

Comment

And the second shall be first

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA.
Email: petsko@brandeis.edu

Published: 26 February 2007

Genome Biology 2007, **8**:103 (doi:10.1186/gb-2007-8-2-103)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2007/8/2/103>

© 2007 BioMed Central Ltd

It's one of the most famous photographs in history - almost iconic. It shows a jubilant US President Harry Truman, fresh from his unexpected victory in the Presidential election of 1948 over New York Governor Thomas Dewey, holding up that morning's edition of the *Chicago Daily Tribune*, whose headline boldly - and incorrectly - proclaims "Dewey defeats Truman" (Figure 1).

When Truman went to bed in the Elms Hotel in Excelsior Springs, Missouri, on the night of 2 November 1948, he was losing the election, as most of the polls predicted he would. In its haste to get the scoop on its rival papers, the *Tribune* printed an early edition with a headline giving the expected result based on the early returns. When Truman woke the next morning, he learned he had in fact won, so he took the train to Washington, DC that same day. On a short stop in St Louis, he was presented with one of the "Dewey defeats Truman" papers while on the back platform of the train. It was at this moment that the now famous photo of Truman holding up the paper was taken.

I first saw that picture in the 1960s, when I worked as a stringer for a big-city newspaper. My editor had it framed on the wall of his office. And underneath it, in his own hand, he had placed these words: "It's nice to be first, but it's better to be right."

I was reminded of this photo when I read about the recent retraction by a group from Scripps Research Institute in San Diego, California, of five papers describing protein crystal structures and their corresponding atomic coordinate sets that had been published in *Science*, *Molecular Biology*, and *Proceedings of the National Academy of Sciences* (see Miller: *A scientist's nightmare: software problem leads to five retractions. Nature* 2006, **314**:1856-1857 and Chang *et al.*: *Retraction. Science* 2006, **314**:1875). Apparently, a computer program used by the group had changed the sign of the anomalous differences, which are the differences between X-ray intensity data measured with the X-ray beam



Figure 1

Harry Truman after his unexpected victory over Thomas Dewey in the Presidential election of 1948 with the *Chicago Daily Tribune*, whose early edition boldly and incorrectly proclaimed Dewey's victory.

hitting the front and back of the crystal, in data sets collected for five different membrane protein crystal structures. Anomalous differences can be a powerful aid to solving the structure of a protein by X-ray crystallography. The head of the lab states that "...our MsbA structures were incorrect in both the hand of the structure and the topology. Thus, our biological interpretations based on these inverted models for MsbA are invalid. ...The error in the topology of the original MsbA structure (published in 2001) was a consequence of the low resolution of the data as well as breaks in the electron density for the connecting loop regions. Unfortunately, the use of the multicopy refinement procedure still allowed us to obtain reasonable refinement values for the wrong structures."

The problem might have gone unrecognized for some time longer had Kaspar Locher (Swiss Federal Institute of Technology, Zurich) not determined the correct structure of

a related protein (Locher and Dawson: *Nature* 2006, **443**: 180-185). Locher's structure was completely consistent with the body of biochemical and biophysical data on that class of protein (the so-called ABC transporter superfamily), in contrast to the structures generated at Scripps, which were notably inconsistent.

Their mistake has consequences beyond the damage to the unfortunate young investigator and his team. For five years, other labs have been interpreting their biophysical and biochemical data in terms of the wrong structures. A number of scientists have been unable to publish their results because they seemed to contradict the published X-ray structures. I personally know of at least one investigator whose grant application was turned down for funding because his biochemical data did not agree with the structures. One could argue that an entire sub-field has been held back for years due to the inordinately persuasive power of the pretty pictures that structural biology produces.

In retrospect, there were a number of serious red flags in the work (a major one being the low resolution of the first structure determination, 4.5 Å, a resolution at which it is all too easy to make major mistakes in interpretation). But why on earth *Science*, which published the original paper, and its referees didn't worry from the get-go about the failure of the structure to explain what was already known about this type of protein is beyond me. In an era when experimental details are relegated to 'Supplemental material', especially in the vanity journals, and when canned software makes it easy for people without a deep understanding of the method to determine structures and to referee the structure papers of others, it may be too much to expect that technical errors can be caught reliably, but that isn't the best criterion to use for the correctness of a protein structure anyway. As my Brandeis colleague Chris Miller notes, in a pithy letter on the retractions, "This case highlights the dangers of ignoring biochemical results, conventional but logically solid" (Miller: *Science* 2007, **315**:459). I've said it before, but it bears repeating: the only reliable test for the correctness of a macromolecular structure is if it makes sense in terms of what is already known about the molecule. If it is consistent with the body of experimental data about the protein or its family, it is probably right. If it is not consistent, it is very likely wrong.

But it seems to me that in all the hoopla about this incident, another point that needs to be made has gotten lost. The hero of the story, of course, is Kaspar Locher, who was not deterred from completing and publishing his own structure even though he had apparently been scooped five years previously. That was easier for him in the field of membrane proteins, with rather few structures compared with their soluble brethren and where even structures of closely related ones are perceived as worthwhile. Imagine how much more difficult it would be for someone to decide to carry on in the

case of an area that is not as hot? Yet if this story should convince us of anything, it is of the value of the second report, and the danger of overvaluing the first. Journals like *Nature*, *Science* and *Cell* place so much importance on being the first to publish something of general interest that they create enormous pressure on people to rush to print. No structural biologist would ordinarily settle for 4.5 Å resolution and expect to get things right unless they felt the need to beat their competitors and to have their work published in a journal with as high a profile as possible.

I think that feeling is often self-defeating, as it clearly was here. The high-profile journals don't *ipso facto* do the best job of reviewing manuscripts - in fact, in my experience, they are often a little worse than the so-called trade journals. And very often the first report of something is incomplete, hasty in its judgments, and not nearly as informative as the second paper, which has the advantage not only of calmer consideration, but also has the first paper to use for target practice. But try telling that to students and post-docs, who clamor to have their work sent to the vanity journals, even when it is patently too specialized. And try telling it to the editors of those journals, who, one worries, may have come to value priority over everything else.

How did we get to this state? It makes no sense to me, because in science what we are supposed to value above all else is reproducibility. The report that confirms a finding should, therefore, be considered of equal value to the one that first announces it, but somehow we have either forgotten that fact or succumbed to a collective frenzy for high-profile publications.

We're all guilty of feeding this beast. I have sat on post-doctoral fellowship panels and listened to people say of candidate X: "She has published five papers from her graduate work, two in *Nature* and one in *Cell*", as if that fact alone is all that needs to be said about the quality of the applicant. Frequently, stating where the papers were published is a surrogate for actually having read them. I'm ashamed to say I've done that myself. Does being the first into print mean more than publishing the best paper, the most thoughtful paper, or a more useful paper on a subject? Does content mean so little any more?

Right now the field of genomics is somewhat insulated from this problem. Nearly all the major genome sequences have been either collaborative or solo efforts. The one example of fierce competition, to be the first to sequence the human genome, resulted in an arranged dead heat, and as far as we know didn't affect the quality of the finished product. But as the \$1,000 genome sequence edges closer to reality, and as the supply of really interesting organisms whose genomes have yet to be sequenced shrinks, you can bet there will be more races, more pressure to get there first, and more cutting of corners along the way. When that happens, will we

remember that all too often the first report is sketchy, superficial in its analysis, and more prone to error? Will we value, as much as we should, the second report, which is often more thoughtful, more useful, and is essential to the scientific process of validation and self-correction?

I think that's the real lesson of the unfortunate events at Scripps. And it's a lesson that all of us - every student, every post-doc, every faculty member, every journal editor, every pharmaceutical executive and biotech science officer, every referee, every grants administrator, every scientist, everywhere - would do well to remember.

So I'm going to frame a copy of that Harry Truman photo, and put it on the wall in my office - or maybe in my lab. And underneath it, in my own hand, I'm going to write: "It's nice to be first, but it's better to be right."