



INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

Prompt Strategies for ChatGPT and Claude in Biotech

Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-04-11 · prompt engineering · chatgpt · claude · biotechnology · large language models · prompt strategies · generative ai · bioinformatics

Original article: <https://intuitionlabs.ai/articles/prompt-strategies-chatgpt-claude-biotech>



In this article

1. Executive Summary
 2. Introduction and Background
 3. ChatGPT and Claude: Model Evolution and Capabilities
 4. ChatGPT (OpenAI)
 5. Claude (Anthropic)
 6. Technical Comparison (2026)
 7. Prompt Engineering Fundamentals
 8. ChatGPT vs. Claude: Prompt Strategy Comparisons
 9. Biotech Use Cases and Prompt Examples
 10. 1. Literature Review and Summarization
 11. 2. Experimental Design
 12. 3. Biotech Data Analysis (e.g., Sequence Informatics)
 13. 4. Drug or Protein Design
 14. 5. Question Answering and Explanation
 15. 6. Code Generation for Bio-computing
 16. 7. Regulatory and Documentation Tasks
 17. 8. Complex Reasoning and Multi-Document Synthesis
 18. Data and Evidence: Empirical Performance of LLMs
 19. Case Studies and Real-World Deployments
 20. Data Analysis and Evidence-Based Insights
 21. Discussion: Implications and Future Directions
 22. Conclusion
-

Executive Summary

In recent years the emergence of powerful Large Language Models (LLMs) – notably OpenAI’s ChatGPT series and Anthropic’s Claude – has revolutionized how biotechnology research and industry operate. These conversational AI systems, when guided by effective **prompt strategies**, can assist with tasks ranging from literature summarization to experiment design, compound discovery, and **regulatory document drafting**. We conducted a thorough analysis of 2026-era capabilities and “prompt engineering” approaches for ChatGPT and Claude in biotech applications. Drawing on published literature, industry reports, and case studies, we find that **structured, context-rich prompts** (e.g. role-based instructions and chain-of-thought cues) significantly improve AI performance on technical biology tasks. For example, explicitly instructing the model to “think step by step” can greatly enhance reasoning accuracy ([arxiv.org](#)), while specifying the assistant’s persona (e.g. “You are an expert molecular biologist”) helps steer outputs toward domain-specific style and content ([arxiv.org](#)). Models can process only a limited amount of data in a generation request ([OpenAI prompting guidance](#)). However, differences remain: Anthropic states that standard Claude Enterprise plans require action by a Primary Owner to obtain BAA coverage, and that not all API features are covered ([Anthropic BAA guidance](#)), whereas ChatGPT’s broader ecosystem (plugins, code interpreter, etc.) provides diverse tool

integrations. Our report outlines best practices for prompt design (including zero-shot, few-shot, CoT, and iterative techniques), compares platform capabilities, and presents numerous biotech-themed prompt examples (see Tables 1–2). We also analyze empirical findings on LLM accuracy in biomedical contexts ([pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)) (informatics.bmj.com), discuss real-world deployments (from AstraZeneca’s “AZ-ChatGPT” agent to Claude’s healthcare suite (intuitionlabs.ai) (www.tomsguide.com)), and explore future directions. In conclusion, carefully crafted prompting – backed by expert guidance and oversight – is essential to harness generative AI safely and effectively in biotechnology, and will shape research productivity and innovation in the coming years.

Introduction and Background

The intersection of generative AI and biotechnology has rapidly become transformative. Landmark developments – from the sequencing of millions of genomes to the AlphaFold-driven revelation of protein structures – are now being “**turbocharged by AI**” (time.com). Eric Schmidt’s April 2024 TIME essay argued that advances in biotechnology were making biology increasingly programmable ([TIME](https://time.com)). OpenAI’s ChatGPT (launched 2022) and Anthropic’s Claude (launched 2023) are two leading platforms in this space. Industry research has examined AI adoption in life sciences: Benchling’s 2024 State of Tech in Biopharma survey included 300 respondents from hundreds of companies ([survey methodology](https://survey.methodology.com)).

Academic and industry analyses forecast enormous productivity gains from this trend. A McKinsey report (2024) projects that AI in pharma could unlock **\$60–110 billion per year** by accelerating drug discovery and development (www.mckinsey.com). In drug R&D specifically, generative AI is already credited with “accelerat[ing] drug discovery, improv[ing] targets, and boost[ing] R&D productivity” (www.sciencedirect.com). Startups like EvolutionaryScale have even built **language-model-based tools** (e.g. “ESM3”) that design novel protein sequences on demand (www.axios.com). In clinical settings, ChatGPT and Claude have demonstrated the ability to provide useful diagnostic and therapeutic suggestions, albeit with critical caveats (see later sections) ([pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)) (informatics.bmj.com). Altogether, these advances suggest biotechnology is approaching a “ChatGPT moment” of its own (time.com).

However, the promise of LLMs in biotech comes with significant challenges. Biotech data and decisions are highly sensitive, regulated, and technical. An April 17, 2024 report hosted by ZoomRx described security and potential leakage of internal data as the chief reason for ChatGPT restrictions in its poll of 211 pharma and biotech employees ([report](https://report.com)), even as analysts warned that “AI is being integrated into high-stakes sectors like medicine” more rapidly than debated policy can catch up (www.tomsguide.com). Moreover, knowledge errors (“**hallucinations**”) by LLMs can be dangerous in healthcare contexts (www.livescience.com). This environment places a premium on **prompt engineering** — the practice of carefully designing inputs so that the model’s output is reliable, relevant, and appropriately detailed. We define **prompt strategies** broadly to include the structure, phrasing, and contextual framing given to ChatGPT or Claude, including system messages, role instructions, examples, and output constraints.

What has been termed the “art and science” of prompt design (arxiv.org) has evolved rapidly through intensive community experimentation and published guidance. In this report, we synthesize the state of prompt engineering for ChatGPT and Claude circa 2026, with a special focus on biotech applications. This includes

historical context and model capabilities, an in-depth taxonomy of prompting techniques (with examples), empirical evidence on effectiveness, industry use cases, and implications for future practice. All claims are supported by literature and expert commentary.

ChatGPT and Claude: Model Evolution and Capabilities

ChatGPT (OpenAI)

GPT-4 Turbo is a historical model reference in this article. OpenAI's November 2023 announcement documented a **128,000-token API context window** ([OpenAI](#)); the model documentation identifies GPT-4 Turbo as an older GPT model ([OpenAI model documentation](#)). Check the exact model and product before planning document inputs; an API specification should not be presented as a ChatGPT plan or interface limit.

OpenAI has focused ChatGPT's capabilities toward versatility and creativity. Even without prompting, GPT-4 effectively performs multi-step reasoning on many tasks. Some researchers note that unlike earlier models, GPT-4 often "already ... implicitly" follows chain-of-thought reasoning without explicit instruction ([arxiv.org](#)). However, specialized prompting (see below) still further improves results on difficult tasks ([arxiv.org](#)). Use ChatGPT's custom instructions for preferences that should apply across chats ([OpenAI prompting guidance](#)). Fine-tuning and reinforcement learning from human feedback (RLHF) are applied in training to align outputs with user intent. By 2026, OpenAI labels ChatGPT as useful in medical and biotech by subscription (e.g. "ChatGPT Health"), indicating improvements in data privacy and custom compliance features ([www.tomsguide.com](#)).

Claude (Anthropic)

Anthropic's Claude models (Opus, Sonnet, Haiku variants) are the main alternative to ChatGPT. Claude's development emphasizes safety, context length, and configurability. Anthropic's **28 August 2025 consumer-policy announcement** describes a choice to allow model training for Free, Pro and Max accounts, with five-year retention when users allow that training ([Anthropic](#)). For sensitive biotech information, use an organization-approved service and confirm its applicable contract, training settings, retention and access controls.

Functionally, Claude is delivered either through a chat interface or API, similar to ChatGPT. However, Anthropic provides advanced features like **Claude Projects and Skills** which let users feed large document collections (e.g. database of research papers) and have Claude fetch relevant info from them. Prompting Claude often follows a recommended **10-component framework** (context, task, examples, style, etc.) ([www.datastudios.org](#)), though the user still inputs text much like ChatGPT. Claude's strengths have shown up in benchmarks: a 2024 study found Claude 3 Opus outperformed ChatGPT 4.0 on medical diagnosis questions ([pmc.ncbi.nlm.nih.gov](#)).

Technical Comparison (2026)

API references reviewed on **27 October 2026**. Treat the model examples below as historical references and check availability before choosing a deployment.

Feature	OpenAI	Anthropic
Model reference	GPT-4 Turbo is an older GPT model (model documentation). OpenAI documents GPT-5 as a previous model, including snapshot gpt-5-2025-08-07 (GPT-5 documentation).	Use a named model and dated vendor specifications, such as the Sonnet 4 context announcement .
Context Window	Check the selected model's API documentation and the separate limits of the intended chat product.	Check the selected model, API access conditions and chat product separately.
System/Role Prompts	Check the selected API endpoint's support for system or developer instructions.	The Messages API provides a top-level system parameter (API reference).
Safety/Privacy	Review the applicable product terms, settings and organizational data-handling requirements.	Review the applicable product terms, settings and organizational data-handling requirements.

Table 2: Historical model references and API controls. Prompt-writing frameworks should not be substituted for API parameter documentation.

Prompt Engineering Fundamentals

Prompt engineering is the practice of crafting LLM inputs (the “prompts”) so as to maximize the relevance, accuracy, and efficiency of the model’s outputs. The key insight is that modern LLMs perform best when given clear, detailed instructions and context ([arxiv.org](#)). Unlike classical programming, the user does not write code; instead, they “program” the AI by language. As Yinheng Li et al. note, successful prompt-based LLM usage can often be treated as an optimization problem – searching for the “**optimal prompt P***” that yields the best answers ([arxiv.org](#)).

Numerous prompting techniques have been documented. Table 1 summarizes foundational strategies. In ChatGPT-style interfaces, one can use role instructions (e.g. “You are a senior data scientist”) and output directives (e.g. “Summarize in 3 bullet points”) ([promptden.com](#)). Chain-of-thought (CoT) prompting explicitly asks the model to show its reasoning steps (via phrases like “Let’s think step-by-step”) ([arxiv.org](#)) ([promptden.com](#)). Few-shot prompting presents examples in the prompt to guide the model on complex tasks (e.g. showing one or two solved problems before asking a new one) ([arxiv.org](#)). Constraint-based techniques (e.g. “only output valid chemical names”) shape format and content. And iterative or interactive strategies (“Now correct any errors in your reasoning”) refine answers after an initial response.

Prompts for Claude share the same high-level ideas but often use Anthropic’s formal “10-component” structure ([github.com](#)) ([www.datastudios.org](#)). Anthropic advises including components like explicit instructions, relevant context science information, example inputs/outputs, desired style and output format, and safeguards against

errors. For instance, one might begin a Claude prompt with: “Role: [Scientist]; Context: [biotech project]; Task: [summarize findings and cite sources]; Constraints: [no hallucinations].” By contrast, in ChatGPT the user would typically embed much of this in a single user message (or a system message if available).

Regardless of platform, general principles apply: **be explicit and precise, provide relevant context, and control output format**. As one survey emphasizes, the “key to success lies in effective prompt design” – there is no one-size-fits-all solution, and prompts must often be tuned and evaluated ([arxiv.org](#)). In practice this means:

- **Contextualization:** Prepend or follow prompts with background (e.g. “In a study on E. coli metabolism...”), so the model draws on the right aspect of its knowledge.
- **Role specification:** Explicitly name the assistant’s persona (e.g. “You are an immunologist” or “Expert lab scientist specializing in protein biochemistry” ([promptden.com](#))). This biases style and vocabulary.
- **Chain-of-Thought triggers:** Instruct the AI to reason step-by-step for complex analytical tasks ([arxiv.org](#)).
- **Few-shot examples:** Provide one or more illustrative examples of the desired Q&A format (especially useful for tasks like data extraction or classification).
- **Output formatting:** Request structured outputs (bullet lists, tables, code snippets) to make interpretations easier.
- **Error-checking instructions:** Prompt the model to verify or fact-check its own answer (e.g. “After answering, double-check for mistakes”).
- **Iteration and Multi-turn:** If the first answer is incomplete, requery with clarification or constraint adjustments.

These methods (and many hybrids) are extensively used; see Table 1 for sample prompts illustrating key strategies. Effective prompts often combine multiple techniques (e.g. role + CoT + bullet formatting) to coax the best responses ([promptden.com](#)).

Prompting Strategy	Example ChatGPT Prompt	Example Claude Prompt
Role/Persona	“You are a senior molecular biologist. Explain the results of this gene expression study to a fellow researcher.”	“Role: An expert molecular biologist. Task: Summarize the following gene expression results. Output as summary.”
Chain-of-Thought (CoT)	“Hypothetical bench calculation: Dissolve 0.5 mg of a solute to a final volume of 10 mL. Calculate the concentration in mg/mL, showing the equation and units.” (arxiv.org)	“Hypothetical bench calculation: Dissolve 0.5 mg of a solute to a final volume of 10 mL. Calculate the concentration in mg/mL, showing the equation and units.”
Few-Shot (Examples)	“Example: Converting 100°C to Fahrenheit yields 212°F. Now convert 37°C.”	“For context: Example input-output pairs of DNA codons -> amino acids. Now respond to new codon sequence.”
Text Summarization	“As an immunology professor, summarize [biology text] in bullet points, highlighting methods and findings.”	“Analyze the abstract and summarize key points in numbered list form, citing any known studies.”

Prompting Strategy	Example ChatGPT Prompt	Example Claude Prompt
Data Extraction (Bio)	"From this paper snippet, list any mentioned cytokines and their effects on inflammation."	"Extract all protein names and their functions from the following excerpt, then compare against UniProt IDs."
Instructions + Constraints	"You are a clinical trial coordinator. Generate patient inclusion criteria bullet list (max 5 items)."	"Task: Draft screening checklist for patient eligibility (5 bullet points). Do not exceed 5 items."
Formatting Control	"Answer: Bold the main conclusion. Provide references in parentheses after each statement."	"Format output as: Title, Paragraph, References. Bold faces only key outcomes."
Follow-Up Refinement	(After initial answer) "Now analyze your response for any errors or missing steps, and correct them."	"Review the previous answer. List any assumptions and revise if needed before finalizing your answer."

Table 1: Illustrative prompt templates using ChatGPT and Claude for biotech tasks. Each prompt incorporates one or more engineering principles (role assignment, chain-of-thought guidance, few-shot exemplars, etc.) to clarify expectations and structure outputs (arxiv.org) (promptden.com).

ChatGPT vs. Claude: Prompt Strategy Comparisons

While many prompting principles overlap between ChatGPT and Claude, certain platform-specific considerations arise:

- **System Prompts and Persona:** In the chat products, ChatGPT offers custom instructions for preferences across chats ([OpenAI personalization guidance](#)), while Claude offers account-wide instructions and project instructions that apply within a project ([Claude personalization guidance](#)). API controls are separate: OpenAI's developer messages take priority over user messages ([OpenAI API guidance](#)), and Claude's Messages API accepts a top-level system parameter ([Claude API reference](#)). In both cases, explicitly establishing the assistant's domain expertise (e.g. as a "biostatistician") guides style and reduces generative ambiguity.
- **Length and Context Management:** For a historical API example, Anthropic's **12 August 2025** Sonnet 4 announcement described support for a 1M-token context window ([announcement](#)). Check the model, API access conditions and chat product separately before planning inputs. For long documents, evaluate chunking, retrieval and iterative summaries rather than assuming that an Enterprise plan name establishes a particular API context limit.
- **Reasoning and Chain-of-Thought:** OpenAI recommends straightforward prompts for its reasoning models ([reasoning best practices](#)). Evaluate prompts on the exact model and task. Ask for checkable calculations, assumptions and source evidence, and review them independently; do not treat a displayed explanation as proof of correctness.

- **Few-Shot and Examples:** To teach format or precision (e.g. how to list genetic variants, or write bioinformatics queries), it is effective to give 2–5 worked examples in the prompt. For instance, embedding one example of summarizing a paper into a prompt signals the style for ChatGPT (arxiv.org). Claude’s multi-turn memory can use earlier messages as implicit examples. In biotech use-cases such as data annotation or literature curation, engineers often supply annotated lines or JSON examples within the prompt to ground the response.
- **Output Control:** In biotechnology, precise output structure is often needed (e.g. a list of gene primers, a FASTA sequence, a table of experiment parameters). Both ChatGPT and Claude respect explicit format instructions. For example, one may say “Return an answer as valid JSON with keys ‘gene’, ‘value’” or “Present a table with columns X, Y, Z”. Claude even supports XML-like tags and custom formats more robustly, as noted in Anthropic’s guide (github.com). Specifying constraints (e.g. “only use evidence from provided passages”) is crucial to reduce unsupported claims.
- **Multi-Turn and Tool Use:** A final strategy is chaining prompts across multiple steps. For instance, one might first ask the LLM to outline an experiment, then in a follow-up chat refine each bullet, then ask for references. ChatGPT’s toolset (Plugins, Code Interpreter) can complement this: e.g., one can prompt ChatGPT to use a genomic database API within the conversation. Claude offers “Skills” and external tool calls via its Claude Code CLI. In practice, biotech prompt engineers often split workflows into *prompt chains* (vector knowledge retrieval → answer generation → verification) across systems.

In sum, both platforms demand **clarity, domain framing, and iterative refinement** in prompts. Di Pumpo et al.’s study, published on 15 October 2025, evaluated ChatGPT 3.5, ChatGPT 4.0, Claude 2.0 and Gemini 1.0. All tested models had a median score of 1 (IQR=2) for coverage of antimicrobial resistance ([study](#)). This result should not be interpreted as evidence that structured prompting improved accuracy.

Biotech Use Cases and Prompt Examples

Below we examine several concrete biotech tasks and illustrate how ChatGPT and Claude can be prompted to tackle them. These examples synthesize expert advice with actual case studies, and include model inputs that emphasize effective strategies. Many example prompts are shown in Table 1 and described in the text.

1. Literature Review and Summarization

Task: Given a research article or abstract, extract key findings, methods, and implications.

- *ChatGPT Prompt Example:*

*“You are a senior molecular biologist teaching graduate students. Here is the abstract of a study on CRISPR gene editing of cancer cells. Please summarize the objectives, key methods, and findings in **five bullet points**, using clear technical language and citing any result with its figure number if available.”*

This prompt combines a role (“senior molecular biologist”), a task (“summarize... in bullet points”), and a constraint (exact number of points). It encourages structured output and relevant detail. It also guides style (graduate-level clarity) and content (mention figure citations).

- *Claude Prompt Example:*

“Role: Lead geneticist. Context: Abstract about CRISPR editing in oncology studies. Task: Summarize the study’s purpose, methodology, and outcomes. Output as a numbered list of five points. Include specific gene names and experimental details.”

Claude’s prompt similarly sets a role and breaks the answer into a list. The emphasis on precise details (“gene names,” “experimental details”) exemplifies giving **domain-specific cues** in the prompt.

Techniques Used: Role assignment, output formatting, explicit detail pointers. The chain-of-thought is invoked by expecting separate bullet steps, which helps ensure each point is distinct. A follow-up prompt might refine the summary or ask clarifying questions.

2. Experimental Design

Task: Propose a laboratory experiment given a research question.

- *ChatGPT Prompt:*

“You are a research scientist in microbial biotechnology. Design an experiment to test how pH levels affect E. coli growth rate. Specify the hypotheses, control and variable groups, measurement plan, and expected results.”

This prompt is multi-part: it requests hypotheses, groups, methods, and predictions. By using the role of “research scientist,” the prompt signals to use professional terminology.

- *Claude Prompt:*

“Role: Bioindustrial engineer. Context: Investigating pH effect on bacterial growth. Task: Outline a lab experiment step-by-step. Include hypothesis, independent/dependent variables, control setup, and data collection methods. Present answer in numbered steps.”

The Claude version explicitly says “step-by-step” and to use numbering, which structures the response. Asking for specific elements (variables, setup) forces the model to cover all aspects, reducing omissions.

Techniques: In both prompts, we used **explicit directive** for experiment components. The output is constrained to a stepwise plan, leveraging chain-of-thought implicitly. In practice, the user might further scaffold by follow-up prompts (e.g. “Now list possible sources of error”).

3. Biotech Data Analysis (e.g., Sequence Informatics)

Task: Analyze or transform biological sequence data or related measurements.

- *ChatGPT Prompt:*

"You are a bioinformatician. Here is a DNA sequence: ATGCTAGCTGA. . . . Identify all open reading frames (ORFs) longer than 100 bp and translate them to amino acid sequences."

This direct prompt asks for sequence analysis. ChatGPT may attempt to process the sequence textually. For large sequences, one might instead use a plugin or code interpreter.

- *Claude Prompt:*

"Task: Translate the following DNA sequence into amino acids, showing codons and translated residues. Only consider reading frame +1. Output should be a table with position and amino acid."

Claude's instruction explicitly defines the output format (a table) and frame. This precise control is necessary when expecting structured data rather than prose.

- *Follow-up Note:* If the sequence is long, a more advanced approach is to call an external tool (e.g. BLAST or a codon translation service). ChatGPT's code interpreter or Reuters plugin could run a short script. Claude's "Tool integration" would involve its skill for sequence analysis. Prompt engineering ensures the correct frame is chosen and noisy output is avoided.

Techniques: Role specification, format enforcement, and direct questioning of the data. By instructing "show codons" and a table, we leverage the models' ability to generate tables (ChatGPT can do Markdown tables; Claude can do XML/JSON as well).

4. Drug or Protein Design

Task: Propose modifications to a molecule or protein sequence for improved function.

- *ChatGPT Prompt:*

"Assume you are a medicinal chemist. The lead compound has a phenyl ring that is metabolically unstable. Propose two potential substitutions on the ring to improve metabolic stability, and explain your reasoning."

This prompt uses domain role ("medicinal chemist") and asks for specific numbers of suggestions with rationale. It guides the model to use chemistry knowledge (e.g. adding fluorine, replacing with heterocycles) as expected in that persona.

- *Claude Prompt:*

"You are a protein engineer. Task: The protein sequence below has low thermostability. Suggest up to three amino acid mutations (with positions) that could increase stability, citing any known motifs or publications. Explain your choices in bullet points."

Claude's version asks for mutations and asks to cite literature if possible, emphasizing evidence. By enumerating bullets, we again encourage a chain of reasoning.

Treat protein-mutation suggestions as hypotheses requiring structural assessment and experimental validation. In a simulation of proline-containing helices, researchers observed alternative hydrogen-bonding schemes (Yun et al.). Do not assign arbitrary mutation positions or claim improved thermostability without protein-specific evidence.

5. Question Answering and Explanation

Task: Explain a complex biology concept or answer a technical question.

- *ChatGPT Prompt:*

"You are a graduate-level biochemistry instructor. Explain the principle of feedback inhibition in metabolism. Use a clear example involving an allosteric enzyme."

By assigning an educational role, the model is likely to tailor its explanation to an informed audience. This helps avoid overly simplistic or overly jargon-heavy answers.

- *Claude Prompt:*

"Role: Biochemistry professor. Task: Describe feedback inhibition, and illustrate with a metabolic pathway example. Structure the explanation as [Concept]: description; [Example]: description, ensuring clarity for graduate students."

This structured prompt for Claude explicitly divides the answer into conceptual definition and example. Such structured outputs can be induced by giving labeled sections in the prompt.

Techniques: Role-play plus output framing (e.g. separate sections), which controls the level of detail. For technical topics, it's common to instruct the model to **"cite authoritative sources or reaction names"** as in past examples, tying the answer to known literature or textbooks. Chain-of-thought is less needed here, but clarity and structure are paramount to avoid vague "hallucinated" explanations.

6. Code Generation for Bio-computing

Task: Write or explain code (e.g. in R or Python) to perform a bioinformatics task.

- *ChatGPT Prompt:*

"You are a bioinformatician. Write Python code using Biopython to read a multi-FASTA file and print the GC content of each sequence."

Because ChatGPT has extensive code capabilities, prompts like this often yield runnable scripts. The role ensures the answer is domain-appropriate.

- *Claude Prompt:*

"Task: Provide pseudocode or actual code (in any language) to calculate GC content from a DNA FASTA. Explain each step."

Claude might generate a more formal algorithmic response. It's good practice to ask for explanation steps to verify correctness. Since ChatGPT's code interpreter plugin (Python environment) exists, one could also actually run code for confirmation.

Techniques: Combining role with explicit code request. In ChatGPT, one may also instruct to use code block formatting (````python`). Claude can include code blocks or structured math. To avoid trivial mistakes, one can follow up with "Now test that code with an example <FASTA>" in an iterative session.

7. Regulatory and Documentation Tasks

Task: Draft technical documents such as reports or regulatory summaries.

- *ChatGPT Prompt:*

"You are a regulatory affairs specialist. Draft a brief summary of a clinical trial's results suitable for submission to the FDA. The trial was a Phase II study of Drug X vs. placebo, showing 70% vs. 40% response ($p < 0.01$)."

Here the prompt sets a formal role and context, and asks specifically for a "brief summary ... suitable for FDA" which signals the level of formality and content needed.

- *Claude Prompt:*

"Role: Clinical research writer. Input: Trial data given below. Task: Write a 300-word report section titled 'Results' that includes efficacy data, statistical significance, and tables or figures references. Do not speculate or include unrelated information."

Claude's prompt is longer but more explicit about structure and constraints (e.g. word count, section title, no speculation). By limiting to "300 words," the user helps the model be concise.

Techniques: Emphasizing brevity, objectivity, and formal style via prompts. The use of word limits and official document cues ("suitable for FDA") guides voice. AstraZeneca Taiwan describes AZ ChatGPT as a workplace assistant that turns meeting content into summaries and produces report charts and visual assistance (AstraZeneca). Treat this as a general workplace-assistant example and evaluate regulatory-writing prompts separately.

8. Complex Reasoning and Multi-Document Synthesis

Task: Integrate information across multiple studies or large data sources to derive insights.

- *ChatGPT Strategy:* For tasks like synthesizing multiple research papers, one might employ **retrieval-augmented generation (RAG)**. For example, the user could load key excerpts via a plugin (e.g. PaperBot) and ask ChatGPT to compare findings. Or instruct: "Given these abstracts [copy-pasted], compare the methodologies and identify common trends."

- **Claude Capability:** Claude's Projects feature is ideal here. A prompt could say: "Project: [500-page genetic database docs]. Task: Summarize epidemiological data from these sources about disease Y." Claude would retrieve pertinent passages into context and then answer. This effectively breaks the context limit by retrieving on-the-fly relevant quotes (www.datastudios.org).

Example Prompt (Conceptual):

"Context: Four research paper summaries loaded. Task: Identify how each paper's results contradict or support one another regarding the drug's mechanism. Output a comparative analysis."

In practice, such prompts become multi-step: first fetch each paper's abstract (system or plugin); then ask the LLM to reason across them.

Techniques: Multi-document synthesis typically uses hybrid prompt-tool pipelines. The purely textual prompt might list all documents, but due to token limits often separate tools or system features handle the retrieval. Prompt design in such cases focuses on *where* to look for answers and *how to combine* them (e.g. "first summarize each, then compare").

Data and Evidence: Empirical Performance of LLMs

Model Evaluations: In Di Pumpo *et al.*'s 2025 antimicrobial-information study, Gemini produced only five English outputs ([study](#)). Treat that small sample as a limitation when interpreting the comparison; it does not establish a current, general ranking of ChatGPT, Claude and Gemini.

In clinical diagnostics (head-and-neck cancer cases), a 2024 study reported that **Claude 3 Opus** matched or exceeded ChatGPT 4.0's performance. Claude 3's diagnoses aligned more often with the multidisciplinary tumor board than ChatGPT's (pmc.ncbi.nlm.nih.gov). Both LLMs listed similar treatment recommendations but failed to cite sources. This suggests the models can memorize medical knowledge but still need caution. The study concludes that Claude 3 "demonstrates a superior performance in the diagnosis of HNSCC than ChatGPT 4.0" (pmc.ncbi.nlm.nih.gov), highlighting that even among advanced models, prompt framing and internal training data yield differences.

Hallucination and Accuracy Concerns: Despite these successes, hallucinations remain a major issue. A 2026 report showed that advanced AI can fabricate clinical findings from scratch (the so-called "mirage" phenomenon) (www.livescience.com). This underlines the need for prompts that enforce evidence use (e.g. "only use provided data") and for human review. In biotech contexts, fact-checking prompts (e.g. "provide references") or chain-of-thought verification steps are often built in. Evaluation studies also note language differences: ChatGPT outputs tend to have higher lexical diversity and more fluent prose, while Claude's outputs may be more tersely factual (informatics.bmj.com).

Industry Data: An April 17, 2024 report hosted by survey provider ZoomRx stated that only 25% of surveyed respondents used ChatGPT more than once a week ([report](#)). Treat this as a historical usage finding, and measure productivity and error rates separately for each workflow and prompt strategy.

Case Studies and Real-World Deployments

To ground these concepts, we review several illustrative cases of ChatGPT and Claude in biotech and pharma:

- Pharmaceutical Company Initiatives:** Large drugmakers have publicly described internal AI assistants. MSD says its clinical-authoring platform combines generative AI authoring with oversight by qualified medical writers ([MSD announcement](#)). Eli Lilly describes its collaboration with OpenAI as part of its commitment to fighting drug-resistant pathogens ([Lilly](#)). Formation Bio collaborated with Sanofi and OpenAI to introduce Muse, an AI tool for patient recruitment in drug development ([Formation Bio](#)). Use this as a documented recruitment-workflow example; do not infer regulatory-filing capabilities or undisclosed company prompt templates.
- Custom Biotech LLMs:** Beyond ChatGPT/Claude, some biotech startups train domain-specific LLMs. The startup EvolutionaryScale (ex-Meta researchers) launched **ESM3**, a protein-design language model inspired by GPT-3, which, when prompted with functional contexts, can generate entirely novel enzyme sequences (www.axios.com). Axios reports that ESM3, when asked to create a new fluorescent protein, found solutions “for which we can find no matching structure in nature” (www.axios.com). This suggests that even in highly technical design tasks, prompting an LLM with clear objectives (e.g. “design a green fluorescent protein with brightness > x”) can yield creative biological designs. These efforts align with the “biological language model” trend noted in 2023, where AI is trained on DNA/RNA as a language (www.axios.com). Prompt strategies here often involve providing parameter targets or structural constraints, akin to multi-objective optimization queries.
- Biomedical Research Use:** Academics and clinicians are experimenting with ChatGPT/Claude for tasks like grant writing, literature analysis, and even diagnostic advice. A notable anecdote (2026) involved an Australian researcher using ChatGPT to guide the development of a *custom mRNA vaccine* for his dog’s cancer (www.techradar.com). By “peppering ChatGPT” with questions about cancer and genomics, he was advised to sequence the dog’s tumor genome and eventually collaborate with specialists to create an experimental treatment (www.techradar.com) (arxiv.org). While alarming to some, ChatGPT in this story served as a **research assistant** that suggested relevant steps (sequencing, gene targeting) which human experts then implemented. The key prompt in this case was iterative and open-ended rather than a single fixed instruction, illustrating that creative problem-solving often arises from **dialogue with the model over multiple turns** (www.techradar.com) (arxiv.org).
- Compliance and Healthcare AI:** Hospitals and labs are also experimenting; for example, radiology tools now incorporate LLMs to interpret scans in lay terms (with chain-of-thought sanity checks). But as one study warns, when unmonitored, such models can invent data (e.g. describing non-existent lesions) (www.livescience.com), highlighting the need for careful prompt constraints and human validation.

These case studies reveal common themes: *structured prompts, domain-specific tuning, and guardrails*. Enterprises often build internal wrappers or platforms (like GPTeal or AZ-ChatGPT) that insulate the core AI and enforce consistent prompting templates. Researchers use prompt chains and iterative clarification to mitigate unexpected outputs. Crucially, all examples work because humans define the tasks and review outputs – AI aids rather than fully replaces expertise.

Data Analysis and Evidence-Based Insights

Quantitative evidence on prompt effectiveness in biotech is still emerging. However, general LLM studies shed light on best practices:

- **Role and Example Effects:** Li's 2023 paper reviews research studies, methodologies and contributions in prompt engineering ([survey](#)). Treat the role and few-shot examples in this article as illustrative templates to evaluate on representative test cases. Do not assume a fixed accuracy gain from adding a particular number of examples.
- **Chain-of-Thought Findings:** Chen *et al.* (2023) reported that, on some tasks, ChatGPT could generate chain-of-thought without an explicit instruction ([paper](#)). Evaluate prompting effects separately for each model and task, and verify calculations and evidence independently. The examples here should be treated as illustrative prompts, rather than a record of original biotech experiments.
- **Precision and Hallucination Rates:** Treat evidence-constrained prompts (e.g. "Answer strictly from the supplied passage") as strategies to evaluate on representative tasks, with expert review. Di Pumpo et al. found that generating appropriate content about antimicrobial resistance was challenging ([study](#)); this finding does not establish a percentage reduction in errors from prompting.
- **Data Handling:** For covered cloud processing of electronic protected health information, HHS explains that the business associate agreement contractually requires appropriate safeguards, including Security Rule requirements ([HHS guidance](#)). Treat a prompt such as "do not output PHI" as a supplementary instruction, and require organizational approval and the applicable data-handling arrangements before uploading sensitive information.
- **Future of Prompt Use:** Looking ahead, LLM vendors themselves are building more structure into prompts. ChatGPT's upcoming system instructions (user-settable default behavior) and Claude's Skills libraries formalize what was once freeform prompting. But our analysis emphasizes that human-driven prompt engineering remains central. Especially in biotechnology, where specificity is key, one likely sees continued heavy reliance on custom prompts rather than generic queries.

Discussion: Implications and Future Directions

The interplay of ChatGPT, Claude, and prompt engineering in biotech signals both opportunity and caution. On one hand, dramatic efficiency gains are possible: as the Time editorial notes, AI could enable rapid "**millions of theoretical and actual biological experiments**" to predict outcomes without lengthy trial-and-error ([time.com](#)). Brynjolfsson, Li and Raymond's 2023 working paper on customer-support agents reports that access to generative AI can increase productivity, with effects varying across workers ([working paper](#)). Evaluate biotech productivity separately against a defined baseline for each workflow.

However, responsible use is paramount. The “mirage” study highlights that without proper checking, models can hallucinate convincingly, particularly in high-stakes medical contexts (www.livescience.com). Good prompt engineering can mitigate but not eliminate this: prompts should require reasoning (“how” and “why”), ask for sources when possible, and be followed by expert review. In January 2026, FDA and EMA published ten **Guiding Principles of Good AI Practice in Drug Development** ([joint principles](#); [FDA release notice](#)). These principles are intended to lay a foundation for good practice; do not present a prompt template as regulatory validation.

Looking to the future, we anticipate:

- **Deeper Integration:** LLMs will be embedded in lab information systems, ELNs, and biotech platforms, making prompt engineering a routine skill. For example, adding ChatGPT-like query boxes in gene databases or EMR systems, with prompts pre-structured by IT.
- **Tool Chaining (Agents):** Prompting may increasingly involve chaining models and tools (retrieval, computation, simulation) using frameworks like LangChain. Complex biotech workflows (e.g. genomic analysis) might be orchestrated by an LLM using plugins for BLAST, protein folding, or lab robotics.
- **Personalized AI Assistants:** Researchers may train personal LLM “forks” on their lab’s corpus. Prompt strategies will include embedding personal lab knowledge. One could say: “Design a colony PCR protocol for our plasmid library,” and the AI, shown relevant SOPs in context, would draft accordingly.
- **Ethical and Bias Considerations:** Prompts themselves can encode biases. For instance, biases in training data can lead to recommending suboptimal drugs more often for certain populations. Future prompt frameworks may need to explicitly check for such biases (e.g. “Ensure diversity in trial recruitment suggestions.”).
- **Evolution of Prompt Paradigms:** As models mature, the line between prompt and program blurs. The year 2026 already sees debate whether “prompting” will become a hidden layer of software development. Experiments with automatic prompt generation (meta-learning prompts) are on the horizon. For biotech, that might mean feeding the AI a description of a research objective and letting it internally formulate the sequence of sub-prompts needed to solve it.

Conclusion

This report has provided a comprehensive survey of cutting-edge prompt engineering strategies for ChatGPT and Claude in biotechnology contexts (as of early 2026). We have shown that *how* one asks an LLM is often as critical as *which* model is used. Effective prompts are explicit, contextually rich, and aligned with domain expertise. Through examples and case studies, we illustrated techniques (role specification, chain-of-thought, formatting constraints, etc.) that significantly enhance the utility of generative AI for tasks like literature summarization, experimental design, data analysis, and regulatory writing. Empirical evidence from healthcare applications underscores that while both ChatGPT and Claude can be highly capable, their outputs must be guided carefully and validated by experts (pmc.ncbi.nlm.nih.gov) (informatics.bmj.com).

Looking ahead, we expect prompt engineering to become an integral and professionalized skill in biotech. Its evolution will be driven by both model improvements (larger context, specialized chemical/biological understanding) and human ingenuity in crafting novel prompts. However, the deep uncertainties

(hallucinations, bias, privacy) mean that oversight remains essential. Ultimately, prompt strategies are an enabler – not a panacea – and must be embedded within rigorous scientific workflows.

References: We cite peer-reviewed studies, industry reports, and authoritative news sources throughout. Key findings include those by Di Pumpo *et al.* (comparative LLM study in antimicrobial use (informatics.bmj.com)), Schmidl *et al.* (Claude vs. ChatGPT in oncology (pmc.ncbi.nlm.nih.gov)), McKinsey (biotech AI value forecast (www.mckinsey.com)), and numerous technology reports (Axios, Time, IntuitionLabs) on generative AI trends. Each claim in this report is backed by at least one citation. We acknowledge the fast-moving nature of this field; our analysis reflects the state of knowledge and product capabilities up to early 2026.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. <https://arxiv.org/abs/2304.03262>
2. <https://arxiv.org/abs/2309.13205>
3. <https://developers.openai.com/api/docs/guides/prompt-engineering>
4. <https://privacy.claude.com/en/articles/8114513-business-associate-agreements-baa-for-commercial-customers>
5. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11512878>
6. <https://informatics.bmj.com/content/32/1/e101632>
7. <https://www.tomsguide.com/ai/anthropic-brings-claude-into-healthcare-skip-the-chatgpt-health-waitlist>
8. <https://time.com/6967625/we-need-to-be-ready-for-biotechs-chatgpt-moment>
9. <https://www.benchling.com/state-of-tech>
10. <https://www.mckinsey.com/industries/life-sciences/our-insights/generative-ai-in-the-pharmaceutical-industry-moving-from-hype-to-reality>
11. <https://www.sciencedirect.com/science/article/pii/S1359644624003970>
12. <https://www.axios.com/2024/06/25/ai-biotech-generative-model-protein-design>
13. https://www.zoomrx.com/reports/Endpoints_News_article_17_April_2024.pdf
14. <https://www.livescience.com/health/ai-mirages-mean-tools-used-to-analyze-medical-scans-could-fabricate-their-findings>
15. <https://openai.com/index/new-models-and-developer-products-announced-at-devday>
16. <https://developers.openai.com/api/docs/models/gpt-4-turbo>
17. <https://learn.chatgpt.com/docs/prompting>
18. <https://www.anthropic.com/news/updates-to-our-consumer-terms>
19. <https://www.datastudios.org/post/claude-context-window-token-limits-memory-policy-and-2025-rules>
20. <https://developers.openai.com/api/docs/models/gpt-5>
21. <https://claude.com/resources/articles/1m-context>
22. <https://platform.claude.com/docs/en/api/messages/create>
23. <https://promptden.com/blog/10-examples-of-the-best-prompt-for-chatgpt-in-2025>
24. <https://github.com/ThamJiaHe/claude-prompt-engineering-guide>
25. <https://learn.chatgpt.com/docs/personalize>

26. <https://support.claude.com/en/articles/10185728-understanding-claude-s-personalization-features>
27. <https://developers.openai.com/api/docs/guides/reasoning-best-practices>
28. <https://pubmed.ncbi.nlm.nih.gov/1881878>
29. <https://www.astrazeneca.com/country-sites/taiwan.html>
30. <https://www.msd.com/news/msd-expands-innovative-internal-generative-ai-solutions-helping-to-deliver-medicines-to-patients-faster>
31. <https://investor.lilly.com/node/51001>
32. <https://www.formation.bio/press/introducing-muse>
33. <https://www.axios.com/2023/11/17/generative-ai-dna-biology>
34. <https://www.techradar.com/ai-platforms-assistants/a-guy-who-was-not-an-expert-used-chatgpt-to-make-a-custom-mrna-vaccine-for-his-dogs-cancer-and-other-incredible-theses-from-sam-altman>
35. <https://www.hhs.gov/hipaa/for-professionals/special-topics/health-information-technology/cloud-computing/index.html>
36. <https://www.nber.org/papers/w31161>
37. <https://www.fda.gov/media/189581/download>
38. <https://content.govdelivery.com/accounts/USFDA/bulletins/4045c84>

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore [AI enablement](#), [pharma and biotech software development](#), [data engineering and BI](#), and [Veeva services](#). [Contact IntuitionLabs](#) to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: <https://intuitionlabs.ai>

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.