PKC PHARMACEUTICALS, INC.

Bosutinib Report #1—February 20, 2012

I. Introduction: multiple "bosutinibs"

Accumulated evidence that seems absolutely unambiguous indicates that at least two different chemical compounds have been described as "bosutinib" in the scientific literature and/or offered for sale under that name in the biochemical research reagent marketplace. Also, there is NMR evidence for the possible existence of a third variant.

The purpose of Section I is to provide a basic perspective for this event – the emergence of more than one version of bosutinib into the scientific literature and the biochemical research reagent marketplace . Further details are presented in Sections II-VII below.

Revisions and corrections of initially published chemical structures, especially for natural products and other complex compounds, are not uncommon in the organic chemical literature and have appeared regularly for many decades in *J. Am. Chem. Soc.*, *J. Org. Chem.* and other journals. Further examples of revisions of structural er-

165 New Boston Street • Woburn, MA 01801 USA (781) 935-3900 • Fax (781) 935-5379

rors continue to appear occasionally in organic chemistry journals.

For an example, chosen at random, see Ookura, R. et al., "Structure Revision of Circumdatins A and B, Benzodiazepine Alkaloids Produced by Marine Fungus Aspergillus ostianus, by X-ray Crystallography." *J. Org. Chem.* **73**: 4245-4247 (2008)].

Such errors occur in the first place because the initially published structures were based only on chemical and spectroscopic reasoning, not X-ray crystal structure analysis, and the reasoning proved to be incorrect. X-ray crystallography provides rigorous, unambiguous, 100% certainty for small-molecule chemical structures. Such certainty cannot be achieved via chemical and spectroscopic techniques alone for many types of small molecules.

In contrast to occasional structural revisions in the organic chemical literature, instances where biochemical research reagent vendors have offered a small-molecule chemical of incorrect structure for worldwide sale and distribution appear to be extremely rare.

Indeed, although we have not attempted a thorough search, we are only aware of one other instance beyond the current bosutinib issue during the past 32 years. We would be interested to hear of any other such instances that have occurred.

> [Please note that the comments above about the rarity of compounds of incorrect structure being offered for worldwide sale apply only to the biochemical research reagent market. At present there appear to be about 170-200 such vendors. Lists of many typical vendors in this marketplace can be viewed in the "price comparison tables" on the LC Laboratories website (**www.LCLabs.com/PriceComparisonsbyAlpha.php4**—case-sensitive).
>
> We have no historical knowledge of whether or not incorrect compounds may have been offered for sale in other major chemical markets, such as those for (i) organic chemicals, of the type sold by *e.g.*, Aldrich and Alfa, and (ii) active pharmaceutical ingredients ("APIs") that are incorporated into finished dosage forms of drugs for human or veterinary use. Each of these two areas may have hundreds or even thousands of active vendors worldwide.
>
> Notably, in the context of the present bosutinib issue, it may turn out that one or more incorrect isomers of dichloromethoxyaniline may recently have been sold in the organic chemical marketplace. Further details on this point appear in Section V below.]

The previous instance we know of, where an incorrect compound was sold in the worldwide biochemical research reagent marketplace, involved two different compounds being sold under the same name, "H7". The H7 event (described in Section III below) serves as a significant and useful historical precedent for the current bosutinib issue.

II. Summary and current status: evidence for multiple "bosutinibs", and work in progress

Research scientists at PKC Pharmaceuticals, Inc. have obtained what we believe is unambiguous physical (melting point), HPLC (two different stationary phases), TLC (two different mobile phases) and spectroscopic evidence that at least two different compounds are currently being offered or have been offered under the name bosutinib by vendors serving the biochemical research reagent marketplace.

Though our firm did not cause this overall event, we are spending quite a bit of time, money and energy to help ourselves and others to understand and address it, including timely informing the scientific community of our findings.

To date, we have obtained data on physical samples of bosutinib from ten apparently independent vendors of bosutinib – that is, as best we can tell, none of the ten vendors is a distributor for any of the other nine. We find that two vendors are offering what we have designated as "Version S1" bosutinib and eight of them are offering "Version S2". (The "S" denotes "samples in hand".)

Moreover, published data from two independent X-ray crystallographic studies of bosutinib bound to enzymes show two different chemical structures, which we have designated as "Version X1" and "Version X2". Version X1 unambiguously has the structure originally published for bosutinib in 2001 by a group at Wyeth-Ayerst

(now Pfizer) [Boschelli, D.H. et al., *J. Med. Chem.* **44**: 3965-3977 (2001)].

Finally, a discrepancy in NMR data relating to bosutinib indicate the possible existence of a variant that we have designated "Version N1" because it seems to be different spectroscopically from both Version S1 and Version S2.

Thus, according to all information presently available to us, there are five "possible" bosutinib entities, characterized by the following three independent sets of data:

- from our testing of physical samples in hand, two clearly different compounds:
  - Version S1
  - Version S2
- from published X-ray studies only, two clearly different compounds:
  - Version X1 (structure matching the original published structure of bosutinib)
  - Version X2
- from published NMR data only, potentially different from S1 and S2:
  - Version N1.

As more data become available to our group and others, the five prospective bosutinib versions listed above might well "merge" into only two genuinely different versions. But as of this writing there seem to be three versions, and we are not aware of any evidence to rule out a total of three or more distinct bosutinibs having been described in the scientific literature and/or offered for sale by biochemical research reagent vendors.

We are presently attempting to obtain samples of Versions X1, X2 and N1 so that we can make direct physical, chromatographic and spectroscopic comparisons among them and other samples already in hand.

We are also attempting to ascertain the original synthetic sources and routes used to produce (i) each of the bosutinib samples giving rise to the data for each bosutinib version, especially X1, X2 and N1, and (ii) the dichloromethoxyaniline used in each case as a raw material in the respective bosutinib syntheses. Please see below for extensive details on these points.

Overall, analytical work on samples of bosutinib from a total of 13 apparently independent sources (including our LC Laboratories division) is in progress in our laboratories. We may be able to obtain and analyze bosutinib samples from an additional 2-4 sources.

In addition to analytical work here, we are also pursuing chemical synthesis studies to examine some aspects of bosutinib raw materials and synthetic routes.

The various efforts undertaken by our group, along with our literature review and communications with other institutions, are intended to further define the scope of the bosutinib issue and facilitate its resolution. Specifically, our goals are to:

(i) determine how many different compounds have been described as bosutinib in the scientific literature and/or offered under the name bosutinib by biochemical research reagent vendors in the recent past;

(ii) obtain further chemical, chromatographic and spectroscopic information about those compounds;

(iii) in light of apparently conflicting X-ray crystal structures and conflicting NMR data already published, contribute our efforts to those of others to establish what actual chemical structure the name bosutinib should apply to [the starting assumption would of course be that the structure published in 2001 by Wyeth-Ayerst (now Pfizer) researchers should be correct, but please see Sections IV-VI below for important further discussion];

(iv) do what we can to assist all parties interested in bosutinib and bosutinib analogs in sorting out all of the issues arising from incorrect bosutinib having entered the biochemical research reagent marketplace and to enable all biochemical research reagent vendors to offer only correct bosutinib going forward;

(v) to the extent that one or more incorrect bosutinibs arose because a starting material(s) of incorrect structure was incorporated into the final compound, establish what the correct starting material(s) should be and how they can be obtained, so that only correct starting materials will be incorporated into bosutinib and the many bosutinib analogs now being prepared in many medicinal chemistry laboratories around the world;

(vi) promptly publish our work in a peer-reviewed journal and/or on the Internet as appropriate.

Much of our data may be subject to pre-publication embargo, but within these limits we will endeavor to

promptly release and post on the internet as much of our data as we can.

This Bosutinib Report #1 report is being sent to all LC Laboratories bosutinib customers, and it will also be posted will be posted at **www.pkcpharma.com/TwoOrMoreBosutinibs.html** (this URL is case-sensitive—it must be entered with capital letters as shown). Please visit that web page for periodic updates to this report.

To the best of our knowledge, in its entire 32 year history our LC Laboratories division has never offered nor sold a compound that did not match the chemical structure provided to customers in our catalog, website and other informational materials. However, if the compound that LC Laboratories has been offering as bosutinib proves not to match the chemical structure we have provided, then (i) according to data now in hand, some other vendors who we believe did not obtain this compound from LC Laboratories (and thus are apparently independent sources) are also selling bosutinib that does not match the initially published structure, and (ii) we will of course fully honor our warranty [**www.LCLabs.com/Warranties.php4**—case-sensitive].

III. Historical precedent: "H7"

The only previous instance we know of in the past 32 years where an incorrect compound was sold in the worldwide biochemical research reagent marketplace serves as a significant and useful historical precedent for the current bosutinib topic.

That compound, a protein kinase inhibitor based on an isoquinolinesulfonamide scaffold, carries the commonly used trivial name "H7".

In 1984 one of our scientists read a paper from the Hidaka laboratory [Hidaka, H. et al., *Biochemistry* **23**: 5036-5041 (1984)] [http://pubs.acs.org/doi/abs/10.1021/bi00316a032], disclosing a series of isoquinolinesulfonamide-based protein kinase inhibitors.

One notable compound in the paper was "H7", depicted in the Hidaka paper as having the structure shown here in Figure 1.A., with the piperazinyl methyl group (arrow) positioned close to the sulfonyl group.

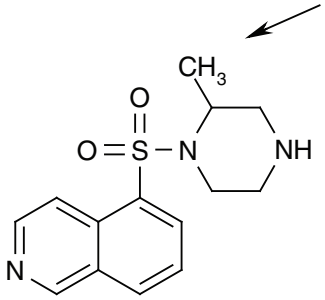


Figure 1.A. Structure of H7, as depicted in the 1984 Hidaka paper.

However, it appeared that the synthetic scheme given for H7 in this paper was almost certainly incorrect, in a way that an astute organic chemistry student would probably recognize. The direct condensation reported (*Biochemistry, op. cit.*, p. 5037, middle of right hand column) should theoretically put the piperazinyl methyl group at the more distant position from the sulfonyl group, as shown in Figure 1B, arrow.

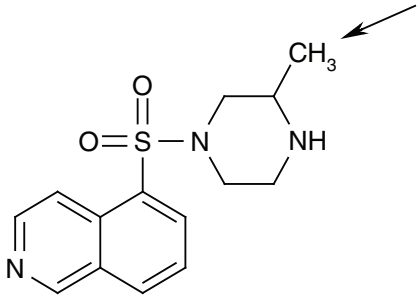


Figure 1.B. "Iso-H7" -- the structure expected to result from the H7 synthesis in the 1984 Hidaka paper.

Data in this paper showed that "H7" had potencies of 3-6 μM for inhibiting protein kinases A, C and G ("PKA", "PKC" and "PKG"). At that time and continuing for a couple of decades or more, inhibitors of PKC were of particularly high interest worldwide, and the compound H7 entered into wide use as a biochemical research reagent.

Because our company was neither working with H7 nor offering it for sale to other researchers, we did not pursue the matter further at that time.

However, in view of the structural ambiguity, the article obviously posed a basic problem: What is H7?

There were several possible answers.

"H7" (of empirical molecular formula $C_{14}H_{17}N_3O_2S$) could be taken to be one or more of:

(i) a particular arrangement of those atoms and bonds depicted graphically, such as Figure 1.A.;

(ii) a substance of that molecular formula in a bottle in Prof. Hidaka's lab;

(iii) a substance of that molecular formula made via the actual starting materials and the particular synthetic reaction(s) and reaction conditions in the Hidaka paper; or

(iv) a substance of that molecular formula with the specific inhibitory potencies reported in the Hidaka article for rabbit skeletal muscle PKA, rabbit brain PKC, and pig lung PKG.

About six years later the H7 topic arose again. While reviewing PKC inhibitor literature, our group came across an article noting that H7 from two different suppliers, Seikagaku and Sigma, differed in their PKC inhibitory potencies by a factor of about four. This stimulated our group to do further literature searching, which produced a second article that reported conflicting results between Seikagaku and Sigma H7 as well as five additional articles that reported anomalous H7 results but did not compare material between vendors.

As before, our firm (named "Alder Research Center" at the time) was neither using nor selling H7. However, essentially as a public service, we undertook an investigation of this problem and published the results [Quick, J. et al., "The Structure and Biological Activities of the Widely Used Protein Kinase Inhibitor, H7, Differ Depending on the Commercial Source." *BBRC* **187**: 657-663 (1992)] [**http://www.sciencedirect.com/science/article/pii/0006291X9291245L**].

In this study, the combination of synthetic chemistry tests in conjunction with a bioassay confirmed that the Hidaka synthetic route was very likely wrong. Our work showed that the synthetic route in the 1984 Hidaka paper produced a compound that was 4.6-fold weaker as a PKC inhibitor than the product of an indirect route that should, by simple theory, match the graphical structure depicted as H7 (Figure 1.A.) in the 1984 Hidaka article, with the piperazinyl methyl group adjacent to the sulfonyl group.

Thus, based on both chemical and bioassay results, we could confirm that Sigma was indeed selling the wrong isomer. The article of 1989 date reporting the anomalous results, combined with our 1991-1992 purchases and chemical and bioassay results, suggest that Sigma had probably been selling the wrong isomer for about 3-4 years and perhaps longer. Because Sigma is a large, worldwide supplier of biochemical research reagents, it seems likely that the incorrect H7 was supplied to at least hundreds of biomedical laboratories around the world.

NOTE, however, that our work did not rigorously prove the structure and associated pharmacological properties of "H7". Though we "assigned" the structure as that depicted in the Hidaka paper (Figure 1.A.), the rather unlikely possibility nonetheless remained that for unanticipated reasons (i) the "direct condensation" synthetic route in the 1984 Hidaka paper placed the methyl group next to the sulfonyl group in spite of steric considerations, and (ii) mono-BOC protection of 2-methylpiperazine took place on N1 instead of being sterically directed to N4 (please see the full 1992 *BBRC* article for details).

Specifically, the possibility remained that the real "H7", having the higher potency as a PKC inhibitor, had in fact the graphical structure of "Iso-H7" (Figure 1.B.) as depicted in the 1992 *BBRC* article.

[A further independent, unlikely possibility also remained, namely, that the long-known "isoquinoline-5-sulfonyl" starting material employed in one or another synthesis in this story was the wrong structure (there are 14 possible quinoline- and isoquinoline-sulfonyl isomers).]

A structure for H7 in the form of its complex with the catalytic subunit of PKA was eventually analyzed by X-ray crystallography and published [Engh, R.A. et al., "Crystal Structures of Catalytic Subunit of cAMP-dependent Protein Kinase in Complex with Isoquinolinesulfonyl Protein Kinase Inhibitors H7, H8 and H89." *J. Biol. Chem.* **271**: 26157-26164 (1996)], four years after our 1992 *BBRC* paper was published. This H7 structure matched the graphical structure depicted in the 1984 Hidaka paper (Figure 1.A.).

PLEASE NOTE FURTHER, however, that neither the 1992 *BBRC* paper nor the 1996 X-ray crystal structure paper rigorously, incontrovertibly established the true structure of H7 and its associated pharmacological properties with 100% certainty. The X-ray paper does not mention its source of H7, nor were confirmatory measurements of the PKA, PKC and PKG inhibitory potencies from the same sample of H7 reported in the paper.

Thus, the remote possibility remains that (i) as above, the "H7" and "Iso-H7" syntheses yielded, as a doubly anomalous scenario, the opposite of the expected products, and (ii) the lab that did the 1996 X-ray studies inadvertently purchased the wrong compound, since both isomers had been available in the marketplace in the 1990's.

To establish true 100% structural and pharmacological certainty for H7, namely, the $C_{14}H_{17}N_3O_2S$ compound (i) made from isoquinoline-5-sulfonyl chloride and 2-methylpiperazine, (ii) by whatever synthetic route, and (iii) having the ~4-fold higher potency as a PKC inhibitor, it would be necessary to do both the X-ray structure determination and confirmatory biological assays on material taken from the same lot, "the same bottle" (see Section IV.B. below for further discussion of this key point).

Admittedly, the probabilities for these potential artefacts for H7 are extremely low, and the structure of H7 and its associated specific biological activity are essentially established well beyond a reasonable chemical doubt. But lacking kinase inhibition assays and X-ray structure studies done on material from the same batch of H7, this structural and pharmacological conclusion relies entirely on chemical and spectroscopic reasoning, not incontrovertible, 100% proof.

These same potential artefacts and structural and pharmacological uncertainties are present to much more serious degree in the case of bosutinib. In the bosutinib case, chemical and spectroscopic reasoning are not sufficient to achieve chemical and pharmacological certainty, and potential artefacts of the type described above for H7 (incorrect starting materials, anomalous reaction results) are much more likely possibilities.

IV. Chemical and pharmacological certainty

IV.A. Certainty for chemical structures

It is common practice in synthetic and medicinal chemistry to characterize synthetic and semi-synthetic products by chemical and spectroscopic means only, not by X-ray crystallography. Conclusions about chemical structures are generally based only on chemical reasoning (starting materials, reaction conditions, theoretically expected products, elemental analyses, etc.) and spectroscopic reasoning (NMR, IR, MS, etc.).

Moreover, in synthetic and medicinal chemistry there is commonly a substantial reliance upon the correctness of the structures of raw materials giving rise to the substituents and substructures that are installed in the final molecule. This is a particular concern for portions of such starting materials that cannot be easily distinguished from potential isomers by NMR and other routine and inexpensive analytical techniques.

X-ray crystallography is an extremely powerful and definitive technique for establishing chemical structures. However, X-ray studies are relatively expensive as an analytical technique and are sometimes very difficult or impossible to carry out in practice. Consequently, X-ray analysis is not done on more than a vanishingly small fraction of new medicinal chemistry compounds.

Thus, for the many cases where X-ray analysis has not been done and for which NMR and other spectroscopic techniques cannot rigorously rule out various isomeric possibilities, there is no absolute, 100% certainty of the actual structures of some compounds coming out of the hundreds or perhaps thousands of medicinal chemistry laboratories that are active worldwide.

In many cases only an X-ray crystal structure determination can provide 100% structural and pharmacological certainty and rule out certain potential artefacts in synthetic and semi-synthetic compounds, such as (i) starting materials of (unrecognized) incorrect structure, (ii) subtle, unexpected structural results, isomerizations, rearrangements and the like, and (iii) artefacts from altered conditions in synthetic reactions,.

Indeed, revisions and corrections of initially published chemical structures, especially for natural products and other complex compounds, occur occasionally and have appeared regularly for decades in *J. Am. Chem. Soc.*, *J. Or. Chem.* and other journals. Further examples continue to appear now and then in the organic chemical literature.

Such errors occur in the first place because the initially published structures were based only on chemical and spectroscopic reasoning, not X-ray analysis, and the reasoning proved to be incorrect. Even the extremely powerful and successful methods of modern chemical and spectroscopic structure determination are not sufficient to provide 100% certainty in all cases.

IV.B. Pharmacological certainty

At the most fundamental level, biomedical interest in chemical compounds is primarily directed towards their pharmacological properties.

In cases where a useful pharmacological property is observed for an individual compound or a crude extract from naturally-occurring biomass, the primary focus is on the pharmacological or therapeutic utility of the material at hand.

Determination of the correct chemical structure of a pharmacologically useful compound is of course intrinsically desirable and is essential for successful further development of analogs of known structure.

Modern practice of non-X-ray structural chemical characterization is extraordinarily powerful and successful – only in relatively rare cases do errors in structural assignments arise. However, once uncertainties have arisen for the structure of a chemical compound as obtained from different sources, there is also a large pharmacological impact:

Once multiple, chemically distinct versions of a compound of pharmacological, biomedical or therapeutic interest enter into use, as was the case for H7, this also necessarily calls into doubt any pharmacological properties that may have been published in the original or later studies, because the identity of the compounds used in those bioassays may be unknown or in question.

These ambiguities can only be rigorously resolved by publication of 100% proof of the intended chemical structure using material from the same batch of material that shows the expected pharmacological properties in at least one but preferably in many bioassays. In many cases for even moderately complex molecules, 100% certainty can only be achieved by X-ray crystal structure determination of material from the same batch used to establish a baseline pharmacological profile.

It would be extremely unfortunate if a chemical of misidentified structure were to enter human or veterinary clinical trials. However, if a given compound used for preclinical toxicology and Phase I safety studies is exactly the same as the compound that enters Phase II and later efficacy studies, then at least the primary concern has been satisfied, namely, that patients have been protected by the earlier toxicological characterizations.

V. Multiple bosutinibs

V.A. Background

All of the problems and uncertainties that were synthetically and structurally trivial in the H7 case are present in the bosutinib case, but with much higher degrees of complexity, importance, difficulty in structural analysis and likelihood of error.

We were alerted to the likely presence of two bosutinibs in the biochemical research reagent marketplace by a customer of our LC Laboratories division.

Published evidence (see Sections V.B. and V.C, below) and our unpublished results to date show that at least two compounds have been described as bosutinib, and there is spectroscopic evidence for a third version (see Section V.D. below).

As was the case for H7, there are multiple ways that incorrect versions of bosutinib or any other similarly complex molecule could inadvertently be produced, without being recognized as incorrect by common methods of chemical and spectroscopic analysis. This is particularly the case for the regions of the bosutinib structure that carry the five aromatic protons.

Sources of error could be:

(i) use of one or more starting materials of (unrecognized) incorrect structure;

(ii) expectation that a reaction would provide a particular product based on theory, but in actuality it gives a different structural result because the theory is wrong; and/or,

(iii) expecting that a reaction would provide a particular product based on past reaction runs, but in actuality it gives a different structure because of altered and unrecognized changes in conditions or presence or absence of trace contaminants in a subsequent reaction.

Given that there are now potentially two or more structures that have been described as bosutinib (see Sections V.B. and V.C. below for details), obviously the question arises, "What is bosutinib?".

Again as was the case for H7, there are several possible answers.

Bosutinib (of empirical formula $C_{26}H_{29}Cl_2N_5O_3$) could be taken to be one or more of:

(i) a particular arrangement of those atoms and bonds depicted graphically, such as in Figure 2, as was published in the 2001 *J. Med. Chem.* article.

(ii) a substance in a bottle at Wyeth-Ayerst (the company that originated bosutinib and was later acquired by Pfizer) at the time that company first made the compound;

(iii) a substance made by a particular synthetic sequence and reaction conditions, such as in *J. Med. Chem., op. cit.* or by one of the three other published synthetic routes that we are aware of;

(iv) a substance with (i) certain specific inhibitory potencies or inactivities when tested on a panel of eleven rat and human kinases [*Cancer Res.* **66**: 11314-11322 (2006)] or (ii) other activity profiles as measured in other set(s) of bioassays;

or, a new, very important alternative:

(v) the substance that is being administered to human patients in Pfizer-sponsored clinical trials.

Structural variations for multiple substances described as bosutinib are most likely to occur in two regions, "A" and "B", shown enclosed in boxes in Figure 2.

For substructure "A", the quinoline substructure, alternative versions with different substitution patterns in this quinoline portion could plausibly arise in a number of synthetic steps among those used in the four published syntheses of bosutinib that we have found.

Many of these potential alternative substructures for "A" can in concept be tested and ruled out by careful chemical and spectroscopic analyses. Moreover, the two available X-ray studies of bosutinib (see Section V.B. below) appear to agree in depicting identical versions of the "A" substructure that comport with the 2001 *J. Med. Chem.* report.

However, there is much greater potential for unrecognized variants to have arisen in substructure "B" in Figure 2, the anilinic portion,:

- there are 16 possible isomers for the substituted aniline derivative that would be used in making bosutinib, as shown in Figure 3, which shows all possibilities for a carbocyclic six-membered ring of formula $C_7H_5Cl_2NO$, with methoxy, amino, and two chloro substituents;
- twelve of these isomers have been reported and have been assigned CAS numbers;
- eight of these anilinic isomers are listed in ChemExper [**http://www.chemexper.com**] as having at least one commercial source (enter exactly "3,4-dichloro-5-methoxyaniline", not "2,4-dichloro-5-methoxyaniline", into the ChemExper search box);

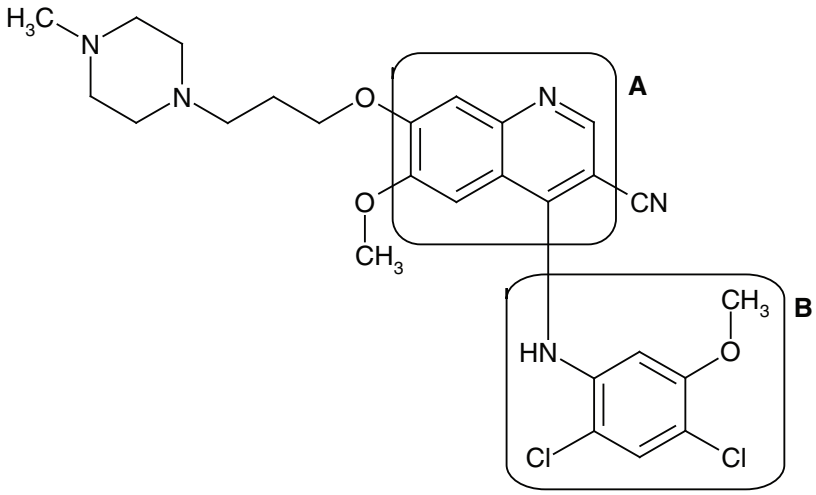
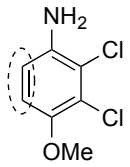


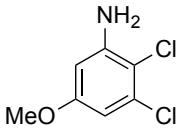
Figure 2. The structure of bosutinib as depicted in the 2001 *J. Med. Chem.* article. A = the bicyclic quinoline substructure. B = the anilinic substructure. We have designated this structure as "Version X1".

Figure 3: Isomers of Dichloromethoxyaniline

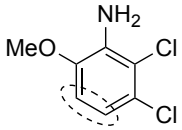
***o*-dichloro**



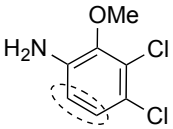
105630-40-8



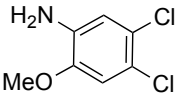
1221716-03-5



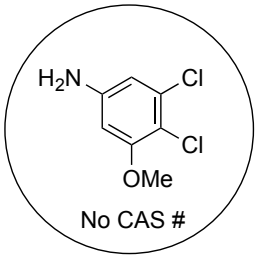
No CAS #



No CAS #

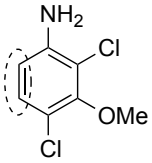


60468-21-5

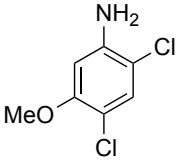


No CAS #

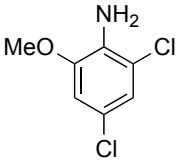
***m*-dichloro**



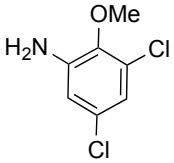
80026-11-5



98446-49-2

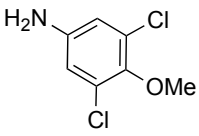


93839-14-6

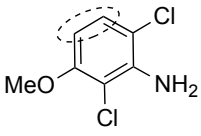


33353-68-3

symmetrical

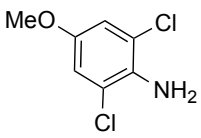


32407-11-7



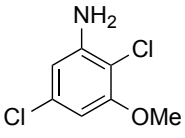
55285-43-3

symmetrical

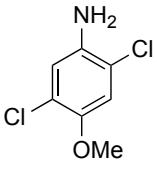


6480-66-6

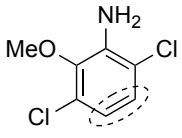
***p*-dichloro**



No CAS #



101251-23-4



101252-85-1

- we are aware of three different published syntheses of the 2,4-dichloro-5-methoxyaniline isomer putatively found in bosutinib (isomer enclosed in a square in Figure 3);
- two of these syntheses can each formally yield six different final isomers that have no ortho hydrogens (which would be easily and unambiguously spotted and distinguished from the other six potential isomeric products by NMR);
- two of the syntheses contain conflicting data for an intermediate, tending to call into question the identicality of the respective final products;
- one of the syntheses provides no structural "proof" at all, other than that "NMR data were consistent with the proposed structure" of a derivative, but NMR data may very well not be able to distinguish among all of the isomers that could be produced by this route;
- the structural "proof" in a second synthesis was merely a melting point similar to that reported for this isomer in a third, decades-earlier synthesis;
- the structural "proof" for that isomer in the third, decades-earlier synthesis was merely similar melting points for (i) a non-anilinic derivative obtained by a two-step transformation of that isomer, compared with (ii) the same non-anilinic derivative obtained in an earlier publication;
- there are the two separate sets of published data noted in Section V.B and V.C. below, along with unpublished data in our hands, that each raise the possibility of a third bosutinib structure in circulation, and at this point it is not clear if they represent the same third version or third and fourth versions of bosutinib.

It appears that proton and C13 NMR can rigorously distinguish or rule out only some of the 16 anilinic isomers, such as for the six isomers that have ortho hydrogens (dashed ovals in Figure 3), and perhaps for some isomers with meta hydrogens that happen to show splitting. And chemical NMR shifts probably cannot reliably prove the structure of one isomer over some or all of the other isomers.

Thus, it may be quite difficult to achieve 100% certainty for the bosutinib structure using just the common chemical and spectroscopic techniques, because it is necessary to rigorously establish which anilinic isomer has been incorporated into the final molecule and to rigorously rule out possible positional isomerism in the quinoline nucleus that some synthetic schemes might allow.

In our opinion, for a molecule of the particular complexity of the present case, this essential certainty can only be done by X-ray crystallography. And for pharmacological and therapeutic certainty, it is necessary to do that X-ray determination on material taken from the same lot that was used for the primary or major pharmacological profiling of the compound.

At first blush, it may seem extreme to advocate for X-ray crystal studies and biological tests on a single relevant batch of bosutinib to establish the structure and associated pharmacological questions with 100% certainty. But please consider the following:

The H7 issue described in Section III was the result of errors in a 1984 publication combined with apparent errors on the part of synthetic organic chemists at one vendor and/or its supplier. The scope of the H7 issue was rather limited – all other suppliers we knew of at the time got the chemistry right; there were only a couple of likely alternatives; and the starting materials were basically not at all in question.

In contrast, the problems surrounding bosutinib appear to reach more deeply into the worldwide chemical enterprise. Among other considerations, a seemingly simple bosutinib starting material, namely a dichloromethoxyaniline isomer whose supposedly correct synthesis and structural "proof" were reported almost a hundred years ago, provides many opportunities for error.

As noted above, there are 16 dichloromethoxyaniline isomers of likely similar chemical and spectroscopic properties. Moreover, there are apparently many commercial suppliers of these isomers, and there are multiple (potentially incorrect) syntheses of at least one and perhaps several of the isomers to contend with.

Thus, an incorrect anilinic isomer could easily find its way into bosutinib by:

- choice of an incorrect synthesis of the anilinic isomer;
- mistakenly purchasing an isomerically incorrect raw material for which the manufacturer's or distributor's chemical substance and labeling both correctly correspond to one another, but the chosen compound is an incorrect isomer for the desired product of the synthesis;

- purchasing a bottle of material labeled by the manufacturer or distributor as the 2,4-dichloro-5-methoxy isomer but containing a different, incorrect isomer.

V.B. Conflicting X-ray structures for bosutinib have been published

An X-ray crystal structure of bosutinib bound to a protein has recently been published [Chao, L.H. et al., *Cell* **146**: 732-745 (2011)] [**http://www.ncbi.nlm.nih.gov/pubmed/21884935**, and, on the Protein Data Bank website, **http://www.rcsb.org/pdb/explore/explore.do?structureId=3SOA**]. This study shows the same chemical structure (Figure 2) for the inhibitor as that described in the initial publication disclosing this compound [*J. Med. Chem.* **44**: 3965-3977 (2001)] [**http://pubs.acs.org/doi/abs/10.1021/jm0102250**]. We have designated this structure as "Version X1". (The sample of bosutinib used in this X-ray study was provided by Prof. Giulio Superti-Furga of Ce-M-M in Austria. We have inquired with Prof. Superti-Furga as to the ultimate source of that bosutinib but have not yet received a reply.)

This chemical structure includes a substructure derived from a substituted aniline, "**2**,4-dichloro-5-methoxyphenyl)amino. . .", which is the anilinic isomer enclosed in a rectangular box in Figure 3.

Not incidentally, this X-ray study proves that (i) the indicated structure can be synthesized successfully, and (ii) the 2,4-dichloro-5-methoxyanilinic substructure can also be synthesized.

However, a second X-ray structure of bosutinib has also appeared on the Protein Data Bank website [**http://www.rcsb.org/pdb/explore.do?structureId=3ZZ2**]. The full publication has not appeared as of this writing. (We have inquired with Prof. Knapp as to the ultimate source of that bosutinib but also have not yet received a reply.)

Unambiguously, the structure of the inhibitor in this study is different from the one in the *Cell* paper.

Specifically, although the title of this second X-ray study cites the inhibitor as including the partial structure "**2**,4-dichloro-5-methoxyphenyl)amino. . .", examination of the 3-dimensional X-ray-generated structure itself and the compound name cited later on the web page under "Ligand Chemical Component" shows that the structure of actually found for the inhibitor in this second bosutinib X-ray study is that depicted in Figure 4. That structure includes the partial structure "**3**,4-dichloro-5-methoxyphenyl)amino. . ." (arrow), which is the circled isomer in Figure 3. We have designated this second X-ray crystal structure as "Version X2".

Notably, the 3,4-dichloro-5-methoxyaniline (circled isomer in Figure 3) that would be utilized in a synthesis

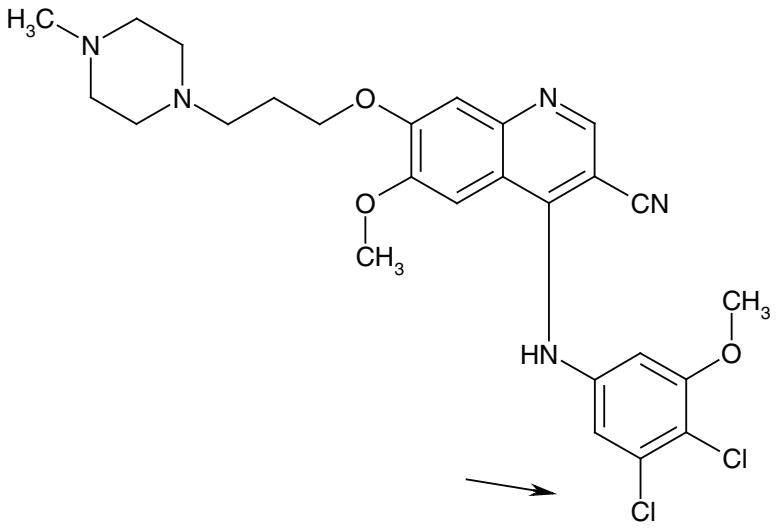


Figure 4. Chemical structure corresponding to Version X2, as determined by X-ray crystallography. Note the altered position of one of the chlorine atoms, arrow, compared to Figure 2. We have designated this structure as "Version X2".

resulting in the bosutinib structure in the second X-ray study apparently has not been reported in the literature. It also has apparently not been knowingly offered for sale by any chemical reagent vendors, and it does not have a CAS number as of this writing. However, 3,4-dichloro-5-methoxyaniline is one of the isomers that could plausibly be formed as an unrecognized alternative product in two of the three published syntheses we have found for "2,4-dichloro-5-methoxyaniline".

V.C. Conflicting NMR data for bosutinib have been published

In 2010 Yin et al. reported a new synthesis of bosutinib, which included use of putative 2,4-dichloro-5-methoxyaniline [Yin, X.J. et al., *Molecules* **15**: 4261-4266 (2010)] [http://www.mdpi.com/1420-3049/15/6/4261/].

The text of the published paper lists numerical proton NMR data for the bosutinib product (first row in the Table below). However, Prof. Fei Li was kind enough to provide us with the actual NMR spectrum itself. As the second row of the Table below shows, the spectrum itself shows a minor but clear difference from the numerical data in the text of the article. The slight difference between these two assessments of the spectrum are shown in bold. However, this minor difference does not affect the basic conclusions in the following discussion.

The Table shows that there are numerous discrepancies between the proton NMR data for bosutinib in the *Molecules* paper (first and second rows) compared to the data in the 2001 *J. Med. Chem.* article (third row).

Specifically, although the aliphatic proton data are very closely concordant in the two publications, the aromatic + NH region shows distinct differences.

Thus, the compound in the *Molecules* paper is clearly different from the compound in the *J. Med. Chem.* paper.

Curiously, the *J. Med. Chem.* and *Molecules* papers both report exactly the same melting range for these two different compounds, 116-120 °C.

The relationship between the compound described in this Section V.C. and the compounds that gave rise to the two X-ray structures discussed in Section V.B. remains to be determined.

V.D. Our analyses of bosutinib samples from 10 sources show two different compounds

Our scientists have obtained what we believe is unambiguous physical (melting point), HPLC (two different stationary phases), TLC (two different mobile phases) and spectroscopic evidence that at least two different compounds are currently being offered or have been offered in the very recent past under the name bosutinib by vendors serving the biochemical research reagent marketplace.

Specifically, data that we have obtained to date on physical samples of bosutinib from ten apparently independent vendors (as best we can tell, none of the ten is a bosutinib distributor for any of the other nine) show that two of them are offering what we have designated as "Version S1" and eight of them are offering "Version S2".

Source	**ArH + NH**				
	A	**B**	**C**	**D**	**E**
Molecules, data from text	9.51 (s, 1H)	8.52 (s, 1H)	7.64 (s, 1H)	**7.43 (s, 1H)**	**7.31 (brs, 2H)**
Molecules, directly from spectrum	9.51 (s, 1H)	8.52 (s, 1H)	7.65 (s, 1H)	**7.36 (s, 2H)**	**7.31 (brs, 1H)**
J. Med. Chem.	9.60 (s, 1H)	8.40 (s, 1H)	7.82 (s, 1H)	7.73 (s, 1H)	7.31 (brs, 2H)

Furthermore, the *Molecules* NMR data for the aromatic protons in their product (Table, first and second rows) differ from what we have found in the NMR spectra for Versions S1 and S2.

The published NMR data taken together with NMR data we have obtained thus appear to show that Versions S1 and S2 and the compound reported in the *Molecules* paper are three different "bosutinibs". Therefore we have designated the spectroscopically distinct compound in the *Molecules* paper as "Version N1".

In sum, according to all information presently available to us, as collected in Sections V.B., V.C. and this Section V.D., there are five prospective bosutinib entities, characterized by the following three independent sets of data:

- from our testing of physical samples in hand, two clearly different compounds:
  - Version S1
  - Version S2
- from published X-ray studies only, two clearly different compounds:
  - Version X1 (structure matches the original published structure of bosutinib)
  - Version X2
- from published NMR data only, potentially different from S1 and S2:
  - Version N1.

As more data become available to our group and others, the five prospective versions listed above might "merge" into only two genuinely different versions. But as of this writing there seem to be three versions, and we are not aware of any evidence to rule out a total of three or more distinct bosutinibs having been described in the scientific literature and/or offered for sale by biochemical reagent vendors.

V.E. What is bosutinib?

For the question, "What is bosutinib?", our opinion is that, in view of:

> (i) the substantial structural ambiguities (for example, the 16 possible isomers for the simple anilinic raw material used to make bosutinib, many of whose structures cannot be rigorously established, distinguished or ruled out in the final molecule by NMR analysis);
>
> (ii) the general and specific difficulty of rigorously proving many complex chemical structures such as bosutinib by any methods other than x-ray crystallography; and
>
> (iii) the fact that bosutinib has been or is now being administered to humans in clinical trials,

bosutinib should now be rigorously defined as that x-ray crystal structure that results from x-ray crystallography carried out on a sample taken directly from the actual "active pharmaceutical ingredient" supplies used to prepare the drug product administered to humans in the Pfizer-sponsored clinical trials – "from the same lot".

Put another way, whatever the structure is of the compound being used in the Pfizer clinical trials, it should now be taken by definition to be "the right stuff", and its structure is by definition the correct structure.

We do not have access to any X-ray crystallography information proving the chemical structure of Pfizer's clinical bosutinib active ingredient. Nor do we know how it is being manufactured, nor from what starting materials, nor its connection to reported pharmacological properties. Therefore, although it is very likely to be as depicted in the 2001 *J. Med. Chem.* publication, we have no direct and certain knowledge of its actual chemical structure and associated pharmacological properties. We are presently attempting to obtain sufficient data to completely remove any possible doubt on this point.

V.F. "Surrogate markers" or "fingerprints" for a chemical compound to enable easy, absolute, non-x-ray structural and pharmacological certainty

As emphasized above, the only way to establish structural and pharmacological certainty for many somewhat complex compounds is to obtain the X-ray crystal structure on material taken from the same lot as is used to determine the compound's basic pharmacological profile.

Such certainty then applies only to the material in that lot. New material from a different supplier, or a new lot from the same supplier, or from a contract laboratory or corporate synthesis group, might unexpectedly be of a different structure. This could occur from possible use of one or more new lots of starting materials of incorrect structure,

or from a reaction giving an incorrect product because of inadvertently altered conditions or contaminants in the reaction compared to previous runs.

Obviously it is impossible to do X-ray crystallography on even a small fraction of synthetic or semi-synthetic chemical lots of a given compound, let alone doing it for the hundreds of thousands of lots of all compounds of biomedical interest.

In our view, to enable others to replicate the chemical and pharmacological results obtained with a compound from a synthesis, it is necessary to do "sufficient" additional, unambiguous, relatively inexpensive, non-X-ray characterizations of material from the same lot used for X-ray structure determination and pharmacological profiling and to then publicize the surrogate marker data resulting from the additional non-X-ray characterizations.

This theoretically would provide sufficient and easily obtained "surrogate marker" or "fingerprint" data to offer a degree of certainty essentially equal to the 100% certainty of X-ray structure determination, without having to have an authentic reference sample from the "base lot" used in the X-ray structure determination.

What might constitute "unambiguous characterizations and measurements"? This will vary from case to case.

Many chemical compounds have sufficiently simple and/or easily characterized structures that they do not need X-ray structure analysis to establish their structures with 100% certainty.

Compounds with more complex or ambiguous structural elements do require x-ray analysis for 100% certainty. Bosutinib is a good test case, because starting materials and synthetic routes are subject to uncertainties in the present case.

Here is the challenge: An X-ray study and same-lot pharmacological profiling can prove a given structure and associated pharmacological properties with absolute certainty. But X-ray crystallography is prohibitively expensive as a routine analytical tool. Thus, what non-X-ray profiling, what "surrogate markers" can other parties use to independently determine, with 100%, ironclad certainty, that the structure of a new, independent batch of this compound from whatever source is correct?

In the case of bosutinib, perhaps proton NMR data in 4-7 solvents, in conjunction with melting point or some other integrative property, could rule out any possible chemical shift and splitting ambiguity or confounding among the 16 possible anilinic isomers in the final structure.

How many more tests would have to be added, such as C13 NMR and MS fragmentation patterns (which can be hard to reproduce free of artefacts), etc., to reach a point where all highly qualified chemists would agree that a match of the data for bosutinib samples from two different sources is 100% accurate and unambiguous?

This would of course vary from compound to compound, but there are many compounds like bosutinib that might require extensive fingerprinting to achieve certainty.

Part of what we are calling for here is publication of an X-ray structure determination done on a genuine sample of bosutinib from Pfizer's clinical material, along with determination and publication of "surrogate marker" data. That would mean sufficient additional NMR and other chemical and spectroscopic profiling data to enable other parties not in possession of the base reference material to conclude with 100% certainty that a sample of the compound from some other source is indeed chemically identical to the Pfizer clinical material.

As a matter of speculation, it seems unlikely that the kind of surrogate marker work suggested above will ever be applied to most pharmacologically important chemicals, not even to those entering clinical trials. We do hope that such data will emerge for bosutinib itself, because it would provide a speedy avenue towards resolution of this entire issue.

Rather, the worldwide chemical enterprise seems likely to continue its past practices – relying on (i) knowledge of and theoretical predictions for a synthetic route, (ii) generally accurate knowledge of the structures of starting materials, (iii) control of reagent purity, contaminants and reaction conditions, and (iv) various chemical and spectroscopic measurements on the products.

Thus, for many compounds not subjected to X-ray crystal structure determinations, some degree of structural uncertainty will always remain to one or another extent. This is especially true for multiply-substituted aromatic compounds whose structural "proofs" rely on work published in the 1800's and early 1900's and/or for which NMR is a relatively weak analytical tool.

VI. Possible contagion of the anilinic isomer problem to bosutinib analogs and other compounds

A potentially large and serious additional problem is that the medicinal chemistry community continues to create and report many other bosutinib analogs, including some that putatively contain a 2,4-dichloro-5-methoxyaniline substituent.

It seems possible that incorrect starting anilinic components could be in current use as starting materials in these syntheses. A search of chemical supplier websites indicates that there may be as many as 50 or more suppliers of the "2,4-dichloro-5-methoxyaniline" raw material, some of which would be original manufacturers. One or more of these suppliers may be offering one or more different, incorrectly labeled isomers.

If incorrect "2,4-dichloro-5-methoxyaniline" material is indeed being offered in the organic chemical market, that would lead to potential further propagation of structural errors to new bosutinib analogs now being reported by several medicinal chemistry groups.

Such contagion might be quite significant in view of the use of combinatorial chemistry syntheses in many laboratories. One could envision that an incorrect isomer might become incorporated into dozens or even hundreds of compounds in combinatorial libraries in the medicinal chemistry world.

VII. The small-molecule biochemical research reagent market

There is no possibility that batches of even a tiny fraction of the 100,000-odd small-molecule compounds offered by 170-200 or so vendors in the biochemical reagent market will ever carry the certainty of an X-ray crystal structure determination, with or without concomitant bioassay confirmation of pharmacological properties.

In the worldwide organic chemical and biochemical research reagent marketplaces, the overwhelming majority of chemical structures will always be correctly assigned and provided to customers. However, despite the best efforts of all concerned, the potential for problems like those that have arisen for H7, bosutinib, and possibly for new bosutinib analogs can never be completely eliminated, because of the prohibitive cost of X-ray structure analysis as an analytical tool.

###