Drug normalization Summary

Interventions

Biological Agents (NCI -> ChEBI -> WD)

Small molecules (NCI -> inchikey -> ChEBI -> WD)

Therapeutic Procedure (includes surgical, resection, transplant) (NCI -> WD)

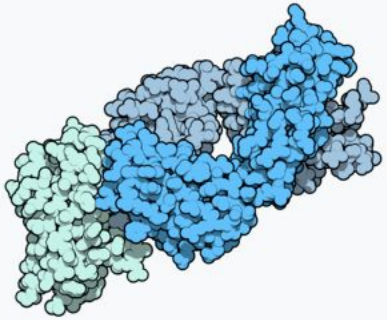
Chemical Class (a class of compounds) (NCI -> ChEBI -> WD)

Regimens (should these be entered as the regimen itself, or as its parts?) (NCI -> ChEBI -> WD)

Combination (ordered or unordered set of interventions)

Biological Agents

Ipilimumab



Fab fragment of ipilimumab (blue) binding CTLA-4 (green). From PDB entry 5TRU.

Monoclonal antibody	
Type	Whole antibody
Source	Human
Target	CTLA-4

- Drug, Food, Chemical or Biomedical Material
  - Pharmacologic Substance
    - Antineoplastic Agent
      - Antineoplastic Biological Agent
        - Antineoplastic Antibody
          - Ipilimumab
    - Biological Agent
      - Antineoplastic Biological Agent
        - Antineoplastic Antibody
          - Ipilimumab
      - Monoclonal Antibody
        - Anti-CTLA-4 Monoclonal Antibody
          - Ipilimumab
    - Immunotherapeutic Agent
      - Immune Checkpoint Modulator
        - Anti-CTLA-4 Monoclonal Antibody
          - Ipilimumab
      - Monoclonal Antibody
        - Anti-CTLA-4 Monoclonal Antibody
          - Ipilimumab

Other Examples:

siRNA

T-cell therapy

Transformed cells

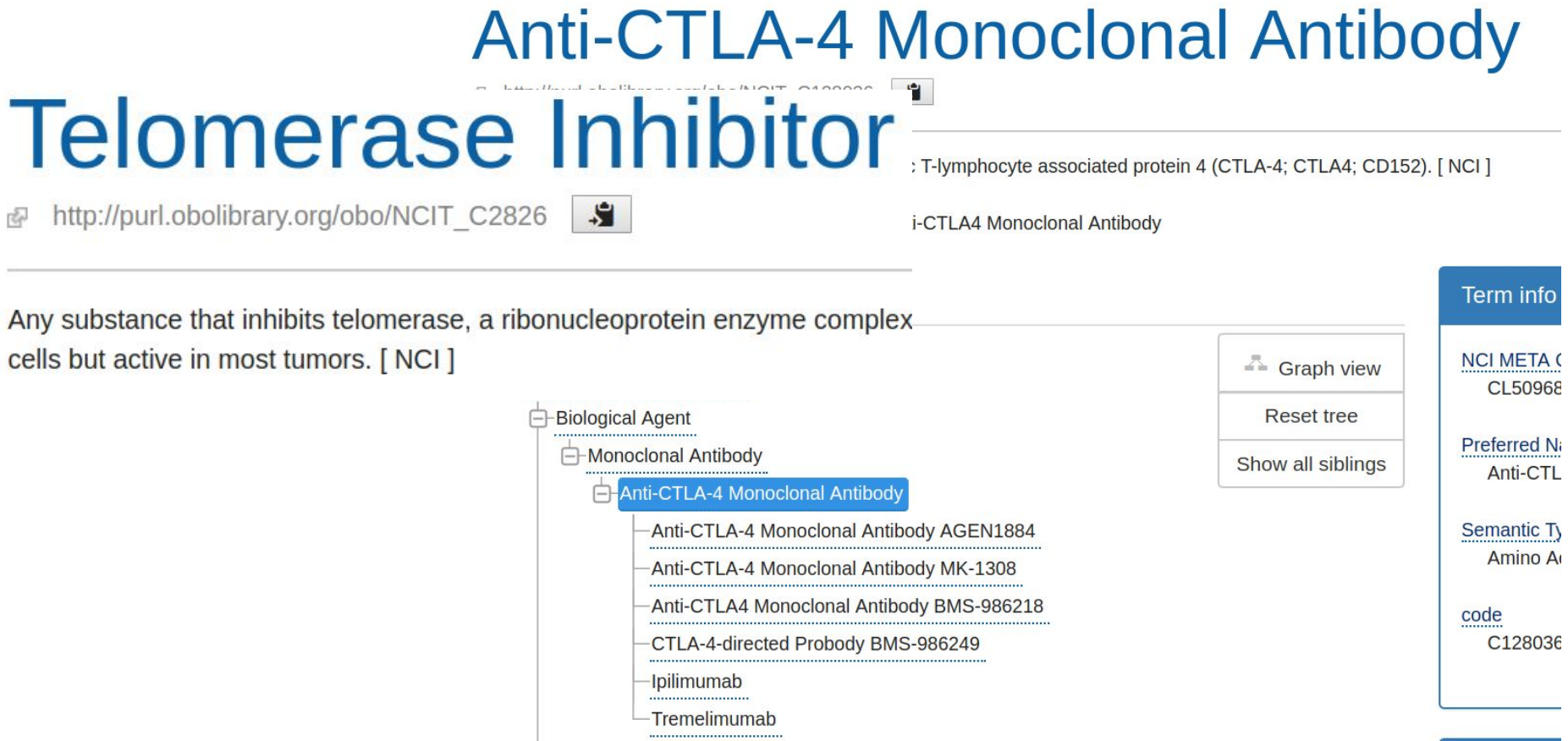
antibody

Small Molecules

Therapeutic Procedure

Radiation Therapy

Chemical class



Ipilimumab/Nivolumab Regimen

http://purl.obolibrary.org/obo/NCIT_C154272

A regimen consisting of ipilimumab and nivolumab that can be used for the treatment of kidney cancer and melanoma.

Synonyms: Ipilimumab Plus Nivolumab, Ipilimumab + Nivolumab, Ipilimumab-Nivolumab, Nivolumab/Ipilimumab, Ipili

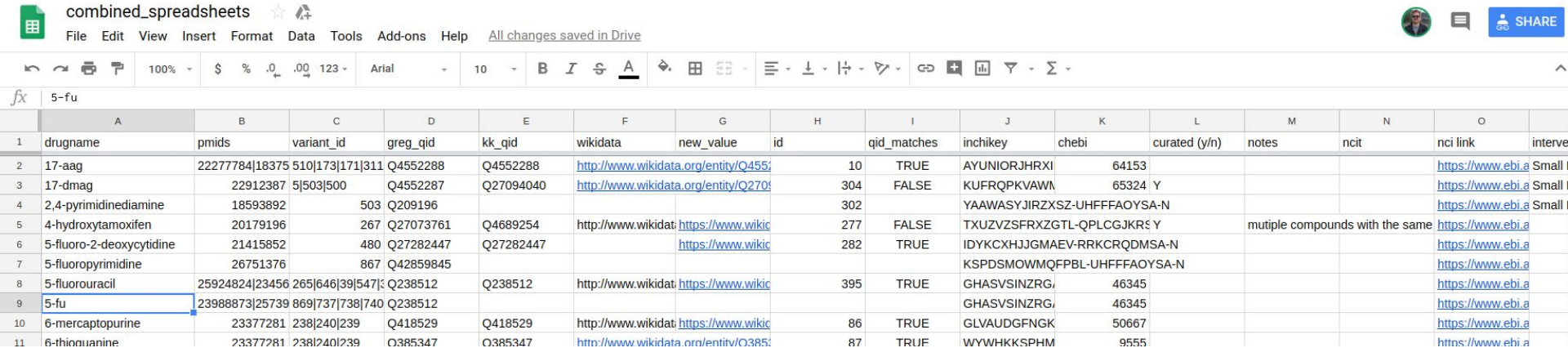
Term relations

Subclass of:

- Chemotherapy Regimen Used to Treat Malignant Renal Neoplasm
- Chemotherapy Regimen Used to Treat Melanoma
- *Chemotherapy_Regimen_Has_Component* some Nivolumab
- *Chemotherapy_Regimen_Has_Component* some Ipilimumab

Implementation Options in CIViC

- WikiData
- NCI Thesaurus
- Different ontologies for different types of interventions (e.g., INCHI Key for small molecules, Chebi for chemical classes)
- Priority list of IDs to use.
- NCIT -> Chebi -> INCHIKey -> .. ?
- Need to handle synonyms
- How to submit new terms to NCIt?



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	drugname	pmids	variant_id	greg_qid	kk_qid	wikidata	new_value	id	qid_matches	inchikey	chebi	curated (y/n)	notes	ncit	nci link	interve
2	17-aag	22277784|18375	510|173|171|311	Q4552288	Q4552288	http://www.wikidata.org/entity/Q455		10	TRUE	AYUNIORJHRXI	64153				https://www.ebi.a	Small
3	17-dmag	22912387	5|503|500	Q4552287	Q27094040	http://www.wikidata.org/entity/Q270		304	FALSE	KUFRQPKVAWN	65324	Y			https://www.ebi.a	Small
4	2,4-pyrimidinediamine	18593892	503	Q209196				302		YAAWASYJIRZXSZ-UHFFFAOYSA-N					https://www.ebi.a	Small
5	4-hydroxytamoxifen	20179196	267	Q27073761	Q4689254	http://www.wikidat	https://www.wikic	277	FALSE	TXUZVZSFRXZGTL-QPLCGJKRS		Y	mutiple compounds with the same		https://www.ebi.a	
6	5-fluoro-2-deoxycytidine	21415852	480	Q27282447	Q27282447		https://www.wikic	282	TRUE	IDYKCXHJJGMAEV-RRKCRQDMSA-N					https://www.ebi.a	
7	5-fluoropyrimidine	26751376	867	Q42859845						KSPDSMOWMQFPBL-UHFFFAOYSA-N					https://www.ebi.a	
8	5-fluorouracil	25924824|23456	265|646|39|547|3	Q238512	Q238512	http://www.wikidat	https://www.wikic	395	TRUE	GHASVSINZRG/	46345				https://www.ebi.a	
9	5-fu	23988873|25739	869|737|738|740	Q238512						GHASVSINZRG/	46345				https://www.ebi.a	
10	6-mercaptopurine	23377281	238|240|239	Q418529	Q418529	http://www.wikidat	https://www.wikic	86	TRUE	GLVAUDGFNGK	50667				https://www.ebi.a	
11	6-thioguanine	23377281	238|240|239	Q385347	Q385347	http://www.wikidata.org/entity/Q385		87	TRUE	WYWHKKSPHM	9555				https://www.ebi.a	

https://docs.google.com/spreadsheets/d/1FL_Mim29q5Osloazj7rbNFoD4sdga247GE__qN1b4yc/edit#gid=14227486

General Issues

Nonspecific terms / non enough information to determine identity of intervention (e.g. "adjuvant chemotherapy")

Drugs or Biological Agents that don't have NCI term (use inchikey or chebi)

Drugs that don't have a NCI, Inchikey or Chebi (use Wikidata)

Treatment Regimen (use NCI term for the regimen, use Wikidata?)

Salt forms, solvent, preparation, formulations (DMSO, mesylate, … ?)

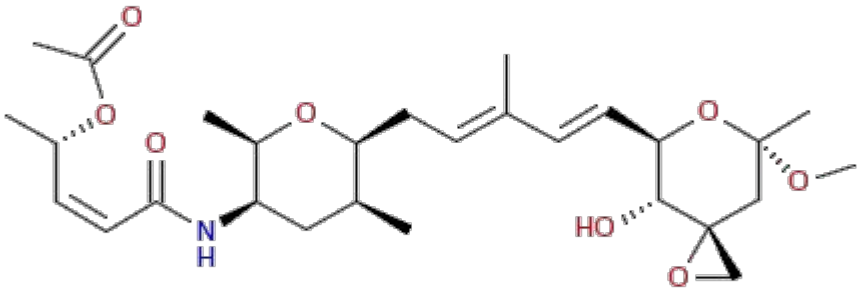
Incorrect data

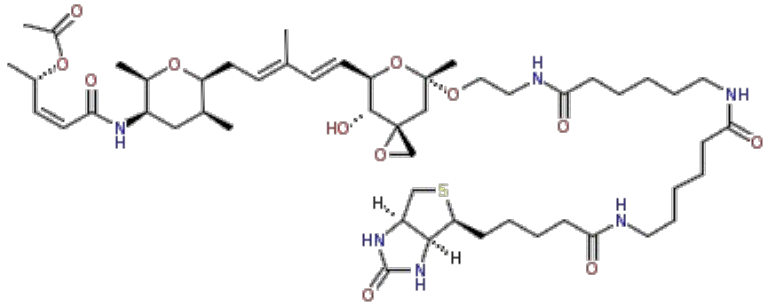
NCI doesn't have inchikeys for small molecules (but usually has ChEBI or UNII), so would have to go through CheBI

Spliceostatin A

Not in NCI, not in Chebi, two different inchikeys in pubchem and wikidata

Use INCHI Key?

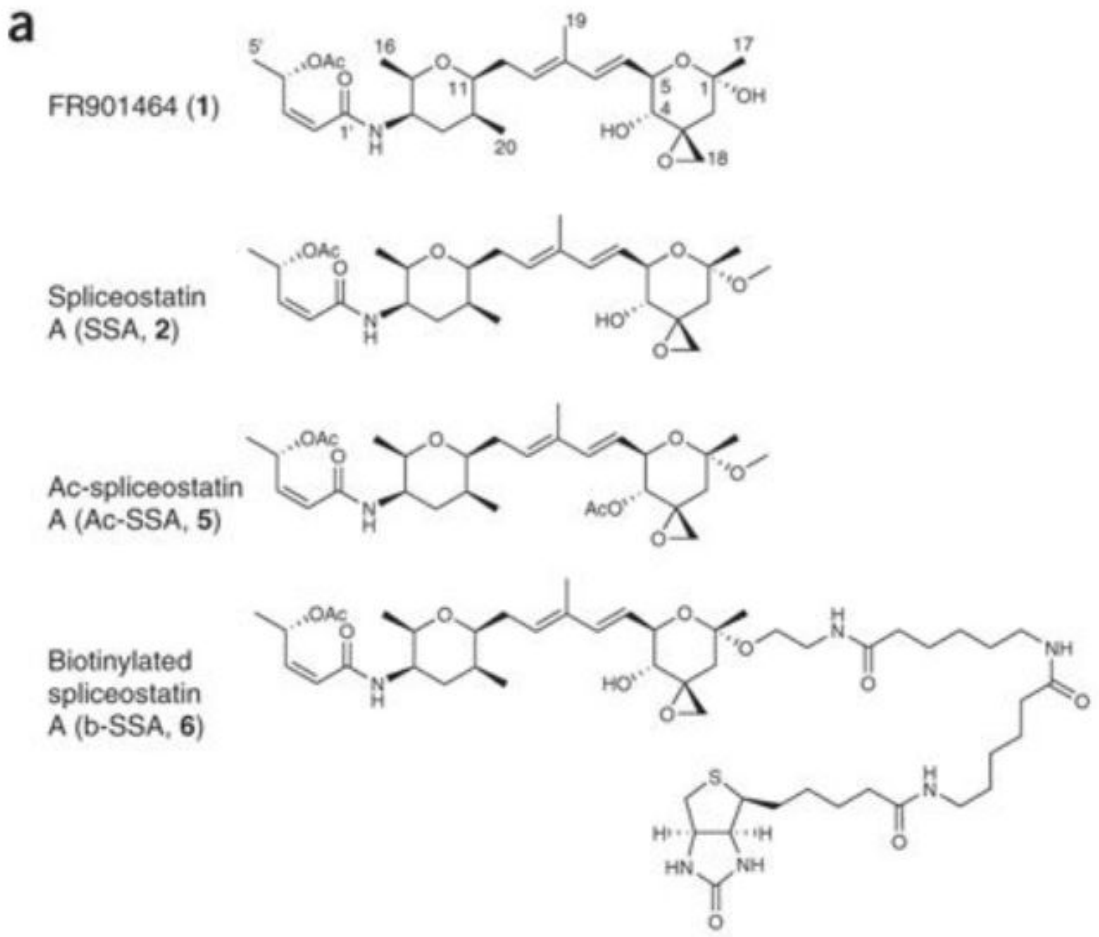




Google it

Result: Use inchikey

Figure 1: FR901464-induced accumulation of p27*.



Non specific terms

adjuvant chemotherapy

Reference actually says: "cisplatin-based adjuvant chemotherapy"

Combinations not in NCI

"brigatinib panitumumab cetuximab"

Each individually exists, but not a combination term.

Options:

- Can create a wikidata item to use as a placeholder, and add NCI term if gets created
- Create a combination intervention using each drug individually

Nonspecific

“VEGF inhibition”

Data extraction

Data extraction was done independently by two authors (BJK and JHK). If these two authors could not reach a consensus, another author was consulted to resolve the dispute. The following data were carefully extracted from all eligible studies: first author's name, year of publication, number of patients, histologic subtype of RCC, alteration status of VHL gene, VEGF-targeted agents used, ORR to anti-VEGF therapy, RR for ORR and PFS stratified by VHL alteration, and HR with 95% CI for PFS and OS.

Anti-VEGF Monoclonal Antibody NCIT:C2496

http://purl.obolibrary.org/obo/NCIT_C2496

A monoclonal antibody directed against human Vascular Endothelial Growth Factor (VEGF), a protein that stimulates angiogenesis, and prevents VEGF from binding to its receptors, causing inhibition of angiogenesis. (NCI)

Ontology: NCI Thesaurus OBO Edition NCIT

VEGFR Tyrosine Kinase Inhibitor NCIT:C93259

http://purl.obolibrary.org/obo/NCIT_C93259

Any substance that inhibits vascular endothelial growth factor receptor, a family of tyrosine kinase receptors involved in the formation of blood vessels.

Ontology: NCI Thesaurus OBO Edition NCIT

Solvents Salt Forms, Preparations

trametinib dmso vs trametinib ? Not explicitly linked together in NCI

Trametinib Dimethyl Sulfoxide

http://purl.obolibrary.org/obo/NCIT_C152711

Synonyms: Trametinib Dimethyl Sulfoxide, TRAMETINIB DIMETHYL S

Trametinib

http://purl.obolibrary.org/obo/NCIT_C77908

ChEBI Ontology

	trametinib dimethyl sulfoxide (CHEBI:75991) **has part** dimethyl sulfoxide (CHEBI:28262)
	trametinib dimethyl sulfoxide (CHEBI:75991) **has part** trametinib (CHEBI:75998)
Outgoing	trametinib dimethyl sulfoxide (CHEBI:75991) **has role** antineoplastic agent (CHEBI:35610)
	trametinib dimethyl sulfoxide (CHEBI:75991) **has role** EC 2.7.11.24 (mitogen-activated protein kina
	trametinib dimethyl sulfoxide (CHEBI:75991) **is a** addition compound (CHEBI:35504)

bitor of mitogen-activated prote

f growth factor-mediated cell si

Imatinib mesylate vs imatinib?

END

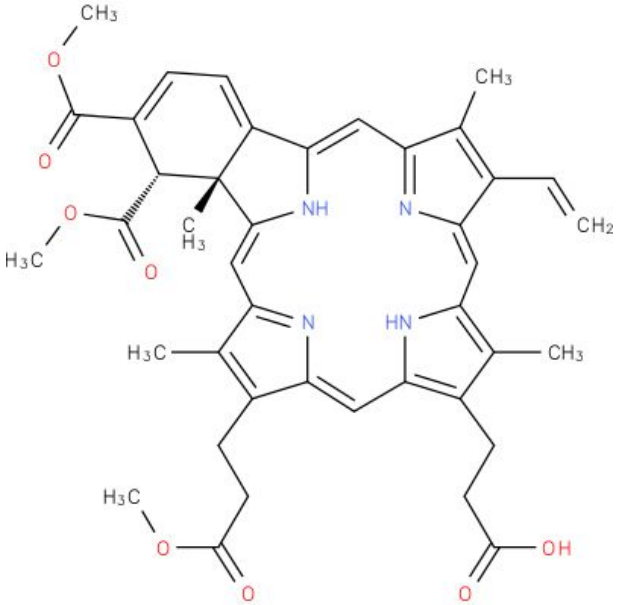
CHEBI:32293 - verteporfin

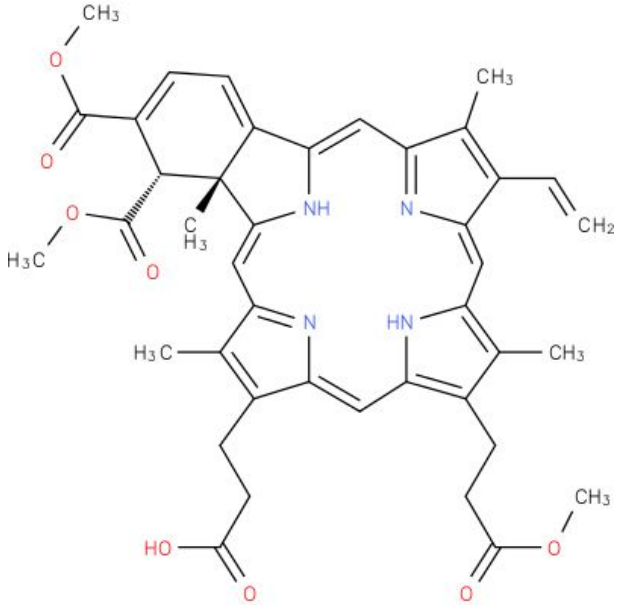
Main | ChEBI Ontology | Automatic Xrefs | Reactions | Pathways | Models

ChEBI Name	**verteporfin**
ChEBI ID	**CHEBI:32293**
	An equimolar mixture of the 9-methyl ester and 13-methyl ester of *trans*-(±)-18-ethenyl-4,4a-dihydro-3,4-bis(methoxycarbonyl)-4a,8,14,19-tetramethyl-23*H*,25*H*-benzo[*b*]porphine-9,13-dipropanoic acid. It is used as a photosensitizer in photodynamic therapy to eliminate the abnormal blood vessels in the eye associated with

Inseperable mixture.
No inchikey

Use chebi instead of inchikey, even though its a small molecule





OLS > NCI Thesaurus OBO Edition NCIT > NCIT:C1014

Verteporfin

http://purl.obolibrary.org/obo/NCIT_C1014

A synthetic light-activated agent with photodynamic activity. Upon s
presence of oxygen, produces highly reactive short-lived singlet oxy
(NCI04) [NCI]

Subsets
NCIT_C63923

Accepted Therapeutic Use
Photosensitizer, subfov

CAS Registry
129497-78-5

CHEBI ID
CHEBI:32293

Contributing Source
FDA

FDA UNII Code
0X9PA28K43

Antibody-Drug Conjugates

OLS > NCI Thesaurus OBO Edition NCIT > NCIT:C82492

Trastuzumab Emtansine

Search NCIT

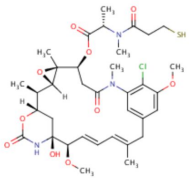
http://purl.obolibrary.org/obo/NCIT_C82492

An antibody-drug conjugate (ADC) consisting of the recombinant anti-epidermal growth factor receptor 2 (HER2) monoclonal antibody trastuzumab conjugated to the maytansinoid DM1 via a nonreducible thioether linkage (MCC) with potential antineoplastic activity. The trastuzumab moiety of this ADC binds to HER2 on tumor cell surfaces; upon internalization, the DM1 moiety is released and binds to tubulin, thereby disrupting microtubule assembly/disassembly dynamics and inhibiting cell division and the proliferation of cancer cells that overexpress HER2. Linkage of antibody and drug through a nonreducible linker has been reported to contribute to the improved efficacy and reduced toxicity of this ADC compared to similar ADCs constructed with reducible linkers. [NCI]

Synonyms: Kadcyla, Trastuzumab-MCC-DM1 Immunoconjugate, Ado Trastuzumab Emtansine, Trastuzumab-MCC-DM1 Antibody-Drug Conjugate, ADO-TRASTUZUMAB EMTANSINE, Trastuzumab Emtansine, Immunoglobulin G1, Anti-(Human p185neu Receptor) (Human-Mouse Monoclonal RhuMab HER2 Gamma1-Chain), Disulfide with Human-Mouse Monoclonal RhuMab HER2 Light Chain, Dimer, Tetraamide with N2'-(3-((1-((4-carboxycyclohexyl)methyl)-2,5-dioxo-3-pyrrolidinyl)thio)-1-oxopropyl)-N2'-deacetylMaytansine, Trastuzumab-DM1, RO5304020, T-DM1, PRO132365, Trastuzumab-MCC-DM1

CHEBI:82755 - mertansine

Main | ChEBI Ontology | Automatic Xrefs | Reactions | Pathways | Models



ChEBI Name	**mertansine**
ChEBI ID	**CHEBI:82755**
Definition	An organic heterotetracyclic compound and 19-membered macrocyclic lactam that is maytansine in which one of the hydrogens of the terminal *N*-acetyl group is replaced by a sulfanylmethyl group.
Stars	⭐⭐⭐ This entity has been manually annotated by the ChEBI Team.
Supplier Information	ZINC000095627837
Download	Molfile XML SDF

- Find compounds which contain this structure
- Find compounds which resemble this structure
- Take structure to the Advanced Search

InChIKey ANZJBCHSOXCCRQ-FKUXLPTCSA-N

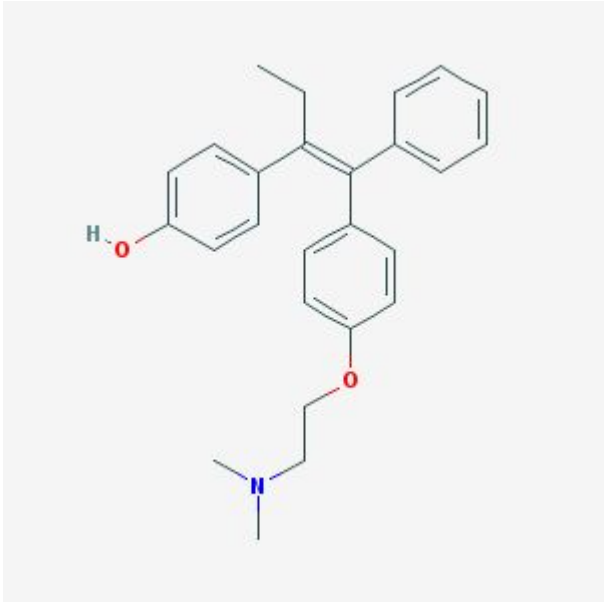
Trastuzumab NCIT:C1647

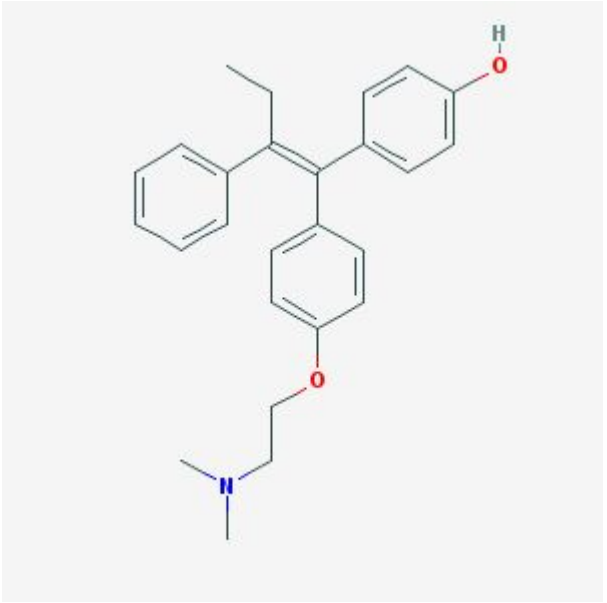
http://purl.obolibrary.org/obo/NCIT_C1647

A recombinant humanized monoclonal antibody directed against the human epidermal growth factor receptor 2 (HER2). After binding to HER2 on the tumor cell surface, trastuzumab induces an antibody-dependent cell-mediated cytotoxicity against tumor cells that overexpress HER2. HER2 is overexpressed by many adenocarcinomas, particularly breast adenocarcinomas. (NCI04)...

Ontology: NCI Thesaurus OBO Edition NCIT

4-hydroxytamoxifen



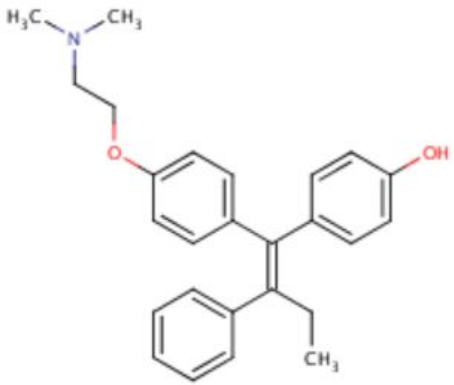


https://pubchem.ncbi.nlm.nih.gov/compound/449459

https://pubchem.ncbi.nlm.nih.gov/compound/5284643

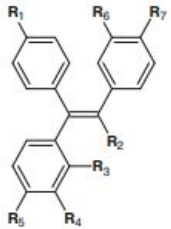
CHEBI:44616 - afimoxifene

Main | ChEBI Ontology | Automatic Xrefs | Reactions | Pathways | Models



ChEBI Name	**afimoxifene**
ChEBI ID	**CHEBI:44616**
Definition	A tertiary amino compound that is tamoxifen in which the phenyl group wh hydroxylated at the *para*- position. It is the active metabolite of tamoxifen.
Stars	★★★ This entity has been manually annotated by the ChEBI Team.
Supplier Information	eMolecules:539295, MolPort-003-848-078, ZINC000000902197
Download	Molfile XML SDF

Table 1 Trivial names, chemical structures and retention times of tamoxifen and metabolites with molecular mass 371.5, 357.5, 373.5 or 387.5 (reprinted and adjusted from Teunissen et al. [5], used with permission)



Trivial name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	Formula	Mol. Mass	Transition *(m/z)*	RT_1 (min)	RT_2 (min)
Tamoxifen	$O–CH_2–CH_2–N(CH_3)_2$	$CH_2–CH_3$	H	H	H	H	H	$C_{26}H_{29}NO$	371.5	372/72	8.00	2.99
N-desmethyltamoxifen	$O–CH_2–CH_2–NH–CH_3$	$CH_2–CH_3$	H	H	H	H	H	$C_{25}H_{27}NO$	357.5	358/58	7.91	2.96
N-desmethyl-α-hydroxytamoxifen[a]	$O–CH_2–CH_2–NH–CH_3$	$CH(OH)–CH_3$	H	H	H	H	H	$C_{25}H_{27}NO_2$	373.5	374/58	6.65	2.81
N-desmethyl-4-hydroxytamoxifen (Endoxifen)	$O–CH_2–CH_2–NH–CH_3$	$CH_2–CH_3$	H	H	OH	H	H	$C_{25}H_{27}NO_2$	373.5	374/58	5.79	2.81
N-desmethyl-3-hydroxytamoxifen[a]	$O–CH_2–CH_2–NH–CH_3$	$CH_2–CH_3$	H	OH	H	H	H	$C_{25}H_{27}NO_2$	373.5	374/58	5.85	2.81
N-desmethyl-4′-hydroxytamoxifen	$O–CH_2–CH_2–NH–CH_3$	$CH_2–CH_3$	H	H	H	H	OH	$C_{25}H_{27}NO_2$	373.5	374/58	6.41	2.81
α-Hydroxytamoxifen[a]	$O–CH_2–CH_2–N(CH_3)_2$	$CH(OH)–CH_3$	H	H	H	H	H	$C_{26}H_{29}NO_2$	387.5	388/72	3.91	2.84
4-Hydroxytamoxifen	$O–CH_2–CH_2–N(CH_3)_2$	$CH_2–CH_3$	H	H	OH	H	H	$C_{26}H_{29}NO_2$	387.5	388/72	6.03	2.84