Sequential multicomponent site-selective synthesis of 4-iodo and 5-

iodopyrrole-3-carboxaldehydes by tuning the conditions from a common set of

starting materials

Sachin Choudhary,^a Jyothi Yadav,^a Mamta,^a Amol Prakash Pawar,^a Satheeshvarma Vanaparthi,^a

Nisar A. Mir,^a Eldhose Iype,^b Ratika Sharma,^c Rajni Kant,^c and Indresh Kumar*^a

^aDepartment of Chemistry, Birla Institute of Technology & Science, Pilani 333 031, India

^bDepartment of Chemical Engineering, BITS Pilani, Dubai Campus, Dubai, UAE

^cX-ray Crystallography Laboratory, Post-Graduate Department of Physics & Electronics, University of Jammu, 180 006, (Jammu & Kashmir) India

Corresponding Email: *indresh.kumar@pilani.bits-pilani.ac.in*, *indresh.chemistry@gmail.com*

Table of Contents

Controlled experimental study	
¹ H and ¹³ C-NMR for synthesized compounds	
Single Crystal X-ray study and data for 5bc and 5bh	S41-S47
DFT-calculation chart and optimized structures and coordinates	

Controlled experimental study for more reaction insight:

To understand whether I₂-mediated C4-iodination of dihydropyrrole intermediate **E** is taking first followed by aromatization or aromatization is taking place first followed by C4-iodinaion. We carried out few controlled experiments, as shown in Scheme S1(b), the selective formation of 5-iodopyrrole (**7ac**, 25% yield) was observed in DMSO with I₂ (1.0 equiv.) through iodination on preformed pyrrole **6ac**. While, the formation of 5-iodopyrrole (**7ac**, 28% yield) and 4,5-diiodopyrole (**12**, 18% yield) were observed with I₂ (3.0 equiv.) in the same solvent. These controlled experiments and the DFT-calculations suggested that the intermediate **E** is more nucleophilic at β -(C4)-position than the pyrrole at C5-position. Thus, iodination at the C4-position of dihydropyrrole-intermediate take place first before the aromatization.



Scheme S1: (a) Understanding on the regioselective iodination on C4-position of pyrrole, (b) controlled experiment for selective iodination.









MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
325.1182	325.1183	0.26	1	161297.91	C18H16N2O4	(M+H)+
326.1215	326.1214	-0.06	1	30724.2	C18H16N2O4	(M+H)+
327.1242	327.124	-0.73	1	4269.75	C18H16N2O4	(M+H)+
328.1241	328.1266	7.6	1	453.82	C18H16N2O4	(M+H)+
347,0984	347,1002	5.38	1	117.52	C18H16N2O4	(M+Na)+

Instrument Info : Agilent Technologies 6545 Q-TOF LC/MS



HRMS Facility, BITS Pilani, Pilani Campus

MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
323.1026	323.1026	0.12	1	126054.85	C18H14N2O4	(M+H)+
324.1058	324.1058	0.1	1	24626.59	C18H14N2O4	(M+H)+
325.1119	325.1083	-11.02	1	3281.99	C18H14N2O4	(M+H)+
345.0845	345.0846	0.31	1	31167.76	C18H14N2O4	(M+Na)+
346.0878	346.0877	-0.05	1	6190.15	C18H14N2O4	(M+Na)+
347.091	347.0903	-2.06	1	947.12	C18H14N2O4	(M+Na)+

Instrument Info : Agilent Technologies 6545 Q-TOF LC/MS



MS Spectrum Peak	List	
------------------	------	--

m/z	Calc m/z	Diff(ppm)	z	Abund	Formula	Ion
448.9961	448.9993	7.01	1	174624.5	C18H13IN2O4	(M+H)+
449.9993	450.0024	7.02	1	34964.94	C18H13IN2O4	(M+H)+
451.0022	451.005	6.27	1	4789.29	C18H13IN2O4	(M+H)+
452.0037	452.0075	8.43	1	567.07	C18H13IN2O4	(M+H)+
470.9777	470.9812	7.47	1	17094.27	C18H13IN2O4	(M+Na)+
471.981	471.9844	7.17	1	3429.8	C18H13IN2O4	(M+Na)+
472.9838	472.9869	6.66	1	466.59	C18H13IN2O4	(M+Na)+
473,9898	473,9895	-0.73	1	54.27	C18H13IN2O4	(M+Na)+

Instrument Info : Agilent Technologies 6545 Q-TOF LC/MS



HRMS Facility, BITS Pilani, Pilani Campus



S8

S19

S26

S31

1D-NOE spectra to confirm the regioselective outcome for 7bc

<u>Crystal structure data for (5bc) and (7bh):</u> 1-(4-chlorophenyl)-4-iodo-2-(4-nitrophenyl)-1H-pyrrole-3-carbaldehyde (5bc):

(CCDC NO. 1471810)

The title compound1-(4-chlorophenyl)-4-iodo-2-(4-nitrophenyl)-1H-pyrrole-3-carbaldehyde (**5bc**), $C_{17}H_{10}CIIN_2O_3$ crystallizes in the Orthorhombic space group*Pca2*with unit cell parameters *a*= 18.4757(12), *b*= 9.7925(5),*c*= 18.6956(13)Å and Z=8. The X-ray analysis reveals that the asymmetric unit

of the title compound comprises of two crystallographically independent molecules, A and B, with similar geometries. The crystal packing of the molecule is governed by four intermolecular hydrogen bonding [one C-H...Cl and three C-H...O]. Block-shaped crystal selected for intensity data collection was of dimensions 0.30 x 0.20 x 0.10 mm. Accurate cell parameters were determined from 2042 reflections with 3.7<0 <27.31°. X-ray intensity data of 8495 reflections (of which 5369 were unique) were collected on a computer controlled single crystal X-ray diffractometer with graphite monochromated Mo Ka radiation ($\lambda = 0.71073$ Å) in ω scan mode. The number of reflections after applying the limiting criterion $1 > 2\sigma(1)$ converged to 3531 which were considered as observed (- $21 \le h \le 22$, $-12 \le k \le 5$, $-23 \le l \le 21$). Data were corrected for Lorentz-polarization and multi-scan absorption corrections.¹ The crystal structure was solved by direct methods using SHELXS97 software.² Full-matrix least-squares refinement was carried out using SHELXL97 software.² All H atoms [except H4b] were positioned geometrically and were treated as riding on their parent C atoms, with C-H distances of 0.93–0.97 Å and with U_{iso} (H) = $1.2U_{eq}$ (C) or $1.5U_{eq}$ (methyl C). The final refinement cycles converged R = 0.0452 and $wR(F^2) = 0.0987$. Atomic scattering factors were taken from International Tables for X-ray Crystallography (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4). The crystallographic data are summarized in Table S1. A general view of the molecule indicating atomnumbering scheme (thermal ellipsoids are drawn at 40% probability level) is shown in Figure S1. ORTEP-3 for Windows software³ was used for making the thermal ellipsoids. The geometry of the molecule was calculated using PLATON⁴ and PARST⁵ software. The asymmetric unit of the title compound comprises of two crystallographically independent molecules, A and B, with similar geometries. Each molecule comprises three rings chlorophenyl, pyrrole and nitrophenyl ring. These rings are labeled as A1, A2, A3 and B1, B2, B3 with respect to molecule A and molecule B. All bond lengths and angles are normal and correspond to those observed in the literature.⁶

CCDC NO.	1471810
Crystal description	Block
Crystal colour	White
Crystal size	0.3 x 0.2 x 0.1 mm
Empirical formula	$C_{17}H_{10}CIIN_2O_3$
Formula weight	452.62
Radiation, Wavelength	Μο <i>Κα,</i> 0.71073 Å
Crystal system	Orthorhombic
Space group	Pca2
Hall symbol	P2c -2ac
No. of molecules per unit cell, Z	8
Unit cell dimensions	<i>a</i> = 18.4757(12), <i>b</i> = 9.7925(5), <i>c</i> = 18.6956(13)Å
Unit cell volume	3382.5(4)Å ³
D _x	1.273 g cm ⁻³
Temperature	293(2) К
Absorption coefficient	2.07 mm ⁻¹
F(000)	1760
θ range for collection of cell parameters	3.716 <θ<27.3100 º
Measurement	X'calibur system–Oxford diffraction make, U.K.
	[Oxford Diffraction, 2010]
Structure determination	Direct methods
Range of indices	<i>h</i> =-21 to 22, <i>k</i> = -12 to 5, <i>l</i> = -23 to 21
Reflections collected / unique	8495 / 5369
Reflections observed (I > 2σ (I))	3531
R _{int}	0.027
R _{sigma}	0.0518
Scan mode	ωscan
θ_{max}	26.00°

 Table S1: Crystal and experimental data for 5bc

θ _{min}	3.73°
T _{min} , T _{max}	0.39569, 1.0000
Absorption correction	multi-scan[CrysAlisRED; Oxford Diffraction, 2010]
Refinement	Full-matrix least squares on F ²
No. of parameters refined	441
Final <i>R</i>	0 .0452
w <i>R</i> (F ²)	0.0987
Weight	w=1/[$\sigma^2(F_o^2)$ +(0.0494P) ² +0.00P]
	where $P = [F_0^2 + 2F_c^2]/3$.
Goodness-of-fit	1.028
(Δ / σ) _{max}	0.001
Final residual electron density	-0.706< Δρ<0.073e Å ⁻³
Software for structure solution	SHELXS97 [Sheldrick, 2008]
Software for refinement	SHELXL97 [Sheldrick, 2008]
Software for molecular plotting	ORTEP-3[Farrugia,2012]; PLATON [Spek,2009]
Software for geometrical calculation	PLATON [Spek, 2009]; PARST [Nardelli, 1995]

Figure S1: ORTEP plot of the molecule (5bc) with 40% probability thermal ellipsoids. H atoms are shown as small spheres of arbitrary radii (CCDC NO. 1471810).

Crystal packing

The crystal packing of the molecule is governed by four intermolecular hydrogen bonding [one C-H...Cl and three C-H...O]. NO₂ group is playing a key role in the stabilization of the crystal structure, as its oxygen atoms [O2B and O3A] acts as the hydrogen bond acceptors in three hydrogen bond formation *via* hydrogen atoms H4B, H6B and H16A, respectively. The O2B acts as a common acceptor to the two hydrogen bond formations i.e. C4B-H4B...O2B and C16A-H16A...O2B, results in the formation of two centered bifurcated hydrogen bonds. There appears to be a distinct molecular layer for both the crystalographically independent molecules forming chain like structure within them. Chains as being formed by both the asymmetric molecules running parallel to each other and are further connected with each other along the *ab*-plane as shown in Figure 2. However, when viewed down the a-axis, the layers of both the crystalographically independent appears to bave adopted a branch type configuration to a structure disposition being guided along the *bc*-plane (Figure 3). Details of the hydrogen bond geometry are given below:

D—HA	D—H (Å)	HA (Å)	DA (Å)	D—HA (°)	
C4B-H4BO2B ⁱ	0.93	2.6	3.65(11)	158	
C6B-H6BO3A ⁱ	0.96	2.6	3.25(1)	127	
C16A-H16AO2B ⁱⁱ	0.93	2.4	3.35(1)	152	
C17A-H17ACL1A ⁱⁱ	ⁱ 0.93	2.7	3.36(8)	126	
Symmetry code: (i) x-1/2, -y+1, z (ii) -x+2, -y+1, z+1/2 (iii) x-1/2,-y+2, z					

Geometry of C-H...O and N-H...O hydrogen bonds

1,2-bis(4-chlorophenyl)-4-iodo-1*H*-pyrrole-3-carbaldehyde (7bh):

(CCDC NO. 1573099)

The title compound1,2-bis(4-chlorophenyl)-4-iodo-1H-pyrrole-3carbaldehyde (**7bh**), $C_{17}H_{10}N_1O_1Cl_2I_1$, crystallizes in the monoclinic space group P2₁/n with unit cell parameters: a = 6.6656(7), b = 19.4376(15), c = 13.2403(13) Å, $\beta = 104.291(11)^{\circ}$ and Z= 4. The crystal structure was solved

by direct methods using single-crystal X-ray diffraction data and refined to R = 0.0383 for 2288 observed reflections. The crystal packing is stabilized by Vander Waal's interactions.

Crystal Structure Determination and Refinement

X-ray intensity data of the crystal of the dimension 0.30 X 0.20 X 0.20 mm³ having well-defined morphology was collected on X'calibur CCD area-detector diffractometer equipped with graphite monochromated MoKa radiation ($\lambda = 0.71073$ Å). X-ray intensity data of 4614 reflections (of which 3039 unique) were collected at 293(2) K. The cell dimensions were determined by a leastsquares fit of angular settings of 1559 reflections in the θ range 3.66 to 29.07 °. The intensities were measured by oscan mode for tranges 3.99 to 27.01 °. 2288 reflections were treated as observed (I> $2\sigma(I)$). Data were corrected for Lorentz and polarization factors. The structure was solved by direct methods using SHELXS97.¹ All non-hydrogen atoms of the molecule were located in the best E-map. Full-matrix least-squares refinement was carried out using SHELXL97.² All the hydrogen atoms were geometrically fixed and allowed to ride on their parent carbon atoms with C-H= 0.93 Å with $U_{iso}(H) = 1.2U_{eq}(C)$. The final refinement cycles converged to an R = 0.0383 and wR (F^2) = 0.0853 for the observed data. Residual electron densities ranged from -0.613 to 0.560 eÅ⁻³. Atomic scattering factors were taken from International Tables for X-ray Crystallography (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4). The crystallographic data are summarized in Table S2. ORTEP-3 for Windows software³ was used for making the thermal ellipsoids. An ORTEP view of the title compound with atomic labeling is shown in Fig.S2. The geometry of the molecule was calculated using the PLATON⁴ and PARST⁵ software's. Bond lengths and bond angles are within expected values.⁶

Table S2: Crystal and experimental data for 7bh

Crystal description	Block shaped
Crystal colour	White
Crystal size	0.3 x 0.2 x 0.2 mm
Empirical formula	$C_{17}H_{10}N_1O_1Cl_2I_1\\$
Formula weight	442.06
Radiation, Wavelength Mo $K\alpha$,	0.71073 Å
Unit cell dimensions	a= 6.6656(7), b= 19.4376(15), c= 13.2403(13) Å,
	β= 104.291(11)°
Crystal system	Monoclinic
Space group	P2 ₁ /n
Unit cell volume	1662.4(3)
No. of molecules per unit cell, Z	4
Temperature	293(2)
Absorption coefficient	2.247 mm ⁻¹
F(000)	856
Scan mode	ωscan
$\boldsymbol{\theta}$ range for entire data collection	3.79<θ<25.99 ⁰
Range of indices	h= -8 to 7, k= -23 to 23, l= -11 to 16
Reflections collected / unique	4614 / 3039
Reflections observed (I > 2σ (I))	2288
R _{int}	0.0507
R _{sigma}	0.0285
Structure determination	Direct methods
Refinement	Full-matrix least-squares on F ²
No. of parameters refined	200
Final R	0.0383
wR(F ²)	0.0853
Weight	1/[σ²(F₀²)+(0.0367 P)²+0.3021P]
	where $P = [F_o^2 + 2F_c^2] / 3$
Goodness-of-fit	1.069
(Δ / σ) _{max}	0.002 (for U33 C7)
Final residual electron density	-0.613 < Δρ< 0.560 eÅ ⁻³

Measurement Software for structure solution: Software for refinement: Software for molecular plotting: Software for geometrical calculation X'calibur system – Oxford diffraction make, U.K. SHELXS97 (Sheldrick, 2008) SHELXL97 (Sheldrick, 2015) ORTEP-3 (Farrugia, 2012) PLATON (Spek, 2009) PLATON (Spek, 2009)

Figure S2: *ORTEP* view of the molecule (**7bh**), showing the atom-labelling scheme. Displacement ellipsoids are drawn at the 50% probability level and H atoms are shown as small spheres of arbitrary radii (**CCDC NO. 1573099**).

References

- (a) G. M. Sheldrick, *Acta Cryst.*, 2008, *A64*, 112-122. (b) Oxford Diffraction, 2010, CrysAlis PRO, Oxford Diffraction, Yarnton, UK.
- 2. G. M. Sheldrick, Acta Cryst., 2015, C71, 3-8.
- 3. L. J. Farrugia, J. Appl. Cryst., 2012, 45, 849-854.
- 4. A. L. Spek, Acta Cryst., 2009, D65, 148-155.
- 5. M. Nardelli, J Appl. Cryst., 1995, 28, 659.
- F. H. Allen, O. Kennard, D. G. Watson, L. Brammer, A. G. Orpen, and R. Taylor, *J. Chem. Soc., Perkin Trans-II*, 1987, S1.

DFT computational details and results

Computational level: Functional: PBE0

Basis set: Triple zeta single polarized (TZP) basis set

A correction for dispersion interaction is applied using Grimme D3 with Becke-Johnson damping. All calculations including geometry optimization and frequencies were performed in the gas phase using the Amsterdam Density Functional (ADF) package.

Optimized structures for the HOMO-LUMO of dihydropyrrole intermediate (E) and pyrrole (6)

DFT-chart for HOMO-LUMO calculations

Figure S3. HOMO-LUMO energy levels diagram Level diagrams for dihydropyrrole enamine-intermediate **E** and pyrrole **6**. The HOMO of enamine-intermediate **E** is 0.89eV higher than the HOMO of pyrrole **6**. Therefore, intermediate **E** is more reactive towards the electrophile at C4-position, as compare to pyrrole **6** at the C5-position.

Coordinates for Intermediate (E)

Atom	Х	Y	Z (Angstrom)
Ν	-0.913075	0.213949	-0.384077
С	-0.725509	-1.154529	-0.580190
С	0.502902	-1.465140	-0.986084
С	1.337329	-0.216774	-1.044702
С	0.264671	0.927176	-0.902524
Н	1.361886	3.256599	-1.717877
Н	-1.544197	-1.830848	-0.403253
Н	0.883346	-2.459783	-1.138818
Н	2.185899	5.128140	-0.336066
Н	1.861992	-0.084698	-1.997301
С	2.439469	-0.170301	0.010583
Н	-5.259884	0.168904	0.985650
0	2.686273	-1.063177	0.777078
Н	-3.074744	-0.882707	0.720420
С	-2.157919	0.823716	-0.231025
С	-2.374440	2.147301	-0.634769
С	-3.617005	2.741895	-0.458943
С	-4.669707	2.042218	0.114709
С	-4.457473	0.729218	0.523576
С	-3.220974	0.126014	0.362796
н	-1.576087	2.721293	-1.076842
Н	-3.757364	3.766416	-0.778367

Н	-5.634712	2.510784	0.249203
С	0.701633	2.089846	-0.043313
С	0.582050	2.043718	1.343017
С	1.039548	3.098773	2.121104
С	1.622044	4.210375	1.522299
С	1.740764	4.263660	0.138494
С	1.278460	3.208418	-0.638533
Н	0.123658	1.181602	1.807257
Н	0.938130	3.054487	3.197290
Н	1.975398	5.032759	2.129946
Н	3.045997	0.755306	0.003935
Н	0.036619	1.300169	-1.907105

Coordinates for compound (6)

Ν	-0.884021	0.113614	-0.384454
С	-0.817602	-1.169928	-0.911834
С	0.491525	-1.499963	-1.086112
С	1.271902	-0.387408	-0.648994
С	0.391658	0.605904	-0.219684
Н	1.999912	2.515862	-1.213209
Н	-1.723116	-1.702631	-1.141015
Н	0.883380	-2.424393	-1.472612
Н	2.508706	4.717825	-0.225798
Н	3.140200	0.560369	-0.082880
С	2.722326	-0.343080	-0.564955
Н	-5.073258	0.395231	1.412922
0	3.461536	-1.219527	-0.969853
Н	-2.922433	-0.772992	1.034908
С	-2.105813	0.805916	-0.159285
С	-2.310023	2.068530	-0.706999
С	-3.507251	2.730624	-0.476099
С	-4.506137	2.132689	0.284737
С	-4.301830	0.865886	0.818400
С	-3.100285	0.202590	0.603755
Н	-1.533560	2.523990	-1.303633
Н	-3.662546	3.713818	-0.899362

Н	-5.439262	2.651017	0.459026
С	0.671608	1.921527	0.363415
С	0.074607	2.317732	1.565277
С	0.349055	3.561486	2.112581
С	1.226547	4.430648	1.472422
С	1.827969	4.046359	0.280338
С	1.550360	2.803197	-0.272854
Н	-0.601738	1.643805	2.071940
Н	-0.118750	3.851727	3.043946
Н	1.439527	5.400917	1.900904