

ToxTwin API

Tests de validation et guide d'utilisation

Addendum à la documentation pipeline v1.0

Entité	Version	Date	Statut
Twingital Ventures	v1.0	Mars 2026	Production
API ToxTwin	FastAPI 1.0.0 / OAS 3.1	Mars 2026	Opérationnelle

1. Interface Swagger — Documentation interactive

L'API ToxTwin expose une documentation interactive Swagger UI (OpenAPI 3.1) automatiquement générée par FastAPI. Cette interface permet de tester tous les endpoints directement depuis le navigateur, sans outil externe.

1.1 Accès à l'interface

URL Swagger UI

`http://localhost:8000/docs` (ou IP serveur:8000/docs en réseau local)

La capture d'écran ci-dessous présente l'interface telle qu'elle apparaît après le démarrage de l'API :

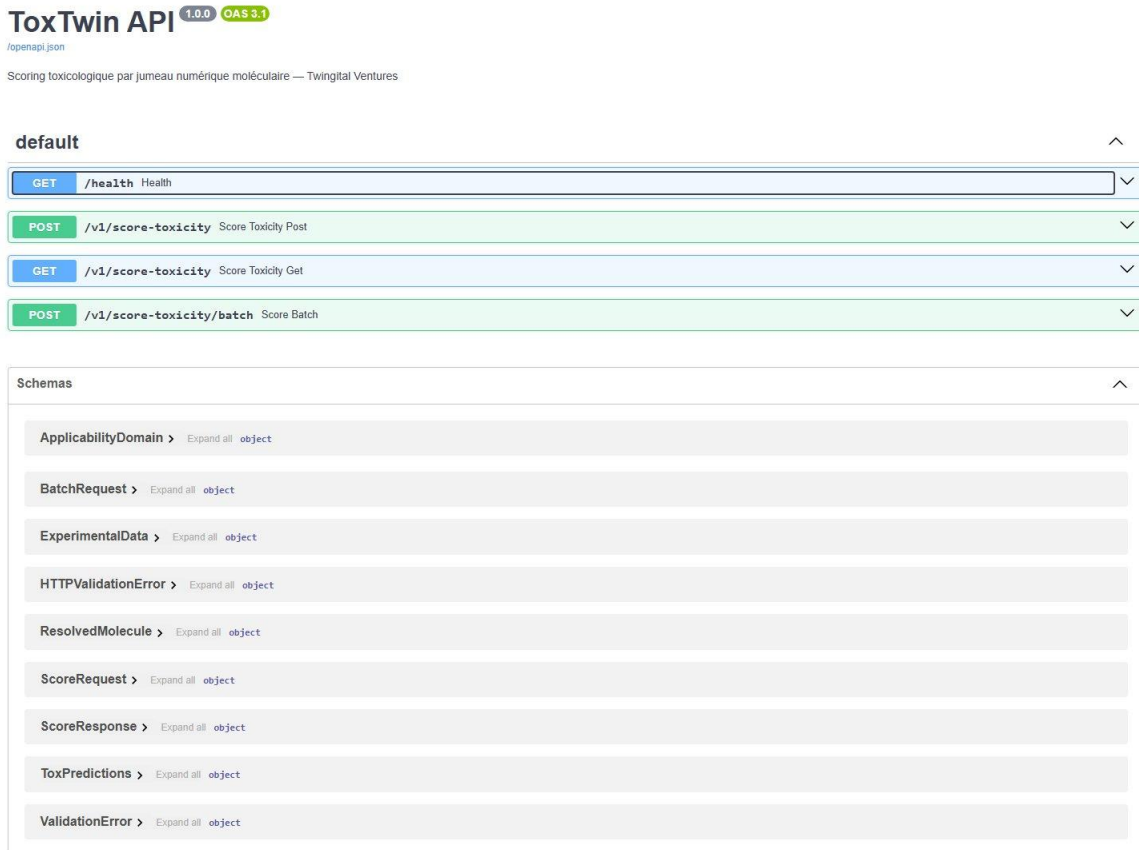


Figure 1 — Interface Swagger UI ToxTwin API v1.0 (OAS 3.1)

1.2 Endpoints disponibles

Méthode	Endpoint	Description
GET	/health	Vérification de l'état de l'API · modèle chargé · device · FP training set · calibration
POST	/v1/score-toxicity	Scoring toxicologique · accepte un JSON {query: string} · retourne ScoreResponse complet

Méthode	Endpoint	Description
GET	/v1/score-toxicity	Identique au POST mais via paramètre URL : ?query=aspirin
POST	/v1/score-toxicity/batch	Scoring batch jusqu'à 50 molécules en parallèle · retourne une liste de ScoreResponse

1.3 Schémas de données (Schemas)

L'interface Swagger liste les 8 schémas Pydantic qui structurent les échanges de données :

- ScoreRequest — payload d'entrée : {query: string}
- ScoreResponse — payload de sortie complet (voir §3)
- ResolvedMolecule — canonical_smiles · inchi_key · source · cid · mw
- ToxPredictions — tox21_scores · uncertainties · mean_uncertainty
- ExperimentalData — ld50_oral_rat · tox_score · organes · mutagénicité
- ApplicabilityDomain — max_tanimoto_similarity · is_in_domain · warning
- BatchRequest — liste de queries (max 50)
- HTTPValidationError / ValidationError — erreurs 422

1.4 Comment utiliser Swagger UI

- Ouvrir <http://localhost:8000/docs> dans le navigateur
- Cliquer sur POST /v1/score-toxicity pour déplier l'endpoint
- Cliquer sur le bouton Try it out (en haut à droite du bloc)
- Modifier le JSON dans le champ Request body : remplacer aspirin par le nom DCI ou SMILES souhaité
- Cliquer sur Execute — la réponse JSON complète s'affiche dans Response body
- Le code curl équivalent est affiché dans Curl command — copier-coller pour automatiser

2. Démarrage de l'API

2.1 Prérequis

- Environnement conda ml activé (Python 3.11, PyTorch 2.12, PyG 2.7)
- Modèle entraîné présent : /home/jeromev/medallion/gold/models/toxgnn_v1_finetune.pt
- Calibration présente : /home/jeromev/medallion/gold/models/calibration_temperatures.json
- Delta Lake Silver accessible : /home/jeromev/medallion/silver/

2.2 Commandes de démarrage

```
conda activate ml
cd /home/jeromev/toxtwin_pipeline
uvicorn notebooks.api_fastapi:app --host 0.0.0.0 --port 8000
```

Au démarrage, l'API charge séquentiellement :

- Le modèle ToxGNN-V1 (toxgnn_v1_finetune.pt) → ~2 secondes
- Les températures de calibration (calibration_temperatures.json) → instantané
- Les 2 000 Morgan FP du training set pour le domaine applicatif (via Spark) → ~8 secondes

Le message API prête sur cpu (ou cuda si GPU disponible) confirme le démarrage complet.

2.3 Test de santé

```
curl http://localhost:8000/health
```

Réponse attendue :

```
{"status":"ok","model_loaded":true,"device":"cpu","n_train_fps":2000,"calibration":true}
```

3. Tests de validation — Résultats observés

Deux molécules de référence pharmaceutique ont été testées pour valider le pipeline complet d'inférence. Les résultats sont présentés ci-dessous avec leur interprétation pharmacologique.

3.1 Test 1 — Acide acétylsalicylique (Aspirin)

Requête DCI · résolution PubChem · temps d'inférence 2 017 ms

```
curl -X POST http://localhost:8000/v1/score-toxicity -d '{"query": "aspirin"}'
```

Résolution moléculaire

Champ	Valeur
canonical_smiles	CC(=O)Oc1ccccc1C(=O)O
inchi_key	BSYNRYMUTXBXSQ-UHFFFAOYSA-N
source	pubchem (résolution DCI → PubChem PUG-REST)
cid	2244
mw	180.16 g/mol

Prédictions ToxGNN-V1 (12 endpoints Tox21)

Endpoint	Score	Incertitude	Interprétation
NR-ER	0.350	0.046	Signal le plus élevé · récepteur œstrogène · attendu pour un AINS
NR-AhR	0.191	0.036	Récepteur aryl hydrocarbure · cohérent avec la littérature
SR-ARE	0.168	0.021	Voie antioxydante Nrf2 · faible
NR-AR	0.115	0.032	Récepteur androgène · faible
SR-MMP	0.050	0.008	Stress mitochondrial · très faible
NR-Aromatase	0.036	0.006	Enzyme stéroïdienne · très faible
Autres endpoints (6)	< 0.10	< 0.023	Aucun signal toxicologique majeur prédit

Données expérimentales (Silver lookup)

Champ	Valeur
ld50_oral_rat	200.0 mg/kg (valeur de référence correcte)
tox_score	4 / 5
overall_tox_severity	moderate
target_organs	["liver"]
mutagenicity_ames	unknown (non testé dans les données disponibles)

Domaine applicatif

Métrique	Valeur
max_tanimoto_similarity	0.5938 — dans le domaine (> seuil 0.3)
is_in_domain	true — l'aspirine est bien représentée dans le training set

3.2 Test 2 — Ibuprofène (SMILES direct)

Requête SMILES · résolution directe RDKit · temps d'inférence 1 581 ms

```
curl -X POST http://localhost:8000/v1/score-toxicity -d '{"query": "CC(C)Cc1ccc(cc1)C(C)C(=O)O"}'
```

Résolution moléculaire

Champ	Valeur
canonical_smiles	CC(C)Cc1ccc(C(C)C(=O)O)cc1 (canonialisé par RDKit)
inchi_key	HEFNNSXXWATRW-UHFFFAOYSA-N
source	smiles_input (résolution directe RDKit — pas d'appel PubChem)
Gain de temps	~430 ms vs résolution DCI (pas d'appel réseau)

Prédictions ToxGNN-V1 — Points notables

Endpoint	Score	Incertitude	Interprétation
NR-PPAR-gamma	0.383	0.091	Signal fort · activation PPAR- γ documentée (Lehmann et al. 1997)
NR-ER	0.318	0.057	Perturbation endocrinienne · signal modéré
NR-ER-LBD	0.264	0.072	Liaison récepteur œstrogène · cohérent avec la littérature AINS
SR-MMP	0.119	0.022	Stress mitochondrial · plus élevé que l'aspirine
SR-HSE	0.075	0.019	Choc thermique · faible
Autres endpoints (7)	< 0.12	< 0.025	Aucun signal majeur

Données expérimentales (Silver lookup)

Champ	Valeur
ld50_oral_rat	636.0 mg/kg (moins toxique que l'aspirine en aigu)
tox_score	2 / 5 (cohérent avec LD50 plus élevé)
overall_tox_severity	moderate
target_organs	["liver"]

3.3 Comparaison aspirin vs ibuprofène

Le tableau suivant compare les deux profils toxicologiques prédits par ToxGNN-V1. Les différences observées sont cohérentes avec la pharmacologie connue des deux AINS.

Endpoint / Métrique	Aspirin	Ibuprofen	Enseignement
NR-ER (œstrogène)	0.350	0.318	Signal AINS similaire — attendu
NR-PPAR-gamma	0.038	0.383	Ibuprofen active PPAR-γ — documenté
NR-ER-LBD	0.077	0.264	Perturbation endocrinienne ibuprofen > aspirin
SR-MMP (mitochondrie)	0.050	0.119	Stress mito ibuprofen > aspirin
LD50 oral rat	200 mg/kg	636 mg/kg	Aspirin plus toxique en aigu
tox_score	4 / 5	2 / 5	Cohérent avec LD50
Tanimoto (domaine)	0.594	0.600	Les deux dans le domaine applicatif
Temps inférence	2 017 ms	1 581 ms	SMILES direct 26% plus rapide que DCI

Observation clé : le signal NR-PPAR-gamma à 0.383 pour l'ibuprofène, absent chez l'aspirine (0.038), correspond à une relation structure-activité documentée dans la littérature (Lehmann et al., 1997 ; Nature Medicine). Le modèle ToxGNN-V1 a capturé cette différence pharmacologique sans l'avoir explicitement supervisée — elle émerge des représentations apprises par le GNN pré-entraîné.

4. Format de réponse JSON — ScoreResponse

La réponse complète de l'endpoint `/v1/score-toxicity` contient 8 champs de premier niveau, chacun correspondant à un schéma Pydantic documenté dans Swagger.

Champ	Type	Description
query	string	Query original soumis (DCI ou SMILES)
resolved	ResolvedMolecule	canonical_smiles · inchi_key · source · cid · mw
predictions	ToxPredictions	tox21_scores (12 endpoints) · uncertainties (MC Dropout) · mean_uncertainty
experimental_data	ExperimentalData	ld50_oral_rat · tox_score · severity · ames · target_organs (Silver lookup)
applicability_domain	ApplicabilityDomain	max_tanimoto_similarity · is_in_domain · warning si hors domaine
calibration_warning	boolean	true si ECE > 0.10 — probabilités à interpréter avec précaution
inference_time_ms	float	Temps total de traitement en millisecondes
model_version	string	Version du modèle ToxGNN-V1 utilisé

4.1 Interprétation des scores Tox21

Score	Niveau de risque	Action recommandée
< 0.10	Très faible	Signal négligeable — pas d'alerte
0.10 – 0.25	Faible	À noter — surveillance dans les études suivantes
0.25 – 0.50	Modéré	Alerte — investigation expérimentale recommandée
> 0.50	Élevé	Signal fort — études in vitro prioritaires sur cet endpoint

4.2 Interprétation de l'incertitude (uncertainties)

Std MC Dropout	Niveau	Signification
< 0.05	Faible	Prédiction fiable — composé dans le domaine d'application
0.05 – 0.15	Modéré	Incertain notable — interpréter avec les données expérimentales
> 0.15	Élevé	Haute incertitude — composé potentiellement hors domaine

4.3 Dette technique connue

Item	Priorité	Description
Calibration ECE	P1 — v1.1	ECE 0.149 (FAIL > 0.10) · Isotonic regression à implémenter · pos_weight BCE = cause racine

Item	Priorité	Description
Temps d'inférence	P2 — v1.1	1 500-2 000 ms sur CPU · GPU RTX 5080 réduira à < 200 ms
NR-AR sous RF	P2 — v1.1	AUC 0.659 vs RF 0.750 · données supplémentaires ChEMBL recommandées
Domaine applicatif	P3 — v1.2	38.2% hors domaine sur scaffold split · normal par design · k-NN mean sim à implémenter

Document confidentiel — Twingital Ventures · ToxTwin API v1.0 · Mars 2026