Supplementary Information

Supplementary Note 1: Prompt Engineering and LLM Configurations

To ensure the precision, traceability, and structured reasoning of the RAG-based Q&A model, we implemented a modular prompt engineering strategy. This section details the specific prompt templates used for query decomposition, intent recognition, and response generation.

1.1 Large Language Model Configurations

All experiments were conducted using the **Qwen-turbo** model. To balance logical consistency with biological factual accuracy, the model was deployed with the following inference parameters:

- **Temperature:** 0.3 (adjusted to minimize hallucination while maintaining reasoning flexibility)
- **Max Tokens:** 1500

1.2 Query Decomposition Prompt

The query decomposition module is designed to atomize complex, multi-entity user queries into independent sub-questions. This process ensures that each biological entity is processed with distinct semantic trajectories, facilitating accurate retrieval from the TDRKB.

Prompt Templates for Query Decomposition

Component	Content
System Role	You are an expert question splitter.
Task Description	- Split the input question into multiple independent questions, each representing one atomic action. - Keep all entities intact; do not remove or replace them. - Ensure each question is complete and can be understood independently. - Only split where there is a semantic separation between distinct actions. - Do NOT split if multiple noun phrases are modified by the same verb and canbe understood together.
Few-shot Examples	**Input**: "What are the targets of Amlodipine and Acarbose?" **Output**: ["What are the targets of Amlodipine?", "What are the targets of Acarbose?"]

	Input: "Are there any tuberculosis repositioning studies involving Mefloquine, and what repurposing methods were used?" **Output**: ["Are there any tuberculosis repositioning studies involving Mefloquine?", "What repurposing methods were used for Mefloquine?"]
Output Format	A valid JSON list of strings: ["Question1?", "Question2?", ...]

1.3 Intent Recognition Prompt

The intent recognition module maps the atomized queries onto the specific schema of the TDRKB, assigning a hierarchical information-seeking intent to guide the RAG pipeline toward the appropriate data nodes and corresponding attributes. By defining explicit associations between keywords (e.g., MIC, genomic targets) and database nodes, the model achieves deterministic mapping of user queries.

Prompt Templates for Intent Recognition

Component	**Content**
System Role	You are an expert tuberculosis drug Q&A assistant and a JSON extractor.
Task Description	- Given a user query, extract the intent and sub_intent in JSON format.
Classification Rules	- **Drug Information**: For general pharmacological properties or clinicaldata. - **Target**: If the query mentions targets, genes, or proteins. - **Pathway**: If the query mentions biological pathways or mechanisms. - **Repurposing Experimental Assay**: If the query mentionsexperiments or drug repositioning. - **Drug Susceptibility Test**: If the query contains MIC, reference strains,or sensitivity data. - **Unknown**: If the query does not align with the above categories.
Few-shot Examples	**Input**: "What are the repurposing experiments for Ipragliflozin?" **Output**: {"intent": "Repurposing Experimental Assay"} **Input**: "What are its targets?" **Output**: {"intent": "Target"} **Input**: "What is the reference strain information?" **Output**: {"intent": "Drug Susceptibility Test"}
Output Format	A valid JSON object: {"intent": "String"}

1.4 Response Generation Prompt (Jinja2 Implementation)

To ensure reliability, we implemented structured Jinja2 templates that constrain LLM to act as a "data synthesizer" rather than a creative generator. This prevents

hallucinations by enforcing an assumption that only information present in the retrieved JSON is included.

Prompt Templates for Response Generation

Component	Content
System Role	You are a TB drug Q&A assistant. Generate a concise, scholarly, and professional answer strictly based on the provided structured data
Task Description	- Do NOT include any information that is not explicitly present in the input data. - If a field is missing, explicitly state 'Not available in the database'."

Full Template Implementation (Jinja2)

The core for the prompt templates is as follows:

```
{% if intent == "drug sensitivity experiment" %}
{% for item in structured_data %}
Entity: {{ item.entity_name }}

{% if item.data and item.data|length > 0 %}
{% for exp in item.data %}
- Reference Strain: {{ exp.reference_strain|default('N/A') }};
- MIC Value: {{ exp.mic_value|default('N/A') }};
- Species: {{ exp.species|default('N/A') }};
- Test Strain ID: {{ exp.test_strain_id|default('N/A') }};
- Test Strain Type: {{ exp.test_strain_type|default('N/A') }};
- Evidence Refs: {{ exp.evidence_refs|default('N/A') }};
- References:
{% for ref in exp.references %}
  - [{{ ref.Ref_ID }}] "{{ ref.Title }}" (PMID: {{ ref.PMID }}). {{ ref.url }}
{% endfor %}
{% endfor %}
{% else %}
No experimental data found in the database.
{% endif %}

{% endfor %}


{% elif intent == "Repurposing Experiment" %}
{% for item in structured_data %}
Entity: {{ item.entity_name|default("Unknown") }}

{% if item.data and item.data|length > 0 %}
Please summarize repurposing experimental results for this entity:
{% for exp in item.data %}
- Experiment Type: {{ exp.experiment_type|default('N/A') }};
- Repurposing Methods: {{ exp.repurposing_methods|default('N/A') }};
- Experiment Effects: {{ exp.experiment_effect|default('N/A') }};
- Probable Mechanisms: {{ exp.probable_mechanisms|default('N/A') }};
- Other Details: {% for k, v in exp.items() %}{% if k not in
["experiment_type","repurposing_methods","experiment_effect","probable_mechanisms"] %}
{{ k|capitalize }}: {{ v|default('N/A') }}; {% endif %}{% endfor %}
{% endfor %}
{% else %}
No repurposing experimental data available for this entity.
{% endif %}
{% endfor %}

{% elif intent == "Target" %}
Information: {{ content_text }}

Please summarize known targets, gene names, Rv_id and their biological roles strictly
based on the provided data.
```

```
{% elif intent == "pathway" %}
{% for item in structured_data %}
Entity: {{ item.entity_name }}

{% if item.data and item.data|length > 0 %}
{% for p in item.data %}
- Pathway ID: {{ p.Pathway_ID|default('N/A') }};
- Name: {{ p.pathway_name|default('N/A') }};
- Class: {{ p.pathway_class|default('N/A') }};
- Description: {{ p.description|default('N/A') }};
- Gene List: {{ p.gene_list|default('N/A') }};
- Gene Count: {{ p.gene_count|default('N/A') }};
- Rv_id: {{ p.Rv_id|default('N/A') }};
- Product: {{ p.Product|default('N/A') }};
- Gene_name: {{ p.Gene_name|default('N/A') }};
{% endfor %}
{% else %}
No pathway data available for this entity.
{% endif %}

{% if item.data and item.data|length > 0 %}
Please summarize pathway information for this entity:
{% for p in item.data %}
- Pathway ID: {{ p.Pathway_ID|default('N/A') }};
- Name: {{ p.pathway_name|default('N/A') }};
- Class: {{ p.pathway_class|default('N/A') }};
- Description: {{ p.description|default('N/A') }};
- Gene List: {{ p.gene_list|default('N/A') }};
- Gene Count: {{ p.gene_count|default('N/A') }};
- Rv_id: {{ p.Rv_id|default('N/A') }};
- Product: {{ p.Product|default('N/A') }};
- Gene_name: {{ p.Gene_name|default('N/A') }};
{% endfor %}
{% else %}
No pathway data available for this entity.
{% endif %}
{% endfor %}

{% else %}
Information: {{ content_text }}
Please summarize content strictly based on the provided data.
{% endif %}
```

Supplementary Note 2: Supplementary Tables

Table S1. A characterization of the Single-Q test set, including coverage, query type, and complexity.

No.	Question	Coverage	Type	Complexity
1	What are the targets of Paromomycin?	Targets	Single-hop	Low
2	Can you list the known gene targets of Triflusal?	Targets	Single-hop	Low
3*	What type of drug is Isoniazid classified as?	Drugs	Single-hop	Low
4	Has Mefloquine been approved yet?	Drugs	Single-hop	Medium
5	What is Flupirtine primarily indicated for?	Drugs	Single-hop	Medium
6	What are the biological pathways associated with nitrofural?	Pathways	Multi-hop	High
7	Are there any repositioning studies for quinacrine in TB treatment?	Repurposing experimental assays	Single-hop	Low
8	What effects were observed in the repurposing assays of Paromomycin?	Repurposing experimental assays	Single-hop	Medium
9	Can you list the repurposing methods applied to Mefloquine?	Repurposing experimental assays	Single-hop	Medium
10	What is the MIC value of Carprofen in drug sensitivity experiments?	Drug susceptibility tests	Single-hop	Medium
11	What test strain types are associated with the artesunate DST results?	Drug susceptibility tests	Single-hop	Medium
12	Which reference strain was used in the idarubicin sensitivity tests?	Drug susceptibility tests	Single-hop	Medium
13*	What are the targets of Xylomycin?	Targets	Single-hop	Low
14	Can you list the UniProt ID of the Isepamicin target?	Targets	Single-hop	Medium

15	What is the pathway class of the KEGG pathway mapped from pknB?	Pathways	Multi-hop	High
16	Can you provide references linked to the Mefloquine repurposing assay?	References	Single-hop	Low
17	Provide references reporting the gene targets of diacerein	References	Single-hop	High
18	Which biological pathways are experimentally associated with Triflusal in the treatment of tuberculosis?	Pathways	Multi-hop	High
19*	Can Primaquine treat malaria?	Drugs	Single-hop	Low
20*	What is your work today?	None	Single-hop	Low

* represents negative samples, which correspond to queries for which the knowledge base contains no answer.

Table S2. Detailed performance of entity parsing, main-intent, and sub-intent recognition within the Single-Q suite.

No.	Questions	Entity	Main-intent	Sub-intent
1	What are the targets of Paromomycin?	✓	✓	✓
2	Can you list the known gene targets of Triflusal?	✓	✓	✓
3*	What type of drug is Isoniazid classified as?	✓	✓	×
4	Has Mefloquine been approved yet?	✓	✓	×
5	What is Flupirtine primarily indicated for?	✓	✓	✓
6	What are the biological pathways associated with nitrofural?	✓	✓	✓
7	Are there any repositioning studies for quinacrine in TB treatment?	✓	✓	✓
8	What effects were observed in the repurposing assays of Paromomycin?	✓	✓	✓
9	Can you list the repurposing methods applied to Mefloquine?	✓	✓	✓
10	What is the MIC value of Carprofen in drug sensitivity experiments?	✓	✓	✓
11	What test strain types are associated with the artesunate DST results?	✓	✓	×
12	Which reference strain was used in the idarubicin sensitivity tests?	✓	✓	✓
13*	What are the targets of Xylomycin?	✓	✓	✓
14	Can you list the UniProt ID of the Isepamicin target?	✓	✓	×
15	What is the pathway class of the KEGG pathway mapped from pknB?	×	✓	✓
16	Can you provide references linked to the Mefloquine repurposing assay?	✓	✓	✓
17	Provide references reporting the gene targets of diacerein	✓	✓	×
18	Which biological pathways are experimentally associated with Triflusal in the treatment of tuberculosis?	✓	✓	✓
19*	Can Primaquine treat malaria?	✓	✓	×
20*	What is your work today?	×	✓	×

* represents negative samples, which correspond to queries for which the knowledge base contains no answer.

Table S3. Detailed performance of entity parsing, main-intent, and sub-intent recognition within the Contextual-Q test suite.

Session	No.	Questions	Entity	Main-intent	Sub-intent
Session 1	1	What are the known targets of dicoumarol?	✓	✓	✓
	2	For each of these targets, which biological pathways are they involved in?	✓	✓	✓
	3	Can you summarize these pathway functions?	✓	✓	×
Session 2	4	What is the MIC value of disulfiram?	✓	✓	✓
	5	Which type of test strain was used in this drug sensitivity test?	✓	✓	×
	6	Were multiple strains tested, or just a single strain?	✓	✓	✓
Session 3	7	What type of drug is tafenoquine?	✓	✓	×
	8	Has tafenoquine been tested in any repurposing experiments for tuberculosis?	✓	✓	✓
	9	What were the observed effects or outcomes from these repurposing experiments?	✓	✓	✓
	10	Can you provide the references or publications reporting these repurposing studies?	✓	✓	✓

Table S4. Detailed performance of entity parsing, main-intent, and sub-intent recognition within the Batch-Q suite.

No.	Questions	Atomic Questions	Entity	Main-intent	Sub-intent
1	What are the targets of Amlodipine and Acarbose?	What are the targets of Amlodipine?	✓	✓	✓
		What are the targets of Acarbose?	✓	✓	✓
2	What is the MIC value of ebselen, and which reference strain was used in the experiment?	What is the MIC value of ebselen?	✓	✓	✓
		Which reference strain was used in the experiment?	✓	✓	✓
3	What is the approved indication of Glimepiride, and is this drug currently approved or still in development?	What is the approved indication of Glimepiride?	✓	✓	✓
		Is this drug currently approved or still in development?	✓	✓	×
4	Are there any tuberculosis repositioning studies involving bromfenac, and what repurposing methods were used?	Are there any tuberculosis repositioning studies involving bromfenac?	✓	✓	✓
		What repurposing methods were used for bromfenac?	✓	✓	×
5	For the target of loperamide, which pathway is it involved in, and is there any supporting reference?	For the target of loperamide, which pathway is it involved in?	✓	✓	✓
		Is there any supporting reference?	✓	✓	×

Table S5. The responses of Single-Q from the Q&A model.

No.	Questions	Answers
1	What are the targets of Paromomycin?	Paromomycin targets the DNA polymerase III subunit beta, encoded by the gene *dnaN* (Rv0002). This protein plays a critical role in DNA replication, providing proofreading activity and acting as a sliding clamp to enhance processivity. The target is identified as T10, with the UniProt ID P9WNU1. The relevant study is cited as Ref_ID R9 (PMID: 39173789).
2	Can you list the known gene targets of Triflusal?	Triflusal is known to target the gene **DciA** (Rv0004), which encodes a conserved hypothetical protein. The gene is associated with the functional category of "Conserved hypotheticals," and its biological role remains unknown. This information is supported by the reference: "Unravelling the potential of Triflusal as an anti-TB repurposed drug by targeting replication protein DciA" (PMID: 38145750).
3*	What type of drug is Isoniazid classified as?	Not available in the database.
4	Has Mefloquine been approved yet?	Yes, Mefloquine has been approved. It is an antiparasitic medication indicated for the treatment and prevention of malaria. Its brand name is Lariam, and it is assigned the ATC codes P01BF02 and P01BC02.
5	What is Flupirtine primarily indicated for?	Flupirtine is primarily indicated as a nonopioid analgesic and has been in phase II trials for the treatment of fibromyalgia.
6	What are the biological pathways associated with nitrofural?	Nitrofural is associated with the following biological pathways in *Mycobacterium tuberculosis H37Rv*: 1. **Biosynthesis of secondary metabolites** (Pathway ID: P15) - Genes involved: *Idi, Rv2245, KasA, InhA, Rv1484, Rv1745c*. 2. **Mycolic acid biosynthesis** (Pathway ID: P8) - Genes involved: *Rv1484, KasA, InhA, Rv2245*. 3. **Metabolic pathways** (Pathway ID: P10) - Genes involved: *canB, Rv2217, Idi, Tmk, Rv2245, KasA, Rv3247c, GlmU, InhA, Rv1284, Rv3588c, Rv1484, canA, Rv1018c, LipB, Rv1745c*. 4. **Biosynthesis of cofactors** (Pathway ID: P17) - Genes involved: *Rv2217, LipB*. 5. **Lipoic acid metabolism** (Pathway ID: P7) - Genes involved: *Rv2217, LipB*. 6. **Nitrogen metabolism** (Pathway ID: P11) - Genes involved: *canB, Rv3588c, canA, Rv1284*. 7. **Terpenoid backbone biosynthesis** (Pathway ID: P19) - Genes involved: *Idi, Rv1745c*. 8. **Biosynthesis of nucleotide sugars** (Pathway ID: P29) -

		Genes involved: *Rv1018c, GlmU*. 9. **Biosynthesis of various nucleotide sugars** (Pathway ID: P25) - Genes involved: *Rv1018c, GlmU*. 10. **Amino sugar and nucleotide sugar metabolism** (Pathway ID: P21) - Genes involved: *Rv1018c, GlmU*. 11. **Pyrimidine metabolism** (Pathway ID: P20) - Genes involved: *Tmk, Rv3247c*. 12. **Nucleotide metabolism** (Pathway ID: P27) - Genes involved: *Tmk, Rv3247c*.
7	Are there any repositioning studies for quinacrine in TB treatment?	Yes, there are repositioning studies for quinacrine in TB treatment. Quinacrine has been identified as a promising candidate compound for TB through in silico and in vitro experiments. These studies utilized phenotypic screening, machine learning-based screening, molecular docking, and experimental validations. The probable mechanisms involve targeting receptor tyrosine kinases (RTKs) and controlling Mycobacterium tuberculosis (Mtb) intracellular survival. This research falls under host-directed therapy (HDT). The findings are supported by evidence from a study published in PubMed (PMID: 29367740).
8	What effects were observed in the repurposing assays of Paromomycin?	In the repurposing assays of Paromomycin, it was observed that the drug effectively limits the growth and replication of M.tb. This makes Paromomycin a promising candidate for repurposing as an anti-TB therapeutic. The probable mechanism involves targeting the protein DnaN, with the effect attributed to the inhibition of DNA replication. These findings were supported by in silico and in vitro experiments, including molecular docking and experimental validations. The study is documented in the reference titled "Revitalizing antimicrobial strategies: paromomycin and dicoumarol repurposed as potent inhibitors of M.tb's replication machinery via targeting the vital protein DnaN" (PMID: 39173789).
9	Can you list the repurposing methods applied to Mefloquine?	The repurposing methods applied to Mefloquine include "Molecular docking," "Phenotypic screening," and "Experimental validation," as identified through in silico, in vitro, and in vivo experiments. These methods were used to investigate its potential against Mycobacterium tuberculosis by targeting specific mechanisms such as FadD32 and Eis.
10	What is the MIC value of Carprofen in drug sensitivity experiments?	The MIC value of Carprofen in drug sensitivity experiments is 250 mg/L, as determined using the reference strain mc2155 of *M. smegmatis*. This information is documented in the study titled "Carprofen elicits pleiotropic mechanisms of bactericidal action with the potential to reverse antimicrobial drug resistance in tuberculosis" (PMID: 32790867).

11	What test strain types are associated with the artesunate DST results?	The test strain type associated with the artesunate DST results is the reference strain "H37Rv" of *M. tuberculosis*.
12	Which reference strain was used in the idarubicin sensitivity tests?	The reference strain used in the idarubicin sensitivity tests was *M. smegmatis* strain mc2155.
13*	What are the targets of Xylomycin?	The provided data does not contain information about the targets, gene names, Rv_ids, or biological roles of Xylomycin. Therefore, no summary can be generated based on the available data.
14	Can you list the UniProt ID of the Isepamicin target?	The UniProt ID of the Isepamicin target is **P9WG29**. The target is identified as **Resuscitation-promoting factor B (RpfB)**, with the gene name **RpfB**, NCBI Gene ID **886048**, and Rv_id **Rv1009**. The biological role of this target is associated with the suppression of Mycobacterium tuberculosis dormancy reactivation, as reported in the study titled "Repurposing of antibacterial compounds for suppression of Mycobacterium tuberculosis dormancy reactivation by targeting resuscitation-promoting factors B" (PMID: 37551014).
15	What is the pathway class of the KEGG pathway mapped from pknB?	The provided data does not contain information to answer the question about the pathway class of the KEGG pathway mapped from pknB.
16	Can you provide references linked to the Mefloquine repurposing assay?	The Mefloquine repurposing assay is linked to the following references: 1. **R14**: "In Silico Drug Repurposing Approach: Investigation of Mycobacterium tuberculosis FadD32 Targeted by FDA-Approved Drugs" (PMID: 35163931, [Link](https://pubmed.ncbi.nlm.nih.gov/35163931/)). 2. **R2**: "Mefloquine and its oxazolidine derivative compound are active against drug-resistant Mycobacterium tuberculosis strains and in a murine model of tuberculosis infection" (PMID: 27364701, [Link](https://pubmed.ncbi.nlm.nih.gov/27364701/)). 3. **R1**: "Discovery and Mechanistic Analysis of Structurally Diverse Inhibitors of Acetyltransferase Eis among FDA-Approved Drugs" (PMID: 36657084, [Link](https://pubmed.ncbi.nlm.nih.gov/36657084/)). These references cover in silico, in vitro, and in vivo studies related to Mefloquine's potential as a repurposed drug for tuberculosis treatment.

17	Provide references reporting the gene targets of diacerein	The provided data does not include any information regarding the gene targets of diacerein. Therefore, no summary of known targets, gene names, Rv_id, or their biological roles can be generated based on the available data.
18	Which biological pathways are experimentally associated with Triflusal in the treatment of tuberculosis?	The provided data does not include information about the biological pathways experimentally associated with Triflusal in the treatment of tuberculosis. Therefore, no specific pathways can be identified based on the available data.
19*	Can Primaquine treat malaria?	Primaquine is a medication used in the treatment of malaria. However, the provided data does not include specific details about its use, mechanism, or efficacy. For comprehensive information, consult a medical reference or healthcare professional.
20*	What is your work today?	Not available in the database.

* represents negative samples, which correspond to queries for which the knowledge base contains no answer.

Table S6. The responses of Contextual-Q from the Q&A model.

No.	Questions	Answers
Session 1		
1	What are the known targets of dicoumarol?	Dicoumarol is known to target the DNA polymerase III subunit beta, encoded by the gene *DnaN* (Rv0002). This protein is involved in DNA replication with proofreading activity and acts as a sliding clamp for processivity. The targeting of DnaN by dicoumarol has been identified as a potential mechanism for its antimicrobial activity against M.tb. The evidence for this interaction is supported by the reference: "Revitalizing antimicrobial strategies: paromomycin and dicoumarol repurposed as potent inhibitors of M.tb's replication machinery via targeting the vital protein DnaN" (PMID: 39173789).
2	For each of these targets, which biological pathways are they involved in?	Dicoumarol is involved in the following biological pathways: 1. **DNA replication - Mycobacterium tuberculosis H37Rv** (Pathway ID: P3) - Pathway class: Genetic Information Processing; Replication and repair - Genes involved: Rv0002, DnaN 2. **Homologous recombination - Mycobacterium tuberculosis H37Rv** (Pathway ID: P14) - Pathway class: Genetic Information Processing; Replication and repair - Genes involved: Rv0002, DnaN 3. **Mismatch repair - Mycobacterium tuberculosis H37Rv** (Pathway ID: P9) - Pathway class: Genetic Information Processing; Replication and repair - Genes involved: Rv0002, DnaN These pathways are all categorized under "Genetic Information Processing; Replication and repair."
3	Can you summarize these pathway functions?	The pathway functions provided are related to "Genetic Information Processing; Replication and repair" and involve the genes Rv0002 and DnaN. These pathways include DNA replication, homologous recombination, and mismatch repair in Mycobacterium tuberculosis H37Rv. Each pathway plays a critical role in maintaining genomic integrity during DNA replication and repair processes.
Session 2		
4	What is the MIC value of disulfiram?	The MIC value of disulfiram against *M. tuberculosis* is 1.56 μg/mL for the reference strains H37Rv and H37Ra. Additionally, for 20 types of clinical isolates, the MIC range was observed to be 0.78–1.56 μg/mL. These values are reported

		in a study referenced as R105 (PMID: 22615274).
5	Which type of test strain was used in this drug sensitivity test?	The drug sensitivity test used three types of test strains: the reference strain H37Rv, the reference strain H37Ra, and 20 types of clinical isolated strains.
6	Were multiple strains tested, or just a single strain?	The data indicates that multiple strains were tested, specifically H37Rv and H37Ra. The entry "/ " suggests that no additional strains were tested beyond these two.
Session 3		
7	What type of drug is tafenoquine?	Tafenoquine is an antiparasitic drug used to treat malaria. Its ATC code is P01BA07, and it is also known by the brand names Arakoda and Krintafel.
8	Has tafenoquine been tested in any repurposing experiments for tuberculosis?	Yes, tafenoquine has been tested in repurposing experiments for tuberculosis. It was identified as a potential new anti-mycobacterial agent, particularly effective against resistant and/or dormant forms of TB, through in vitro experiments. The repurposing methods used included phenotypic screening and knowledge-based approaches. This information is supported by the study titled "Effects of tafenoquine against active, dormant and resistant Mycobacterium tuberculosis" (PMID: 34004588).
9	What were the observed effects or outcomes from these repurposing experiments?	The repurposing experiments involving tafenoquine observed that it represents a new anti-mycobacterial agent, particularly effective against resistant and/or dormant forms of TB. The study utilized phenotypic screening and knowledge-based methods, with evidence supporting its efficacy documented in the reference "Effects of tafenoquine against active, dormant and resistant Mycobacterium tuberculosis" (PMID: 34004588).
10	Can you provide the references or publications reporting these repurposing studies?	The repurposing studies on tafenoquine are reported in the publication titled "Effects of tafenoquine against active, dormant and resistant Mycobacterium tuberculosis," with the PMID 34004588. The reference can be accessed at the following link: https://pubmed.ncbi.nlm.nih.gov/34004588/.

Table S7. The responses of Batch-Q from the Q&A model.

No.	Questions	Answers
1	What are the targets of Amlodipine and Acarbose?	Amlodipine's target is a probable conserved integral membrane transport protein for multidrug efflux, designated as "Tap" (gene name: Tap; Rv_id: Rv1258c). It is a putative transporter involved in the export of an unidentified substrate, possibly a macrolide, and is categorized under "Cell wall and cell processes." Acarbose has four targets: 1. **PpiB** (gene name: PpiB; Rv_id: Rv2582): A peptidyl-prolyl cis-trans isomerase involved in accelerating protein folding, classified under "Information pathways." 2. **PknB** (gene name: PknB; Rv_id: Rv0014c): A transmembrane serine/threonine-protein kinase involved in signal transduction and regulation of cell division/differentiation, categorized under "Regulatory proteins." 3. **MurE** (gene name: MurE; Rv_id: Rv2158c): Involved in cell wall synthesis. 4. **LipU** (gene name: LipU; Rv_id: Rv1076): A lipase, with no specific function provided.
2	What is the MIC value of ebselen, and which reference strain was used in the experiment?	The MIC value of ebselen is 12.5 μg/mL, as determined using the reference strain H37Rv of *M. tuberculosis*. The experiment is described in the study titled "Identifying and Tackling Emergent Vulnerability in Drug-Resistant Mycobacteria" (PMID: 27759382).
3	What is the approved indication of Glimepiride, and is this drug currently approved or still in development?	Glimepiride is approved for the treatment of type 2 diabetes mellitus. It is currently approved.
4	Are there any tuberculosis repositioning studies involving bromfenac, and what repurposing methods were used?	Yes, there are tuberculosis repositioning studies involving bromfenac. These studies utilized repurposing methods such as molecular docking and experimental validations. The research demonstrated that bromfenac showed antitubercular activity levels by targeting Zmp1 and PDF. The findings are detailed in the study titled "A Repurposing Approach for Uncovering the Anti-Tubercular Activity of FDA-Approved Drugs with Potential Multi-Targeting Profiles" (PMID: 31795400).
5	For the target of	Loperamide is involved in the "Mycolic acid

	loperamide, which pathway is it involved in, and is there any supporting reference?	biosynthesis - Mycobacterium tuberculosis H37Rv" pathway, which falls under the category of "Metabolism; Lipid metabolism." Additionally, it is associated with "Metabolic pathways - Mycobacterium tuberculosis H37Rv" and "Biosynthesis of secondary metabolites - Mycobacterium tuberculosis H37Rv." The genes involved in these pathways include "FadD32/FAAL32" and "Rv3801c." However, no specific supporting reference is provided in the data.

Table S8. Performance evaluation of model responses on the Single-Q test set.

No.	Questions	Factual Correctness	KB Alignment	Completeness & Formatting	Final Score
1	What are the targets of Paromomycin?	1.0	1.0	0.5	0.83
2	Can you list the known gene targets of Triflusal?	1.0	1.0	0.5	0.83
3*	What type of drug is Isoniazid classified as?	1.0	1.0	1.0	1.0
4	Has Mefloquine been approved yet?	1.0	1.0	1.0	1.0
5	What is Flupirtine primarily indicated for?	1.0	1.0	1.0	1.0
6	What are the biological pathways associated with nitrofural?	1.0	1.0	1.0	1.0
7	Are there any repositioning studies for quinacrine in TB treatment?	1.0	1.0	0.5	0.83
8	What effects were observed in the repurposing assays of Paromomycin?	1.0	1.0	0.5	0.83
9	Can you list the repurposing methods applied to Mefloquine?	0.5	1.0	0.5	0.67
10	What is the MIC value of Carprofen in drug sensitivity experiments?	1.0	1.0	1.0	1.0
11	What test strain types are associated with the artesunate DST results?	0.5	1.0	1.0	0.83
12	Which reference strain was used in the idarubicin sensitivity tests?	1.0	1.0	1.0	1.0
13*	What are the targets of Xylomycin?	1.0	1.0	0.5	0.83
14	Can you list the UniProt ID of the Isepamicin target?	1.0	1.0	0.5	0.83
15	What is the pathway class of the KEGG pathway mapped from pknB?	0.0	0.0	0.0	0

16	Can you provide references linked to the Mefloquine repurposing assay?	1.0	1.0	1.0	1.0
17	Provide references reporting the gene targets of diacerein	0.0	0.0	0.0	0
18	Which biological pathways are experimentally associated with Triflusal in the treatment of tuberculosis?	0.0	0.0	0.0	0
19*	Can Primaquine treat malaria?	0.5	1.0	0.5	0.83
20*	What is your work today?	1.0	1.0	1.0	1.0

* represents negative samples, which correspond to queries for which the knowledge base contains no answer.

Table S9. Performance evaluation of model responses on the Contextual-Q test set

Sessions	No.	Questions	Factual Correctness	KB Alignment	Completeness & Formatting	Final Score
Session 1	1	What are the known targets of dicoumarol?	1.0	1.0	0.5	0.83
	2	For each of these targets, which biological pathways are they involved in?	1.0	1.0	1.0	1.0
	3	Can you summarize these pathway functions?	0.5	1.0	0.5	0.67
Session 2	4	What is the MIC value of disulfiram?	1.0	1.0	0.5	0.83
	5	Which type of test strain was used in this drug sensitivity test?	1.0	1.0	1.0	1.0
	6	Were multiple strains tested, or just a single strain?	0.5	0.5	1.0	0.67
Session 3	7	What type of drug is tafenoquine?	1.0	1.0	1.0	1.0
	8	Has tafenoquine been tested in any repurposing experiments for tuberculosis?	1.0	1.0	0.5	0.83
	9	What were the observed effects or outcomes from these repurposing experiments?	1.0	1.0	0.5	0.83
	10	Can you provide the references or publications reporting these repurposing studies?	1.0	1.0	1.0	1.0

Table S10. Performance evaluation of model responses on the Batch-Q test set.

No.	Questions	Factual Correctness	KB Alignment	Completeness & Formatting	Final Score
1	What are the targets of Amlodipine and Acarbose?	1.0	1.0	1.0	1.0
2	What is the MIC value of ebselen, and which reference strain was used in the experiment?	1.0	1.0	0.5	0.83
3	What is the approved indication of Glimepiride, and is this drug currently approved or still in development?	1.0	1.0	1.0	1.0
4	Are there any tuberculosis repositioning studies involving bromfenac, and what repurposing methods were used?	1.0	1.0	0.5	0.83
5	For the target of loperamide, which pathway is it involved in, and is there any supporting reference?	0.5	0.5	0.5	0.5