

Supplementary Information: Automated electrosynthesis reaction mining with multimodal large language models (MLLMs)

Shi Xuan Leong,^{1,2} Sergio Pablo-García,^{1,3,4} Zijian Zhang,^{3,4} Alán Aspuru-Guzik^{1,3,4,5,6,7,8*}

¹Department of Chemistry, University of Toronto, Lash Miller Chemical Laboratories, 80 St. George Street, ON M5S 3H6, Toronto, Canada

²Division of Chemistry and Biological Chemistry, School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, 21 Nanyang Link, Singapore 637371

³Department of Computer Science, University of Toronto, Sandford Fleming Building, 10 King's College Road, ON M5S 3G4, Toronto, Canada

⁴Vector Institute for Artificial Intelligence, 661 University Ave. Suite 710, ON M5G 1M1, Toronto, Canada

⁵Acceleration Consortium, 80 St George St, M5S 3H6, Toronto, Canada

⁶Department of Materials Science & Engineering, University of Toronto, 184 College St., M5S 3E4, Toronto, Canada

⁷Department of Chemical Engineering & Applied Chemistry, University of Toronto, 200 College St. ON M5S 3E5, Toronto, Canada

⁸Lebovic Fellow, Canadian Institute for Advanced Research (CIFAR), 661 University Ave., M5G 1M1, Toronto, Canada

*Corresponding author email: alan@aspuru.com

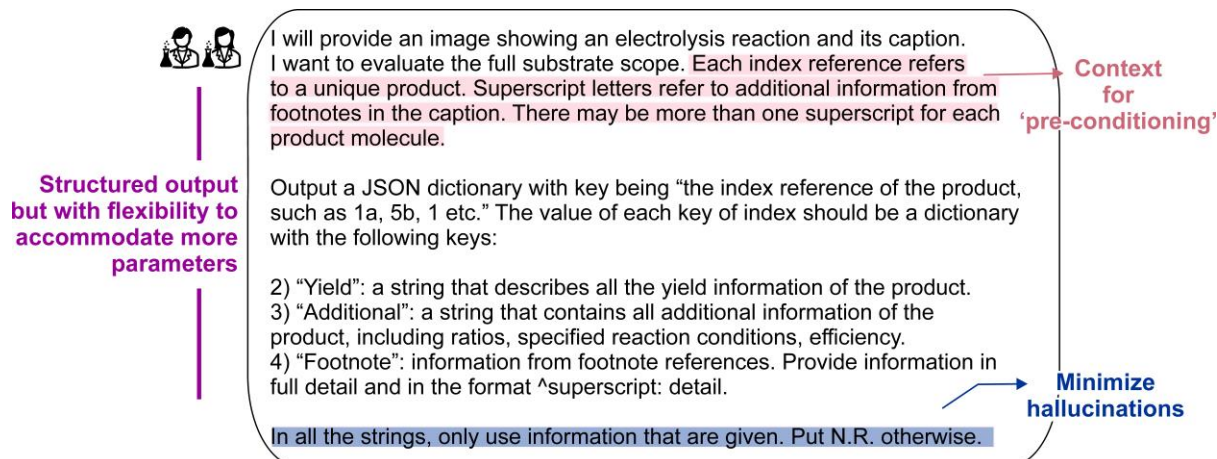
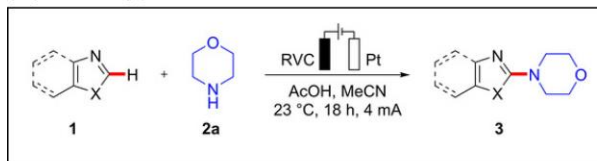


Figure S1. Full formal prompt to extract footnote cross-references description and substrate-specific information including index identifier, yield information, and any additional information related to the investigated reaction for each synthesized product.

Supplementary Note 1.

(A) Error type 1: Misclassification



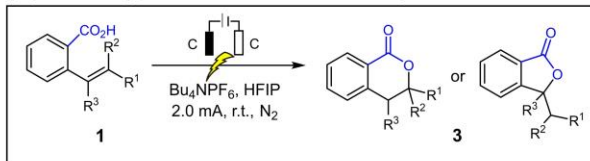
```

1  {}
2  {
3  "anode material": "Pt",
4  "cathode material": "RVC",
5  "electrolytes": "AcOH, MeCN",
6  "solvents": "N.R.",
7  "current": "4 mA",
8  "duration": "18 h",
9  "air/inert": "N.R.",
10 "temperature": "23 °C",
11 "others": "N.R."
12 }
13 ...

```

→ anode and cathode materials are swapped
 → solvent misclassified as electrolyte

(B) Error type 2: Failure to identify



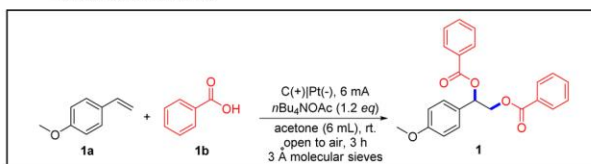
```

1  {}
2  {
3  "anode material": "carbon",
4  "cathode material": "N.R.",
5  "electrolytes": "Bu4NPF6 (2.0 mM) in HFIP",
6  "solvents": "N.R.",
7  "current": "2.0 mA",
8  "duration": "N.R.",
9  "air/inert": "N2",
10 "temperature": "room temperature",
11 "others": "N.R."
12 }
13 ...

```

→ Failure to identify cathode
 → Error type 1 (solvent misclassified as electrolyte)

(C) Error type 3: Cannot recognize contextual abbreviations



```

1  {}
2  {
3  "anode material": "Pt",
4  "cathode material": "C(+)Pt",
5  "electrolytes": "nBu4NOAc(1.2 eq.)",
6  "solvents": "acetone (6 mL, open to air)",
7  "current": "N.R.",
8  "duration": "3 h",
9  "air/inert": "open to air",
10 "temperature": "N.R.",
11 "others": "N.R."
12 }
13 ...

```

→ Error type 2 (cannot identify)

Figure S2. Example of common errors observed in unsuccessful (incorrect) responses outputs for Task 1, namely a) misclassification of identified reaction conditions, as well as failure to identify b) reaction conditions and c) contextual abbreviations. The first two response outputs are from Gemini and the third response output is from LLaVA. The diagrams depicted are adapted with permission from ref. 1 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim; ref. 2 © 2022 Wiley-VCH GmbH; and ref 3. with permission from Royal Society of Chemistry.

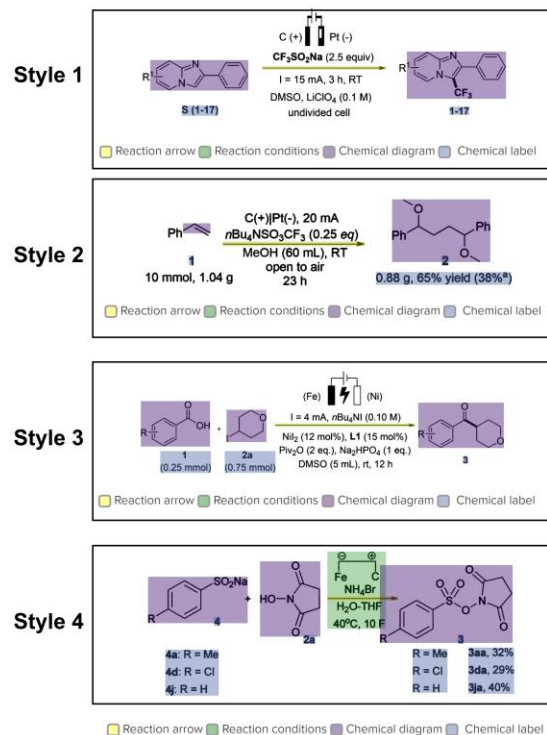
In terms of reaction diagram parsing, GPT-4V,⁴ Gemini⁵ and Claude 3⁶ demonstrate comparable accuracies at identifying and classifying the standard reaction parameters including current, duration, air/inert atmosphere, and temperature, with $\geq 96\%$ accuracy using both hard and soft match evaluation. For domain-specific parameters whose exact roles demand domain

knowledge, the models perform reasonably well but exhibit lower accuracies for solvents and electrode materials, with GPT-4V outperforming the other MLLMs. Gemini and Claude 3 have a higher tendency to misclassify solvents as electrolytes/additives (Figure S2-A), as reflected from their lower hard match accuracy of 59% and 68%, respectively, compared to 83% for GPT-4V. For electrode materials, Gemini and Claude 3 often fail to identify the electrode materials and would return incorrect “Not reported (N.R.)” responses or provide a random material (Figure S2-B), thus resulting in lower soft match anode (cathode) identification accuracies of 66(66) % and 72(69) % for the two models, respectively, compared to 93(93) % using the GPT-4V model.

As for InternVL⁷ and LLaVA,^{8,9} they both exhibit poor contextual awareness of abbreviations that are commonly found in scientific publications, resulting in poor identification accuracies across all categories. For instance, they often fail to recognize ““N₂” and “Ar” as descriptors of an inert atmosphere, as well as “RT” or “r.t.” as abbreviations for room temperature, resulting in the generation of inaccurate ‘Not reported (N.R.)’ responses (Figure S2-C). In addition, they frequently fail to identify the electrode materials and/or fully identify mixed solvents and/or electrolyte systems.

Supplementary Note 2.

(A) ReactionDataExtractor2.0



(B) RxnScribe

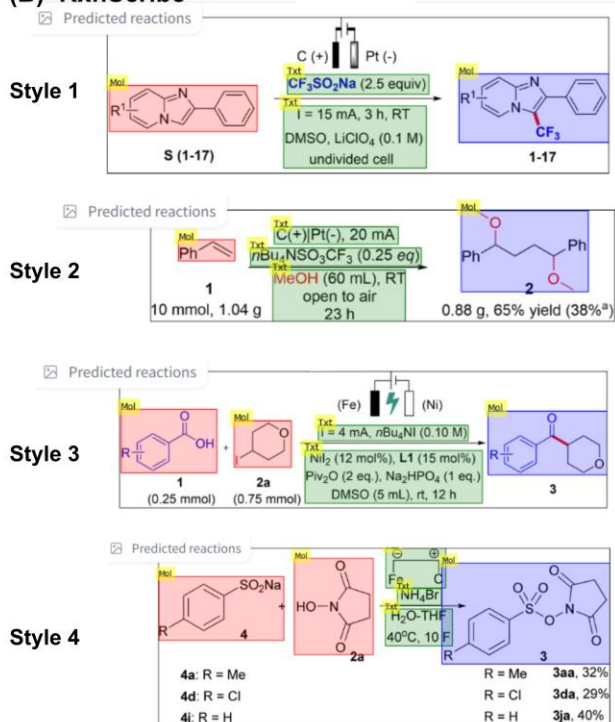


Figure S3. (A) Example visualizations of (A) ReactionDataExtractor2.0 and (B) RxnScribe predictions using the web-based interfaces *via* <https://www.reactiondataextractor.org/demo> and <https://huggingface.co/spaces/yujieq/RxnScribe>, respectively. The diagrams depicted were adapted with permission from ref. 10 (Style 1) © 2023 Wiley-VCH GmbH, ref. 11 (Style 2) with permission from the Royal Society of Chemistry, ref. 12 (Style 3) © 2022 Wiley-VCH GmbH, and ref. 13 with permission from the Royal Society of Chemistry.

We compare the MLLM-based reaction diagram parsing with two reaction diagram parsing software – ReactionDataExtractor2.0¹⁴ and RxnScribe¹⁵. The former is a rule-based pipeline while the latter is a sequence generation model. Considering that these toolkits are trained on chemical reaction schemes and not specifically tailored for electrochemistry reaction diagrams, we performed the equivalent of a soft match evaluation, whereby the identified parameter is assigned as a correct prediction if it is included under “Reaction conditions” for both ReactionDataExtractor2.0 and RxnScribe. The poor performance of ReactionDataExtractor2.0 with < 40% accuracy across all categories is due to its failure at identifying the regions with

reaction conditions for most of the reaction diagrams, despite accurately identifying the location and direction of the reaction arrows (**Figure S2A**). On the other hand, RxnScribe achieves > 90% accuracy for 5 out of 10 categories, comparable to that of MLLM-based reaction diagram parsing (**Figure 3A** in main text). We observe that the high accuracies are for more standard reaction conditions such as duration and atmospheric conditions, and for quantitative categories such as solvent and electrolyte amounts. In comparison, domain-specific information such as the electrode materials and the electrolytes are poorly identified, whereby RxnScribe makes mistakes such as missing one or both electrode materials. We also note that the software is prone to misspelling such as “H₂O” instead of “H₂O”, and “ⁿBuANBF₄” instead of “ⁿBu₄NBF₄”, likely due to model artefacts. These misspellings were disregarded for the purpose of comparison in this study.

Supplementary Note 3.

A sample response output for Task 2 – resolving cross-modality interdependencies, generated using GPT-4V as the MLLM is presented below. Yield and any additional information of each substrate (indicated by its unique index label), are extracted from the figure and captions and organized into a structured dictionary. Notably, full footnote references are accurately identified, e.g. for substrates 3ad, 3ag, 3ba etc.

```
{
  "3aa": {
    "yield": "81%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3ab": {
    "yield": "80%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3ac": {
    "yield": "81%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3ad": {
    "yield": "75%",
    "additional": "N.R.",
    "footnote": "^a: 6 h."
  },
  "3ae": {
    "yield": "84%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3af": {
    "yield": "43%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3ag": {
    "yield": "87%",
    "additional": "N.R.",
    "footnote": "^b: 7 h."
  },
}
```

```

"3ah": {
  "yield": "trace",
  "additional": "N.R.",
  "footnote": "N.R."
},
"3ai": {
  "yield": "95%",
  "additional": "N.R.",
  "footnote": "^c: 8 h."
},
"3aj": {
  "yield": "88%",
  "additional": "N.R.",
  "footnote": "^c: 8 h."
},
"3ak": {
  "yield": "81%",
  "additional": "N.R.",
  "footnote": "N.R."
},
"3ba": {
  "yield": "91%",
  "additional": "N.R.",
  "footnote": "^d: 3aa as a by-product was isolated in 40% yield."
},
"3ca": {
  "yield": "72%",
  "additional": "N.R.",
  "footnote": "^c: 8 h."
},
"3da": {
  "yield": "85%",
  "additional": "N.R.",
  "footnote": "^a: 6 h."
},
"3ea": {
  "yield": "66%",
  "additional": "N.R.",
  "footnote": "^b: 7 h."
},
"3fa": {
  "yield": "44%",
  "additional": "3aa as a by-product was isolated in 40% yield.",
  "footnote": "N.R."
},
"3ga": {
  "yield": "42%",
  "additional": "N.R.",

```



```

    "footnote": "N.R."
  },
  "3ha": {
    "yield": "92%",
    "additional": "N.R.",
    "footnote": "^a: 6 h."
  },
  "3ia": {
    "yield": "88%",
    "additional": "N.R.",
    "footnote": ^b: 7 h."
  },
  "3ja": {
    "yield": "42%",
    "additional": "N.R.",
    "footnote": "N.R."
  },
  "3ka": {
    "yield": "90%",
    "additional": "N.R.",
    "footnote": {}
  },
  "3la": {
    "yield": "87%",
    "additional": "N.R.",
    "footnote": {}
  },
  "3ma": {
    "yield": "53%",
    "additional": "3aa as a by-product was isolated in 40% yield.",
    "footnote": {
      ^d: "3aa as a by-product was isolated in 40% yield."
    }
  },
  "3na": {
    "yield": "92%",
    "additional": "N.R.",
    "footnote": {}
  },
  "3oa": {
    "yield": "90%",
    "additional": "N.R.",
    "footnote": {}
  },
  "3pa": {
    "yield": "91%",
    "additional": "N.R.",
    "footnote": {}
  }

```

```
},  
"3qa": {  
  "yield": "83%",  
  "additional": "N.R.",  
  "footnote": {}  
},  
"3ra": {  
  "yield": "18%",  
  "additional": "N.R.",  
  "footnote": {}  
}  
}
```

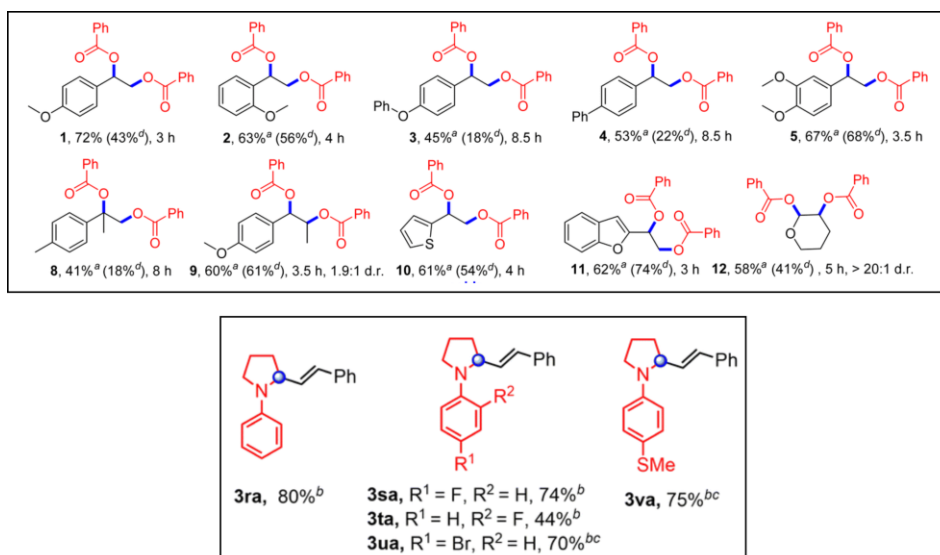


Figure S4. Example images with false negative assignments (i.e., mis-assignment or overlooking of footnote references). Adopted with permission from ref. 3 (top) and ref. 16 (bottom), with permission from the Royal Society of Chemistry.

Table S1. List of DOI identifiers

#	DOI identifier	#	DOI identifier	#	DOI identifier
1	10.1039/d1ob00079a	23	10.1002/adsc.202301071	45	10.1002/anie.201909951
2	10.1039/d1qo00038a	24	10.1002/adsc.202300118	46	10.1002/anie.202201595
3	10.1039/D3GC02701E	25	10.1002/adsc.202300730	47	10.1002/adsc.202200003
4	10.1039/C80B03162B	26	10.1002/adsc.202301343	48	10.1002/ejoc.202000311
5	10.1039/C9GC01474H	27	10.1002/adsc.202000997	49	10.1002/ejoc.202300927
6	10.1039/D1GC00661D	28	10.1002/anie.201910077	50	10.1002/anie.202212131
7	10.1039/D2GC04399H	29	10.1002/anie.202207660	51	10.1002/ajoc.202100620
8	10.1039/D3GC02345A	30	10.1002/ajoc.202200719	52	10.1002/cssc.201900378
9	10.1039/D20B01402E	31	10.1002/ajoc.202200425	53	10.1039/d3ob00671a
10	10.1039/D3NJ05205B	32	10.1002/ajoc.202300294	54	10.1039/D2GC02086F
11	10.1039/D3GC03389A	33	10.1002/chem.201802832	55	10.1039/D1GC01069G
12	10.1039/D0GC00771D	34	10.1002/chem.202201654	56	10.1039/D30B00831B
13	10.1039/D3SC00527E	35	10.1002/celc.202200989	57	10.1039/C8CC09899A
14	10.1039/C8GC01411F	36	10.1002/celc.201800422	58	10.1039/C9CC00975B
15	10.1039/C7GC00468K	37	10.1002/ejoc.201901928	59	10.1039/C8CC06451B
16	10.1039/D3Q000204G	38	10.1002/ejoc.202300553	60	10.1002/anie.201808555
17	10.1039/D2CC03883H	39	10.1002/chem.202100960	61	10.1002/asia.202200780
18	10.1039/D2SC05507D	40	10.1002/celc.202101155	62	10.1002/cctc.202300258
19	10.1039/C9CC03789F	41	10.1002/anie.201610715	63	10.1002/celc.201900138
20	10.1039/C80B00500A	42	10.1039/C9CC03789F	64	10.1002/adsc.202200932
21	10.1002/ejoc.202300063	43	10.1002/anie.202013478	65	10.1039/D2GC00457G
22	10.1002/adsc.202200847	44	10.1002/celc.201900080		

- 1 **Table S2.** Hard and soft match % precision of electrosynthesis reaction diagram parsing across different MLLMs without any visual prompting.
- 2 Each model prediction was assigned as follows: true positive for correct identification of reaction parameter; false positive for incorrect assignment
- 3 or redundant information; true negative for correct identification of unspecified reaction parameters; and false negative for missing information.

	Precision									
	Anode (+)	Cathode (-)	Electrolyte/additives	Electrolytes /additives amounts	Solvents	Solvent amount	Current	Duration	Air/ Inert	Temperature
GPT-4V	100% (100%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)	99% (100%)	100% (100%)	100% (100%)	100% (100%)
Gemini	100% (100%)	100% (100%)	99% (100%)	93% (92%)	100% (100%)	93% (85%)	100% (100%)	100% (100%)	100% (97%)	100% (100%)
Claude 3	100% (100%)	100% (100%)	99% (99%)	79% (79%)	100% (100%)	78% (83%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)
LLaVA	100% (100%)	94% (97%)	96% (96%)	60% (62%)	100% (100%)	48% (50%)	93% (95%)	92% (92%)	70% (72%)	97% (98%)
InternVL	100% (100%)	97% (98%)	100% (100%)	51% (52%)	100% (100%)	39% (39%)	91% (92%)	78% (78%)	64% (64%)	95% (95%)
ReactionData Extractor 2.0	- (100%)	- (100%)	- (96%)	- (40%)	- (100%)	- (20%)	- (84%)	- (59%)	- (27%)	- (92%)
RxnScribe	- (100%)	- (100%)	- (100%)	- (93%)	- (100%)	- (86%)	- (100%)	- (100%)	- (97%)	- (99%)

5 **Table S3.** Hard and soft match % recall of electrosynthesis reaction diagram parsing across different MLLMs without any visual prompting. Each
6 model prediction was assigned as follows: true positive for correct identification of reaction parameter; false positive for incorrect assignment or
7 redundant information; true negative for correct identification of unspecified reaction parameters; and false negative for missing information.

	Recall									
	Anode (+)	Cathode (-)	Electrolyte/additives	Electrolytes/additives amounts	Solvents	Solvent amount	Current	Duration	Air/Inert	Temperature
GPT-4V	85% (93%)	85% (93%)	100% (100%)	100% (100%)	83% (100%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)
Gemini	61% (66%)	50% (59%)	89% (89%)	87% (100%)	59% (89%)	88% (88%)	97% (92%)	95% (90%)	97% (97%)	100% (96%)
Claude 3	56% (72%)	54% (69%)	96% (98%)	100% (100%)	68% (92%)	94% (94%)	99% (99%)	95% (95%)	100% (100%)	99% (99%)
LLaVA	16% (29%)	20% (37%)	54% (58%)	72% (79%)	30% (46%)	75% (81%)	17% (26%)	73% (78%)	60% (66%)	39% (52%)
InternVL	46% (60%)	43% (58%)	50% (51%)	53% (55%)	26% (51%)	59% (59%)	55% (57%)	48% (48%)	83% (83%)	52% (52%)
ReactionData Extractor 2.0	- (24%)	- (23%)	- (32%)	- (42%)	- (33%)	- (33%)	- (30%)	- (36%)	- (38%)	- (31%)
RxnScribe	- (49%)	- (49%)	- (79%)	- (83%)	- (83%)	- (100%)	- (95%)	- (97%)	- (100%)	- (100%)

9 **Table S4.** Hard and soft match % F1 score of electrosynthesis reaction diagram parsing across different MLLMs without any visual prompting.
 10 Each model prediction was assigned as follows: true positive for correct identification of reaction parameter; false positive for incorrect assignment
 11 or redundant information; true negative for correct identification of unspecified reaction parameters; and false negative for missing information.

	F1 score									
	Anode (+)	Cathode (-)	Electrolyte/additives	Electrolytes /additives amounts	Solvents	Solvent amount	Current	Duration	Air/ Inert	Temperature
GPT-4V	92% (96%)	92% (96%)	100% (100%)	100% (100%)	91% (100%)	100% (100%)	99% (100%)	100% (100%)	100% (100%)	100% (100%)
Gemini	76% (79%)	67% (74%)	94% (94%)	90% (96%)	74% (94%)	90% (86%)	99% (96%)	97% (95%)	99% (97%)	100% (98%)
Claude 3	72% (84%)	70% (82%)	98% (98%)	88% (88%)	81% (96%)	85% (88%)	99% (99%)	97% (97%)	100% (100%)	99% (99%)
LLaVA	28% (45%)	33% (54%)	69% (72%)	65% (69%)	46% (63%)	59% (62%)	29% (41%)	81% (85%)	65% (69%)	56% (68%)
InternVL	63% (75%)	60% (73%)	67% (68%)	52% (54%)	42% (67%)	47% (47%)	68% (70%)	60% (60%)	73% (73%)	67% (67%)
ReactionData Extractor 2.0	- (38%)	- (37%)	- (48%)	- (41%)	- (49%)	- (25%)	- (44%)	- (44%)	- (32%)	- (46%)
RxnScribe	- (66%)	- (66%)	- (88%)	- (87%)	- (90%)	- (93%)	- (97%)	- (98%)	- (99%)	- (99%)

12

13

14 **Table S5.** Hard and soft match % specificity of electrosynthesis reaction diagram parsing across different MLLMs without any visual prompting.
 15 ‘Not applicable (NA)’ is indicated when there are no instances in the figure or when there is a division by zero error. Each model prediction was
 16 assigned as follows: true positive for correct identification of reaction parameter; false positive for incorrect assignment or redundant information;
 17 true negative for correct identification of unspecified reaction parameters; and false negative for missing information.

	Specificity									
	Anode (+)	Cathode (-)	Electrolyte/additives	Electrolytes/additives amounts	Solvents	Solvent amount	Current	Duration	Air/Inert	Temperature
GPT-4V	N.A.	100% (100%)	100% (100%)	100% (100%)	N.A.	100% (100%)	90% (100%)	100% (100%)	100% (100%)	100% (100%)
Gemini	N.A.	100% (100%)	50% (100%)	93% (90%)	N.A.	96% (91%)	100% (100%)	100% (100%)	100% (98%)	100% (100%)
Claude 3	N.A.	100% (100%)	50% (50%)	69% (69%)	N.A.	89% (85%)	100% (100%)	100% (100%)	100% (100%)	100% (100%)
LLaVA	N.A.	0% (0%)	0% (0%)	43% (43%)	N.A.	53% (53%)	90% (90%)	85% (85%)	83% (83%)	88% (88%)
InternVL	N.A.	0% (0%)	100% (100%)	40% (40%)	N.A.	44% (44%)	60% (60%)	70% (73%)	69% (69%)	75% (75%)
ReactionData Extractor 2.0	N.A.	- (100%)	- (50%)	- (27%)	N.A.	- (32%)	- (60%)	- (42%)	- (33%)	- (71%)
RxnScribe	N.A.	- (100%)	- (100%)	- (93%)	N.A.	- (91%)	- (100%)	- (100%)	- (98%)	- (99%)

19 **Table S6.** Hard match % identification accuracy of electrosynthesis reaction diagram parsing
20 using GPT-4V, with and without the integration of different single-shot visual prompts.

	Zero-shot (no prompting)	Single-shot visual prompts			
		Style 1 only	Style 2 only	Style 3 only	All styles
Anode	85%	96%	96%	93%	99%
Cathode	85%	98%	96%	93%	99%
Electrolyte/ additive	89%	100%	98%	99%	100%
Electrolyte/additive amount	89%	100%	100%	99%	100%
Solvents	94%	100%	93%	100%	100%
Solvent amounts	94%	100%	100%	100%	100%
Current	99%	100%	100%	100%	100%
Duration	100%	100%	100%	99%	100%
Air/inert	100%	100%	100%	100%	100%
Temperature	100%	100%	100%	100%	100%

21

22

23 **Table S7.** Hard match % identification accuracy for electrode materials across the four
 24 presentation styles with and without the integration of different single-shot visual prompts.

	Zero-shot (no prompting)	Single-shot visual prompts			
		Style 1 only	Style 2 only	Style 3 only	All styles
Anode (style 1)	100%	100%	100%	100%	100%
Cathode (style 1)	92%	100%	100%	100%	100%
Anode (style 2)	100%	100%	100%	96%	100%
Cathode (style 2)	100%	100%	100%	96%	100%
Anode (style 3)	75%	98%	95%	98%	100%
Cathode (style 3)	78%	100%	95%	98%	100%
Anode (style 4)	70%	80%	90%	90%	89%
Cathode (style 4)	70%	80%	89%	89%	89%

25 In the absence of single-shot visual prompts, analysis of GPT-4V’s responses underscores its
 26 struggle to handle more specialized chemical tasks wherein the presentation style significantly
 27 deviates from standard norm, despite including pre-conditioning textual prompts. For an
 28 example figure whereby the model returns “N.R.” for the two electrode materials, the provided
 29 explanation is as follows: “The image provided does not depict a standard electrolysis reaction
 30 scheme or its typical conditions; instead, it shows a chemical reaction scheme. Therefore, it
 31 does not contain specific information about electrolysis conditions such as anode or cathode
 32 material, current, or duration.”

33 **Table S8.** Precision, recall, F1 scores, specificity, and accuracy of each figure without image
 34 cropping: resolving footnote cross-references using GPT-4V. ‘Not applicable (NA)’ is indicated
 35 when there are no instances in the figure or when there is a division by zero error.

Feeding entire figure					
Figure	Precision	Recall	F1	Specificity	Accuracy
1.	100%	100%	100%	100%	100%
2.	100%	100%	100%	100%	100%
3.	0%	0%	NA	NA	0%
4.	100%	100%	100%	100%	100%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	50%	50%	50%	100%	83%
8.	19%	16%	18%	NA	16%
9.	100%	100%	100%	100%	100%
10.	100%	100%	100%	100%	100%
11.	100%	100%	100%	100%	100%
12.	NA	NA	NA	100%	100%
13.	67%	100%	80%	100%	100%
14.	NA	NA	NA	100%	100%
15.	NA	NA	NA	100%	100%
16.	100%	90%	95%	96%	95%
17.	67%	100%	80%	100%	100%
18.	NA	NA	NA	100%	100%
19.	75%	100%	86%	89%	91%
20.	100%	100%	100%	100%	100%
21.	100%	100%	100%	100%	100%
22.	100%	100%	100%	100%	100%
23.	100%	100%	100%	100%	100%
24.	100%	100%	100%	100%	100%
25.	83%	71%	77%	100%	92%
26.	100%	100%	100%	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	100%	100%	100%	100%	100%
30.	100%	100%	100%	100%	100%
31.	125%	100%	111%	100%	100%
32.	100%	100%	100%	100%	100%
33.	80%	67%	73%	96%	86%
34.	NA	NA	NA	100%	100%
35.	94%	100%	97%	100%	100%
36.	67%	67%	67%	NA	67%
37.	83%	88%	86%	86%	87%
38.	100%	100%	100%	100%	100%

39.	100%	100%	100%	100%	100%
40.	100%	100%	100%	100%	100%
41.	100%	100%	100%	100%	100%
42.	100%	100%	100%	100%	100%
43.	87%	87%	87%	100%	92%
44.	NA	NA	NA	100%	100%
45.	NA	NA	NA	100%	100%
46.	17%	18%	17%	90%	52%
47.	83%	83%	83%	94%	92%
48.	100%	100%	100%	100%	100%
49.	100%	100%	100%	100%	100%
50.	75%	75%	75%	97%	94%
51.	NA	NA	NA	100%	100%
52.	100%	100%	100%	88%	93%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	67%	75%	71%	97%	93%
56.	100%	100%	100%	92%	94%
57.	100%	100%	100%	100%	100%
58.	NA	NA	NA	100%	100%
59.	NA	NA	NA	100%	100%
60.	100%	100%	100%	100%	100%
61.	19%	18%	19%	NA	18%
62.	100%	100%	100%	100%	100%
63.	100%	100%	100%	100%	100%
64.	100%	100%	100%	100%	100%
65.	67%	67%	67%	100%	94%
66.	100%	100%	100%	100%	100%
67.	100%	100%	100%	100%	100%
68.	83%	100%	91%	100%	100%
69.	100%	100%	100%	100%	100%
70.	100%	100%	100%	100%	100%
71.	113%	100%	106%	100%	100%
72.	67%	100%	80%	77%	84%
73.	100%	100%	100%	100%	100%
74.	100%	100%	100%	100%	100%

37 **Table S9.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
 38 cropping: resolving footnote cross-references using GPT-4V. ‘Not applicable (NA)’ is indicated
 39 when there are no instances in the figure or when there is a division by zero error.

With image cropping					
Figure	Precision	Recall	F1	Specificity	Accuracy
1.	100%	100%	100%	100%	100%
2.	100%	100%	100%	100%	100%
3.	100%	100%	100%	NA	100%
4.	100%	100%	100%	100%	100%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	100%	100%	100%	100%	100%
8.	100%	84%	91%	NA	84%
9.	100%	100%	100%	100%	100%
10.	100%	100%	100%	100%	100%
11.	100%	100%	100%	100%	100%
12.	NA	NA	NA	100%	100%
13.	67%	100%	80%	91%	92%
14.	NA	NA	NA	100%	100%
15.	NA	NA	NA	100%	100%
16.	89%	80%	84%	96%	92%
17.	67%	100%	80%	95%	96%
18.	NA	NA	NA	100%	100%
19.	75%	100%	86%	95%	95%
20.	100%	100%	100%	100%	100%
21.	100%	100%	100%	100%	100%
22.	100%	100%	100%	100%	100%
23.	100%	100%	100%	100%	100%
24.	100%	100%	100%	100%	100%
25.	100%	86%	92%	100%	96%
26.	100%	100%	100%	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	100%	100%	100%	100%	100%
30.	100%	100%	100%	100%	100%
31.	100%	80%	89%	100%	95%
32.	100%	100%	100%	100%	100%
33.	100%	83%	91%	100%	94%
34.	NA	NA	NA	100%	100%
35.	94%	100%	97%	96%	98%
36.	100%	100%	100%	NA	100%
37.	83%	88%	86%	86%	87%
38.	100%	100%	100%	100%	100%

39.	100%	100%	100%	100%	100%
40.	100%	100%	100%	100%	100%
41.	100%	100%	100%	100%	100%
42.	100%	100%	100%	100%	100%
43.	100%	100%	100%	100%	100%
44.	NA	NA	NA	100%	100%
45.	NA	NA	NA	100%	100%
46.	75%	82%	78%	70%	76%
47.	100%	100%	100%	100%	100%
48.	100%	100%	100%	100%	100%
49.	100%	100%	100%	100%	100%
50.	100%	100%	100%	100%	100%
51.	NA	NA	NA	100%	100%
52.	100%	100%	100%	100%	100%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	89%	100%	94%	97%	98%
56.	100%	100%	100%	100%	100%
57.	100%	100%	100%	100%	100%
58.	NA	NA	NA	100%	100%
59.	NA	NA	NA	100%	100%
60.	100%	100%	100%	100%	100%
61.	100%	97%	99%	NA	97%
62.	100%	100%	100%	100%	100%
63.	100%	100%	100%	100%	100%
64.	100%	100%	100%	100%	100%
65.	100%	100%	100%	100%	100%
66.	100%	100%	100%	100%	100%
67.	100%	100%	100%	100%	100%
68.	83%	100%	91%	89%	93%
69.	100%	100%	100%	100%	100%
70.	100%	100%	100%	100%	100%
71.	100%	89%	94%	100%	94%
72.	67%	100%	80%	77%	84%
73.	100%	100%	100%	100%	100%
74.	100%	100%	100%	100%	100%

41 **Table S10.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
 42 cropping: resolving footnote cross-references using Gemini. ‘Not applicable (NA)’ is indicated
 43 when there are no instances in the figure or when there is a division by zero error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	100%	50%	67%	100%	92%
2.	100%	100%	100%	100%	100%
3.	NA	0%	NA	NA	0%
4.	100%	100%	100%	100%	100%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	NA	0%	NA	100%	67%
8.	100%	6%	12%	NA	6%
9.	9%	100%	17%	63%	64%
10.	100%	100%	100%	100%	100%
11.	77%	100%	87%	88%	91%
12.	NA	NA	NA	100%	100%
13.	0%	0%	NA	91%	77%
14.	NA	NA	NA	100%	100%
15.	NA	NA	NA	100%	100%
16.	100%	70%	82%	100%	92%
17.	100%	100%	100%	100%	100%
18.	NA	NA	NA	100%	100%
19.	75%	100%	86%	95%	95%
20.	NA	0%	NA	100%	67%
21.	100%	100%	100%	100%	100%
22.	100%	60%	75%	100%	90%
23.	100%	100%	100%	100%	100%
24.	100%	29%	44%	100%	67%
25.	100%	43%	60%	100%	83%
26.	60%	100%	75%	88%	90%
27.	100%	67%	80%	100%	72%
28.	100%	20%	33%	100%	60%
29.	100%	75%	86%	100%	94%
30.	100%	75%	86%	100%	83%
31.	100%	40%	57%	100%	86%
32.	100%	33%	50%	100%	88%
33.	100%	17%	29%	100%	71%
34.	NA	0%	NA	100%	82%
35.	85%	69%	76%	92%	83%
36.	100%	100%	100%	NA	100%
37.	100%	100%	100%	100%	100%
38.	100%	100%	100%	100%	100%
39.	100%	100%	100%	100%	100%

40.	75%	100%	86%	93%	94%
41.	67%	100%	80%	92%	93%
42.	83%	100%	91%	95%	96%
43.	94%	100%	97%	89%	96%
44.	NA	NA	NA	100%	100%
45.	NA	NA	NA	100%	100%
46.	56%	45%	50%	60%	52%
47.	50%	33%	40%	89%	75%
48.	100%	100%	100%	100%	100%
49.	89%	100%	94%	96%	97%
50.	100%	100%	100%	100%	100%
51.	NA	NA	NA	100%	100%
52.	100%	14%	25%	100%	60%
53.	80%	80%	80%	95%	92%
54.	100%	100%	100%	100%	100%
55.	89%	100%	94%	97%	98%
56.	50%	67%	57%	67%	67%
57.	100%	67%	80%	100%	95%
58.	NA	NA	NA	100%	100%
59.	NA	NA	NA	100%	100%
60.	100%	100%	100%	100%	100%
61.	100%	42%	59%	NA	42%
62.	100%	100%	100%	100%	100%
63.	100%	69%	82%	100%	86%
64.	75%	86%	80%	91%	90%
65.	100%	100%	100%	100%	100%
66.	100%	100%	100%	100%	100%
67.	100%	75%	86%	100%	88%
68.	NA	0%	NA	100%	64%
69.	100%	100%	100%	100%	100%
70.	100%	100%	100%	100%	100%
71.	100%	78%	88%	100%	88%
72.	43%	43%	43%	67%	58%
73.	100%	100%	100%	100%	100%
74.	100%	100%	100%	100%	100%

45 **Table S11.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
 46 cropping: resolving footnote cross-references using Claude 3. ‘Not applicable (NA)’ is
 47 indicated when there are no instances in the figure or when there is a division by zero error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	100%	100%	100%	100%	100%
2.	71%	100%	83%	91%	93%
3.	100%	94%	97%	NA	94%
4.	50%	100%	67%	93%	94%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	100%	100%	100%	100%	100%
8.	100%	81%	89%	NA	81%
9.	100%	100%	100%	100%	100%
10.	100%	100%	100%	100%	100%
11.	77%	100%	87%	88%	91%
12.	NA	NA	NA	100%	100%
13.	67%	100%	80%	91%	92%
14.	NA	NA	NA	100%	100%
15.	NA	NA	NA	100%	100%
16.	89%	80%	84%	96%	92%
17.	100%	100%	100%	100%	100%
18.	NA	NA	NA	100%	100%
19.	100%	100%	100%	100%	100%
20.	36%	100%	53%	10%	40%
21.	67%	100%	80%	75%	83%
22.	100%	100%	100%	100%	100%
23.	100%	100%	100%	100%	100%
24.	100%	100%	100%	100%	100%
25.	75%	86%	80%	88%	88%
26.	60%	100%	75%	88%	90%
27.	100%	93%	97%	100%	94%
28.	90%	90%	90%	90%	90%
29.	75%	75%	75%	92%	88%
30.	100%	100%	100%	100%	100%
31.	100%	100%	100%	100%	100%
32.	100%	100%	100%	100%	100%
33.	100%	92%	96%	100%	97%
34.	100%	100%	100%	100%	100%
35.	100%	100%	100%	100%	100%
36.	100%	100%	100%	NA	100%
37.	85%	100%	92%	86%	92%
38.	100%	100%	100%	100%	100%
39.	100%	100%	100%	100%	100%

40.	100%	100%	100%	100%	100%
41.	18%	100%	31%	31%	40%
42.	80%	80%	80%	95%	92%
43.	100%	100%	100%	100%	100%
44.	NA	NA	NA	100%	100%
45.	0%	NA	NA	88%	88%
46.	75%	27%	40%	90%	57%
47.	100%	100%	100%	100%	100%
48.	50%	100%	67%	94%	94%
49.	100%	100%	100%	100%	100%
50.	80%	100%	89%	97%	97%
51.	NA	NA	NA	100%	100%
52.	100%	71%	83%	100%	87%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	100%	100%	100%	100%	100%
56.	100%	83%	91%	100%	94%
57.	100%	100%	100%	100%	100%
58.	NA	NA	NA	100%	100%
59.	NA	NA	NA	100%	100%
60.	100%	100%	100%	100%	100%
61.	100%	29%	45%	NA	29%
62.	100%	100%	100%	100%	100%
63.	100%	85%	92%	100%	93%
64.	100%	100%	100%	100%	100%
65.	100%	67%	80%	100%	94%
66.	100%	100%	100%	100%	100%
67.	100%	75%	86%	100%	88%
68.	100%	40%	57%	100%	79%
69.	100%	100%	100%	100%	100%
70.	100%	92%	96%	100%	96%
71.	80%	89%	84%	71%	81%
72.	100%	100%	100%	100%	100%
73.	100%	67%	80%	100%	88%
74.	100%	100%	100%	100%	100%

49 **Table S12.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
50 cropping: resolving footnote cross-references using LLaVA. ‘Not applicable (NA)’ is indicated
51 when there are no instances in the figure or when there is a division by zero error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	0%	NA	100%	85%
2.	0%	0%	NA	35%	29%
3.	100%	19%	32%	NA	19%
4.	NA	0%	NA	100%	94%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	0%	0%	NA	40%	27%
8.	NA	0%	NA	NA	0%
9.	0%	0%	NA	0%	0%
10.	NA	0%	NA	100%	88%
11.	NA	0%	NA	100%	71%
12.	0%	NA	NA	0%	0%
13.	NA	0%	NA	100%	85%
14.	NA	NA	NA	100%	100%
15.	0%	NA	NA	47%	47%
16.	NA	0%	NA	100%	74%
17.	0%	0%	NA	45%	42%
18.	0%	NA	NA	41%	41%
19.	NA	0%	NA	100%	86%
20.	NA	0%	NA	100%	67%
21.	NA	0%	NA	100%	67%
22.	NA	0%	NA	100%	80%
23.	NA	0%	NA	100%	70%
24.	0%	0%	NA	38%	20%
25.	0%	0%	NA	35%	25%
26.	NA	0%	NA	100%	85%
27.	100%	40%	57%	100%	50%
28.	NA	0%	NA	100%	50%
29.	NA	0%	NA	100%	75%
30.	100%	75%	86%	100%	83%
31.	NA	0%	NA	100%	77%
32.	0%	0%	NA	79%	65%
33.	0%	0%	NA	61%	40%
34.	0%	0%	NA	50%	41%
35.	100%	6%	12%	100%	63%
36.	NA	0%	NA	NA	0%
37.	0%	0%	NA	82%	46%
38.	NA	0%	NA	100%	91%
39.	NA	0%	NA	100%	50%

40.	NA	0%	NA	100%	82%
41.	0%	0%	NA	38%	33%
42.	0%	0%	NA	29%	23%
43.	43%	20%	27%	56%	33%
44.	0%	NA	NA	56%	56%
45.	NA	NA	NA	100%	100%
46.	0%	0%	NA	0%	0%
47.	0%	0%	NA	72%	54%
48.	0%	0%	NA	94%	89%
49.	0%	0%	NA	0%	0%
50.	5%	25%	8%	28%	27%
51.	NA	NA	NA	100%	100%
52.	0%	0%	NA	0%	0%
53.	NA	0%	NA	100%	79%
54.	0%	0%	NA	94%	71%
55.	0%	0%	NA	89%	73%
56.	NA	0%	NA	100%	67%
57.	0%	0%	NA	78%	67%
58.	0%	NA	NA	29%	29%
59.	0%	NA	NA	92%	92%
60.	0%	0%	NA	35%	33%
61.	NA	0%	NA	NA	0%
62.	NA	0%	NA	100%	79%
63.	0%	0%	NA	73%	39%
64.	NA	0%	NA	100%	76%
65.	NA	0%	NA	100%	83%
66.	NA	0%	NA	100%	67%
67.	NA	0%	NA	100%	50%
68.	NA	0%	NA	100%	64%
69.	NA	0%	NA	100%	63%
70.	100%	100%	100%	100%	100%
71.	NA	0%	NA	100%	44%
72.	0%	0%	NA	69%	47%
73.	0%	0%	NA	0%	0%
74.	0%	0%	NA	75%	71%

53 **Table S13.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
54 cropping: resolving footnote cross-references using InternVL. ‘Not applicable (NA)’ is
55 indicated when there are no instances in the figure or when there is a division by zero error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	0%	NA	100%	85%
2.	NA	0%	NA	100%	82%
3.	NA	0%	NA	NA	0%
4.	0%	0%	NA	67%	63%
5.	NA	NA	NA	100%	100%
6.	NA	NA	NA	100%	100%
7.	0%	0%	NA	80%	53%
8.	100%	3%	6%	NA	3%
9.	0%	0%	NA	0%	0%
10.	0%	0%	NA	0%	0%
11.	0%	0%	NA	80%	57%
12.	0%	NA	NA	0%	0%
13.	0%	0%	NA	55%	46%
14.	NA	NA	NA	100%	100%
15.	0%	NA	NA	0%	0%
16.	NA	0%	NA	100%	74%
17.	0%	0%	NA	0%	0%
18.	0%	NA	NA	41%	41%
19.	0%	0%	NA	32%	27%
20.	100%	80%	89%	100%	93%
21.	NA	0%	NA	100%	67%
22.	0%	0%	NA	63%	48%
23.	0%	0%	NA	29%	20%
24.	0%	0%	NA	0%	0%
25.	NA	0%	NA	100%	71%
26.	NA	0%	NA	100%	85%
27.	100%	53%	70%	100%	61%
28.	NA	0%	NA	100%	50%
29.	NA	0%	NA	100%	75%
30.	0%	0%	NA	0%	0%
31.	0%	0%	NA	59%	45%
32.	0%	0%	NA	36%	29%
33.	14%	8%	11%	74%	51%
34.	NA	0%	NA	100%	82%
35.	8%	13%	10%	8%	10%
36.	100%	100%	100%	NA	100%
37.	0%	0%	NA	0%	0%
38.	0%	0%	NA	80%	73%
39.	0%	0%	NA	0%	0%

40.	NA	0%	NA	100%	82%
41.	NA	0%	NA	100%	87%
42.	0%	0%	NA	67%	54%
43.	0%	0%	NA	0%	0%
44.	0%	NA	NA	0%	0%
45.	NA	NA	NA	100%	100%
46.	0%	0%	NA	50%	24%
47.	0%	0%	NA	67%	50%
48.	0%	0%	NA	0%	0%
49.	0%	0%	NA	4%	3%
50.	3%	25%	6%	3%	6%
51.	0%	NA	NA	67%	67%
52.	0%	0%	NA	88%	47%
53.	0%	0%	NA	32%	25%
54.	NA	0%	NA	100%	76%
55.	0%	0%	NA	56%	45%
56.	0%	0%	NA	17%	11%
57.	0%	0%	NA	44%	38%
58.	0%	NA	NA	0%	0%
59.	0%	NA	NA	77%	77%
60.	0%	0%	NA	0%	0%
61.	NA	0%	NA	NA	0%
62.	0%	0%	NA	0%	0%
63.	67%	15%	25%	93%	57%
64.	NA	0%	NA	100%	76%
65.	0%	0%	NA	80%	67%
66.	0%	0%	NA	0%	0%
67.	NA	0%	NA	100%	50%
68.	0%	0%	NA	78%	50%
69.	0%	0%	NA	80%	50%
70.	NA	0%	NA	100%	48%
71.	NA	0%	NA	100%	44%
72.	NA	0%	NA	100%	63%
73.	NA	0%	NA	100%	63%
74.	0%	0%	NA	0%	0%

57 **Table S14.** Recall scores of each figure with and without image cropping: identifying index-
 58 yield pairs using GPT-4V. ‘Not applicable (NA)’ is indicated when there are no instances in the
 59 figure or when there is a division by zero error.

Figure	Recall (feeding entire figure)	Recall (with image cropping)	Figure	Recall (feeding entire figure)	Recall (with image cropping)
1.	100%	100%	38.	100%	100%
2.	100%	100%	39.	100%	100%
3.	0%	100%	40.	100%	100%
4.	94%	100%	41.	100%	100%
5.	100%	100%	42.	85%	100%
6.	100%	100%	43.	100%	100%
7.	27%	100%	44.	83%	94%
8.	100%	100%	45.	100%	100%
9.	100%	100%	46.	38%	100%
10.	100%	100%	47.	100%	100%
11.	100%	100%	48.	100%	100%
12.	100%	100%	49.	100%	100%
13.	100%	100%	50.	79%	100%
14.	100%	100%	51.	100%	100%
15.	100%	100%	52.	100%	100%
16.	97%	89%	53.	100%	100%
17.	100%	96%	54.	100%	100%
18.	100%	100%	55.	100%	98%
19.	100%	100%	56.	100%	100%
20.	100%	100%	57.	100%	100%
21.	100%	100%	58.	100%	100%
22.	100%	100%	59.	100%	100%
23.	100%	100%	60.	95%	95%
24.	100%	100%	61.	45%	100%
25.	92%	100%	62.	89%	100%
26.	100%	100%	63.	100%	100%
27.	89%	100%	64.	97%	100%
28.	100%	95%	65.	100%	100%
29.	100%	100%	66.	100%	100%
30.	100%	100%	67.	100%	100%
31.	100%	95%	68.	100%	100%
32.	100%	100%	69.	100%	100%
33.	97%	100%	70.	100%	100%
34.	100%	100%	71.	100%	100%
35.	98%	100%	72.	89%	95%
36.	78%	100%	73.	100%	100%
37.	100%	100%	74.	95%	95%

60 **Table S15.** Precision, recall, specificity, and F1 scores of each figure without image cropping:
61 identifying additional substrate-specific information using GPT-4V. ‘Not applicable (NA)’ is
62 indicated when there are no instances in the figure or when there is a division by zero error.

Feeding entire figure					
Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	NA	NA	100%	100%
2.	NA	NA	NA	100%	100%
3.	100%	3%	6%	NA	3%
4.	100%	100%	100%	NA	100%
5.	100%	100%	100%	NA	100%
6.	100%	100%	100%	NA	100%
7.	100%	77%	87%	NA	77%
8.	100%	100%	100%	NA	100%
9.	96%	100%	98%	50%	96%
10.	NA	NA	NA	100%	100%
11.	100%	100%	100%	NA	100%
12.	100%	100%	100%	NA	100%
13.	100%	100%	100%	NA	100%
14.	100%	100%	100%	NA	100%
15.	100%	100%	100%	NA	100%
16.	100%	100%	100%	100%	100%
17.	100%	100%	100%	100%	100%
18.	100%	100%	100%	100%	100%
19.	NA	NA	NA	100%	100%
20.	NA	NA	NA	100%	100%
21.	NA	NA	NA	100%	100%
22.	NA	NA	NA	100%	100%
23.	NA	NA	NA	100%	100%
24.	100%	100%	100%	100%	100%
25.	NA	NA	NA	100%	100%
26.	NA	NA	NA	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	NA	NA	NA	100%	100%
30.	NA	NA	NA	100%	100%
31.	100%	100%	100%	100%	100%
32.	NA	NA	NA	100%	100%
33.	NA	NA	NA	100%	100%
34.	100%	100%	100%	100%	100%
35.	NA	NA	NA	100%	100%
36.	NA	NA	NA	100%	100%
37.	100%	100%	100%	100%	100%
38.	100%	100%	100%	100%	100%

39.	100%	100%	100%	100%	100%
40.	NA	NA	NA	100%	100%
41.	NA	NA	NA	100%	100%
42.	NA	NA	NA	100%	100%
43.	NA	NA	NA	100%	100%
44.	100%	100%	100%	NA	100%
45.	100%	100%	100%	NA	100%
46.	NA	NA	NA	100%	100%
47.	100%	100%	100%	NA	100%
48.	100%	100%	100%	100%	100%
49.	100%	100%	100%	100%	100%
50.	100%	100%	100%	100%	100%
51.	100%	100%	100%	100%	100%
52.	100%	100%	100%	100%	100%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	NA	NA	NA	100%	100%
56.	100%	100%	100%	NA	100%
57.	NA	NA	NA	100%	100%
58.	100%	100%	100%	NA	100%
59.	100%	100%	100%	NA	100%
60.	100%	100%	100%	NA	100%
61.	NA	NA	NA	100%	100%
62.	NA	NA	NA	100%	100%
63.	NA	NA	NA	100%	100%
64.	NA	NA	NA	100%	100%
65.	NA	NA	NA	100%	100%
66.	NA	NA	NA	100%	100%
67.	NA	NA	NA	100%	100%
68.	NA	NA	NA	100%	100%
69.	NA	NA	NA	100%	100%
70.	NA	NA	NA	100%	100%
71.	NA	NA	NA	100%	100%
72.	NA	NA	NA	100%	100%
73.	NA	NA	NA	100%	100%
74.	100%	100%	100%	NA	100%

64 **Table S16.** Precision, recall, specificity, and F1 scores of each figure with image cropping:
65 identifying additional substrate-specific information using GPT-4V. ‘Not applicable (NA)’ is
66 indicated when there are no instances in the figure or when there is a division by zero error.

With image cropping					
Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	NA	NA	100%	100%
2.	NA	NA	NA	100%	100%
3.	100%	100%	100%	NA	100%
4.	100%	100%	100%	NA	100%
5.	100%	100%	100%	NA	100%
6.	100%	100%	100%	NA	100%
7.	100%	100%	100%	NA	100%
8.	100%	100%	100%	NA	100%
9.	100%	100%	100%	NA	100%
10.	NA	NA	NA	100%	100%
11.	100%	100%	100%	NA	100%
12.	100%	100%	100%	NA	100%
13.	100%	92%	96%	NA	92%
14.	100%	100%	100%	NA	100%
15.	100%	100%	100%	NA	100%
16.	100%	100%	100%	100%	100%
17.	75%	100%	86%	95%	96%
18.	100%	100%	100%	100%	100%
19.	0%	NA	NA	95%	95%
20.	NA	NA	NA	100%	100%
21.	NA	NA	NA	100%	100%
22.	NA	NA	NA	100%	100%
23.	NA	NA	NA	100%	100%
24.	100%	100%	100%	100%	100%
25.	NA	NA	NA	100%	100%
26.	NA	NA	NA	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	NA	NA	NA	100%	100%
30.	0%	NA	NA	83%	83%
31.	100%	100%	100%	100%	100%
32.	NA	NA	NA	100%	100%
33.	NA	NA	NA	100%	100%
34.	100%	100%	100%	100%	100%
35.	NA	NA	NA	100%	100%
36.	NA	NA	NA	100%	100%
37.	100%	25%	40%	100%	92%
38.	100%	100%	100%	100%	100%

39.	100%	100%	100%	100%	100%
40.	NA	NA	NA	100%	100%
41.	NA	NA	NA	100%	100%
42.	NA	NA	NA	100%	100%
43.	NA	NA	NA	100%	100%
44.	100%	94%	97%	NA	94%
45.	100%	100%	100%	NA	100%
46.	NA	NA	NA	100%	100%
47.	100%	100%	100%	NA	100%
48.	100%	100%	100%	100%	100%
49.	67%	100%	80%	97%	97%
50.	100%	100%	100%	100%	100%
51.	100%	100%	100%	100%	100%
52.	100%	100%	100%	100%	100%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	NA	NA	NA	100%	100%
56.	100%	100%	100%	NA	100%
57.	NA	NA	NA	100%	100%
58.	100%	100%	100%	NA	100%
59.	100%	100%	100%	NA	100%
60.	100%	100%	100%	NA	100%
61.	NA	NA	NA	100%	100%
62.	NA	NA	NA	100%	100%
63.	NA	NA	NA	100%	100%
64.	NA	NA	NA	100%	100%
65.	NA	NA	NA	100%	100%
66.	NA	NA	NA	100%	100%
67.	NA	NA	NA	100%	100%
68.	NA	NA	NA	100%	100%
69.	NA	NA	NA	100%	100%
70.	NA	NA	NA	100%	100%
71.	NA	NA	NA	100%	100%
72.	NA	NA	NA	100%	100%
73.	NA	NA	NA	100%	100%
74.	100%	100%	100%	NA	100%

68 **Table S17.** Recall scores of each figure with image cropping: identifying index-yield pairs
 69 using Gemini. ‘Not applicable (NA)’ is indicated when there are no instances in the figure or
 70 when there is a division by zero error.

Figure	Recall	Figure	Recall
1.	92%	38.	100%
2.	100%	39.	100%
3.	100%	40.	100%
4.	100%	41.	100%
5.	100%	42.	100%
6.	100%	43.	100%
7.	100%	44.	100%
8.	100%	45.	100%
9.	100%	46.	100%
10.	100%	47.	100%
11.	100%	48.	100%
12.	100%	49.	100%
13.	100%	50.	100%
14.	100%	51.	100%
15.	100%	52.	100%
16.	100%	53.	100%
17.	100%	54.	100%
18.	100%	55.	100%
19.	100%	56.	100%
20.	100%	57.	100%
21.	100%	58.	100%
22.	100%	59.	100%
23.	100%	60.	100%
24.	100%	61.	100%
25.	75%	62.	100%
26.	100%	63.	100%
27.	100%	64.	100%
28.	100%	65.	100%
29.	100%	66.	100%
30.	100%	67.	100%
31.	100%	68.	100%
32.	100%	69.	100%
33.	100%	70.	100%
34.	100%	71.	100%
35.	100%	72.	100%
36.	100%	73.	100%
37.	97%	74.	100%

71 **Table S18.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
72 cropping: identifying additional substrate-specific information using Gemini. ‘Not applicable
73 (NA)’ is indicated when there are no instances in the figure or when there is a division by zero
74 error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	NA	NA	100%	100%
2.	NA	NA	NA	100%	100%
3.	100%	100%	100%	NA	100%
4.	100%	100%	100%	NA	100%
5.	100%	100%	100%	NA	100%
6.	100%	100%	100%	NA	100%
7.	100%	100%	100%	NA	100%
8.	100%	52%	68%	NA	52%
9.	100%	100%	100%	NA	100%
10.	NA	NA	NA	100%	100%
11.	100%	100%	100%	NA	100%
12.	100%	100%	100%	NA	100%
13.	100%	100%	100%	NA	100%
14.	100%	100%	100%	NA	100%
15.	100%	100%	100%	NA	100%
16.	100%	50%	67%	100%	97%
17.	100%	100%	100%	100%	100%
18.	100%	100%	100%	100%	100%
19.	NA	NA	NA	100%	100%
20.	NA	NA	NA	100%	100%
21.	NA	NA	NA	100%	100%
22.	NA	NA	NA	100%	100%
23.	NA	NA	NA	100%	100%
24.	100%	100%	100%	100%	100%
25.	NA	NA	NA	100%	100%
26.	NA	NA	NA	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	NA	NA	NA	100%	100%
30.	NA	NA	NA	100%	100%
31.	100%	100%	100%	100%	100%
32.	NA	NA	NA	100%	100%
33.	NA	NA	NA	100%	100%
34.	NA	NA	NA	100%	100%
35.	NA	NA	NA	100%	100%
36.	NA	NA	NA	100%	100%
37.	100%	100%	100%	100%	100%

38.	100%	100%	100%	100%	100%
39.	100%	100%	100%	100%	100%
40.	NA	NA	NA	100%	100%
41.	NA	NA	NA	100%	100%
42.	NA	NA	NA	100%	100%
43.	NA	NA	NA	100%	100%
44.	100%	100%	100%	NA	100%
45.	100%	100%	100%	NA	100%
46.	NA	NA	NA	100%	100%
47.	100%	100%	100%	NA	100%
48.	100%	100%	100%	100%	100%
49.	100%	100%	100%	100%	100%
50.	100%	100%	100%	100%	100%
51.	100%	100%	100%	100%	100%
52.	100%	100%	100%	100%	100%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	NA	NA	NA	100%	100%
56.	100%	100%	100%	NA	100%
57.	NA	NA	NA	100%	100%
58.	100%	100%	100%	NA	100%
59.	100%	100%	100%	NA	100%
60.	100%	100%	100%	NA	100%
61.	NA	NA	NA	100%	100%
62.	NA	NA	NA	100%	100%
63.	NA	NA	NA	100%	100%
64.	NA	NA	NA	100%	100%
65.	NA	NA	NA	100%	100%
66.	NA	NA	NA	100%	100%
67.	NA	NA	NA	100%	100%
68.	NA	NA	NA	100%	100%
69.	NA	NA	NA	100%	100%
70.	NA	NA	NA	100%	100%
71.	NA	NA	NA	100%	100%
72.	NA	NA	NA	100%	100%
73.	NA	NA	NA	100%	100%
74.	100%	33%	50%	NA	33%

76 **Table S19.** Recall scores of each figure with image cropping: identifying index-yield pairs
 77 using Claude 3. ‘Not applicable (NA)’ is indicated when there are no instances in the figure or
 78 when there is a division by zero error.

Figure	Recall	Figure	Recall
1.	100%	38.	100%
2.	100%	39.	100%
3.	100%	40.	100%
4.	100%	41.	100%
5.	100%	42.	100%
6.	100%	43.	100%
7.	100%	44.	83%
8.	100%	45.	100%
9.	100%	46.	67%
10.	100%	47.	100%
11.	100%	48.	100%
12.	100%	49.	100%
13.	100%	50.	100%
14.	100%	51.	100%
15.	100%	52.	87%
16.	100%	53.	100%
17.	100%	54.	100%
18.	100%	55.	100%
19.	100%	56.	100%
20.	100%	57.	100%
21.	100%	58.	93%
22.	100%	59.	100%
23.	100%	60.	100%
24.	100%	61.	95%
25.	92%	62.	100%
26.	100%	63.	100%
27.	100%	64.	100%
28.	95%	65.	100%
29.	100%	66.	100%
30.	100%	67.	100%
31.	100%	68.	100%
32.	100%	69.	75%
33.	97%	70.	100%
34.	100%	71.	100%
35.	100%	72.	100%
36.	100%	73.	100%
37.	100%	74.	100%

79 **Table S20.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
80 cropping: identifying additional substrate-specific information using Claude 3. ‘Not applicable
81 (NA)’ is indicated when there are no instances in the figure or when there is a division by zero
82 error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	NA	NA	NA	100%	100%
2.	NA	NA	NA	100%	100%
3.	100%	100%	100%	NA	100%
4.	100%	100%	100%	NA	100%
5.	100%	100%	100%	NA	100%
6.	100%	100%	100%	NA	100%
7.	100%	100%	100%	NA	100%
8.	100%	100%	100%	NA	100%
9.	100%	100%	100%	NA	100%
10.	NA	NA	NA	100%	100%
11.	100%	100%	100%	NA	100%
12.	100%	100%	100%	NA	100%
13.	100%	100%	100%	NA	100%
14.	100%	100%	100%	NA	100%
15.	100%	100%	100%	NA	100%
16.	100%	100%	100%	100%	100%
17.	100%	100%	100%	100%	100%
18.	100%	100%	100%	100%	100%
19.	NA	NA	NA	100%	100%
20.	NA	NA	NA	100%	100%
21.	NA	NA	NA	100%	100%
22.	NA	NA	NA	100%	100%
23.	NA	NA	NA	100%	100%
24.	100%	100%	100%	100%	100%
25.	NA	NA	NA	100%	100%
26.	NA	NA	NA	100%	100%
27.	100%	100%	100%	100%	100%
28.	100%	100%	100%	100%	100%
29.	NA	NA	NA	100%	100%
30.	NA	NA	NA	100%	100%
31.	100%	100%	100%	100%	100%
32.	NA	NA	NA	100%	100%
33.	NA	NA	NA	100%	100%
34.	NA	NA	NA	100%	100%
35.	NA	NA	NA	100%	100%
36.	NA	NA	NA	100%	100%
37.	100%	100%	100%	100%	100%

38.	100%	100%	100%	100%	100%
39.	100%	100%	100%	100%	100%
40.	NA	NA	NA	100%	100%
41.	NA	NA	NA	100%	100%
42.	NA	NA	NA	100%	100%
43.	NA	NA	NA	100%	100%
44.	100%	100%	100%	NA	100%
45.	100%	100%	100%	NA	100%
46.	0%	NA	NA	67%	67%
47.	100%	100%	100%	NA	100%
48.	100%	100%	100%	100%	100%
49.	100%	100%	100%	100%	100%
50.	100%	100%	100%	100%	100%
51.	100%	100%	100%	100%	100%
52.	100%	100%	100%	100%	100%
53.	100%	100%	100%	100%	100%
54.	100%	100%	100%	100%	100%
55.	NA	NA	NA	100%	100%
56.	100%	100%	100%	NA	100%
57.	NA	NA	NA	100%	100%
58.	100%	93%	96%	NA	93%
59.	100%	100%	100%	NA	100%
60.	100%	100%	100%	NA	100%
61.	NA	NA	NA	100%	100%
62.	NA	NA	NA	100%	100%
63.	NA	NA	NA	100%	100%
64.	NA	NA	NA	100%	100%
65.	NA	NA	NA	100%	100%
66.	NA	NA	NA	100%	100%
67.	NA	NA	NA	100%	100%
68.	NA	NA	NA	100%	100%
69.	NA	NA	NA	100%	100%
70.	NA	NA	NA	100%	100%
71.	NA	NA	NA	100%	100%
72.	NA	NA	NA	100%	100%
73.	NA	NA	NA	100%	100%
74.	100%	100%	100%	NA	100%

84 **Table S21.** Recall scores of each figure with image cropping: identifying index-yield pairs
 85 using LLaVA. ‘Not applicable (NA)’ is indicated when there are no instances in the figure or
 86 when there is a division by zero error.

Figure	Recall	Figure	Recall
1.	77%	38.	82%
2.	29%	39.	88%
3.	71%	40.	100%
4.	94%	41.	33%
5.	100%	42.	85%
6.	100%	43.	54%
7.	43%	44.	39%
8.	68%	45.	88%
9.	75%	46.	0%
10.	81%	47.	54%
11.	69%	48.	89%
12.	42%	49.	0%
13.	8%	50.	88%
14.	88%	51.	100%
15.	93%	52.	0%
16.	71%	53.	83%
17.	88%	54.	76%
18.	94%	55.	73%
19.	41%	56.	89%
20.	100%	57.	76%
21.	83%	58.	29%
22.	88%	59.	92%
23.	100%	60.	33%
24.	60%	61.	13%
25.	25%	62.	93%
26.	90%	63.	43%
27.	67%	64.	97%
28.	75%	65.	100%
29.	88%	66.	86%
30.	100%	67.	100%
31.	73%	68.	79%
32.	65%	69.	100%
33.	66%	70.	100%
34.	94%	71.	81%
35.	93%	72.	100%
36.	56%	73.	0%
37.	79%	74.	76%

87 **Table S22.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
88 cropping: identifying additional substrate-specific information using LLaVA. ‘Not applicable
89 (NA)’ is indicated when there are no instances in the figure or when there is a division by zero
90 error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	0%	NA	NA	77%	77%
2.	0%	NA	NA	29%	29%
3.	NA	0%	NA	NA	0%
4.	100%	94%	97%	NA	94%
5.	100%	100%	100%	NA	100%
6.	100%	100%	100%	NA	100%
7.	100%	27%	42%	NA	27%
8.	100%	68%	81%	NA	68%
9.	100%	36%	53%	NA	36%
10.	NA	NA	NA	100%	100%
11.	100%	66%	79%	NA	66%
12.	100%	42%	59%	NA	42%
13.	100%	85%	92%	NA	85%
14.	100%	100%	100%	NA	100%
15.	100%	70%	82%	NA	70%
16.	NA	0%	NA	100%	95%
17.	15%	67%	25%	48%	50%
18.	NA	0%	NA	100%	88%
19.	0%	NA	NA	86%	86%
20.	NA	NA	NA	100%	100%
21.	NA	NA	NA	100%	100%
22.	NA	NA	NA	100%	100%
23.	NA	NA	NA	100%	100%
24.	67%	100%	80%	43%	73%
25.	0%	NA	NA	25%	25%
26.	NA	NA	NA	100%	100%
27.	NA	0%	NA	100%	22%
28.	NA	0%	NA	100%	60%
29.	NA	NA	NA	100%	100%
30.	NA	NA	NA	100%	100%
31.	100%	100%	100%	100%	100%
32.	0%	NA	NA	94%	94%
33.	0%	NA	NA	69%	69%
34.	NA	NA	NA	100%	100%
35.	0%	NA	NA	93%	93%
36.	NA	NA	NA	100%	100%
37.	0%	0%	NA	89%	79%

38.	100%	88%	93%	100%	91%
39.	100%	100%	100%	100%	100%
40.	NA	NA	NA	100%	100%
41.	0%	NA	NA	33%	33%
42.	0%	NA	NA	85%	85%
43.	0%	NA	NA	38%	38%
44.	NA	0%	NA	NA	0%
45.	100%	100%	100%	NA	100%
46.	0%	NA	NA	76%	76%
47.	NA	0%	NA	NA	0%
48.	50%	100%	67%	94%	94%
49.	0%	0%	NA	0%	0%
50.	60%	100%	75%	85%	88%
51.	100%	100%	100%	100%	100%
52.	0%	0%	NA	0%	0%
53.	100%	40%	57%	100%	88%
54.	0%	0%	NA	80%	76%
55.	0%	NA	NA	73%	73%
56.	100%	100%	100%	NA	100%
57.	0%	NA	NA	76%	76%
58.	100%	29%	44%	NA	29%
59.	100%	92%	96%	NA	92%
60.	100%	29%	44%	NA	29%
61.	0%	NA	NA	13%	13%
62.	0%	NA	NA	93%	93%
63.	0%	NA	NA	79%	79%
64.	0%	NA	NA	97%	97%
65.	NA	NA	NA	100%	100%
66.	0%	NA	NA	95%	95%
67.	NA	NA	NA	100%	100%
68.	0%	NA	NA	79%	79%
69.	NA	NA	NA	100%	100%
70.	NA	NA	NA	100%	100%
71.	0%	NA	NA	81%	81%
72.	NA	NA	NA	100%	100%
73.	0%	NA	NA	0%	0%
74.	100%	76%	86%	NA	76%

92 **Table S23.** Recall scores of each figure with image cropping: identifying index-yield pairs
 93 using InternVL. ‘Not applicable (NA)’ is indicated when there are no instances in the figure or
 94 when there is a division by zero error.

Figure	Recall	Figure	Recall
1.	77%	38.	82%
2.	0%	39.	0%
3.	52%	40.	100%
4.	63%	41.	100%
5.	100%	42.	73%
6.	100%	43.	17%
7.	13%	44.	22%
8.	6%	45.	100%
9.	0%	46.	24%
10.	13%	47.	75%
11.	57%	48.	39%
12.	17%	49.	3%
13.	69%	50.	45%
14.	100%	51.	67%
15.	7%	52.	80%
16.	79%	53.	25%
17.	0%	54.	81%
18.	41%	55.	52%
19.	27%	56.	22%
20.	93%	57.	38%
21.	100%	58.	0%
22.	48%	59.	77%
23.	20%	60.	19%
24.	33%	61.	0%
25.	46%	62.	25%
26.	95%	63.	64%
27.	61%	64.	76%
28.	60%	65.	83%
29.	63%	66.	0%
30.	50%	67.	88%
31.	45%	68.	86%
32.	29%	69.	63%
33.	54%	70.	88%
34.	94%	71.	88%
35.	54%	72.	68%
36.	44%	73.	88%
37.	0%	74.	0%

95 **Table S24.** Precision, recall, F1 scores, specificity, and accuracy of each figure with image
 96 cropping: identifying additional substrate-specific information using InternVL. ‘Not applicable
 97 (NA)’ is indicated when there are no instances in the figure or when there is a division by zero
 98 error.

Figure	Precision	Recall	F1	Specificity	Accuracy
1.	0%	NA	NA	85%	85%
2.	0%	NA	NA	86%	86%
3.	100%	32%	49%	NA	32%
4.	100%	63%	77%	NA	63%
5.	100%	86%	92%	NA	86%
6.	100%	100%	100%	NA	100%
7.	100%	13%	24%	NA	13%
8.	100%	6%	12%	NA	6%
9.	NA	0%	NA	NA	0%
10.	0%	NA	NA	0%	0%
11.	100%	54%	70%	NA	54%
12.	100%	17%	29%	NA	17%
13.	100%	69%	82%	NA	69%
14.	100%	88%	93%	NA	88%
15.	NA	0%	NA	NA	0%
16.	NA	0%	NA	100%	95%
17.	0%	0%	NA	0%	0%
18.	0%	0%	NA	33%	29%
19.	0%	NA	NA	0%	0%
20.	0%	NA	NA	93%	93%
21.	NA	NA	NA	100%	100%
22.	0%	NA	NA	48%	48%
23.	0%	NA	NA	20%	20%
24.	25%	14%	18%	63%	40%
25.	0%	NA	NA	75%	75%
26.	0%	NA	NA	95%	95%
27.	50%	29%	36%	0%	22%
28.	100%	50%	67%	100%	80%
29.	0%	NA	NA	63%	63%
30.	0%	NA	NA	0%	0%
31.	0%	0%	NA	48%	45%
32.	0%	NA	NA	12%	12%
33.	0%	NA	NA	54%	54%
34.	0%	NA	NA	94%	94%
35.	0%	NA	NA	0%	0%
36.	0%	NA	NA	44%	44%
37.	0%	0%	NA	0%	0%

38.	100%	38%	55%	100%	55%
39.	0%	0%	NA	0%	0%
40.	NA	NA	NA	100%	100%
41.	NA	NA	NA	100%	100%
42.	0%	NA	NA	73%	73%
43.	0%	NA	NA	0%	0%
44.	NA	0%	NA	NA	0%
45.	NA	0%	NA	NA	0%
46.	0%	NA	NA	0%	0%
47.	100%	17%	29%	NA	17%
48.	0%	0%	NA	0%	0%
49.	0%	0%	NA	3%	3%
50.	5%	17%	8%	33%	30%
51.	83%	63%	71%	75%	67%
52.	75%	43%	55%	88%	67%
53.	0%	0%	NA	21%	17%
54.	0%	0%	NA	75%	71%
55.	0%	NA	NA	52%	52%
56.	100%	22%	36%	NA	22%
57.	0%	NA	NA	38%	38%
58.	NA	0%	NA	NA	0%
59.	NA	0%	NA	NA	0%
60.	100%	19%	32%	NA	19%
61.	0%	NA	NA	0%	0%
62.	0%	NA	NA	25%	25%
63.	0%	NA	NA	64%	64%
64.	0%	NA	NA	76%	76%
65.	0%	NA	NA	83%	83%
66.	0%	NA	NA	0%	0%
67.	0%	NA	NA	88%	88%
68.	0%	NA	NA	86%	86%
69.	0%	NA	NA	63%	63%
70.	NA	NA	NA	100%	100%
71.	NA	NA	NA	100%	100%
72.	0%	NA	NA	68%	68%
73.	0%	NA	NA	88%	88%
74.	NA	0%	NA	NA	0%

101 References

- 102 1 Y. Qiu, J. Struwe, T. H. Meyer, J. C. A. Oliveira, L. Ackermann, Catalyst- and Reagent-Free
103 Electrochemical Azole C–H Amination, *Chem. Eur. J.* 2018, **24**, 12784.
- 104 2 E. Yu, H. Kim, C.-M. Park, Metal- and Oxidant-Free Electrosynthesis of Heterocycles from 1,2-
105 Diarylalkene Derivatives, *Adv. Synth. Catal.* 2022, **364**, 4088.
- 106 3 Y.-F. Tang, D. Yang, Y.-N. Zhao, J.-F. Lv, Z. Guan, Y.-H. He, Electrochemical radical–polar crossover
107 diesterification of alkenes with carboxylic acids, *Green Chem.*, 2023, **25**, 1540-1545.
- 108 4 OpenAI, et al. GPT-4 Technical Report. arXiv, 2023, DOI: 10.48550/arXiv.2303.08774.
- 109 5 Gemini Team et al., Gemini: A Family of Highly Capable Multimodal Models. arXiv, 2023, DOI:
110 10.48550/arXiv.2312.11805.
- 111 6 Anthropic, The Claude 3 Model Family: Opus, Sonnet, Haiku, 2024, [https://www-](https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf)
112 [cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf](https://www-cdn.anthropic.com/de8ba9b01c9ab7cbabf5c33b80b7bbc618857627/Model_Card_Claude_3.pdf).
- 113 7 Z. Chen, J. Wu, W. Wang, W. Su, G. Chen, S. Xing, M. Zhong, Q. Zhang, X. Zhu, L. Lu, B. Li, P.
114 Luo, T. Lu, Y. Qiao and J. Dai, InternVL: Scaling up Vision Foundation Models and Aligning for
115 Generic Visual-Linguistic Tasks. *arXiv*, 2024, DOI: 10.48550/arXiv.2312.14238.
- 116 8 H. Liu, C. Li, Q. Wu and Y. J. Lee, Visual Instruction Tuning. *arXiv*, 2023, DOI:
117 10.48550/arXiv.2304.08485.
- 118 9 H. Liu, C. Li, Y. Li and Y. J. Lee, Improved Baselines with Visual Instruction Tuning. *arXiv*, 2023,
119 DOI: 10.48550/arXiv.2310.03744.
- 120 10 A. Halder, S. Ghosh and S. De Sarkar, Transition Metal and Base-Free Electro-Oxidative
121 Regioselective Trifluoromethylation of Imidazo[1,2-*a*]pyridines, *Asian J Org Chem*, 2023, **12**,
122 e202300294.
- 123 11 Y.-F. Tan, D. Yang, Y.-H. Yang, J.-F. Lv, L.-X. Zong, Z. Guan and Y.-H. He, Electricity-driven 1,4-
124 alkoxydimerization of alkenes *via* radical–polar crossover, *Green Chem.*, 2023, **25**, 9388–9393.
- 125 12 X. Zhou, L. Guo, H. Zhang, R. Y. Xia, C. Yang and W. Xia, Nickel-Catalyzed Reductive Acylation
126 of Carboxylic Acids with Alkyl Halides and *N*-Hydroxyphthalimide Esters Enabled by
127 Electrochemical Process, *Adv Synth Catal*, 2022, **364**, 1526–1531.
- 128 13 A. O. Terent'ev, O. M. Mulina, V. D. Parshin, V. A. Kokorekin and G. I. Nikishin, Electrochemically
129 induced oxidative S–O coupling: synthesis of sulfonates from sulfonyl hydrazides and *N*-
130 hydroxyimides or *N*-hydroxybenzotriazoles, *Org. Biomol. Chem.*, 2019, **17**, 3482–3488.
- 131 14 D. Wilary and J. M. Cole, ReactionDataExtractor 2.0: A Deep Learning Approach for Data
132 Extraction from Chemical Reaction Schemes. *J. Chem. Inf. Model.* 2023, **63**, 6053–6067.
- 133 15 Y. Qian, J. Guo, Z. Tu, C. W. Coley and R. Barzilay, RxnScribe: A Sequence Generation Model for
134 Reaction Diagram Parsing, *J. Chem. Inf. Model.*, 2023, **63**, 4030–4041.
- 135 16 K. Liang, D. Zhang, Y. Su, L. Lu, J. Hu, Y.-H. Chen, X. Zhang, A. Lei and H. Yi, Fragile intermediate
136 identification and reactivity elucidation in electrochemical oxidative α -C(sp³)-H functionalization
137 of tertiary amines, *Chem. Sci.*, 2023, **14**, 4152–4157.