

AI-SHIPS 統合的毒性予測システム ユーザーガイド 第2章

AI-SHIPS 統合的毒性予測システム

操作説明書

(第 1.1 版)

2026 年 7 月
AI-SHIPS プロジェクト

目次

1.	はじめに	1
1.1.	本システムの機能	1
1.2.	動作環境	1
1.3.	基本的な利用の流れ	2
2.	ログイン、ログアウト、環境設定	4
2.1.	ログイン	4
2.2.	ログアウト	5
2.3.	環境設定	5
3.	体内動態予測、毒性予測、類似物質検索の条件入力と実行	6
3.1.	予測や類似物質検索の対象物質の入力	6
3.2.	物質 ID、解析 ID 等の入力	9
3.3.	予測や類似物質検索の対象物質の適用範囲に関する情報の確認	9
3.4.	予測・類似物質検索対象の選択	10
3.5.	予測や検索の実行	13
4.	予測・検索のステータス、履歴表示	14
4.1.	検索条件入力	14
4.2.	検索の実行	14
4.3.	検索結果の表示	14
4.4.	検索結果の閲覧	16
5.	予測対象の適用範囲に関する情報（CS 画面）	17
5.1.	主要な mordred 記述子の範囲	17
5.2.	ケミカルスペース	18
6.	予測結果の表示	21
6.1.	毒性予測結果	24
6.2.	体内動態予測結果	32
6.3.	類似物質検索結果	34
7.	その他の機能	39
7.1.	データベース検索	39
7.2.	ユーザーデータの登録	41
7.3.	モデル・データベース情報	42
7.4.	DeepSnap	43
8.	ファイルフォーマット	45
8.1.	毒性予測や類似物質検索をする際の入力ファイルのフォーマット	45
8.2.	ユーザーデータ登録フォーマット	48

1. はじめに

AI-SHIPS 統合的毒性予測システムの主な機能は、「毒性予測」、「体内動態予測」、「類似物質検索」です。AI-SHIPS 統合的毒性予測システムは、ユーザーが保持しているデータを登録することが可能ですが、登録したユーザーデータは、全てのユーザーが閲覧できる状態となりますので、企業秘密にあたる情報は登録しないでください。

1.1. 本システムの機能

本システムは主に以下の機能を備えています。

1. 毒性予測

予測対象とする化学物質の構造情報を入力値とし、28 日間ラット反復投与毒性試験の結果を予測します。毒性試験の結果は、NOEL（毒性試験においていかなる影響も認められない最高投与量）や LOEL(毒性試験において何らかの影響が認められた最小投与量)の値が得られますが、本プロジェクトでは、本毒性試験について、肝臓、血液、腎臓に関する 9 つの毒性エンドポイントで投与量 30mg/kg/day 以下ならびに 300mg/kg/day 以下で陽性が陰性を予測します。併せて、各種 *in vitro* 試験の予測結果を表示します。

2. 体内動態予測

予測対象とする化学物質の構造情報と投与量等を入力値とし、これをラットに 28 日間反復投与した際の血液、肝臓、腎臓での経時的な物質濃度を予測します。あわせて吸収や蓄積性に係るパラメータの予測値も出力します。

3. 類似物質検索

予測対象の化学物質の構造情報と類似性の条件を入力値とし、データベース内の類似物質を検索し、類似物質の毒性予測結果等を比較可能な形式で表示します。

下記の機能もあります。本書では「7. その他の機能」で説明しています。

4. データベース検索

AI-SHIPS プロジェクトで収集した毒性試験のデータや、ユーザーが登録した毒性試験のデータを検索、表示します。

5. DeepSnap

分子の 3D 画像データを入力情報として学習し、毒性判別を実施するモデルを搭載しています。初代培養ラット肝細胞を用いた 7 種類の CYP 分子種の阻害活性について、対象物質の毒性判別を行うとともに、予測モデルに基づく活性に寄与する分子特徴領域を図示します。

1.2. 動作環境

本システムは Web ブラウザで動作します。動作環境として下記の環境で使用してください。

動作 Web ブラウザ Chrome (97 以降)

1.3. 基本的な利用の流れ

体内動態予測や毒性予測の基本的なフローは下記の通りです。

まず、予測対象物質を入力します。併せて投与量などの条件を入力します（⇒本書3）。実行する前に、予測対象物質の毒性予測の適用領域の確認をします（⇒本書 3.3）。具体的には、主要分子記述子に関して学習データの最小～最大の範囲内に予測対象物質の値が入っているかの確認や、AI-SHIPS で作成したケミカルスペース（多くの一般化学物質について、その化学構造を多数の構造記述子で表現し、記述子の数と同じ次元数の空間に化学物質をプロットしたものをGTM手法¹で二次元平面に写像した図で、本システムではCSと言います）で学習データ群と予測対象物質の位置関係を確認します。

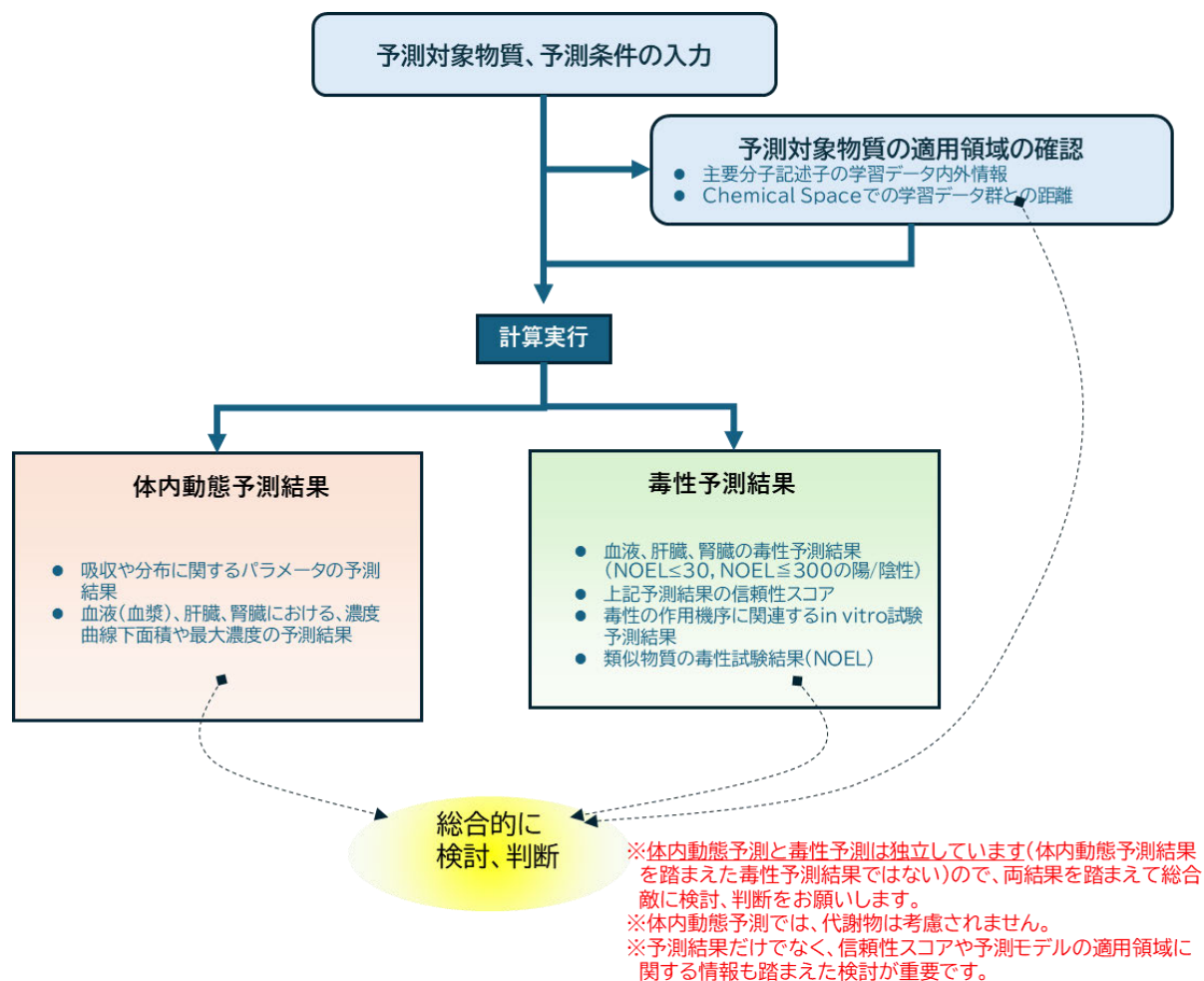
予測を実行します。複数物質を対象としている場合などは計算に時間を要しますので、予測結果の検索画面から当該予測計算が終了していることを確認してから、予測結果を表示させます（⇒本書 4 および 6）。

本システムでは、体内動態予測と毒性予測は独立したアプリケーションとなっています。体内動態予測と毒性予測は同時計算が可能で、両者の結果が出力されますが、毒性予測結果に体内動態予測結果が考慮されていないことにご留意ください。体内動態予測結果と毒性予測結果の両方を踏まえた総合的な検討、判断が必要です。

毒性予測については、「類似物質検索類似物質検索」機能を用いて、データベースに収載されている物質の中から予測対象物質と類似性の高い物質を抽出し、毒性試験結果（実測値）を表示することができます。類似性の高い物質の情報は、予測対象物質の信頼性の検討に活用できます。

さらに、表示毒性予測の変数として使用されている各種 *in vitro* 試験の結果も表示されます。この情報は、毒性発現の作用機序の解釈などの参考として活用ください。

¹ GTM : Generative topographic mapping。多次元空間の物質を、確率モデルを用いて二次元座標に投影する次元圧縮手法の1つ



2. ログイン、ログアウト、環境設定

2.1. ログイン

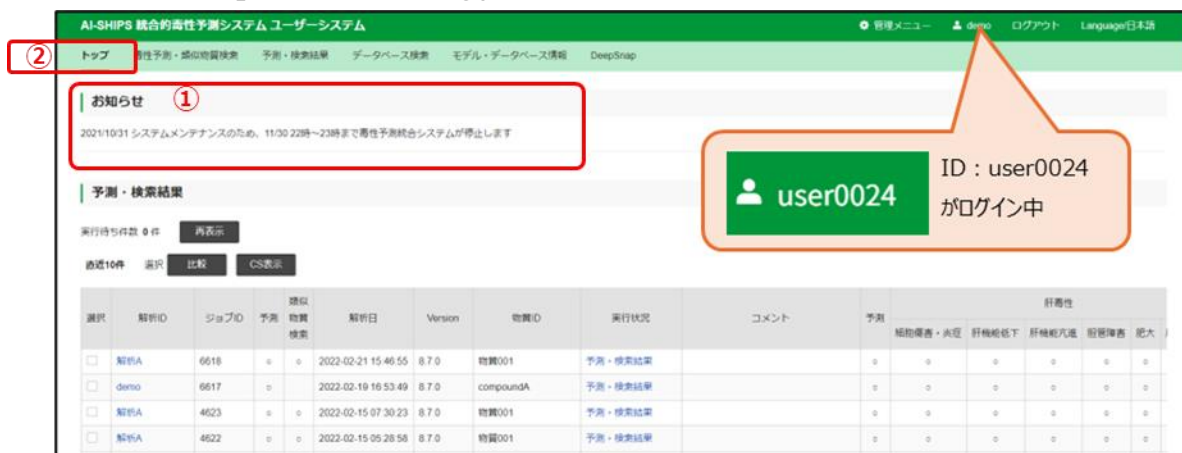
ユーザーID、パスワードを入力し、[ログイン]をクリックします。



ログインすると、トップ画面が表示されます。ヘッダーにログインしている ID が表示されます。

「お知らせ (①)」には、システムメンテナンスなどのアナウンス情報が表示されます。

メニューバーから「トップ」を選択した場合 (②) も本画面に遷移します。



※ID/パスワードが分からなくなった場合はシステム管理者までご連絡ください。

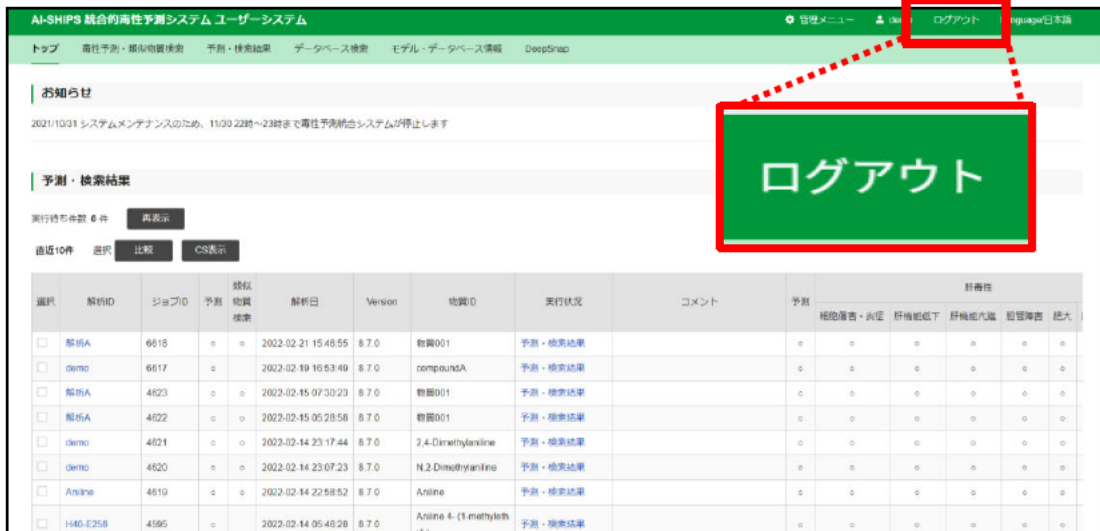
※同一ユーザーID で複数ログインした場合には、後でログインしたユーザーが優先されます。

先にログインされていたユーザーは認証エラーとなります。



2.2. ログアウト

ヘッダーの[ログアウト]をクリックします。



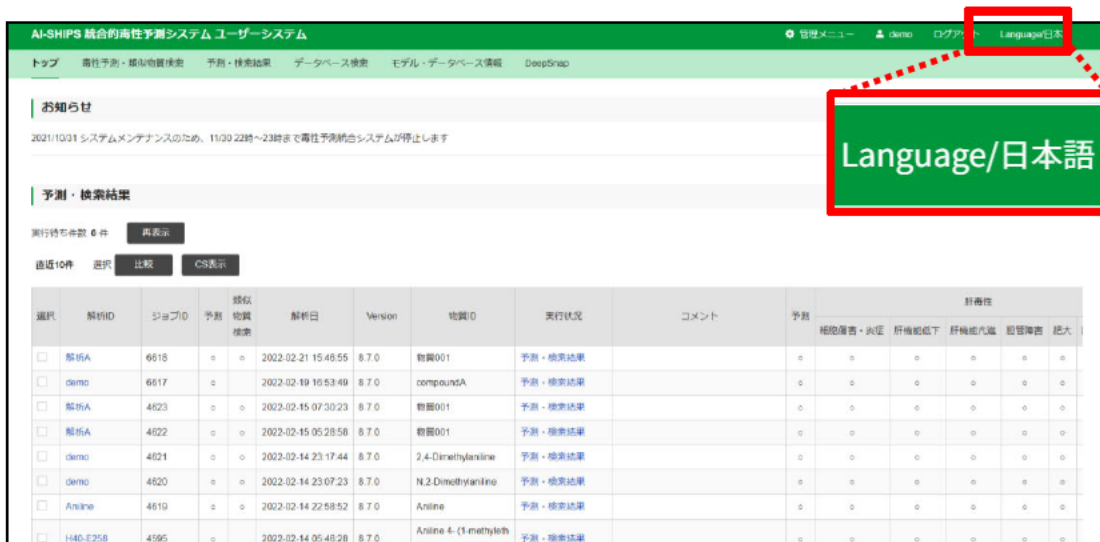
The screenshot shows the top navigation bar of the AI-SHIPS system. The 'ログアウト' (Logout) button is highlighted with a red box. A red dashed line points from this button to a larger green box with the text 'ログアウト' (Logout) inside, indicating the action to be taken.

選択	解析ID	ジョブID	類似 物質 検索	解析日	Version	物質ID	実行状況	コメント	予測	副作用				
										細胞障害・炎症	肝機能低下	肝機能亢進	胆管障害	総大
<input type="checkbox"/>	解析A	6618	○ ○	2022-02-21 15:48:55	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	6617	○	2022-02-19 16:53:49	8.7.0	compoundA	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	解析A	4623	○ ○	2022-02-15 07:30:23	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	解析A	4622	○ ○	2022-02-15 05:28:58	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	4621	○ ○	2022-02-14 23:17:44	8.7.0	2,4-Dimethylaniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	4620	○ ○	2022-02-14 23:07:23	8.7.0	N,2-Dimethylaniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	Aniline	4619	○ ○	2022-02-14 22:58:52	8.7.0	Aniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	H40-E258	4595	○	2022-02-14 05:48:28	8.7.0	Aniline 4-(1-methylph	予測・検出結果		○	○	○	○	○	○

2.3. 環境設定

2.3.1. 言語の切り替え

ログイン後に、ヘッダーの[Language/日本語]をクリックすると[Language/English]に変更できます。



The screenshot shows the top navigation bar of the AI-SHIPS system. The 'Language/日本語' button is highlighted with a red box. A red dashed line points from this button to a larger green box with the text 'Language/日本語' inside, indicating the action to be taken.

選択	解析ID	ジョブID	類似 物質 検索	解析日	Version	物質ID	実行状況	コメント	予測	副作用				
										細胞障害・炎症	肝機能低下	肝機能亢進	胆管障害	総大
<input type="checkbox"/>	解析A	6618	○ ○	2022-02-21 15:48:55	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	6617	○	2022-02-19 16:53:49	8.7.0	compoundA	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	解析A	4623	○ ○	2022-02-15 07:30:23	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	解析A	4622	○ ○	2022-02-15 05:28:58	8.7.0	物質001	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	4621	○ ○	2022-02-14 23:17:44	8.7.0	2,4-Dimethylaniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	demo	4620	○ ○	2022-02-14 23:07:23	8.7.0	N,2-Dimethylaniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	Aniline	4619	○ ○	2022-02-14 22:58:52	8.7.0	Aniline	予測・検出結果		○	○	○	○	○	○
<input type="checkbox"/>	H40-E258	4595	○	2022-02-14 05:48:28	8.7.0	Aniline 4-(1-methylph	予測・検出結果		○	○	○	○	○	○

3. 体内動態予測、毒性予測、類似物質検索の条件入力と実行

メニューバーの[毒性予想・類似物質検索]をクリックして、毒性予測・類似物質検索画面に遷移します。



3.1. 予測や類似物質検索の対象物質の入力

必須

毒性や体内動態の予測、類似物質検索の対象物質の化学構造を入力します。化学構造は、以下の3つの方式で入力することができます。

使用したい入力方法のトグルボタンをクリックします。

- ① 構造式エディタ (Molecule Editor) による構造式の描画
- ② Mol ファイルまたは、SD ファイルまたは tsv(SMILES)ファイルのアップロード
- ③ SMILES の入力

②の方法では、一度に多くの物質を指定することが可能です。

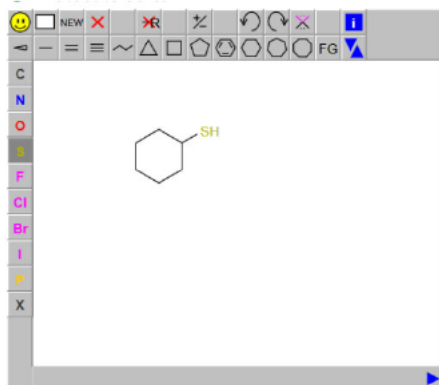
※制限事項

いずれの入力方式においても 81 文字以上の SMILES で 構成される物質は入力できません。



① 構造式エディタ (Molecule Editor) による構造式の描画

ツールバーから原子（炭素、酸素など）や結合（単結合、二重結合など）を選び、キャンバス上でドラッグして分子構造を描きます。



周囲に配置されている各ボタンの説明

ボタン	説明
	SMILES を作成
	立体結合（選択後、原子または結合をクリック。クリックを繰り返すことでタイプ（上、下、方向）が切り替わる）
	特定原子の入力
	上記以外の原子の入力
	描画領域をクリア
	新しい分子コンポーネントを開始
	特定箇所の削除（選択後、原子または結合をクリック）
	官能基を削除（選択後、基を主骨格に接続する結合をクリック）
	無機原子や非標準の原子価状態の原子を入力
	戻す、次
	スピロ環（選択すると次の環がスピロ環として追加）
	特定官能基の追加（クリックすると表示される選択肢*から選択） *-C(=O)OH, -NC=O, CF3 など
	結合線の入力
	環状構造の入力
	SMILES や MOL, InChI, InChI Key 等の生成 またはこれからの読み込み

② Mol ファイルまたは、SD ファイルまたは tsv(SMILES)ファイルのアップロード

[ファイル選択]をクリックして、対象のファイルを指定します。

この方法では、物質 ID はファイルに記載して指定するため、画面からの指定はできません。

入力ファイルのフォーマットについては、8.1 を参照ください。

③ SMILES の入力

SMILES を入力します。

例) CCCCC(C)(CC)c1ccc(O)cc1

3.2. 物質 ID、解析 ID 等の入力 任意

1) 物質 ID を入力します。

※2.1 で、「②Mol ファイルまたは、SD ファイルまたは tsv(SMILES)ファイルのアップロード」を選択している場合、本項目は表示されません。

物質 ID は、検索や予測の対象物質を識別する ID であり、予測・結果一覧画面で計算結果を識別するのに役立ちます。

数値や文字をつかって、後から判別しやすい ID をつけるとよいでしょう。

未入力の場合は、日時（予測・検索の開始時刻）が ID として自動的に設定されます。

2) 解析 ID を入力します。

解析 ID は、検索処理をジョブ管理する際の ID であり、予測・結果一覧画面で計算結果を識別するのに役立ちます。

未入力の場合は、日時（予測・検索の開始時刻）が ID として自動的に設定されます。

3) コメントを入力します。

必要に応じて計算条件などを入力して記録しておきます。

The screenshot shows the 'AI-SHIPS 統合的毒性予測システム ユーザーシステム' interface. The main heading is '毒性予測・類似物質検索'. Under '構造データ', there is a 'CSチェック' button. Below that is a 'Molecule editor' section with a toolbar and a text area. At the bottom, there are radio buttons for 'Mol, SD or tsv (SMILES) file' and 'SMILES'. Below these are input fields for '物質ID', '解析ID', and a 'コメント' text area. A red box highlights the '物質ID', '解析ID', and 'コメント' fields.

3.3. 予測や類似物質検索の対象物質の適用範囲に関する情報の確認

構造データを入力後に、[CS チェック]をクリックすると、入力した物質の適用範囲に関する情報が新規ウィンドウで表示されます。結果の見方は 5 を参照ください。

This screenshot is similar to the previous one, but the 'CSチェック' button is highlighted with a red box, indicating the next step in the process.

3.4. 予測・類似物質検索対象の選択 必須

入力した対象物質についての実行内容を指示します。本システムでは、下記が可能です。

- 毒性予測 (→3.4.1)
- 体内動態予測 (→0)
- 類似物質検索 (→**エラー! 参照元が見つかりません。**)

複数を同時に実行することができますので、実行したい項目を全て指定します。

3.4.1. 毒性予測

1) [毒性モデル]の+ボタンをクリックし、項目を展開します。

2) 予測を行いたい毒性項目（エンドポイント）のチェックボックスにチェックを入れます。

「毒性モデル」をチェックすると、全ての毒性項目にチェックが入ります。



留意点 予測対象外の物質について

下記の物質は毒性予測モデルの学習データの対象外としているため、適用範囲外となります。

- ・ 脱塩処理の後で炭素を含まない物質
- ・ プロジェクトで実施した *in vitro* 試験の対象物質に含まれていない元素である、B, Al, Si, V, Cr, Mn, Fe, Ni, Cu, Zn, Sr, Zr, Sn, Te, Pt を含む物質
- ・ 化審法の対象外の物質
- ・ 高分子 (AI-SHIPS では、入力物質の SMILES として 80 文字まで、分子量は 1000 に制限されている)

3.4.2. 体内動態予測

1) [体内動態予測]の+ボタンをクリックし、項目を展開します。

2) 体内動態予測のチェックボックスにチェックを入れます。

3) 投与量を入力します。 **必須**
0～10000の範囲の数値で投与量を入力します。単位は、mg/kg/day(1日あたり、体重1kgあたりの量(kg))です。体内動態予測を選択した際は、必須の入力項目です。

4) その他の入力パラメータの値を入力します。 **任意**
実験値や推定モデルの推測値を入力します。入力は任意ですが、情報が得られている場合は入力を推奨します。そうすることでより正確に予測することができます。入力されなかった場合は、構造式から自動的に推測値が計算されます。

※対象物質の化学構造の入力をSDファイルやtsv(SMILES)ファイルで指定する場合は、本項目も同ファイル内で指定できます。

- LogP : 水-オクタノール分配係数。入力形式は、-10～100の範囲の数値。
- fu,p : 血漿タンパク質結合率。入力形式は、0～1の範囲の数値。
- LogD (pH 6.0) : 小腸環境(pH=6.0)での分配係数で、値が高いと化学物質が小腸の膜を通過する能力が高いことを示す。入力形式は、-10～100の範囲の数値。
- LogD (pH 7.5) : 各臓器環境(pH=7.5)での分配係数で、値が高いと化学物質が小腸の膜を通過する能力が高いことを示す。入力形式は、-10～100の範囲の数値。

一酸塩基/一塩基酸の場合

- pKa acid : 一酸塩基の酸解離定数。入力形式は、-1～100の範囲の数値。
- pKa base : 一塩基酸の酸解離定数。入力形式は、-1～100の範囲の数値。

二酸塩基/二塩基酸の場合

二酸塩基/二塩基酸をチェックします。チェックすると入力項目が右図のように変わります。

- pKa acid1 : 二酸塩基の第1酸解離定数。入力形式は、-10～100の範囲の数値。
- pKa base1 : 二塩基酸の第1酸解離定数。入力形式は、-10～100の範囲の数値。
- pKa acid2 : 二酸塩基の第2酸解離定数。入力形式は、-10～100の範囲の数値。
- pKa base2 : 二酸塩基の第2酸解離定数。入力形式は、-10～100の範囲の数値。

※二酸塩基/二塩基酸を選択した場合にはpKa acid1とpKa acid2、またはpKa base1とpKa base2は入力必須

3.4.3. 類似物質検索

1) [類似物質検索]の+ボタンをクリックし、項目を展開します。

2) OR 検索とするか、AND 検索とするかを選択します。

OR 検索とは、4)で複数の条件を選択した際に、いずれかの条件に合致した物質を表示します。

AND 検索とは、4)で複数の条件を選択した際に、全ての条件に合致した物質を表示します。

※4)で1つの条件を選択した場合は、OR 検索とAND 検索の結果は同じです。

※デフォルトではAND 検索が選択されています。

3) 検索対象のデータベースを選択します。

データベースは、以下の2つのうち、1つまたは両方を検索対象とすることができます。

※デフォルトではAI-SHIPSデータが選択されています。

➤ AI-SHIPS データ

Hazard evaluation support system integrated platform (以後、HESS という)、REACH 登録情報、ToxRef および化審法新規審査シートを情報源として、下記の条件に合致する約 2,200 のラット毒性反復投与毒性試験データが収録されています。

- ・ OECD TG407 (28 日間反復経口投与毒性試験)、OECD TG422 (反復投与毒性試験と生殖/発生毒性スクリーニング試験の併合試験) またはこれに相当する毒性試験
- ・ 投与期間が 28 日以上 90 日未満 (垂急性毒性に該当する)
- ・ 雄のデータ (ただし、雌雄が不明の場合も採用)

なお、このうち、1,900 物質程度が毒性予測モデル構築の学習データとして使用されています。

➤ ユーザーデータ

ユーザーが登録した毒性情報データ。

データの登録は、管理メニューのユーザーデータ管理から実施します。登録方法は 7.2 を、登録フォーマットは、8.2 を参照ください。

※登録したユーザーデータは、全てのユーザーが閲覧 (検索結果として) できる状態となりますので、企業秘密にあたる情報は登録しないでください。

4) 類似物質検索アルゴリズムを選択します。

必須

類似性の判定アルゴリズムが複数用意されているので、アルゴリズムを1つ以上選択します。全てのアルゴリズムを選択することも可能です。

さらに、選択したアルゴリズムについて、類似性の判定の閾値を設定します。

類似性の判定の閾値を厳しく設定すると、検索の結果として抽出される物質は少なくなりますので、結果を見ながら調整してください。

類似性アルゴリズム名		説明	値の幅と意味
構造類似性	Mordred 記述子	rdkitの mordred 記述子（304種）の値での多次元空間における物質間のユークリッド距離を指標とした類似性評価	0～1の範囲で、0に近いほど類似性が高くなる。0～0.5の数値を入力する。
	Fingerprint Topological	rdkitの Topological fingerprint ² を用い、その類似性をタニモト係数で評価する類似性評価 Topological fingerprint：分子内の部分構造パターンを列挙してビット化する。	0～1の範囲で、1に近いほど類似性が高い。0.5～1の数値を入力する。
	Fingerprint Morgan	rdkitの Morgan fingerprint ³ を用い、その類似性をタニモト係数で評価する類似性評価 Morgan fingerprint：原子を中心に局所環境を半径で拡張して表現したもの。	0～1の範囲で、1に近いほど類似性が高い。0.5～1の数値を入力する。
インビトロ実験プロファイル類似性		<i>in vitro</i> 試験結果予測値の陽性/陰性のパターンの類似性をタニモト係数で評価する類似性評価	0～1の範囲で、1に近いほど類似性が高い。0.5～1の数値を入力する。
PKプロファイル類似性		体内動態パラメータ(吸収速度、分布容積、肝代謝消失)を用いた三次元空間における物質間のユークリッド距離を指標とした類似性評価	0～1の範囲で、0に近いほど類似性が高い。0～0.5の数値を入力する。
分子量類似性		予測対象物質の分子量との差分を指標とした類似性評価	小さいほど分子量の類似性が高い。1～500範囲の数値を入力可能。

3.5. 予測や検索の実行

[予測・検索]をクリックします。

² <https://www.rdkit.org/docs/GettingStartedInPython.html#topological-fingerprints>

³ <https://www.rdkit.org/docs/GettingStartedInPython.html#morgan-fingerprints-circular-fingerprints>

4. 予測・検索のステータス、履歴表示

メニューバーの[予測・検索結果]をクリックして、予測・検索結果一覧画面に遷移します。

毒性予測・類似性検索画面の[予測・検索]をクリックした際も、予測・検索結果一覧画面に遷移します。



4.1. 検索条件入力

ステータスや履歴を表示するための検索条件を入力します。

- ・予測や検索を実施した日の範囲 **必須**
5 年前から検索日までの間で指定が可能です。
デフォルトでは、最新 1 カ月に設定されています。



- ・解析 ID、物質 ID **任意**
解析 ID や物質 ID で検索したい場合に入力します。
※ワイルドカードを使った指定はできません。
未入力の場合は、上記で指定した期間の全データが検索されます。

4.2. 検索の実行

[検索]をクリックします。

4.3. 検索結果の表示

入力された条件に合致する (AND 検索) 結果が表示されます。

検索一覧

実行待ち件数 1件 再表示

比較 CS表示 削除

表示件数 10 1 2 3 244 245

選択	解析ID	ジョブID	予測物質検索	解析日	Version	物質ID	実行状況	コメント	予測	肝毒性							
										細胞傷害・炎症	肝機能低下	肝機能亢進	胆管障害	肥大	脂質代謝異常	腎毒性	
<input type="checkbox"/>	202509042147	2823	○	2025-09-04 21:47:40	8.7.0	20250904214740399	実行待ち		○	○	○	○	○	○	○	○	○
<input type="checkbox"/>	202509042138	2822	○	2025-09-04 21:38:24	8.7.0	20250904213824264	予測・検索結果		○	○	○	○	○	○	○	○	○
<input type="checkbox"/>	202503061652	2821	○	2025-03-06 16:52:18	8.7.0	244303-78-4	予測・検索結果		○	○	○	○	○	○	○	○	○
<input type="checkbox"/>	202503061652	2820	○	2025-03-06 16:52:18	8.7.0	97-64-3	予測・検索結果		○	○	○	○	○	○	○	○	○
<input type="checkbox"/>	202503061652	2819	○	2025-03-06 16:52:18	8.7.0	41395-83-9	予測・検索結果		○	○	○	○	○	○	○	○	○
<input type="checkbox"/>	202503061652	2818	○	2025-03-06 16:52:18	8.7.0	132900-75-5	予測・検索結果		○	○	○	○	○	○	○	○	○

- ・ 実行待ち数(①)
計算中のステータスの件数が表示されます。
- ・ [再表示]ボタン(②)
実行状況が「実行待ち」の際に、状況を更新して最新化したいときに利用します。クリックすると、検索が再実行されます。
- ・ [比較]ボタン(③)
複数の解析結果を比較したいときに利用します。比較したい解析 ID の「選択」チェックボックスを ON にした後にクリックすると、新規ウィンドウにて選択された解析 ID の結果を並列して表示します。
- ・ [CS 表示]ボタン(④)
解析結果のケミカルスペース (CS) 情報を見たいときに利用します。表示したい解析 ID の「選択」チェックボックスを ON にした後にクリックすると、当該解析 ID のケミカルスペース情報が新規ウィンドウで表示されます。結果の見方は 5 を参照ください。
※複数の解析 ID を選択し、それらについての CS 情報を表示することもできます。
- ・ [削除]ボタン(⑤)
不要となった解析結果を削除する際に利用します。削除したい解析 ID の「選択」チェックボックスを ON にした後にクリックすると、当該解析が削除されます。

検索一覧表の表示内容は下記の通りです。

項目	表示内容	
解析 ID	予測・検索の実行時にユーザーが指定した解析 ID (未入力の場合は、日時 (予測・検索の開始時刻) が自動設定されている) クリックすると同じ解析 ID について選択にチェックが入る。	
ジョブ ID	自動付与	
予測	体内動態や毒性の予測をした場合に○が表示される。	
類似物質検索	類似物質検索を実行した場合に○が表示される。	
解析日	解析日	
Version	モデル構築のデータマトリクスバージョン	
物質 ID	予測・検索の実行時にユーザーが指定した物質 ID (未入力の場合は、日時 (予測・検索の開始時刻) が自動設定されている)	
実行状況	予測・検索の計算が完了していない場合は「実行待ち」、完了していれば「予測・検索結果」と表示される。 処理がエラーとなった場合は「検索エラー」と表示される。マウスオーバーでエラー内容が表示される。 「予測・検索結果」をクリックすると、新規ウィンドウとして、当該解析の結果表示画面に遷移する。	
コメント	実行時にユーザーが入力したコメント	
肝毒性	細胞傷害・炎症	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肝機能低下	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肝機能亢進	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	胆管障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肥大	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。

項目		表示内容
	脂質代謝異常	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
腎毒性	腎障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
血液毒性	凝固障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	貧血	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
体内動態	投与量	体内動態予測時に指定した条件。設定値が表示される。
	LogP	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	fu,p	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	LogD (pH 6.0)	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	LogD (pH 7.5)	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	pKa acid	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	pKa base	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	二酸塩基/二塩基酸	体内動態予測時に設定された条件。チェックされた場合に○が表示される。
	pKa acid1	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	pKa acid2	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	pKa base1	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	pKa base2	体内動態予測時に設定された条件。設定値が表示される（設定されなかった場合は空欄）。
	構造類似性	条件
AI-SHIPS データ		類似物質検索時に指定した検索条件。検索対象に AI-SHIPS データが選択されていた場合に○が表示される。
ユーザーデータ		類似物質検索時に指定した検索条件。検索対象にユーザーデータが選択されていた場合に○が表示される。
mordred 記述子類似性		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。
Topological FP		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。
Morgan FP		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。
インビトロ実験プロファイル類似性		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。
PK プロファイル類似性		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。
分子量類似性		類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値が表示される（選択されていない場合は空欄）。

4.4. 検索結果の閲覧

表示したい解析 ID の[予測・検索結果]をクリックします。複数の結果を並列して表示したい場合は、対象の「選択」チェックボックスを ON にし、[比較]をクリックします。



5. 予測対象の適用範囲に関する情報 (CS画面)

「予想・検索結果一覧」画面で、[CS表示]をクリックした場合や、「毒性予想・類似物質検索」画面で[CS表示]をクリックした場合に表示される画面です。

5.1. 主要な mordred 記述子の範囲

画面上部には選択した物質の化学構造と共に、主要な mordred 記述子の値が表示されます (①)。併せて、主要な mordred 記述子 (*in vivo* 毒性予測モデルの説明変数として使用した 15 の記述子) について、AI-SHIPS に含まれる物質の最大値 (②の Max) 及び最小値 (②の Min) が表示されます。

Descriptor	Max	Min	value
MW_2D	2284.0546	31.0422	192.0374
AMW_2D	208.9904	4.3566	10.6687
SLogP_2D	37.8040	-15.2306	2.0286
SMR_2D	696.5154	0	29.3760
TPSA_3D	740.9473	0	213.2725
nHBAcc_2D	38	0	0
nHBDon_2D	21	0	0
nRot_2D	112	0	0
RotRatio_2D	0.9091	0	0
nAcid_2D	8	0	0
nBase_2D	8	0	0
apol_2D	422.3382	4.1728	17.3628
bpol_2D	296.4017	0	12.6972
nAromAtom_2D	60	0	0
nAromBond_2D	82	0	0

mordred 記述子	説明 ⁴
W_2D, AMW_2D	exact molecular weight(2D), averaged exact molecular weight (2D)
SLogP_2D	Wildman-Crippen LogP(2D)
SMR_2D	Wildman-Crippen MR(2D)
TPSA_3D	total polar surface area(3D)
nHBAcc_2D, nHBDon_2D	number of hydrogen bond acceptor(2D), number of hydrogen bond donor(2D)
nRot_2D, RotRatio_2D	rotatable bonds count(2D), rotatable bonds ratio (2D)
nAcid_2D, nBase_2D	acidic group count(2D), basic group count(2D)
apol_2D, bpol_2D	atomic polarizability(2D), bond polarizability(2D)
nAromAtom_2D, nAromBond_2D	aromatic atoms count(2D), aromatic bonds count(2D)

⁴ <https://mordred-descriptor.github.io/documentation/master/index.html>

読み方

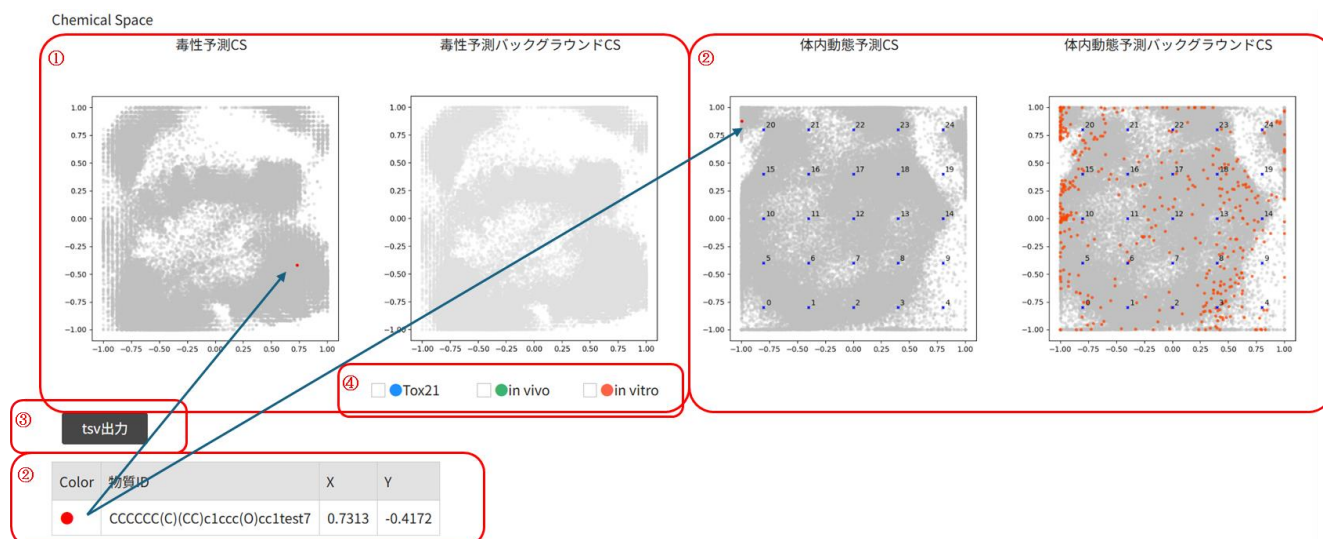
AI-SHIPS *in vivo* 毒性予測モデルは、本画面で表示されている 15 の mordred 記述子を説明変数として、AI-SHIPS に収載されている約 1,900 の物質の毒性情報を学習データとして構築されたモデルです。この mordred 記述子についての予測対象物質の値が、学習データの最大値と最小値の間であれば、学習データ群の中にあると言えます。毒性予測結果の信頼性が高いと判断できる材料の 1 つになります。

5.2. ケミカルスペース

CS 画面の下部には AI-SHIPS プロジェクトで作成したケミカルスペース (CS) が表示されます。

ケミカルスペースとは、化学構造を多数の構造記述子で表現し、その記述子の数と同じ次元数の空間に化学構造を GTM 手法で⁵二次元平面に写像したものです。AI-SHIPS プロジェクトで作成したケミカルスペースを“CS”として表示しています。

以下の 4 つのマップのうち、左の 2 つが毒性予測用、右の 2 つが体内動態予測用の CS です。



【毒性予測の CS (①)】

PubChem(アメリカ NIH 傘下の NCBI が運営する化学物質のデータベース) より無作為に抽出した約 10 万の化学物質について、866 種の mordred 記述子を算出し、確率モデルを用いて二次元平面に写像したものです。

毒性予測 CS (①の左図)

灰色のプロット : CS を構成した際の 10 万件の物質

カラー (1 物質の場合は赤色) で示されるプロット : 予測対象の物質

予測対象の物質 ID と、予測対象の CS 上の座標位置が示されます (③)。

[tsv 出力]ボタン (②) で、座標の値をエクスポート (tsv 形式) できます。


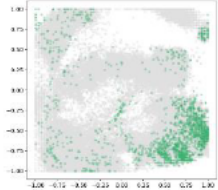

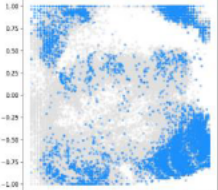

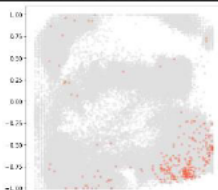
⁵ GTM : Generative topographic mapping。多次元空間の物質を、確率モデルを用いて二次元座標に投影する次元圧縮手法の 1 つ

毒性予測バックグラウンド CS (①の右図)

灰色のプロット：CSを構成した際の10万件の物質

チェックボックス(④)で指定された物質が灰色のプロットの上に重ねて表示される。

チェックボックスの内容は以下の通りです。

チェックボックス	毒性予測バックグラウンド CS に重ねてプロットされる内容
<input checked="" type="checkbox"/>  in vivo	チェックすると、AI-SHIPS の <i>in vivo</i> 毒性予測モデルの構築において学習データとして使用した 1,979 物質が緑色のプロットで示される。 
<input checked="" type="checkbox"/>  Tox21	チェックすると、Tox21 の予測モデル構築において学習データとして使用した約 12,974 物質が青色のプロットで示される。 
<input checked="" type="checkbox"/>  in vitro	チェックすると、AI-SHIPS の <i>in vitro</i> 予測モデル構築において学習データとして使用した約 326 物質がオレンジ色のプロットで示される。 

読み方

AI-SHIPS *in vivo* 毒性予測モデルは、画面上部で表示されている 15 の mordred 記述子ならびに、AI-SHIPS で実施した多様な *in vitro* 試験に関する 95 のパラメータの予測値、Tox21 の MIE アッセイに関する 101 のパラメータの予測値を説明変数として、AI-SHIPS データベースに含まれているデータを精査し、約 1,900 の物質の毒性情報を学習データとして構築されたモデルです。AI-SHIPS で実施した多様な *in vitro* 試験は、AI-SHIPS データベースから、多様性等に配慮して選定された 326 物質に対して実施されました（したがって上記の「*in vitro*」のオレンジ色のプロットは、「*in vivo*」の緑のプロットの一部と重なります）。そして、これらの実験結果を学習データとして、実験結果を予測する 95 の予測モデルが構築されました。また、米国の Tox21 プロジェクトで 1 万を超える物質に対して実施された多様なハイスループット *in vitro* 試験の結果を学習データとして、101 のパラメータに対する予測モデルが構築されました。

CS で近接した位置に図示される物質同士は、一定の構造類似性が認められます。①の右図で、「*in vivo*」のチェックボックスを ON にすると、*in vivo* 毒性予測モデル構築において学習データとして使用した物質が緑色でプロットされます。このプロット群の中に予測対象物質のプロットがあれば、予測対象物質が学習データ群との構造類似性が高く、予測結果の信頼性が高いと考えられます。逆に青色のプロット群から外れた位置に予測対象物質のプロットがあれば、学習データとの類似性が低い物質であり、予測結果の信頼性は低いと考えられます。さらに、上記で「*in vitro*」のオレンジ色のプロット群や「Tox21」の青色のプロット群の中に予測対象物質のプロットがあれば、説明変数としているパラメータの予測値の信頼性も高いと考えられます。

【体内動態予測 CS (②)】

複数の一般化学物質データベースから抽出した約 5 万の化学物質について、rdkit(version 2019.3.1)の 196 記述子について確率モデルを用いて二次元平面に写像したものです。

体内動態予測 CS (②の左図)

灰色のプロット：CS を構成した際の 5 万件の物質

カラー (1 物質の場合は赤色) で示されるプロット：予測対象の物質

体内動態予測バックグラウンド CS (②の右図)

灰色のプロット：CS を構成した際の 5 万件の物質

オレンジプロット：体内動態の予測モデルの学習データとして使用した約 370 物質

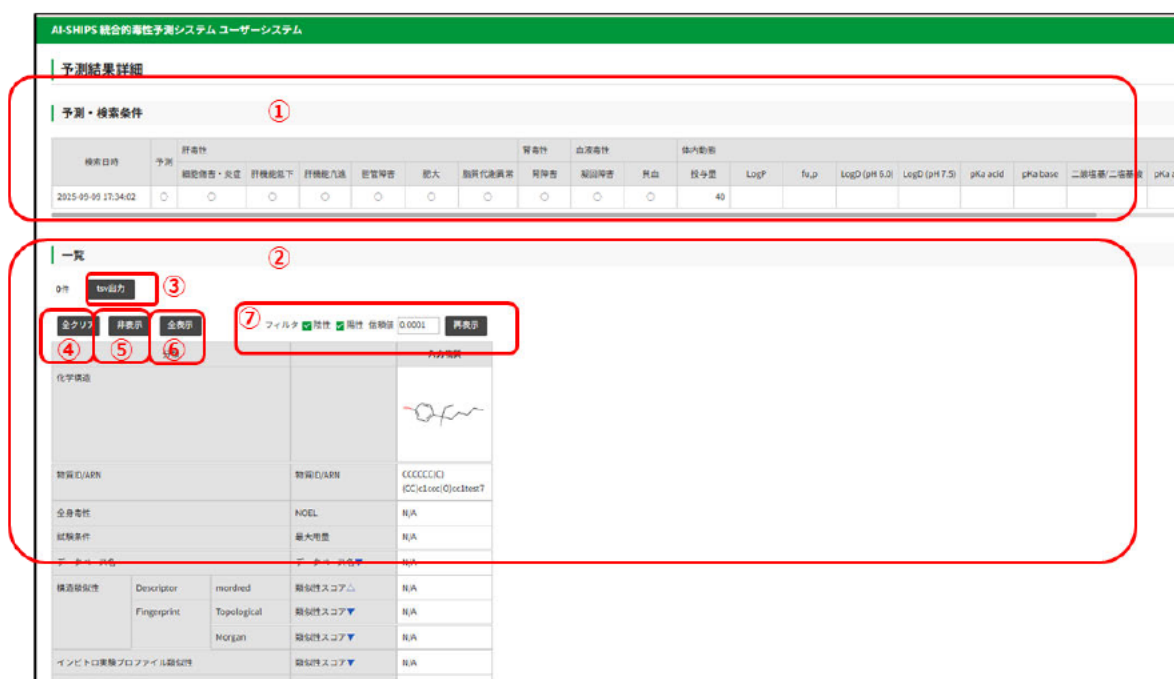
読み方

CS 上に示されているドットは、体内動態に関わる 4 つのパラメータ、すなわち化学物質の消化管利用率 (FaFg)、吸収速度定数 (k_a)、分布容積 (V_1)、および肝代謝固有クリアランス ($CL_{h,int}$) の予測モデルの構築に用いた 372 件の学習データの位置を示しています。学習データが CS を 5×5 に分割した 25 区画の全域に広く分布していることから、本予測モデルは多様な化学物質に適用可能であることを示しています。

なお、体内動態予測モデルでは、これら 4 つのモデル予測値を入力値として体内動態シミュレーションを実行し、その結果として、血中、肝臓および腎臓における経時的な濃度推移に加え、濃度-時間曲線下面積 (AUC) および最高臓器中濃度 (C_{max}) が算出されます (ユーザーガイド第 1 章参照)。

6. 予測結果の表示

実行状況の[予測・検索結果]をクリックすると、選択された解析 ID に対する予測・検索結果が新規ウィンドウで表示されます。



上部に予測・検索条件（①）、下部に予測や検索を実行した一覧（②）が表示されます。

選択した解析 ID での予測や類似物質検索の条件が表示されます。

項目	内容	
解析日	解析した日	
予測	体内動態や毒性の予測を実行したものに○	
肝毒性	細胞傷害・炎症	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肝機能低下	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肝機能亢進	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	胆管障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	肥大	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	脂質代謝異常	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。

腎毒性	腎障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
血液毒性	凝固障害	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
	貧血	体内動態予測時に指定した条件。毒性予測対象として選択された場合に○が表示される。
体内動態	投与量	体内動態予測時に指定した条件。設定値を表示
	LogP	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	fu,p	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	LogD (pH 6.0)	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	LogD (pH 7.5)	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	pKa acid	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	pKa base	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
	二塩基/二塩基酸	体内動態予測時に設定された条件。チェックされた場合に○が表示される。
	pKa acid1	体内動態予測時に設定された条件。設定値を表示（設定されなかった場合は空欄）
類似物質検索		類似物質検索を実行したものに○
構造類似性	条件	類似物質検索時に指定した検索条件。OR または AND
	AI-SHIPS データ	類似物質検索時に指定した検索条件。検索対象に AI-SHIPS データが選択されていた場合に○
	ユーザーデータ	類似物質検索時に指定した検索条件。検索対象にユーザーデータが選択されていた場合に○
	mordred 記述子類似性	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値
	Topological FP	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値
	Morgan FP	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値
	インビトロ実験プロファイル類似性	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値
	PK プロファイル類似性	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値
	分子量類似性	類似物質検索時に指定した検索条件。当該類似性が選択された場合に設定した閾値

【一覧】(②)

予測・検索結果の一覧を表示します。上から順に、化学構造、物質 ID/ARN^{*1}、全身毒性 NOEL^{*2}、最大用量、データベース、類似物質検索結果、毒性予測結果、体内動態予測結果が表示しています。

詳細は、0 毒性予測結果、6.2 体内動態予測結果、6.3 類似物質検索を参照ください。

^{*1}ARN : AI-SHIPS プロジェクトにおける毒性試験 ID

^{*2}全身毒性 NOEL : 毒性試験における全ての検査項目を対象とした試験全体の NOEL

また、この画面では以下の操作ができます。

- ・ [tsv 出力]ボタン(③)
予測・検索結果を tsv 形式にて出力します。
- ・ [全選択]ボタン(④)

類似物質検索結果の表示時に有効です。
物質の全てのチェックボックスにチェックを入れます。

- [非表示]ボタン(⑤)

類似物質検索結果の表示時に有効です。
チェックボックスにチェックが入っている物質（右図の⑧）を非表示にします。

- [全表示]ボタン(⑥)

類似物質検索結果の表示時に有効です。
非表示にしたデータを表示します。

- フィルタ(⑦)

類似物質検索結果の表示時に有効です。

表示する物質をフィルタリングします。フィルタリングの条件は、陰性／陽性（片方、または両方を選択可）、信頼性スコアです。条件を指定したのちに[再表示]ボタンで変更します。

The screenshot displays the '予測結果詳細' (Prediction Results Detail) page. At the top, there is a table for '予測・検索条件' (Prediction/Search Conditions) with columns for various properties like '予測' (Prediction), '有害性' (Toxicity), '環境影響' (Environmental Impact), etc. Below this, there is a list of similar substances. Each entry includes a chemical structure, a name, and a score. The entry for 'H40-E258' is highlighted with a red box and a circled 8, indicating it is the target of the '非表示' (Hide) action. Below the list, there are buttons for '全表示' (Show All) and '再表示' (Refresh).

6.1. 毒性予測結果

6.1.1. 毒性予測結果

毒性予測を実行した際は、下記の情報が表示されます。

属性エリア (①) の表示項目は下記の通りです。

表示項目	説明
化学構造	予測対象化学物質の構造が表示される。
物質 ID/ARN	物質 ID/ARN
全身毒性	NOEL (毒性予測結果では N/A と表示)
試験条件	最大用量 (毒性予測結果では N/A と表示)
データベース名	データベース名 (毒性予測結果では N/A と表示)



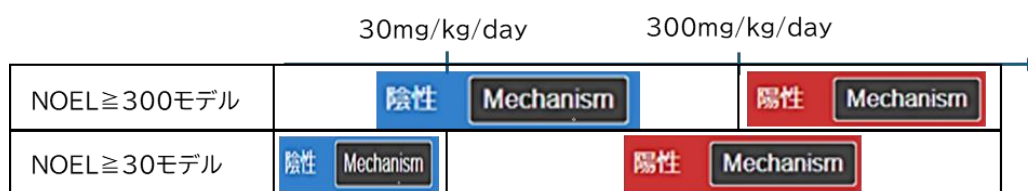
毒性予測結果エリア (②) の表示項目は下記の通りです。

表示項目		説明	
腎毒性	腎障害	NOEL ≥ 30 モデル ※文字列をクリックすると各 <i>in vitro</i> 試験の結果ページに遷移する (6.1.2 参照)	<p>予測結果</p> <p>腎障害に関する NOEL ≥ 30 かどうかの 2 値判定モデルでの予測結果が陽性/陰性で示される。陽性の場合には 陽性 Mechanism、陰性の場合には 陰性 Mechanism と表示される。</p> <p>陽性とは腎障害の NOEL が 30mg/kg/day 以上であることを示す。</p> <p>陰性とは腎障害の NOEL が 30mg/kg/day 未満であることを示す。</p> <p>Mechanism をクリックすると <i>in vitro</i> の試験結果が表示される (→6.1.3 章参照)。</p>
		信頼性スコア	上記の予測結果の信頼性スコア。0 ~ 1 の範囲の値で、1 に近いほどモデルの予測が明確であり、0 に近いほど予測が曖昧であることを表す。
	試験値	NOEL(mg/kg/day)	(毒性予測結果では N/A と表示)
…以下、予測対象として選択された数だけ繰り返し		(省略)	
肝毒性 (細胞傷害・炎症)			
肝毒性 (肝機能低下)、肝毒性 (肝機能亢進)、肝毒性 (胆管障害、肥大)、肝毒性 (脂質代謝異常)			
腎毒性 (腎障害)、血液毒性 (凝固障害、貧血)			

読み方 2つのモデル予測結果の解釈

1つの毒性エンドポイントに対して、NOEL \geq 30 モデルと NOEL \geq 300 モデルでの推定結果が示されます。両方が陽性であれば、当該物質の NOEL は 300mg/kg/day 以上であると解釈できます。NOEL \geq 30 モデルと NOEL \geq 300 モデルの両方が陰性であれば、当該物質の NOEL は 30mg/kg/day 未満であると解釈できます。NOEL \geq 30 モデルで陽性、NOEL \geq 300 モデルで陰性であれば、当該物質の NOEL は 30mg/kg/day 以上、300mg/kg/day 未満であると解釈できます。

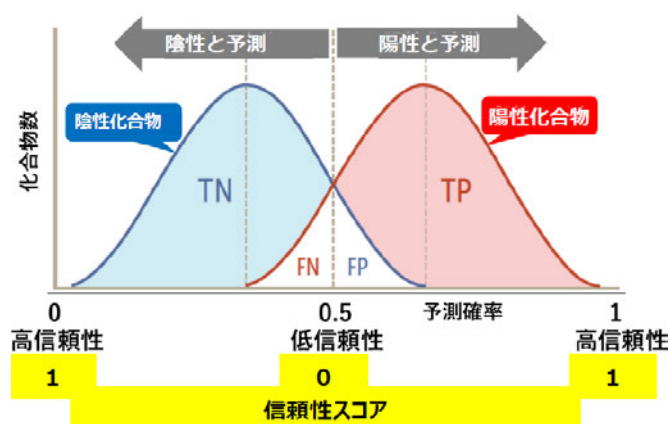
NOEL \geq 30 モデルで陰性、NOEL \geq 300 モデルで陽性という矛盾した結果が出ることがあります。このような予測については、信頼性が低いと解釈してください。



読み方 信頼性スコア(reliability score)について

モデルの出力として得られる 0~1 の範囲の予測確率(probability)から Youden 法で最適化したカットオフ値を算出し、式 1 を用いて補正することによってカットオフ値が 0.5 となる補正済予測確率を算出しました。さらに式 2 を用いてこの補正済予測確率と 0.5 の差の絶対値の 2 倍(区間[0,1]に変換するため)を信頼性スコアとしました。

各モデルの予測確率が 0 (negative)もしくは 1 (positive)であれば、信頼性スコアは 1.0 となります。また、補正済予測確率が 0.5 であれば、信頼性スコアは 0 となります。



$$\text{corrected probability} = \text{probability}^{-\frac{\log 2}{\log \text{cutoff}}} \quad (\text{式 1})$$

probability: モデルの出力した予測確率

cutoff: 2 値化の閾値

$$\text{reliability score} = 2 \times |\text{corrected probability} - 0.5| \quad (\text{式 2})$$

信頼性スコアが 1 に近いほど予測結果の信頼性が高いと解釈できますが、下記に留意してください。

- ① 同じモデルでは、物質間の信頼性スコアの比較は可能である。
例えば、「肝毒性/細胞障害・炎症」で予測結果が 1 (毒性あり) である 2 種類の物質について、片方が信頼性 0.8、もう片方が信頼性 0.1 の時、前者は毒性ありの可能性が高いが、後者は毒性ありの予測が誤っている可能性も考えられる。
- ② 異なるモデル間での信頼性スコアの比較はできない。
例えば、「肝毒性/細胞障害・炎症」の信頼性 0.7 と「血液毒性/凝固/NOEL」の信頼性 0.7 の意味合いは異なる。

また、予測結果の信頼性を検討する際は、信頼性スコアだけでなく、下記の情報も踏まえて複合的に判断してください。

- ・ケミカルスペース上にマッピングした学習データ群と予測対象物質の位置関係
- ・主要な mordred 記述子についての学習データ群の最大—最小値と、予測対象物質の値
- ・類似物質の毒性試験の結果

読み方 **【重要】信頼性スコアを用いた適用範囲の考え方**

モデルの外部評価において、信頼性スコアが一定の値（閾値）より大きい物質の予測精度は、全体の予測精度より高くなる傾向が見られました。そこで、この閾値より大きい領域を適用範囲（applicability domain; AD）と捉え、最適な閾値を設定しました。具体的には、閾値を大きくすると物質のカバー率（適用範囲に入る物質数）が低下するため、信頼性スコアが 0~0.4 の範囲で、かつカバー率がある程度確保できる範囲で、AD 内の精度が最も高くなる値を AD の閾値と設定しました。この AD 閾値は、精度評価に用いる物質群によってばらつくため、25 回の試行を行った。その結果に対して Bootstrap 検証を実施した結果を以下に示します。Mean, Standard deviation, Range は、25 回の試行で得られた 25 個の AD 閾値の平均値とその分散、AD 閾値の幅（最小値と最大値）を示しています。BA(balanced accuracy)はこの閾値を求めた際の各モデルの予測精度を表します。予測結果の信頼性スコアが以下に示す AD 閾値より大きければ AD 内と捉えることができ、その時の精度は下記で示す BA 程度ということです。

モデル名	AD 内外の閾値となる信頼性スコア			Mean BA
	Mean	Standard deviation	Range	
EP_BL01_NOEL_300	0.30	0.09	[0.09, 0.40]	0.78
EP_BL02_NOEL_300	0.34	0.07	[0.11, 0.40]	0.77
EP_KD01_NOEL_300	0.33	0.08	[0.04, 0.40]	0.74
EP_LV01_NOEL_300	0.22	0.06	[0.07, 0.30]	0.80
EP_LV02_NOEL_300	0.16	0.06	[0.05, 0.28]	0.77
EP_LV03_NOEL_300	0.32	0.10	[0.04, 0.40]	0.71
EP_LV04_NOEL_300	0.35	0.05	[0.23, 0.40]	0.76
EP_LV05_NOEL_300	0.36	0.05	[0.13, 0.40]	0.84
EP_LV06_NOEL_300	0.36	0.06	[0.14, 0.40]	0.76
EP_BL01_NOEL_30	0.13	0.08	[0.02, 0.27]	0.80
EP_BL02_NOEL_30	0.13	0.13	[0.00, 0.40]	0.77
EP_KD01_NOEL_30	0.15	0.14	[0.01, 0.40]	0.73
EP_LV01_NOEL_30	0.21	0.11	[0.07, 0.38]	0.80
EP_LV02_NOEL_30	0.20	0.12	[0.00, 0.39]	0.77
EP_LV03_NOEL_30	0.06	0.03	[0.00, 0.13]	0.82
EP_LV04_NOEL_30	0.09	0.09	[0.00, 0.34]	0.80
EP_LV05_NOEL_30	0.23	0.10	[0.01, 0.36]	0.82
EP_LV06_NOEL_30	0.24	0.11	[0.01, 0.39]	0.77

6.1.2. インビトロ実験予測結果

予測結果詳細画面において、[モデル名]をクリックするとインビトロ実験予測結果画面に遷移します。

The image shows two screenshots from the AI-SHIPS system. The left screenshot shows a search results table with a red box highlighting a model name 'NOELs30モデル' and a red arrow pointing to a '予測・検索結果' (Prediction/Search Results) button. A blue arrow points from this button to the right screenshot. The right screenshot shows the 'インビトロ実験予測値' (In vitro Experiment Prediction Value) page for the selected model. It features a table with columns for 'カテゴリ' (Category), 'モデル名' (Model Name), '重要度' (Importance), '予測結果' (Prediction Result), and '20220131142227255'. The table lists various models and their predicted results, with the top row (NOELs30モデル) circled in red and labeled with a circled '4'.

カテゴリ	モデル名	重要度	予測結果	20220131142227255
化学構造				
	④ 肝毒性細胞傷害・炎症 NOELs30モデル		予測結果・信頼性スコア	陰性 0.5992
薬物代謝_P450	rCYP1A1-A1	0.0000	予測結果・信頼性スコア	陰性 0.7572
薬物代謝_P450	rCYP1A2-A1	39.6068	予測結果・信頼性スコア	陰性 0.3742
薬物代謝_P450	③ rCYP2B1-A1	0.0000	予測結果・信頼性スコア	陰性 0.4458
薬物代謝_P450	rCYP2C6-A1	46.3618	予測結果・信頼性スコア	陰性 0.9745
薬物代謝_P450	rCYP2D1-A1	41.2516	予測結果・信頼性スコア	陰性 0.7740
薬物代謝_P450	rCYP2E1-A1	0.0000	予測結果・信頼性スコア	陰性 0.6021
薬物代謝_P450	rCYP3A2-A1	36.7373	予測結果・信頼性スコア	陰性 0.6332
薬物代謝_UGT	rUGT-A1	56.2589	予測結果・信頼性スコア	陽性 0.2091
酸化還元_GSH含量	GSH-A1	0.0000	予測結果・信頼性スコア	陽性 0.1760
酸化還元_GSH含量	GSH-A2	0.0000	予測結果・信頼性スコア	陽性 0.5508

この画面では、選択された毒性予測モデル（最上段（④）にエンドポイント名とモデル名が表示されます）の予測の説明変数である各 *in vitro* 試験やmordred 記述子の重要度、予測結果、予測結果の信頼性スコアが示されます。

※毒性予測モデルにより、説明変数として使用した *in vitro* 試験の予測モデルの数は異なります。採用していない *in vitro* 試験の重要度は 0.000 と表示されます。

※*in vitro* 試験の予測モデル名（③）をクリックすると、モデルの情報とパラメータ説明がポップアップ表示されます。

The image shows a popup window for the model 'rCYP1A1-A1'. It contains the following information:

- モデル情報
- ・ P450_r1A1_MAXの値が15以上であれば1
- パラメータ説明
- ・ 組織スルホンと発光基質を用いたインビトロ試験におけるラット CYP1A1に対する阻害作用% (最大100、最低0) (3濃度における最大値)

- ・ 重要度：モデルを構築した際の説明変数が、予測にどれだけ影響しているかを数値で示したものの、0～100の値で、大きいほど重要度が高いことを示す。また、重要度が高いほど背景色が濃く表示される。
- ・ 予測結果：陽性（背景色:赤）／陰性（背景色:青）で示される。
- ・ 信頼性スコア：0～1の値。数値が大きいほど信頼性が高い。信頼性スコアが高いと背景色の色は濃く表示される。

また、この画面では以下の操作ができます。

- ・ [tsv 出力] ボタン(①)
in vitro 試験の結果を tsv ファイル形式で出力します。
- ・ [重要度]の▼(②)
 各説明変数を重要度順にソートできます。

読み方 ***in vitro* 試験と重要度について**

毒性予測モデルは、以下の 1.~4.を説明変数としていますが、いくつかのモデルでは変数選択をしています。重要度はこれらの各説明変数が毒性予測にどれだけ影響しているかを数値で示したものです。本画面の表示対象は下記の 2.と 3.です。

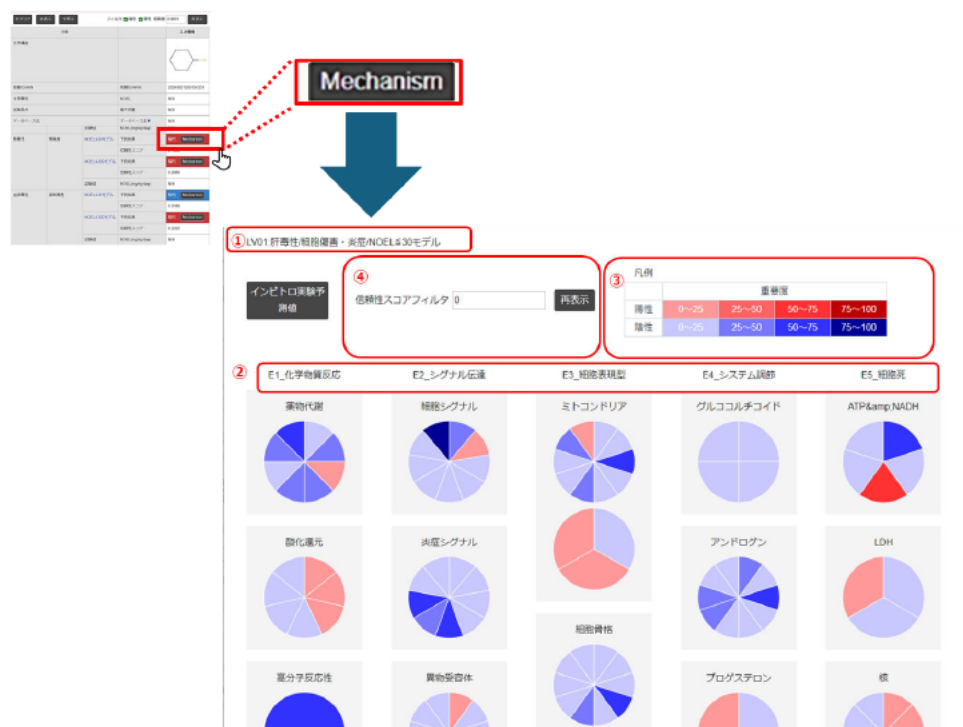
1. mordred 記述子 15 種
2. AI-SHIPS プロジェクトで実施した *in vitro* 試験結果 95 個（予測値）
3. Tox21 *in vitro* 試験結果 101 個（予測値）
4. NOEL \leq 300 モデルの予測値（NOEL \leq 30 モデルのみ）

2.や 3.の *in vitro* 試験は、毒性発現の作用機序に関連する多様な試験です。各試験結果は、パラメータ化して 2 値化しました。*in vitro* 試験の予測結果が陽性であった場合、必ずしも毒性予測結果の陽性側に寄与するというわけではありません。

6.1.3. 毒性メカニズム

予測結果詳細画面において、[Mechanism]をクリックすると毒性メカニズム画面に遷移します。

毒性メカニズム画面では、インビトロ実験予測結果(6.1.2)の内容が、毒性発現の段階別にもた、試験のカテゴリーでまとめて視覚的に表示されます。



画面左上 (①) に、表示対象のエンドポイント名とモデル名が表示されます。

上部 (②) には、左から右に向かってより最終的な毒性発現に至る作用機序に関わる5つのグループ (E1_化学物質反応～E5_細胞壊死) が表示されています。左から、化学物質反応、シグナル伝達、細胞表現型、システム調節、細胞死と、毒性経路 (パスウェイ) に沿って配置されています。

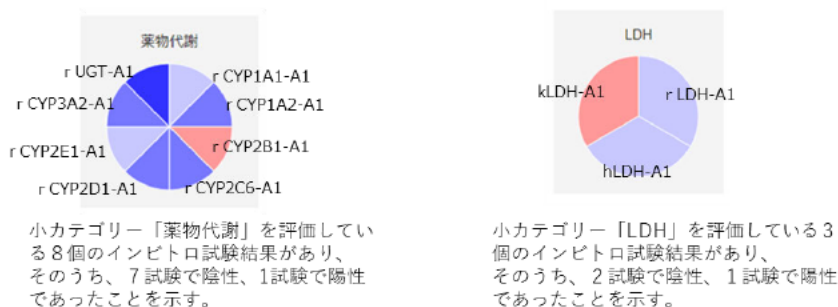


さらに、それぞれのグループの下に、関連する小カテゴリーを示す円グラフが縦に並んでいます。グループと小カテゴリーの対応関係と、小カテゴリーごとに含まれる *in vitro* 試験のエンドポイントは下表のとおりです。

グループ	小カテゴリー	小カテゴリーに含まれる <i>in vitro</i> 試験のエンドポイント
化学物質反応	薬物代謝	シトクロム P450 阻害、UDP-グルクロノシルトランスフェラーゼ阻害
	酸化還元	細胞内 GSH レベル、Nrf2 転写活性
	高分子反応性	ADRA
	胆汁排泄	BSEP 阻害、MRP2 阻害
	DNA 損傷	p53 転写活性、ATAD5 レベル、H2AX リン酸化
シグナル伝達	細胞シグナル	TGFβ/SMAD 転写活性、SHH/GLI1 転写活性、細胞内 cAMP レベル
	炎症シグナル	NF-κB/AP-1 転写活性
	異物受容体	AHR 転写活性、PXR 転写活性、CAR 転写活性

グループ	小カテゴリー	小カテゴリーに含まれる <i>in vitro</i> 試験のエンドポイント
	酵素誘導	CYP2B1 誘導、CYP3A1 誘導
	転写調節	RXRα 転写活性、HDAC 活性
細胞表現型	細胞小器官	ミトコンドリア、細胞骨格（F-アクチン）、小胞体ストレス
	脂質代謝	PPARα 転写活性、PPARδ 転写活性、LXRα 転写活性、FXR 転写活性
	脂質蓄積	中性脂肪蓄積、リン脂質蓄積
	エネルギー代謝	PPARγ 転写活性
	ストレス応答	HSF 転写活性、HIF 転写活性
システム調節	内分泌	グルココルチコイド受容体転写活性、アンドロゲン受容体転写活性、プロゲステロン受容体転写活性、エストロゲン受容体 α 転写活性、エストロゲン受容体 β 転写活性、ERR 転写活性、甲状腺ホルモン受容体転写活性、アロマターゼ活性、甲状腺刺激ホルモン受容体作用、サイトロピン放出ホルモン受容体作用
	免疫	RORγ 転写活性
	ビタミン	RAR 転写活性、VDR 転写活性
細胞死	細胞内 ATP	CellTiter アッセイ
	細胞内 NADH	WST アッセイ
	逸脱酵素	LDH 漏出
	核	ハイコンテント解析
	アポトーシス	Caspase3/7
	転写抑制	レポーターアッセイ

小カテゴリーは円グラフで示されます。円グラフはカテゴリーに含まれる *in vitro* 試験の数で等分されており、分割数が多いほど本カテゴリーに含まれる *in vitro* 試験が多いことを示します。円グラフにマウスオーバーすると、対象の *in vitro* 試験のモデル名が表示されます。それぞれの試験の結果は、陽性であれば赤、陰性は青で示されます。濃度が濃いほど重要度が高いことを示します（凡例(③)参照）。



信頼性スコアフィルタ (④) に数値(0~1) を入力して[再表示]をクリックすると、画面下部において、入力した値より大きな信頼性スコアを有する *in vitro* 試験予測モデルの予測結果のみが表示されます。

読み方 小カテゴリーと *in vitro* 試験結果の見方

下記を考慮して検討ください。

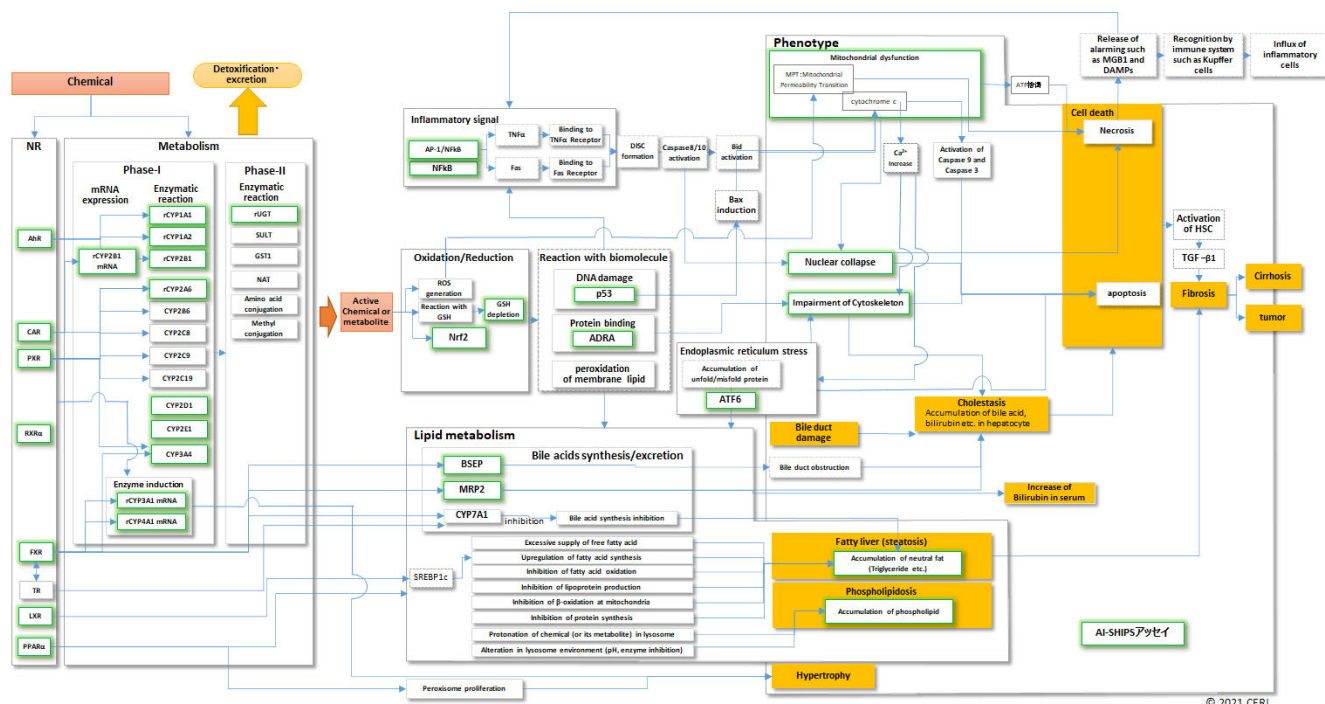
- 重要度を異なる *in vitro* 試験で比較することはできません。
- それぞれの *in vitro* 試験が何を「陽性」としているか(方向性)を確認してください。
- 陽性が強ければ、毒性発現に作用していると言えるわけではありません。陰性が強いことが毒性発現に作用していることもあります。

読み方 肝毒性に関する一般的なメカニズム

化学物質によって発現する毒性の機序は①標的組織への輸送、②細胞内分子との反応または組織内微小環境の変化、③②に起因する細胞内シグナル、細胞機能または細胞や細胞小器官の構造変化、④修復と適応過程の4段階に大別されます。

まず、①標的組織への輸送の前後で毒性を発現する物質の除去や解毒が行われます。解毒には薬物代謝酵素が重要であり、シトクロム P450 (CYP)、UGT (UDP-glucuronosyltransferase) や GST (Glutathione-S-transferase) 等があります。この他、加水分解、酸化等があります。薬物代謝酵素の誘導には AhR (Aryl Hydrocarbon Receptor) や種々の核内受容体 (CAR, PXR, FXR 等) が関与しています。標的組織に輸送されると、「②細胞内分子との反応」として、DNA、生体内タンパク質、脂質を主な標的とした反応が起こります。これら反応には、共有結合を起こす反応、受容体や酵素等への結合、酸化・還元等があり、こうした反応の結果として細胞内分子の機能が変化します。変化の種類には、活性化、阻害、破壊、免疫応答があります。細胞内分子とは反応せず、「②細胞内微小環境の変化」を起こさせる場合もあります。こうした変化の後に③および④の工程へと進み、毒性が発現します。

AI-SHIPS プロジェクトで整理した肝毒性の一般的なメカニズムを下図に示します。肝毒性における急性肝障害の障害タイプとして、肝細胞死 (壊死 (Necrosis) およびアポトーシス (Apoptosis))、脂肪肝 (Fatty liver)、胆汁うっ滞 (Cholestasis) などがあり、慢性肝障害の障害タイプとして、肝炎、肝硬変 (Cirrhosis)、リン脂質症 (Phospholipidosis)、肥大 (Hypertrophy)、肝細胞がん、血管肉腫などがあります。図中の緑で囲んだ部分は、AI-SHIPS プロジェクトで関連する *in vitro* 試験を実施していることを示します。



6.2. 体内動態予測結果

6.2.1. 体内動態予測結果

属性エリア (①) の表示項目は下記の通りです。

表示項目		説明
化学構造		予測対象化学物質の構造が表示される。
物質 ID/ARN	物質 ID/ARN	予測対象化学物質の物質 ID が表示される。
全身毒性	NOEL	(毒性予測結果では N/A と表示)
試験条件	最大用量	(毒性予測結果では N/A と表示)
データベース名	データベース名	(毒性予測結果では N/A と表示)



毒性予測結果エリア (②) の表示は下記の通りです。

表示項目			説明
膜透過	膜透過係数	PappAtoB	nm/s 消化管から血液側への膜透過係数(nm/s)
		PappBtoA	nm/s 血液側から消化管への膜透過係数(nm/s)
体内動態	吸収	FaFg	割合 消化管利用率
		インピボ吸収速度定数	予測値(1/h) ラットに物質を経口投与した際の吸収速度定数(1/h)
	分布	インピボ分布容積	予測値(L) 物質の分布容積 (血中から臓器移行の目安を物質が及ぶ体積で提示) (L)
		肝代謝消失	インピボ肝固有クリアランス 予測値(L/h) ラット肝での物質の代謝変換効率(L/h)
	腎排泄	インピボ腎クリアランス 予測値(L/h) ラット腎での物質の代謝変換効率(L/h)	
	血液 (血漿)	単回投与 AUC	予測値((μg/L)/h) ラット単回投与における血中濃度曲線下面積((μg/L)/h)
		単回投与 Cmax	予測値(μg/L) ラット単回投与における最高血中濃度 (μg/L)
		28日間反復投与 AUC	予測値((μg/L)/h) ラット 28日間反復投与における血中濃度曲線下面積((μg/L)/h)
		28日間反復投与 Cmax	予測値(μg/L) ラット 28日間反復投与における最高血中濃度(μg/L)
	肝臓	単回投与 AUC	予測値((μg/L)/h) ラット単回投与における肝臓中濃度曲線下面積((μg/L)/h)
		単回投与 Cmax	予測値(μg/L) ラット単回投与における最高肝臓中濃度 (μg/L)
		28日間反復投与 AUC	予測値((μg/L)/h) ラット 28日間反復投与における肝臓中濃度曲線下面積((μg/L)/h)
		28日間反復投与 Cmax	予測値(μg/L) ラット 28日間反復投与における最高肝臓中濃度(μg/L)
	腎臓	単回投与 AUC	予測値((μg/L)/h) ラット単回投与における腎臓中濃度曲線下面積((μg/L)/h)
		単回投与 Cmax	予測値(μg/L) ラット単回投与における最高腎臓中濃度 (μg/L)
		28日間反復投与 AUC	予測値((μg/L)/h) ラット 28日間反復投与における腎臓中濃度曲線下面積((μg/L)/h)
28日間反復投与 Cmax		予測値(μg/L) ラット 28日間反復投与における最高腎臓中濃度(μg/L)	

※ クリックするとメニュー画面に移ります。

読み方

例えばインビボ肝固有クリアランス（クリアランス）については、数値が低ければ、体内からの消失が遅いことを意味します。

また、臓器中の濃度について単回投与と比較して 28 日間反復投与で最大濃度の予測値が大きければ、蓄積性が高い物質であることを意味し、毒性発現のリスクが高いと判断できます。

本システムでは、体内動態予測と毒性予測は独立したモデルであり、毒性予測結果に体内動態予測の結果が考慮されていません。体内動態予測結果をみて、そもそも蓄積せず体内から早急に消失する物質については毒性発現のリスクは小さいと判断できます。

また、体内動態予測では代謝物の考慮は非常に重要ですが、本モデルでは考慮されていません。実験あるいは別途代謝物予測システムで予測した物質を本システムの入力対象として、代謝物の体内動態を検討することを強く推奨します。代謝候補物質生成が可能なシステムは商用を含めると数多く存在します。無償のものでは、GLORYx⁶、SyGMA⁷、BioTransformer⁸などがあります。精度や得意領域などを踏まえて利用することを推奨します。

6.2.2. 体内動態予測結果の可視化

予測結果詳細画面において、[体内動態]をクリックすると体内動態可視化の画面に遷移します。

入力物質がラットに投与された際に、各臓器でどのように濃度が変化するかをアニメーションで確認することができます。

- 1) 投与のパターンを選択します。
単回投与か 28 日間投与のどちらかを選択します。
- 2) 再生ボタンをクリックします。

データベース名	検索連関	単位	データベース名	単位
Papp A to B	nm/s		N/A	
Papp B to A	nm/s		1.04.0004	
体内動態	吸収	FaFg	割合	0.9221
	分布	インビボ吸収速度定数	予測値(L/h)	1.5644
	肝代謝消失	インビボ分布容積	予測値(L)	0.4470
		インビボ肝固有クリアランス	予測値(L/h)	17.9102

【画面左側】

下記項目の時間的推移が濃淡でアニメーション表示されます。

[濃度表示](①)をチェックすると数値も表示されます。

【画面右側】

下記項目の経時変化がグラフで表示されます。



⁶ <https://nerdd.univie.ac.at/gloryx/>

⁷ <https://github.com/3D-e-Chem/sygma>

⁸ <https://bitbucket.org/djombou/biotransformerjar/>

- ・ 小腸 (Gut) の吸収量 (単位 : μg)
- ・ 血液 (血漿) (Central (plasma))の濃度 (単位 : $\mu\text{g/L}$)
- ・ 肝臓 (Liver) の濃度 (単位 : $\mu\text{g/L}$)
- ・ 腎臓 (Kidney) の濃度 (単位 : $\mu\text{g/L}$)

6.3. 類似物質検索結果

類似物質検索を実行した際の「一覧」画面では、入力物質の右に類似物質検索の結果が表示されます。類似物質については、各毒性エンドポイントの NOEL 値が表示されます。

予測結果詳細







予測・検索条件

検索日時	肝毒性							腎毒性			血液毒性		体内動態			
	予測	細胞傷害・炎症	肝機能低下	肝機能亢進	胆管障害	肥大	脂質代謝異常	腎障害	凝固障害	貧血	投与量	LogP	fu,p	LogD (pH 6.0)	Log	
2025-02-23 05:56:03	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	300					

一覧

403件 top出力

全選択 非表示 全表示
フィルタ 陰性 陽性 信頼値 0.0001 再表示

分類		入力物質	<input type="checkbox"/> H40-Z132	<input type="checkbox"/> H40-Z129	<input type="checkbox"/> H40-Z115	<input type="checkbox"/> H40-Y864	<input type="checkbox"/> H		
化学構造									
物質ID/ARN	物質ID/ARN	ACN02275	H40-Z132	H40-Z129	H40-Z115	H40-Y864	H40-N14		
全身毒性	NOEL	N/A	0.0999	0.0155	4	1000	8.3837		
試験条件	最大用量	N/A	0.9990000000000001	0.38799999999999996	24	1000	838.372		
データベース名	データベース名	N/A	CSCl_new	CSCl_new	CSCl_new	REACH	HESS		
構造類似性	Descriptor	mordred	類似性スコア△	N/A	N/A	N/A	N/A		
	Fingerprint	Topological	類似性スコア▼	N/A	1.0000	0.6452	0.6216	0.6122	0.6053
		Morgan	類似性スコア▼	N/A	1.0000	0.3684	0.1667	0.0294	0.1538
インビトロ実験プロファイル類似性	類似性スコア▼	N/A	1.0000	0.7143	0.6286	0.7188	0.5682		
PKプロファイル類似性	類似性スコア▲	N/A	N/A	N/A	N/A	N/A	N/A		
分子重類似性	分子重差分▲	N/A	N/A	N/A	N/A	N/A	N/A		
肝毒性	細胞傷害・炎症	NOEL ≤ 30モデル	予測結果	陰性 Mechanism	N/A	N/A	N/A	N/A	
			信頼性スコア	0.1401	N/A	N/A	N/A	N/A	
		NOEL ≤ 300モデル	予測結果	陰性 Mechanism	N/A	N/A	N/A	N/A	
		信頼性スコア	0.3694	N/A	N/A	N/A	N/A		
	試験値	NOEL (mg/kg/day)	N/A	0.9990000000000001	0.38799999999999996	24	1000	838.372	
	肝機能低下	NOEL ≤ 30モデル	予測結果	陽性 Mechanism	N/A	N/A	N/A	N/A	
		信頼性スコア	0.1689	N/A	N/A	N/A	N/A		
NOEL ≤ 300モデル		予測結果	陽性 Mechanism	N/A	N/A	N/A	N/A		

属性エリア (①) の表示項目は下記の通りです。

表示項目		説明	
		入力物質	入力物質より右側に示される類似物質
化学構造		N/A	データベースから抽出された物質の構造が表示される。
物質 ID/ARN	物質 ID/ARN	N/A	データベースから抽出された類似物質の ARN が表示される。

表示項目			説明		
			入力物質	入力物質より右側に示される類似物質	
全身毒性		NOEL	N/A	データベースから抽出された類似物質の NOEL が表示される。	
試験条件		最大用量	N/A	データベースから抽出された類似物質の毒性試験での最大投与量が示される。	
データベース名		データベース名	N/A	データベースから抽出された類似物質のデータベース名が示される。	
構造類似性	Descriptor	mordred	類似性スコア	N/A	データベースから抽出された類似物質と入力物質の mordred の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。
	Fingerprint	Topological	類似性スコア	N/A	データベースから抽出された類似物質と入力物質の Fingerprint(Topological) の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。
		Morgan	類似性スコア	N/A	データベースから抽出された類似物質と入力物質の Fingerprint(Morgan) の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。
インビトロ実験プロファイル類似性		類似性スコア	N/A	データベースから抽出された類似物質と入力物質の Fingerprint(Morgan) の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。	
PK プロファイル類似性		類似性スコア	N/A	データベースから抽出された類似物質と入力物質の Fingerprint(Morgan) の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。	
分子量類似性		分子量差分	N/A	データベースから抽出された類似物質と入力物質の Fingerprint(Morgan) の類似性が 0～1 の範囲で示される (1 に近いと類似性が高い)。	

毒性予測結果エリア (2) の表示項目は下記の通りです。

表示項目			説明		
			入力物質	入力物質より右側に示される類似物質	
腎毒性	腎障害	NOEL ≥ 30 モデル ※文字列をクリックすると各 <i>in vitro</i> 試験の結果ページに遷移する (6.1.2 参照)	予測結果	腎障害に関する NOEL ≥ 30 かどうかの 2 値判定モデルでの予測結果が陽性/陰性で示される。陽性の場合には 陽性 Mechanism、陰性の場合には 陰性 Mechanism と表示される。 陽性とは腎障害の NOEL が 30mg/kg/day 以上であることを示す。 陰性とは腎障害の NOEL が 30mg/kg/day 未満であることを示す。 Mechanism をクリックすると <i>in vitro</i> 試験の結果が表示される (→ 6.1.3 章参照)	N/A
			信頼性スコア	上記の予測結果の信頼性スコア。0～1 の範囲の値で、1 に近いほどモデルの予測が明確であり、0 に近いほど予測が曖昧であることを表す。	N/A
		NOEL ≥ 300 モデル	予測結果	腎障害に関する NOEL ≥ 300 かどうかの 2 値判定モデルでの予測結果が陽性/陰性で示される。陽性の場合には 陽性 Mechanism、陰性の場合には 陰性 Mechanism と表示される。 陽性とは腎障害の NOEL が 300mg/kg/day 以上であることを示す。 陰性とは腎障害の NOEL が 300mg/kg/day 未満であることを示す。 Mechanism をクリックすると <i>in vitro</i> 試験の結果が表示される (→ 6.1.3 参照)。	N/A

表示項目				説明	
				入力物質	入力物質より右側に示される類似物質
		信頼性スコア		上記の予測結果の信頼性スコア。0～1の範囲の値で、1に近いほどモデルの予測が明確であり、0に近いほど予測が曖昧であることを表す。	N/A
		試験値	NOEL(mg/kg/day)	N/A	物質の有害性試験の腎障害に関するNOEL(mg/kg/day)
…以下、予測対象として選択された数だけ繰り返し 肝毒性（細胞傷害・炎症） 肝毒性（肝機能低下）、肝毒性（肝機能亢進）、 肝毒性（胆管障害、肥大）、肝毒性（脂質代謝異常） 腎毒性（腎障害）、血液毒性（凝固障害、貧血）				(省略)	(省略)
膜透過	膜透過係数	PappAtoB	nm/s	消化管から血液側への膜透過係数(nm/s)	N/A
		PappBtoA	nm/s	血液側から消化管への膜透過係数(nm/s)	N/A
体内動態	吸収	FaFg	割合	消化管利用率	N/A
		インビボ吸収速度定数	予測値(1/h)	ラットに物質を経口投与した際の吸収速度定数(1/h)	N/A
	分布	インビボ分布容積	予測値(L)	物質の分布容積（血中から臓器移行の目安を物質が及ぶ体積で提示）(L)	N/A
		肝代謝消失	インビボ肝固有クリアランス	予測値(L/h)	ラット肝での物質の代謝変換効率(L/h)
	腎排泄	インビボ腎クリアランス	予測値(L/h)	ラット腎での物質の代謝変換効率(L/h)	N/A
	血液(漿)	単回投与AUC	予測値((μ g/L)/h)	ラット単回投与における血中濃度曲線下面積((μ g/L)/h)	N/A
		単回投与Cmax	予測値(μ g/L)	ラット単回投与における最高血中濃度(μ g/L)	N/A
		28日間反復投与AUC	予測値((μ g/L)/h)	ラット28日間反復投与における血中濃度曲線下面積((μ g/L)/h)	N/A
		28日間反復投与Cmax	予測値(μ g/L)	ラット28日間反復投与における最高血中濃度(μ g/L)	N/A
	肝臓	単回投与AUC	予測値((μ g/L)/h)	ラット単回投与における肝臓中濃度曲線下面積((μ g/L)/h)	N/A
		単回投与Cmax	予測値(μ g/L)	ラット単回投与における最高肝臓中濃度(μ g/L)	N/A
		28日間反復投与AUC	予測値((μ g/L)/h)	ラット28日間反復投与における肝臓中濃度曲線下面積((μ g/L)/h)	N/A
		28日間反復投与Cmax	予測値(μ g/L)	ラット28日間反復投与における最高肝臓中濃度(μ g/L)	N/A
	腎臓	単回投与AUC	予測値((μ g/L)/h)	ラット単回投与における腎臓中濃度曲線下面積((μ g/L)/h)	N/A
		単回投与Cmax	予測値(μ g/L)	ラット単回投与における最高腎臓中濃度(μ g/L)	N/A
		28日間反復投与AUC	予測値((μ g/L)/h)	ラット28日間反復投与における腎臓中濃度曲線下面積((μ g/L)/h)	N/A
28日間反復投与Cmax		予測値(μ g/L)	ラット28日間反復投与における最高腎臓中濃度(μ g/L)	N/A	

読み方 類似物質情報の活用について

予測対象物質の各毒性エンドポイントの予測値は「NOEL \leq 300 か否か」、「NOEL \leq 30 か否か」ですが、類似物質については、各毒性エンドポイントの毒性試験の NOEL 値（実測値）が表示されます。実測値を「NOEL \leq 300 か否か」、「NOEL \leq 30 か否か」に当てはめて、比較すると良いでしょう。予測結果と類似物質で、2 値判定結果が異なるモデルについては、その信頼性スコアを確認し、その値及びその付近の値以下のスコアの結果は用いるべきではないと考えられます。

読み方 各エンドポイントの NOEL について

AI-SHIPS データベースには、HESS、REACH 登録情報、ToxRef および化審法新規審査シートを情報源として、下記の条件に合致するものを収載しています。

- OECD TG407（28 日間反復経口投与毒性試験）、OECD TG422（反復投与毒性試験と生殖/発生毒性スクリーニング試験の併合試験）またはこれに相当する毒性試験
- 投与期間が 28 日以上 90 日未満（亜急性毒性に該当する）
- 雄のデータ（ただし、雌雄が不明の場合も採用）

HESS には、試験データごとに、反復投与毒性試験結果の詳細情報、すなわち血液生理化学検査、病理組織学的検査等 400 を超える項目についての NOEL 値が収載されています。肝臓、血液、腎臓の各毒性エンドポイントと関連のある毒性所見を抽出し、対応する毒性所見の NOEL 値の最低値を毒性エンドポイントの NOEL 値としました。REACH 登録情報、ToxRef および化審法新規審査シートについても HESS と同じ項目で整理し、各毒性エンドポイントの NOEL 値を得ています。

各毒性エンドポイントに対応付けた毒性所見は下記の通りです。

毒性エンドポイント	HESS の <i>in vivo</i> の毒性所見 *
肝毒性：細胞傷害・炎症 (EP_LV01)	Blood chemical examination·Blood serum (Enzyme)·<->·ALT(GPT) ↑
	Histopathological findings·Liver·<->·Cell infiltration/Infiltration
	Histopathological findings·Liver·<->·Cell infiltration/Inflammation
	Histopathological findings·Liver·Hepatocyte·Necrosis hepatocyte (Centrilobular)
	Histopathological findings·Liver·Hepatocyte·Necrosis hepatocyte (Other)
肝毒性：肝機能低下 (EP_LV02)	Histopathological findings·Liver·Hepatocyte·Necrosis hepatocyte (Periportal)
	Blood chemical examination·Blood serum (Enzyme)·<->·A/G ↓
	Blood chemical examination·Blood serum (Enzyme)·<->·Alb ↓
	Blood chemical examination·Blood serum (Enzyme)·<->·T. protein ↓
	Organ weights·Liver·<->·Absolute organ weight ↓
肝毒性：肝機能亢進 (EP_LV03)	Organ weights·Liver·<->·Relative organ weight ↓
	Blood chemical examination·Blood serum (Enzyme)·<->·A/G ↑
	Blood chemical examination·Blood serum (Enzyme)·<->·Alb ↑
	Blood chemical examination·Blood serum (Enzyme)·<->·T. protein ↑
	肝毒性：胆管障害 (EP_LV04)
Blood chemical examination·Blood serum (Enzyme)·<->·ALP ↑	
Blood chemical examination·Blood serum (Enzyme)·<->· γ -GTP ↑	
Histopathological findings·Liver·Bile duct·Bile thrombus	
Histopathological findings·Liver·Hepatocyte·Necrosis hepatocyte (Periportal)	
肝毒性：肥大 (EP_LV05)	Histopathological findings·Liver·Hepatocyte·Hypertrophy/Swelling hepatocyte (Centrilobular)
	Histopathological findings·Liver·Hepatocyte·Hypertrophy/Swelling hepatocyte (Other)
	Histopathological findings·Liver·Hepatocyte·Hypertrophy/Swelling hepatocyte (Periportal)
	Necropsy·Liver·<->·Enlarged (Necropsy)
	Organ weights·Liver·<->·Absolute organ weight ↑
	Blood chemical examination·Blood serum (Lipid)·<->·T. cholesterol ↑

毒性エンドポイント	HESS の <i>in vivo</i> の毒性所見 *
肝毒性：脂質代謝異常 (EP_LV06)	Blood chemical examination·Blood serum (Lipid)·<>·T. cholesterol ↓
	Blood chemical examination·Blood serum (Lipid)·<>·Triglyceride ↑
	Blood chemical examination·Blood serum (Lipid)·<>·Triglyceride ↓
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↑ (Centrilobular)
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↑ (Other)
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↑ (Periportal)
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↓ (Centrilobular)
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↓ (Other)
	Histopathological findings·Liver·Hepatocyte·Fatty change/Vacuolization ↓ (Periportal)
血液毒性：凝固 (EP_BL01)	Hematological examination·Blood cell (Coagulation)·<>·APTT ↑
	Hematological examination·Blood cell (Coagulation)·<>·CT ↑
	Hematological examination·Blood cell (Coagulation)·<>·Fibrinogen ↓
	Hematological examination·Blood cell (Coagulation)·<>·PT ↑
血液毒性：貧血 (EP_BL02)	Hematological examination·Blood cell (Erythrocyte)·<>·HCT ↓
	Hematological examination·Blood cell (Erythrocyte)·<>·HGB ↓
	Hematological examination·Blood cell (Erythrocyte)·<>·RBC ↓
	Hematological examination·Blood cell (Erythrocyte)·<>·Reticulocyte ↑
腎毒性：腎障害・機能障害 (EP_KD01)	Blood chemical examination·Blood serum (Nitrogen compound)·<>·BUN ↑
	Blood chemical examination·Blood serum (Nitrogen compound)·<>·Creatinine ↑
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Atrophy
	Histopathological findings·Kidney·Renal tubule (Other)·Atrophy
	Histopathological findings·Kidney·Renal tubule (Proximal)·Atrophy
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Basophilic change/Regeneration/Degeneration
	Histopathological findings·Kidney·Renal tubule (Other)·Basophilic change/Regeneration/Degeneration
	Histopathological findings·Kidney·Renal tubule (Proximal)·Basophilic change/Regeneration/Degeneration
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Casts
	Histopathological findings·Kidney·Renal tubule (Other)·Casts
	Histopathological findings·Kidney·Renal tubule (Proximal)·Casts
	Histopathological findings·Kidney·<>·Cell infiltration/Infiltration
	Histopathological findings·Kidney·<>·Cell infiltration/Inflammation
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Desquamation tubular epithelium cell
	Histopathological findings·Kidney·Renal tubule (Other)·Desquamation tubular epithelium cell
	Histopathological findings·Kidney·Renal tubule (Proximal)·Desquamation tubular epithelium cell
	Histopathological findings·Kidney·Pelvis/Papilla·Enlarged
	Histopathological findings·Kidney·Pelvis/Papilla·Hyperplasia epithelium
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Hypertrophy epithelium
	Histopathological findings·Kidney·Renal tubule (Other)·Hypertrophy epithelium
	Histopathological findings·Kidney·Renal tubule (Proximal)·Hypertrophy epithelium
	Histopathological findings·Kidney·Pelvis/Papilla·Necrosis
	Histopathological findings·Kidney·Renal tubule (Distal/Collecting duct)·Necrosis
	Histopathological findings·Kidney·Renal tubule (Other)·Necrosis
Histopathological findings·Kidney·Renal tubule (Proximal)·Necrosis	
Histopathological findings·Kidney·<>·Nephropathy	
Necropsy·Kidney·<>·Enlarged (Necropsy)	
Organ weights·Kidney·<>·Absolute organ weight ↑	

*「Examination items」、「Organ(Tissue)」、「Tissue」、「Effect」をハイフン「-」で連結して示している。空欄の場合は該当項目を「<>」とした。

7. その他の機能

7.1. データベース検索

メニューバーの[データベース検索]をクリックして、データベース検索画面に遷移します。



7.1.1. データベース検索の条件入力

- 1) 検索対象データベースを選択します。
- 2) 検索したい情報を入力します。

以下の項目で検索できます。いずれかの方法で検索条件を入植します。

- ① 物質名
- ② InChIKey
- ③ ARN (AI-SHIPS プロジェクトにおける毒性試験番号)
- ④ CAS 登録番号 (CAS RN[®])
- ⑤ Molecule Editor を使用した部分化学構造入力 (部分構造一致)
- ⑥ Mol ファイルまたは SD ファイルのアップロード (部分構造一致)
- ⑦ SMILES (部分構造一致)

* ARN、CAS 登録番号 (CAS RN[®]) については、検索した履歴が画面上に残ります。履歴をクリアしたい場合には、[clear]ボタンをクリックしてください。

- 3) [検索]をクリックします。



7.1.2. データベース検索結果

画面下部に入力した部分構造をもつ物質のデータが表示されます。

表示される内容は以下の通りです：化学構造、データベース名、ARN、物質名、InChIKey、CAS 登録番号 (CAS RN[®])、SMILES、最大用量 (毒性試験での最大投与量)、各毒性 EP の NOEL (mg/kg/day)

※AI-SHIPS データベースには、1 物質で複数の試験結果を有する物質が含まれています。

表示件数 (①)：表示する件数を指定して変更します。

- [tsv 出力]ボタン (②)：検索結果を tsv 形式にて出力します。



分類	48671	48674	48676	48677	48682	48729	48741	48766	48806	48807	
化学構造											
データベース名	HES5	HES5	HES5	HES5	HES5	HES5	HES5	HES5	HES5	HES5	
ARN	H4D-E066	H4D-E069	H4D-E061	H4D-E062	H4D-E067	H4D-E114	H4D-E126	H4D-E151	H4D-E191	H4D-E192	
物質名	1-Chloronaphthalene	1-Methoxynaphthalene	1-Naphthyl-4-sulfonic acid sodium salt	1-Naphthylacetic acid	1-methylnaphthalene	2,6-Naphthalenedicarboxylic acid, dimethyl ester	2-Amino-1-naphthalenesulfonic acid	2-Isocytosynaphthalene	3-Hydroxy-2-naphthol	3-Hydroxy-2-naphthol	
InChIKey	JTPNRXUJCXHCXMH-UHFFFAOYSA-N	NQMJGNUMFTYCHC-UHFFFAOYSA-N	IAAVNVCARVEFS-UHFFFAOYSA-M	PRPNYUQVPRX-UHFFFAOYSA-N	QRUYECJOLPKSFR-UHFFFAOYSA-N	QYUJMLEYMPKZAZ-UHFFFAOYSA-N	QVIAALJASRVOA-UHFFFAOYSA-N	XCHIHDSOMWMS-UHFFFAOYSA-N	JFQQAHLVUBOP-UHFFFAOYSA-N	JFQQAHLVUBOP-UHFFFAOYSA-N	
CAS RN	CAS90-13-1	CAS2216-69-6	CAS64-07-7	CAS66-07-3	CAS90-12-0	CAS640-65-3	CAS61-16-3	CAS2173-67-1	CAS92-77-3	CAS92-77-3	
SMILES	Clc1ccc2c1cccc2	COc1ccc2c1cccc2	O=C(O)(=O)c1ccc2c1cc(O)=O	OC(=O)c1ccc2c1cc(O)=O	Cc1ccc2c1cccc2	CCOC(=O)c1ccc2c1cc(O)=O	Nc1ccc2c1c(S(=O)(=O)O)cc2	OC1C=CC2=CC=CC=C2O1	O=C1C=CC2=CC=CC=C2O1	O=C1C=CC2=CC=CC=C2O1	
最大用量	257.1	991	871.5607	623.125	243	999	969	495.5	1012	1012	
社会性	経口経食 - 公定	NOEL	55.7	297.3	871.5607	124.625	243	999	969	99.1	1012
	経口経食 - 公定	NOEL	257.1	99.1	871.5607	623.125	243	999	969	495.5	1012
	経口経食 - 公定	NOEL	257.1	99.1	871.5607	623.125	243	999	969	495.5	1012
	経口経食 - 公定	NOEL	257.1	99.1	871.5607	124.625	243	999	197.8	495.5	1012
	経口経食 - 公定	NOEL	257.1	99.1	871.5607	623.125	48.6	999	197.8	99.1	1012

7.2. ユーザーデータの登録

メニューバーの[管理メニュー]の[ユーザーデータ管理]をクリックして、ユーザーデータ管理画面に遷移します。



ユーザーが保有する毒性試験の情報をシステムのデータベースに登録します。登録したデータは、類物質検索やデータベース検索で利用できます。

登録したユーザーデータは、全てのユーザーが閲覧（検索結果として）できる状態となりますので、企業秘密にあたる情報は登録しないでください。

- 1) [ファイルを選択]をクリックし、登録するデータを選択します。
ユーザーデータのファイル形式は8.2を参照ください。

- 2) データベース名を入力します（必須）。

The screenshot shows the 'ユーザーデータ管理' (User Data Management) page. The '新規登録' (New Registration) form is visible. Four red boxes with numbers 1 through 4 point to specific elements: 1) 'ファイルを選択' (Select File) button, 2) 'データベース名' (Database Name) input field, 3) 'コメント' (Comment) input field, and 4) '登録' (Register) button.

- 3) コメントを入力します（任意）。
- 4) [登録]をクリックします。

登録が完了すると、ユーザーデータ一覧の状態（①）に「登録済み」または「登録エラー」と表示されます（ユーザーデータ登録には時間がかかります）。「登録エラー」となった場合は、各項目の並びや構造の SMILES 等を確認の上、再登録してください。



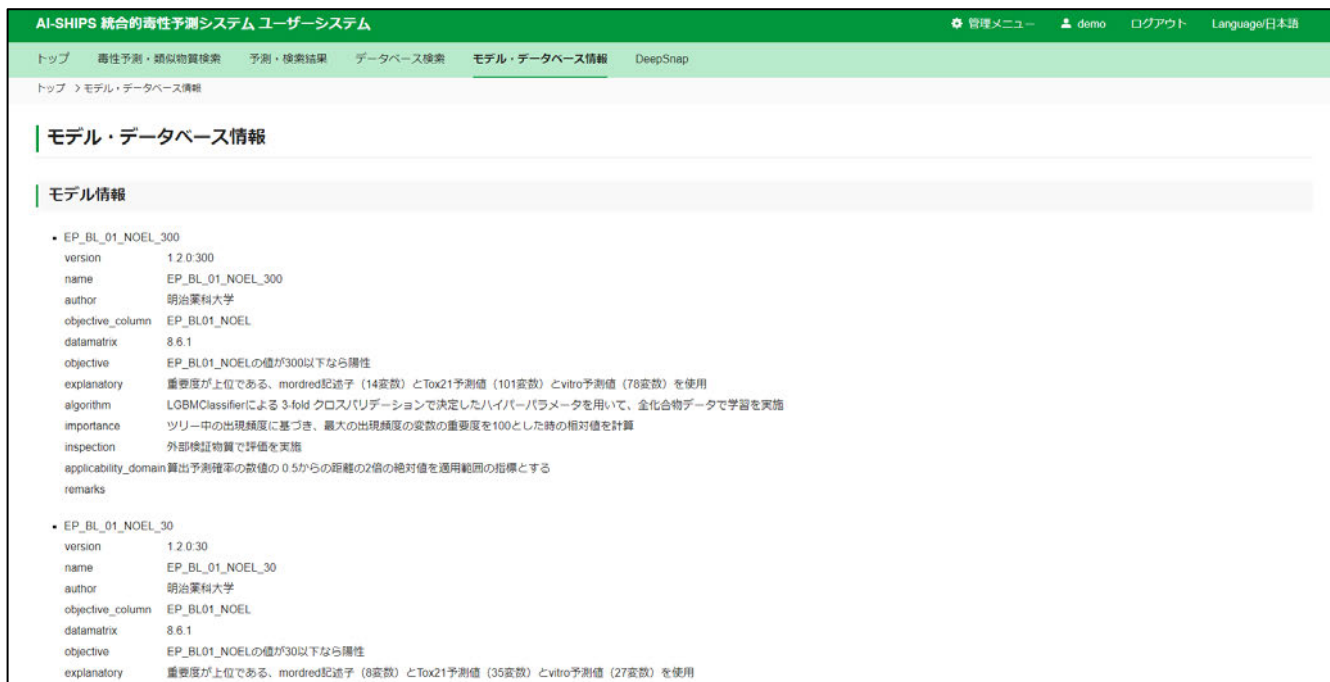
登録したデータを削除したい場合は、データベースを選択 (②) し、[削除]をクリック (③) します。

7.3. モデル・データベース情報

メニューバーの[モデル・データベース情報]をクリックして、モデル・データベース情報画面に遷移します。



システムで利用されている予測モデル、データベース情報が表示されます。



7.4. DeepSnap

「DeepSnap」とは、分子全体の特徴を画像データとして学習させる新規の構造情報入力法です。分子の立体構造をXYZ軸方向に回転させ複数の分子構造画像を取得し、それをDeep Learningに入力する事で、予測モデルを作成しています。

本機能では、初代培養ラット肝細胞を用いた7種類のCYP分子種の阻害活性について、DeepSnapの手法を用いて対象物質の毒性判別を行うとともに、予測モデルに基づく活性に寄与する分子特徴領域を图示します。

メニューバーの[DeepSnap]をクリックして、DeepSnap画面に遷移します。



1) 対象物質の化学構造を入力します。

化学構造は、以下の3つの方式で入力することができます。詳細は、3.1を参照ください。

- ① 構造式エディタ (Molecule Editor) による構造式の描画
- ② Molファイルまたは、SDファイルまたはtsv(SMILES)ファイルのアップロード
- ③ SMILESの入力



2) 予測対象 (エンドポイント)、CAM、Rotation を選択します。

CYPモデル：予測対象 (エンドポイント) を選択

ラット CYP1A1 (r1A1)、ラット CYP1A2 (r1A2)、ラット CYP2B1 (r2B1)、ラット CYP2C6 (r2C6)、ラット CYP2D1 (r2D1)、ラット CYP3A2 (r3A2)の阻害作用予測モデルのうち、1つを選択する。

CAM：Deep Learningの手法を選択

ScoreCAM：小さな変化に対する過剰反応を対策したアルゴリズムであり、従来のアルゴリズムではノイズが出ていた部分を除去し、特徴領域をより絞り込むことが可能。

GuidedGradCAM: 大まかな領域を捉えることができる GradCAM と詳細を捉える guided backpropagation を組み合わせ、両手法の欠点を補ったアルゴリズム。

Rotation：入力した構造の回転角度をX軸、Y軸、Z軸それぞれについて、0～360の範囲で指定

3) [実行]をクリックします。

4) 画面下部に実行結果が表示されます。

下記の情報が表示されます。

- 予測結果（総合的な判定結果）（①）
予測結果（陽性／陰性）、信頼度（信頼性スコア）、対象モデルが表示されます。予測対象分子に対する総合的な判定結果（多様な角度からの総合判定）を示します。
- 予測結果（単一分子画像からの判定結果）（②）
予測結果（陽性／陰性）、信頼度（信頼性スコア）、回転角、対象モデルが表示されます。画像上部に表示される予測結果は、対象分子が特定の角度の単一分子画像からの予測値に基づいており、総合的な判定結果と一致するとは限りません。
- 特徴領域（③）
指定された回転角、モデル情報と Deep Learning にて判定を行った際の特徴領域の画像が表示されます。
オリジナル：特徴抽出前の構造
ScoreCAM：赤色に近いほど、判定結果に対する寄与の強い部位であることを示す。
GuidedGradCAM：濃色であるほど、判定結果に対する寄与の強い部位であることを示す。
※評価アルゴリズムが異なるため ScoreCAM と GuidedGradCAM の結果は必ずしも一致しません。

The screenshot displays the 'Results' section of the AI-SHIPS interface. It contains two rows of analysis results for a molecule, each enclosed in a red box. Each row consists of a summary bar, a rotation angle, and three visualizations: the original molecule, ScoreCAM heatmap, and GuidedGradCAM heatmap.

Row 1:

- ① 予測結果：陽性 信頼度：0.987 モデル：r1A1
- ② 予測結果：陽性 回転角 (X: 90, Y: 0, Z: 0) モデル：r1A1
- ③ オリジナル、ScoreCAM、GuidedGradCAM

Row 2:

- ② 予測結果：陽性 回転角 (X: 0, Y: 0, Z: 0) モデル：r1A1
- ③ オリジナル、ScoreCAM、GuidedGradCAM

8. ファイルフォーマット

本システムで利用可能なファイルのフォーマットを説明します。

8.1. 毒性予測や類似物質検索をする際の入力ファイルのフォーマット

8.1.1. Mol ファイル

MDL (Molecular Design Limited、現在は BIOVIA 社の一部) の MDL Mol ファイル形式で記載ください。拡張子は、「.mol」です。V2000、V3000 の両バージョンが取り込み可能です。

- 文字コードは UTF8 とします。

```

CCCCC(C)(CC)c1ccc(O)cc1
JME 2017-02-26 Tue Sep 09 15:17:45 GMT+900 2025

16 16 0 0 0 0 0 0 0 0999 V2000
 6.0622 7.7000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 4.8498 7.0000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 4.8498 5.6000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 3.6374 4.9000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 3.6374 3.5000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2.4250 2.8000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2.4250 1.4000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1.2124 3.5000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0.0000 2.8000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 3.6374 2.1000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 3.6374 0.7000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 4.8498 0.0000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 6.0622 0.7000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 7.2747 0.0000 0.0000 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 6.0622 2.1000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 4.8498 2.8000 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

2 1 1 0 0 0 0
3 2 1 0 0 0 0
4 3 1 0 0 0 0
5 4 1 0 0 0 0
6 5 1 0 0 0 0
7 6 1 0 0 0 0
8 6 1 0 0 0 0
9 8 1 0 0 0 0
10 6 1 0 0 0 0
11 10 2 0 0 0 0
12 11 1 0 0 0 0
13 12 2 0 0 0 0
14 13 1 0 0 0 0
15 13 1 0 0 0 0
16 15 2 0 0 0 0
16 10 1 0 0 0 0

M END
    
```

1-3 行目: ヘッダーブロック
 ※1 行目の内容が物質 ID にセットされます。

←4 行目: カウント行。原子のレコード数、結合のレコード数

原子ブロック。
 x 座標、y 座標、z 座標、
 元素記号

結合ブロック。
 原子 1 のインデックス、原子 2 のインデックス、結合次数

←終端行

8.1.2. SD ファイル

SD ファイル (Structure Data File)は、Mol ファイルに物質名や物性などのフィールドを付与したものです。Mol ファイルは 1 ファイル 1 物質しか記述できませんが、SD ファイルは複数の物質を記載できます。拡張子は「.sdf」となります。標準の SD フォーマットに準じて記載ください。

- ・ 文字コードは UTF8 とします。
- ・ 体内動態予測でのパラメータの指定時には、“> <PBPK>”を記述した項目に下記パラメータを記載してください。
(省略可)
- ・ 二酸塩基/二塩基酸を選択した場合には pKa acid1 と pKa acid2、または pKa base1 と pKa base2 は必須入力となります。
* 体内動態パラメータを省略する場合には、改行をいれてください。(画面入力での体内動態パラメータと同等の入力となります。)

N o.	カラム名	説明	必須/省略可	入力形式
1	> <PBPK>	体内動態パラメータを入力する旨の宣言	省略可	体内動態パラメータを設定する場合は、必須。
2	LogP	水-オクタノール分配係数	省略可	任意の文字列
3	fu,p	血漿タンパク質結合率	省略可	任意の文字列
4	LogD pH6.0	小腸環境(pH=6.0)での分配係数。 値が高いと化学物質が小腸の膜を通過する能力が高いことを示す。	省略可	任意の文字列
5	LogD pH7.5	各臓器環境(pH=7.5)での分配係数	省略可	任意の文字列
6	二酸塩基/二塩基酸		省略可	1 (該当) / 0 (該当しない)
8	pKa acid	一酸塩基の酸解離定数	省略可	任意の文字列 (二酸塩基/二塩基酸=0の場合に設定)
9	pKa base	一塩基酸の酸解離定数	省略可	任意の文字列 (二酸塩基/二塩基酸=0の場合に設定)
10	pKa acid1	二酸塩基の第1酸解離定数	省略可 (※)	任意の文字列 (二酸塩基/二塩基酸=1の場合に設定)
11	pKa acid2	二酸塩基の第2酸解離定数	省略可 (※)	任意の文字列 (二酸塩基/二塩基酸=1の場合に設定)
12	pKa base1	二塩基酸の第1酸解離定数	省略可 (※)	任意の文字列 (二酸塩基/二塩基酸=1の場合に設定)
13	HpKa base2	二塩基酸の第2酸解離定数	省略可 (※)	任意の文字列 (二酸塩基/二塩基酸=1の場合に設定)

(※) 二酸塩基/二塩基酸、1(該当する)の場合は、pKa acid1 と pKa acid2、または pKa base1 と pKa base2 は必須

```

16 16 0 0 0 0 0 0 0 0999 V2000
.
.
.
.
<Mol ファイル形式>
M END
> <PBPK>
10      ←logP
0.2     ←fu,p
10      ←LogD pH6.0
11      ←LogD pH7.5
0       ←二酸塩基/二塩基酸
0       ←pKa acid
50      ←pKa base
$$$    ← 1 物質目の終了記号。次の行から2 物質目を記載可能。
    
```

8.1.3. tsv(SMILES)ファイル

SMILES と物質 ID、PBPK パラメータをタブ区切りで並べたものとします。フォーマットは以下の通りで、タブ区切りで並べる。1 解析 1 行で指定します。

```

SMILES <tab> 物質 ID <tab> LogP <tab> fu,p <tab> LogD pH6.0 <tab> LogD pH7.5
<tab>二酸塩基/二塩基酸 <tab> pKa acid <tab> pKa base <tab> pKa acid1 <tab> pKa
acid2 <tab> pKa base1 <tab> pKa base2
    
```

- ・ 文字コードは UTF8 とします。
- ・ <tab> はタブ区切りを意味する。省略可のパラメータを省略する場合、<tab> を連続入力し、カラム位置を合わせます。
- ・ 物質 ID は省略可です。
- ・ LogP 以降の体内動態予測のためのパラメータであり、省略可です。入力パラメータは条件入力画面と同じ順番となっており、内容の説明は 0 を参照ください。
- ・ 二酸塩基/二塩基酸は、1 (該当する) または 0 (該当しない) で記載します。1 (該当する) の場合は、pKa acid1 と pKa acid2、または pKa base1 と pKa base2 は必須です。
- ・ 行頭の文字が「#」の行はコメント行です。

N o.	カラム名	説明	必須/省 略可	入力形式
1	SMILES	SMILES	必須	任意の文字列
2	物質 ID		省略可	任意の文字列
3	LogP	水-オクタノール分配係数	省略可	任意の文字列
4	fu,p	血漿タンパク質結合率	省略可	任意の文字列

N o.	カラム名	説明	必須/省略可	入力形式
5	LogD pH6.0	小腸環境(pH=6.0)での分配係数。 値が高いと化学物質が小腸の膜を通過する能力が高いことを示す。	省略可	任意の文字列
6	LogD pH7.5	各臓器環境(pH=7.5)での分配係数	省略可	任意の文字列
7	二酸塩基/二塩基酸		省略可	1 (該当) /0 (該当しない)
8	pKa acid	一酸塩基の酸解離定数	省略可	任意の文字列
9	pKa base	一塩基酸の酸解離定数	省略可	任意の文字列
10	pKa acid1	血漿タンパク質結合率	省略可 (※)	任意の文字列
11	pKa acid2	小腸環境(pH=6.0)での分配係数。 値が高いと化学物質が小腸の膜を通過する能力が高いことを示す。	省略可 (※)	任意の文字列
12	pKa base1	投与量の単位	省略可 (※)	任意の文字列
13	HpKa base2	HESS DB の番号	省略可 (※)	任意の文字列

(※) 二酸塩基/二塩基酸、1 (該当する) の場合は、pKa acid1 と pKa acid2、または pKa base1 と pKa base2 は必須

SMILES(tsv)ファイル例 1

```
# SMILES <tab> 物質 ID <tab> LogP <tab> fu,p <tab> LogD pH6.0 <tab>
LogD pH7.5 <tab>二酸塩基/二塩基酸 <tab> pKa acid <tab> pKa base <tab>
pKa acid1 <tab> pKa acid2 <tab> pKa base1 <tab> pKa base2
CC(=O)Nc1ccc(cc1)O      test1      1      1      1      1      1
      1      1      1      1      1      1      1
```

SMILES(tsv)ファイル例 2

```
CCCCC(C)(CC)c1ccc(O)cc1      test10
C=O      50-00-0
```

8.2. ユーザーデータ登録フォーマット

ユーザーが毒性試験の情報をシステムのデータベースに登録する際のフォーマットです。ユーザーデータの登録は、管理メニューから行うことができます。7.2 を参照ください。

下記の入力項目を全てタブ区切りで並べます。文字コードは UTF8 とします。

N o.	カラム名	説明	必須/省略可	入力形式
1	ARN	プロジェクト固有の毒性試験番号	必須	任意の文字列 システム内で採番しますが、何らかの文字列を入力してください。
2	AI-SHIPS_ACN	AI-SHIPS で利用する物質番号	必須	任意の文字列 システム内で採番しますが、何らかの文字列を入力してください。
3	H_NAME	毒性試験データの出典における物質名	省略可	任意の文字列
4	H_CASno	毒性試験データの出典における CAS 登録番号 (CAS RN [®])	省略可	任意の文字列
5	H_SMILES	毒性試験データの出典における SMILES	省略可	任意の文字列
6	H_InChI	H_SMILES を InChi に変換したもの	省略可	任意の文字列
7	H_InChIKey	H_SMILES を InChiKey に変換したもの	省略可	任意の文字列
8	H_Day	毒性試験データの出典における投与日数	省略可	任意の文字列
9	H_Route	投与方法	省略可	任意の文字列
10	H_Max_dose_num	最大投与量	省略可	任意の文字列
11	H_Min_dose_num	最小投与量	省略可	任意の文字列
12	H_Unit	投与量の単位	省略可	任意の文字列
13	H_Chemical_No.(Link_to_HESS_DB)	HESS DB の番号	省略可	任意の文字列
14	Data_source	毒性試験データの出典 : "AI-SHIPS_in_vitro"、"HESS"、"ToxRef"、"REACH"、"CSCL_new"	省略可	任意の文字列 * システムではデータベース名として使用されます。省略時には、ユーザーデータ登録画面での【データベース名】が使用されます。
15	AI-SHIPS_CASno	システムで利用する CAS 登録番号 (CAS RN [®])	省略可	任意の文字列
16	AI-SHIPS_SMILES	システムで利用する SMILES	必須	入力物質の SMILES
17	AI-SHIPS_tested	<i>in vitro</i> 試験実施の有無	省略可	0 (無し)または 1 (有り)
18	molW	被験物質のモル質量	省略可	任意の数値
19	molW_after_Desalting	被験物質を脱塩化したもののモル質量	省略可	任意の数値
20	mol_conv	被験物質の脱塩化後の脱塩化前に対するモル質量比	省略可	任意の数値
21	Purity_m	純度に記載されている不等号	省略可	0 (不等号なし)または 1 (不等号あり)
22	Purity_min	純度の最小値	省略可	任意の数値
23	Purity_max	純度の最大値	省略可	任意の数値
24	Purity_conv	換算に利用した純度 (最大値と最小値の算術平均)	省略可	任意の数値
25	EP_LV01_NOEL	EP_LV01 (肝臓の細胞障害・炎症) の	省略可	任意の数値

N o.	カラム名	説明	必須/省略可	入力形式
		NOEL		
26	EP_LV02_NOEL	EP_LV02 (肝臓の肝機能低下) の NOEL	省略可	任意の数値
27	EP_LV03_NOEL	EP_LV03 (肝臓の肝機能亢進) の NOEL	省略可	任意の数値
28	EP_LV04_NOEL	EP_LV04 (肝臓の胆管障害) の NOEL	省略可	任意の数値
29	EP_LV05_NOEL	EP_LV05 (肝臓の肥大) の NOEL	省略可	任意の数値
30	EP_LV06_NOEL	EP_LV06 (肝臓の脂質代謝異常) の NOEL	省略可	任意の数値
31	EP_BL01_NOEL	EP_BL01 (血液の凝固障害) の NOEL	省略可	任意の数値
32	EP_BL02_NOEL	EP_BL02 (貧血の) NOEL	省略可	任意の数値
33	EP_KD01_NOEL	EP_KD01 (腎障害) の NOEL	省略可	任意の数値
34	EP_LV01_LOEL	EP_LV01 (肝臓の細胞障害・炎症) の LOEL	省略可	任意の数値
35	EP_LV02_LOEL	EP_LV02 (肝臓の肝機能低下) の LOEL	省略可	任意の数値
36	EP_LV03_LOEL	EP_LV03 (肝臓の肝機能亢進) の LOEL	省略可	任意の数値
37	EP_LV04_LOEL	EP_LV04 (肝臓の胆管障害) の LOEL	省略可	任意の数値
38	EP_LV05_LOEL	EP_LV05 (肝臓の肥大) の LOEL	省略可	任意の数値
39	EP_LV06_LOEL	EP_LV06 (肝臓の脂質代謝異常) の LOEL	省略可	任意の数値
40	EP_BL01_LOEL	EP_BL01 (血液の凝固障害) の LOEL	省略可	任意の数値
41	EP_BL02_LOEL	EP_BL02 (貧血の) LOEL	省略可	任意の数値
42	EP_KD01_LOEL	EP_KD01 (腎障害) の LOEL	省略可	任意の数値
43	Systemic_NOEL	試験全体における NOEL	省略可	任意の数値