Jan 15, 2026

Anthropic / Allen Inst Plasmid Design

Invited Alek Kemeny Gabi Martins grace.huynh@alleninstitute.org Jonah Cool Justin Wei mialy.defelice@alleninstitute.org ~~Alec .~~ ~~Ben Lehrburger~~ ~~jesse.gray@alleninstitute.org~~

Attachments Anthropic / Allen Inst Plasmid Design

Meeting records Recording

Summary

Gabi Martins, Mialy DeFelice, someone in 741-Salmonberry (San Francisco, 7) (Alec), Justin Wei, and Jonah Cool met to discuss developing AI tools for DNA and plasmid design, focusing on integrating these tools with Claude. Mialy DeFelice presented on the existing basic tools (SnapGene, Ven, and Benchling) and the challenges encountered with generalist multi-agent tools like Biomnia, which struggled with *de novo* design and sequence hallucination, but also showed success with specialist platforms like Crispr GPT on discrete tasks, achieving 96% accuracy on a full sequence generation. Justin Wei and Alec emphasized the goal of achieving near 100% accuracy by using Claude for tool use and creating evaluation sets to benchmark against Crispr GPT, while Mialy DeFelice's team will focus on vector types and the scoring rubric, and Alec and Justin Wei will pursue agentic orchestration with Claude, starting with basic plasmid design.

Details

- **Introductions and Meeting Context** Gabi Martins opened the meeting, introducing Mialy DeFelice as new to the group and noting their role in organizing the discussion threads. Mialy DeFelice introduced themself as leading the AI applications team at the Allen Institute and a data scientist, having joined in December. Someone in 741-Salmonberry (San Francisco, 7) (Alec) and Justin

Wei from the applied AI team introduced themselves; Alec is the key technical resource for building on Claude, and Justin is supporting them. Jonah Cool, Head of Life Sciences work focused on deployment and partnerships, also introduced themself. The discussion will focus on plasmids, and Alec noted their background is in quantum computing, not life sciences, requiring clear explanations of the topic.

- **Meeting Objectives and Preparations** Justin Wei outlined the goal for the meeting is to get a granular view of the tools currently used for DNA and plasmid design, including examples of data sets and ground truth, to figure out how these tools can interface with Claude. They emphasized the need for an itemization of tasks, flow, and tools, even if tactical. Mialy DeFelice planned to share slides detailing existing work with current plasmid tools. Justin Wei also sent a framework to Gabi Martins and requested the meeting be shortened, if possible, to a half-hour.
- **Initial Prototype and Current Tools for Plasmid Design** Mialy DeFelice presented the initial prototype goal: constructing a basic expression plasmid by combining a gene and a backbone. The current tools used by the team are extremely basic, relying on SnapGene or Ven, with cataloging in Benchling, and involving simple copy-paste for design. A simple usable example is adding GSP to a constitutive promoter.
- **Challenges with Generalist Multi-agent Tools (Biomnia)** Mialy DeFelice detailed attempts to use Biomnia, a multi-agent tool, for plasmid design, noting unsuccessful initial attempts. For example, when prompted to design a GFP expression plasmid, Biomnia attempted a denovo design without necessary sequences, resulting in a non-functional plasmid. A second attempt using a commercially available sequence (PCDNA 3.1) led Biomnia to hallucinate a backbone sequence and ultimately fail to provide an output. In a third attempt, Mialy DeFelice provided exact sequence files, and Biomnia claimed to have created a file, but the output was incorrect upon inspection.
- **Success with Discrete Tasks and Specialist Platforms (Crispr GPT)** Mialy DeFelice found that Biomnia was successful on a very discreet task, which was pulling an exact plasmid sequence by name and outputting the correct sequence file, indicating success on individual tasks when leveraging its internal functions. They also discussed adapting Crispr GPT, a specialist platform originally for generating CRISPR guide RNAs, for plasmid design. Crispr GPT uses pre-designed steps for its agents.

- **Crispr GPT Testing and Validation Rubric** Mialy DeFelice tested Crispr GPT, which uses OpenAI in the background, by supplying both the backbone and insert sequences. Using a step-by-step guidance mode, the tool outputted a full sequence, which Mialy DeFelice validated using Benchling, confirming the plasmid was valid, though off by one nucleotide (96% accuracy). The validation was performed manually by inserting the sequence into the backbone in Benchling and comparing against the Crispr GPT output. Mialy DeFelice also created a rubric for scoring, where all critical failures result in a failure, and a score over 90% is currently targeted.
- **Challenges with Crispr GPT Using Names and Future Accuracy Goals** Mialy DeFelice's second test with Crispr GPT, providing only the plasmid name and details along with the full insert sequence, failed because it couldn't pull the sequence from the catalog number or URL, unlike Biomnia. Mialy DeFelice noted a goal is to achieve near 100% accuracy, as making simple plasmids is easy for humans, and reliability is key for users. They aim to eventually allow users to input simple natural language, such as "create a constitutive expression plasmid for this gene," and have the tool select the best backbone.
- **Improving Tool Performance and Developing Evaluation Sets** Alec emphasized that Claude is proficient at tool use, such as calling APIs or searching databases, which can help improve accuracy compared to GPT. Alec proposed creating evaluation sets and modifying the agent in a loop to achieve "essentially certifiable" or deterministic accuracy, starting with simple basic plasmid design. Mialy DeFelice provided examples of tasks at varying difficulty, starting with fluorescent protein expression vectors, and suggesting using the Addgene catalog as ground truth, possibly mixing and matching components to avoid memorization. Justin Wei suggested focusing on basic plasmid design, ensuring the model is not hallucinating sequences and correctly orienting them.
- **Roles, Responsibilities, and Next Steps** Alec inquired about the scope of work and roles, and Mialy DeFelice explained that their team would focus on continuing to develop vector types and subject matter expertise, including the scoring rubric, while looking to Anthropic for agentic orchestration using Claude. Alec proposed starting with a prototype, defining evaluation criteria, and benchmarking against Crispr GPT. Justin Wei suggested a follow-up conversation with Alec and a show-and-tell in about two weeks to keep the feedback loop tight. They agreed that if simple expression vectors can be reliably generated, they can move on to more difficult questions, such as predicting

transfection and expression efficiency. A Slack channel will be set up for quick communication.

Suggested next steps

- [ ] Someone in 741-Salmonberry (San Francisco, 7) will prototype something useful for the group on the agentic effectiveness of the tools.
- [ ] Gabi Martins will set up a Slack channel for the people on the call to keep the feedback loop tight and ask little questions.
- [ ] Someone in 741-Salmonberry (San Francisco, 7) and Justin Wei will schedule a meeting in about two weeks to share progress on the prototype and allow for a show-and-tell.
- [ ] Someone in 741-Salmonberry (San Francisco, 7) and Justin Wei will talk about creating an evaluation set and try to figure out what a good output looks like for the plasmid design problem, focusing on basic sequence examples to ensure the agent is not hallucinating.