ANTHROP\C

Blind tests of Claude on unpublished total syntheses

Published

November 9, 2026

Author

David Kamber

Overview **3**
The test 3
What the professors found 3
The prompt given to each model 4
GDGT-4, a macrocyclic lipid target from the Burns laboratory **6**
The target and the test 6
What the models proposed 7
Professor Burns's assessment 8
Limitations and next steps 9
Two targets from the Newhouse laboratory **10**
The targets and the test 10
What the models proposed 10
Professor Newhouse's assessment 12
Limitations and next steps 12
An unpublished target from the Reisman laboratory **13**
Why an unpublished target 13
How the test was run 13
What the models proposed 13
Professor Reisman's assessment 15
Limitations and next steps 17

Overview

Professors Noah Burns of Stanford, Timothy Newhouse of Yale and Sarah Reisman of Caltech shared four targets from total syntheses that their laboratories have not yet published. Three Claude models planned a route to each target from the same prompt, with one attempt each and no access to the literature, and each professor reviewed the routes without knowing which model had written which route.

The professors saw the routes labeled only Model A, Model B and Model C, which we have kept in their quotes. Model A is an earlier pre-release version of Claude Opus 5.5, Model B is an earlier pre-release version of Claude Mythos 5.2 and Model C is Claude Mythos 5.1.

The test

Each target here comes from a synthesis that the professor's group has completed or is pursuing and has not yet published, so the professor can compare every proposal with chemistry the group has tried at the bench. For each target, every model received one image of the structure and the prompt reproduced below, which was the same in every run. The models had no internet, literature or reaction-database access, and their only tool was local code such as RDKit. Each model had one attempt, and the runs took 26 to 84 minutes. Nothing was edited, selected or corrected, and the professors received the route documents labeled only A, B and C. The letters stand for the same three models in every section.

Professor	Target	The group's own synthesis	Length of the three routes
Noah Burns, Stanford University	GDGT-4, a membrane lipid of archaea with a 68-membered macrocycle and 22 stereocenters	Completed, not yet published	29 to 35 steps
Timothy Newhouse, Yale University	Triconnarin D, a natural product with six stereocenters, three of them all-carbon quaternary centers	Completed, not yet published	27 to 31 steps
	A bicyclic intermediate with five contiguous stereocenters	In progress	10 to 15 steps
Sarah Reisman, California Institute of Technology	A highly oxygenated tetracyclic compound with ten stereocenters (structure withheld at her request)	Nearly completed, not yet published	26 to 29 steps

Summary of the three tests. Route length is the total number of steps in each of the three routes.

What the professors found

Each of the twelve runs returned a complete route. Professor Burns called the routes "the kinds of initial retros I expect from early-year grad students," Professor Newhouse wrote that they "appear to be at the level of a first-year graduate student in synthetic chemistry,"

and Professor Reisman wrote that the model “seems like it could be helpful as a tool for the types of total synthesis efforts we pursue.” In the Burns and Reisman tests, parts of several routes matched chemistry that the group had used or investigated itself, and some of the steps that the professor doubted were among those the models had listed as most likely to fail.

The tests also showed clear limits. In the Newhouse and Reisman tests the models stayed with well-precedented reactions, so those tests do not show whether they can find a less obvious disconnection that shortens a route, and the routes to GDGT-4 are long, at 29 to 35 steps. One of the twelve routes was written for the wrong diastereomer, because the model had read two hashed bonds in the drawing as wedges. None of the routes has been run in a laboratory, so all twelve remain proposals.

The prompt given to each model

Every model received the following prompt together with the image of the target.

> The target is the compound in the attached image.
>
> Propose a complete total synthesis of this compound, as the single enantiomer drawn, from commercially available starting materials. Commit to one route.
>
> Start with a retrosynthetic overview of five lines or fewer: the key disconnections and the order in which the rings are formed.
>
> Number every compound (1, 2, 3, ...) in the order it first appears, starting materials included. Name or describe each one clearly enough that a chemist could draw it, stereochemistry included. Write the route as numbered steps in the forward direction, each in this form:
>
> Step N
>
> Reactants (the compounds drawn as structures in a scheme): number and name of each
>
> Product: number and name
>
> Conditions: reagents, solvent and temperature, written the short way they appear over an arrow in a paper
>
> Why: one sentence on why it should work, including how any new stereocenter is set
>
> Before you write the final answer, check your own route and fix what you find:
>
> - Re-read the image and confirm the structure you planned for matches it: ring sizes, every substituent, and the wedge or dash at every stereocenter.
>
> - Walk through the route step by step and confirm each product follows from its reactants with every atom accounted for, and that the last product is the target. Use RDKit or other local code tools where they help, for example to confirm each

intermediate's molecular formula.

- For each step, check that the other functional groups and protecting groups in the molecule survive the conditions.

- Check that every stereocenter in the target is set somewhere in the route and is not scrambled later.

Show only the corrected route, then add one line saying what you changed after checking.

End with the total step count, the longest linear sequence, and the three steps most likely to fail, each with one fallback.

Give the final answer in the chat and also as a PDF file named target_route.pdf. In the PDF, first draw the route as a reaction scheme: use RDKit (or another local drawing tool) to draw every numbered compound with its stereochemistry shown as wedges and dashes, with the compound number under each structure and the conditions over each arrow. Draw the structures large enough to read easily, a few steps per row. After the scheme, include the retrosynthetic overview, the step-by-step text and the closing summary. Check that every drawn structure matches its name in the text. Do not put the name of any AI model or company anywhere in the PDF.

Do not use web search, literature connectors or any internet access. Local code tools such as RDKit are fine.

Do not ask me any questions. If something is ambiguous, make the most reasonable assumption, state it in one line, and continue.

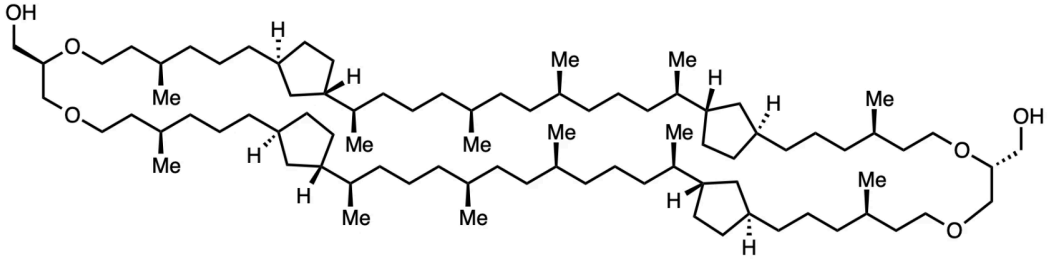
GDGT-4, a macrocyclic lipid target from the Burns laboratory

Professor Noah Burns of Stanford shared the structure of GDGT-4, a target his group has synthesized and not yet published. Three Claude models each planned a route to it, with one attempt and no access to the literature, and he compared the routes with what his group actually does.

> "*Overall I must say I'm impressed! I did not see anything totally nonsensical. They are all the kinds of initial retros I expect from early-year grad students.*"
>
> Noah Burns, Associate Professor of Chemistry, Stanford University

The target and the test

GDGT-4 (Figure 1), a glycerol dialkyl glycerol tetraether, is a structurally unusual membrane lipid of archaea. Two C_{40} chains, each bearing two trans-1,3-disubstituted cyclopentane rings and six methyl-bearing stereocenters, are joined at both ends to glycerol units through four ether linkages. The result is a 68-membered macrocycle with 22 stereocenters and no functional groups beyond the ethers and two primary alcohols. The Burns group has completed a synthesis that is not yet published. Each model received the drawing, with its SMILES string printed underneath, and the prompt reproduced above. The models had no internet, literature or reaction-database access, and their only tool was local code such as RDKit. The runs took 26 to 51 minutes. Nothing was edited or selected, and Professor Burns received the documents labeled only A, B and C. All three models arrived at the same structure, with the same configuration at all 22 stereocenters.



[H][C@@]12CCC[C@H](CCO[C@H](CO)COCC[C@H](C)CCC[C@@](C3)([H])CC[C@]3([H])[C@H](C)CCC[C@H](C)CC[C@@H](C)CCC[C@@H](C)[C@](CC4)([H])C[C@@]4([H])CCC[C@@H](C)CCOC[C@@H](CO)OCC[C@H](C)CCC[C@@](C5)([H])CC[C@]5([H])[C@H](C)CCC[C@H](C)CC[C@@H](C)CCC[C@@H](C)[C@](CC2)([H])C1)C

Figure 1. The structure as Professor Burns drew it, in the image the models received.

What the models proposed

The routes have 35, 29 and 31 steps, with longest linear sequences of 24, 24 and 25. All three used the twofold symmetry the same way, building one half-chain with a single cyclopentane and five stereocenters and using it four times. All three also bought the ring, as an enantiopure commercial monoacetate of cis-cyclopent-4-ene-1,3-diol, and converted its cis arrangement into the trans ring of the target by substitutions of known stereochemical course (Table 1): two palladium-catalyzed allylic alkylations around a Mitsunobu inversion (Opus 5.5); a copper-catalyzed Grignard substitution of the allylic acetate, a Mitsunobu inversion and a malonate displacement of the ring mesylate (Mythos 5.2); or a palladium-catalyzed allylic alkylation with a nitroalkane, removal of the nitro group with tributyltin hydride and a cyanide displacement (Mythos 5.1).

They differ in how the C_{40} chains and the macrocycle are made. Opus 5.5 puts one half-chain on each glycerol first, joins those two sn-2 chains by a metathesis dimerization, then adds the sn-3 pair and closes the ring by ring-closing metathesis. Mythos 5.2 runs the same dimerization on the free half-chain alcohol (Figure 2), hangs the resulting C_{40} chain between two glycerols and adds two more copies of the half-chain to set up the ring-closing metathesis. Mythos 5.1 couples an iodide with its own Grignard reagent to make the C_{40} diol and closes the ring by a double Williamson etherification under slow addition. All three listed the ring closure among their three most likely failures. The self-checks requested in the prompt changed details of the routes. Mythos 5.2, for example, moved every alkene reduction next to a stereocenter from palladium on carbon to diimide or platinum oxide, so that double-bond migration could not scramble the center. “This goes beyond identifying a plausible chemical reaction,” Professor Burns wrote. “It really shows awareness of how certain reduction conditions could negatively affect nearby stereochemistry.”

	Opus 5.5	Mythos 5.2	Mythos 5.1
Ring: from the cis monoacetate to trans	Steps 1 to 3: dimethyl malonate (1.5 equiv), NaH (1.3 equiv), $Pd(PPh_3)_4$ (2 mol %), THF, 25 to 50 °C; 4-$O_2NC_6H_4CO_2H$, PPh_3, DIAD, THF, 0 °C; dibenzyl malonate (3 equiv), NaH (2.5 equiv), $Pd(PPh_3)_4$ (5 mol %), THF, 50 °C	Step 4: side-chain RMgCl (3 equiv), CuCN (30 mol %), THF, -18 °C. Step 6: 4-$O_2NC_6H_4CO_2H$, PPh_3, DIAD, THF, 0 °C to rt. Step 9: dimethyl malonate (3 equiv), NaH, THF/DMF 3:1, 70 °C, on the ring mesylate	Step 6: nitroalkane (2.5 equiv), $Pd_2(dba)_3{\cdot}CHCl_3$ (2.5 mol %), dppf (10 mol %), DBU (2.5 equiv), THF, 25 °C. Step 11: NaCN (5 equiv), DMSO, 80 °C, on the ring mesylate
Methyl center C11	Step 21: pseudoephedrine amide, LDA (2.1 equiv), LiCl (6 equiv), THF, -78 to 0 °C; then the alkyl iodide (1.5 equiv), 0 to 25 °C (Myers)	Step 13: oxazolidinone imide, NaHMDS, THF, -78 °C; then MeI, -78 to -40 °C (Evans)	Step 14: enethioate, MeMgBr (1.5 equiv), $CuBr{\cdot}SMe_2$ (5 mol %), (R,S)-Josiphos (6 mol %), t-BuOMe, -78 °C

	Opus 5.5	Mythos 5.2	Mythos 5.1
Copper-catalyzed alkyl-alkyl coupling	Step 12 (C4 to C5): RMgBr, then the tosylate, Li_2CuCl_4 (3 mol %), NMP (4 equiv), THF, -10 to 20 °C	Step 20 (C12 to C13): RMgBr (2.5 equiv), then the tosylate, Li_2CuCl_4 (5 mol %), THF, -20 °C to rt	Step 24 (joins two C_{20} halves): RMgI from half of the iodide, then the other half, Li_2CuCl_4 (5 mol %), THF, -20 to 25 °C
Closure of the macrocycle	Step 33: Hoveyda-Grubbs II (10 mol %, portionwise), 1,4-benzoquinone (20 mol %), toluene, 0.5 mM, 80 °C	Step 28: Hoveyda-Grubbs II (2 × 10 mol %), 1,4-benzoquinone (20 mol %), toluene, 1 mM, 80 °C	Step 30: NaH (10 equiv), THF/DMF 1:1, 60 °C; diol and C_{40} bis-triflate (1:1) by syringe pump over 12 h, final 1 mM
Fallback named for that closure	Double Williamson etherification of the diol dianion with the C_{40} bis-triflate at high dilution	$OsO_4/NaIO_4$ to the dialdehyde, then intramolecular McMurry coupling	The bis-iodide with KH/18-crown-6 in THF at 60 °C, or isolate the mono-ether and cyclize it separately

Table 1. Conditions for key steps, as written in the three documents.

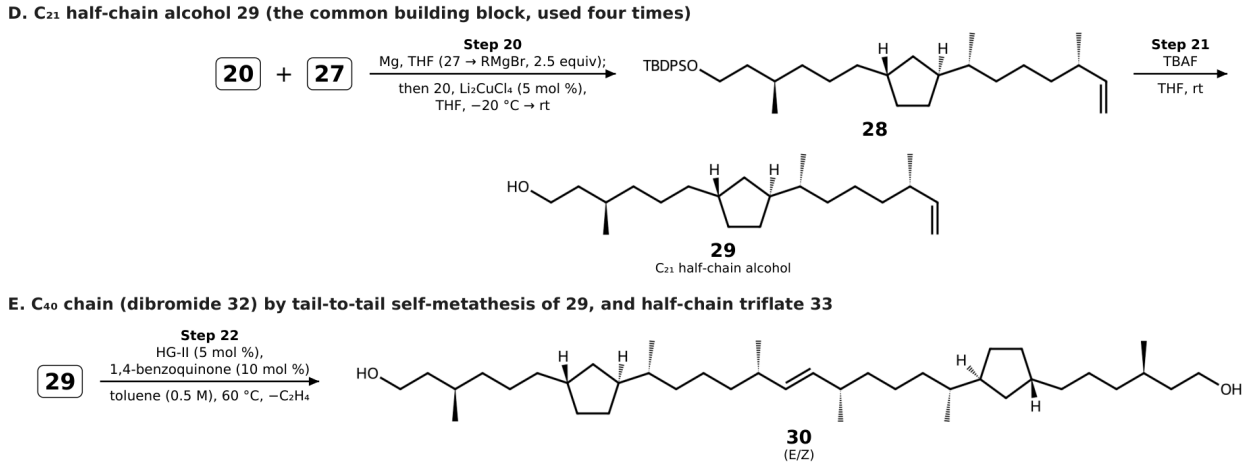


Figure 2. Steps 20 to 22 of Mythos 5.2's route, cropped from its scheme exactly as the model drew it. Compound 20 is the tosylate that carries the ring and the C1 to C6 side chain, 27 is (S)-5-bromo-3-methylpent-1-ene, and alcohol 29 is the half-chain that the route uses four times.

Professor Burns's assessment

Big picture. The routes are "steppy, for sure, but avoiding that has been one of the biggest challenges in these molecules," and "none of the models blow the others out of the water, so 'best' is a tough call."

Ranking. For now he leans toward Mythos 5.2, "from a strategic perspective," for its "nice approach for making 29 then using it as a dimer then as two monomers to set up a ring-closing metathesis," although the route to tosylate 20 "definitely has some issues." Its Evans auxiliaries "could work, but not my first choice," and using one to make the small bromide 27 "is a bit overkill." Opus 5.5 has "the most attractive strategy for putting together the cyclopentane of all the models." Its allylic alkylations "have potential issues, but there is

nothing fundamentally wrong with them," and its "Myers alkylations are definitely a better strategy than Evans in Model B." Mythos 5.1 is his least favorite "because of use of multiple Wittig-like olefinations for couplings," a Horner-Wadsworth-Emmons reaction in step 13 and a Wittig reaction in step 20.

Against his own route. Mythos 5.1 "has the most strategic similarity to what we actually do," because its C_{40} diol "is an exact intermediate we use and couple with a protected glycerol." Mythos 5.2's synthesis of the C1 to C6 side chain from citronellol is "nearly identical to what we used at one stage," and the copper-catalyzed Grignard couplings of Opus 5.5 and Mythos 5.2 (steps 12 and 20, Table 1) are "very similar to something we used at several points."

Steps he questioned. Mythos 5.1's macrocyclic Williamson etherification "failed in our hands for a similar system, but it was an early thing on our list to try." Opus 5.5 names the same closure as its fallback. Of Opus 5.5's ring-closing metathesis he noted that "RCM has been done on a related macrocycle." Mythos 5.2's malonate displacement of the ring mesylate (step 9) is "unlikely to work (model correctly flagged this)," and the fallback the model named is a palladium-catalyzed allylic alkylation with malonate, the reaction Opus 5.5 used. Mythos 5.2's copper-catalyzed step 4 "wastes material 4," the side-chain chloride used at 3 equiv. "There are ways to avoid it, but this reaction will likely give regioisomeric mixtures if it works." Of Mythos 5.2's Mitsunobu inversion he wrote that "Mitsunobu reactions look good on paper, but we always tend to avoid them." Opus 5.5's route has one as well. The monoacetate is commercial but "too expensive a starting material (>$100/g)," a point he made for Opus 5.5 and Mythos 5.2, and Mythos 5.1 starts from it as well.

Surprises. "It struck me that none of the models used asymmetric hydrogenation to set stereochemistry." He called Mythos 5.1's allylic alkylation with a nitroalkane "interesting": "I know similar things have been done with nitromethane. We always rather avoid having to use functional groups that then need to be removed, but it is still interesting."

Limitations and next steps

None of these routes has been run in the laboratory. At 29 to 35 steps they are long, and none sets a stereocenter by asymmetric hydrogenation. The models could not look up what a starting material costs or know that one of the proposed ring closures had already failed on a similar system in the Burns laboratory, and both points came only from the expert. The same three models will next be given a problem of his choosing from his group.

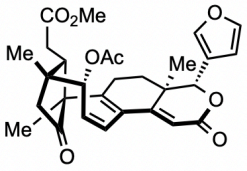
Two targets from the Newhouse laboratory

Professor Timothy Newhouse of Yale shared two structures from an unpublished natural product synthesis, the natural product itself and a bicyclic intermediate that his group is working on now. Three Claude models each planned a route to both, with one attempt per target and no access to the literature.

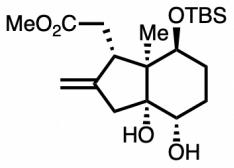
> *"I'm incredibly impressed by the synthetic planning capability of Claude. These models are already delivering synthetic routes that appear to be at the level of a first-year graduate student in synthetic chemistry."*
>
> Timothy Newhouse, Professor of Chemistry, Yale University

The targets and the test

The natural product, triconnarin D (Figure 1, left), has been made in the Newhouse laboratory, and that synthesis is not yet published. It has six stereocenters, three of them all-carbon quaternary centers, and a cyclopentanone bridged across a nine-membered ring that contains a Z alkene. The intermediate (Figure 1, right) is a hydrindane with five contiguous stereocenters. The group is actively pursuing a synthesis of this intermediate within the context of the natural product. Each model received the drawing and the prompt reproduced above, with no internet, literature or reaction-database access, and its only tool was local code such as RDKit. The runs took 30 to 84 minutes. Nothing was edited or selected, and Professor Newhouse received the documents labeled only A, B and C.



methyl 2-((4*R*,4a*R*,7*S*,10*R*,14*S*,*Z*)-11-acetoxy-4-(furan-3-yl)-4a,7,10-trimethyl-2,8-dioxo-2,4,4a,5,6,7,8,9,10,11-decahydro-7,10-methanocyclonona[*f*]isochromen-14-yl)acetate
Chemical Formula: $C_{29}H_{32}O_8$



methyl 2-((1*S*,3a*R*,4*S*,7*S*,7a*S*)-7-((*tert*-butyldimethylsilyl)oxy)-3a,4-dihydroxy-7a-methyl-2-methyleneoctahydro-1*H*-inden-1-yl)acetate
Chemical Formula: $C_{20}H_{36}O_5Si$

Figure 1. The two drawings as Professor Newhouse sent them and as the models received them.

What the models proposed

For the intermediate the three routes have 10, 14 and 15 linear steps and differ in strategy. Opus 5.5 starts from 2-methylcyclohexane-1,3-dione, sets the quaternary carbon and the neighboring carbinol in one step with a baker's yeast reduction, and closes the five-membered ring last with an intramolecular Hosomi-Sakurai reaction that forms the

tertiary alcohol and the exocyclic methylene together (Figure 2). Mythos 5.2 buys the five-membered ring as (R)-4-hydroxycyclopent-2-en-1-one, builds the quaternary carbon with a samarium-mediated Reformatsky-type aldol reaction, and closes the six-membered ring with a samarium-mediated ketyl-aldehyde cyclization that gives the cis diol directly. Mythos 5.1 read two hashed bonds in the drawing as wedges and stated the resulting configuration in its one-line assumption. It wrote a complete 15-step route to that diastereomer, with both ring-fusion centers inverted, which Professor Newhouse spotted at once. The other five of the six routes in this test were written for the structure as drawn.

For the natural product the routes have 31, 27 and 29 steps, with longest linear sequences of 27, 18 and 20. All three join a furan-bearing fragment to a cyclopentanone that carries two of the quaternary carbons, close the medium ring near the end and rate that closure as the step most likely to fail. Opus 5.5 and Mythos 5.1 chose the same closure, a Stork-Zhao olefination followed by an intramolecular Nozaki-Hiyama-Kishi reaction at 2 mM, and Mythos 5.2 chose ring-closing metathesis at 1 mM (Table 1). Mythos 5.2 and Mythos 5.1, independently, built the quaternary carbon C10 with the same copper-catalyzed enantioselective conjugate addition to 3-methylcyclopentenone.

	Opus 5.5	Mythos 5.2	Mythos 5.1
Intermediate: quaternary carbon and neighboring carbinol	Baker's yeast, sucrose, H_2O, 30 °C, on 2-allyl-2-methylcyclohexane-1,3-dione	SmI_2 (2.2 equiv), THF, -78 °C, α-bromoketone plus pent-4-enal	The same yeast reduction, then PhCHO, SmI_2 (15 mol %), THF, -10 °C (Evans-Tishchenko)
Intermediate: ring closed last	Five-membered: $EtAlCl_2$ (3 equiv), CH_2Cl_2, -78 to -20 °C (Hosomi-Sakurai)	Six-membered: SmI_2 (2.5 equiv), t-BuOH (2 equiv), THF, -78 to -20 °C	Five-membered: pyrrolidine, AcOH, benzene, reflux (aldol condensation)
Natural product: quaternary carbon C10	CH_2=CHMgBr, CuBr·SMe_2, TMSCl, HMPA, THF, -78 °C	Alkenylalane, $CuCl_2 \cdot 2H_2O$ (5 mol %), chiral NHC-Ag (2.5 mol %), THF, 22 °C	Alkenylalane, $CuCl_2 \cdot 2H_2O$ (2 mol %), chiral NHC-Ag (1 mol %), THF, -30 °C
Natural product: closure of the medium ring	$CrCl_2$ (10 equiv), $NiCl_2$ (1 mol %), DMSO/THF, 2 mM, rt	Grubbs II (10 mol %), toluene, 1 mM, 80 °C	$CrCl_2$ (10 equiv), $NiCl_2$ (5 mol %), DMF, 2 mM, rt
Fallback named for that closure	Ring-closing metathesis, Hoveyda-Grubbs II, 1 mM	Relay ring-closing metathesis	Ring-closing metathesis, or a samarium-mediated Barbier closure

Table 1. Conditions for key steps, as written in the six documents.

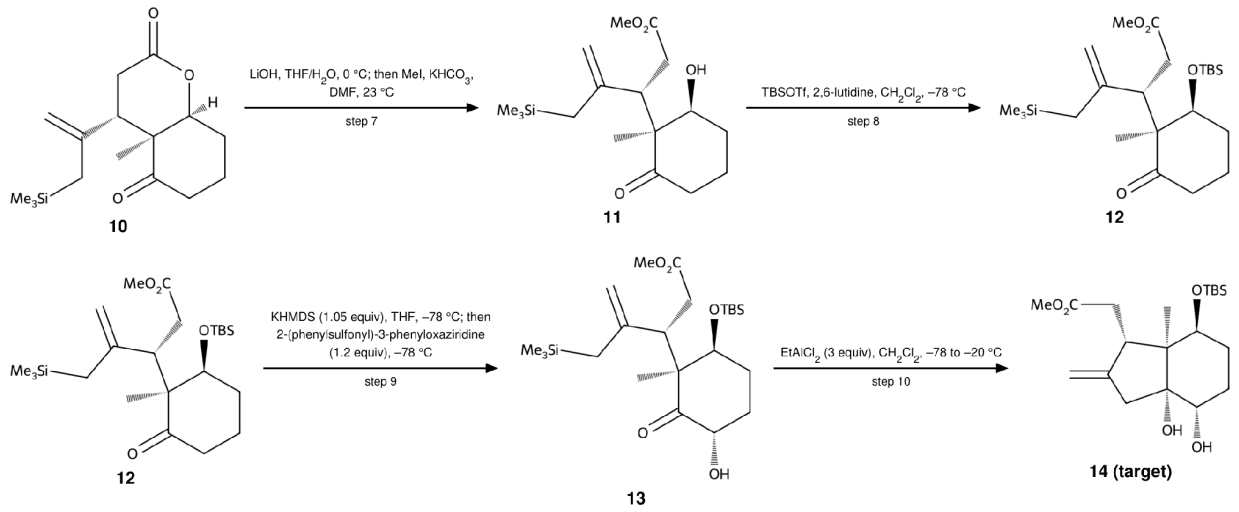


Figure 2. The last four steps of Opus 5.5's ten-step route to the intermediate, exactly as the model drew them.

The self-checks that the prompt asks for changed the routes. Opus 5.5, for example, replaced $TiCl_4$ with $EtAlCl_2$ in the ring closure so that the free hydroxyl group could not release HCl onto the allylsilane.

Professor Newhouse's assessment

He found that the strategies apply "robust chemical transformations that are well-known and established" and that the methods are "generally low risk and appear to be derived from precedented chemical reactions." He rated the capability of the model to select and apply transformations to be similar to the level of a first-year graduate student and added that "developing retrosynthetic tools for more complex secondary metabolites remains an outstanding challenge." One of the most important challenges in applying an AI-based approach to organic synthesis is the interpretation of line drawings. Mythos 5.1 made an error in how the structure was interpreted, underscoring that this remains an outstanding challenge in the area.

Limitations and next steps

These routes have not been attempted in a laboratory, but these models provide compelling insights into how a viable and actionable plan could be developed. The models stayed with well-precedented reactions, so this test does not show whether they can find a less obvious disconnection that shortens a route. One of the six routes was written for the wrong structure, and it was caught because the model stated its assumption and an expert looked.

An unpublished target from the Reisman laboratory

Professor Sarah Reisman of Caltech shared a target from an unpublished total synthesis in her laboratory, and three Claude models each planned a route to it. Each model had one attempt and no access to the literature, and she judged the routes against what her group has tried at the bench.

> *"This retro synthesis model seems like it could be helpful as a tool for the types of total synthesis efforts we pursue."*
>
> Sarah Reisman, Bren Professor of Chemistry, California Institute of Technology

Why an unpublished target

Synthesis planning is usually tested on molecules whose syntheses are already published, where planning cannot be separated from recall. Professor Reisman offered a target with no reported total synthesis, a highly oxygenated tetracyclic compound with ten stereocenters, one of them an all-carbon quaternary center, which her group has nearly completed. At her request the structure is withheld until her group publishes, so we show the conditions each model chose for the key steps (Table 1) and the opening steps of one route (Figure 1).

How the test was run

Each model received one image of the structure and one prompt, which is reproduced above. The models had no internet, literature or reaction-database access. Their only tool was local code, mainly RDKit, which they used to draw every intermediate, to check the atom balance of every step and, like a chemist with a model kit, to build three-dimensional models that show which face of a carbonyl group is open or whether a bridged ring can close. Each model had one attempt of 32 to 74 minutes. Nothing was edited, selected or corrected, and Professor Reisman received the three documents labeled only A, B and C. All three models planned for the same structure, with the same configuration at every stereocenter.

What the models proposed

All three returned complete routes of 26, 27 and 29 steps (longest linear sequence 19 to 21, with seven to nine protecting-group operations each) and, working independently, chose the same strategy. Two ring fragments are made separately and joined by an aldol reaction that creates the quaternary carbon and the neighboring carbinol. The central ring is closed by ring-closing metathesis, and the last ring by an alcohol-directed epoxidation followed by acid-promoted opening of the epoxide. Table 1 shows how closely the three agreed, down

to the metathesis concentration. Each model named the aldol union and the metathesis as its two most likely failures, and each capped the new aldol hydroxyl group in the very next step to suppress the retro-aldol reaction.

	Opus 5.5	Mythos 5.2	Mythos 5.1
Aldol union of the two fragments	Lithium enolate from the silyl enol ether (MeLi, THF, 0 °C), then the aldehyde, –78 °C	LDA, THF, –78 °C, then the aldehyde, –78 °C	LDA (1.1 equiv), THF, –78 °C, then the aldehyde, –78 °C, 1 h
Next step (to block the retro-aldol reaction)	Ac_2O, Et_3N, DMAP	Ac_2O, DMAP, pyridine	TBSOTf, 2,6-lutidine, –78 to 0 °C
Ring-closing metathesis	Hoveyda-Grubbs II (10 mol %), toluene, 1 mM, 80 °C	Grubbs II (10 to 15 mol %), toluene, 2 mM, 80 °C	Grubbs II (10 mol %), toluene, 1 mM, 80 °C, 6 h
Alcohol-directed epoxidation	$VO(acac)_2$ (10 mol %), TBHP, CH_2Cl_2, rt	$VO(acac)_2$ (10 mol %), TBHP, CH_2Cl_2, 0 °C to rt	$VO(acac)_2$ (5 mol %), TBHP (2 equiv), CH_2Cl_2, 0 °C
Epoxide opening that closes the last ring	CSA (10 mol %), CH_2Cl_2, 0 °C to rt	DDQ to free the alcohol, then CSA (cat.)	CSA (10 mol %), CH_2Cl_2, 0 °C to rt
Fallbacks named by the model	Mukaiyama aldol ($BF_3{\cdot}OEt_2$ or $MgBr_2{\cdot}OEt_2$); relay metathesis or the Stewart-Grubbs catalyst	Mukaiyama aldol ($BF_3{\cdot}OEt_2$); Stewart-Grubbs catalyst, 100 °C, slow addition	Mukaiyama aldol ($BF_3{\cdot}OEt_2$); Hoveyda-Grubbs II or Stewart-Grubbs with an ethylene purge, or a Nozaki-Hiyama-Kishi closure

Table 1. Conditions chosen for the shared key steps, as written in the three documents.

The routes differed in how the fragments were made, and here the models drew on unrelated chemistry. One fragment came from a commercial chiral cyclohexene (Opus 5.5), an oxazaborolidinium-catalyzed Diels-Alder reaction (Mythos 5.2) or a Diels-Alder reaction of a chiral 1-amino-3-siloxydiene (Mythos 5.1, Figure 1). The other came from an Evans syn aldol followed by ring-closing metathesis (Opus 5.5), a cyclopentenol known from the entecavir literature (Mythos 5.2) or a silyl cuprate addition followed by Fleming-Tamao oxidation (Mythos 5.1).

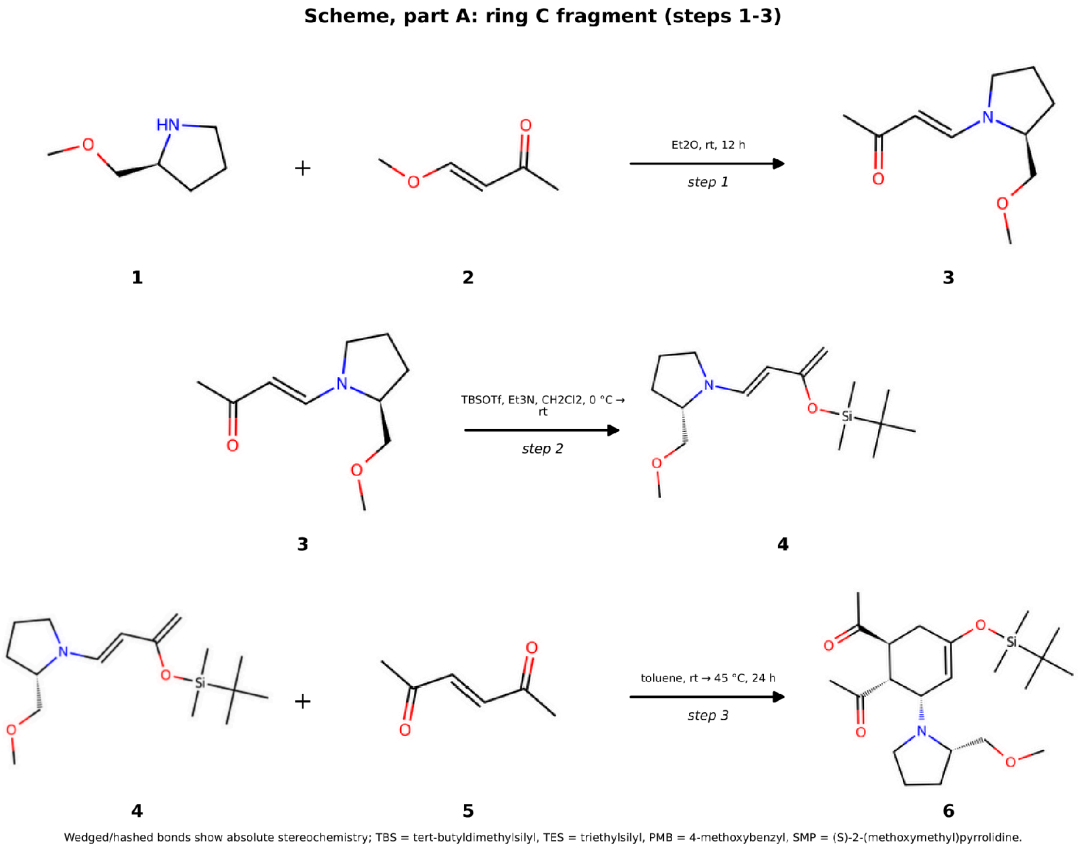


Figure 1. The first three steps of Mythos 5.1's route, exactly as the model drew them. The Diels-Alder reaction in step 3 is a variant of the chiral amino siloxy diene chemistry of Kozmin and Rawal (J. Am. Chem. Soc. 1999, 121, 9562), which the model used without access to the literature and did not cite.

The self-checks changed the routes. Opus 5.5 found that its first plan left a stereocenter between two carbonyl groups, where it would epimerize, and replaced that fragment synthesis. Mythos 5.2 replaced a Wittig methylenation of a 1,4-diketone with the Petasis reagent, because the basic ylide would set off an intramolecular aldol reaction.

Professor Reisman's assessment

Big picture. "The reasoning is generally sound, good use of protecting groups and redox chemistry," and "the stereochemical considerations appear well reasoned." She liked that the models gave a reason for each step, including how any new stereocenter is set, as the prompt asks, and called this "an important aspect of how a human would evaluate a route." She found the three routes very similar in strategy, differing "in the details of the fragments, rather than at a true strategic level," and with that caveat preferred Mythos 5.1. The routes differ from her group's current route and include variations of reactions the group has tried. She looked more at tactics than at the specifics of conditions, such as concentration.

The end game. All three routes finish the same way (Table 1, last rows), then remove the remaining protecting groups and acetylate. “The end game sequences are reasonable,” she wrote. “Model A converges on the same tactics (although not necessarily conditions) that we have investigated.”

Steps that look wrong. Her main concern is the aldol union. Her group has not run this exact reaction, but similar aldol reactions have failed in their hands, and in such products the reverse reaction “can be very facile.” She added that “the prevailing logic of the model makes sense.” Asked whether capping the new alcohol in the next step, as all three routes do, would be enough, she answered that the challenge, as she sees it, is that the aldol reaction is slow at low temperature, and competing processes, such as elimination of the β-alkoxy group, then happen at warm temperatures. Her group designed systems in which a nearby protecting group might trap the alkoxide in situ, and this did not work, so she thinks that “if you get the product, protection of the alcohol is critical to carry through subsequent steps.” Even so, “the bigger challenge is just the reactivity of the hindered system itself,” and the reversibility is part of that problem. The only protecting group her group has been able to cleanly put on the aldol-type products it has made is a TBS ether. Mythos 5.1 capped the alcohol as a TBS ether, and Opus 5.5 and Mythos 5.2 capped it as an acetate (Table 1). She rated the metathesis “a challenging RCM with a reasonable fail possibility,” a tactic her group has avoided but one that could work “if the goal is simply, make the compound.” The models had flagged the same two steps themselves.

Fallbacks. The prompt asked for one fallback for each of the three steps a model judged most likely to fail. “One critique: the fallback suggestions are about conditions, rather than tactics. Typically in a proposal, I like to see other tactics as fallbacks rather than just conditions.”

Surprises. “I was surprised by the quality of the reasoning about possible diastereoselectivity in the sequences. It seemed consistent with what an expert might predict, in the absence of prior experimental data.” In a separate, smaller exercise on a building block, one route used an intermediate her group had not evaluated, in an approach it had set aside. She plans to have a student try it, because it made her interested in “re-evaluating something we had deprioritized.”

Overall. She would have ranked the model “at the level of a solid first year graduate student,” often the person her group gets the “first draft” of a synthesis plan from. She has not recently tried Synthia or every recently disclosed model, and her group has not rigorously evaluated other retrosynthesis algorithms apart from one collaboration, so her comment at the top of this section is not a head-to-head comparison with other models.

Limitations and next steps

These routes are proposals, and none has yet been run in the laboratory. All three models chose a well precedented strategy, and Professor Reisman viewed its key ring closure as a workable way to make the compound but not a strategically interesting one, so this test does not show whether the models can find the less obvious disconnection that shortens a route, or innovate a new tactic. Her main practical request is that every step based on known chemistry carry its literature reference, as added in Figure 1. The models worked without the literature in this test, and supplying those references is a straightforward addition once they work with it. A later test can also ask for fallbacks that change the tactic, which she prefers to fallbacks that change only the conditions.