

Additional file 3

Table S3: Examples of generated text for selected models, selected compounds as input, and selected prompts

Model	Input	Prompt	Generated Text	Comment	Factuality
BioGPT	Aphidicolin	{entity} is	Aphidicolin is a potent inhibitor of DNA polymerase alpha and beta, but not of DNA polymerase gamma.	Aphidicolin is a potent inhibitor of both host cell DNA polymerase alpha and herpes simplex virus (HSV)-induced DNA polymerase but has no effect on DNA polymerases beta and gamma of host cells. [ref]	0
	Ergosterol	{entity} is	Ergosterol is a natural product that has been shown to have a wide range of biological activities.	General description which was repeated by the model several times.	1
	Aphidicolin	{entity} is a compound	Aphidicolin is a compound that inhibits DNA replication by inhibiting the enzyme DNA polymerase alpha.		1
	Aphidicolin	{entity} is a substance	Aphidicolin is a substance that inhibits DNA replication by blocking the replicative DNA polymerase.		1
	Flavasperone	{entity} is a compound	Flavasperone is a compound with a wide range of pharmacological properties.	General description which was repeated by the model several times.	1
	Verrucar A	{entity} is a compound	Verrucar A is a compound with a unique structure and a unique mechanism of action.	General description which was repeated by the model several times.	0
	Fumitremorgin C	{entity} is a compound	Fumitremorgin C is a compound with a unique structure and a unique mode of action.	General description which was repeated by the model several times.	0
	Alternariol	{entity} is a compound	Alternariol is a compound with a wide range of biological activities.	General description which was repeated by the model several times.	1
	Fumitremorgin C	{entity} is a substance	Fumitremorgin C is a substance that is produced by the fungus Aspergillus terreus.	Fumitremorgin C is not produced by the fungus Aspergillus terreus.	0
	Ergosterol	{entity} is a substance	Ergosterol is a substance that is produced by the fungus Ergoderma lucidum.	Model generates new name of fungus. There is no Ganoderma lucidum in the literature. [ref]	0
	Conidiogenone	{entity} is a substance	Conidiogenone is a substance isolated from the fungus Aspergillus versicolor.	Conidiogenone is not isolated from the fungus Aspergillus versicolor.	0
	(+)-Chloromonilicin	{entity} is a substance	(+)-Chloromonilicin is a substance that inhibits the growth of the fungus Fusarium oxysporum f.	(+)-Chloromonilicin doesn't inhibit the growth of the fungus Fusarium oxysporum f.	0
BioGPT-Large	Verrucar A	{entity} is a substance	Verrucar A is a substance that is produced by the fungus Aspergillus versicolor.	Verrucar A is not produced by the fungus Aspergillus versicolor.	0
	(+)-Chloromonilicin	{entity} is	(+)-Chloromonilicin is a new antibiotic produced by a strain of Streptomyces. It is active against Gram-positive bacteria and fungi. The structure of (+) -chloromon	Chloromonilicin is a natural product found in Alternaria sonchi and Monilinia fructicola .	0
ChatGPT			Aphidicolin is a naturally occurring compound that was originally isolated from the fungus Cephalosporium aphidicola. It is a potent inhibitor of DNA synthesis and is commonly used as a research tool to study DNA replication.		
			Aphidicolin inhibits the activity of DNA polymerases, which are the enzymes responsible for synthesizing new DNA strands during DNA replication. By blocking this process, aphidicolin prevents cells from dividing and proliferating. This makes it useful for studying the effects of DNA replication inhibition on cellular processes such as DNA repair, cell cycle progression, and gene expression.		1
			Aphidicolin has also been investigated as a potential anti-cancer agent, as many cancer cells rely on rapid DNA replication to support their growth and survival. However, its clinical use has been limited by its toxicity and side effects.		
			Overall, aphidicolin is a valuable research tool for studying DNA replication and has potential applications in cancer treatment, but further research is needed to better understand its mechanisms of action and potential therapeutic uses.		
	Aphidicolin	{entity} is	Aphidicolin is a naturally occurring compound that is commonly used as a research tool to study DNA replication. It is a potent inhibitor of DNA synthesis and works by blocking the activity of DNA polymerases, which are the enzymes responsible for synthesizing new DNA strands during DNA replication. By inhibiting DNA replication, aphidicolin prevents cells from dividing and proliferating, making it useful for studying the effects of DNA replication inhibition on cellular processes such as DNA repair, cell cycle progression, and gene expression. Aphidicolin has also been investigated as a potential anti-cancer agent, but its clinical use has been limited by its toxicity and side effects.		1
	Aphidicolin	{entity} is a compound	Yes, aphidicolin is a compound. It is a naturally occurring tetracyclic diterpene, which means it is a type of organic molecule made up of four fused rings and derived from the isoprene unit. The chemical structure of aphidicolin consists of 18 carbon atoms, 28 hydrogen atoms, and 2 oxygen atoms.	The chemical formula was incorrect	0
	Aphidicolin	{entity} is a substance	Yes, aphidicolin is a substance. It is a naturally occurring organic compound that is found in certain fungi, including Cephalosporium aphidicola. It can also be synthesized chemically in the laboratory. As a substance, aphidicolin is used as a research tool to study DNA replication and as a potential anti-cancer agent, although its clinical use has been limited due to its toxicity and side effects.		1

In the 'Comment' column there are correct answers or comments by the authors referring to semantic coherence or factuality of generated text.