Electronic Supplementary Information to: Automated LLM based Extraction of Standardized Synthesis Procedures: an All-Domain, Zero-Shot Approach

Daniel P. Costa[1], Matteo Manica[2], Teodoro Laino[2,3], M. Filipa Ribeiro[1], and Pedro S. F. Mendes[1]

[1]Chemical Engineering Department Centro de Química Estrutural - Institute of Molecular Sciences, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal
[2]IBM Research Europe, Saümerstrasse 4, 8803 Rüschlikon, Switzerland
[3]NCCR Catalysis, Switzerland

October 14, 2025

Contents

1 Preliminary Tests

We performed our preliminary tests on 10 procedures about zeolite synthesis using four different models: microsoft/Phi-3-medium-4k-instruct, microsoft/Phi-3-mini-4k-instruct, meta-llama/Meta-Llama-3-8B-Instruct, openchat/openchat-3.6-8b-20240522. You can find multiple examples of the four different LLMs responses without defining any action in the Figures 1.1 to 1.4.

Phi-3-mini-4k-instruct

Action Graph for Synthesis Procedure:
1. Convert H+-(or partial Na+/H+)-CHA into Na+ form:
  a. Mix 150 mL of 0.5 M NaCl (>99%, VWR) solution with 1 g of solids
  b. Stir the mixture at ambient conditions for 16 hours
  c. Centrifuge the samples and decant the liquid
2. Repeat the ion exchange process two more times with the following exchange times:
  a. 16 hours b. 8 hours c. 16 hours
3. Separate the solids by centrifugation and wash at least three times with deionized water (>150 mL/(g solids), 18.2 MΩ)
4. Oven-dry the Na+-form solids at 100 °C
5. Convert Na-exchanged zeolite into (Co,Na)-form:
  a. Mix 150 mL of 0.05 M Co(NO3)2 (>99%, Acros Organics) solution with 1 g of Na-exchanged zeolite
  b. Stir the mixture at ambient conditions for 16 hours
  c. Centrifuge the samples and decant the liquid
6. Repeat the ion exchange process two more times with the following exchange times:
  a. 16 hours b. 8 hours c. 16 hours
7. Separate the solids by centrifugation and wash at least three times with deionized water (>150 mL/(g solids), 18.2 MΩ)
8. Oven-dry the (Co,Na)-form zeolite at 100 °C overnight

openchat-3.6-8b-20240522

1. Convert H+-(or partial Na+/H+)-CHA into the Na+ form:
  a. Add 150 mL of an aqueous 0.5 M NaCl solution per g of solids.
  b. Stir at ambient conditions for 16 h.
  c. Centrifugate and decant the liquid.
  d. Repeat steps a-c three times with 16, 8, and 16 h of exchange times.
2. Separate solids and wash with deionized water:
  a. Centrifugate the solids.
  b. Wash at least three times with deionized water (>150 mL/(g solids), 18.2 MΩ).
3. Dry Na+-form solids:
  a. Oven-dry at 100 °C.
4. Perform Co-exchange procedure:
  a. Add 150 mL of an aqueous 0.05 M Co(NO3)2 solution per g Na-exchanged zeolite.
  b. Stir at ambient conditions for 16, 8, and 16 h exchange times.
  c. Centrifugate and wash the solids at least three times (>150 mL/(g solids), 18.2 MΩ).
5. Dry (Co,Na)-form zeolite:
  a. Oven-dry at 100 °C overnight.

Figure ESI 1.1: LLMs response examples showing how Centrifugate is presented inside a main action Separate

Phi-3-mini-4k-instruct

Action Graph:

1. Set up a 250 cm^3 round-baker with a reflux-condenser.
2. Fill the reflux-condenser with 181 cm^3 of aqueous 0.2m NaOH.
3. Heat the mixture to 338 K.
4. Add 6 g of zeolite sample to the mixture.
5. Stir magnetically at 500 rpm for 30 minutes.
6. Quench the suspension in an ice–water mixture.
7. Filter the resulting solid.
8. Wash the solid.
9. Dry the solid at 373 K for 12 hours.
10. Calcine the dried solid.
11. Perform three successive ion exchanges in a 0.1m NH4NO3 aqueous solution.
12. Calcine the sample after ion exchange.
13. Set up a 20 cm^3 glass tube with 17 cm^3 of aqueous 0.1m HCl.
14. Heat the mixture to 343 K.
15. Add 0.55 g of zeolite sample to the mixture.
16. Stir magnetically at 500 rpm for 6 hours.
17. Filter the solid.
18. Wash the solid.
19. Dry the solid.
20. Calcine the dried solid.

Meta-Llama-3-8B-Instruct

Here is the action graph:
Step 1: Mixing
* 1 g of prepared catalyst (F-HS) is mixed with 1 M KOH solution in a mass ratio of 1:10
Step 2: Ion-Exchange
* The slurry mixture is refluxed at 60 °C for 24 h using a stirring speed of 800 rpm
Step 3: Filtration
* The modified product is filtered
Step 4: Washing
* The filtered product is washed
Step 5: Drying
* The washed product is dried overnight at 120 °C
Step 6: Calcination
* The dried solid product is calcined in a muffle furnace at 550 °C for 1.5 h

openchat-3.6-8b-20240522

1. Prepare 1 g of the catalyst (F-HS).
2. Mix the catalyst with 1 M KOH solution in a mass ratio of 1:10.
3. Ion-exchange the slurry mixture by refluxing at 60 °C for 24 h with a stirring speed of 800 rpm.
4. Filter, wash, and dry the modified product overnight at 120 °C.
5. Calcinate the dried solid product in a muffle furnace at 550 °C for 1.5 h.

Figure ESI 1.2: LLMs response examples that show that Reflux and Quench are linked with a change of temperature

Phi-3-mini-4k-instruct

1. Stir for 2 hours
2. Hydrothermal crystallization at 1608C for 5 days
3. Wash
4. Filtrate
5. Dry at 1108C for 12 hours
6. Calcine in air at 5508C for 10 hours
7. Crush
8. Sieve to 40–60 mesh granules

Phi-3-medium-4k-instruct

1. Prepare Solution A:
   - Mix NaOH (1.36 g), sodium aluminate (0.227 g), template, and distilled water (14.95 g)
2. Prepare Solution B:
   - Mix silicic acid (22.91 g) and distilled water (59.81 g)
3. Add Solution A to Solution B:
   - Add dropwise over 3 hours with vigorous stirring
4. Adjust gel pH:
   - Add concentrated H2SO4 to the mixture
5. Crystallization:
   - Place the mixture in stainless-steel Teflon-lined autoclaves
   - Heat at 180 °C for 48 hours without stirring under autogenous pressure
6. Recover solid product:
   - Filter the mixture and wash several times with deionized water until pH neutral
7. Dry the product:
   - Dry overnight at 105 °C
8. Calcination:
   - Calcine catalyst samples in a muffle furnace under a flow of air at 530 °C for 12 hours

Meta-Llama-3-8B-Instruct

Action Graph:
Step 1: Na+ Ion Exchange
1.1. Add 150 mL of 0.5 M NaCl solution per g of solids
1.2. Stir at ambient conditions for 16 hours
1.3. Centrifuge and decant the liquid
1.4. Repeat steps 1.1-1.3 three times with 16, 8, and 16 hour exchange times
Step 2: Washing
2.1. Wash the solids at least three times with deionized water (>150 mL/(g solids), 18.2 MΩ)
2.2. Centrifuge and decant the liquid after each washing step
Step 3: Drying
3.1. Oven-dry the Na+-form solids at 100 °C
Step 4: Co-Exchange
4.1. Add 150 mL of 0.05 M Co(NO3)2 solution per g Na-exchanged zeolite
4.2. Stir at ambient conditions for 16 hours
4.3. Centrifuge and decant the liquid
4.4. Repeat steps 4.1-4.3 three times with 16, 8, and 16 hour exchange times
Step 5: Washing
5.1. Wash the solids at least three times with deionized water (>150 mL/(g solids), 18.2 MΩ)
5.2. Centrifuge and decant the liquid after each washing step
Step 6: Drying
6.1. Oven-dry the (Co,Na)-form zeolite at 100 °C overnight

Figure ESI 1.3: LLMs response examples that show the need to create a Sieve, Repeat and NewSolution actions.

Phi-3-medium-4k-instruct

1. Prepare Solution A:
  - Mix NaOH (1.36 g), sodium aluminate (0.227 g), template, and distilled water (14.95 g)
2. Prepare Solution B:
  - Mix silicic acid (22.91 g) and distilled water (59.81 g)
3. Add Solution A to Solution B:
  - Add dropwise over 3 hours with vigorous stirring
4. Adjust gel pH:
  - Add concentrated H2SO4 to the mixture
5. Crystallization:
  - Place the mixture in stainless-steel Teflon-lined autoclaves
  - Heat at 180 °C for 48 hours without stirring under autogenous pressure
6. Recover solid product:
  - Filter the mixture and wash several times with deionized water until pH neutral
7. Dry the product:
  - Dry overnight at 105 °C
8. Calcination:
  - Calcine catalyst samples in a muffle furnace under a flow of air at 530 °C for 12 hours

Meta-Llama-3-8B-Instruct

1. Stirring
  - Time: 2 h
2. Hydrothermal crystallization
  - Temperature: 160 C
  - Time: 5 days
3. Washing
4. Filtration
5. Drying
  - Temperature: 1108C
  - Time: 12 h
6. Calcination
  - Temperature: 5508C
  - Time: 10 h
7. Crushing
8. Sieving
  - Granule size: 40-60 mesh
9. Catalyst preparation for further use

Figure ESI 1.4: LLMs response examples that show that the Crystallization appears linked with a temperature and time, and not to the entire process of Crystallization.

2 Post-Processing

The pipeline was tested with and without post-processing on seven different LLMs. The post-processing involves complex actions, transformations based on action parameters, and the removal of chemicals containing banned words. The results of each LLM are presented in Table 3.1. All the rules applied in the post-processing are in Table 2.4

General Rules	Type of correction
Remove every action containing chemical substances with banned words in their names such as: newsolution, teflon-lined, N/A,...	Improve on Chemical Name

Table ESI 2.1: General rules applied in the post-processing of the laboratory action pipeline

Zeolite Synthesis Rules	Type of correction
•Add a SetTemperature action before actions containing a temperature value in the context that are not: ThermalTreatment, Dry, or Crystallization	Improve on Action Conditions
•Add a Set Atmosphere action before an Add action containing the atmosphere composition information	Improve on Action Conditions
•Add a NewSolution action before an Add action if there is no NewSolution action between this Add action and the beginning of the procedure or a Wash/Separate actions	Improve on Action Conditions
•Add a repeat action after a wash if the amount of repetitions are mentioned in the context	Improve on Action Conditions
•Convert IonExhange/AlkalineTreatment/AcidTreatment into a sequence of NewSolution-Add-Add-SetTemperature-Stir-Repeat whenever possible	Improve on Treatment Definition
•Convert SetTemperature with duration into Crystallization	Improve on Action Conditions
•Remove Wait actions at the end of the procedures if the previous action is a Dry, Wait, ThermalTreatment, Wash, Separate	Improve on Action Conditions
•Eliminate the Repeat action with a lower amount of repetitions if there are two consecutive actions	Improve on Action Conditions

Table ESI 2.2: Zeolite synthesis rules applied in the post-processing of the laboratory action pipeline

Organic Synthesis Rules	Type of correction
•Add a SetTemperature action before actions containing a temperature value in the context	Improve on Action Conditions
•temperature reductions without a value got the temperature value set to room temperature	Improve on Action Conditions
•Temperature value changed to 0 ^{o}C when an ice bath was mentioned in the context without a temperature value	Improve on Action Conditions
•Add a SetTemperature action before actions containing a temperature value in the context	Improve on Action Conditions
•Converted MakeSolution action into a sequence of Add actions	Improve on Action Conditions
•Set the phase to keep in the Filter action to filtrate when there is no mention of the phase kept in the context	Improve on Action Conditions
•Transform SetTemperature actions into Reflux action when no temperature value is found and the word "reflux" was detected in the context	Improve on Action Conditions
•Transform SetTemperature actions into Reflux action when no temperature value is found and the word "reflux" was detected in the context	Improve on Action Conditions
•Add a Wait action after a SetTemperature action if a duration value was found in the context	Improve on Action Conditions

Table ESI 2.3: Organic synthesis rules applied in the post-processing of the laboratory action pipeline

SAC Synthesis Rules	Type of correction
•Temperature reductions without a value got the temperature value set to room temperature	Improve on Action Conditions
•Set the phase to keep in the Filter action to precipitate when there is no mention of the phase kept in the context	Improve on Action Conditions
•Convert SetTemperature action with atmopshere information converted to ThermalTreatment	Improve on Action Conditions
•Add a Wait or a Stir action after a SetTemperature action if a duration value was found in the context depending if the stirring speed is mentioned in the context	Improve on Action Conditions

Table ESI 2.4: SAC synthesis rules applied in the post-processing of the laboratory action pipeline

Model	Action Sequence Similarity	Action Content f-score	Chemical Info f-score
With Post Processing			
microsoft/Phi-3-medium-4k-instruct	0.85	0.82	0.85
upstage/SOLAR-10.7B-Instruct-v1.0	0.82	0.74	0.78
ibm-granite/granite-3.2-8b-instruct	0.85	0.80	0.85
meta-llama/Meta-Llama-3-8B-Instruct	0.87	0.81	0.82
openchat/openchat-3.5-0106	0.79	0.73	0.79
Qwen/Qwen3-4B-Instruct-2507	0.85	0.81	0.84
google/gemma-3-12b-it	0.89	0.84	0.85
microsoft/Phi-3-medium-4k-instruct	0.76	0.73	0.83
upstage/SOLAR-10.7B-Instruct-v1.0	0.76	0.64	0.72
ibm-granite/granite-3.2-8b-instruct	0.76	0.69	0.78
meta-llama/Meta-Llama-3-8B-Instruct	0.75	0.70	0.76
openchat/openchat-3.5-0106	0.65	0.60	0.70
Qwen/Qwen3-4B-Instruct-2507	0.71	0.68	0.80
google/gemma-3-12b-it	0.82	0.76	0.84

Table ESI 2.5: Performance metrics (action sequence similarity, action content f-score, and chemical information f-score) for the seven models with and without the post-processing of the pipeline.

3 Pipeline Optimization and Ablation Studies

17 models were tested on an optimization dataset containing 75 paragraphs about zeolite synthesis. Each model was tested in 4 different temperature values (0, 0.25, 0.5, 0.7). In Figure 3.1, each model is compared in the best parameter configuration. Note that all models were run in a deterministic setting. However, the model google/gemma-3-12b-it still produced non-deterministic results due to the inference Python library. Hence, for this model only, 5 runs were made at each stage, and the average and deviation of these 5 runs were computed and used for comparison with other models.

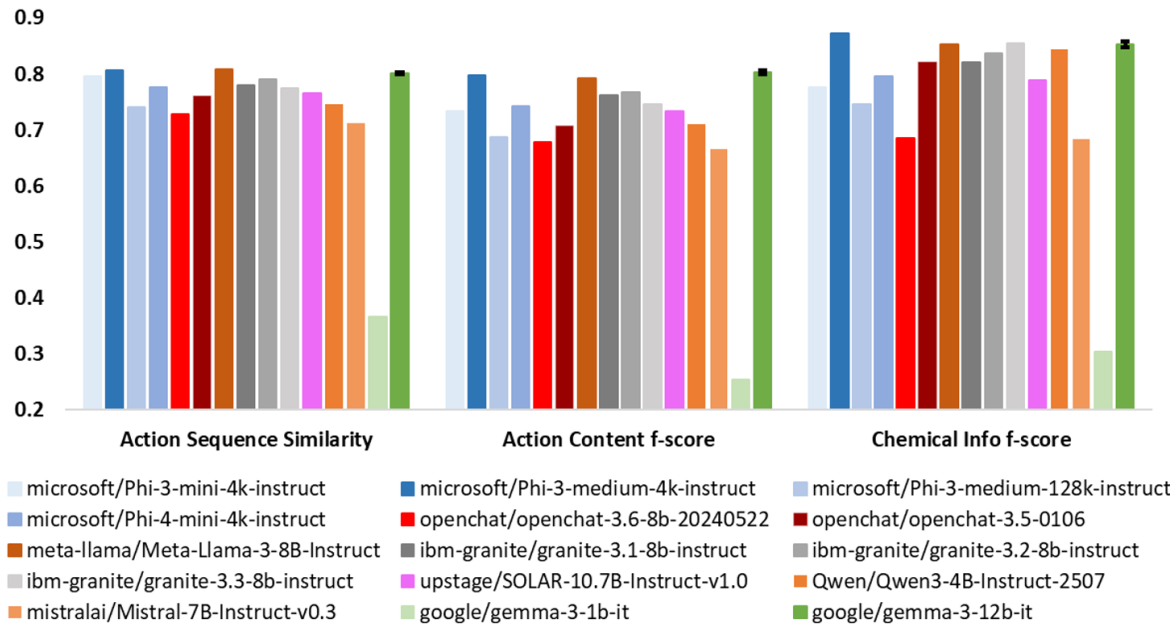

Figure ESI 3.1: Performance of the 15 models tested on the optimization dataset on 3 different metrics (action sequence similarity, action content f-score, and chemical information f-score) with the best temperature value. The optimization of each model can be found in 10.5281/zenodo.17241571.

Ablation studies were conducted to identify the specific parts of the prompt necessary for eliciting coherent responses from the LLMs. 14 different action prompts and 8 chemical extraction prompts were tested across the seven different LLMs. A description of the best prompt on each LLM can be found in Table **??**. You can see the LLMs performance before and after prompt optimization in Figure 4.1

Model	Prompt Description
microsoft/Phi-3-medium-4k-instruct	Asking for an action list and giving an action template to the model Not defining the AcidTreatment action
upstage/SOLAR-10.7B-Instruct-v1.0	sking for an action list and giving an action template to the model
ibm-granite/granite-3.2-8b-instruct	Asking for an action list and giving an action template to the model
meta-llama/Meta-Llama-3-8B-Instruct	Asking for an action list and giving an action template to the model
openchat/openchat-3.5-0106	Asking for an action list and giving an action template to the model Not defining the AlkalineTreatment action
Qwen/Qwen3-4B-Instruct-2507	Asking for an action list and giving an action template to the model
google/gemma-3-12b-it	Asking for an action list and giving an action template to the model and clarify that the model should output one action per line

Table ESI 3.1: Description of the best prompt instruction for each LLM. The full prompts are available in 10.5281/zenodo.17241571.

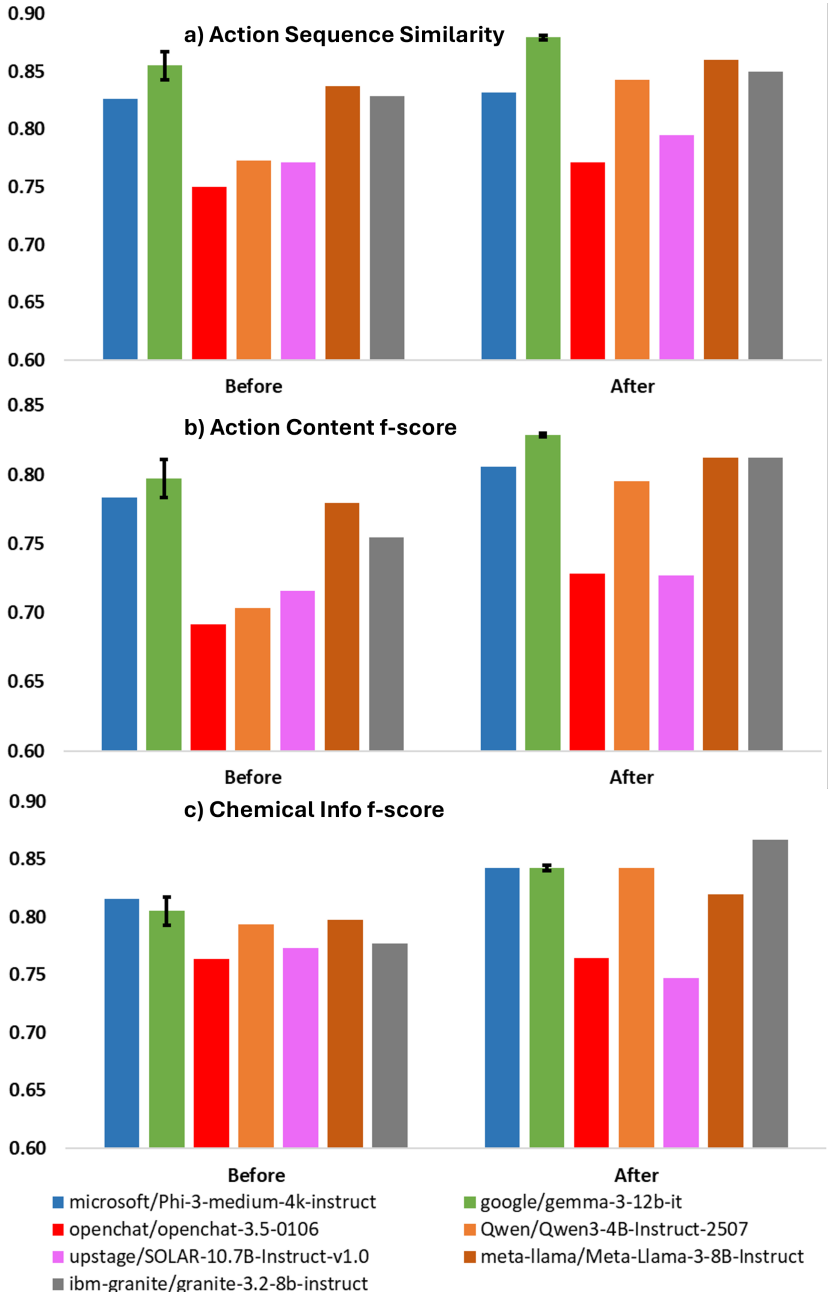

Figure ESI 3.2: Performance of the 7 models tested on the optimization dataset on 3 different metrics: a) action sequence similarity, b) action content f-score, and c) chemical information f-score before and after prompt optimization

4 Ultimate Performance on Different Chemistry Domains

The seven LLMs with their optimized prompts were tested at the extraction laboratory for activities related to zeolite, organic, and SAC synthesis procedures. The results for zeolite and organic synthesis procedures are presented in Figure ?? using a dataset containing 110 paragraphs and 132 sentences, respectively.

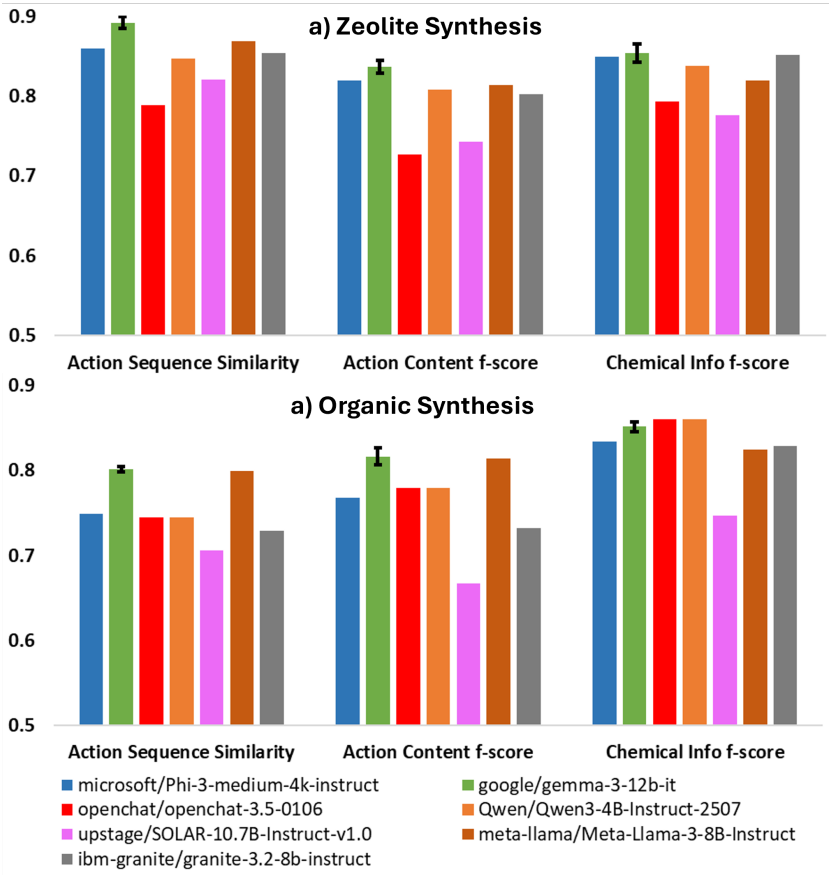

Figure ESI 4.1: Performance of the 7 models tested on the optimization dataset on 3 different metrics(action sequence similarity, action content f-score, and chemical information f-score) for a) zeolite synthesis and b) organic synthesis procedures