Identification, synthesis and structure assignment of two impurities of Erlotinib, a drug used for the treatment of some types of cancer.

María Guadalupe Morales-Vilchis[1], Patricia Flores-Sánchez[1], and Jaime Escalante[2,*]

[1]Signa S.A de C.V., Jiutepec, Morelos 62500, México
[2]Centro de Investigaciones Químicas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos 62209.

*Corresponding author: Jaime Escalante, e-mail: jaime@uaem.mx

Abstract. In the synthesis and industrial production of Erlotinib, the formation of organic impurities has been detected, so the identification and control of these is a critical point within the process. Standards for the consumption of pharmaceutical products are well established for active substances. In this sense, the synthesis and unequivocal identification of two impurities of Erlotinib as structural isomers was developed in the present study. The structural characterization of both isomers was confirmed by ^{1}H and ^{13}C-NMR as well as by their X-ray diffraction study.
**Keywords:** Erlotinib; Cancer; Isomers; Impurity.

Resumen. Durante la síntesis y producción industrial del Erlotinib, se ha detectado la formación de impurezas orgánicas. La identificación y control de éstas es un punto crítico dentro del proceso. Los estándares de subtancias activas para el consumo de productos farmacéuticos están bien establecidos. En este sentido, en el presente estudio se desarrolló la síntesis e identificación inequívoca de dos impurezas del Erlotinib como isómeros estructurales. La caracterización de ambos isómeros se confirmó por ^{1}H- y ^{13}C-RMN, así como por su estudio de difracción de rayos-X.
Palabras clave: Erlotinib; Cáncer; Isómeros; Impurezas.

Introduction

Erlotinib is a 4-anilinequinazoline drug used in the treatment of non-small cell lung cancer, pancreatic cancer and other types of cancer. [1,2] This drug was approved by Food and Drug Administration in 2004 for the treatment to cancer. Its mechanism of action is inhibition of epidermal growth factor receptor (EGFR), phosphorylation and blocking of tumor cell signal transduction by competing with the substrate and thus inhibiting growth of the tumor cells and inducing their apoptosis.[3-5]

One of the methods for the synthesis of Erlotinib **3** includes the coupling of 4-chloro-6,7-bis-(2-methoxyethoxy) quinazoline **1** with 3-ethynylaniline **2**, under reflux conditions (Scheme 1).
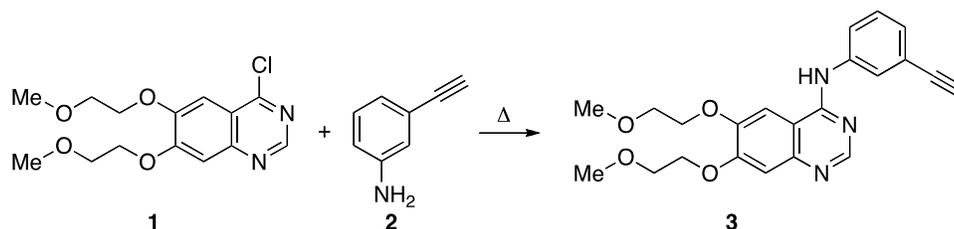
Scheme 1. Synthesis of Erlotinib **3**.

In the heating process and the acid conditions of the reaction produces various impurities, some of which are shown in Fig. 1. [6] International Conference on Harmonization (ICH) [7] indicates than all impurities than exceed levels of 0.1% must be identified and quantified using validated analytical procedures because they can have adverse toxicological effects. Therefore, the quality and safety of the drug substances is an important regulatory requirement. [8]

During the synthesis of Erlotinib at industrial level, impurity **4** was identified and controlled at percentages not greater than 0.1% by HPLC; during the same process, another unknown impurity was detected in percentages of almost 0.1%, which was isolated by column chromatography and analyzed by mass spectrometry, observing that it had the same molecular weight as **4**, so it was proposed that both molecules could be isomers. In this sense, the goal of this work was to synthesize and characterize both impurities and determining the correct structure for each of them in order to have pure substances as references in the chromatographic analysis of the industrial process.
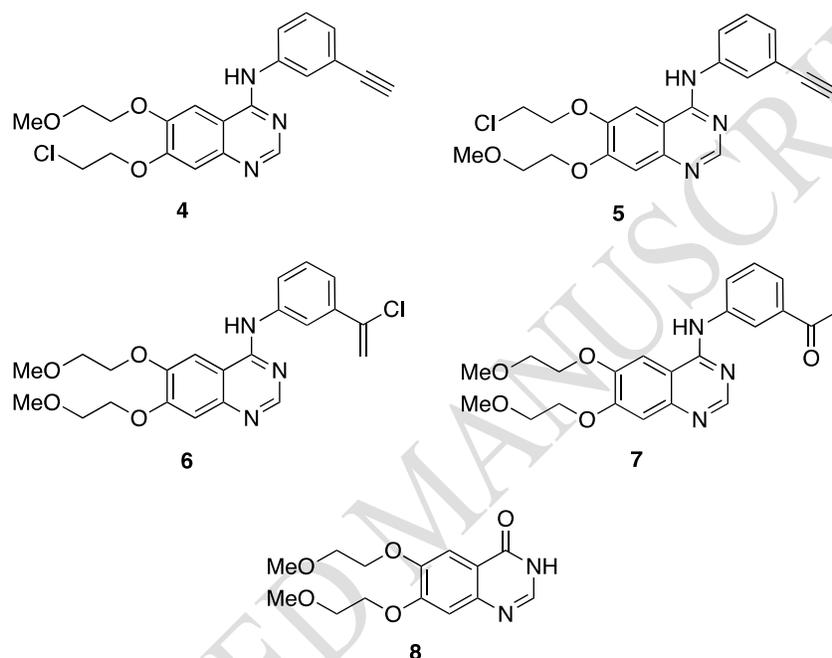

Fig. 1. Some Erlotinib impurities identified in the synthesis process.

Indeed, we have synthetized, separated and identified two impurities of Erlotinib: 7-(2-chloroethoxy)-*N*-(3-ethynylphenyl)-6-(2-methoxyethoxy)quinazoline-4-amine **4** and its structural isomer 6-(2-chloroethoxy)-*N*-(3-ethynylphenyl)-7-(2methoxyethoxy)quinazoline-4-amine **5**. Both impurities are derivative of Erlotinib in which the chlorination of methyl ether could take place in the aliphatic chain below to give rise to **4**, or in the upper part to yield **5**. It is noteworthy that both compounds could be detected by thin layer chromatography (TLC) and isolated by column chromatography (*vide infra*). The structure for each isomer was confirmed by ^{1}H and ^{13}C-NMR, X-ray, and MS.

Experimental

General

All chemicals were purchased from Aldrich and used as received. Melting points were determinated on a Buchi B-540 melting point apparatus and were not corrected. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Oxford, at 400 and 100 MHz respectively in CDCl$_3$ or CD$_3$OD. Mass spectra were recorded on a JEOL JMS-AX50SHA spectrometer. The proportions of each product in the reaction mixture were determine by gas cromatography-mass spectrometry on an Agilent GC system 6890 series coupled to a Agilent mass selective detector 5973N with an HP-5MS column length of 30 m, internal

size of 0.25 mm, film thickness of 0.25 mm, split radio of 5:1, 40°C, min^{-1}, 10°C cm^{-1} to 250°C cm^{-1}. Thin layer chromatography (TLC) was performed on Merck TLC aluminum sheets silica 60 F254. Visualization was made by UV ligth at 254 and 366 nm. Column chomatography was performed on silica gel 60 (0.040-0.63 mm). Single crystal X-ray diffraction studies were performed on a Bruker-APEX diffractometer with a CCD area detector (lMoKa = 0.71073 Å, monochromator: graphite).

2-(4-Chloro(bromine)-6(7)-(2-methoxyethoxy)quinazolin-7(6)-yloxy)ethanol (9). To a solution of 4-chloro-6,7-bis-(2-methoxy-ethoxy)-quinazoline **1** (3.0 g, 9.5 mmol) in CH$_2$Cl$_2$ (45 mL) was added dropwise a solution of BBr$_3$ 1M in CH$_2$Cl$_2$ (5.7 mL, 3.8 mmol) during 3 h at -55 °C. The temperature was maintained at -55 °C during 1 h after the addition of BBr$_3$. The reaction mixture was allowed to warm to room temperature and stirred for 30 h. The reaction was cooled to 10-15°C and adjusted to pH 7-8 with saturated aqueous NaHCO$_3$ solution. The product was extracted with CH$_2$Cl$_2$ (3 x 10 mL), the organic layer was dried with anhydrous Na$_2$SO$_4$ and evaporated to give a brown crude product which was purified by column chromatography (inner diameter 35 mm and length 60 cm) using EtOAc/MeOH 9:1, as eluent. Rf: 0.27 (AcOEt/MeOH, 98:2). Yield 28%; white solid; We only reported the major signals in NMR. m.p. 118-120 °C; ^{1}H NMR (CDCl$_3$, 400 MHz): δ 3.45 (s, 3H), 3.86-3.84 (m, 2H), 4.05-4.02 (m, 2H), 4.26-4.23 (m, 2H), 4.31- 4.28 (m, 2H), 7.28 (s, 1H), 7.36 (s, 1H), 8.82 (s, 1H); ^{13}C NMR (CDCl$_3$, 100 MHz): δ 59.4, 61.0, 68.9, 70.4, 71.5, 104.9, 108.2, 119.7, 149.0, 152.7, 153.7, 156.3, 159.3.

2-(4-(3-Ethynylphenylamine)-6(7)-(2-methoxyethoxy) quinazolin-7(6)-yloxy) ethanol (10). To a solution of **9** (1.4 g, 4.6 mmol) in a mixture of IPA (30 mL), ethyl acetate (30 mL) and CH$_2$Cl$_2$ (15 mL) was added 3-ethynylaniline (0.660 g, 5.6 mmol) at 25 ºC and the reaction mixture was heated under reflux for 18 h. After completing the reaction, the mixture was cooled and adjusted to pH ~7 with saturated aqueous NaHCO$_3$ solution. The product was extracted with ethyl acetate (3 x 10 mL), the organic layer was dried with anhydrous Na$_2$SO$_4$ and evaporated to give a light brown crude product. The crude product was purified by column chromatography (inner diameter 25 mm and length 50 cm) using EtOAc/MeOH 9:1, as eluent. Rf: 0.28 (AcOEt/MeOH, 95:5). Yield 91%; light orange solid; We only reported the major signals in NMR m.p. 157-160°C; MS/FAB$^+$: *m/z* 380 (M + H$^+$) HRMS (FAB$^+$): *m/z* 380.1627 [M + H$^+$] (Calculated for C$_{21}$H$_{22}$N$_3$O$_4$ H$^+$ 380.1610); ^{1}H NMR (CDCl$_3$, 400 MHz): δ 3.01 (s, 1H), 3.33 (s, 1H), 3.74-3.67 (m, 2H), 3.91-3.84 (m, 2H), 4.13-4.03 (m, 4H), 7.00 (s, 1H), 7.27-7.15 (m, 2H), 7.67-7.65 (m, 2H), 7.77 (s, 1H), 8.51 (s, 1H); ^{13}C NMR (CDCl$_3$, 100 MHz): □ 59.3, 60.8, 68.2, 69.5, 70.8, 70.7, 71.1, 77.7, 83.5, 103.5, 109.1, 109.3, 122.5, 122.9, 125.2, 127.9, 129.1, 138.9, 147.6, 148.7, 153.8, 154.5, 156.5.

2-(4-(3-Ethynylphenylamino)-6(7)-(2-methoxyethoxy) quinazolin-7(6)-yloxy)ethyl methanesulfonate (11). To a cool suspension of **10** (1.7 g, 4.4 mmol) in CH$_2$Cl$_2$ (50 mL) under nitrogen atmosphere were added triethylamine (0.726 mL, 4.9 mmol) and mesyl chloride (0.522 mL, 4.9 mmol). The reaction mixture was stirred at room temperature for 2 h. After completed the reaction the pH was adjusted with saturated aqueous NaHCO$_3$ solution at ~7. The product was extracted with ethyl acetate (3 x 10 mL), the organic layer was dried with anhydrous Na$_2$SO$_4$ and evaporated to give a light brown crude product. The crude product was purified by column chromatography (inner diameter 25 mm and length 50 cm) using EtOAc/MeOH 9:1, as eluent. Rf: 0.66 (AcOEt/MeOH, 95:5). Yield 95%; light brown solid; We only reported the major signals in NMR. m.p. 40°C (dec.); MS/FAB$^+$: *m/z* 458 (M + H$^+$) HRMS (FAB$^+$): *m/z* 458.1380 [M + H$^+$] (Calculated for C$_{22}$H$_{24}$N$_3$O$_6$SH$^+$ 458.1386); ^{1}H NMR (CDCl$_3$, 400 MHz): □ 3.03 (s, 1H), 3.09 (s, 3H), 3.33 (s, 3H), 3.65 (t, *J* = 4.0 Hz, 2H), 4.19-4.05 (m, 4H), 4.52-4.48 (m, 2H), 6.95 (s, 1H), 7.25-7.14 (m, 3H), 7.82-7.66 (m, 2H), 8.49 (s, 1H). ^{13}C NMR (CDCl$_3$, 100 MHz): □ 37.9, 59.2, 66.5, 68.2, 68.7, 70.8, 77.7, 83.5, 102.2, 108.6, 109.7, 122.5, 122.6, 125.2, 127.9, 129.1, 147.0, 147.6, 153.6, 156.4.

7-(2-Chloroethoxy)-*N*-(3-ethynylphenyl)-6-(2-methoxyethoxy) quinazolin-4-amine (4) and 6-(2-chloroethoxy)-*N*-(3-ethynylphenyl)-7-(2-methoxyethoxy) quinazolin-4-amine (5). To a solution of **11** (1 g, 2.1 mmol) in THF-DMF (4:1, 50 mL) LiCl (4 g) was added. The mixture was irradiated under microwave (130 W, 90°C) for 1.5 h. The solvent was evaporated and water (25 mL) was poured into the reaction mixture, the solution was extracted with EtOAc (3 x 15 mL), the organic layer was separated, dried with anhydrous Na$_2$SO$_4$ and evaporated to give a white crude product. The crude product was purified by column chromatography (inner diameter 25 mm and length 50 cm) using EtOAc/hexane, 9:1, as eluent to obtain isomer **4**. Rf: 0.78 (AcOEt/MeOH, 95:5). Yield 25%; white solid; m.p. 164-166 °C; MS/FAB$^+$: *m/z* 398 (M+H$^+$) HRMS (FAB$^+$): *m/z* 398.1256 [M+H$^+$] (Calculated for C$_{21}$H$_{21}$ClN$_3$O$_3$H$^+$ 398.1271); ^{1}H NMR (CDCl$_3$, 400 MHz): □□3.50 (s, 3H), 3.52 (s, 1H), 3.87 (t, *J* = 4.8 Hz, 2H), 3.97 (t, *J* = 5.6 Hz, 2H), 4.34 (t, *J* = 4.4 Hz, 2H), 4.42 (t, *J* = 5.2 Hz, 2H), 7.16 (s, 1H), 7.27-7.24 (m, 1H), 7.37 (t, *J* = 1.6 Hz, 1H), 7.76–7.76 (m, 1H), 7.78 (s, 1H), 7.91 (t, *J* = 1.6Hz, 1H), 8.43 (s, 1H); ^{13}C NMR (CDCl$_3$, 100 MHz): □ 43.0, 59.7, 70.5, 70.6, 72.1, 84.4, 104.7, 108.8, 124.4, 127.2, 128.9, 130.0, 150.6, 154.2, 155.8. X-ray crystallographic structure in Fig. 2.[10] Isomer **5**, Rf: 0.67 (AcOEt/MeOH, 95:5). Yield 30%; white solid; m.p. 110-112 °C; MS/FAB$^+$: *m/z* 398 (M + H$^+$) HRMS (FAB$^+$): *m/z* 398.1279 [M + H$^+$] (Calculated for

$C_{21}H_{21}ClN_3O_3H^+$ 398.1271); ^1H NMR (CDCl$_3$, 400 MHz): ☐ 3.47 (s, 3H), 3.49 (s, 1H), 3.87-3.83 (m, 4H), 4.25 (t, *J* = 4.4 Hz, 2H), 4.31 (t, *J* = 6.0 Hz, 2H), 7.22 (s, 1H), 7.24 (s, 1H), 7.26-7.28 (m, 1H), 7.34 (t, *J* = 7.6 Hz, 1H), 7.74-7.72 (m, 1H), 7.84 (s, 1H), 8.66 (s, 1H); ^{13}C NMR (CDCl$_3$, 100 MHz): ☐☐41.8, 59.5, 68.6, 70.0, 70.6, 77.7, 83.4, 103.9, 109.2, 122.5, 123.0, 125.3, 128.1, 129.2, 138.7, 148.3, 154.0, 156.5. X-ray crystallographic structure in Fig. 3. [9]

Results and discussion

The fact that sometimes the synthesis of impurities is a difficult process, with poor yield, costly and also with long routes, suggests strongly looking for other alternatives. For example, the recovery of these impurities from the mother liquors of the same industrial process could be an option.

On the other hand, at industrial level a possible origin of synthesis of impurities **4** and **5** is when the reaction of hydrolysis of the compound **1** is carried out by action of the water present in the process, generating the impurity **8** and hydrochloric acid. The ion chloride, across a nucleophilic substitution reaction would attack in either of the two methylenes of **1**, to generate the precursors of **4** (Path A) or the precursor **5** (Path B) respectively, as described in Scheme 2.
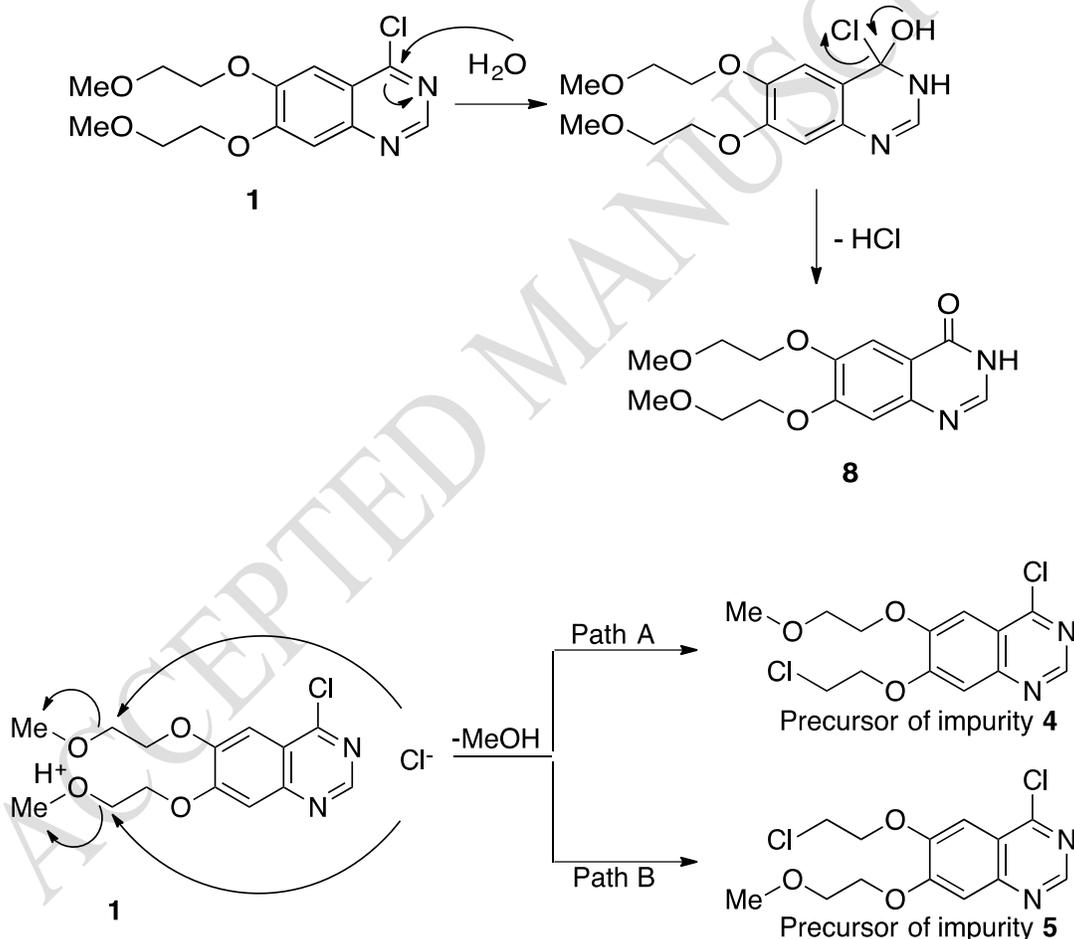

Scheme 2. Possible origin of the precursors of impurities **4** and **5**.

Therefore, the strategy of synthesis that was followed at laboratory level for impurities **4** and **5** is shown in Scheme 3. Thus, raw material **1** was treated with BBr$_3$ in DCM to obtain an inseparable mixture by TLC of four compounds of **9**: two of them correspond to the O-dealkylated products in the lower and upper part of the aliphatic chain of the molecule and

which contain the chlorine or bromine atom respectively at the position of the quinazoline ring. It should be noted that TLC shown only a single spot, and according to ^{1}H-NMR analysis showed duplicated signals, one of them was identified as chlorinate compounds and the other signal corresponded to the brominate compounds in a portion of 80:20 respectively.

To confirm the presence and proportions of each O-dealkylated compound in the reaction mixture a GC-MS analysis was performed: the results showed the presence of the chlorinated compound at a retention time of 25.6 min in a proportion of 86.3% and the brominated compound at 27.1 min in a proportion of 13.6%.
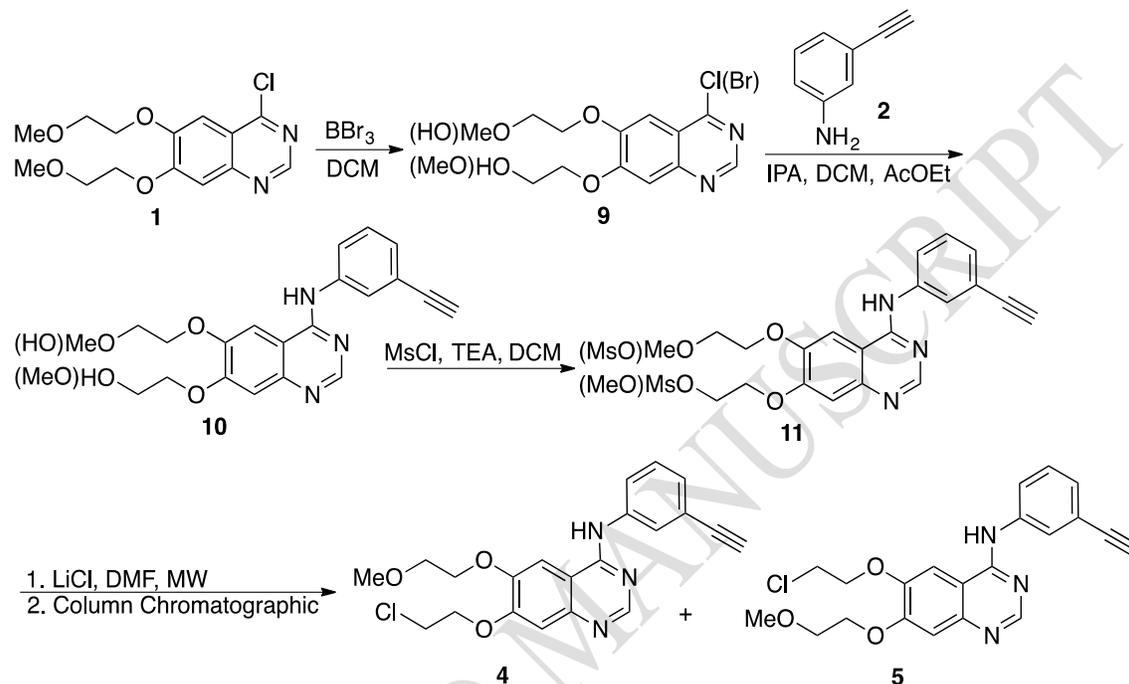
Scheme 3. Synthesis of Erlotinib impurities **4** and **5**.

It is noteworthy the fact that a part of the molecule had been replaced by bromine, and this did not pose a problem in the strategy since in the next reaction both, chlorine and bromine, would be replaced by the amine of the 3-ethynylaniline **2**.

Thus, the mixture of brominated and chlorinated products (**9**) was coupled with 3-ethynylaniline (**2**) to obtain **10** in 91% yield. Again, a spot was observed by TLC that corresponded to a mixture of two isomeric alcohols, which showed duplicated signals by ^{1}H-NMR in a proportion of 54:46.

Next, the mixture of mesylates **11** was obtained by conversion of **10** using methanesulfonyl chloride and triethylamine in DCM [10] at room temperature to obtain a 95% yield of both isomers in a proportion of 52:48, determined by ^{1}H-NMR. Again, this mixture was impossible to be separated by TLC.

Finally, impurities **4** and **5** were obtained by treatment of **11** with LiCl in DMF at 90 °C and with microwave irradiation at 130 watts during 1.5 h. Fortunately, on this occasion it was possible to separate these compounds by means of silica gel column chromatography, obtaining **4** in a yield of 25% and a purity by GC of 99.2%, while isomer **5** was obtained with a yield of 30% and a purity by GC of 98.7%.

After separating by column chromatography the impurities **4** and **5,** and in order to identify and assign unequivocally both isomers, they were recrystallized in methanol, obtaining in both cases monocrystals of suitable size for single-crystal X-ray diffraction analysis.

In Fig. 2, the X-ray structure for impurity **4** is displayed and as we can see, this isomer corresponds to the one with the chlorine atom in the lower part of the molecule. Also, as shown in Fig. 3, it was clear that compound **5** is structural isomer of **4**, the X-ray structure of **5** revealed that the chloride atom was located in opposite side of the molecule comparing to **4**. In the crystal structure of **5**, the molecule is stabilized by O-H···O hydrogen bond with a water molecule.
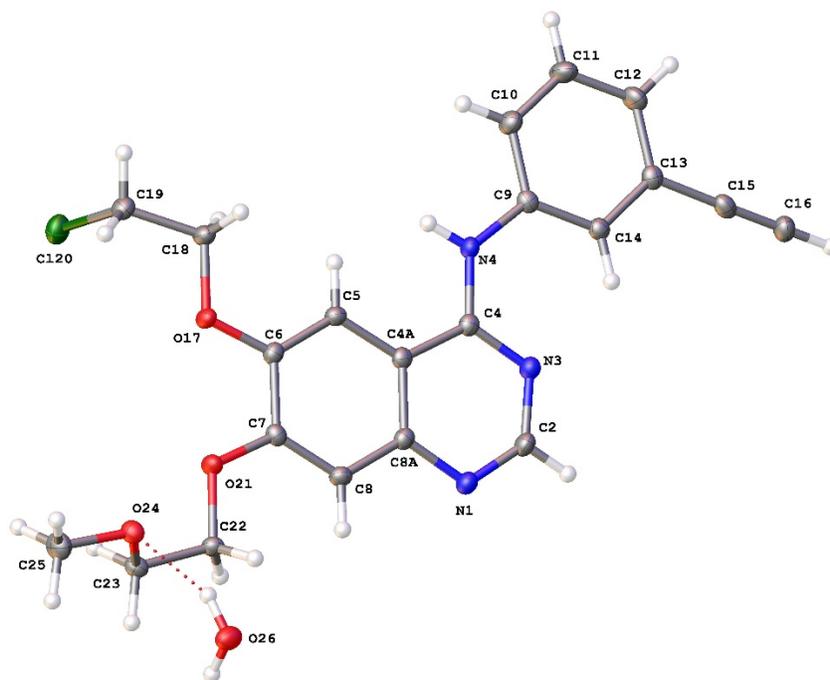
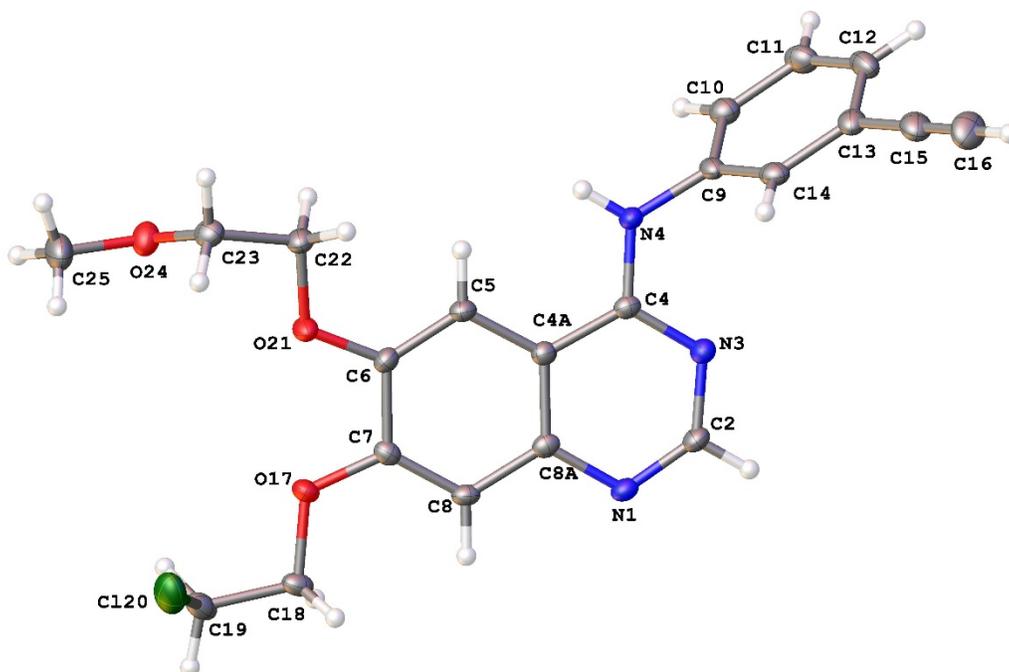**Fig. 2.** Crystal structure of **4**.

Fig. 3. Crystal structure of **5**.

Conclusions

The severe requirements in the development of drugs indicate the need to isolate and structurally identify the impurities generated in the manufacturing process. The correct structures of two impurities of Erlotinib was confirmed by of ^{1}H, ^{13}C-NMR spectroscopy and X-ray diffraction analysis. It was demostrated that chlorination can occur in both O-methylenes groups to generate structural isomers which could be separated by column chromatography.

Acknowledgments

References

1. Harris, C.; Hennequin, L.; Willerval, O. *Tetrahedron Lett.***2009**, *50*, 1600-1602.
2. Chandregowda, V.; Rao, V. G.; Reddy, G. C. *Org. Process Res. Dev.* **2007**, *11*, 813-816.
3. Zhang, G.; Zha, L. *Res. Chem. Intermed.* **2013**, *39*, 2303–2309.
4. Bridjes, A. J. *Chem. Rev.* **2001**, *101*, 2541-2572.
5. Bagul, C.; Rao, G. K.; Makani, V. K. K.; Tamboli, J. R.; Pal-Bhadra, M.; Kamal, A. *Med. Chem. Commun.* **2017**, *8*, 1810-1816.
6. Schnur, R. C.; Arnold, L. D. from PCT Int. Appl. (1996), WO 9639347 A1 Oct 03, 1996.
7. Q3A(R2) Impurities in New Drug Substances, ICH Tripartite Guideline.
8. Volk, B.; Dancsó, A.; Bakó, T.; Peltz, C.; Simig, G. *Org. Process Res. Dev.* **2011**, *15*, 339-342.
9. Crystallographic data was deposited at Cambridge Crystallographic Data Center (CCDC 1535571 for **4** and 1535572 for **5**).
10. Crossland, R. K.; Servis, K. L. *J. Org. Chem.* **1970**, *35*, 3195-3196.