



Article

Subscriber access provided by UB + Fachbibliothek Chemie | (FU-Bibliothekssystem)

# Origins of regioselectivity in the Fischer indole synthesis of a selective androgen receptor modulator

Elizabeth L. Noey, Zhongyue Yang, Yanwei Li, Hannah Yu, Rachel N Richey, Jeremy M Merritt, Douglas P Kjell, and Kendall N. Houk

J. Org. Chem., Just Accepted Manuscript • Publication Date (Web): 03 May 2017 Downloaded from http://pubs.acs.org on May 3, 2017

#### Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



The Journal of Organic Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

## Origins of regioselectivity in the Fischer indole synthesis of a selective androgen receptor modulator

Elizabeth L. Noey<sup>†</sup>, Zhongyue Yang<sup>†</sup>, Yanwei Li<sup>†</sup>, Hannah Yu<sup>‡</sup>, Rachel N. Richey<sup>‡</sup>, Jeremy M. Merritt<sup>‡</sup>, Douglas P. Kjell<sup>‡</sup>, K. N. Houk<sup>†</sup>\*

<sup>†</sup>Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States.

<sup>\*</sup>Eli Lilly and Co. Lilly Corporate Center, Indianapolis, Indiana 46285, United States.

E-mail: houk@chem.ucla.edu



#### Abstract

The selective androgen receptor modulator, (*S*)-(7-cyano-4-(pyridin-2-ylmethyl)-1,2,3,4-tetrahydrocyclopenta[*b*]indol-2-yl)-carbamic acid isopropyl ester, LY2452473, is a promising treatment of side effects of prostate cancer therapies. An acid-catalyzed Fischer indolization is a central step in its synthesis. The reaction leads to only one of the two possible indole regioisomers, along with minor decomposition products. Computations show that the formation of the observed indole is most favored energetically, while the potential pathway to the minor isomer leads instead to decomposition products. The disfavored [3,3]-sigmatropic rearrangement, that would produce the unobserved indole product, is destabilized by the electron-withdrawing phthalimide substituent. The most favored [3,3]-sigmatropic rearrangement transition state is bimodal, leading to two reaction intermediates from one transition state, which is confirmed by molecular dynamics simulations. Both intermediates can lead to the observed indole product, albeit through different mechanisms.

#### Introduction

Indole derivatives have extensive biological activity, and are important in pharmaceuticals.<sup>1,2</sup> The venerable Fischer indole synthesis<sup>3</sup> continues to be a widely applicable approaches to the synthesis of indoles.

Indole **1**,<sup>4</sup> named LY2452473 (**Figure 1**), is a selective androgen receptor modulator (SARM), and has been brought to phase 1 clinical trials for the treatment of secondary hypogonadism induced by androgen deprivation therapy (ADT). ADT reduces the levels of androgens, or male hormones, which stimulate prostate cancer cells. However, ADT has a variety of side effects,<sup>5</sup> which indole **1** shows promise in combating.



**Figure 1**. Indole **1**, LY2452473, a SARM is a clinical candidate to treat symptoms of androgen deprivation therapy.<sup>4</sup>

Çelebi-Ölçüm et al.<sup>6</sup> showed that in the acid catalyzed Fischer indole synthesis, either the arene or enamine nitrogen, N $\alpha$  or N $\beta$ , respectively, can be protonated to accelerate the [3,3]-sigmatropic rearrangement (**Figure 2**). Protonation of N $\beta$  is favored

#### The Journal of Organic Chemistry

and leads to the lowest energy rearrangement. Protonation of N $\alpha$  is disfavored by 1.5 kcal/mol, and the thermal reaction of this species is disfavored by 12.6 kcal/mol.



Figure 2. The [3,3]-sigmatropic rearrangement in the Fischer indolization.

In the current study, as shown in **Figure 3**, the hydrazine **2** reacts with the cyclopentanone **3** under acidic conditions to form only one indole regioisomeric product, **4**. A variety of reaction conditions were screened, resulting in differing yields and side products **5** and **7**, among other uncharacterized side products. The isolation of hydrazone **5** indicates the reaction does not go to completion under certain conditions, and the isolation of phthalimide **7**, suggests the decomposition reaction shown. No cyclopentenone is observed, indicating that the **7** is not the result of the decomposition of starting material **3**. In all cases, the isomeric indole, **8**, is not formed (**Figure 4**).



Figure 3. Experimental Fischer indolization, and observed side products.



Figure 4. Unobserved indole regioisomer.

#### **Results & Discussion**

Figure 5 shows the reaction pathways explored computationally. The hydrazone intermediate, 5, can form two indole regioisomers, 4 and 8, as well as decomposition products, 6 and 7. Free energies of each intermediate and several relevant transition states

are shown. The hydrazone intermediate, **5**, is preferentially protonated at N $\beta$ . This will be in equilibrium with the hydrazone protonated at N $\alpha$ , or at the phthalimide O. Either of the two carbons  $\beta$  to the hydrazone N on the cyclohexane ring (i.e. C2 or C5) can be deprotonated to form two regioisomeric enamine intermediates, which can also be protonated at N $\alpha$ , N $\beta$ , or O and are in equilibrium with the hydrazone. For the formation of the observed indole product, **4**, the deprotonation occurs at C5 and forms enamine **9**. Enamine **9**-N $\alpha$  undergoes the [3,3]-sigmatropic rearrangement via **TS2**-N $\alpha$ \*, which has an energy of 33.4 kcal/mol relative to hydrazone **5**-N $\beta$ . This forms **10**-N $\alpha$  or **13**-N $\alpha$ , which eventually lead to the observed indole product, **4**.

The enamine formed from deprotonation at C2, **11**, decomposes by a retro-Michael reaction through **TS4-O** to form **6** and **7**. Decomposition of **11** is favored over the [3,3]-sigmatropic rearrangement via **TS3-N** $\beta$  by 1.8 kcal/mol. This explains why the indole regioisomer, **8**, is not observed. Decomposition of enamine **11** is disfavored by 1.8 kcal/mol compared to the formation of the observed indole **4**, by the amount discussed above.



Figure 5. Competing pathways from hydrazone, 5-N $\beta$ , to the two possible regioisomeric indoles, 4 and 8, and decomposition products, 6 and 7.  $\Delta G$  is given in kcal/mol

**Figure 6** shows the most favored transition structure for [3,3]-sigmatropic rearrangement, **TS2-N** $\alpha$ \*, which is on a bifurcating potential surface, and leads to two products. Because the intermediates share a common transition structure, dynamics control the product distribution.<sup>7</sup> This bifurcation happens when N $\alpha$  is protonated, and the phthalimide is oriented with one O near the reacting moieties. This is because the arene carbon and the phthalimide oxygen are competing nucleophiles toward the enamine motif. The competing nucleophiles stabilize the partially positively-charged  $\beta$ -carbon. On the bifurcating surface, the first transition state, **TS2-N\alpha\***, involves primarily breaking of the N-N bond, while the second, a transition state for the interconversion of the products, either forms the C-C or C-O bond. The formation of the C-C bond gives the

Page 7 of 17

 rearrangement product, **10-N** $\alpha$ , and the formation of the C-O bond gives a [3.2.1] bridged heterocycle, **13-N** $\alpha$ .



Figure 6. The bifurcating energy surface from 9-N $\alpha$ .  $\Delta G$  is given in kcal/mol. Bond lengths are given in Ångstroms.

Reaction dynamics simulations were performed to investigate the bifurcation of the [3,3]-sigmatropic rearrangement of enamine 9-N $\alpha$ . Normal mode sampling was conducted on the TS2-N $\alpha$ \* to obtain an ensemble of TS geometries, as shown in Figure 7. The formations of bond 2 and 4 lead to 10-N $\alpha$  and 13-N $\alpha$ , respectively. The distributions of bond lengths for these two bonds are identical, which suggests that 10N $\alpha$  and 13-N $\alpha$  should have equal chance of formation. The formation of bond 3, on the other hand, also results in 10-N $\alpha$ , through a chair-cope like pathway. Bond 3 is 0.4 Å longer than bond 2 or 4 in the transition structure, leaving the question of whether 10-N $\alpha$  is also produced by the formation of bond 3. This bond has a distribution of 3.50 ± 0.38 Å, which partially overlaps with that of bond 2.



Figure 7. The structure of TS2-N $\alpha$ \* and the overlay of TS geometries obtained from normal model sampling.

One hundred and eighty trajectories were initiated from the TS geometries. Trajectories were propagated forward and backward until the formation of the product **10-N** $\alpha$  (bond 1 > 3.5 Å and bond 2 < 1.5 Å, or bond 3 < 1.5 Å), or **13-N** $\alpha$  (bond 1 > 3.5 Å and bond 4 < 1.5 Å), or the formation of the reactant **9-N** $\alpha$  (bond 1 < 1.5 Å, bond2 > 3.5 Å, and bond 4 > 3.5 Å). Trajectories were terminated after 500 fs if neither criteria was met. Both **10-N** $\alpha$  and **13-N** $\alpha$  were observed in the simulation. One hundred and sixty trajectories were productive, and twenty were recrossing. Among these reactive trajectories, seventy nine (49.3%) give **10-N** $\alpha$  by forming bond 2 and one (0.7%) by bond 3, seventy eight (48.7%) form **13-N** $\alpha$ , and two (1.3%) bridge **10-N** $\alpha$  and **13-N** $\alpha$ . The trajectories bridging 10-N $\alpha$  and 13-N $\alpha$  were observed because the sampled transition state geometries slightly overlap the transition state TS7-N $\alpha$ . Figure 8 shows the distribution of ten representative trajectories leading to 10-N $\alpha$  and 13-N $\alpha$ , and typical snapshots for the formation of these two products. Labeled in red represents the transition zone, which was previously defined as the region that embraces 98% of the transition state geometries. As the elongation of bond 1 (N-N bond), trajectories pass the transition zone and start to bifurcate when bond 1 is about 3.0 Å.



Figure 8. Distribution of ten representative trajectories and snapshots for the formation of  $10-N\alpha$ , and  $13-N\alpha$ . The black dots represent the sampled transition state geometries, and the red circle represents the transition zone.

**Figure 9** shows how the products of the bifurcation lead to the observed product through different routes. The separate molecules of 13-N $\alpha$ , likely do not completely react with each other. From intermediate **13-N\alpha** there is a low barrier (23.7 kcal/mol, **TS7-N\alpha**) for the attack of the cyano aniline to form **14-N\alpha**. The transition structure for the competing deprotonation of **13-N\alpha** to form the enimine (not shown) is 32.7 kcal/mol, **8**.9 kcal/mol higher in energy than **TS7-N\alpha**. Intermediate **14-N\alpha** can undergo proton transfer, followed by intramolecular attack of the arene on the imine to form **15**. This is followed by sequential deprotonation and loss of NH<sub>4</sub> to form the product.



Figure 9. Competing pathways from 9-N $\alpha$ , including the bifurcation and competing paths to the observed indole product, 4.  $\Delta G$  is given in kcal/mol.

The [3,3]-sigmatropic rearrangement transition state has many degrees of freedom: N $\alpha$ , or N $\beta$  may be positively charged, it can adopt a chair or boat conformation, and the arene can form a bond on the same or opposite side of the phthalimide. When the

transition state is in a boat conformation the arene is over the cyclohexane ring, and it is not sterically possible for the phthalimide to be on the same side in this conformation. This means there were six conformations to explore. Additionally, the cyclohexane ring may adopt one of two envelope conformations, and only the lowest energy conformation is reported in here. These transition states are reported in **Table 1**. There are two transition structures for the [3,3]-sigmatropic rearrangement for the observed regioisomer that are comparable in energy. The second one, only 0.4 kcal/mol higher in energy, is not bimodal and only leads to **10-N** $\alpha$ . The energy of this transition structure is shown in **Table 1** highlighted in bold text. It is in a chair conformation with N $\alpha$  protonated, and the phthalimide on the same side as the arene.

**Table 1**.  $\Delta G$  for the [3,3]-sigmatropic rearrangement transition states. "Chair" and "boat" indicate the conformation of the atoms involved in the [3,3]-sigmatropic rearrangement, and N $\alpha$  and N $\beta$  refer to where the positive charge resides. The energy of the lowest energy transition structures for the observed and unobserved regioisomers, **TS2-N\alpha** and **TS3-N\alpha**, respectively, are shown in bold-face and also reported in **Figure 5**.

	Observed Regioisomer				Unobserved Regioisomer			
relationship	chair		boat		chair		boat	
to phthalamide	Να	Νβ	Να	Nβ	Να	Νβ	Να	Nβ
same side	34.4	38.5	n.a.	n.a.	41.4	43.3	n.a.	n.a.
opposite side	36.2*	38.5	34.4	39.0	39.8	40.9	38.6	42.1

\*bimodal

The phthalimide group is electron-withdrawing, which has a greater destabilizing effect the closer it is to the positively-charged reactive moiety. In the [3,3]-sigmatropic

rearrangement for the formation of the observed regioisomer (**TS2**) the phthalimide is two carbons from the atoms involved in the rearrangement. In the formation of the unobserved regioisomer (**TS3**) it is only one carbon away. Therefore the phthalimide destabilizes **TS3** more so than **TS2**, as indicated qualitatively in **Figure 10**.



**Figure 10**. Transition states for the [3,3]-sigmatropic rearrangement. The phthalimide is further from the positive charge in the favored transition state (**TS2**) than it is in disfavored transition state (**TS3**).

The analogous [3,3]-sigmatropic rearrangement transition states en route to **8** are not bimodal. This is because the attack of the phthalimide oxygen on the enamine would form a strained 5-membered heterocycle. In the [3,3] transition state, the oxygen of the phthalimide is far from the  $\beta$ -C.

#### Conclusion

Our calculations explain the formation of only one regioisomeric indole product in the Fischer indolization used for the synthesis of **1**, and accounts for the presence of Page 13 of 17

#### The Journal of Organic Chemistry

decomposition product **6**. In the pathway to the observed indole product, **4**, several [3,3]sigmatropic rearrangement transition states are possible. The lowest energy path (**TS2-** $N\alpha^*$ ) is favored by 1.8 kcal/mol compared to the competing decomposition pathway through **TS4-O**, and 3.6 kcal/mol compared to the formation of the unobserved indole, **8** via **TS3-N\alpha**. The [3,3]-sigmatropic rearrangements from enamine **9**, **TS2-N** $\alpha$ , experiences less destabilization than the disfavored **TS3-N** $\alpha$  because the atoms involved in the rearrangement are further from the phthalimide. The favored [3,3]-sigmatropic rearrangement, **TS2-N** $\alpha^*$ , is bimodal, leading to two products from one transition state. Of the molecular dynamics simulations from **TS2-N** $\alpha^*$  50.0% form **10-N** $\alpha$ , 48.7% form **13-N** $\alpha$ , and 1.3% bridge **10-N** $\alpha$  and **13-N** $\alpha$ . The two intermediates lead to the observed indole product by different paths. This work elucidates the driving force for the observed regioselectivity of this significant reaction. These calculations enable design of novel regioselective Fischer indolization reactions.

#### Experimental

All calculations were performed with Gaussian09.<sup>8</sup> Structures were optimized with M06-2X/6-31G(d) in the gas phase. Truhlar's quasiharmonic approximation was used to correct for errors in the estimation of vibrational entropies by setting all positive frequencies below 100 cm<sup>-1</sup> to 100 cm<sup>-1</sup>.<sup>9</sup> Single point energy calculations used M06- $2X/6-311+G(2d,p)^{10}$  with the SMD solvation model<sup>11</sup> for acetic acid. Relative Gibb's Free Energies are reported in kcal/mol.

Direct molecular dynamics (MD) simulations were performed in the gas phase. Because of the demanding quantum calculations required in the dynamics simulations, these were performed at the M06-2X/6-31G(d) level of theory. Quasiclassical trajectories (QCTs) were initialized in the region of the potential energy surface near the transition state.<sup>12</sup> Normal mode sampling involves adding zero-point energy for each real normal mode in the TS, and then Boltzmann sampling to afford a set of geometries reflecting the thermal energy available at 300 K with a random phase.<sup>13</sup> No additional velocities added other than vibrations along the mode perpendicular to the reaction coordinate. The trajectories were propagated forward and backward, until the formation of product, or the reactant. The classical equations of motion were integrated with a velocity-Verlet algorithm using Singleton's program Progdyn,<sup>14</sup> with the energies and derivatives computed on the fly by the M06-2X method using Gaussian 09. The step length for integration was 1 fs.

#### **Supporting Information**

High throughput liquid chromatography trace from the Fischer indolization reaction, absolute energies, imaginary frequencies, and Cartesian coordinates for all structures, energies, and bond lengths of transition states and intermediates involved in the bifurcating surfaces.

#### Acknowledgement

We are grateful to the National Science Foundation (CHE-1361104) for financial support of this research. Calculations were performed on the Hoffman2 cluster at UCLA and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported

 by the NSF (OCI-1053575). We thank Natalie C. James and Hao Wang for preliminary computational results.

#### References

- Indole Ring Synthesis From Natural Products to Drug Discovery," G.W. Gribble, Wiley, 2016.
- 2 Humprey, G. R.; Kuethe, J. T. Chem. Rev. 2006, 106, 2875–2911.
- 3 (a) Fischer, E.; Jourdan, F. Ber. Dtsch. Chem. Ges. 1883, 16, 2241–2245. (b) Fischer,
  E.; Hess, O. Ber. Dtsch. Chem. Ges. 1884, 17, 559–568.
- 4 (a) Fales, K. R.; Green, J. E.; Jadhav, P. K.; Matthews, D. P.; Neel, D. A.; Smith, E. C. R.; Tetrahydrocarbazole derivatives useful as androgen receptor modulators, WO2007002181 A2, Jan 4, 2007. (b) Jadhav, P. K.; Krishnan, V.; Kim, E. J., Tetrahydrocyclopenta[b]indole androgen receptor modulators. WO2009140448 A1, Nov 19, 2009. (c) Benson, C. T., Treatment of Androgen Deprivation Therapy Associated Symptoms. WO 2016040234 Al, Sept 8, 2015.
- 5 (a) Nguyen, P. L.; Alibhai, S. M.; Basaria ,S.; D'Amico, A.V.; Kantoff, P.W.; Keating, N. L.; Penson D. F.; Rosario D. J.; Tombal B.; Smith M. R. *Eur. Urol.* 2015, *67*, 825-836. (b) Krupski, T. L.; Smith M. R.; Lee, W. C.; Pashos, C. L.; Brandman, J.; Wang, Q.; Botteman, M.; Litwin, M. S. *Cancer*. 2004, *101*, 541-549. (c) Wilke, D. R.; Parker, C.; Andonowski, A.; Tsuji, D.; Catton, C.; Gospodarowicz, M.; Warde, P. *BJU Int*. 2006, *97*, 963-968.

### ACS Paragon Plus Environment

- 6 Çelebi-Ölçüm, N.; Boal, B. W.; Huters, A. D.; Garg, N. K.; Houk, K. N. J. Am. Chem. Soc. 2011 133, 5752-5755.
- 7 (a) Ess, D. H.; Wheeler, S. E.; Iafe, R. G.; Xu, L.; Çelebi- Ölçüm, N.; Houk, K. N. *Angew. Chem., Int. Ed.* 2008, 47, 7592-7601. (b) Thomas, J. B.; Waas, J. R.; Harmata, M, Singleton, D. *J. Am. Chem. Soc.* 2008, *130*, 14 544-14 555. (c) Martin-Somer, A.; Yanez, M.; Hase, W. L.; Gaigeot M-P.; Spezia R. *J. Chem. Theory Comput.* 2016, *12*, 974–982.
- 8 Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonneerg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, M. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.
- 9 Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. J. Phys. Chem. B, 2011, 115, 14556-14562.

10 (a) Zhao, Y.; Truhlar, D. G. Theor. Chem. Acc. 2008, 120, 215-241. (b) Zhao, Y.;
Truhlar, D. G. Acc. Chem. Res. 2008, 41, 157-167.
11 Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. J. Phys. Chem. B, 2009, 113, 6378-
6396.
12 (a) Yang, Z.; Doubleday, C.; Houk, K. N. J. Chem. Theory Comput. 2015, 11,
5606–5612. (b) Yang, Z.; Yu, P.; Houk, K. N. J. Amer. Chem. Soc. 2016, 12, 4237-
4242.
13 (a) Doubleday, C.; Bolton, K.; Hase, W. L. J. Phys. Chem. A 1998, 102, 3648-3658.
(b) Black, K.; Liu, P.; Xu, L.; Doubleday, C.; Houk, K. N. Proc. Natl. Acad. Sci. U. S.

A. 2012, 109, 12860–12865. (c) Bunker, D. L.; Hase, W. L. J. Chem. Phys. 1973, 59, 4621–4632.

14 Oyola, Y.; Singleton, D. A. J. Am. Chem. Soc. 2009, 131, 3130-3131.