# **Supporting Information**

# Synthesis and Crystal Characteristics Research of Nirmatrelvir

Bibo Jiang, <sup>†</sup> Gang Li, <sup>†</sup> Jun Yu, <sup>†</sup> Xiaoyan Xu, <sup>†</sup> Hongjuan Pan, <sup>†</sup> Chuanmeng Zhao, <sup>†</sup> Jialiang Zhong <sup>\*, †</sup> and Fuli Zhang<sup>\*, †</sup>

<sup>†</sup>Shanghai Institute of Pharmaceutical Industry, China State Institute of Pharmaceutical Industry, 285 Gebaini Road, Pudong District, Shanghai 201203, P. R. of China.

Correspondence should be addressed to Jialiang Zhong; zhongjialiang@sinopharm.com

# **Table of Contents**

1. HPLC data, NMR, and MS of Analytical spectra of compounds 14, 8, 10 an	id 10 (IBAC
solvate)	\$03-\$14
2. Analytical Spectrograms and Data of Nirmatrelvir 1	S14-S24
3. Analytical spectra of 15 & 16, 18, 22, 23, 2 in intermediates or nirmatrel	virS24-
S33	
4. Crystal void and surface area calculation of nirmatrelvir	\$34-\$35
5. Interaction energy calculation of nirmatrelvir	\$35-\$38

#### 1. HPLC, HRMS, and NMR Spectra of compounds 14, 8, 10 and 10, IBAC solvate. 1.1 HRMS, and NMR of Compound 14 1.1.1 HRMS



# 1.1.2 <sup>1</sup>H NMR (D<sub>2</sub>O)



# 1.1.3 <sup>13</sup>C NMR (DMSO-d6)







1.2.2 HRMS





# 1.2.3 <sup>1</sup>H NMR (DMSO-d6)



# 1.2.4 <sup>13</sup>C NMR (DMSO-d6)



1.2.5 SXRD



Table 1.Crystal data and structure refinement for cu\_d8v22277\_0m.Identification codecu\_d8v22277\_0mEmpirical formulaC16.33 H24.33 F3 N2 O4.33

Formula weight	375.04					
Temperature	293(2) К					
Wavelength	1.54178 Å					
Crystal system	Trigonal					
Space group	R 3 :H					
Unit cell dimensions	a = 23.7973(8) Å	a= 90°.				
	b = 23.7973(8) Å	b= 90°.				
	c = 10.5714(4) Å	g = 120°.				
Volume	5184.6(4) Å <sup>3</sup>					
Z	9					
Density (calculated)	1.081 Mg/m <sup>3</sup>					
Absorption coefficient	0.806 mm <sup>-1</sup>					
F(000)	1782					
Crystal size	0.200 x 0.150 x 0.120 mm <sup>3</sup>					
Theta range for data collection	7.061 to 67.388°.					
Index ranges	-26<=h<=28, -28<=k<=25, -12<	<=l<=12				
Reflections collected	20262					
Independent reflections	4076 [R(int) = 0.0617]					
Completeness to theta = 67.679°	98.2 %					
Absorption correction	Semi-empirical from equivale	nts				
Max. and min. transmission	0.7533 and 0.4876					
Refinement method	Full-matrix least-squares on F <sup>2</sup>	2				
Data / restraints / parameters	4076 / 136 / 300					
Goodness-of-fit on F <sup>2</sup>	1.059					
Final R indices [I>2sigma(I)]	R1 = 0.0565, wR2 = 0.1578					
R indices (all data)	R1 = 0.0593, wR2 = 0.1635					
Absolute structure parameter	osolute structure parameter 0.02(10)					
Extinction coefficient	0.0107(14)					
Largest diff. peak and hole	0.218 and -0.174 e.Å <sup>-3</sup>					

# 1.3 HPLC data, HRMS, and NMR of Compound 10

# 1.3.1.1 HPLC



1.3.2 HRMS

#### **Qualitative Compound Report**

WOK009-6

Data File	
Sample Type	
Instrument Name	
DA Method	

AcquiredTime

Sample Instrument 1 Default.m 7/29/2022 2:53:07 PM

20220729017-0729-027.d

 Position
 Vial 87

 Acq Method
 Vial 87

 Comment
 lot:220714

 DataFile
 D:\MassHu

Sample Name

D:\MassHunter\Data\2022-07\2022-07-29\20220729017-0729-027.d

Compound Table								
Compound Label	RT	Mass	Abund	Formula	Tgt Mass	Diff (ppm)		
Cpd 1: C23 H34 F3 N5 O5	0.144	517.2525	194751 (	23 H34 F3 N5 O5	517.2512	2.43		
Compound Label		RT Alg	orithm	Mass			<u>C</u>	-
Cpd 1: C23 H34 F3 N5 O5		0.144 Find	l By Formula	517.2525			CF <sub>3</sub>	
MS Zoomed Spectrum							- La	
x10 5 Cpd 1: C23 H34 F3	N5 O5: +	ESI Scan (0.12	7-0.144 min, 2 s	cans) Frag=110.0V 20	0220729017-072	29-027.d Subtr		~
	5	18.2599						0
3.5-		(M+H)+						
3-								/~INH
2.5-				540.3	2408			
2-				(M+I	Na)+		$  \langle \rangle^{m} \sqrt{F}$	1
1.5-								$_{\rm NH_2}$
1-							$  H \vee H H \vee A$	2
0.5-								
0				1.1				
505 510	515	520	525 530	535 54	0 545	550	10	
MS Spectrum Peak List		0	unts vs. mass-to	-Charge (m/z)				
m/z $Calc m/z$	Diff(ppp	a) Abund	Formula	Ton				
518 2599 518 2585	Din(ppi	2 7 373635	C23 H35 E3 N5 (	)5 (M+H)+				
540,2408 540,2404		0.67 194751	C23 H34 E3 N5 M	la O5 (M+Na)+				
End Of Report		10100	02010110101	(1114)	·			

# 1.3.3 <sup>1</sup>H NMR(DMSO-d6)



# 1.3.4 <sup>13</sup>C NMR(DMSO-d6)



# 1.3.5 <sup>1</sup>H NMR(DMSO-d6)



# 1.3.6 <sup>13</sup>C NMR(DMSO-d6)









#### 1.3.9 DSC



#### 1.3.10 TGA



# cu\_20221245\_0m

Table 1 Crystal data and structure refinement for cu_20221245_0m.						
Identification code	cu_20221245_0m					
Empirical formula	C <sub>29</sub> H <sub>46</sub> F <sub>3</sub> N <sub>5</sub> O <sub>7</sub>					
Formula weight	633.71	CF <sub>3</sub>				
Temperature/K	170.00	HN to O				
Crystal system	triclinic					
Space group	P1					
a/Å	14.6382(13)	$H V H H O NH_2$				
b/Å	8.7434(8)	· · · · · · · · · · · · · · · · · · ·				
c/Å	14.8811(14)	10 (IBAC solvate)				
α/°	90.011(2)					
β/°	112.727(4)					
γ/°	90.015(3)					
Volume/Å <sup>3</sup>	1756.7(3)					
Z	2					
ρ <sub>calc</sub> g/cm <sup>3</sup>	1.198					
µ/mm <sup>-1</sup>	0.811					
F(000)	676.0					
Crystal size/mm <sup>3</sup>	0.15 × 0.13 × 0.12					
Radiation	CuKα (λ = 1.54184)					
20 range for data collection/	°6.44 to 149.72					
Index ranges	$-18 \le h \le 18,  -10 \le$	k ≤ 10, -18 ≤ l ≤ 18				
Reflections collected	33066					
Independent reflections	13052 [R <sub>int</sub> = 0.048	4, R <sub>sigma</sub> = 0.0519]				
Data/restraints/parameters	13052/329/905					
Goodness-of-fit on F <sup>2</sup>	1.047					
Final R indexes [I>=2σ (I)]	$R_1 = 0.0428, wR_2 =$	0.1071				
Final R indexes [all data]	$R_1 = 0.0488, wR_2 =$	0.1148				
Largest diff. peak/hole / e $Å^{-3}$	0.44/-0.20					
Flack parameter	0.17(7)					

Datablock cu\_20221245\_0m - ellipsoid plot



# 2. HPLC, HRMS, and NMR Spectra of Nirmatrelvir (1)

# 2.1 HPLC, HRMS, NMR of nirmatrelvir (1) 2.1.1 HPLC data



#### 2.1.2 HRMS







# 2.1.4 <sup>1</sup>H NMR ((DMSO-d6-40℃)



# 2.1.5 $^{1}$ H NMR ((DMSO-d6-60 $^{\circ}$ C)



### 2.1.6 <sup>1</sup>H NMR (CDCl<sub>3</sub>)



# 2.1.7 <sup>1</sup>H NMR (CD<sub>3</sub>OD)



# 2.1.8 <sup>13</sup>C NMR (DMSO-d6)



# 2.1.10<sup>1</sup>H-<sup>1</sup>H COSY (DMSO-d6)



# 2.1.12 HMBC (DMSO-d6)



2.2 UV, FT-IR (KBr), DSC, TGA, PXRD of Nirmatrelvir (1) 2.2.1 UV



Software Informati	on	
Software Name:	UVProbe	
Version:	2.52	
Mode:	LabSolutions	
Data Information		
Data is:	Original	
Analyst	hanru	
Date/Time:	08/10/2022 02:4	5:15 PM
Comments:		
Instrument Informa	ation	
Instrument Name:	ACQ-3F-010	
Instrument Type:	UV-2600 Series	
Model (S/N):	UV-2600 (A116	85000543)
Wavelength Paper	pertiesj	100 00 to 400 00
Soon Spood:	e (nm.):	190.00 to 400.00
Scan Speed.		0.1
Auto Sampling Interval.	anval:	Disabled
Scan Mode:	ar ven.	Auto
oour mode.		
[Instrument Proper	ties]	
Instrument Type:		UV-2600 Series
Measuring Mode:		Absorbance
Silt width:		2.0
Light Source Chan	: wae Wavelength:	0.1 sec. 313.0 pm
Detector Unit:	ige wavelengut.	Direct
S/R Exchange:		Normal
Stair Correction:		OFF
[Attachment Prope	ution 1	
Attachment:	- Contraction of the Contraction	None
Addunient.		None
[Operation]		
Threshold:		0.0100000
Points:		4
InterPolate:		Disabled
Average:		Disabled
[Sample Preparation 1997]	on Properties]	
Weight:		13.03mg
Volume:		25ml
Dilution:		25
Path Length:		Complete and the second second
Additional informat	uon.	Sample name:///OF

# K009

Lot:220721 No:20220804027 Solvent:Methanol

## 2.2.2 FT-IR (KBr)



20220804027\_144307 - RawData







#### 2.2.5 XRPD

7/25/2022 6:08:00 PM





7/25/2022 6:08:00 PM

5.000/44.552/0.020/0.155/40.0KV/40.0IIIA								
Index	Angle	d Value	Net Intensity	Gross Intensity	Rel. Intensity	EWHM		
1	7.577 °	11.65886 Å	869.145	1100.54	19.1%	0.105		
2	9.730 °	9.08326 Å	831.810	1207.50	18.3%	0.108		
3	10.287 °	8.59259 Å	292.138	729.403	6.4%	0.108		
4	10.766 *	8.21069 Å	174.729	660.766	3.8%	0.193		
5	11.363 °	7.78127 Å	753.586	1294.71	16.6%	0.105		
6	11.849 *	7.46314 Å	2769.60	3351.13	60.8%	0.113		
7	12.675 °	6.97841 Å	4553.15	5194.07	100.0%	0.113		
8	13.327 °	6.63840 Å	284.646	964.154	6.3%	0.100		
9	14.564 °	6.07704 Å	299.374	1032.07	6.6%	0.314		
10	15.628 °	5.66574 Å	4248.56	5005.96	93.3%	0.118		
11	17.263 °	5.13256 Å	1221.81	1979.35	26.8%	0.108		
12	17.798 °	4.97961 Å	1091.26	1838.91	24.0%	0.108		
13	18.270 °	4.85203 Å	2942.99	3677.82	64.6%	0.110		
14	18.819 °	4.71167 Å	641.644	1356.74	14.1%	0.110		
15	19.791 °	4.48225 Å	2199.45	2866.90	48.3%	0.122		
16	20.466 °	4.33611 Å	2192.51	2817.42	48.2%	0.113		
17	20.931 °	4.24079 Å	990.151	1581.17	21.7%	0.117		
18	21.583 °	4.11415 Å	666.637	1214.89	14.6%	0.147		
19	22.177 °	4.00523 Å	1498.02	2015.01	32.9%	0.115		
20	22.454 °	3.95634 Å	225.152	725.481	4.9%	0.142		
21	23.002 *	3.86340 Å	169.267	648.115	3.7%	0.109		
22	23.518 *	3.77972 Å	507.188	977.983	11.1%	0.122		
23	24.234 °	3.66972 Å	131.138	591.107	2.9%	0.100		
24	24.699 *	3.60158 Å	560.629	1008.83	12.3%	0.109		
25	25.286 *	3.51938 Å	369.786	812.381	8.1%	0.103		
26	26.945 °	3.30637 Å	300.770	777.976	6.6%	0.174		
27	27.206 *	3.27515 Å	359.588	848.364	7.9%	0.158		
28	27.985 *	3.18572 Å	116.940	633.041	2.6%	0.174		
29	28.622 *	3.11628 Å	158.250	688.880	3.5%	0.121		
30	29.425 °	3.03302 Å	448.539	991.643	9.9%	0.122		
31	29.898 *	2.98612 Å	305.497	853.775	6.7%	0.209		
32	30.206 *	2.95635 Å	139.579	689.164	3.1%	0.286		
33	31.718 *	2.81883 Å	95.8101	628.206	2.1%	0.266		
34	32.622 °	2.74275 Å	279.750	783.028	6.1%	0.105		
35	33.576 °	2.66699 Å	262.176	724.289	5.8%	0.120		
36	35.629 *	2.51787 Å	209.480	632.959	4.6%	0.150		

#### 20220725061(WOK009 220721) 3.000/44.992/0.020/0.15s/40.0kV/40.0mA

# 3. Analytical spectra of 15 & 16, 18, 22, 23, 2 in intermediates or

# Nirmatrelvir

3.1 HPLC data, HRMS, and NMR of Compound 15&16 (compounds 15 and 16 are not separated)

3.1.1 HPLC



100.00

100.00

#### Total:

3.1.2 HRMS

#### **Qualitative Compound Report** Data File 20220722051-0725-003.d WOK009-5CH3 Sample Name Sample Type Vial 2 Sample Position Inst Instrument 1 Default.m Acq Method Instrument DA Method lot:220721 Comment iredTim 7/25/2022 11:04:59 AM DataFile D:\MassHunter\Data\2022-07\2022-07-25\20220722051-0725-003.d Compound Table Diff Compound Label Cpd 1: C17 H25 F3 N2 O RT 0.174 Abund 22527 Formula C17 H25 F3 N2 O4 Tgt Mass 378.1766 (ppm) 0.15 Mass 378.1767 Compound Label RT Algorithm 0.174 Find By Formula Mass 378.1767 Cpd 1: C17 H25 F3 N2 O4 MS Zoomed Spectrum CF<sub>3</sub> x10 5 Cpd 1: C17 H25 F3 N2 O4: +ESI Scan (0.124-0.360 min, 15 scans) Frag=120.0V 20220722051-0725-003.d Subt... 4.5 HN 0 4 3.5 379.1839 (M+H)+ 3 2.5 2 1.5 401.1660 (M+Na)+ 370 410 415 375 385 390 395 Counts vs. Mass-to-Charge (m/z) 400 405 38 MS Spectrum Peak List z Abund Formula 379.1839 379.1839 -0.02 1 225276 C17 H26 F3 N2 04 401.166 401.1659 0.45 1 180210 C17 H25 F3 N2 Na 04 Ion 379.1839 401.166 End Of Report (M+H)+ (M+Na)+ 15 & 16

# 3.1.3 <sup>1</sup>H NMR (DMSO-d6)



# 3.1.4 <sup>13</sup>C NMR (DMSO-d6)



#### 3.2 HPLC data, HRMS, and NMR of Compound 18

#### 3.2.1 HPLC data



#### 3.2.2 HRMS



# 3.2.3 <sup>1</sup>H NMR (DMSO-d6)



3.2.4 <sup>13</sup>C NMR (DMSO-d6)



#### 3.3 HPLC data, HRMS, and NMR of Compound 22

#### 3.3.1 HPLC



3.3.2 HRMS

#### **Qualitative Compound Report**



# 3.3.3 <sup>1</sup>H NMR (DMSO-d6)



# 3.3.4 <sup>13</sup>C NMR (DMSO-d6)



# 3.4 HPLC data, HRMS, and NMR of Compound 23

# 3.4.1 HPLC



3.4.2 HRMS

#### Qualitative Compound Report



# 3.4.3 <sup>1</sup>H NMR (DMSO-d6)



<sup>3.4.4 &</sup>lt;sup>13</sup>C NMR (DMSO-d6)



# 3.5 purity data of Compound 2 3.5.1 GC (chemical purity)



# 3.5.2 HPLC (chiral purity)



Integr	ntegration Results										
No.	Peak Name	Retention Time	Area	Height	Relative Area	Relative Height	Amount				
		min	mAU*min	mAU	%	%	n.a.				
1		5.190	126.741	992.882	99.99	99.99	n.a.				
2		5.905	0.007	0.079	0.01	0.01	n.a.				
Total:			126.748	992.961	100.00	100.00					

4 Crystal void and surface calculation of nirmatrelvir



Figure 1. Crystal form A molecular void map

#### Crystal form A

Type Crystal Voids Resolution High (Standard) Isovalue 0.002 Volume 481.04 Å<sup>3</sup> Area 1267.06 Å<sup>2</sup> Globularity 0.234 Asphericity 0.090



Figure 2. Crystal form B molecular void map

#### **Crystal form B**

Type Crystal Voids Resolution High (Standard) Isovalue 0.002 Volume 386.86 Å<sup>3</sup> Area 1068.44 Å<sup>2</sup> Globularity 0.240 Asphericity 0.105

#### 5 Interaction energy calculation of nirmatrelvir

The pairwise intermolecular interaction energy was calculated using CrystalExplorer software (version 17.5) with experimental crystal geometry.1 Considering the uncertainty of hydrogen position by single crystal X-ray diffraction, the hydrogen positions were normalized to standard neutron diffraction values during the calculation. The total intermolecular interaction energy for given molecule, is summed up the electrostatic, polarization, dispersion, and exchange-repulsion components terms with scale factors of 1.057, 0.740, 0.871, and 0.618. The intermolecular interaction is neglected with molecule-molecule distance more than 3.8 Å.

Ref. 1 M. J. Turner, J. J. McKinnon, S. K. Wolff, D. J. Grimwood, P. R. Spackman, D. Jayatilaka and M. A. Spackman, CrystalExplorer17 (2017). University of Western Australia.
Crystal form A

Interaction Energies (kJ/mol) R is the distance between molecular centroids (mean atomic position) in Å.Total energies, only reported for two benchmarked energy models, are the sum of the four energy components, scaled appropriately (see the scale factor table below)

Ν	Symop	R	Electron Density	E_ele	E_pol	E_dis	E_rep	E_tot
2	x+1/2, -y+1/2, -z	11.24	B3LYP/6- 31G(d,p)	-7.0	-1.7	-26.9	16.7	-21.8
2	-x, γ+1/2, -z+1/2	10.53	B3LYP/6- 31G(d,p)	-3.7	-1.6	-18.7	5.3	-18.1
2	х, ү, z	9.21	B3LYP/6- 31G(d,p)	-15.8	-6.4	-52.8	36.2	-45.0
2	-х+1/2, -у, z+1/2	10.21	B3LYP/6- 31G(d,p)	-30.3	-8.9	-18.9	39.9	-30.5
2	x+1/2, -y+1/2, -z	10.30	B3LYP/6- 31G(d,p)	-39.2	-12.2	-29.9	31.7	-56.9
2	-х+1/2, -у, z+1/2	11.78	B3LYP/6- 31G(d,p)	0.4	-0.9	-15.2	4.2	-11.0
2	-х, у+1/2, -z+1/2	7.79	B3LYP/6- 31G(d,p)	-5.2	-3.2	-45.8	16.8	-35.0

Table S1. Intermolecular interaction energies (kJ mol<sup>-1</sup>) of crystal form A estimated using B3LYP/6-31G (d, p) dispersion-corrected DFT models (nirmatrelvir as the central molecule)







**Dispersion Energy** 

Total Energy



**Coulomb Energy** 

# Figure 3. Energy frame diagram

#### **Crystal form B**

Interaction Energies (kJ/mol) R is the distance between molecular centroids (mean atomic position) in Å. Total energies, only reported for two benchmarked energy models, are the sum of the four energy components, scaled appropriately (see the scale factor table below)

Ν	Symop	R	Electron Density	E_ele	E_pol	E_dis	E_rep	E_tot
2	-x, y+1/2, -z+1/2	10.65	B3LYP/6-31G(d,p)	-1.1	-10.0	-18.2	1.2	-23.6
2	x, γ, z	9.69	B3LYP/6-31G(d,p)	-36.8	-16.2	-41.6	23.5	-72.6
2	-x+1/2, -y, z+1/2	9.84	B3LYP/6-31G(d,p)	-65.8	-28.8	-28.7	24.2	-100.9
2	x+1/2, -γ+1/2, -z	11.74	B3LYP/6-31G(d,p)	7.0	-4.5	-18.6	4.5	-9.3
2	x+1/2, -γ+1/2, -z	10.75	B3LYP/6-31G(d,p)	-9.8	-6.8	-27.5	19.1	-27.5
2	-x, y+1/2, -z+1/2	7.85	B3LYP/6-31G(d,p)	-11.8	-27.4	-43.4	11.9	-63.1
2	-x+1/2, -y, z+1/2	11.79	B3LYP/6-31G(d,p)	-13.8	-9.6	-16.8	2.3	-34.9

------

Table S2. Intermolecular interaction energies (kJ mol<sup>-1</sup>) of crystal form B estimated using B3LYP/6-31G (d, p) dispersion-corrected DFT models (nirmatrelvir as the central molecule)





**Dispersion Energy** 



Total Energy



Coulomb Energy Figure 4. Energy frame diagram