスティック、スペースフィル、タンパク質のバックボーン表示、チューブ表示、リボン表示、カートゥーン表示、サーフェス表示（ポリゴン表示）、Cavity表示、ラベル表示、糖鎖のSNFG表記、スラブ表示（zクリッピングプレーン機能）、センターリセット3種類、水素結合表示、静電ポテンシャル面表示、アナグリフによるステレオ表示、投影モードのperspective（透視投影）またはparallel（平行投影）切り替え	○	○
色指定	モノクローム表示、CPK表示、シェイプリー表示、グループ表示、チェーン表示、温度表示、構造表示、電荷表示、残基表示、ポリゴン透明度指定	○	○
分子操作	マウスホイールによる拡大・縮小、マウスによる回転・移動・削除	○	○

ライセンス

[製品名] 機能	価格（税別）
[MolDesk Screening] MolDesk Basicのすべての機能 + バーチャルスクリーニング・高精度ポケット探索・ 部分構造検索・類似構造計算・活性予測	[お問い合わせください] 1年 1ノード
	1年経過後の使用は更新が必要です。

※ 動作OSは、Windows 11 / 10 (64bit), Linux (64bit), macOS 10.11 以上
ただし、スクリーニング計算は、8並列で16GB、48並列で32GBメモリが必要。

詳しくは、MolDesk のホームページをご覧ください www.moldesk.com



株式会社 情報数理バイオ

○お問合せ先 (株)情報数理バイオ 営業部
〒170-0013 東京都豊島区東池袋4-21-1 アウルタワー6F
TEL 03-6907-0315 FAX 03-6907-0316
EMAIL: info@imsbio.co.jp URL: http://www.imsbio.co.jp