



American Chemical Society

OFFICE OF THE PRESIDENT

William F. Carroll, Jr.
President-Elect, 2004
President, 2005
Immediate Past President, 2006

1155 SIXTEENTH STREET, N.W.
WASHINGTON, D.C. 20036
Phone 202-872-4461
Fax 202-872-6338

August 23, 2005

Dear Colleague:

I want to update you on the status of our work toward a resolution of the NIH "PubChem" database issue, which was the subject of my June 20 letter to members (http://acswebcontent.acs.org/PDF/PubChem_open_letter.pdf).

According to NIH, PubChem has two purposes: to disseminate data from the molecular libraries screening center network and to create a comprehensive database of chemical structures with biological activities. While the first has drawn wide support, the second puts PubChem on course to fully duplicate a widely available database offered by ACS: the CAS Registry.[®] ACS has invested heavily over decades to build and maintain the CAS Registry,[®] which already includes the kind of data NIH is assembling in PubChem. We believe NIH should take advantage of and not replicate the CAS Registry to advance its molecular libraries initiative (<http://nihroadmap.nih.gov>).

ACS has been striving to work with NIH to arrive at a mutually beneficial solution. We have initiated more than a dozen good faith meetings, letters, e-mails, and phone calls with NIH, and I personally have had several face-to-face and telephone communications with NIH Director Elias Zerhouni. Despite assurances from NIH that PubChem will focus only on a limited set of "biomedically relevant" compounds, PubChem was recently expanded from 600,000 to more than 3 million compounds, most of which have no particular biomedical focus or affiliated data. We believe there is a better way to serve science, and we have proposed a truly collaborative approach to NIH.

ACS has offered to build, manage, and make available *for free* a database for NIH of data created both by its screening centers and all other compounds with associated bioassay data. ACS would contribute up to 15 staff and costs worth \$10 million over 5 years. Because ACS already analyzes and manages data on virtually all compounds, this approach would save money and guarantee the high-quality data verification and standardization scientists expect. ACS and NIH would work together to disseminate any additional compound data in a way that utilizes, and avoids competition with, existing databases. Unfortunately, NIH has just rejected our offer, but we are studying their newly proposed alternative structure for resolving this issue.

In the end, by taking advantage of the CAS Registry, we believe NIH can:

- Avoid unnecessary duplication and gain cost savings that can be used for research grants.
- Satisfy federal policies that direct agencies to rely on substantially similar private services that can meet agency missions and avoid unfair government competition.
- Accelerate its molecular libraries initiative by utilizing a well-established, comprehensive and high quality tool for disseminating data on chemical compounds.

The issue is not *whether* to advance the NIH Molecular Libraries Initiative but *how*. Both the House and Senate appropriations committees have urged NIH to focus PubChem to avoid unnecessary duplication and competition with private sector databases. We are firmly committed to working toward a resolution that advances NIH's Roadmap, respects longstanding private efforts, and best serves science. I genuinely hope we can achieve this soon.

Sincerely,

William F. Carroll, Jr.
President, American Chemical Society