

Supplementary Note S3: Verbatim LLM Extraction Outputs and Qualitative Behavior Comparison

This supplementary note provides full, verbatim extraction outputs from all evaluated literature-mining platforms in response to an identical prompt applied to a single benchmark paper. The purpose of this note is to transparently substantiate the qualitative model behaviors discussed in the main text and illustrated schematically in Fig. 3a, allowing readers to directly inspect raw outputs and independently assess differences in extraction style, scope, and grounding.

Prompt: “List all perovskite compounds mentioned in the Experimental or Results sections.”

Paper: Paper 9 (Nano-Micro Letters, 2021)

Run conditions:

- Single extraction run
- Identical prompt
- No follow-up queries
- Raw outputs reproduced verbatim

ChatGPT4o Exact Conversation:

You said:

List all perovskite compounds mentioned in the Experimental or Results sections.

ChatGPT said:

The perovskite compound explicitly mentioned in the Experimental and Results sections of the uploaded paper is:

Perovskite Compound

- $(\text{BA})_2(\text{MA})_4\text{Pb}_5\text{I}_{16}$
 - Type: Two-dimensional Ruddlesden–Popper perovskite ($n = 5$)

Materials Used in Synthesis and Their IUPAC Names

Component	IUPAC Name
BAI	Butylammonium iodide (1-butylaminium iodide)
MAI	Methylammonium iodide (methanaminium iodide)
PbI ₂	Lead(II) iodide
NH ₄ SCN	Ammonium thiocyanate
Ti ₃ C ₂ T _x	Titanium carbide MXene with surface terminations (–O, –OH, and/or –F), obtained by

Component	IUPAC Name
MXene	etching Ti_3AlC_2 using LiF and HCl

Synthesis Details

- **Solvents:** 1:1 mixture of Dimethyl sulfoxide (DMSO, IUPAC: dimethyl sulfoxide) and N,N-Dimethylformamide (DMF, IUPAC: N,N-dimethylmethanamide)
- **Stoichiometry:** BAI : MAI : PbI_2 : NH_4SCN = 2 : 4 : 5 : 2
- **Ti_3C_2Tx concentrations tested:** 0.1 mM, 0.3 mM, 0.5 mM, 0.7 mM
- **Pb^{2+} concentration in precursor:** 1.2 M
- **Precursor stirring:** 8 hours at 50 °C
- **Spin coating:** 4000 rpm for 30 seconds
- **Antisolvent dripping:** 300 μ L toluene during spin coating
- **Annealing:** 100 °C for 10 minutes in DMF atmosphere

If you need a table of photovoltaic performance metrics or synthesis outputs for each Ti_3C_2Tx doping level, let me know.

Observed behavior (ChatGPT)

ChatGPT organizes extracted information into clearly labeled categorical sections and tabular formats, including compound identification, precursor inventories, and synthesis conditions. In addition to listing the perovskite compound explicitly mentioned in the Experimental and Results sections, the model provides ontological classification (e.g., perovskite type and dimensionality) and aggregates detailed synthesis parameters such as stoichiometry, processing conditions, and concentration ranges. While these additions are factually consistent with the source text, they extend beyond the strict scope of the prompt, illustrating a tendency toward comprehensive, assistant-style responses rather than minimal extraction.

Perplexity Exact Conversation:

You said:

List all perovskite compounds mentioned in the Experimental or Results sections

Perplexity said:

Here is a list of all perovskite compounds mentioned in the **Experimental** or **Results** sections of the provided article:

Perovskite Compounds Mentioned

1. $(\text{BA})_2(\text{MA})_4\text{Pb}_5\text{I}_{16}$

- **Section:** Experimental (2.3 Preparation of 2D Perovskite Precursor Solution), Results and Discussion
- **Details:** This is the main 2D Ruddlesden–Popper (RP) perovskite used as the absorber layer in the study. The precursor solution is prepared by mixing BAI (butylammonium iodide), MAI (methylammonium iodide), PbI_2 , and NH_4SCN in DMSO/DMF.

2. $\text{CH}_3\text{NH}_3\text{PbI}_3$ (also written as MAPbI_3)

- **Section:** Results and Discussion (Background, referencing previous work)
- **Details:** Mentioned as a 3D perovskite in the context of previous studies where $\text{Ti}_3\text{C}_2\text{T}_x$ MXene was used as a dopant to enhance crystal size and charge transfer.

Summary Table

Compound Name	Formula	Context/Section	Notes
Butylammonium-methylammonium lead iodide	$(\text{BA})_2(\text{MA})_4\text{Pb}_5\text{I}_{16}$	Experimental, Results	Main 2D RP perovskite studied
Methylammonium lead iodide	$\text{CH}_3\text{NH}_3\text{PbI}_3$ (MAPbI_3)	Results (background)	3D perovskite, mentioned for comparison

Notes

- $(\text{BA})_2(\text{MA})_4\text{Pb}_5\text{I}_{16}$ is the only perovskite compound actually synthesized, processed, and characterized in the experimental work of this paper.
- $\text{CH}_3\text{NH}_3\text{PbI}_3$ is only referenced in the context of prior studies and is not part of the current experimental procedures.

If you need more details (e.g., full chemical formulas, roles in device architecture, or references to figures/tables), let me know!

Observed behavior (Perplexity)

Perplexity produces narrative, explanatory responses that contextualize extracted entities with background descriptions, section references, and repeated summaries. In addition to listing the perovskite compounds mentioned in the Experimental and Results sections, the model elaborates on their roles, literature context, and prior usage, and reformulates the same information into multiple descriptive formats (bullet lists and summary tables). While the extracted compounds are correctly identified, the output exhibits high verbosity and redundancy relative to the prompt, illustrating an explanatory retrieval style rather than concise information extraction.

DeepSeek Exact Conversation:

You said:

List all perovskite compounds mentioned in the Experimental or Results sections.

DeepSeek said:

The perovskite compounds mentioned in the Experimental or Results sections are:

1. **2D Ruddlesden-Popper (RP) perovskite** with the general chemical formula $(A')_2A_n-1MnX_{3n+1}$, where:
 - A' is a bulky long-chain organic spacer.
 - A is a monovalent cation.
 - M is a divalent metal cation (e.g., Pb^{2+}).
 - X is a halide anion (e.g., I^-).
2. **Specific 2D perovskite composition:**
 - $(BA)_2(MA)_4Pb_5I_{16}$ (where BA is butylammonium and MA is methylammonium).
3. **3D perovskite:**
 - $CH_3NH_3PbI_3$ (mentioned in the context of comparison with 2D perovskites).
4. **Other related materials:**
 - $MAPbI_3$ (mentioned in the stability comparison section).

These compounds are discussed in the context of their structural properties, stability, and performance in solar cells.

Observed behavior (DeepSeek)

DeepSeek produces compact, enumerated outputs that prioritize concise representation of chemical information. In addition to listing explicitly named perovskite compositions, the model introduces generalized structural formulas and normalized chemical representations. While these generalizations are

chemically reasonable and common in the literature, they are not stated verbatim in the source text. Referenced materials discussed for comparison are included without explicit distinction from synthesized compounds. This behavior illustrates a distilled extraction style that favors abstraction and compression, which can increase information density but may introduce inferred or generalized content beyond strict textual grounding.

NotebookLM Exact Conversation:

You said:

List all perovskite compounds mentioned in the Experimental or Results sections.

NotebookLM:

Based on the information found in the Experimental Section and Results and Discussion sections of the provided source, the perovskite compounds and phases mentioned are:

- **2D perovskite $\text{BA}_2\text{MA}_4\text{Pb}_5\text{I}_{16}$ ($n = 5$)** This is the specific 2D Ruddlesden-Popper perovskite absorber used in the study, fabricated by mixing BAI, MAI, PbI_2 , and NH_4SCN in solvents1....
- Perovskite phases with $n = 2, 3, 4,$ and ∞ These phases are noted as coexisting within the 2D perovskite films, identified by multiple absorption peaks in the UV/Vis spectra4.... Separate peaks from large- n phases were observed upon front excitation, while phases with different small- n values were observed upon back excitation6. An ordered distribution of n values is illustrated7....
- **MAPbI_3** This compound is shown in Fig. 6 within the Results section for comparison regarding device stability9.

Observed behavior (NotebookLM)

NotebookLM produces conservative, source-grounded extractions that explicitly reference the sections and figures from which information is drawn. The model clearly distinguishes the synthesized 2D perovskite absorber from coexisting phases discussed in the Results and from reference compounds included only for comparison. Inclusion of additional n -phase components is justified by explicit spectroscopic evidence reported in the source text, and no generalized formulas or inferred synthesis conditions are introduced. This behavior reflects a retrieval-grounded extraction strategy that prioritizes traceability and minimizes unsupported assumptions.

Note:

The numbers highlighted in red are from NotebookLM's RAG function, on their model it references exactly where in the uploaded PDFs with an interactable button that is the numbers. They are in the output as the response were directly copied from the chat.

Elicit Exact Conversation:

You said:

List all perovskite compounds mentioned in the Experimental or Results sections.

Elicit said:

Here are the perovskite compounds mentioned in the Experimental or Results sections of the paper:

(BA)₂(MA)₄Pb₅I₁₆

This is the main 2D Ruddlesden-Popper perovskite compound studied in the paper. It is mentioned multiple times, including:

"In this study, we for the first time fabricate PSCs employing the 2D (BA)₂(MA)₄Pb₅I₁₆ RP perovskite absorber in which Ti₃C₂T_x nanosheets were added as a nano-dopant." 1

"The precursor solution of 2D perovskite BA₂MA₄Pb₅I₁₆ (n = 5) was prepared by mixing BAI, MAI, PbI₂, and NH₄SCN in dimethylsulfoxide (DMSO) and dimethylformamide (DMF) solvents (5:5 volume ratio) with a stoichiometric ratio of 2:4:5:2; the concentration of Pb²⁺ in the precursor solution was 1.2 M." 2

CH₃NH₃PbI₃

This 3D perovskite compound is briefly mentioned when discussing previous work:

"Guo et al. reported that Ti₃C₂T_x doping in the CH₃NH₃PbI₃ perovskite layer efficiently enhanced the crystal size and charge transfer of the film." 3

The paper focuses primarily on the 2D (BA)₂(MA)₄Pb₅I₁₆ perovskite, with the goal of improving its performance by incorporating Ti₃C₂T_x MXene nanosheets as dopants. No other specific perovskite compounds are mentioned in the Experimental or Results sections.

Observed behavior (Elicit)

Elicit produces text-grounded extractions that directly quote relevant passages from the source document to justify identified compounds. The model correctly identifies both the synthesized 2D perovskite absorber and the 3D perovskite referenced for comparison, and explicitly qualifies the context in which each appears. Compared to more conservative extractors, Elicit provides extended quoted context and explanatory prose, resulting in moderate verbosity and less compact structure. This behavior is consistent with its intermediate performance and higher variance observed in quantitative benchmarking.

You said:

List all perovskite compounds mentioned in the Experimental or Results sections.

Publication Analyzer said:

The perovskite compound specifically mentioned in the Experimental and Results sections of the document is:

- (BA)₂(MA)₄Pb₅I₁₆

This compound is a two-dimensional Ruddlesden–Popper (RP) perovskite and forms the core of the study's photovoltaic absorber layer. It was doped with varying concentrations of Ti₃C₂T_x MXene nanosheets to enhance its properties.

You said:

For each perovskite, list all materials used in its synthesis and their IUPAC names.

Publication Analyzer said:

- For the synthesis of the perovskite $(\text{BA})_2(\text{MA})_4\text{Pb}_5\text{I}_{16}$, the materials used and their corresponding IUPAC names are as follows:
 1. **BAI (Butylammonium iodide)**
 - **IUPAC Name:** Butan-1-aminium iodide
 2. **MAI (Methylammonium iodide)**
 - **IUPAC Name:** Methanaminium iodide
 3. **PbI₂ (Lead(II) iodide)**
 - **IUPAC Name:** Plumbane(2+) diiodide
 4. **NH₄SCN (Ammonium thiocyanate)**
 - **IUPAC Name:** Azanium thiocyanate

These were mixed in a solvent mixture of DMSO and DMF to create the perovskite precursor solution. Additionally, $\text{Ti}_3\text{C}_2\text{T}_x$ MXene was used as a nano-dopant, but it's a material with variable composition (surface terminations), not defined by a single IUPAC name.

Observed behavior (Publication Analyzer)

Publication Analyzer produces minimal, high-precision extractions that adhere strictly to the scope of the prompt. The model lists only the perovskite compound explicitly synthesized and characterized in the study and avoids inclusion of referenced or inferred materials. Precursor components are enumerated concisely with correct IUPAC names, and chemically ambiguous entities such as MXene additives are explicitly identified as lacking a single canonical IUPAC designation. This conservative extraction strategy prioritizes precision and reproducibility over completeness, consistent with the model's low variance and high stability observed in quantitative benchmarking.

Summary and Relation to Quantitative Benchmarking

Taken together, these verbatim outputs demonstrate that the qualitative behaviors discussed in the main text differences in structure, verbosity, abstraction, and grounding are directly observable and systematic rather than anecdotal. The examples contextualize the quantitative performance metrics reported in Fig. 3 and Supplementary Fig. S1, illustrating how stylistic tendencies and extraction strategies influence recall, precision, verbosity, and reproducibility across models.