Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.
This journal is © The Royal Society of Chemistry 2023

Supplementary Information for

Optimized synthesis of the anti-COVID-19 drugs aided by retrosynthesis software

Wentao Qi,† Dong Zhai,† Danna Song, Chengcheng Liu, Junxia Yang, Lei Sun, Youyong Li,* Xingwei Li, and Weiqiao Deng*

E-mail: yyli@suda.edu.cn (Y.L.) and dengwq@sdu.edu.cn (W.D.)

Table of Contents

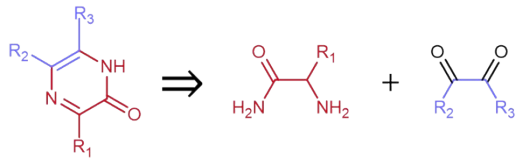
S1. Details of the digital-intellectual retrosynthesis algorithm 2
S2. Experimental details.....9
S3. Price calculation of our route and the original route.....13
S4. Synthetic routes of other potential drugs for treating COVID-19 15
S5. Similar reactions containing pyrazine ring in the products 16

S1. Details of the digital-intellectual retrosynthesis algorithm

1.1 Method

To correctly predict the synthetic route of favipiravir, we introduced the expertise into the retrosynthesis templates based on ASKCOS retrosynthesis program. First, we searched the chemical reactions containing pyrazine ring in the products on SciFinder database (The reactions can be found in S4 section). By analyzing these reactions, we found that they follow a unified reaction rule. The product containing pyrazine ring can be obtained by the cyclization reaction shown in Figure S1a. We wrote this generalized reaction rule into the retrosynthesis template using SMARTS expression (Figure S1b). This template can convert any target complex containing the pyrazine motif (Figure S1c) into its corresponding cyclization precursors by using the RDKit module. The original ASKCOS program cannot output the cyclization precursors for the target containing pyrazine motif. After testing, our template is valid for all targets containing pyrazine motif.

a) Generalized retrosynthesis rule:



b) Retrosynthesis template:

```
[O]~[c;H0;D3;+0:1]1:[c:2]:[n:3]:[c:4]:[c:5]:[n;D2;+0:6]:1>>[N;H2;D1;+0:6]-
[C;H0;D3;+0:1](=[O;D1;H0])-[C:2]-[NH2;D1;+0:3].[O;H0;D1;+0]=[C;D2;+0:4]-
[C;D2;+0:5]=[O;H0;D1;+0]
```

c) Pyrazine motif:

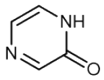


Figure S1. Generalized retrosynthesis rule (a) for the cyclization reaction of the target complex containing the pyrazine motif (c) and its retrosynthesis template using SMARTS expression (b).

Then we added this template and its usage condition as expert knowledge to the ASKCOS program. In the original ASKCOS program, the possible multi-step synthetic routes for a target complex were obtained by its Tree Builder module using the Monte Carlo Tree Search algorithm. Our improvement is as follow. For a given target complex, we first judge whether it contains pyrazine motif. If not, the Tree Builder module is used to obtain the possible synthetic routes for the target. If the pyrazine motif is included, our template is applied to the target to obtain its cyclization precursors. The possible synthetic routes for each precursor are obtained using the Tree Builder module. Finally, the complete synthetic routes are connected as the results. The overall flowchart is shown in Figure S2.

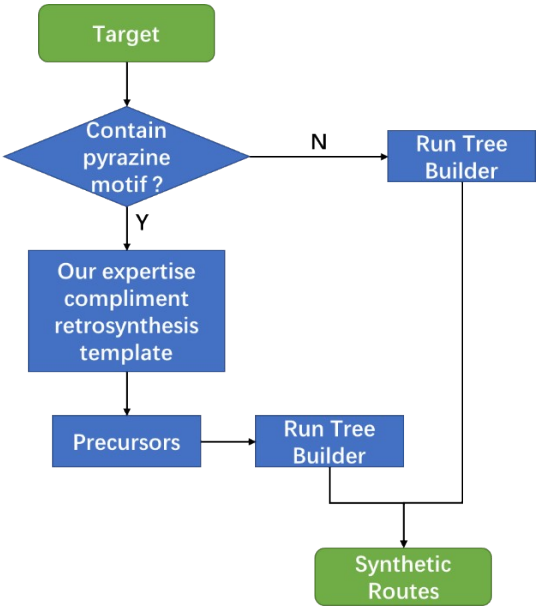


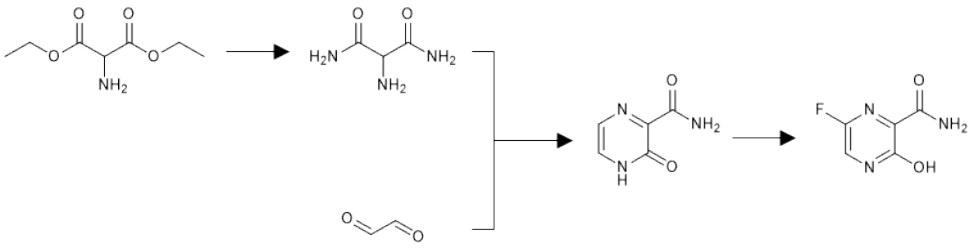
Figure S2. The flowchart of our digital-intellectual retrosynthesis algorithm.

1.2 Synthetic routes aided by AI

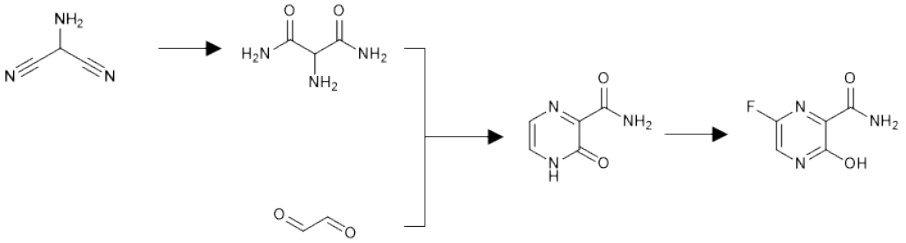
Total 9 synthetic routes of favipiravir less than 5 steps were obtained using our digital-intellectual retrosynthesis algorithm, as shown below (Table S1). We chose the one of the shortest routes for experimental synthesis.

Table S1. Synthetic routes of favirpiravir designed by digital-intellectual retrosynthesis algorithm

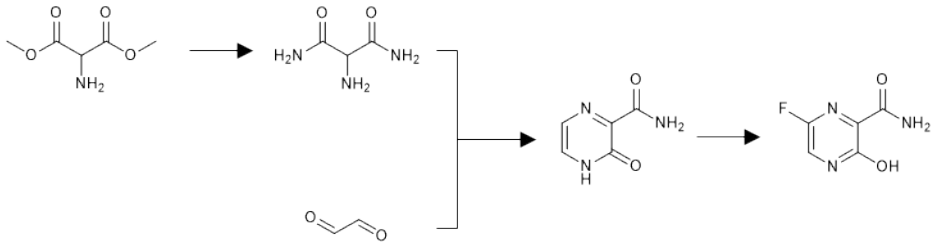
Route 1:



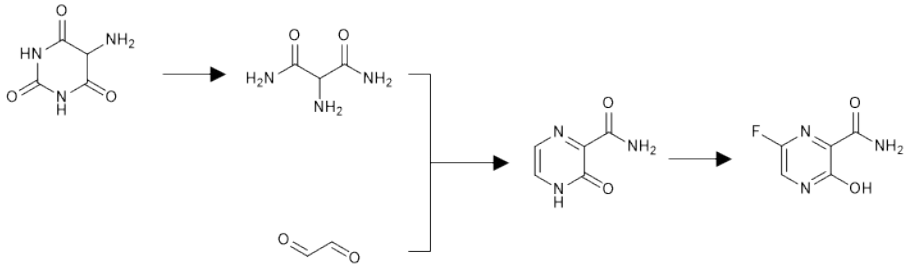
Route 2:



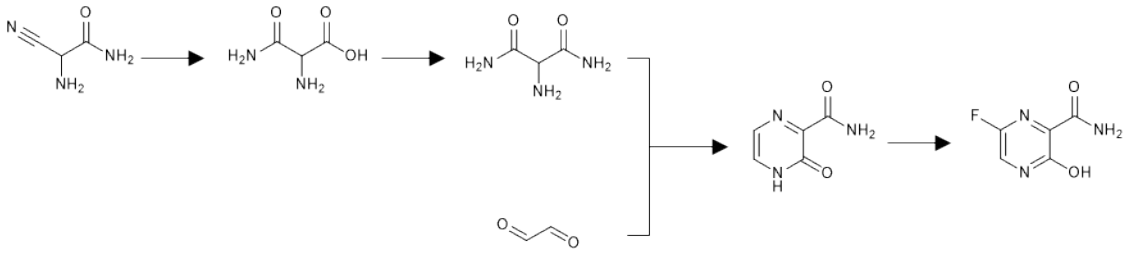
Route 3:



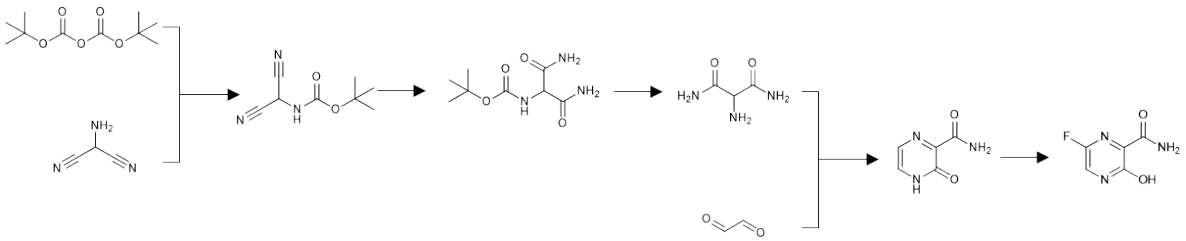
Route 4:



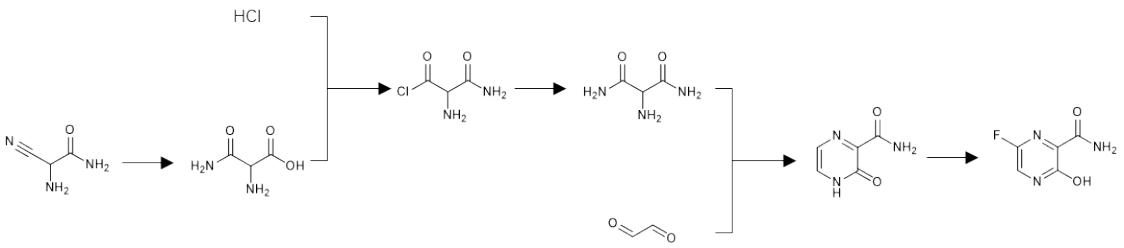
Route 5:



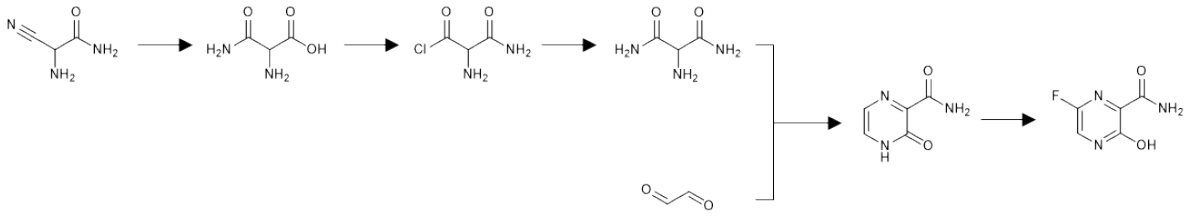
Route 6:



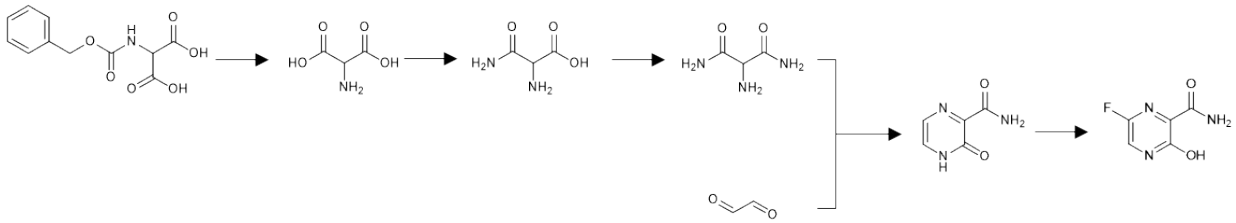
Route 7:



Route 8:



Route 9:



1.3 Computation protocol

Detailed **protocol** describing how to obtain synthetic routes using our digital-intellectual retrosynthesis algorithm:

Step 1. Clone the askcos-core git repository from the github:

```
git clone https://github.com/ASKCOS/askcos-core.git
```

Step 2. Build the askcos-core docker image according to the instructions in the README.md file of the askcos-core git repository.

Then check the name and ID of the askcos-core docker image using the following command:

```
docker images
```

An output example is shown below. The ID of askcos-core image is “2502685a2ff1”.

REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
askcos/askcos-core	latest	2502685a2ff1	4 days ago	7.82GB
askcos/askcos-data	latest	eb55ef3e23a9	12 months ago	5.09GB
askcos/askcos-base	2020.03.6-gh2855-py37-conda	a585aec9656b	2 years ago	1.6GB
askcos	latest	273c1299a407	3 years ago	4.1GB
<none>	<none>	b00ce64e29bc	3 years ago	1.73GB
<none>	<none>	1d6f1a5b931b	3 years ago	117MB
<none>	<none>	45444e1018d8	3 years ago	117MB
<none>	<none>	4ea9475600ba	3 years ago	117MB
redis	latest	9b188f5fb1e6	3 years ago	98.2MB
nginx	latest	f7bb5701a33c	3 years ago	126MB
debian	stretch	f6c68e2ad82a	3 years ago	101MB
rabbitmq	latest	6addf4b6a4ef	3 years ago	151MB

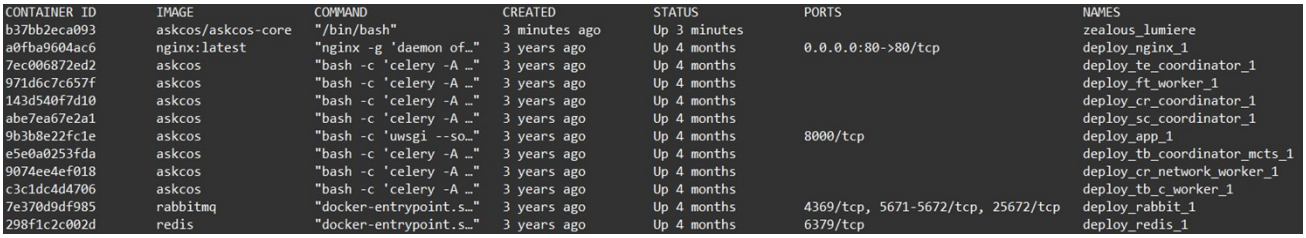
Step 3. Start a new docker container using the askcos-core image and run the /bin/bash command inside it:

```
docker run -it askcos/askcos-core /bin/bash
```

Step 4. Open another terminal and get the container ID of askcos-core using the following command:

```
docker ps
```

An output example is shown below. Get the CONTAINER ID “b37bb2eca093” of the “askcos/askcos-core” container.



CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
b37bb2eca093	askcos/askcos-core	"/bin/bash"	3 minutes ago	Up 3 minutes		zealous_lumiere
a0fba9604ac6	nginx:latest	"nginx -g 'daemon of…"	3 years ago	Up 4 months	0.0.0.0:80->80/tcp	deploy_nginx_1
7ec006872ed2	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_te_coordinator_1
971d6c7c657f	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_ft_worker_1
143d540f7d10	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_cr_coordinator_1
abe7ea67e2a1	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_sc_coordinator_1
9b3b8e22fc1e	askcos	"bash -c 'uwsgi --so…"	3 years ago	Up 4 months	8000/tcp	deploy_app_1
e5e0a0253fda	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_tb_coordinator_mcts_1
9074ee4ef018	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_cr_network_worker_1
c3c1dc4d4706	askcos	"bash -c 'celery -A …"	3 years ago	Up 4 months		deploy_tb_c_worker_1
7e370d9df985	rabbitmq	"docker-entrypoint.s…"	3 years ago	Up 4 months	4369/tcp, 5671-5672/tcp, 25672/tcp	deploy_rabbit_1
298f1c2c002d	redis	"docker-entrypoint.s…"	3 years ago	Up 4 months	6379/tcp	deploy_redis_1

Step 5. Copy our modified tree_builder.py file to the appropriate location in the “askcos/askcos-core” container using the following command:

```
docker cp mod_tree_builder.py b37bb2eca093:/usr/local/askcos-core/askcos/retrosynthetic/mcts/
```

Step 6. In the “askcos/askcos-core” container, write the SMILES expression of the target molecule in “target_smiles” variable in the “if __name__ == '__main__'” section of our mod_tree_builder.py file. e.g.

```
target_smiles = ‘NC(=O)c1nc(F)cnc1O’    # this example expression represents favirpiravir
```

Step 7. Run the modified tree builder module to obtain the synthetic routes using the following command:

```
cd /usr/local/askcos-core/askcos/retrosynthetic/mcts/

python mod_tree_builder.py
```

The result of the synthetic routes is saved in the “routes.txt” file.

S2. Experimental details

2.1 General details

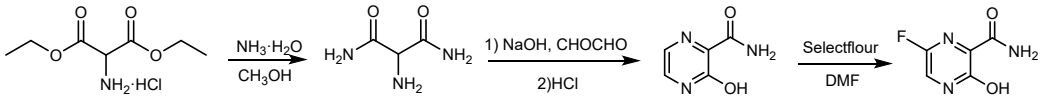
All reactions were carried out in constant temperature oil-bath pans with magnetic stirring. Solvents and reagents were purchased from commercial companies and used as received without further purification. Visualization was accomplished with UV light. 1H NMR spectra were recorded on Bruker AVANCE NEO 600MHz spectrometer and are reported in ppm using solvent as an internal standard (DMSO-d6 at 2.50 ppm). Accurate masses were obtained using a Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer.

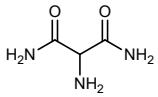
2.2 Materials

Diethyl aminomalonate hydrochloride (>98% purity; CAS. no. 13433-00-6) was purchased from Shanghai Jinsui Biotechnology Co., LTD. Glyoxal (40 wt. % in H_2O; CAS. no. 107-22-2) was obtained from Acmec. Selectfluor (>95% purity; CAS. no. 140681-

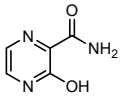
55-6) was obtained from Meryer. Aqua ammonia ($NH_3 \cdot H_2O$, 28%; CAS. no. 7664-41-7), Methanol (MeOH, >99% purity; CAS. no. 67-56-1), N,N-Dimethylformamide (DMF, >99% purity; CAS. no. 68-12-2), sodium hydroxide (NaOH, >99% purity; CAS. no. 1310-73-2), and hydrochloric acid (HCl, 37%; CAS. no. 7647-01-0) were purchased from Sinopharm.

2.3 Synthesis of favipiravir

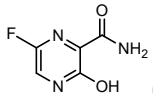




2-aminomalonamide: In a dry two-necked flask, diethyl aminomalonate hydrochloride (5 g) was dissolved in methanol (5 mL), then ammonia (10 mL) was added and the mixture was stirred at room temperature for 24 h. After the reaction, ethanol was added to get precipitate, and the precipitate was washed with ethanol and dried under vacuum to get a bright yellow solid. (Yield: 66.48%). 1H NMR (600 MHz, DMSO-d6) δ 7.32 (d, J = 108.0 Hz, 4H), 3.75 (s, 1H), 2.37 (s, 2H). 13C NMR (151 MHz, DMSO-d6) δ 171.65, 57.65; HRMS (ESI) Calcd for $C_3H_8N_3O_2$ [M+H] 118.06, found 118.0614.



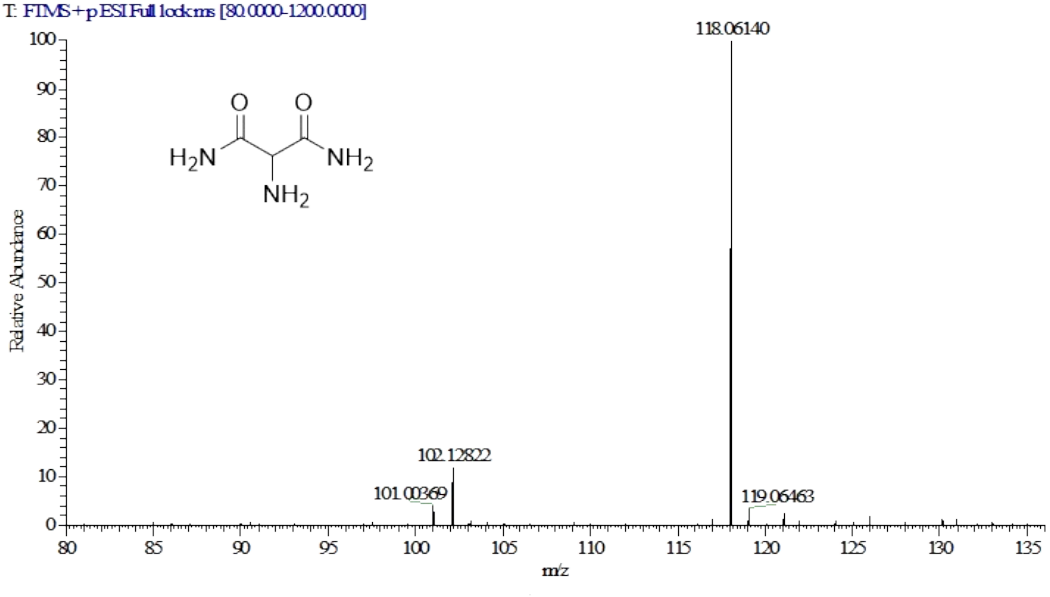
3-hydroxypyrazine-2-carboxamide: Aminomalonamide (1 g) was dissolved in 20% NaOH solution (5 mL). Then 40% glyoxal (1.16 mL) was added dropwise and stirred at rt for 4 h. The mixture was adjusted to pH=2 with 6 M HCl and filtered. Washed with 95% ethanol and dried under vacuum, the precipitate was a canary yellow solid. (Yield: 77.45%). 1H NMR (600 MHz, DMSO-d6) δ 13.23 (s, 1H), 8.71 (s, 1H), 8.11 (d, J = 25.3 Hz, 2H), 7.92 (s, 1H). 13C NMR (151 MHz, DMSO-d6) δ 169.53, 167.55, 159.37, 155.32, 128.61; HRMS (ESI) Calcd for $C_5H_4N_3O_2$ [M-H] 138.04, found 138.02943.



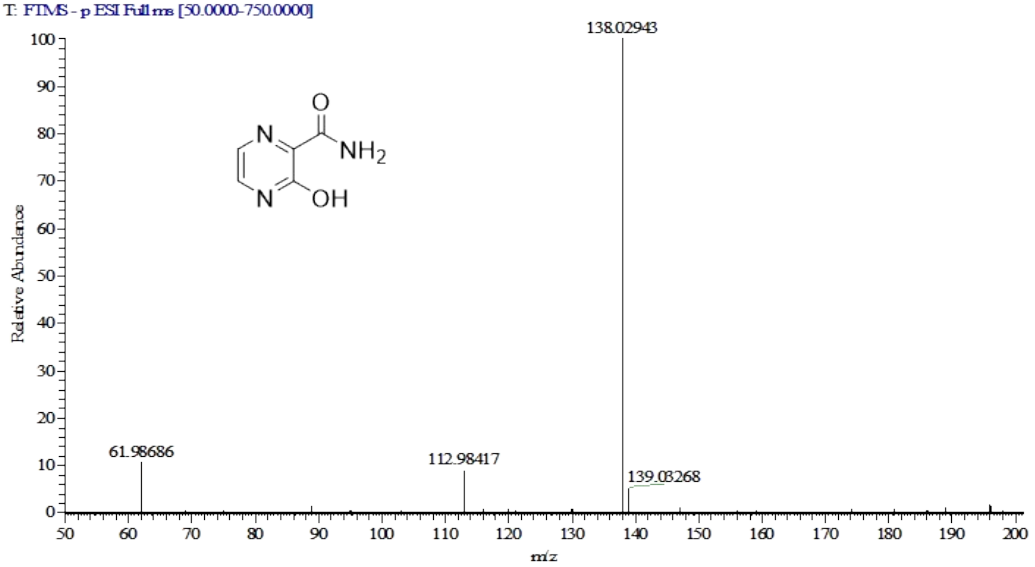
6-fluoro-3-hydroxypyrazine-2-carboxamide (favipiravir): 3-hydroxypyrazine-2-carboxamid (0.1391 g) and selectflour (0.531 g) were dissolved in 10 mL DMF and stirred at 50 ℃ for 60 h. Most solvents were removed using a rotary evaporator and the resulting mixture was extracted with ethyl acetate (20 mL×3) and washed with brine (10 mL×3). The combined organic phase was dried over Na_2SO_4. After filtration and removal of the solvent, a canary yellow solid product was obtained. (Yield: 62.89%). 1H NMR (600 MHz, DMSO-d6) δ 13.40 (s, 1H), 8.73 (s, 1H), 8.54 – 8.44 (m, 2H). 13C NMR (151 MHz, DMSO-d6) δ 169.11, 160.17, 152.85 (d, J = 243.1 Hz), 136.13 (d, J = 41.3 Hz), 122.93; HRMS (ESI) Calcd for $C_5H_3FN_3O_2$ [M-H] 156.03, found 156.0203.

2.4 Raw spectra data

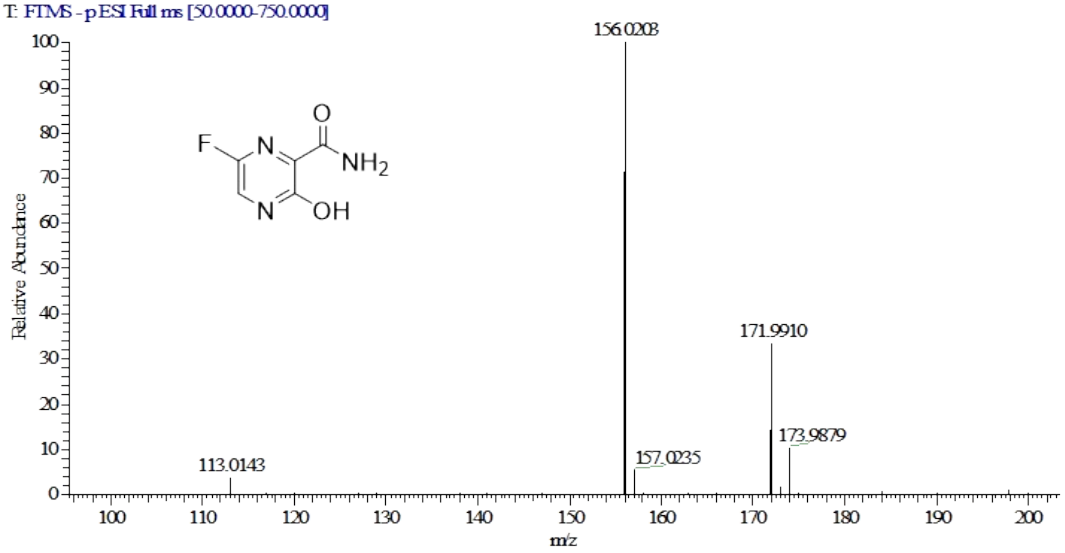
Qiwentao-1 #20 RT: 0.11 AV: 1 SB: 2 0.07 , 0.19 NL: 7.65E7
T: FTMS + p ESI Full lock ms [80.0000-1200.0000]

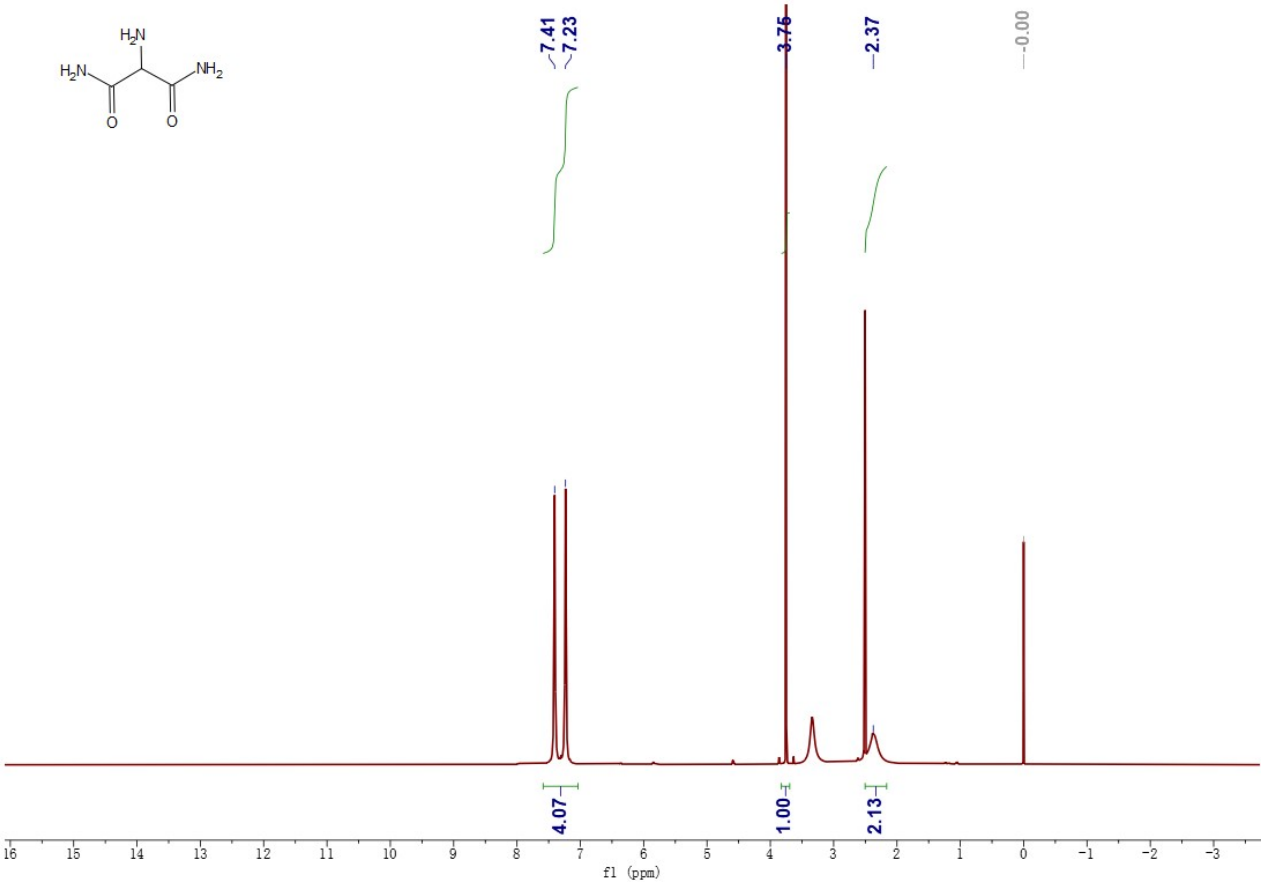


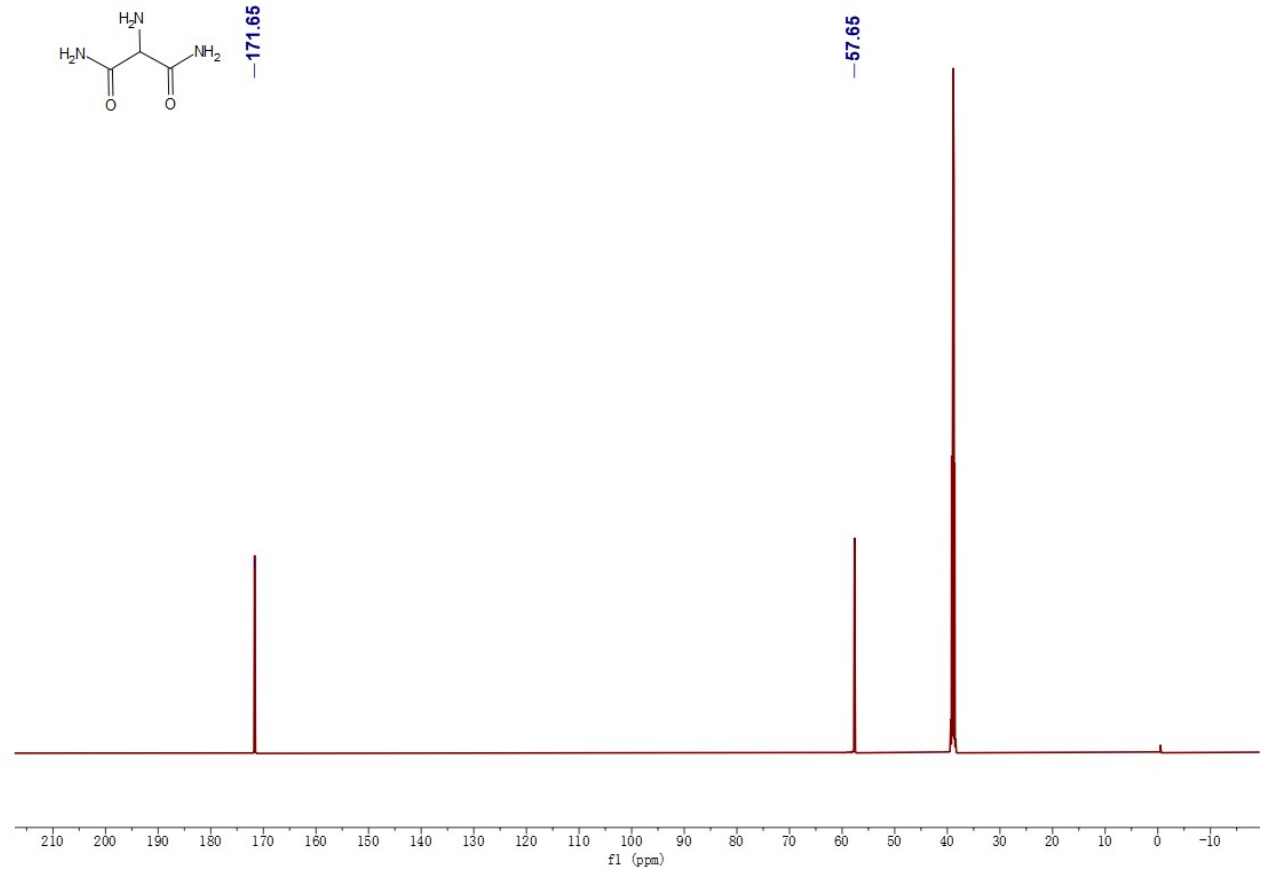
Qiwentao-2_210319133701 #10 RT: 0.10 AV: 1 NL: 9.37E7
T: FTMS - p ESI Full ms [50.0000-750.0000]

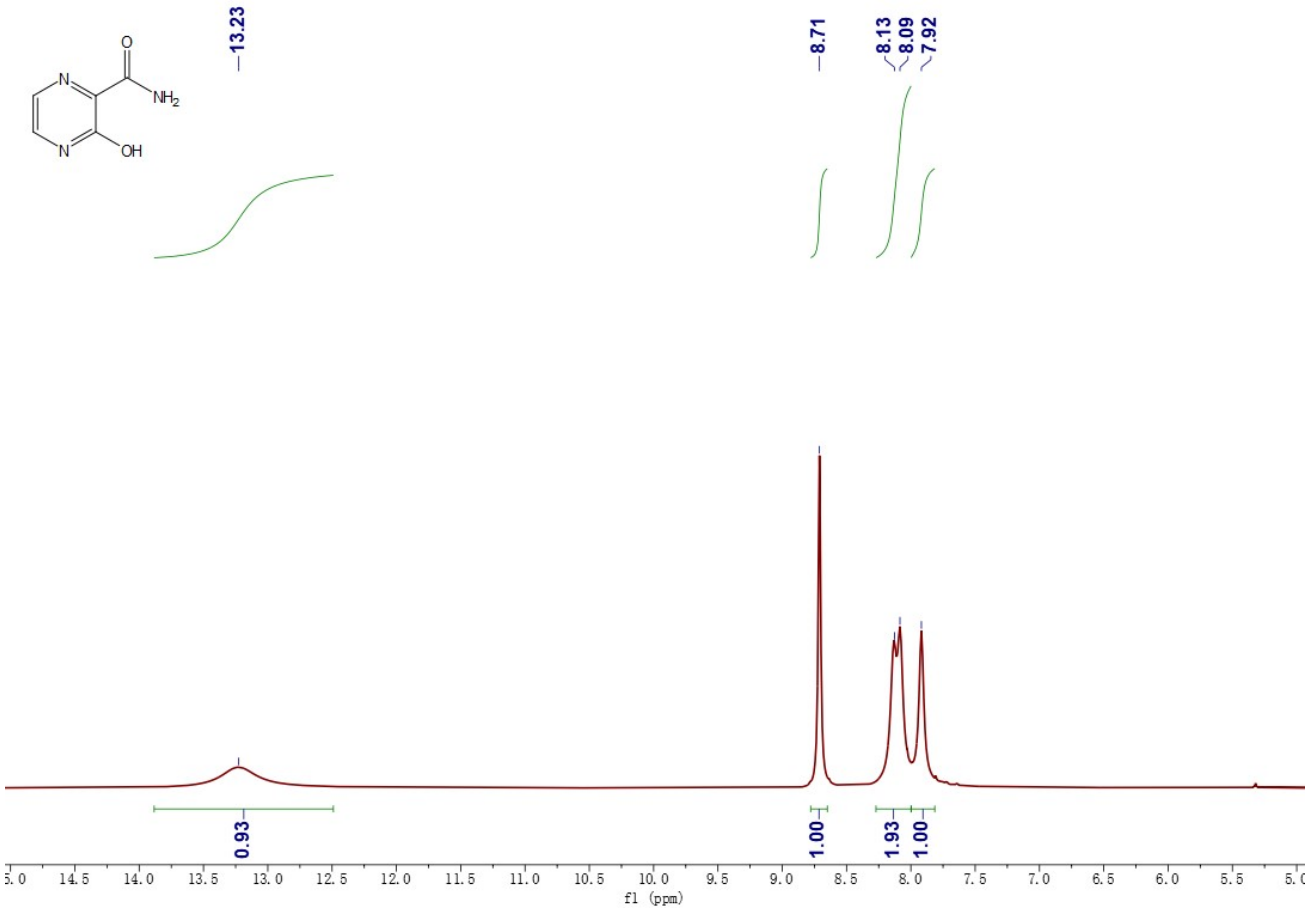


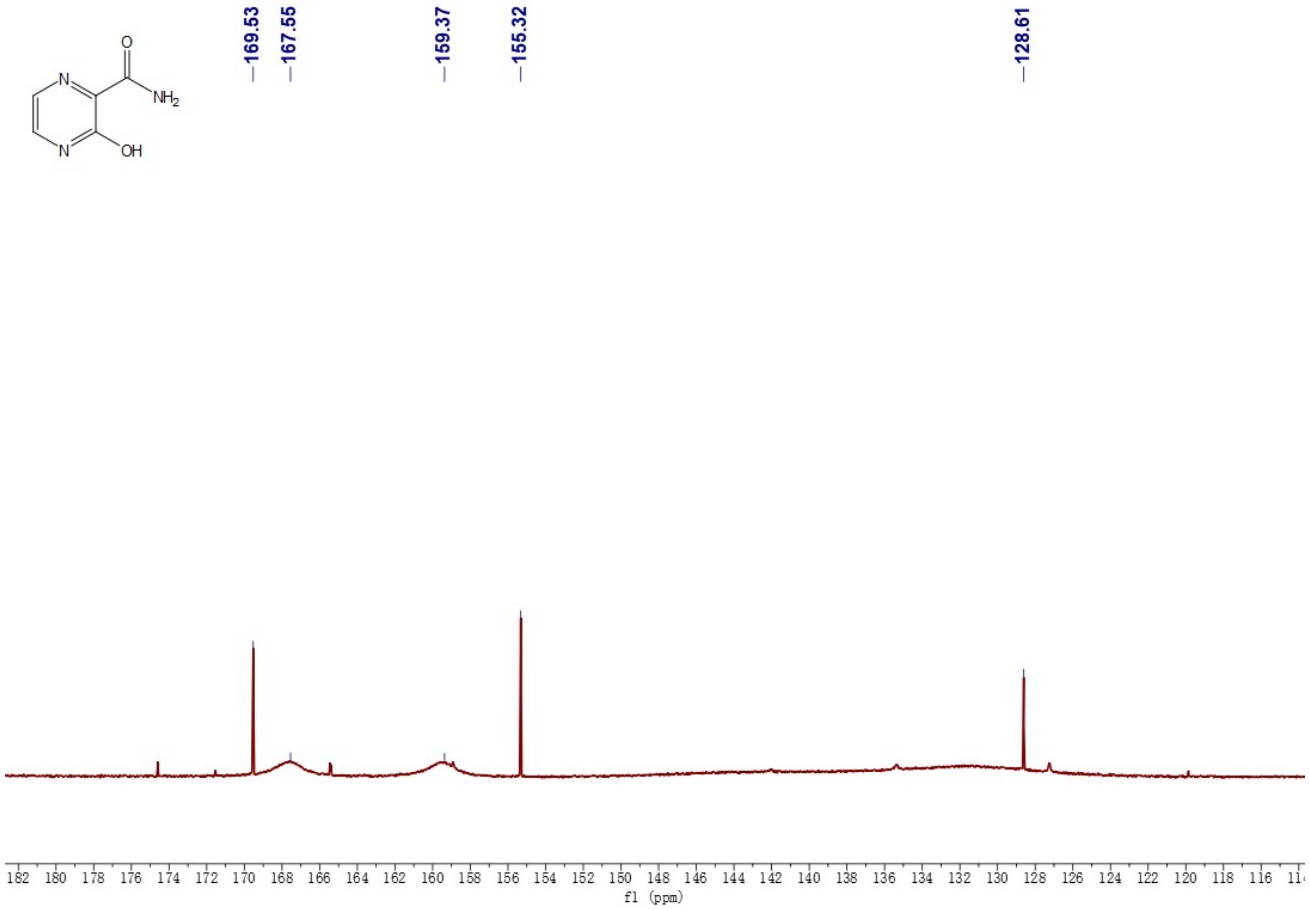
Qiwentao-1 #116 RT: 1.44 AV: 1 NL: 5.54E8
T: FTMS - p ESI Full ms [50.0000-750.0000]

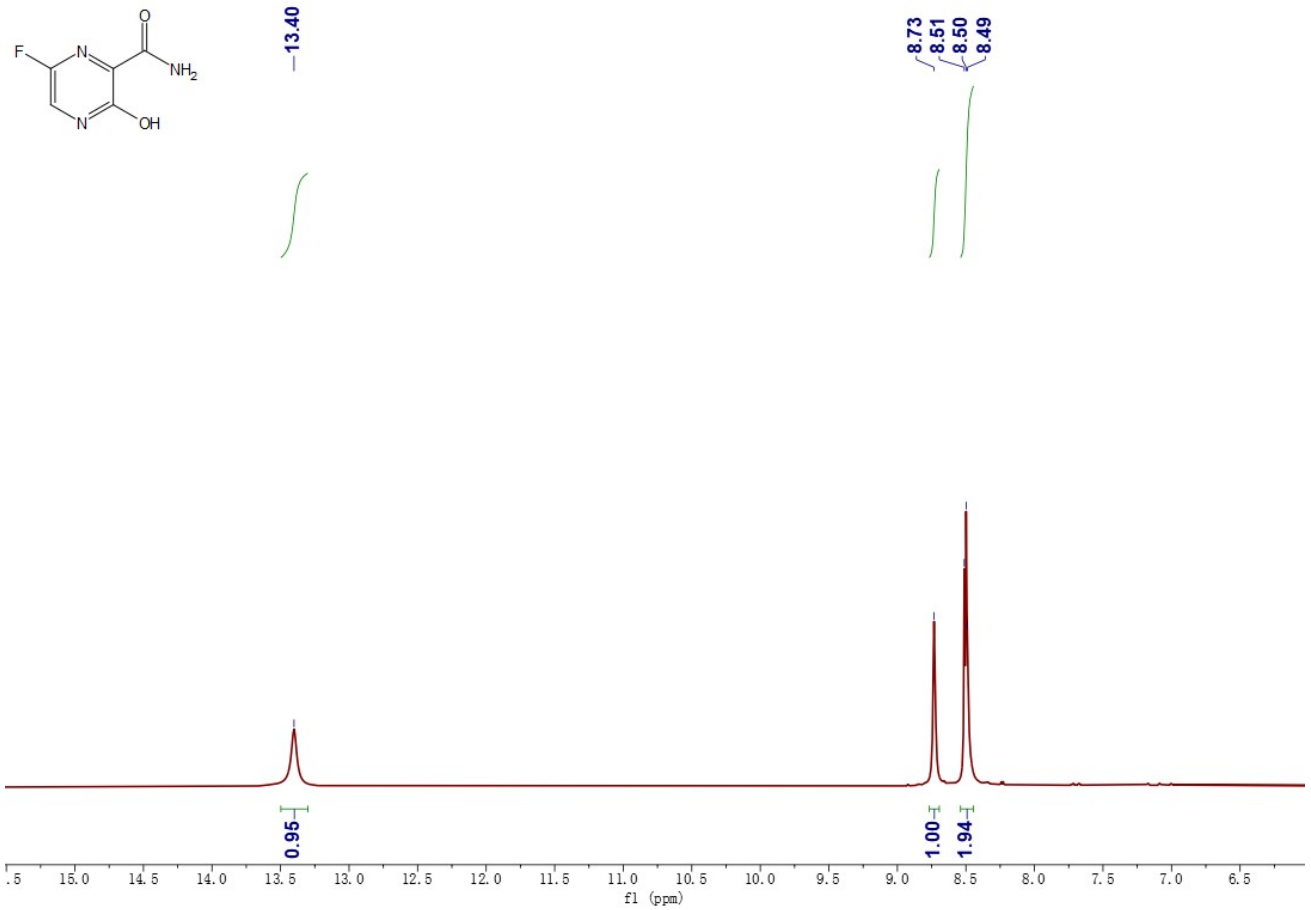


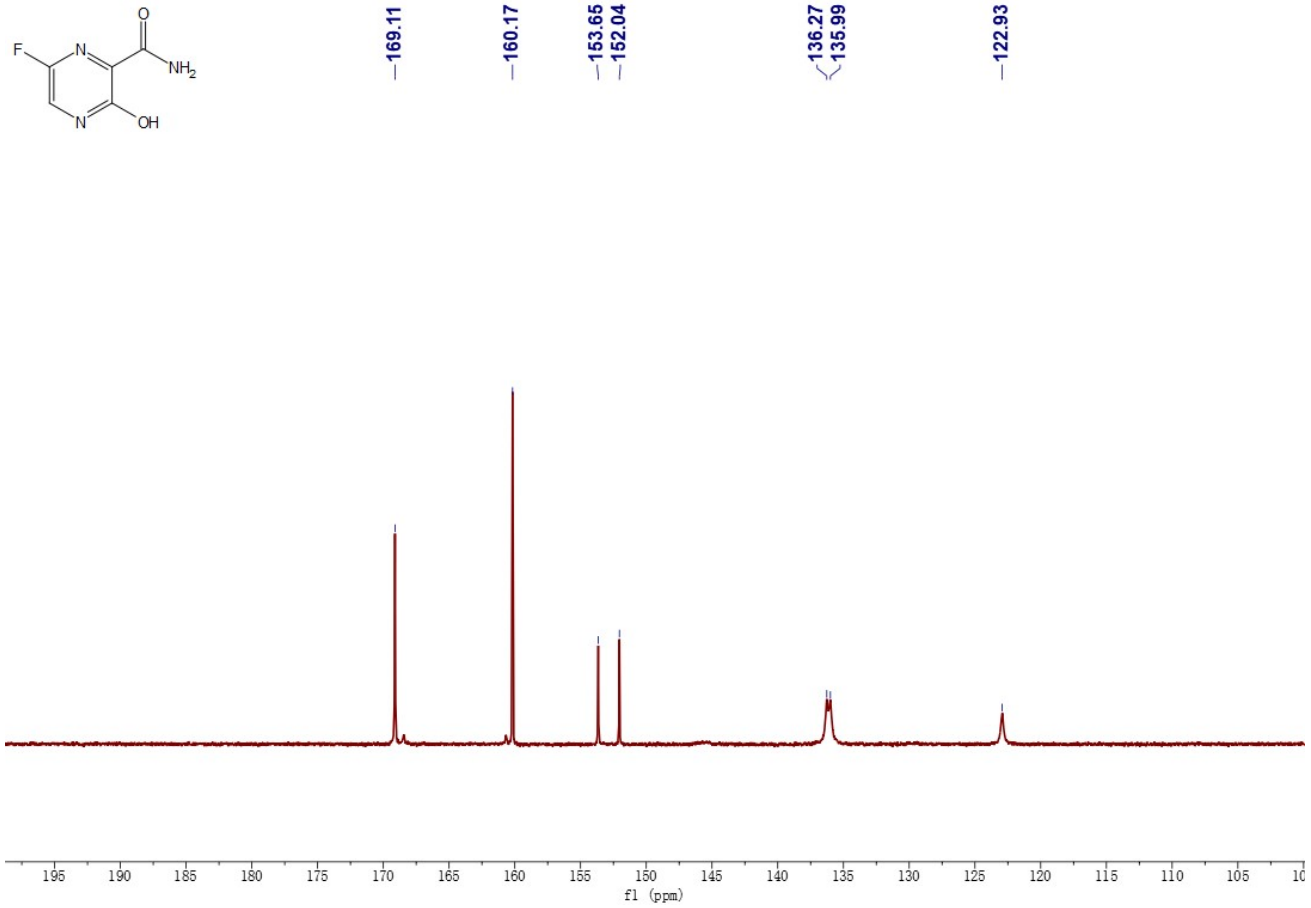

7.41
7.23
3.75
2.37
0.00
4.07
1.00
2.13
f1 (ppm)


171.65
57.65
f1 (ppm)


13.23
8.71
8.13
8.09
7.92
0.93
1.00
1.93
1.00
f1 (ppm)


169.53
167.55
159.37
155.32
128.61
f1 (ppm)


F
N
O
NH2
N
OH
13.40
8.73
8.51
8.50
8.49
0.95
1.00
1.94
f1 (ppm)


F
N
O
NH2
N
OH
169.11
160.17
153.65
152.04
136.27
135.99
122.93
f1 (ppm)

S3. Price calculation of our route and the original route

We calculated the cost of our route and the original route based on the assumption that 1 mol of starting materials was used. The price and quantities of the reagent/reactant involved are shown below (Table S2). All manufacturers and price information source from MOLBASE (www.molbase.com) except some controlled products.

Table S2. All material the original route used (a) and our route used (b).

a) All material the original route used

Reagent/reactant	CAS NO.	Price ($)	Manufacturers	Consumption
3-Aminopyrazine-2-carboxylic acid	5424-01-1	333.33/1kg	Ester (Chengdu) Biopharmaceutical Co., Ltd.	139.11g
Methanol	67-56-1	300/1t	Xi'an Lichang Chemical Technology Co., Ltd.	32.04g
N-Bromosuccinimide	128-08-5	20.00/1kg	Shanghai Angsi Biology Techology Co.,Ltd	156.6g
Sodium nitrite	7632-00-0	30.00/1kg	Shanghai Dumi Bioteechnology Co.,Ltd	71g
Sulfuric acid	7664-93-9	0.164/1L	Sinopharm Chemical Reagent Co., Ltd.	0.66L
Benzophenone imine	1013-88-3	66.00/25kg	HeBei GuanLang Biotechnology Co.,Ltd	37.2g
Tris(dibenzylideneacetone)dipalladium	51364-51-3	7091.67/1kg	Tianjin Leju Technology Co., Ltd.	1.53g
2,2'-Bis(diphenylphosphino)-1,1'-binaphthalene	98327-87-8	2,000.00/1kg	Tianjin Leju Technology Co., Ltd.	3.1g
Sodium tert-butoxide	865-48-5	78.00/1kg	Beijing Huawei Ruike Chemical Co., Ltd.	22.39g
Ammonia	7664-41-7	66.00/25kg	HeBei GuanLang Biotechnology Co.,Ltd	0.092L
Pyridine hydrofluoride	62778-11-4	183.33/1kg	Shanghai Jiangge Chemical Co., Ltd.	0.089L
Sodium iodide	7681-82-5	40.00/1kg	Jinan Wenyi Chemical Technology Co.,Ltd.	34g
Chlorotrimethylsilane	75-77-4	11.67/1kg	Rizhao Power Texaco Chemical Co. Ltd.	24.75g
Yield of favipiravir: 1.287g				

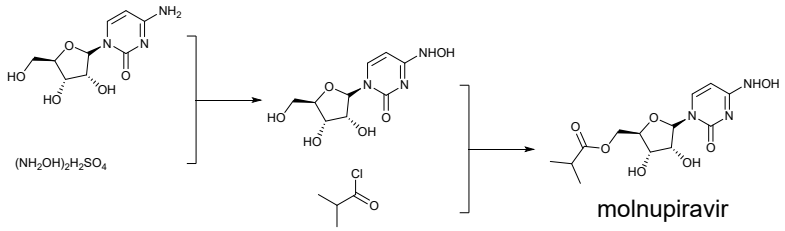
b) All material our route used

Reagent/reactant	CAS NO.	Price ($)	Manufacturers	Consumption
Diethyl Aminomalonate Hydrochloride	13433-00-6	165.00/500g	Nanjing Chemlin Chemical Industry	211.6g

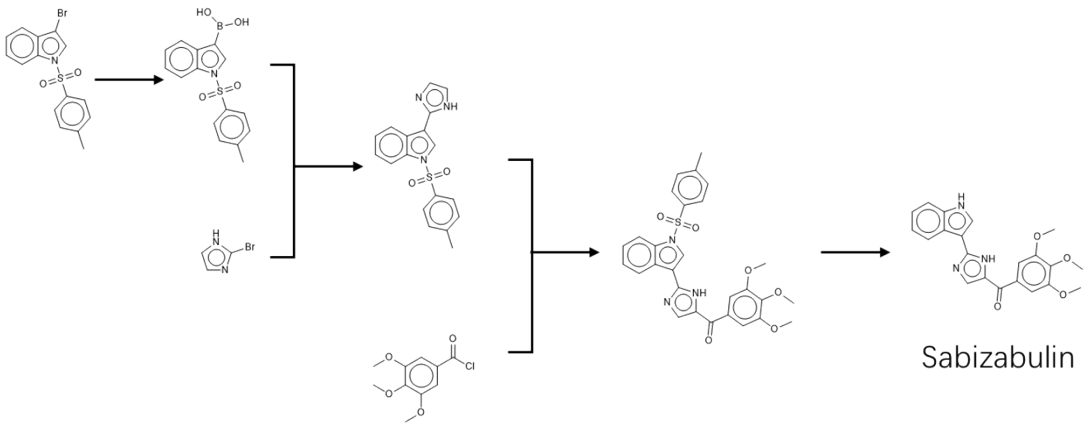
			Co.,Ltd.	
Ammonia	7664-41-7	66.00/25kg	HeBei GuanLang Biotechnology Co.,Ltd	0.422L
Sodium hydroxide	1310-73-2	6,770.00/1t	Lianyungang Longtaiwei Food Ingredients Co., Ltd.	0.0284g
Glyoxal	107-22-2	5.83/500mL	shandong xiya chemical technology Co.,LTD	0.091L
Selectflour	140681-55-6	23.00/1kg	Aobo Rui (Tianjin) Technology Co., Ltd.	273.4g
Yield of favipiravir: 50.836g				

S4. Synthetic routes of other potential drugs for treating COVID-19

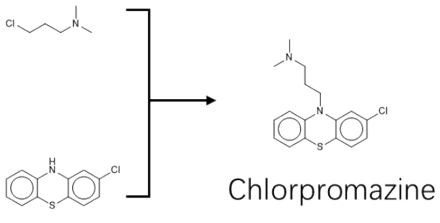
Molnupiravir:



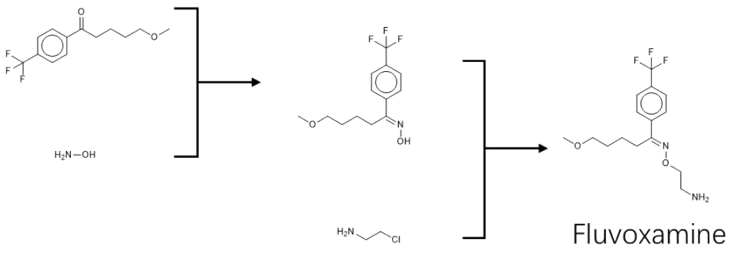
Sabizabulin:



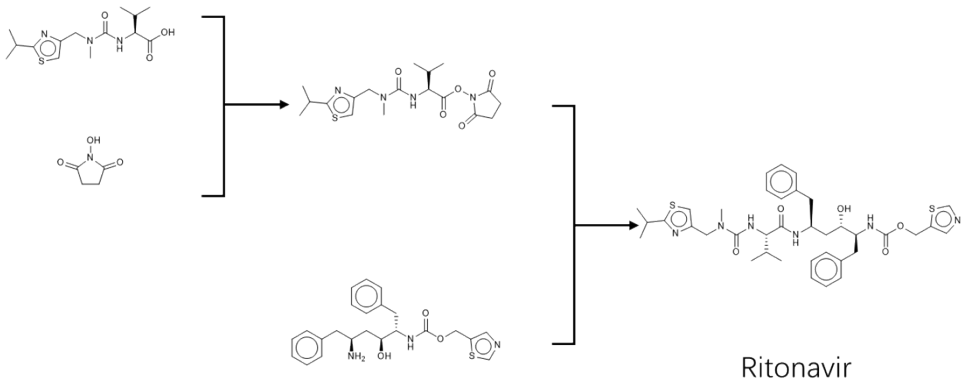
Chlorpromazine:



Fluvoxamine:

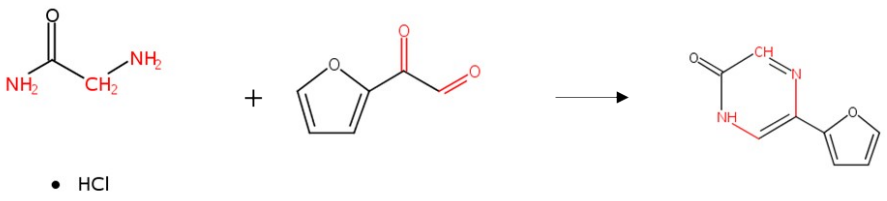


Ritonavir:

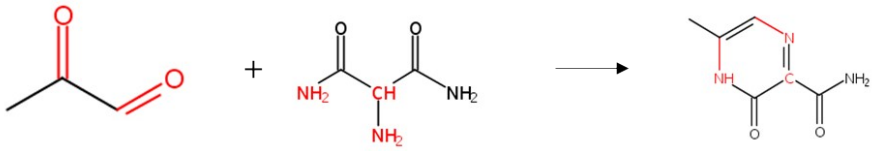


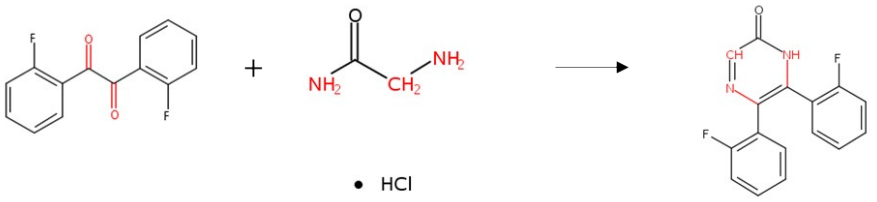
S5. Similar reactions containing pyrazine ring in the products

Some similar reactions containing pyrazine ring in the products we found are on SciFinder database shown below. The expertise template is obtained accordingly.

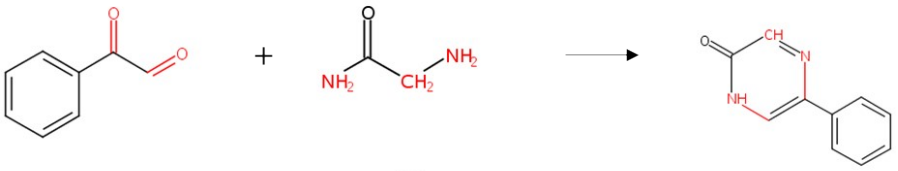


• HCl

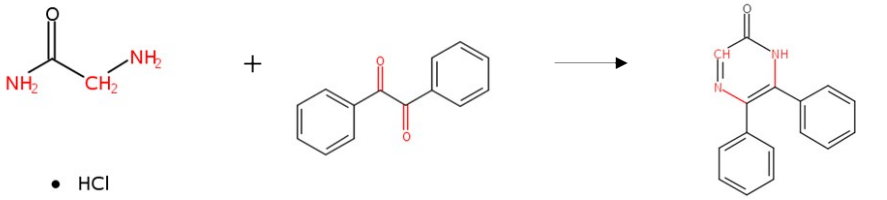




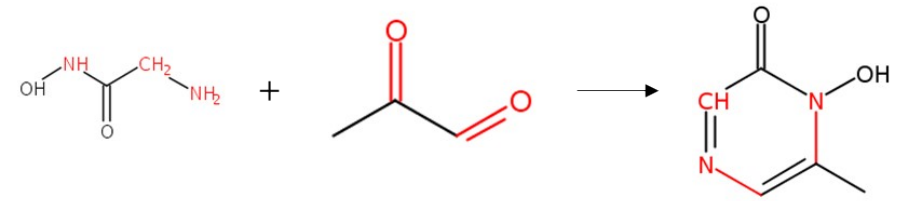
• HCl

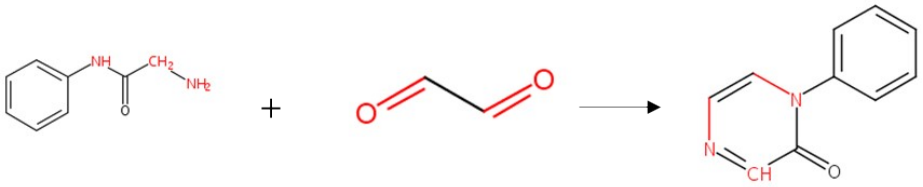


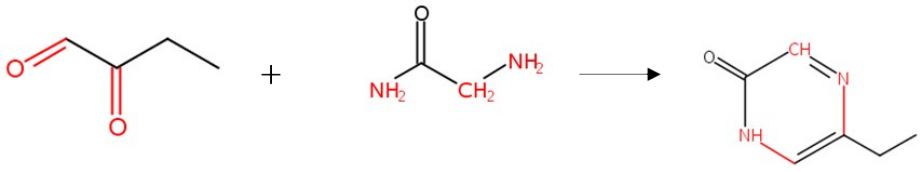
• HCl



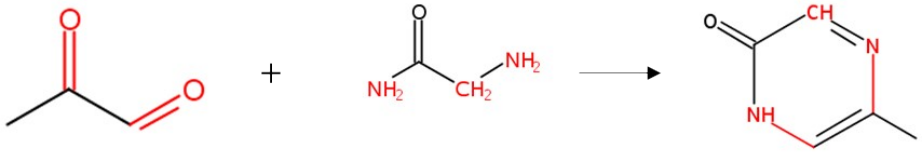
• HCl



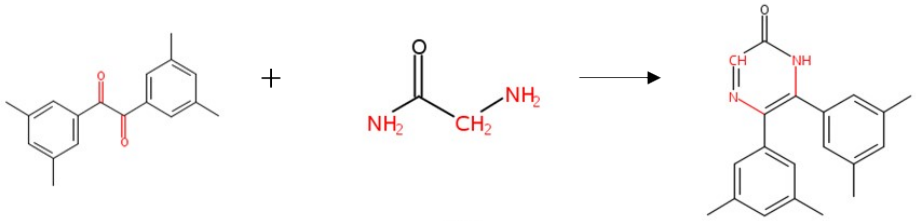




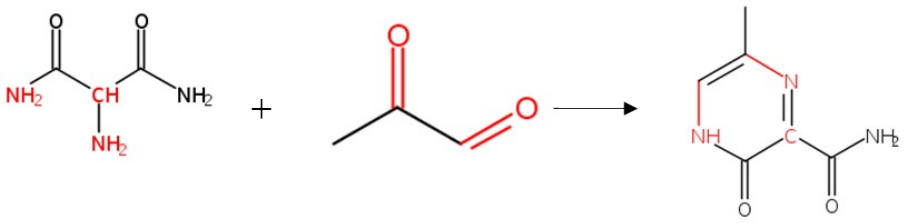
- HCl

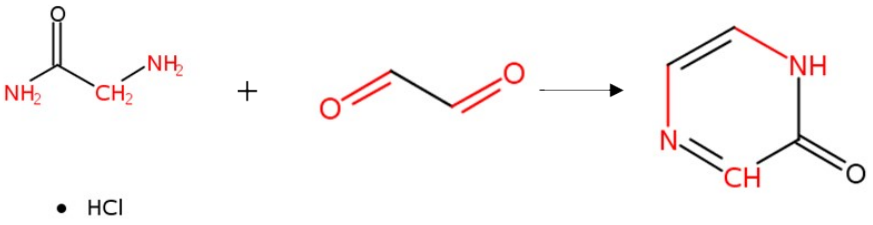


- HCl

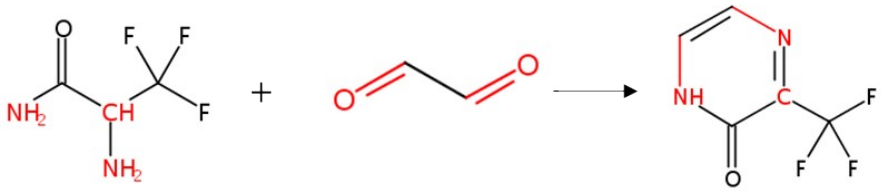


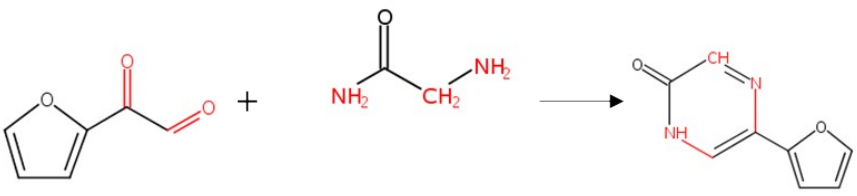
- HCl



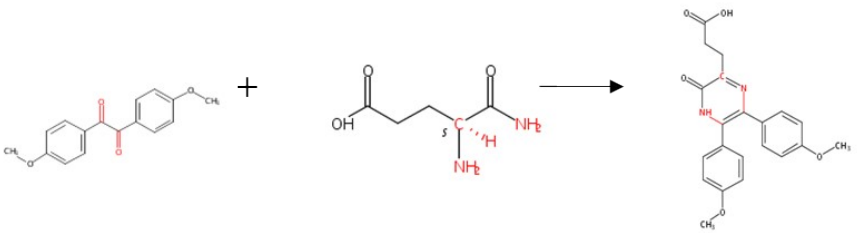


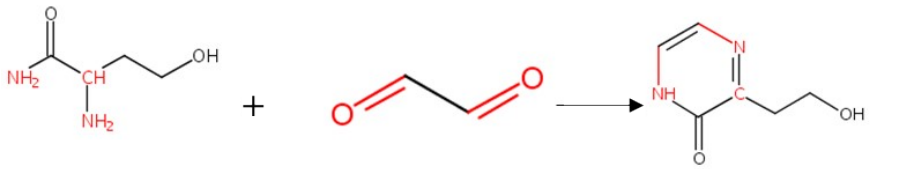
• HCl

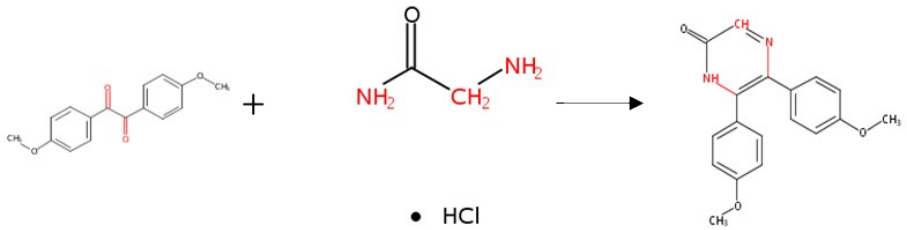




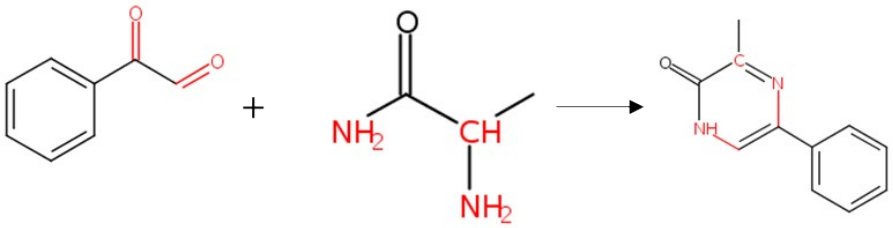
• HCl

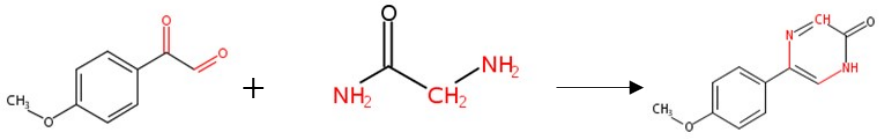




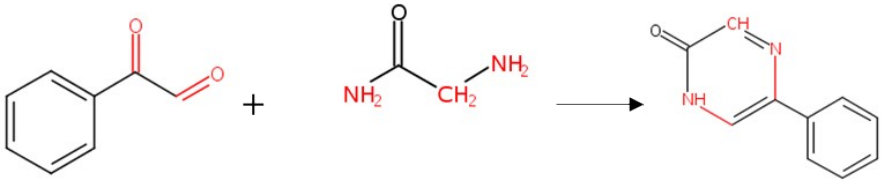


• HCl

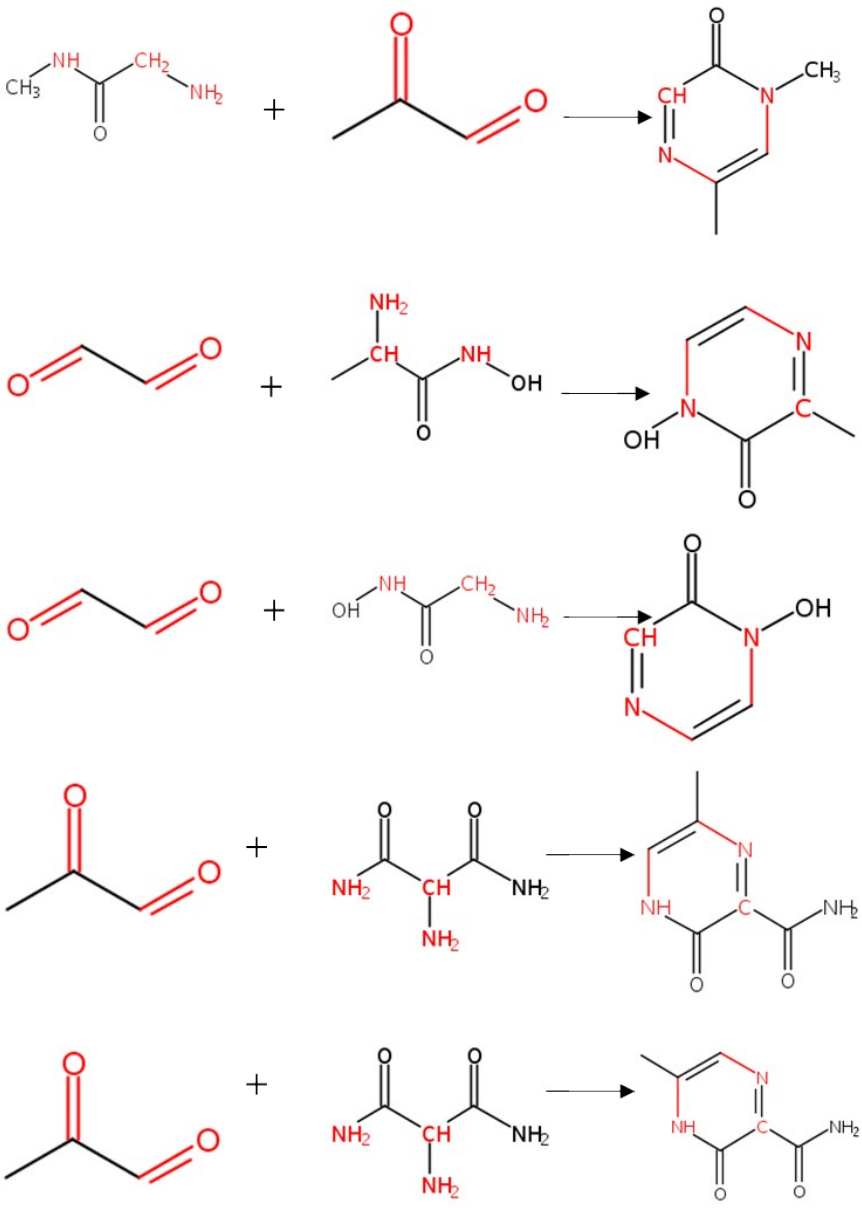


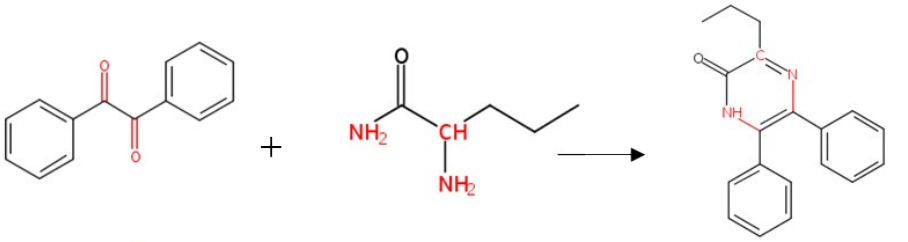


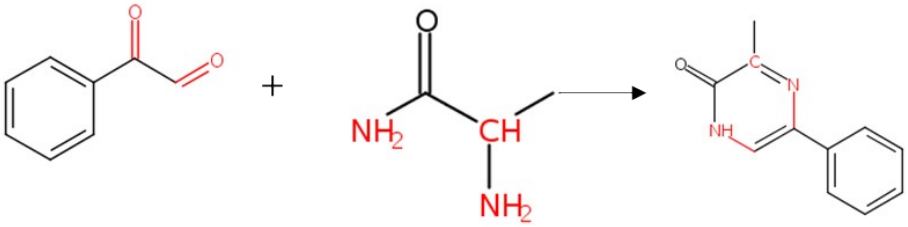
• HCl

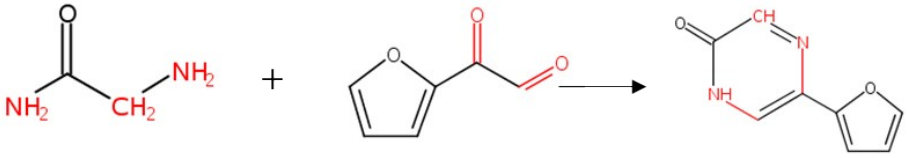


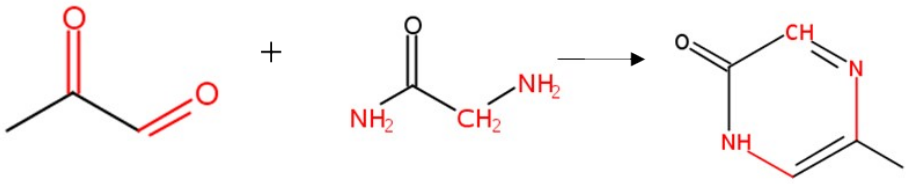
• HCl

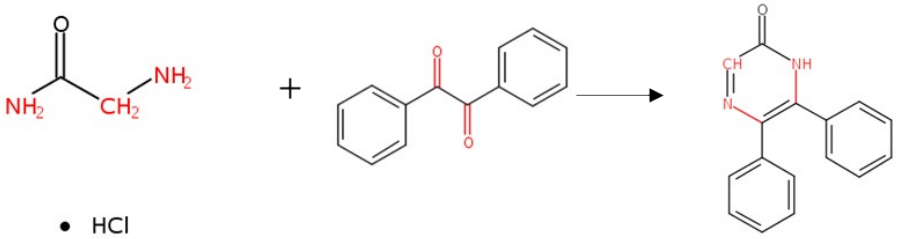




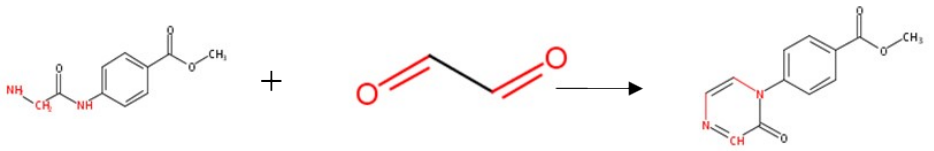


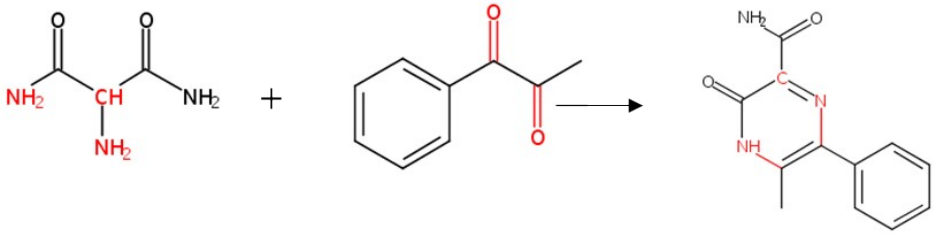


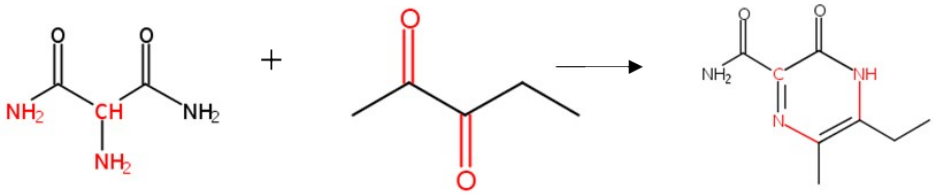


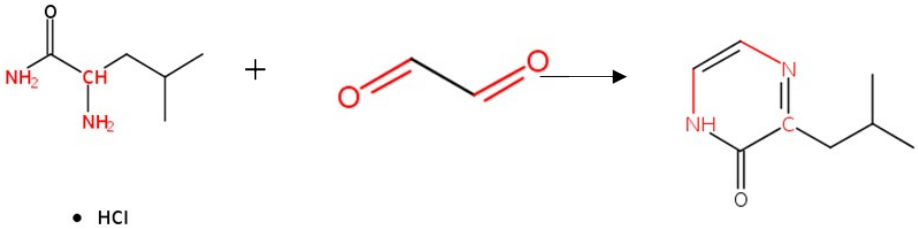


• HCl

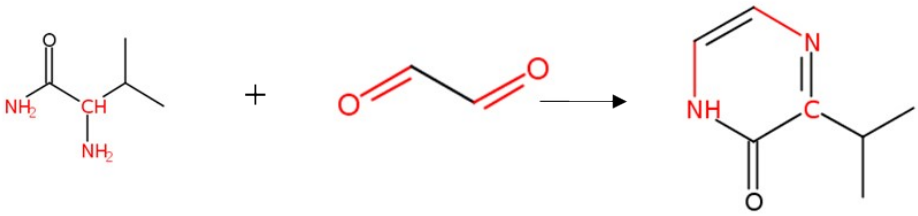




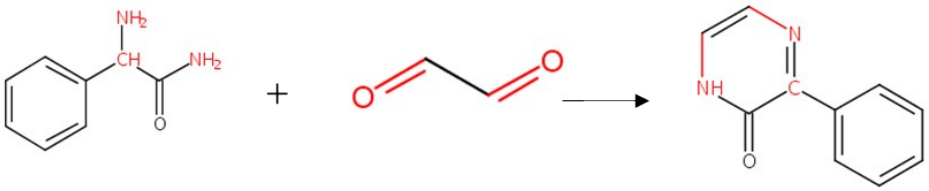


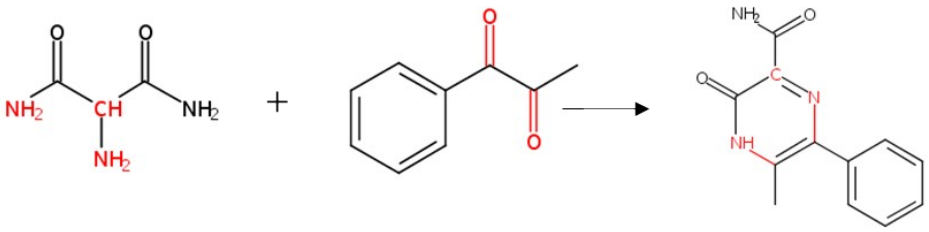


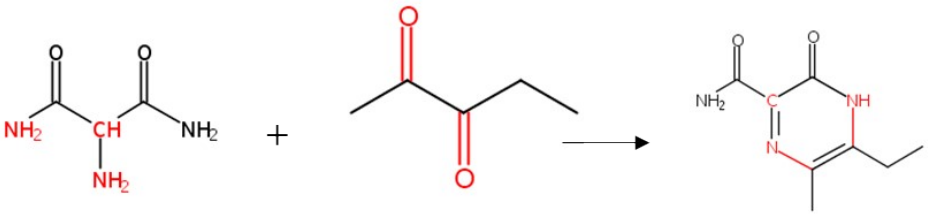
• HCl

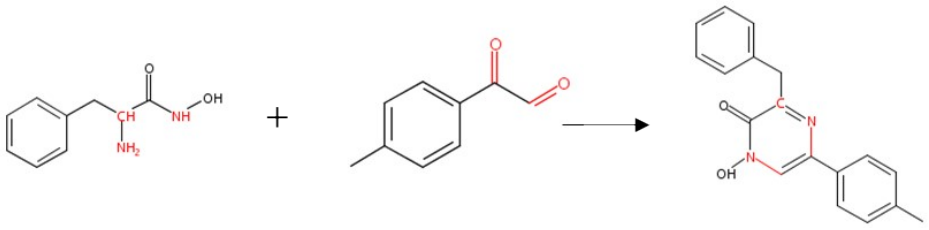


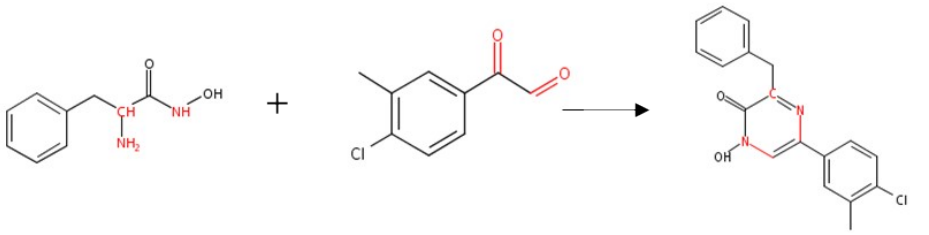
• HCl

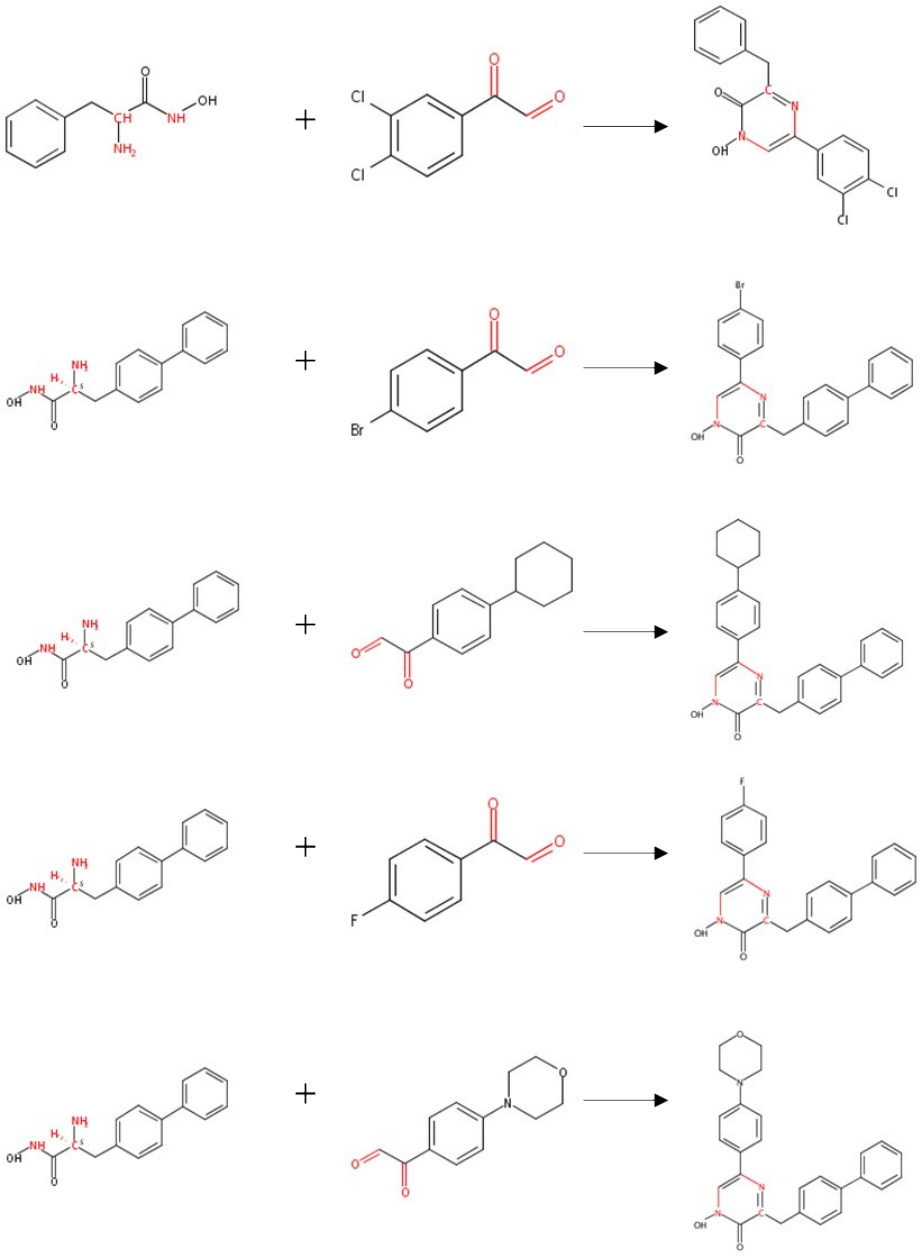


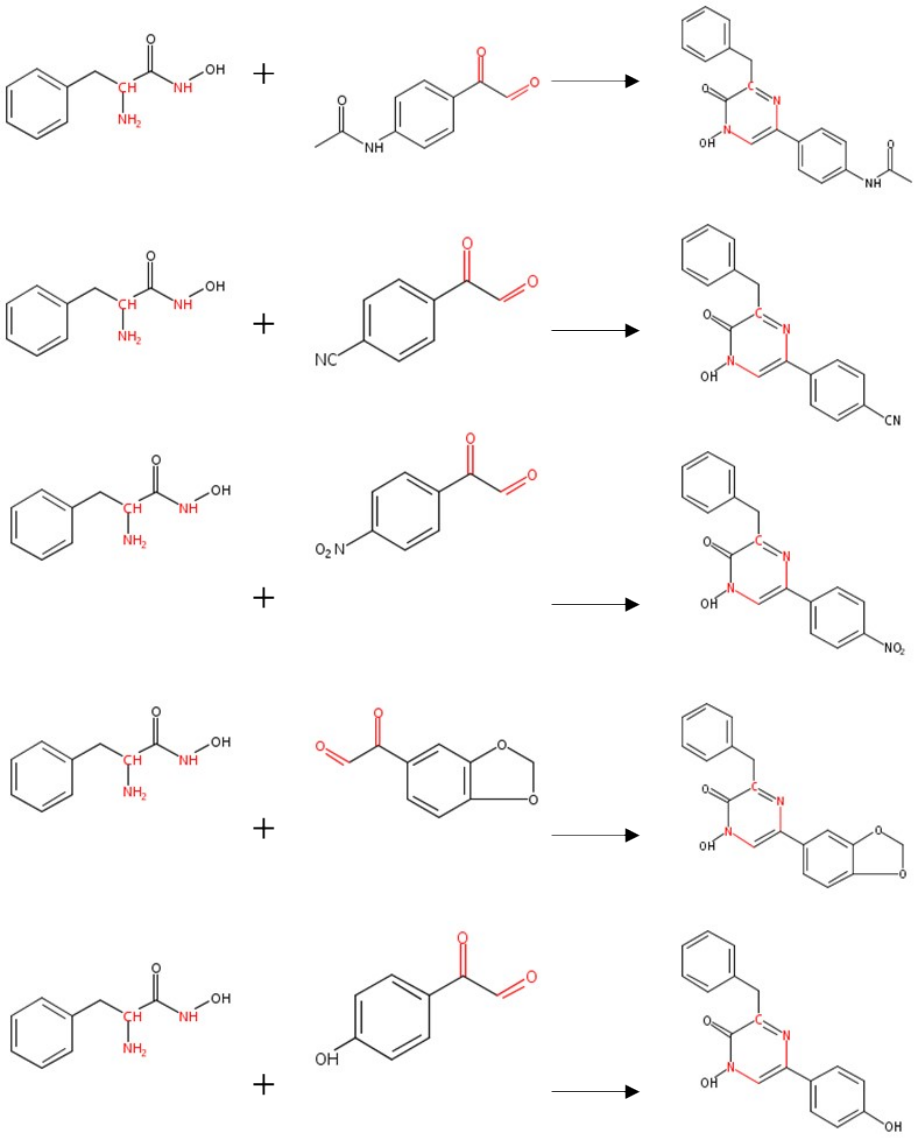


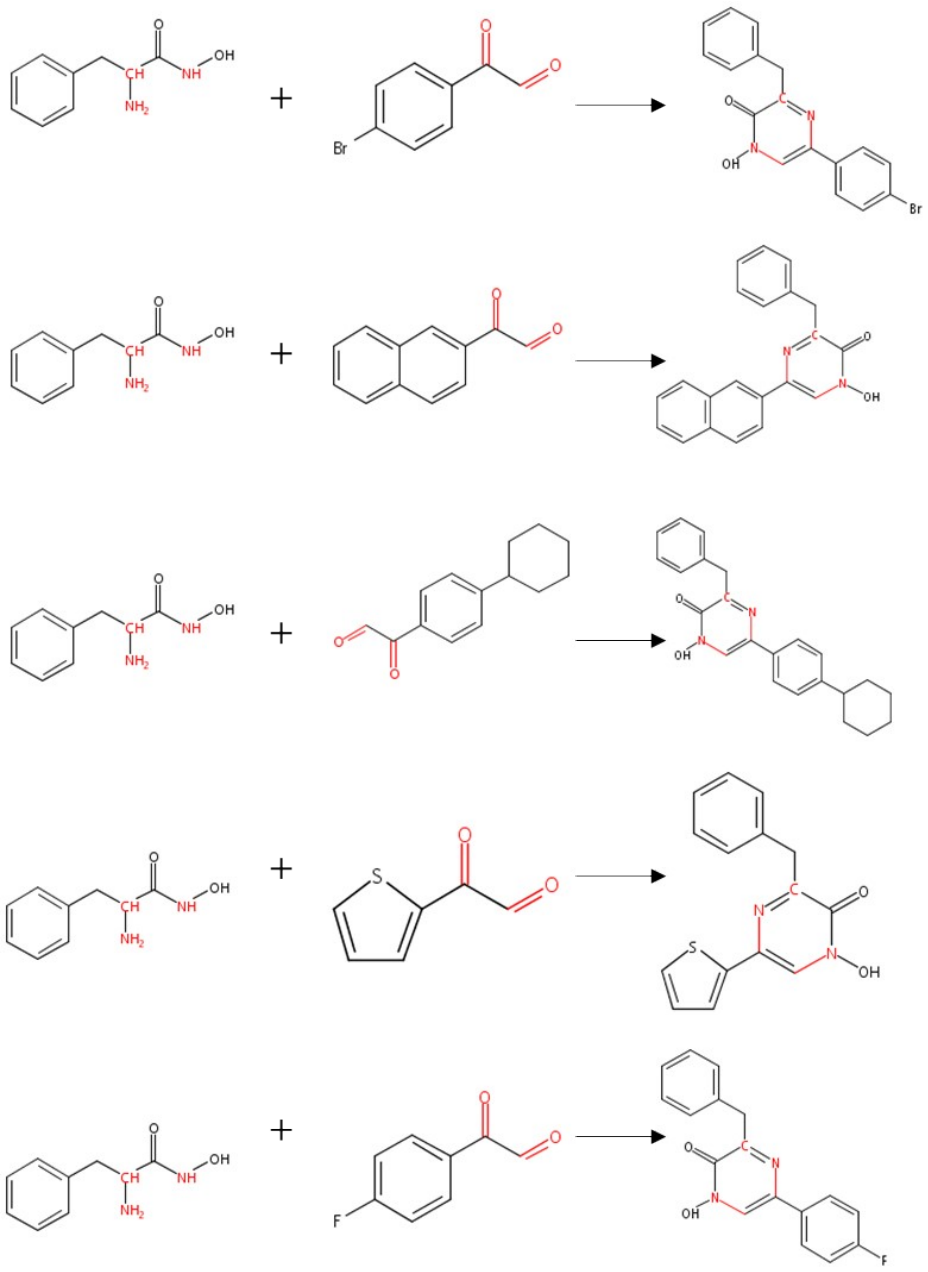


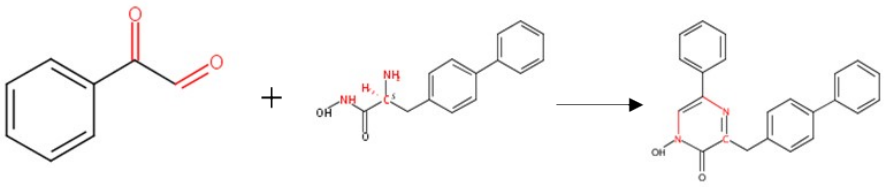


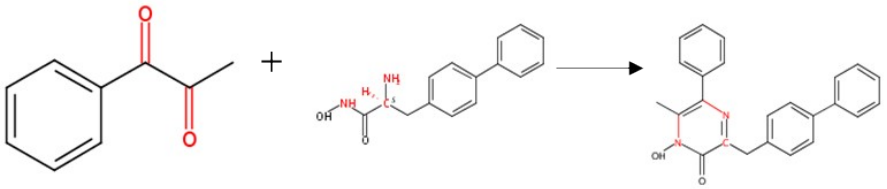


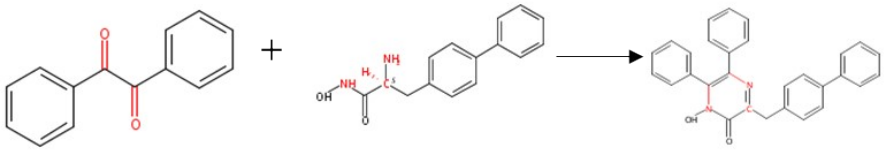


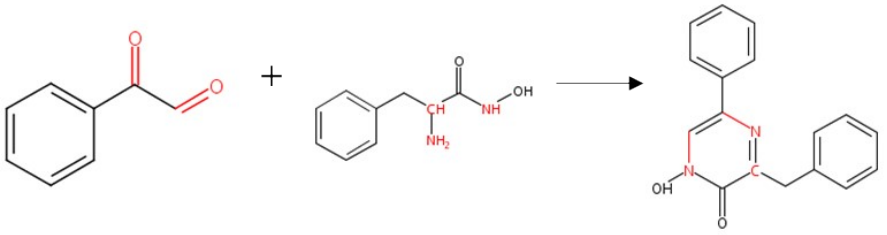


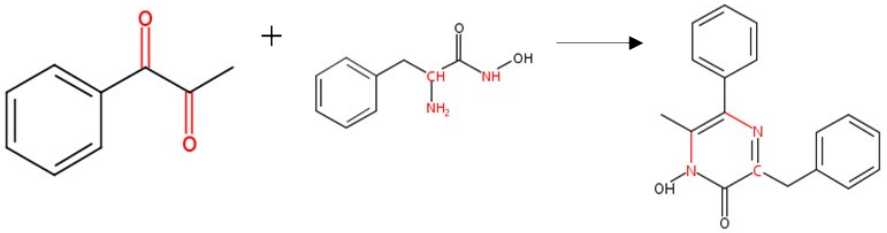


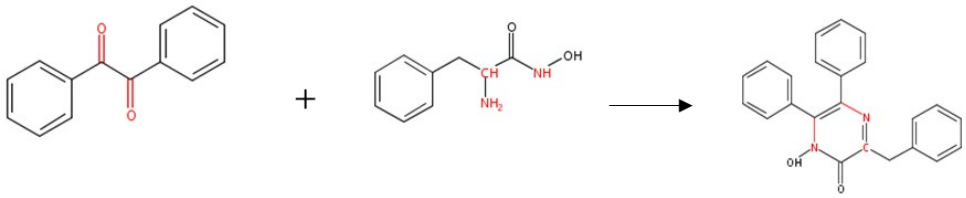


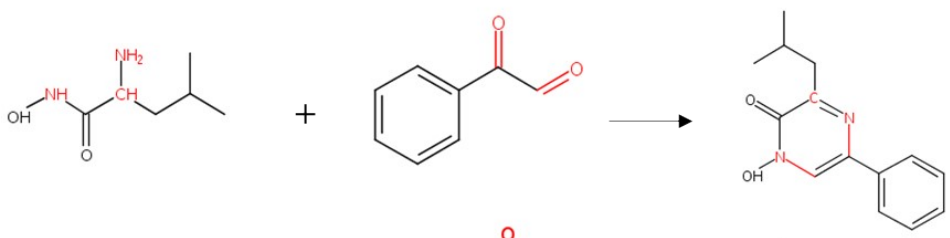


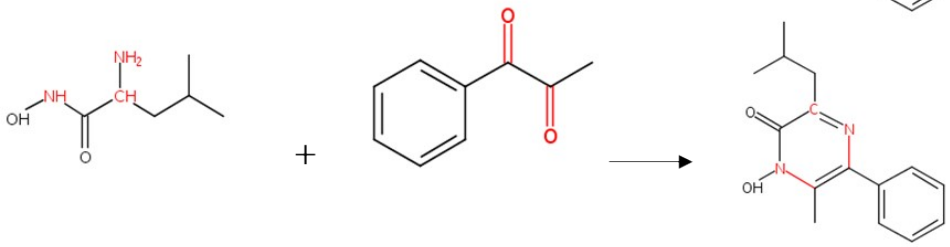


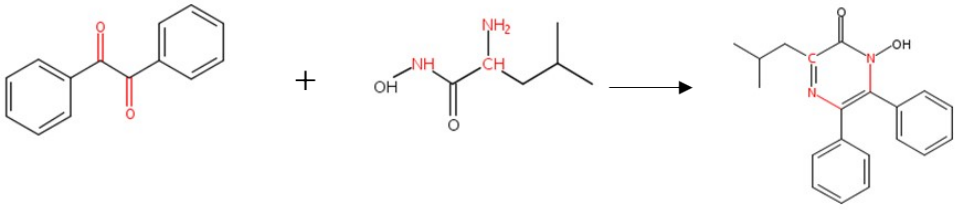


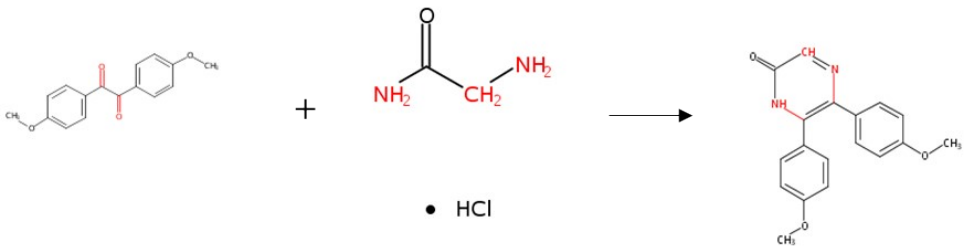




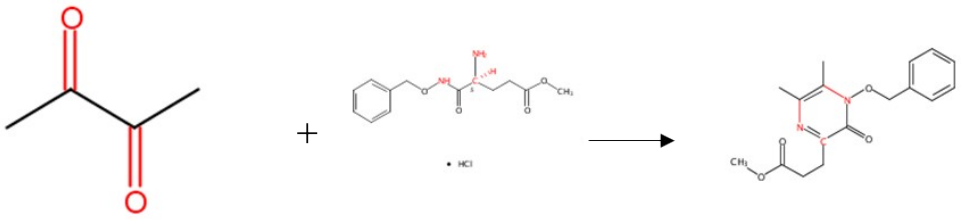


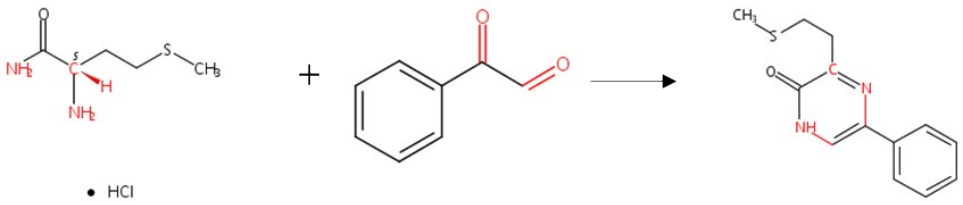


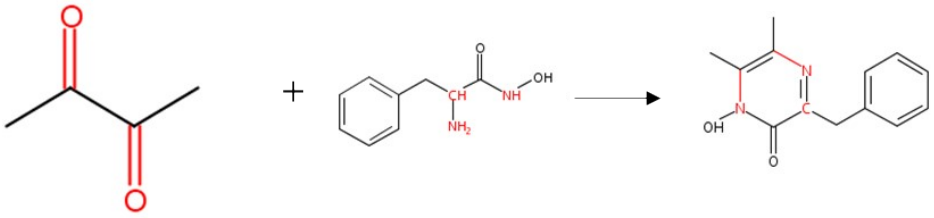


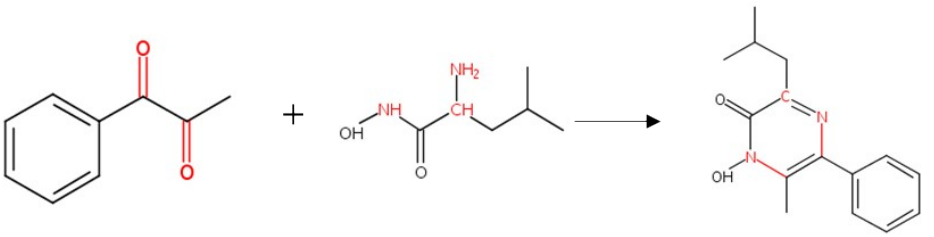


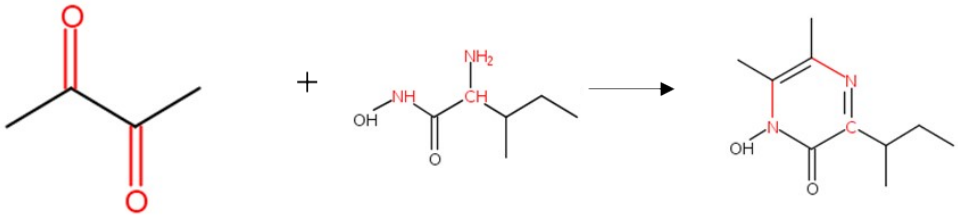
• HCl

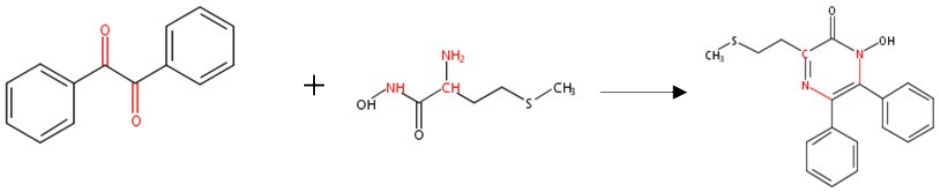


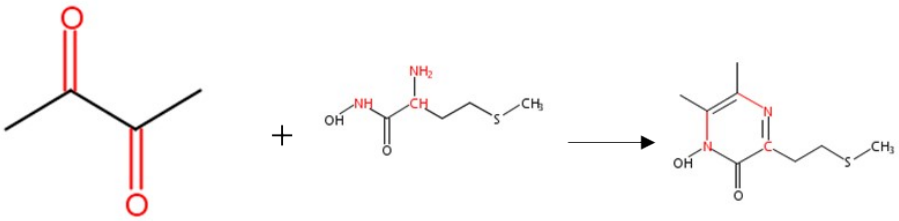


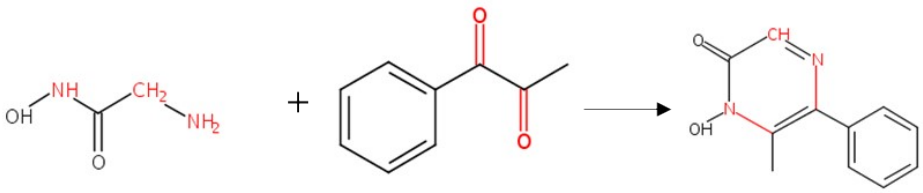


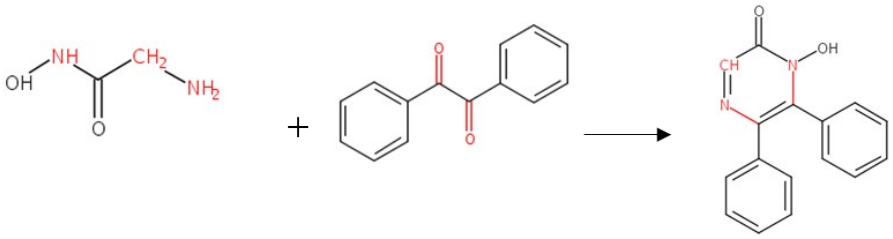


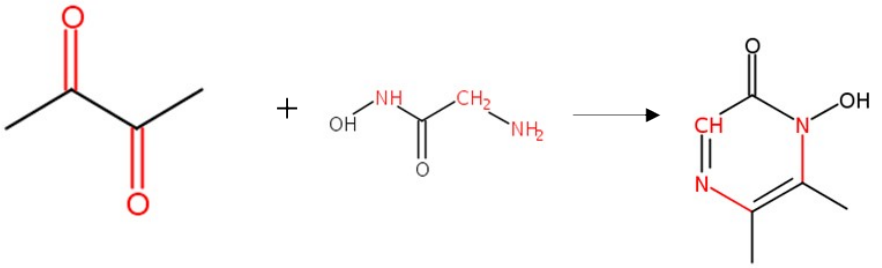


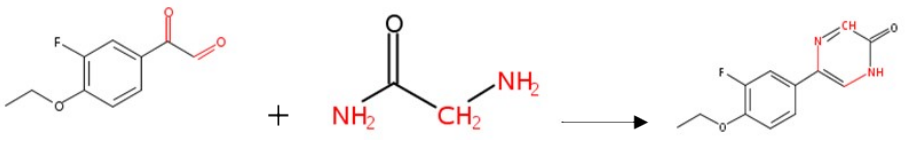


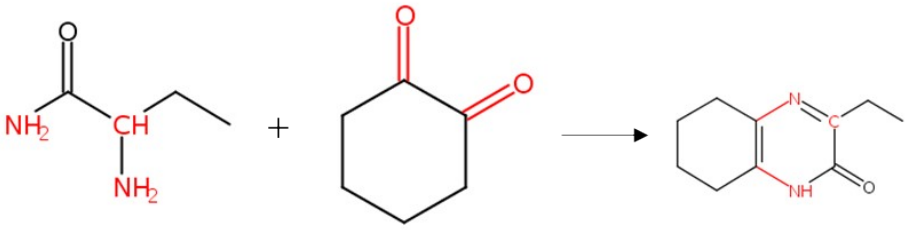


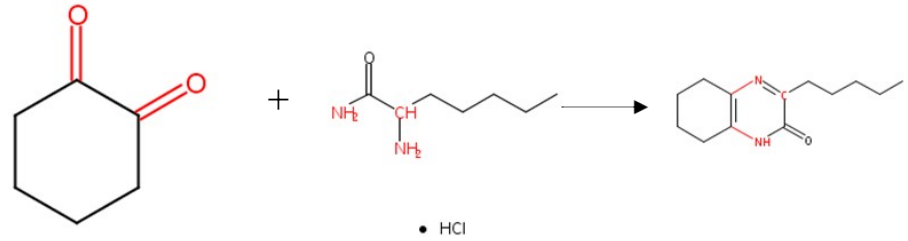


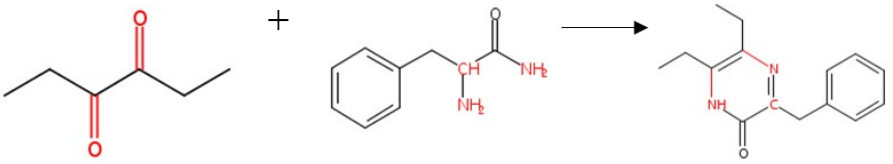


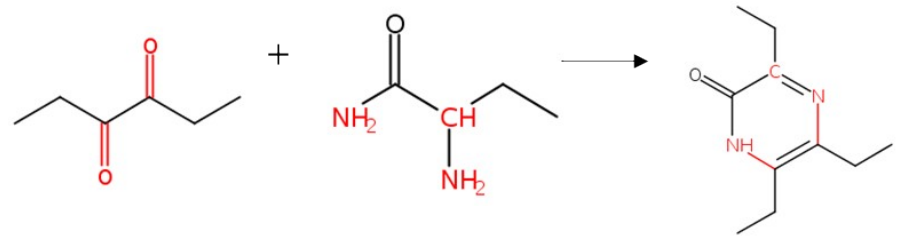


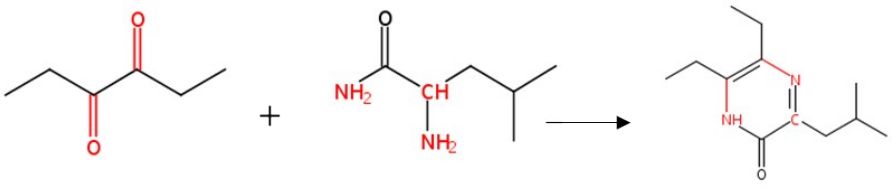


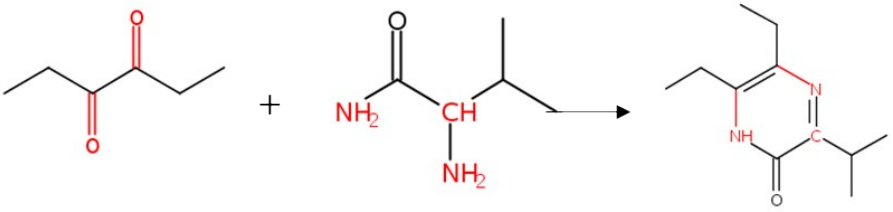


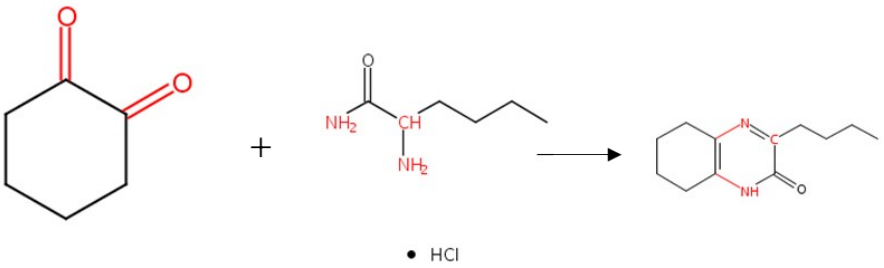

• HCl



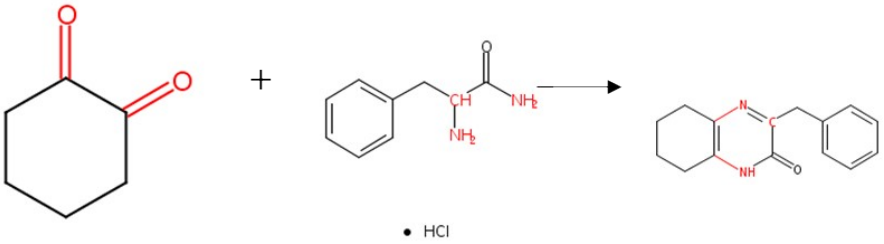




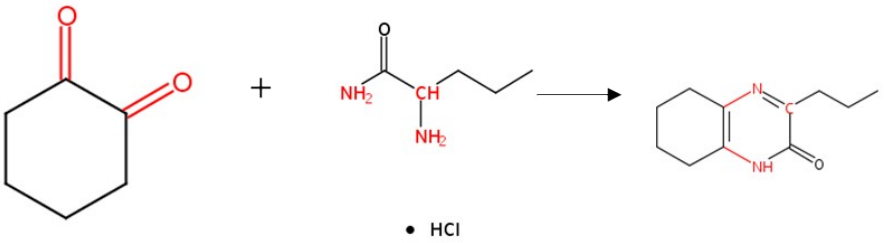




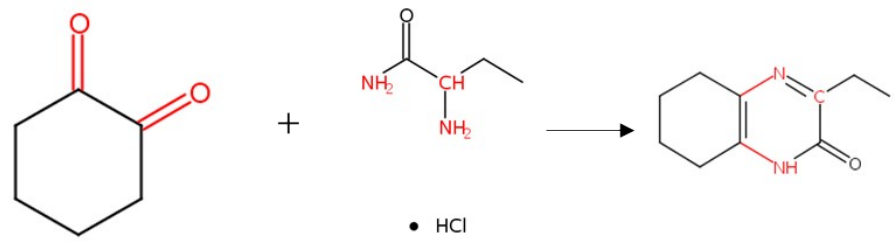
• HCl

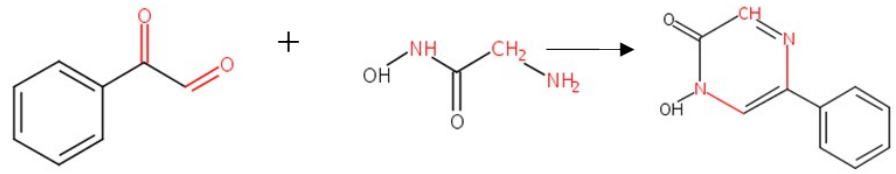


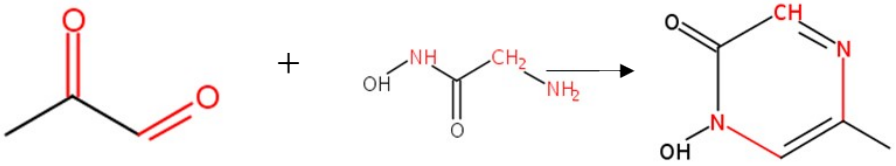
• HCl

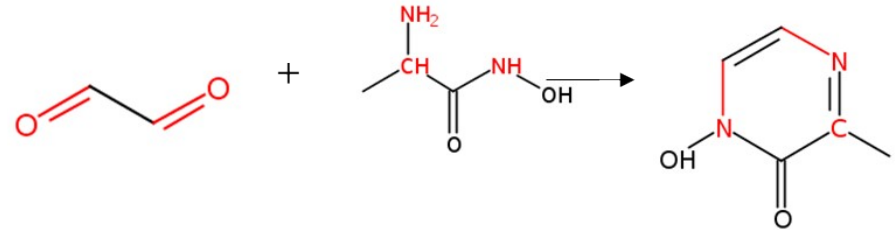


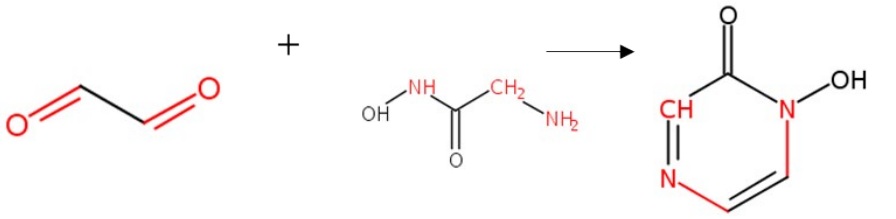
• HCl

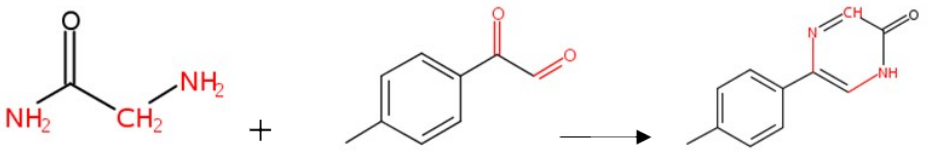

O
O
+
O
NH2
CH
NH2
• HCl
N
C
NH
O

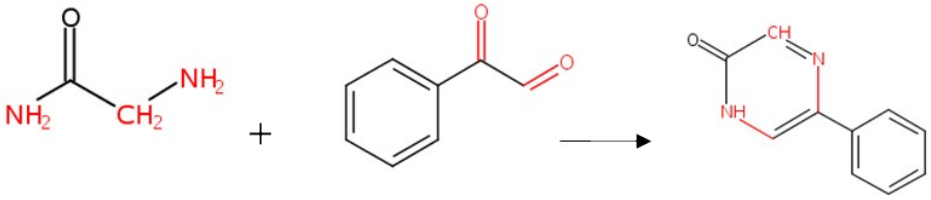

O
O
+
NH
OH
CH2
NH2
O
CH
N
O
N
OH

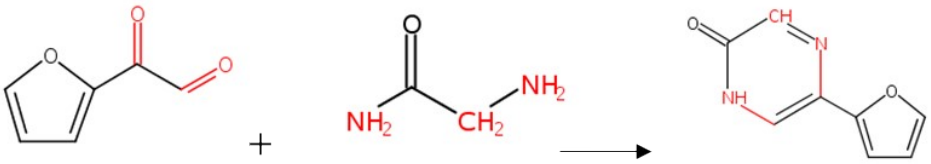

O
O
+
NH
OH
CH2
NH2
O
CH
O
N
N
OH

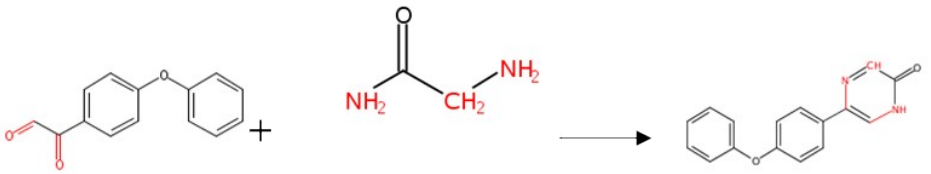
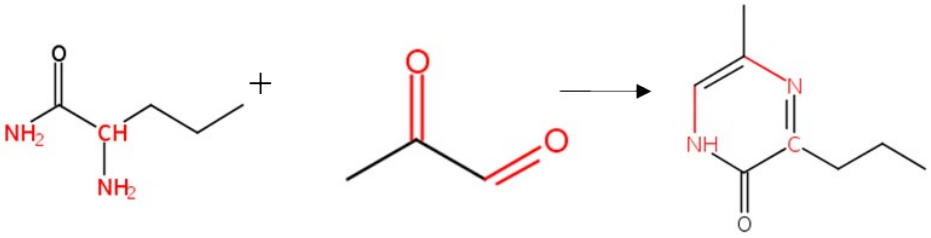
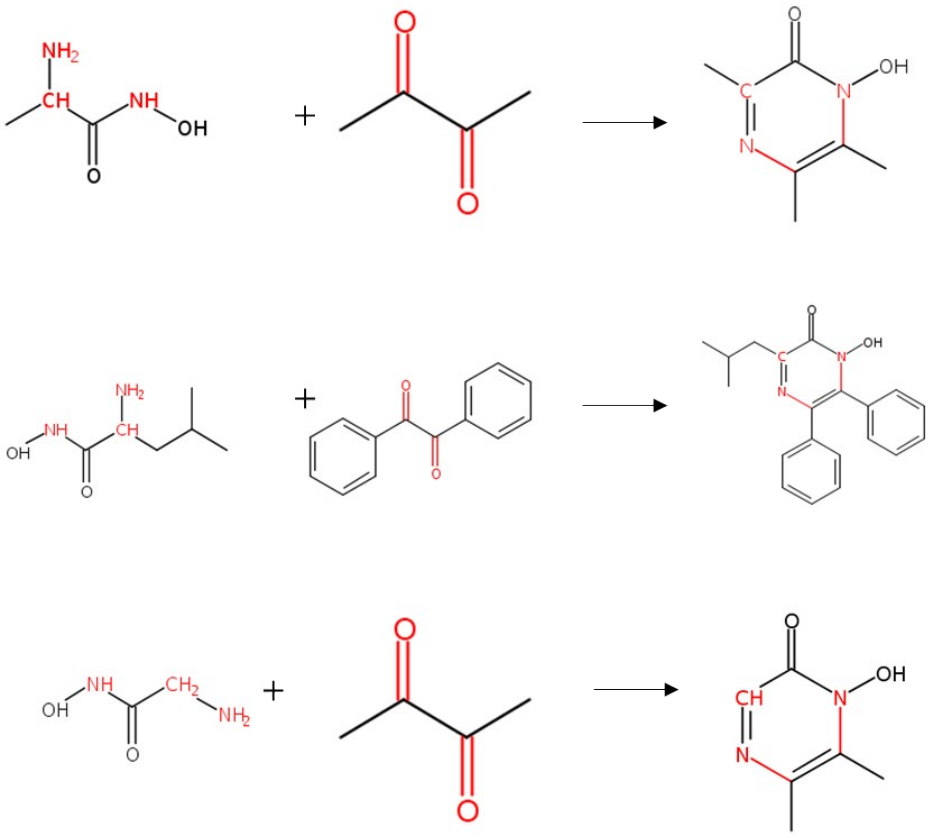

O
O
+
NH2
CH
NH
OH
O
N
N
C
OH
O


O
O
+
NH
OH
CH2
NH2
O
O
OH
CH
N
N









- HCl